

# Der Katalog des Deutschen Literaturarchivs Marbach – Metadaten für einen „breiten Kanon“

*Karin Schmidgall*

„Archiv und Kanon“ war der Titel eines Workshops, zu dem sich vom 10. – 12. Mai 2023 Mitglieder des interdisziplinären Netzwerks #breiterkanon trafen. Neben allgemeineren Fragestellungen nach Repräsentation und Marginalisierungen im Archiv beschäftigte sich die Gruppe im Rahmen konkreter Forschungsfragen mit einzelnen Autorinnen und Autoren – u.a. Therese Huber, Charlotte von Ahlefeld, Gabriele Tergit, Louise von Ploennies, Friederike Hauffe und Wilhelmine von Hillern – sowie mit thematischen Fragestellungen zu Metadaten und Mehrsprachigkeit, der Neu-Edierung von Autorinnen des 20. Jahrhunderts, u.a.m.<sup>1</sup>

Für einen Impulsbeitrag aus dem DLA zum Thema Metadaten wurden kleinere Auswertungen auf Basis der Katalogdaten des Marbacher Katalogs erstellt, die in diesem Blogbeitrag vorgestellt werden.

Der Onlinekatalog des DLA umfasst 3,8 Millionen Einträge, er enthält Daten zu:

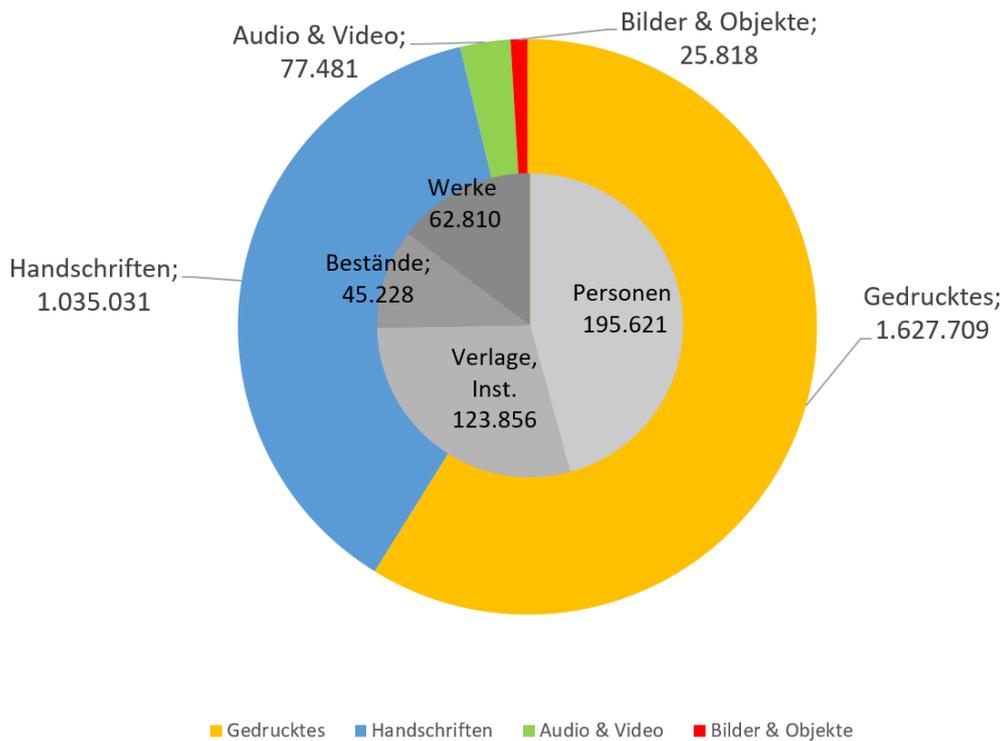
- Autor:innen, wie Lebensdaten, Geburts-, Sterbe-, Wirkungsorte und -land, Geschlecht, Beruf, Beziehungen zu weiteren Personen und Verlagen
- Verlagen (Verlagsort, Gründungs- und Auflösungsdaten)
- Werken (Gattungen, Erscheinungsjahre, Informationen zur Werkgeschichte, Beziehungswerke wie Verfilmungen, Vertonungen)
- Beständen/einzelnen Objekten (strukturierte und normierte Beschreibungsdaten von literarischen Nachlässen, Verlags- und Redaktionsarchiven, Manuskripten, Korrespondenzen, Ton- und Bildträgern, Pressematerialien, in Zeitschriften und

---

<sup>1</sup> Tagungsankündigung auf der Webseite des DLA Marbach: <https://www.dla-marbach.de/kalender/detail/573> [abgerufen am 18.10.2023].

Sammelwerken veröffentlichten Beiträgen, Büchern, Provenienzexemplaren, Zeitschriften, Fotos, Bildnissen und Erinnerungsstücken.)

Jede Entität (Autorin und Autor, Verlag, Werke, Bestand/Einzelobjekt) hat im Katalog einen systeminternen persistenten Identifier und ist darüber mit anderen Entitäten in Beziehung gesetzt.



*Das Schaubild zeigt die Verteilung der Medien im Katalog (Stand 2023), die Objekte beschreibenden Titeldaten sind mit Normdaten (Personen, Verlage, Werke) verknüpft. Für Nachlässe, Sammlungen und geschlossen aufgestellte Autoren- und Sammlerbibliotheken gibt es sogenannte Bestandsdatensätze, die die Sammlung als Ganzes beschreiben.*

Über das zentrale Suchfeld des Katalogs ist die benutzerfreundliche gleichzeitige Suche nach Handschriften, Gedrucktem, Bildern, Gegenständen, Audio- und Video-Dokumenten sowie digitalen Objekten möglich. Der Anspruch des Katalogs ist es, Forschende bei ihrer Arbeit zu unterstützen, sie aber auch auf neue Ideen und Pfade zu bringen.

Für den Workshop wurde auf der Startseite im Katalog ein Teaser ergänzt, der zu einer [Liste von Autorinnen \(deutschsprachige Literatur\)](#) führt<sup>2</sup>, die Liste (11776 Einträge<sup>3</sup>) ist

<sup>2</sup> Suchkriterium: die Person hat den Status „literarische Autorin“, eine Zuordnung zu einer Signaturgruppe (Epoche) erhalten und das Feld „Geschlecht“ ist im Personennormsatz mit weiblich belegt.

<sup>3</sup> Die Anzahl der Einträge Stand 19.10.2023.

sortiert nach Häufigkeit der verknüpften Titelnachweise im Katalog. Die Einträge mit den meisten Verknüpfungen findet man auf „Seite 1“: Hilde Domin (7997 Einträge), Gertrud von LeFort (6452 Einträge), Ricarda Huch (4747 Einträge), Sarah Kirsch (4715 Einträge), Elisabeth Borchers (4609 Einträge), Claire Goll (3919 Einträge). Blättert man die Liste ausgehend von der letzten Seite durch, kann man unbekanntere Autorinnen entdecken, zu denen es im DLA Bestand gibt, oft nur einen Beitrag in einer literarischen Zeitschrift oder wenige Dokumente, wie z.B. Alma Johanna König (68 Einträge), Ilona Bodden (65 Einträge), Ossip Schubin (71 Einträge) oder Lioba Happel (63 Einträge).

**Kallias** – der Online-Katalog des Deutschen Literaturarchivs Marbach

Merkliste (0) Feedback

Gesamtbestand Gedrucktes Handschriften Bilder & Objekte Audio & Video Namen & Werke

Therese Huber

Erweiterte Suche Bestandsübersicht Hilfe & Auskunft

**Huber, Therese (1764-1829)**  
Schriftstellerin, Übersetzerin, Redakteurin

Weitere Details

**von Huber, Therese ...**

- Gedrucktes (41)
- Handschriften (533)
- Bilder & Objekte (0)
- Audio & Video (0)
- Daten (0)

**an Huber, Therese ...**

- Gedrucktes (6)
- Handschriften (94)
- Bilder & Objekte (0)
- Audio & Video (0)
- Daten (0)

**über Huber, Therese ...**

- Gedrucktes (64)
- Handschriften (4)
- Bilder & Objekte (0)
- Audio & Video (2)
- Daten (0)
- Bestände (0)

**unter Huber, Therese ...**

- Bestände (4)
- Bilder & Objekte (0)
- Provenienzexemplare (0)

Mögliche weitere Treffer

- Gedrucktes (0)
- Handschriften (18)
- Bilder & Objekte (0)
- Audio & Video (0)
- Bestände (0)
- Daten (0)

Nicht die richtige Kategorie dabei?

Johannes Schreiber, see [1] | Wikimedia Commons | Public domain

Werkbeziehungen  
Werke (5)

Andere Fundorte  
Wikipedia  
NDB  
GND  
Wikidata

Die Abbildung zeigt den Eintrag zu Therese Huber im Katalog des DLA (<https://www.dla-marbach.de/find/opac/id/PE00008249>). Über „weitere Details“ gelangt man zu biographischen Informationen, über „Werkbeziehungen“ erhält man eine Liste ihrer Werke, zu denen es bereits Werknormsätze gibt, andere Fundorte verbinden den Katalogeintrag mit externen Nachschlagewerken.

Folgende Suchlinks<sup>4</sup> bieten weitere Auswahllisten im Katalog an: Frauen, zu denen es im Archiv Vorlass- oder Nachlassbestand gibt, sogenannte [Bestandsbildnerinnen](#) (246 Einträge), [Künstlerinnen in der Sammlung Bilder & Objekte](#) (132 Einträge), deutschsprachige Autorinnen mit Erstveröffentlichungsjahr [bis 1945](#) (2014 Einträge) und [ab 1945](#) (9758 Einträge), [fremdsprachig schreibende Autorinnen](#) (1572).

Eine wichtige Methode im Katalog sind Facetten, mit denen explorative Suchen möglich sind, damit können große Treffermengen eingeschränkt werden, z.B. auf eine Erscheinungsform, Sprache, Gattung oder Zeitspanne. Facetten eignen sich auch für statistische Datenanalysen und können zur Visualisierung der Katalogdaten genutzt werden. Mit der Anwendung von Facetten nähert man sich aus der Vogelperspektive den Sammlungen. Der Onlinekatalog stößt aber an Grenzen, wenn die Fragestellung eine Verschneidung von biographischen und bibliographischen Informationen enthält, ebenso bei vergleichenden Analysen. Als Beispiele für derartige Fragestellungen, wie sie sich aus aktuellen Forschungsprojektzusammenhängen ergeben, seien genannt: Wie ist die Gender-Verteilung in Verlagsproduktionen? Wie hoch ist der Anteil an publizierenden Frauen bei den Verlagen Cotta und Reclam im Vergleich zu Suhrkamp und Insel oder Rotbuch? Wie sieht die Genderverteilung in expressionistischen Zeitschriften aus?

Die Datenbasis für den Katalog ist eine OpenRefine<sup>5</sup> Umgebung, die eine von der Katalog-Oberfläche unabhängige Sicht auf die Daten ermöglicht. Für die folgenden Auswertungen<sup>6</sup> wurde direkt mit den Daten in OpenRefine gearbeitet, dort können bibliographische Daten um biographische Daten angereichert und anschließend z.B. genderspezifische Auswertungen durchgeführt werden. Zur Zeit arbeitet das DLA im Rahmen eines Kooperationsprojekts Data+<sup>7</sup> an einer offenen Schnittstelle, die Ende des Jahres Anwender:innen ermöglichen wird, selbst Abfragen zu formulieren. Ein Export der Datensets in verschiedenen Formaten macht es möglich, die Daten dann mit eigenen Tools zu analysieren oder in weitere Umgebungen zu integrieren.<sup>8</sup>

---

<sup>4</sup> Direkt im Passus als Hyperlinks hinterlegt.

<sup>5</sup> <https://openrefine.org/> [eingesehen am 18.10.2023]. Mit OpenRefine kann man Daten analysieren, bereinigen, konvertieren und anreichern. Über den Reconciliation Service können zudem die eigenen Daten mit weiteren Datenquellen im Internet abgeglichen und ergänzt werden, z.B. mit der GND oder Wikidata.

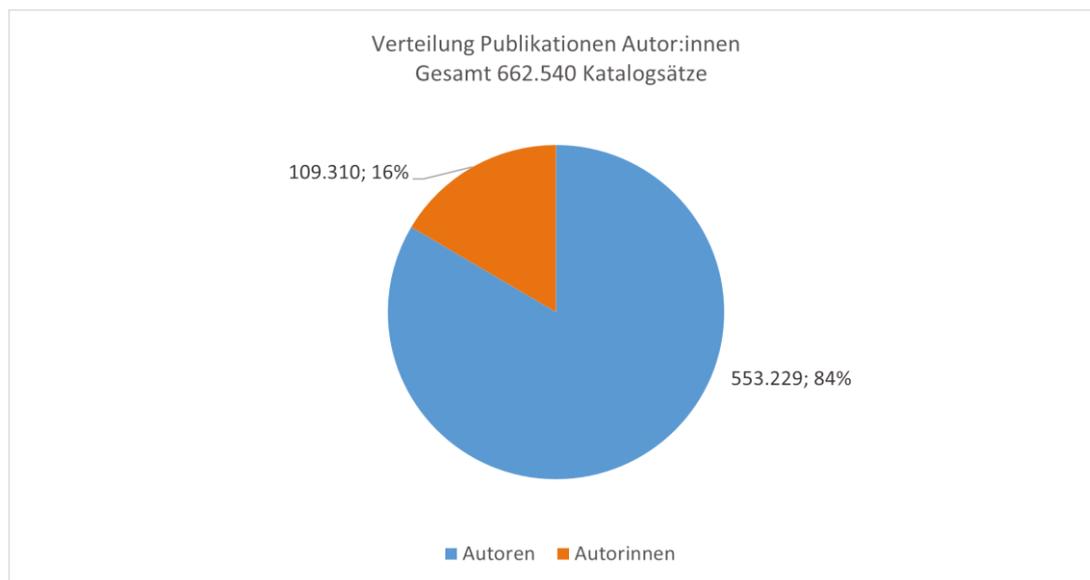
<sup>6</sup> Die Daten für die Auswertung hatten den Stand 02/2023.

<sup>7</sup> Finanziert über das NFDI Konsortium Text+.

<sup>8</sup> <https://textplus.hypotheses.org/6774> [eingesehen am 20.10.2023].

Von den 195.620 im Katalog enthaltenen Personen-Normsätzen sind rund 60.000 als literarische Autor:innen (deutschsprachige Literatur) gekennzeichnet und haben eine Zuordnung zu einer Signatur (Epochenzuordnung)<sup>9</sup>, die sich aus dem Jahr der ersten Veröffentlichung ergibt. In den rund 60.000 Einträgen haben 36% die Belegung männlich, 20% weiblich, 44% keine Angabe<sup>10</sup>. Die Angabe „divers“ kann bislang nicht in codierter Form<sup>11</sup> angegeben werden, auch der Erfassungsleitfaden der Gemeinsamen Normdatei (GND) lässt bislang als codierte Werte nur „Frau“ oder „Mann“ zu und weist darauf hin, dass eine Geschlechterangabe wie „intersexuell“ in einem Freitextfeld erfasst werden kann.

Für die folgenden Auswertungen wurden nur Personen-Normsätze mit einer Angabe im Feld Geschlechterangabe berücksichtigt, in die bibliographischen Metadaten der Verfasserverwerke<sup>12</sup> wurde das Geschlecht der Verfasser:innen zugetragen. Berücksichtigt werden konnten 662.540 Katalogsätze (Bibliotheksdaten: Gedrucktes, Audio und Video, unselbständig und selbständig erschienen).



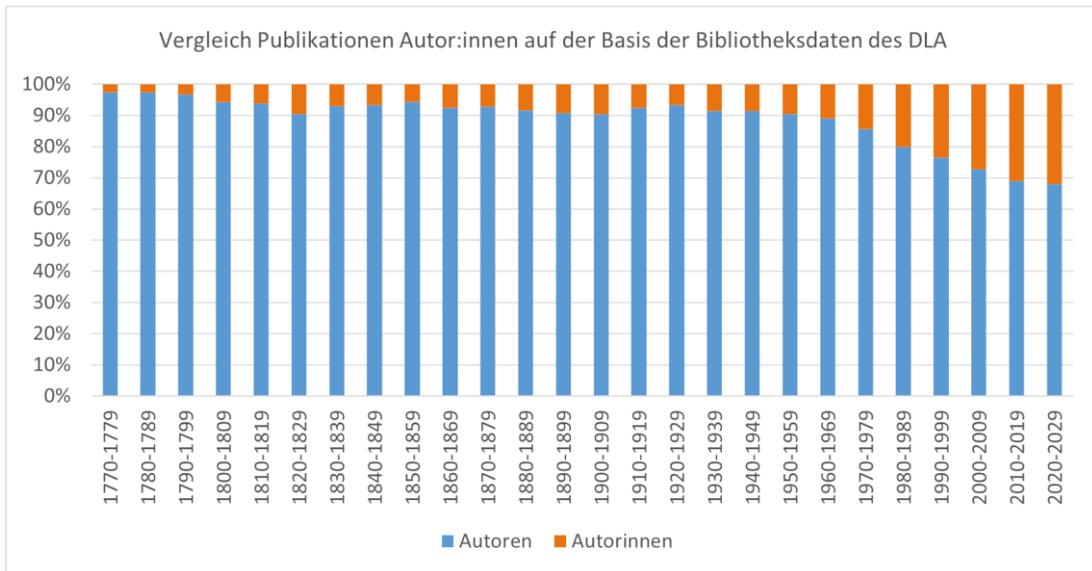
*Die Abbildung zeigt die Genderverteilung bezogen auf die Gesamtmenge der betrachteten Katalogsätze: 553.229 Publikationen von 21.830 Autoren und 109.310 Publikationen von 11.580 Autorinnen. Auf einen Autor kommen im Katalog im Durchschnitt 25 Publikationen, auf eine Autorin 9 Publikationen.*

<sup>9</sup> Deutschsprachig schreibende Autorinnen mit einer Erstveröffentlichung in den Jahren 1880 bis 1909 haben z.B. die Signatur „K“.

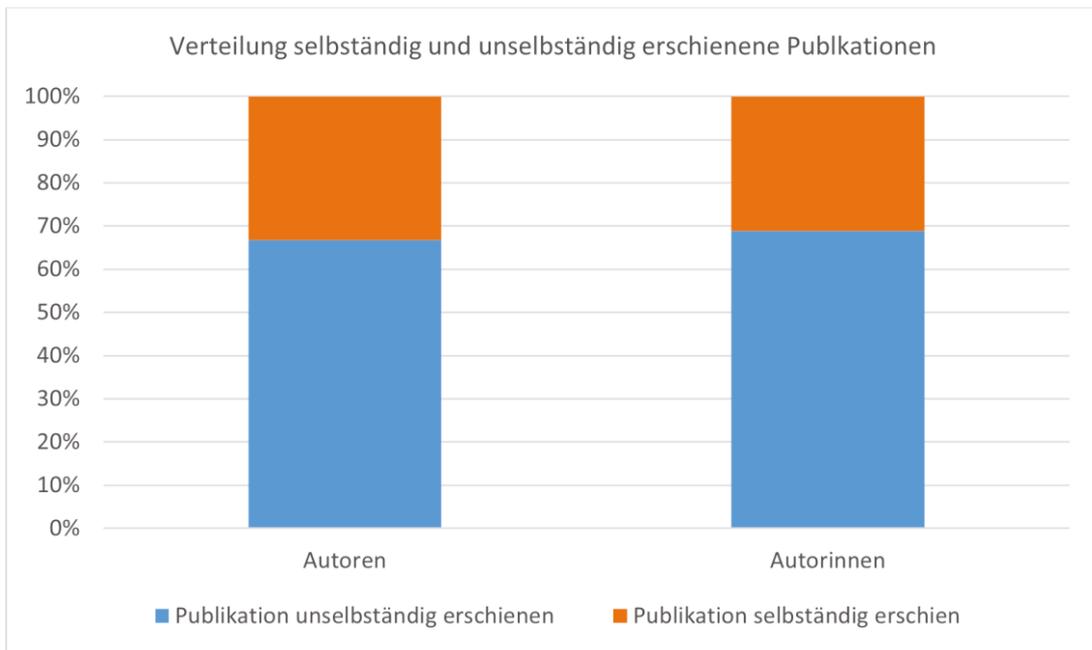
<sup>10</sup> Das Feld „Geschlecht“ wird erst seit 1998 erfasst und konnte in Altdaten bislang nicht systematisch nachgetragen werden. Es ist geplant, mit halbmaschinellen Verfahren das Feld rückwirkend zu belegen.

<sup>11</sup> Ein dringendes Desiderat, da nur codierte Angaben eine Auswertung ermöglichen.

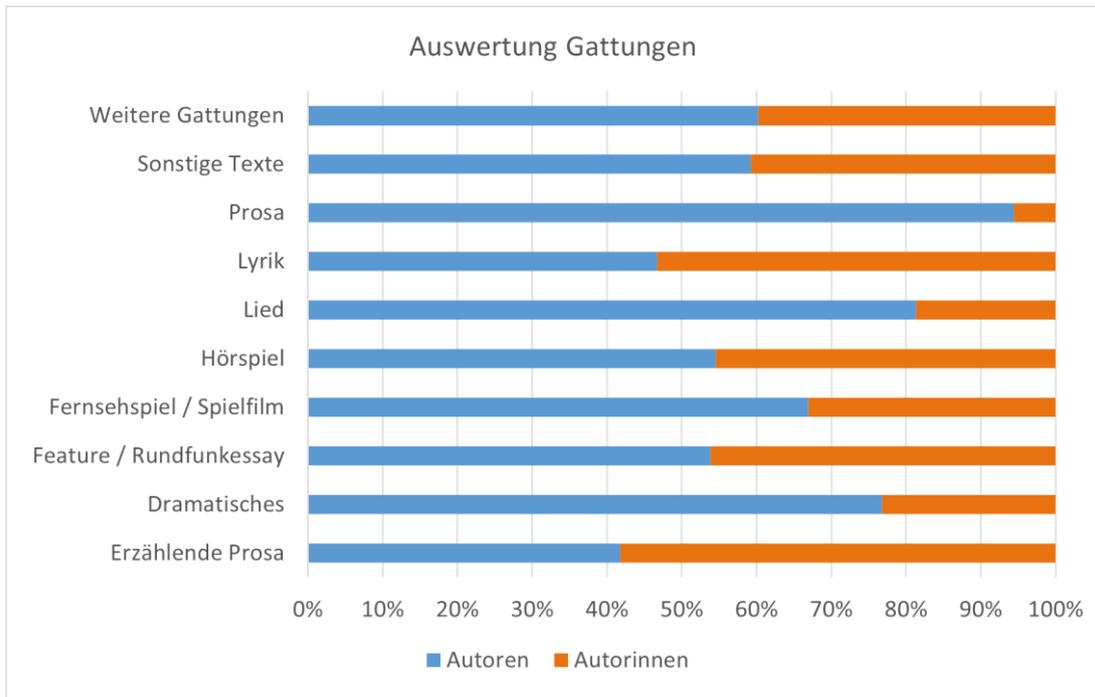
<sup>12</sup> Ausgeklammert wurden Daten von Herausgeber-Werken, diese können bei manchen Verlagspublikationen einen recht hohen Anteil ausmachen.



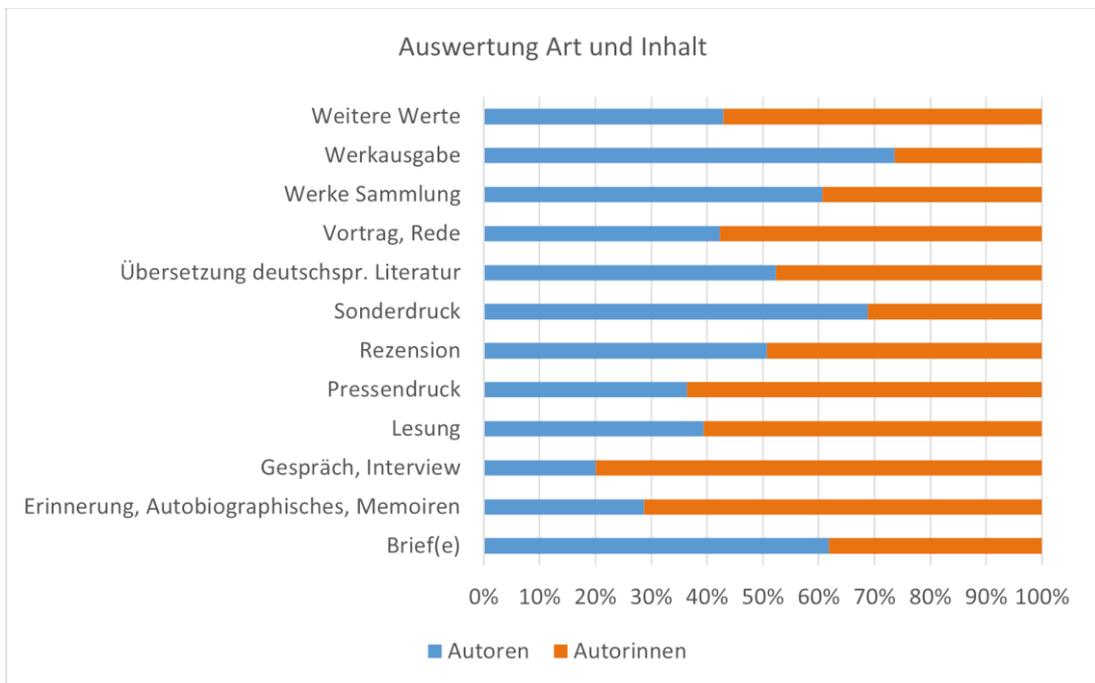
Die Abbildung zeigt eine kontinuierliche Steigerung der Publikationen von Autorinnen im Katalog, ab den 1950iger Jahren signifikant.



Die Abbildung zeigt die Verteilung auf selbständig und unselbständig erschienene Publikationen, hier sind keine genderspezifischen Unterschiede erkennbar.



*Die Abbildung zeigt den Vergleich im Bereich Gattung, nur etwa jeder dritte bis vierte Datensatz enthält Gattungsangaben.*



*Die Abbildung zeigt den Vergleich im Bereich Art & Inhalt, nur etwa jeder dritte bis vierte Datensatz enthält Angaben.*

<b>Verlage mit den meisten Publikationen im Katalog von Autoren</b>	
Suhrkamp	4125
S. Fischer	3928
Reclam	2735
Insel-Verl.	2358
Rowohlt	2033
Cotta	1736
Dt. Taschenbuch-Verl.	1599
Aufbau-Verl.	1504
Piper	1326
Diogenes	1167

<b>Verlage mit den meisten Publikationen im Katalog von Autorinnen</b>	
S. Fischer	1085
Dt. Taschenbuch-Verl.	721
Suhrkamp	646
Piper	614
Insel-Verl.	344
Rowohlt	311
Luchterhand	306
Aufbau-Verl.	268
Ullstein	247
Kiepenheuer & Witsch	247

<b>Autoren mit den meisten Publikationen (selbständig und unselbständig erschienen)</b>	
Goethe, Johann Wolfgang von (1749-1832)	4983
Schiller, Friedrich von (1759-1805)	4699
Hesse, Hermann (1877-1962)	4621
Döblin, Alfred (1878-1957)	4370
Kracauer, Siegfried (1889-1966)	3660
Tucholsky, Kurt (1890-1935)	3254
Rilke, Rainer Maria (1875-1926)	2760
Brecht, Bertolt (1898-1956)	1921
Mann, Thomas (1875-1955)	1753
Ettlinger, Karl (1882-1939)	1588
Jünger, Ernst (1895-1998)	1564
Noder, Anton Alfred (1864-1936)	1486
Heine, Heinrich (1797-1856)	1430
Fontane, Theodor (1819-1898)	1366
Kracauer, Siegfried (1889-1966)	1303
Mörike, Eduard (1804-1875)	1266
Kunert, Günter (1929-2019)	1240
Kafka, Franz (1883-1924)	1223
Zweig, Stefan (1881-1942)	1193
Hofmannsthal, Hugo von (1874-1929)	1168

<b>Autorinnen mit den meisten Publikationen (selbständig und unselbständig erschienen)</b>	
Mayröcker, Friederike (1924-2021)	709
Seghers, Anna (1900-1983)	548
Huch, Ricarda (1864-1947)	514
Bäumer, Gertrud (1873-1954)	490
Lasker-Schüler, Else (1869-1945)	476
Wohmann, Gabriele (1932-2015)	471
Kaschnitz, Marie Luise (1901-1974)	450
LeFort, Gertrud von (1876-1971)	420
Wolf, Christa (1929-2011)	375
Rinser, Luise (1911-2002)	349
Jelinek, Elfriede (1946-)	349
Domin, Hilde (1909-2006)	322
Kirsch, Sarah (1935-2013)	317
Bachmann, Ingeborg (1926-1973)	307
Seidel, Ina (1885-1974)	301
Ebner-Eschenbach, Marie von (1830-1916)	298
Frischmuth, Barbara (1941-)	292
Krechel, Ursula (1947-)	289
Drewitz, Ingeborg (1923-1986)	283
Droste-Hülshoff, Annette von (1797-1848)	282

Die Auswertungen sind lediglich Beobachtungen auf Basis bestandbezogener Katalogdaten und noch ohne weitere Bewertungen. Die Daten des DLA Marbach haben auf Grund ihrer Dichte, Tiefe, Medienvielfalt und hohem Vernetzungsgrad viel Potential für vergleichende Fragestellungen, dabei ist aber zu berücksichtigen, dass Bestandskataloge nicht homogen sind. Regelwerke und Standards ändern sich, Inhalte können in unterschiedlichen Feldern abgelegt sein, nur zum Teil strukturiert und normiert vorliegen. Insbesondere für Auswertungen relevante Felder wie Gattungen, Sprache, Art/Inhalt oder

Genderangaben sind oft nicht durchgängig belegt. Die Kuratierung von Katalogdaten ist daher für Archive und Bibliotheken eine Daueraufgabe, die der Auffindbarkeit, Verständlichkeit und Nutzbarkeit von Datensätzen dient. Immer wieder neu ist zu hinterfragen, wann ein Katalogeintrag qualitativ ausreichend ist. Die Antwort kann projektspezifisch unterschiedlich ausfallen. Um auf Impulse und Forschungsinteressen reagieren zu können, müssten Katalogdaten (insbesondere aus Kartenkatalogen konvertierte Altdaten) rückwirkend überarbeitet und über halbmaschinelle Abgleich-Verfahren angereichert werden. Archivarische und bibliothekarische Regelwerke sollten kontinuierlich am Bedarf der Forschung weiterentwickelt werden und Forschungsergebnisse (zu Sammlungen und einzelnen Objekten, Erkenntnisse aus Datenanalysen) in die Katalogdaten zurückfließen oder verlinkt werden<sup>13</sup>.

Zum Schluss noch eine Bemerkung zur retrospektiven Bestandsergänzung im DLA, da letztendlich Katalogdaten immer ein Spiegel (unikaler) Sammlungen und damit eng mit dem Sammlungsprofil und dem Bestandsaufbau einer Institution verbunden sind. In den letzten Jahren ergänzte die Bibliothek des DLA ihre Sammlung mit Erstausgaben von Autorinnen und baut damit auch kontinuierlich den „breiten Kanon“ im Katalog aus. Als Hilfsmittel diente u.a. dabei die Bibliografie »Die deutschsprachigen Schriftstellerinnen des 18. und 19. Jahrhunderts« von Elisabeth Friedrichs (1981), die zahlreiche auch heutzutage noch bekannte, teils aber auch in Vergessenheit geratene, für ihre eigene Zeit jedoch nicht unbedeutende Schriftstellerinnen versammelt. Im Laufe der letzten Jahre wurden u.a. Werke von Susanne von Bandemer, Frieda und Margarethe von Bülow, Elise Bürger, Annette Kolb, Rosa Mayreder und Sophie Mereau erworben. Darunter sind auch seltene, deutschlandweit nur vereinzelt nachgewiesene Titel von Autorinnen.<sup>14</sup>

---

<sup>13</sup> Im Projekt Literatur im Raum Frauen <https://www.literatursehen.com/projekt/frauenliteratur-im-archiv/> wurden einzelne Objekte aus den Sammlungen des DLA im Rahmen von Forschungspraxisseminaren an der Universität Stuttgart genauer „unter die Lupe“ genommen, auf die im Projektraum veröffentlichten Beiträge der Student:innen könnte z.B. perspektivisch aus dem Katalog des DLA verlinkt werden.

<sup>14</sup> <https://blog.dla-marbach.de/2023/02/09/marbach-ergaenzt-seine-sammlung-mit-erstausgaben-von-autorinnen/> [abgerufen am 19.10.2023].