

Andreas Nestl und Dr. Michael Unger^{*}

Bewertung von Zivilprozessakten bayerischer Amtsgerichte

– Werkstattbericht aus einem Pilotprojekt zur angewandten Datenanalyse
am Beispiel des Fachverfahrens forumSTAR –

Gerichtsurteile stehen einer interessierten (Fach-)Öffentlichkeit als Einzeldokumente zur Verfügung, sei es in Form von gerichtlichen Urteilssammlungen, sei es – zumeist anonymisiert – in publizierter Form¹. Archive eröffnen darüber hinaus Zugänge zu Gerichtsentscheidungen in deren jeweiligem Entstehungskontext, durch den Zugang zu Verfahrensakten, die neben der Evidenz prozessualer Abläufe auch ein Maximum an Informationswert für den jeweiligen Fall bieten. Das ist freilich überhaupt nur für diejenigen Akten möglich, denen ein bleibender Wert zugemessen wird. Als massenhaft gleichförmige Unterlagen stellen gerichtliche Verfahrensakte die archivische Bewertung vor besondere methodische Herausforderungen. Soll nur „Besonderes“ archiviert werden oder auch „Typisches“, und wenn ja, in welchem Umfang und nach welcher Stichprobenmethode? Überlieferungsziele bedingen die jeweils anzuwendende Methodik maßgeblich, aber auch organisatorische und technische Rahmenbedingungen. Insbesondere die Digitalisierung eröffnet der archivischen Bewertung neue Potentiale. Diese Erkenntnis ist in der Archivwelt nicht neu und gerade in Bezug auf Justizakten liegen bereits Erfahrungen zur Nutzung digitaler Bewertungstools

^{*} Dr. Michael Unger ist Ltd. Archivdirektor und in der Generaldirektion der Staatlichen Archive Bayerns Leiter der Abteilung 2 „Überlieferungsbildung, Erschließung, Beratung und Archivpflege“. Andreas Nestl ist Archivoberrat und in dieser Abteilung zuständig für die Bewertung von Unterlagen der Justiz.

- 1 Die Online-Datenbank BAYERN.RECHT² etwa bietet aktuelle wichtige Entscheidungen bayerischer Gerichte sowohl der ordentlichen als auch der Fachgerichtsbarkeit, derzeit im Umfang von 28.463 Dokumenten. Im Internet unter <https://www.gesetze-bayern.de/Search/Hitlist>, zuletzt abgerufen am 05.05.2024.
- 2 Zu archivischen Stichprobenverfahren bei Massenakten siehe v.a. Buchholz, Überlieferungsbildung bei massenhaft gleichförmigen Einzelfallakten im Spannungsverhältnis von Bewertungsdiskussion, Repräsentativität und Nutzungsperspektive. Eine Fallstudie am Beispiel von Sozialhilfeakten der oberbergischen Gemeinde Lindlar (Landschaftsverband Rheinland, Rheinisches Archiv- und Museumsamt, Archivberatungsstelle – Archivhefte 35), Köln 2001, 2011. – Büttner/Kretzschmar/Stahlschmidt, Der archivistische Umgang mit großen Fallaktenserien (Veröffentlichungen der Archivschule Marburg 34), Marburg 2001. – Zu Massenakten der Justiz einschlägig Empfehlungen zur Archivierung von Massenakten der Rechtspflege. Abschlussbericht der Bund-Länder-Arbeitsgruppe zu Fragen der Bewertung und Archivierung von Massenakten der Justiz in Deutschland (Der Archivar, Beiheft 2), Düsseldorf 1999.

vor³. Auch die Staatlichen Archive Bayerns haben vor wenigen Jahren schon einmal entsprechende Testläufe unternommen, wobei die sogenannten C-Verfahren, also die Zivilprozessakten der Amtsgerichte als Versuchsgegenstand dienten. Daran schließt ein Kooperationsprojekt an, dessen Ergebnis derzeit zur Anwendungsreife gebracht werden soll.

Im Folgenden sollen die Rahmenbedingungen dieses Projekts und die darin erarbeiteten Erkenntnisse dargestellt werden. Ausgehend von der aktuellen Praxis bei der Bewertung von Zivilprozessverfahren werden die Möglichkeiten dargestellt, die Stichprobenziehungen aus den Metadaten zu den einzelnen Verfahren für die Bildung einer aussagekräftigen, transparenten und methodisch nachvollziehbaren Überlieferung für diese Aktengruppe bieten. Diese Möglichkeiten führten unmittelbar zu einem neuen Bewertungsmodell für Zivilgerichtsverfahren bei den Amtsgerichten, nach dessen Grundsätzen zukünftig Verfahrensakten archiviert werden sollen. In einem Ausblick gehen die Autoren auf die technischen und organisatorischen Herausforderungen bei der praktischen Umsetzung ein und skizzieren Perspektiven auch über den unmittelbaren Anwendungsfall hinaus.

I. Das Projekt: Rahmenbedingungen, Ziele und Struktur

Den Rahmen für dieses Projekt öffnete das Bayerische Staatsministerium für Digitales, das Ende 2022 eine Initiative zur Förderung von Data Science in der bayerischen Staatsverwaltung startete. In Pilotprojekten sollten dazu konkrete Problemstellungen der Verwaltung mithilfe moderner Datenanalysetechniken gelöst werden. Die Expertise und fachliche Unterstützung stellte der Lehrstuhl für Statistik und Data Science an der Ludwig-Maximilians-Universität München zur Verfügung, dessen Inhaberin Frauke Kreuter auf einschlägige Erfahrungen in den USA zurückgreifen kann⁴.

Die Generaldirektion der Staatlichen Archive Bayerns definierte als Projektziel die Ermittlung archivwürdiger Akten aus einer sehr großen Anzahl gleichförmiger Unterlagen anhand der zu jedem Einzelfall vorliegenden Metadaten mithilfe datenanalytischer Verfahren. Dadurch sollte eine qualitative Verbesserung des bisherigen Auswahlverfahrens in zweierlei Hinsicht erzielt werden: Im Interesse einer transparenten, möglichst auswertungsoffenen archivalischen Überlieferung und zugleich im Interesse von Effizienzgewinnen sowohl für die anbietepflichtigen Justizstellen, als

3 Ziwes, Bewertung zwischen Fingerspitzengefühl und e-Skills. Strategien zur Bewältigung einer archivischen Kernaufgabe. In: Ernst/Müller (Hrsg.), Aktuelle Fragen der Überlieferungsbildung. Vorträge des 79. Südwestdeutschen Archivtags am 16. und 17. Mai 2019 in Ludwigsburg, Stuttgart 2020, S. 37–45, hier v.a. S. 41–43. – Koch/Naumann, Bewertungsautomat statt Autopsie. Neue Möglichkeiten und ihre Konsequenzen am Beispiel der Strafjustiz. In: Massenakten – Massendaten. Rationalisierung und Automatisierung im Archiv. 87. Deutscher Archivtag in Wolfsburg (Tagungsdokumentation zum Deutschen Archivtag 22), Fulda 2018, S. 37–52.

4 Alle Informationen zu der hier vorgestellten Workshopserie sind frei zugänglich unter <https://ada-oeffentliche-verwaltung.de>, zuletzt abgerufen am 06.05.2024.

auch für die staatlichen Archive, denen die Bewertungskompetenz zusteht. Dabei lag es nahe, an die Erfahrungen mit den C-Verfahren der Amtsgerichte anzuknüpfen, die mit einem großen Aufkommen pro Jahrgang eine hinreichend große Grundgesamtheit bilden. Umfängliche Metadaten zu diesen Akten verwaltet die bayerische Justiz in dem Fachverfahren forumSTAR. Ein Auszug zu den 126.945 Verfahren, die im Jahr 2018 angelegt worden waren, wurde vom Bayerischen Staatsministerium der Justiz zur Verfügung gestellt. Wegen der Sensibilität der Daten waren an die gemeinsame Arbeitsumgebung der interdisziplinären Projektgruppe aus Archivarinnen und Archivaren, Justizmitarbeitern und Data Scientists besonders hohe Sicherheitsanforderungen zu stellen. Die eigentliche Datenanalyse einschließlich der grafischen Darstellung der Ergebnisse erfolgte mit der Open-Source-Software R.

Den organisatorischen Rahmen bildete eine modulare Workshop-Reihe, in der die folgenden Arbeitsschritte zu durchlaufen waren:

- Frage definieren
- Analysestrategie finden
- Daten erfassen
- Daten betrachten und bereinigen
- Daten analysieren
- Ergebnisse aufbereiten und präsentieren
- Entscheidungen treffen

II. Aktuelle Bewertungspraxis

Ausgangspunkt für das archivfachliche Anliegen nach einer Optimierung der Überlieferungsbildung stellte das derzeitige Bewertungsmodell dar. Im Sinne einer Archivierungsvereinbarung gemäß Art. 6 Abs. 2 BayArchivG regelt seit 1994 die Aussonderungsbekanntmachung Justiz nicht nur das Verfahren, sondern auch den Umfang der von allen Dienststellen des Justizressorts anzubietenden Unterlagen⁵. Als Positivkatalog werden dort diejenigen Registerzeichen und die darunter geführten Aktengruppen aufgeführt, die den zuständigen Archiven umfänglich anzubieten sind. Während das Registerzeichen C an dieser Stelle mehrfach genannt wird, allerdings nur für sehr spezifische Verfahrensgegenstände, finden sich die eigentlichen Zivilprozessakten nur indirekt und pauschal adressiert unter einer Sammelbestimmung zur Anbietung ersichtlich bedeutsamer Unterlagen, die aufgrund einer Einzelfallprüfung aus allen Geschäftsbereichen als Archivsachen gekennzeichnet und behandelt werden sollen. Die dabei maßgeblichen Kriterien umfassen dabei den folgenden Katalog (Nr. 10.2. der Aussonderungsbekanntmachung Justiz):

5 Aussonderungsbekanntmachung Justiz vom 27. April 1994 (JMBL. S. 71), die zuletzt durch Bekanntmachung vom 27. Mai 2013 (JMBL. S. 44) geändert worden ist.

- Akten über Verfahren, an denen bekannte Persönlichkeiten des öffentlichen Lebens (Politiker, Wissenschaftler, Künstler usw.), bedeutende Familien, Stiftungen, Firmen oder sonstige Unternehmen beteiligt waren;
- Akten über Verfahren, in denen Landesrecht aus der Zeit vor 1900 angewendet wurde;
- Akten über Verfahren mit wertvollen Hinweisen auf Volkssitte und Brauchtum, auch wenn es sich um Bagatellfälle handelte;
- Akten über zeittypische Zivilrechtsstreitigkeiten (Mietstreitigkeiten, Reisevertrags-sachen, Sachmängelprozesse usw.)

Primär richtet sich dieser Kriterienkatalog an die bei jeder Justizstelle obligatorisch zu bestimmenden Archivpfleger, also Justizmitarbeiter, an die die Erstbewertung faktisch delegiert ist. Es liegt auf der Hand, dass der Ermessensspielraum in der Anwendung dieser unspezifischen Kriterien breit und das Ergebnis in der Praxis entsprechend vielfältig ausfallen. In der Regel ist der Umfang der angebotenen Unterlagen sehr gering und diese werden ohne Weiteres und mit einem Minimalset an Metadaten zur Nachnutzung für Erschließungszwecke von den zuständigen Archiven ohne Vor-nahme einer weiteren Bewertung übernommen.

III. Analyse und Bereinigung der Metadaten

Im Projekt galt es zunächst, den sehr umfänglichen Metadatenatz aus forumSTAR intellektuell zu erfassen, also die Bedeutung der einzelnen Attribute kennenzulernen und diese anschließend auf ihre Relevanz für die archivfachliche Bewertung zu prüfen. Eine erste Herausforderung stellte dabei die Bereinigung des Datenmodells dar, indem alle Metadaten auf die Entität „Akte“ hin zu orientieren waren. Nach der Klärung, welche Informationen in welchen Feldern enthalten sind, folgten erste gezielte Abfragen und die Analyse der dadurch erzielten Ergebnisse. Wie viele Akten liegen pro Amtsgericht vor? Wie viele Personen waren an den Verfahren beteiligt? Und wie hoch ist absolut bzw. im Mittel der Streitwert? Immer wieder festzustellende Ausreißer legten die Vermutung von Erfassungsfehlern nahe, die nicht immer so offensichtlich in Erscheinung treten wie im Fall einer negativen Verfahrensdauer. Jedenfalls sensibilisierten gerade solche Fälle für die Datenqualität, bei der Unschärfen in einem für den hier verfolgten Zweck vertretbaren Umfang einzuplanen sind. Besonders interessant fiel das Ergebnis auf die Abfrage aus, wie viele und welche Verfahren bereits als Archivsachen gekennzeichnet worden waren. Dies traf auf 372 Akten zu, die ihrer-seits nur aus 48 von 73 Amtsgerichtsbezirken (einschließlich Nebenstellen) gemeldet worden waren. Aus dem Freitextfeld in forumSTAR lagen die Begründungen für die Anbietung vor, deren sehr individueller Charakter jenseits des Umstands, dass es sich um besondere Fälle handelt, wenig Regelmäßigkeit erkennen ließ. Der relativ höchste Anteil des Amtsgerichts München an dieser Teilmenge mit rund 3 Prozent von dessen Akten korrespondiert mit der überragenden Gesamtzahl der dort verhandelten Verfahren. Dass das Amtsgericht Schweinfurt mit ebenfalls 3 Prozent seiner Akten

zahlenmäßig auf Platz zwei folgt, entzieht sich allerdings dieser Logik und dürfte in einer besonders engagierten Anwendung der Aussonderungsbekanntmachung Justiz durch die dortigen Mitarbeitenden begründet sein. Letztlich führen diese Ergebnisse die methodischen Schwächen des aktuellen Bewertungsverfahrens vor Augen, dass einer planvollen und gleichförmigen Überlieferungsbildung nur wenig gerecht wird.

IV. Einfache Stichprobenziehung⁶: Bildung einer repräsentativen Überlieferung

Unter dem Gesichtspunkt einer qualitativen Verbesserung der Überlieferung selbst stand von vornherein auch das Überlieferungsziel auf dem Prüfstand⁷. Sollte etwa an einer Dokumentation von „Besonderem“ festgehalten, oder in einem stärkeren Maße auch „Typisches“ übernommen werden? Dafür waren Stichprobenziehungen zu testen, wobei zunächst einfache Zufallsstichproben gezogen wurden. Während die Grundgesamtheit von vornherein feststand und die Gesamtzahl aller C-Verfahren des Jahrgangs 2018 umfasste, war die Stichprobengröße variabel festzulegen. Je größer die Stichprobe ausfällt, desto genauer kann die Grundgesamtheit abgebildet werden. Genauigkeit bedeutet in diesem Zusammenhang die Nähe eines Stichprobenmittels zum Mittelwert, der auf der Grundlage der Grundgesamtheit errechnet ist. Für eine Stichprobengröße im Umfang der als Archivsachen gekennzeichneten 372 Akten ergab die Zufallsauswahl eine deutlich gleichmäßigere Verteilung auf alle Amtsgerichtsbezirke, wobei die Wahrscheinlichkeit, dass dabei ein Amtsgericht auch nur 1 Prozent seiner Akten anbieten würde, äußerst gering war. Der Vorteil dieser Art der Zufallsauswahl liegt in der Repräsentativität des Ergebnisses, wobei der Vorteil, dass alle Akten dieselbe Auswahlwahrscheinlichkeit haben, zugleich als ein Nachteil gewertet werden kann. Eine inhaltliche oder regionale Schwerpunktsetzung lässt diese Methode beispielsweise nicht zu.

6 Alle Stichprobenziehungen wurden konzipiert, umgesetzt und visualisiert von Dr. Heidi Seibold (Digital Research Academy), Dr. Malte Schierholz und Dr. Marcel Neunhoffer (beide Institut für Statistik der Ludwig-Maximilians-Universität München), denen wir hier einen besonderen Dank aussprechen.

7 Vgl. Arbeitskreis „Archivische Bewertung“ im VdA – Verband deutscher Archivarinnen und Archivare e.V., Kriterien zur Evaluierung von Bewertungsdokumenten. Eine Handreichung zur archivischen Überlieferungsbildung. In: Evaluierung von Bewertungsdokumenten. Beiträge zur archivischen Überlieferungsbildung, hrsg. vom Arbeitskreis „Archivische Bewertung im VdA, Stuttgart 2018, S. 18–33. Wendenburg, Prüfen! Prüfen! Prüfen! Sind inhaltliche Ziele in der Überlieferungsbildung erreichbar? Oder: Drum prüfe, wer sich ewig bindet! Evaluierung von Bewertungsdokumenten. Beiträge zur archivischen Überlieferungsbildung, hrsg. vom Arbeitskreis „Archivische Bewertung im VdA, Stuttgart 2018, S. 44–47.

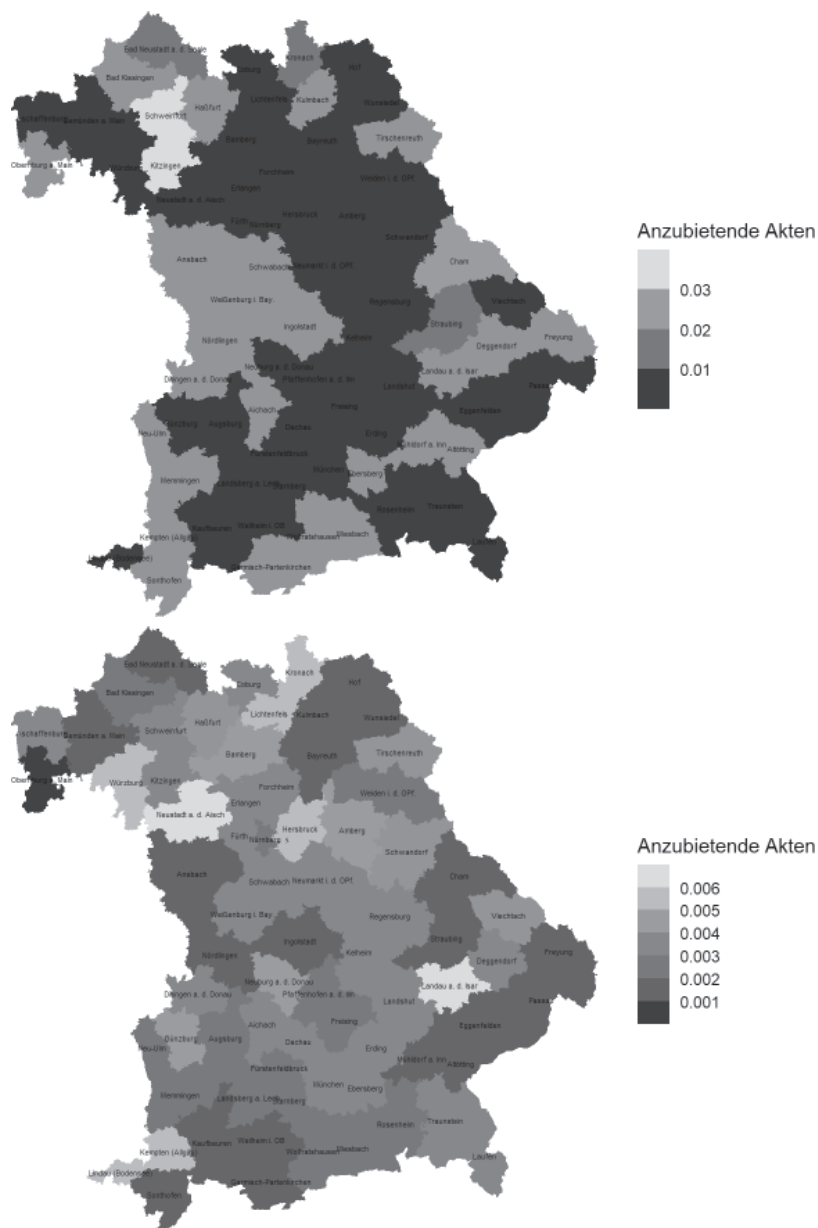


Abbildung 1: Verteilung einer Auswahl von 372 Akten (= Anzahl der von den Gerichten als archivwürdig gekennzeichneten Verfahren); links nach Archivsachenvermerk, rechts nach einfacher Stichprobenziehung.

V. Stratifizierte Stichprobenziehung: Bildung einer repräsentativen und inhaltlich zielgenauen Überlieferung

Dieses Manko kann mit einer stratifizierten bzw. geschichteten Stichprobenziehung umgangen werden. Wollte man etwa sicherstellen, dass Historiker etwa regionale Vergleiche anstellen können und dazu Akten aus allen Amtsgerichtsbezirken archivieren, würde der Amtsgerichtsbezirk als Stratum bzw. Schicht gewählt werden können. Ein Testlauf, in dem zufällig 10 Akten pro Amtsgerichtsbezirk gezogen wurden, führte zu dem Ergebnis, dass Akten aus kleinen Amtsgerichtsbezirken mit einer deutlich höheren Wahrscheinlichkeit Teil der Stichprobe wurden als etwa Akten des Amtsgerichts München. Statistisch valide Aussagen lassen sich aus einem solchen Ergebnis nur ableiten, wenn die unterschiedlichen Auswahlwahrscheinlichkeiten berücksichtigt und hinreichend transparent gemacht werden. Jede Schicht erhielt dann im weiteren Verlauf eine eigene Auswahlwahrscheinlichkeit, wobei sich zusätzlich das Konfidenzintervall errechnen ließ, das mit einer Wahrscheinlichkeit von 95 Prozent den wahren Wert der Grundgesamtheit abdeckt. Ein Testlauf zielte darauf ab, Akten mit einer besonders langen Verfahrensdauer besonders häufig zu archivieren, da diese Akten als besonders aussagekräftig gelten können. Mittels einer gewillkürten Einteilung der Grundgesamtheit in kurze, mittlere und lange Verfahren anhand der Verfahrensdauer in Tagen wurden die Stichprobengrößen für jede Schicht festgelegt, wobei alle langdauernden Verfahren ausgewählt wurden. Aus den Verfahren mittlerer Dauer wurden 100 und aus denen kürzerer Dauer 50 Akten ausgewählt. Der naheliegende Gedanke dahinter: Je kürzer die Verfahrensdauer, desto geringer der Informationswert der Verfahrensakten. Alternativen, die zu einer besseren Aussagekraft hinsichtlich der Grundgesamtheit gelangen, bieten sich an. Die Größe der Stichproben kann proportional zur Größe der in der Schicht vorhandenen Verfahren berechnet, jeder Schicht kann die gleiche Auswahlwahrscheinlichkeit zugeordnet oder die Genauigkeit der Schätzung statistisch optimiert werden.

Zur genaueren Steuerung der auszuwählenden Verfahren hinsichtlich der Zielsetzung der Überlieferung können auch mehrere Schichten miteinander kombiniert werden. Diese Vorgehensweise erhöht einerseits die Komplexität der Auswahl, die im Rahmen der technischen Umsetzung berücksichtigt werden muss. Zudem steigert sich der Dokumentationsaufwand, da eine Rückbeziehung von der stratifizierten Auswahl auf die Grundgesamtheit nur dann gewährleistet ist, wenn die Auswahlkriterien und -wahrscheinlichkeiten nachvollziehbar bleiben. Andererseits erlaubt die Unterscheidung in mehrere Schichten eine wesentlich bessere Abbildung der mit der Archivierung intendierten Zielsetzung. Im Versuch wurde in einem ersten Schritt die Grundgesamtheit reduziert und auf Verfahren beschränkt, die mit einer Sachentscheidung geendet haben. Damit blieben aus der Grundgesamtheit nur die Verfahren übrig, die mit den Erledigungsarten Endurteil, Beschluss und Vergleich beendet worden sind. Verfahren, die ohne eine Sacheinlassung seitens des erkennenden Richters erledigt wurden, weisen nach unserer Annahme einen erheblich geringeren Informationswert auf. Aus der solcherart reduzierten Grundgesamtheit wurde anschließend eine in zwei

Schichten unterteilte stratifizierte Stichprobe gezogen: Die erste vergab an alle Verfahren eine Auswahlwahrscheinlichkeit von 1, die bereits vom Gericht als archivwürdig gekennzeichnet waren oder eine Verfahrensdauer von mehr als 1.500 Tagen aufwiesen; alle anderen Verfahren wurden gleichmäßig berücksichtigt. Nach diesen Vorgaben wurde aus jedem Amtsgericht ein geringer Anteil gezogen, um eine flächendeckende Überlieferung abbilden zu können. Insgesamt wurden vermittels dieser Methode 1.910 Verfahren ausgewählt – und damit deutlich mehr als archivfachlich erwünscht. Dennoch wurde das Ziel erreicht, die Vor- und Nachteile gegenüber der einfachen Zufallsstichprobe beurteilen zu können.

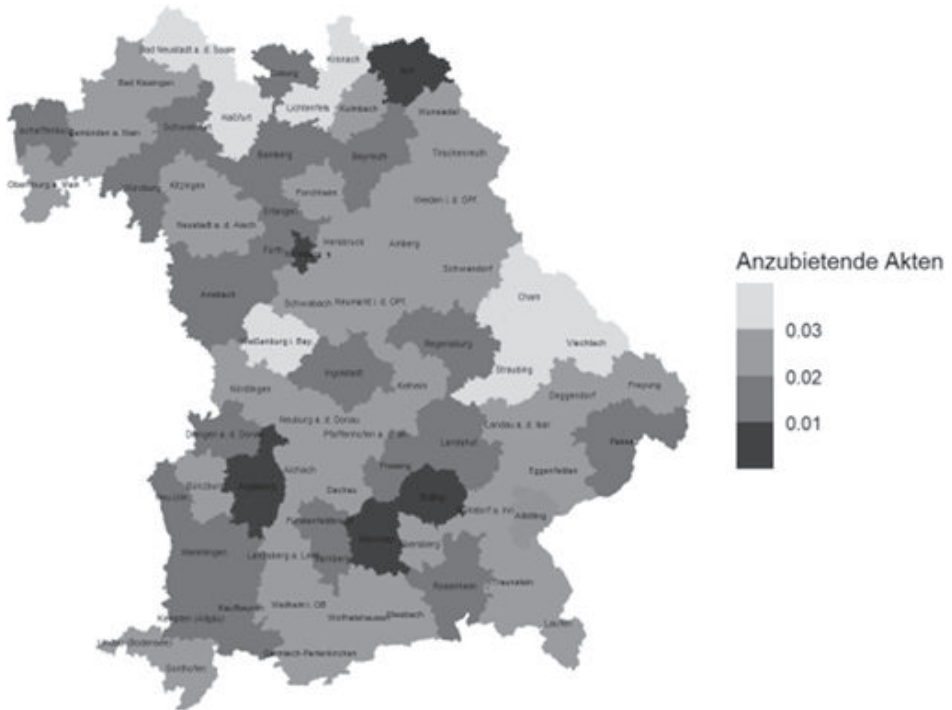


Abbildung 2: Verteilung einer stratifizierten Auswahl von 1.910 Akten.

Letztlich erwies sich die Methode der stratifizierten Stichprobe für die Bewertung der gegenständlichen Verfahrensakten als am besten geeignet, um zielgerichtet inhaltsreiche Verfahren auszuwählen und gleichzeitig – eine geeignete Stichprobengröße vorausgesetzt – Rückschlüsse auf die Grundgesamtheit zuzulassen. Voraussetzung für das Erreichen der letzteren Zielsetzung bleibt freilich eine genaue Dokumentation der vorgenommenen Schritte und deren Zugänglichmachung für die zukünftigen Nutzerinnen und Nutzer. Der weiteren archivfachlichen Entscheidung vorbehalten blieb dabei die Entscheidung über eine Reihe grundlegender Fragen, vor allem die, anhand welcher

Variablen konkret einzelne Schichten definiert und wie viele Akten aus jeder Schicht gezogen werden sollten.

VI. Abbildung des Besonderen

Entsprechend dem Kriterienkatalog der Aussonderungsbekanntmachung Justiz dominiert das „Besondere“ die gängige Bewertungspraxis. Darauf aufbauend wurde der Versuch unternommen, anhand von personenbezogenen Daten aus dem Fachverfahren prominente Personen zu identifizieren und als archivwürdig zu kennzeichnen. Als Hilfsmittel wurde dazu eine externe Ressource, die dem Online-Lexikon „Wikipedia“ zu Grunde liegende Datenbank „Wikidata“ herangezogen. Ein Abgleich des dort verfügbaren Datensatzes zu dem Fußballspieler Thomas Müller, nebenbei die statistisch am häufigsten vorkommende Namenskombination in Deutschland, mit den Einträgen in forumSTAR ergab ebenso zahlreiche wie bei näherer Betrachtung der Daten allesamt unplausible Ergebnisse. Das Problem: die in dieser Hinsicht mangelnde Datenqualität der Fachdaten, die eine eindeutige Identifizierung nicht zulassen: Das Geburtsdatum als womöglich eindeutigstes Merkmal zur Identifikation wird in der Datenbank nicht erfasst. Aktuell steht für die Auswahl des Besonderen im Sinne der o.a. Kriterien nur der Vermerk des aufmerksamen Richters bzw. Archivpflegers zur Verfügung.

VII. Aktualisierung des Bewertungsmodells

Die im Projekt gewonnen Erkenntnisse machten das Potential der Datenanalyse für Zwecke der archivischen Bewertung deutlich. Um dieses Potential für die Praxis nutzbar zu machen, war es notwendig, das bestehende und vorerst noch über die Aussonderungsbekanntmachung Justiz formalisierte Bewertungsmodell für die Archivierung von Zivilprozessakten der Amtsgerichte in Bayern grundlegend zu überarbeiten. Diesen Unterlagen kommt ein besonderer Quellenwert für Fragestellungen aus der Alltags- und Regionalgeschichte zu. Wie vielleicht keine andere Quellengattung versprechen Zivilprozessakten Einblicke in das (streitige) Alltagsgeschehen der Bürgerinnen und Bürger der jeweiligen Amtsgerichtssprengel. Ziel der Bewertung ist künftig daher, nicht allein das „Besondere“, wie bisher in Nr. 10. 2 Aussonderungsbekanntmachung Justiz niedergelegt (s. o.), sondern auch das Alltägliche hinreichend zu dokumentieren, um eine Rekonstruktion der Alltagsgeschichte unter Berücksichtigung regionaler Spezifika zu ermöglichen.⁸ Die alleinige Konzentration auf das Besondere ließe die überwiegende Masse an Beteiligten im Rechts- und Risikoverkehr der Gesellschaft – und

⁸ Treffeisen, Erweitertes Auswahlmodell der staatlichen Archivverwaltung Baden-Württemberg bei Massenakten der Justiz, in: *Archivar*, Jahrgang 2005, S. 188 ff.: Die vorhandenen (digitalen) statistischen Auswertungen sollen danach mit besonderen Einzelfällen beispielhaft ergänzt werden. Warum aber sollen gerade die außergewöhnlichen Fälle die vorhandenen und aussagekräftigen Justizstatistiken mit Leben füllen? Wenn zutrifft, dass vor den Amtsgerichten um

damit auch den überwiegenden Anteil des Geschäftsanfalls der Amtsgerichte – unberücksichtigt. Letztlich soll mit den übernommenen Akten die jährlich erstellte Justizstatistik in Zivilsachen, Familiensachen, Straf- und Bußgeldverfahren sowie in Ermittlungsverfahren, Verfahren nach dem OWiG und sonstigen bei den Staatsanwaltschaften zu erledigenden Geschäften mit Leben gefüllt werden.

Diese Auswahl des Alltäglichen soll durch eine Zufallsstichprobe erfolgen. Das Ziel, die Dokumentation der Alltags- und Regionalgeschichte, erfordert allerdings einen hinreichenden Informationswert der zu übernehmenden Verfahrensakten. Bei einer lediglich zurückgenommenen Klage beispielsweise sind keine oder nur wenige Informationen zum streitigen Lebenssachverhalt zu erwarten, der gerade Gegenstand der Überlieferung sein soll. Daher sollen nur Verfahren zur Auswahl anstehen, die mit einer Entscheidung in der Sache beendet worden sind. In Betracht für die Auswahl kommen nach Rücksprache mit dem Bayerischen Staatsministerium der Justiz vornehmlich Verfahren, die mit den in forumSTAR ausgeworfenen Erledigungsarten Urteil, Vergleich oder Beschluss beendet worden sind.

Aus dieser reduzierten Grundgesamtheit wird jeder Amtsgerichtsbezirk berücksichtigt und damit eine bayernweite Stichprobe aus den Verfahren aller Amtsgerichte gezogen. Um Alltagsgeschichte auch unter Berücksichtigung regionaler Besonderheiten erzählbar zu machen, wird gerade in einem Flächenland wie Bayern mit deutlich unterschiedlichen Landschafts- und Kulturräumen eine regionale Clusterbildung, beispielsweise anhand einzelner Muster-Amtsgerichte mit städtischer und ländlicher Prägung als nicht zielführend angesehen.⁹ Zudem fehlt es aus unserer Sicht an hinlänglichen Kriterien, nach der eine Auswahl einzelner Amtsgerichte vorgenommen werden sollte. Um Entwicklungen und Bewegungen in der Rechtsprechung ohne zeitliche Zäsuren zuverlässig dokumentieren zu können, wird eine Stichprobe für jedes Jahr gezogen.

Für eine auch statistisch nachvollziehbare Überlieferungsbildung muss die Stichprobe groß genug sein, um überhaupt entsprechende Aussagen tätigen zu können. Wie groß die Stichprobe auszufallen hat, ist eine zentrale Frage. Eine in der Bewertungspraxis vielfach übliche Größenbestimmung in Prozent bezogen auf eine Grundgesamtheit ist zwar praktikabel im Sinne einer einfachen Umsetzbarkeit. Sie wird den genannten Anforderungen jedoch nicht gerecht. Wesentlich für eine zufriedenstellende Lösung war dabei die Erkenntnis, dass in der vorliegenden Fallkonstellation einer sehr großen Grundgesamtheit eine Genauigkeit von 95 Prozent erreicht werden kann,

Alltägliches gestritten wird, sollte gerade das Alltägliche, Gewöhnliche, entsprechender Informationsgehalt vorausgesetzt, dokumentiert werden.

⁹ So aber Landesarchiv Baden-Württemberg, Erweitertes Auswahlmodell bei Massenakten der Justiz, Stand 14.11.2022, Nr. 3a, Zivilprozessakten (in Überarbeitung): Mindestens 1 %, maximal 10 % der angebotenen Akten der Jahrgänge mit der Endziffer 0 und 5 bei vier Amtsgerichten, abrufbar unter: https://www.landearchiv-bw.de/sixcms/media.php/120/bewertung-massenakten_justiz_2022_11_14.pdf, zuletzt abgerufen am 06.05.2024.

wenn ein Stichprobenumfang von 385 gewählt wird.¹⁰ Dieser Umfang, der im Bewertungsmodell auf 400 aufgerundet wurde, entspricht ziemlich genau den bisherigen Übernahmemengen und der strategischen Zielsetzung, die künftige Übernahmequote nicht spürbar zu erhöhen. Dies wäre jedoch zwangsläufig der Fall, wenn repräsentative Stichprobenziehungen für jeden Amtsgerichtsbezirk mit vielfach deutlich kleineren Grundgesamtheiten durchgeführt würden, was die Entscheidung zugunsten einer bayernweiten Stichprobe zusätzlich stützte. Diese führt über die Zeit zu einer Aggregation regionaler, jeweils gut dokumentierter Teilstichproben in den einzelnen Staatsarchiven, die ihrerseits eine hinreichende Grundlage auch für rein regionalgeschichtliche Forschungen am Quellenbestand auch nur eines der acht betroffenen Staatsarchive darstellen werden. Ein mögliches Manko der gesamtbayerischen Stichprobe kann auf diese Weise ausgeglichen werden.

Neben der Erledigungsart können sich vor allem aus der Anzahl der vor Gericht wahrgenommenen Termine Hinweise auf die Befassung mit dem Lebenssachverhalt ergeben. Dieses Attribut erscheint im Ergebnis geeigneter als die Verfahrensdauer. Bei der Wahrnehmung von Terminen, ggf. mit Beweiserhebung, ist davon auszugehen, dass eine vertiefte Erörterung des Lebenssachverhalts stattfindet, was genau Zielsetzung der Bewertung ist. Die Länge der Verfahrensdauer korreliert weitgehend mit der Anzahl an Terminen, hat aber einen geringeren Aussagegehalt auf den Inhalt; die Dauer kann vielmehr auch im konkreten Parteibetrieb ihren Grund finden. In einem nächsten Schritt wird daher die Anzahl an Terminen, die sich aus forumSTAR auch auslesen lassen, eine unterschiedliche Gewichtung im Rahmen der Stichprobenziehung erhalten. Verfahren mit 20 oder mehr Terminen werden vollständig übernommen. Verfahren, die zwischen 10 und 20 Termine aufwiesen, werden mit einer Auswahlwahrscheinlichkeit von 20 Prozent übernommen, was in einer ersten Rechnung etwa 100 Verfahren ausmacht. Der Rest der Stichprobe bilden Verfahren, die mit weniger als 10 Terminen ausgekommen sind.

Die Justizstatistik weist eine Übersicht der erledigten Verfahren nach einzelnen Sachgebieten auf, die auch in den Fachverfahrensdaten ausgewiesen werden.¹¹ Um einzelne Sachgebiete, in denen nach unserer Auffassung archivwürdige Verfahren in höherem Maße vorhanden sind, stärker abbilden zu können und gleichzeitig die Dokumentation von Sachgebieten mit einer besonders hohen Zahl gleichförmiger Fälle auf ein angemessenes Maß reduzieren zu können, wurden die Sachgebiete nach thematischen Gesichtspunkten in fünf Gruppen aufgeteilt. Im Rahmen der Stichprobenziehung soll jeder dieser Gruppen die gleiche Auswahlwahrscheinlichkeit zukommen, also jeweils mit 20 Prozent der Stichprobengröße vertreten sein:

10 Vgl. Kauermann/Küchenhoff, Stichproben. Methoden und praktische Umsetzung mit R, Heidelberg 2011, S. 42. Ein besonderer Dank gilt Herrn Prof. Dr. Helmut Küchenhoff für die individuelle Beratung im Rahmen des Projekts.

11 Eine Übersicht der Erledigungsarten bietet die Justizstatistik in Zivilsachen, Familiensachen, Straf- und Bußgeldverfahren sowie in Ermittlungsverfahren, Verfahren nach dem OWiG und sonstigen bei den Staatsanwaltschaften zu erledigenden Geschäfte, hier für 2022 (JStat 2022), Nummern 3600 ff.

- Verkehrsunfallsachen (dieses Sachgebiet weist regelmäßig eine hohe Anzahl erledigter Verfahren auf)
- Körper und Person (Eingriffe in die körperliche Unversehrtheit sollen verstärkt akzentuiert werden)
 - Arzthaftungssachen
 - Körperverletzung
- Leben und Wohnen (Zusammenleben im privaten Raum)
 - Nachbarschaftssachen
 - Wohnungsmietsachen
 - Sonstige Mietsachen
 - Wohnungseigentumssachen
- Wirtschaftsrecht
 - Bau-/Architektensachen (ohne Architektenhonorarsachen)
 - Kaufsachen
 - Reisevertragssachen
 - Kredit-/Leasingsachen
 - Ansprüche aus Versicherungsverträgen (ohne Verkehrsunfallsachen)
 - Gesellschaftsrechtliche Streitigkeiten
 - Honorarforderungen von Personen, für die eine besondere Honorarordnung gilt
- Sonstiges
 - Schuldrechtsanpassungs- und Bodenrechtssachen der neuen Länder
 - Sonstiger Verfahrensgegenstand

Nicht relevant für das Erreichen des Überlieferungsziels ist nach unserer Auffassung eine Berücksichtigung des Streitwerts. Große Streitwerte, die über die gesetzliche Zuweisung von 5.000 Euro hinausgehen, sind vor allem in der besonderen Zuständigkeit der Amtsgerichte für Miet- und WEG-Sachen begründet und dadurch eben kein Indikator für besonders spektakuläre Sachverhalte, die in der Stichprobenziehung hervorgehoben werden sollten. Verfahren mit geringen Streitwerten können vielmehr ebenso geeignet sein zur Dokumentation der Alltagsgeschichte (z. B. in nachbarschaftlichen Streitigkeiten).

Auch wenn der bisherige Fokus auf das „Besondere“ durch das Stichprobenverfahren relativiert wird, soll auch künftig auf dieses Bewertungsmerkmal nicht verzichtet werden. Zu dessen Umsetzung wird am Archivsachenvermerk durch den sachbearbeitenden Richter bzw. die Geschäftsstelle grundsätzlich festgehalten. Nur die Anwendung soll künftig auf eine deutlich engere und formalisiertere Grundlage gestellt werden, um Verfahren von ersichtlicher Bedeutsamkeit im Sinne eines deutlich strengeren Bewertungsmaßstabs und folglich in einer spürbar geringeren Zahl angeboten zu bekommen. Maßgeblich sollen dabei die folgenden Merkmale werden:

- Einschaltung der Pressestelle, mediale Berichterstattung (ggf. vermerkt werden, je nach Übung des jeweiligen Amtsgerichts, erteilte Auskünfte nach Art. 4 des Bayerischen Pressegesetzes),
- Verfahren mit besonderer Bedeutung aus rechtlicher Sicht (Schwierigkeiten rechtlicher Art, grundsätzliche Bedeutung der Rechtssache, vgl. entsprechend § 348 Abs. 3 Nr. 1 und 2 ZPO) oder in denen Landesrecht aus der Zeit vor 1900 angewendet wurde,
- Verfahren, die im Einzelfall nach Würdigung des Gerichts Besonderheiten aufweisen.

Durchaus kritisch zu hinterfragen ist das weit verbreitete und auch bei den Zivilprozessakten angewendete Bewertungsmerkmal der Prominenz einzelner Prozessbeteiligter. Während die Evidenz archivischer Benutzung zweifellos ein dauerhaftes Interesse an prominenten Personen – wie auch immer dieses Attribut definiert wird – bestätigt, wird bei den Lebenssachverhalten, die vor dem Amtsgericht in Zivilsachen verhandelt werden, durchaus infrage zu stellen sein, ob hier Relevantes für biographische Studien zu erwarten ist.

Verfahren mit besonderer Bedeutung aus rechtlicher bzw. rechtshistorischer Sicht können nach aktuellem Kenntnisstand nur mittels Kennzeichnung durch den einzelnen sachbearbeitenden Richter ermittelt werden. Kein zuverlässiges Kriterium für die Auswahl besonderer, gerade in rechtlicher Hinsicht bedeutsamer Verfahren bietet jedenfalls der Abgleich mit der Datenbank „Bayern.Recht“. Nach einer Auskunft des Bayerischen Verwaltungsgerichtshofes werden Urteile und Beschlüsse dort vermehrt veröffentlicht, um den Anspruch der Allgemeinheit auf Zugänglichkeit zu Gerichtsurteilen zu bedienen. Eine Beschränkung auf rechtlich oder tatsächlich bedeutsame Verfahren finde vor diesem Hintergrund nicht mehr statt.

Die als ersichtlich bedeutsam gekennzeichneten Verfahren sollen im Rahmen der Stichprobenziehung selbst Berücksichtigung finden und als eigene Schicht mit einer Auswahlwahrscheinlichkeit von 1 selbst Teil der gesamten Stichprobe des jeweiligen Jahrgangs übernommen werden. Voraussetzung hierfür ist, dass sich die oben artikulierte Erwartung hinsichtlich eines im Gegensatz zur heutigen Praxis strengeren Maßstabes bei der Anbringung des Archivsachenvermerks durch das Amtsgericht erfüllt, was die Anzahl der als besonders gekennzeichneten Verfahren erheblich reduzieren sollte. Es wird sich zeigen, ob sich diese Erwartungshaltung praktisch realisiert. Nicht nur insofern empfiehlt sich eine gewisse Offenheit hinsichtlich der genauen Zusammensetzung der Stichprobe, unter der Voraussetzung einer hinreichenden Dokumentation des Auswahlverfahrens. Flexibel und reagibel zu bleiben, um auf Unerwartetes reagieren zu können und vorhandenes Sonderwissen aus dem Tagesgeschäft der Rechtsprechung nicht zu verschenken, ist auch im Rahmen eines teilautomatisierten Verfahrens geboten.

VIII. Technische und organisatorische Umsetzung

Eine Operationalisierung der gewonnenen Erkenntnisse erfordert technische und organisatorische Anpassungen, und zwar sowohl auf Seiten der Archive, als auch auf Seiten der Justiz. Im Zentrum steht dabei die Verfügbarkeit eines IT-Werkzeugs, das die Stichprobenziehung anhand von Daten aus forumStAR ermöglicht. Dieses Werkzeug entsteht bis Mai 2024 im Rahmen eines Consulting-Projekts zweier Studierender der Statistik an der LMU in Form einer Shiny-App¹². Auch wenn dabei von den im neuen Bewertungsmodell definierten Schichten ausgegangen wird, gehört die Anpassungsfähigkeit an mögliche neue archivfachliche Bedarfe, etwa im Sinne veränderter oder neuer Strata, zu den Kernanforderungen an das Tool. Flankierend arbeitet eine dritte Studierende methodische Fragen archivischer Stichprobenverfahren in ihrer Masterarbeit auf, die das nachhaltige Wissensmanagement im Rahmen der Aus- und Fortbildung auf archivischer Seite unterstützen soll. Offen ist noch die technische Umgebung, in der das Verfahren künftig zu Anwendung kommen soll, wobei die bisher genutzte Cloud-Umgebung als Prototyp einer künftigen bayerischen Behörden-Cloud den sicheren Datentransfer zwischen Justiz und Archiven erleichtern könnte. Alternativ ist eine Implementierung in die eigene Infrastruktur der Staatlichen Archive Bayerns möglich.

Die Bewertung selbst anhand eines gesamt-bayerischen Datensatzes wird künftig zweckmäßigerweise zentral in der Generaldirektion der Staatlichen Archive Bayerns erfolgen. Dazu werden künftig regelmäßige Datenexporte aus forumSTAR durch die Justiz erforderlich, wofür noch keine Routine besteht. Gleiches gilt für die weiteren Schritte des Bewertungsprozesses, der die Übermittlung des Bewertungsergebnisses und dessen korrekte Verarbeitung bis hin zur konkreten Übergabe der ausgewählten Akten durch die verwahrenden Gerichte an die zuständigen Archive nach dem Ende der Aufbewahrungsfrist umfasst. Zwischen der eigentlichen Bewertung und diesem letzten logischen Schritt werden voraussichtlich einige Jahrzehnte liegen.

Die Bewertungsgrundsätze und das Bewertungsverfahren weichen damit perspektivisch deutlich ab von der bisherigen Norm. Die Anpassung der Aussonderungsbeachtlichmachung Justiz ist daher dringend geboten; ein formaler Anlass dazu wird nach der aktuell anstehenden Novellierung des Bayerischen Archivgesetzes gegeben sein.

Nutzerinnen und Nutzer werden aber nicht nur die normative Grundlage der Überlieferungsbildung kennenlernen, sondern auch im Einzelfall nachvollziehen können wollen, ob eine Akte als Teil einer Stichprobe oder aufgrund einer individuellen Einzelentscheidung per Archivsachenvermerk im Archiv vorliegt. Auch die archivische Erschließung wird somit anzupassen sein, um die methodische Sorgfalt bei der Überlieferungsbildung auch für die Forschung unmittelbar nutzbar zu machen.

12 Eine Shiny-App ist eine webbasierte Anwendung, die es ermöglicht, interaktive und benutzerfreundliche Datenvisualisierungen und Analysewerkzeuge mit R zu erstellen und zu teilen, ohne dass Nutzer Programmierkenntnisse benötigen.

IX. Ausblick

Über diese teils noch zu konkretisierenden Anpassungsbedarfe hinaus sind bereits jetzt weitergehende Fragen an die Zukunft zu stellen. Dies betrifft zunächst die dauerhafte Verfügbarkeit einer geeigneten Datengrundlage, auch über das bereits jetzt absehbare Anwendungsende des Fachverfahrens forumSTAR hinaus. Dabei ist anzunehmen, dass auch das bereits bekannte Nachfolgesystem GeFA (Gemeinsames Fachverfahren für die Justiz) identische, jedenfalls ähnliche und brauchbare Daten wird liefern können¹³. Die Bewertung anhand von Fachverfahrensdaten wird dabei zu harmonisieren sein mit einem standardisierten Aussonderungsprozess für elektronische Verfahrensakte der Justiz gemäß dem einschlägigen XÖV-Standard XJustiz¹⁴, dessen Spezifikation derzeit zwischen Fachgremien der Justiz und der Archivverwaltungen des Bundes und der Länder in seinen Grundzügen abgestimmt wird. Nicht auszuschließen ist, dass es geboten oder zweckmäßig sein kann, als Datengrundlage für die Bewertung anstelle von CSV-Daten aus forumSTAR oder GeFa künftig die XML-Daten gemäß XJustiz zu nutzen. Die seit 2016 in der bayerischen Justiz von einer stetig wachsenden Zahl an Gerichten eingesetzte elektronische Akte eröffnet zudem in technischer Hinsicht eine bisher ungeahnte Möglichkeit¹⁵, neben den Meta- auch die Primärdaten für Bewertungszwecke datenanalytisch zu nutzen. Der Einsatz von KI könnte gerade dabei hilfreich sein. Inwieweit die rechtlichen Voraussetzungen dazu vorhanden sein werden, soll hier nicht weiter erörtert werden.

Weit unmittelbarer stellt sich die Frage, wie die erprobte Bewertungsmethode auf der Grundlage angewandter Datenanalyse über das gewählte Fallbeispiel der Zivilprozessakten der Amtsgerichte auf weitere Aktengruppen ausgeweitet werden kann, sei es aus dem Zuständigkeitsbereich der Justiz, sei es darüber hinaus. Eine hinreichend umfängliche Grundgesamtheit an Akten und aussagekräftigen Metadaten dazu auf Seiten einer anbietepflichtigen Stelle vorausgesetzt, können ähnliche Lösungen zur Auswahlarchivierung erwogen werden. Automatismen sind dabei abzulehnen, entscheidend wird immer die jeweilige Zielsetzung in der Überlieferungsbildung sein, von der sich methodische Entscheidungen ableiten.

13 Weiterführende Hinweise im Internet unter <https://gefa-justiz.de>, zuletzt abgerufen am 01.05.2024.

14 Vgl. Ernst, Standards und Normen im Bereich der Langzeitarchivierung. In: *Archivar* 2/74 (2021), S. 62–70, hier S. 66. – Spezifikation XJustiz 3.4.1 Elektronische Datenübermittlung in der Justiz (Fassung 03.03.2023), hrsg. von der Bund-Länder-Kommission für Informationstechnik in der Justiz (BLK), S. 405–416, abrufbar unter: https://xjustiz.justiz.de/system/pdf/Spezifikation_XJustiz_3_4_1.pdf, zuletzt abgerufen am 06.05.2024.

15 Zum aktuellen Stand der E-Akte-Nutzung in der bayerischen Justiz siehe Bekanntmachung des Bayerischen Staatsministeriums der Justiz über die Elektronische Aktenführung bei den Gerichten der ordentlichen Gerichtsbarkeit und Staatsanwaltschaften vom 2. März 2020 (BayMBl. Nr. 119), die zuletzt durch Bekanntmachung vom 26. März 2024 (BayMBl. Nr. 168) geändert worden ist.

X. Fazit

Die Anwendung von Methoden der Datenanalyse in der Bewertung ermöglicht nicht nur die Digitalisierung bisheriger Prozesse bei der Aussonderung, Anbietung und Übernahme archivwürdiger Unterlagen. Vielmehr wirken die Methoden zurück auf die jedem Bewertungsmodell zugrundeliegenden Überlieferungsziele, die ihrerseits immer von der praktischen Umsetzbarkeit und damit letztlich von der Schriftgutverwaltung der akten- bzw. datenproduzierenden Stellen beeinflusst sind. Die inhaltlichen, methodischen und prozesshaften Anpassungen sind in dem vorgestellten Beispiel so tiefgreifend, dass man tatsächlich von einer digitalen Transformation sprechen darf. Diese ist kein Selbstzweck, ebenso wenig wie die archivische Bewertung, deren Ergebnis immer eine möglichst auswertungsoffene Überlieferung sein soll. Im Fall der Zivilprozessakten bayerischer Amtsgerichte bedeutet dies, perspektivisch eine repräsentative Dokumentation gerichtlichen Handelns und der dadurch berührten Sachverhalte aufzubauen, die der Forschung dadurch andere und hoffentlich noch bessere Auswertungsmöglichkeiten eröffnen. Inwieweit dies gelingt, wird vorrangig von künftigen Nutzerinnen und Nutzern zu beurteilen sein.

Zusammenfassung: Staatliche Archive bewerten ausgesonderte Unterlagen unter anderem der Justiz, übernehmen diejenigen mit einem bleibenden Wert in ihre Überlieferung und gewährleisten so die dauerhafte Verfügbarkeit einer Quellengrundlage für vielfältige Forschungen. Der Beitrag beschreibt ein innovatives Projekt der Staatlichen Archive Bayerns zum Einsatz von Methoden der angewandten Datenanalyse für die Bewertung von jährlich massenhaft anfallenden Zivilprozessakten der Amtsgerichte in Bayern. Grundlage dafür sind die Metadaten zu diesen Akten aus dem Fachverfahren forumSTAR der bayerischen Justizverwaltung. Die neu entwickelte Möglichkeit einer stratifizierten, repräsentativen Stichprobenziehung erlaubt zudem eine Veränderung des bisherigen Bewertungsmodells und soll damit das Potential künftiger Überlieferung für die Forschung erhöhen.

Summary: By appraisal of procedural documents state archives identify those of sufficient value to be accessioned. In this way long-term availability of these historical sources for scientific research is ensured. The paper describes an innovative project of the Bavarian States Archives using methods of Data Science for the appraisal of files from civil procedures at local courts (Amtsgerichte) in Bavaria. Descriptive metadata from the IT application forumSTAR, used by bavarian judicial administration, provide the basis of this project. The new chance to build a representative and stratified sample allows to adapt the present principles of appraisal. Thus further archival fonds shall offer greater potential for scientific research.



© Andreas Nestl und Michael Unger