

Transferarbeit im Rahmen der Laufbahnprüfung für den Höheren Archivdienst an der
Archivschule Marburg (57. Wissenschaftlicher Lehrgang)

Maschinelle Methoden in der Bewertung von Zivilprozessakten
—
Aktueller Stand und Ausblick

Dr. Marius Luszek

Betreuer Landesarchiv Baden-Württemberg: Dr. Kai Naumann
Betreuer Archivschule Marburg: Dr. Dominik Haffer

Abgabedatum: 02.04.2024

Inhaltsverzeichnis

1. Einleitung.....	1
2. Metadaten und ihre Bereitstellung durch die Justiz.....	4
3. Modelle und Methoden.....	6
3.1 Bewertungsmodelle.....	6
3.2 ‚Manuelle‘ Methoden.....	9
3.3 Maschinelle Methoden.....	13
3.4 Management von Modellen und Methoden.....	19
4. Zukunftserwartungen.....	20
4.1 Landesarchive: Standardisierung, Vereinheitlichung, technische Innovation.....	20
4.2 Internationale Forschung: technische Innovation.....	23
5. Ergebnisse.....	27
6. Zusammenfassung.....	31
7. Anhang.....	32
7.1 Tabelle zu Merkmalen der Archivwürdigkeit bei Zivilprozessakten.....	32
7.2 Literaturverzeichnis.....	33

1. Einleitung

Könnte die Speicherung von Archivalien günstiger werden, als die Kosten und Komplexitäten der Bewertung? Werden Bewertungen künftig nur noch über große Aktengruppen erfolgen, mit Ausnahmen bei in ‚industriellem Ausmaß‘ produzierten Unterlagen? Sind fast alle Arbeitsschritte bei digitalem Archivgut automatisierbar? Und geht es bald nur noch darum, für den Nutzer das ‚Wichtige‘ hervorzuheben, statt auszuwählen und zu übernehmen/zu vernichten?¹

Diese Thesen berühren auch die Unterlagen der Justiz und im Besonderen die Zivilprozessakten, die das Thema dieser Arbeit sind. In ‚industriellem Ausmaß‘ entstehen sie zwar nicht; dennoch handelt es sich um sehr große Mengen. In großen Bundesländern fallen sehr viele Verfahren an, in Hessen 2022 etwa um die 89.000 Verfahren, und geschätzt 80% des Schriftgutes der Justiz entfallen auf die Zivilprozessakten.² Die Bewertung der Zivilprozessakten ist bisher auch sehr ressourcenintensiv – weshalb eine „vorausschauende, zielgerechte Überlieferungsbildung“ besonders notwendig erscheint.³ Diese erfolgt mithilfe von Bewertungsmodellen, über Registerzeichen statt Einzelakten, aber nicht als kategorische Übernahme oder Vernichtung. Die Mengenbegrenzung spielt immer noch eine Rolle und die Auswahl, nicht nur Hervorhebung, archivwürdiger Verfahren ist weiter zentral. Der Prozess der Auswahl ist das Thema dieser Arbeit, denn alle seine Teile verändern sich zum Verfassungszeitpunkt der Arbeit stark in den deutschen Landesarchiven. Digitale Massenakten benötigen und ermöglichen eine neue Art des archivarischen ‚sense making‘, bei der Computer und Mensch zusammenwirken.⁴ Dabei muss in ganzen Systemen gedacht werden, in „organized collections of hardware, software, supplies, people, policies and procedures“.⁵ Die Systeme, und besonders die Software, haben dabei die Funktion eines Trichters („funnel“) und erleichtern die Analyse, Reduktion und Priorisierung in der Masse der Unterlagen.⁶

Am Beginn des Verfahrens im Zivilprozess steht die Klageerhebung mit der Klageschrift, die den wesentlichen Sachverhalt darstellt und beteiligte Parteien, Streitwerte und Beweise enthält. Im weiteren Verfahren werden Schriftsätze und weitere Beweise eingereicht, Gutachten eingeholt und schließlich ein Vergleich erzielt oder ein

¹ Yeo, Future, S. 45-58.

² Zahlen für Hessen siehe Hessisches Statistisches Landesamt, Zivilgerichte (Tabelle BV15_j2022_Zivilgericht_Tabelle_1). Für die Schätzung siehe Haberkorn/Herge/Laube/Maaß/Petter/Zilien/Gniffke, Teilbewertungsmodell, S. 25.

³ Haberkorn/Herge/Laube/Maaß/Petter/Zilien/Gniffke, Teilbewertungsmodell, S. 3.

⁴ Bunn, Frames, S. 72-75.

⁵ Bunn, Frames, S. 69.

⁶ Bunn, AI, o.S.

Urteil gefällt.⁷ Gegebenenfalls wird das Verfahren an spezialisierte Gerichte weitergegeben oder in Berufung vor der nächsthöheren Instanz gegangen und dort wieder ein Urteil gefällt. Die Zivilprozessakten sind dabei ungeheuer vielfältig, denn im Zivilprozess werden „alle privatrechtlichen Streitigkeiten, die im alltäglichen Leben zwischen Privatpersonen auftreten können“ verhandelt.⁸ Die Zivilprozessakten haben daher vor allem einen Wert für die regionale, lokale und die Mikrogeschichte sowie die Alltagsgeschichte, da sie alltägliche Sachverhalte vor Gerichte mit relativ kleinen Gerichtsbezirken bringen. Die Fälle werden Sachgebietsschlüsseln zugeordnet, die einen „breiten Querschnitt unserer alltäglichen Lebenswelt“ zeigen: Bausachen, Verkehrsunfälle, Kaufsachen, Kredite und Leasing, Nachbarschaftssachen, Wohnmiet- und Wohneigentumssachen, gesellschaftsrechtliche Streitigkeiten etc.⁹

Welche Verfahren zu den Zivilprozessen gehören, hat sich in den letzten Jahren aber auch verändert. Diese Arbeit wird sich dabei auf die größten und wichtigsten Gruppen der Zivilprozessakten konzentrieren. Diese haben die Registerzeichen C (Amtsgerichte) bzw. O, OH (Landgerichte). Die weiteren Gruppen sowie die Zivilprozessakten der Oberlandesgerichte einzubeziehen, würde für den Rahmen dieser Arbeit zu komplex.¹⁰

Diese Transferarbeit schließt sich an die von Susanne Brenneisen an, die sich 2018 ebenfalls mit Zivilprozessakten beschäftigte. Sie fokussierte sich dabei auf die konkreten Auswahlmöglichkeiten des Landesarchivs Baden-Württemberg. Seither gab es in der Digitalisierung in der Justiz und hinsichtlich der für die Auswahl interessanten Methoden große Veränderungen. Diese Arbeit hat deshalb einen leicht anderen Ansatz. Sie wird sich auf die oben angesprochenen ‚Systeme‘ konzentrieren, also das Zusammenspiel der verschiedenen Komponenten und ihre Entwicklung. Dabei kommt der Technik, die immer bessere maschinelle Methoden ermöglicht, eine zentrale Rolle zu, um eine genauere und bessere Überlieferung zu bilden. Der Verwendung von digitalen Auswahlprozessen werden verschiedene Namen gegeben, etwa „technology assisted review“ oder „algorithmically assisted selection“;¹¹ diese Arbeit wird der Einfachheit halber von ‚manuellen‘ und ‚maschinellen‘ Methoden sprechen. Als maschinell wird für diese Arbeit definiert, dass ein Mensch in der Auswahl nicht mehr Einzelakten prüft, sondern nur noch Regeln formuliert, die von einer Maschine umgesetzt werden.

⁷ Brenneisen, Black box, S. 4-5.

⁸ Gehrke, Zivilprozessakten, S. 142-143.

⁹ Brenneisen, Black box, S. 5-6.

¹⁰ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024. Damit wird einer Empfehlung von Fr. Koch gefolgt, die viele Jahre Erfahrung in der Bewertung von Unterlagen der Justiz besitzt.

¹¹ Siehe etwa The National Archives, Application, S. 24-26 oder The National Archives, AI, S. 9.

Die Möglichkeiten und Begrenzungen neuer maschineller Methoden beeinflussen die Umsetzung und Reformulierung zu überarbeitender Bewertungsmodelle. Besonders interessant sind dabei die verschiedenen Stände der Einführung maschineller Methoden, wie auch verschiedene Ansätze an Modellen und Methoden. Die Arbeit will also die Interdependenzen herausarbeiten, Probleme beleuchten und den Gesamtprozess bewerten. Eine Folie bietet dafür der Kriterienkatalog zur Evaluierung von Bewertungsdokumenten des VdA,¹² der für Massenakten und maschinelle Methoden aber wenig umfangreich ist.

Da die Arbeit sich auf die aktuelle Praxis in den Archiven fokussiert, wurden Interviews als beste Methode zur Informationssammlung gewählt. Es wurden Interviews mit vier¹³ Landes-/Staatsarchiven geführt, die alle Erfahrungen mit der Zivilprozessakten-Bewertung haben, sich aber (noch) unterschiedlich stark auf maschinelle Methoden stützen. Das Landesarchiv Hessen und das Landesarchiv NRW sind auf dem Weg der Einführung maschineller Methoden. Das Landesarchiv Baden-Württemberg verfügt als Pionier bei der Einführung maschineller Methoden über langjährige Erfahrung und verfolgt den ‚exemplarischen‘ Ansatz zur gezielten Auswahl von Verfahren nach inhaltlichen Kriterien. Das Bayerische Staatsarchiv erarbeitet hingegen gerade einen ‚repräsentativen‘ Ansatz, der auf die statistische Samplebildung zielt. Zusätzlich wurden zwei Interviews zu KI/Data Science geführt, die die Chancen dieser Technologien/Methodiken erörtern: mit Jenny Bunn von The National Archives in Großbritannien und Basma Makhlof Shabou von der Universität Genf.

Die Arbeit ist dementsprechend wie folgt aufgebaut. Zuerst wird die Justiz betrachtet, die Akten und/oder die Daten bereitstellt. Danach werden die Bewertungsmodelle der Archive verglichen, gefolgt von der Untersuchung ihrer Nutzung von manuellen und maschinellen Methoden. Diesen Block schließt ein Kapitel zum Management von Modellen und Methoden ab. In einem zweiten Block werden die Zukunftserwartungen der Landesarchive sowohl für die Justiz, ihre eigene Praxis und die technischen Neuerungen untersucht. Daran schließt sich die Diskussion der Forscherinterviews an, die technisch noch ‚weiter‘ in die Zukunft sehen. Abschließend werden die Ergebnisse der Arbeit diskutiert.

¹² Arbeitskreis Archivische Bewertung im VdA, Kriterien, S. 18-33.

¹³ Das Sächsische Staatsarchiv, ein Pionier der Verwendung maschineller Verfahren, konnte leider keine Zeit für ein Interview entbehren.

2. Metadaten und ihre Bereitstellung durch die Justiz

Die Justiz erstellt in ihren Fachverfahren die für ihre Arbeit notwendigen Metadaten und stellt sie den Archiven bereit. Die Eigenheiten der Metadaten und ihrer Bereitstellung bestimmen wesentlich, welche Werkzeuge zur Bewertung genutzt werden können. Dieses Kapitel stellt daher den derzeitigen Stand dar, auf dessen Basis die Archive arbeiten können, und beschreibt die derzeitigen Schwachpunkte in diesem Bereich.

Die wichtigsten Metadaten werden zunächst tabellarisch dargestellt:

Metadatum	Beispiele	Kommentar
Registerzeichen	C, H, O, OH	Diese Arbeit konzentriert sich auf die häufigsten Registerzeichen C, O, OH.
Bezeichnung der Sache	X gegen Y	Aufstellung der Parteien, Freitext
„wegen“	Forderung, Schadensersatz, Feststellung, Unterlassung	Freitextfeld, hier extrem viele Varianten; am häufigsten sind Forderung, Schadensersatz und leer
Verfahrensart	Klageverfahren, Arrest/einstweilige Verfügung sonstige Zivilsache, selbständiges Beweisverfahren	7 Varianten, Klageverfahren sind bei weitem am häufigsten
Sachgebietsschlüssel	Miet-, Kredit- und Leasingsachen, Verkehrsunfallsachen, Wohnungsmietsachen, sonstiger Verfahrensgegenstand	18 Sachgebiete, ‚sonstiger Verfahrensgegenstand‘ ist am häufigsten
Wert	Streitwert in Euro	Ab 5000 Euro beim LG
Erledigungsart	Versäumnisurteil, Gerichtlicher Vergleich, Urteil, Rücknahme Klage/Antrag, Beschluss nach § 91 ZPO, Arrest oder einstweilige Verfügung, Abgabe an anderes Gericht	29 Varianten, am häufigsten sind Versäumnisurteil, Vergleich, Urteil, Rücknahmen
Bewertungsvorschlag Gericht	„anzubieten“, leer	Akte wird vom Gericht gekennzeichnet

Tabelle 1: Merkmale für Archivwürdigkeit¹⁴

Allgemein besteht das Problem, dass sich Gerichte und Archive für unterschiedliche Metadaten interessieren. Während die Archivare mehr Details für die Bewertung wünschen, organisiert die Justiz ihre Akten vor allem nach übergeordneten Sachgebieten und Begriffen, was die analoge wie die digitale Bewertung nicht vereinfacht.¹⁵

¹⁴ Tabelle vom Autor erstellt. Ich danke hier Frau Koch vom Landesarchiv Baden-Württemberg, für Einblicke in die derzeit typischen Metadaten in Aussonderungslisten, wie sie das Landesarchiv erhält. Das zugrundeliegende Fachverfahren ist Forumstar. Die weiteren Informationen basieren auf tabellarischen Analysen von Beispieldaten durch das Landesarchiv, etwa die Häufigkeit bestimmter Begriffe. In anderen Archiven, die Daten aus anderen Fachverfahren erhalten sind die Metadaten gleich, ggf. sind die Felder anders benannt.

¹⁵ Interview Gillner (Landesarchiv NRW) am 15.01.2024.

Gerade das letzte in der Tabelle genannte Metadatum (Bewertungsvorschlag war daher im analogen Zeitalter für die Archive entscheidend, wie auch später noch erläutert wird. Auf Seiten des Gerichts ergibt sich dabei jedoch ein besonderer Aufwand für die Richter/Mitarbeiter. Das Archiv mit zu bedenken bedeutet schlicht zusätzlichen Arbeitsaufwand und dessen Kriterien für archivwürdige Akten sind oft abstrakt und umfangreich (siehe das nächste Kapitel) – entsprechend selten wird dieses Feld angekreuzt.¹⁶ Aber auch die Bereitstellung allgemeiner Anbietungslisten ist noch nicht problemlos möglich. So sind die Mitarbeiter teils nicht ausreichend geschult und können sie trotz Vorhandenseins solcher Funktionen oft nicht direkt aus dem System erzeugen. Dadurch erhalten die Archive verschieden ausgestaltete Listen von den Gerichten, die schwieriger zu verarbeiten sind. Eine flächendeckende maschinelle Verarbeitung scheidet bei solchen Gegebenheiten aus, da die Anpassung jeder individuellen Liste zu aufwendig wäre.¹⁷ Allgemein wird ein hoher Beratungsbedarf der Justiz konstatiert, der aber in den letzten Jahren nicht unbedingt Früchte trug.¹⁸

Eine besondere Herausforderung ist auch, dass Schreibweisen in wichtigen Feldern oder auch die Schreibweise von Namen nicht normiert sind. So ist etwa ‚wegen‘ ein Freitextfeld, wodurch verschiedene Mitarbeiter Sachverhalte verschieden schreiben und Varianten erzeugen, die in der Bewertung wieder zusammengeführt werden müssen.¹⁹

Wenn Exporte möglich sind, muss oft auf von den Archiven gewünschte Metadaten (d.i. Felder) im Fachverfahren verzichtet werden. Grund dafür ist entweder das Fehlen dieser Felder (bei Zivilprozessakten etwa Geburtsdaten) oder von Schnittstellen. Schnittstellen hinzuzufügen bedeutet zusätzliche Kosten, für die keine Mittel vorhanden sind.²⁰

Diese Problematiken haben dann auch Auswirkungen auf die inhaltliche Bewertung, da archivwürdige Akten mithilfe der Metadaten oftmals kaum zu erkennen sind.²¹ Stattdessen ist die Kenntnis des Akteninhalts notwendig, mithilfe dessen die diversen, weit gefassten Begriffe (siehe Tabelle 1) konkretisiert werden können und die eindeutige Identifizierung von Personen vereinfacht wird.²² Hier handelt es sich um eine

¹⁶ Siehe hierzu etwa Brenneisen, Black box, S. 11-12. Diese Thematik wird auch im Kapitel zu Kennzeichnung durch die Gerichte und Autopsie genauer besprochen.

¹⁷ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024 ähnlich Interview Gillner (Landesarchiv NRW) am 15.01.2024. Eine wesentliche Hoffnung sind übergeordnete Stellen für die IT in der Justiz, die zusätzlich schulen könnten.

¹⁸ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024, ebenso Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

¹⁹ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

²⁰ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

²¹ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024. In der Diskussion ergab sich immer wieder das Problem, dass von der Justiz erörterte inhaltliche Merkmale archivwürdiger Akten von den Archivaren anhand der gegebenen Metadaten für nicht identifizierbar gehalten wurden.

²² So zumindest Brenneisen, Black box, S. 12 im Jahr 2018. Der ‚Aktentitel‘ bzw. die korrespondierenden Metadaten reichten für die Ermittlung etwa von ‚zeittypischen‘ oder besonderen Einzelfällen nicht aus.

besondere ‚Schwäche‘ der Zivilprozessakten gegenüber den Strafprozessakten, bei denen Geburtsdaten aufgrund ihrer Relevanz für das Strafmaß immer vorhanden sind und IDs den Tatbestand nach Paragraphen des Strafgesetzbuches genauer widerspiegeln.²³ Dementsprechend wurden schon früh (2018) Hoffnungen bezüglich der Nutzung des Volltexts der Zivilprozessakten geäußert.²⁴

Andererseits wird die Bereitstellung digitaler Metadaten auch positiv bewertet. Die digitale Bereitstellung vereinfacht die Bewertung der Akten *aller* Gerichte, die aufgrund der in den folgenden Kapiteln noch beschriebenen Methodiken nun sehr schnell und (zu) einfach erfolgt.²⁵ Fehler, Unschärfen und fehlende Daten behindern die Bewertung, die Bereitstellung der Metadaten nach derzeitigem Stand wird trotzdem als großer Fortschritt gegenüber früheren Zeiten empfunden.²⁶

3. Modelle und Methoden

3.1 Bewertungsmodelle

Das vorige Kapitel wurde der Justiz und den von ihr bereitgestellten Metadaten gewidmet, weil diese die Bewertungsmethoden und möglicherweise auch die Bewertungsmodelle beeinflussen. Aber wie genau sind die Überlieferungsziele für Zivilprozessakten im Modell definiert? Welche Unterlagen sollen zu ihrer Erfüllung übernommen werden?

Ein Meilenstein für die Massenakten der Justiz (inklusive der Zivilprozessakten) war der Bericht der Stahlschmidt-Arbeitsgruppe von 1999.²⁷ Die Arbeitsgruppe formulierte jedoch nur grobe Empfehlungen für ganze Aktengruppen. Bei den Zivilprozessakten (Registerzeichen C, O) etwa eine „kleine Auswahl: herausragende und zeittypische Einzelfälle“ und „Samplebildungen nach den praktischen Möglichkeiten der Aktenüberlieferung“, die für die Praxis zu vage blieben.²⁸ Die Landesarchive erarbeiteten in der Folge konkretere Bewertungsmodelle, auch um sich auseinanderentwickelnde Auswahlkriterien einzelner Abteilungen stärker zusammenzuführen.²⁹

²³ Siehe hierzu etwa Brenneisen, Black box, S. 13, 23. Für die Rolle von IDs, die grob den Paragraphen des Strafgesetzbuches entsprechen siehe Interview Gillner (Landesarchiv NRW) am 15.01.2024.

²⁴ Brenneisen, Black Box, S. 25-26.

²⁵ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024. Frau Koch beschrieb etwa, dass durch die eher schlechte Datenbasis sich eine genauere Aufbereitung und Sichtung der Daten nicht lohnt; dementsprechend nimmt die Bewertung von Zivilprozessakten momentan viel weniger Zeit in Anspruch als von Strafprozessakten. Gleichzeitig sei der heutige Stand der Bewertung aber auch deutlich besser als früher.

²⁶ Interview Unger für das Bayerische Hauptstaatsarchiv am 27.02.2024.

²⁷ Stahlschmidt (Hg.), Empfehlungen.

²⁸ Stahlschmidt (Hg.), Empfehlungen, Punkt 13 sowie Punkt 312ab und b.

²⁹ Landesarchiv NRW, Abschlussbericht, S. 7; siehe für weitere Rekapitulationen der Entwicklungen auch Haberkorn/Herge/Laube/Maaß/Petter/Zilien/Gniffke, Teilbewertungsmodell, S. 3-4.

Bewertungsmodelle sind weiterhin mit das wichtigste Konzept für die gezielte Überlieferungsbildung und sollen „durch vereinheitlichte, koordinierte und fachlich überprüfbare Bewertungsentscheidungen [...] eine effektive Mengensteuerung“ ermöglichen und einen hohen Aussagewert der archivierten Unterlagen sichern.³⁰ Als wichtige Faktoren für ein gutes Bewertungsmodell definierte ein Arbeitskreis des VdA 2018 unter anderem eine Überlieferungsstrategie, ausformulierte Bewertungsziele und die Wahrung der Aktualität dieser Ziele (K 1 bis K 4). Außerdem müsse das Modell Methoden und Instrumente definieren, die konsequent angewendet werden (K 10).³¹

Dadurch sollten große Anbietungsmengen bewältigt werden, die getroffenen Bewertungsentscheidungen transparent und die Überlieferungsbildung insgesamt besser nachvollziehbar sein. Die Modelle selbst müssten evaluiert werden, besonders hinsichtlich der Umsetzbarkeit in der Praxis und der Passung zwischen Überlieferung und dem inhaltlichen Ziel der Überlieferungsbildung. Dabei beeinflussten alle Elemente (Strategie, Methoden) der Überlieferungsbildung einander.³² Aufgrund der Kürze der Arbeit kann hier kein detaillierter Vergleich der Bewertungsmodelle vorgenommen werden, siehe dazu die Tabelle im Anhang der Arbeit. Die Bewertungsmodelle weisen aber eine bemerkenswerte Ausdifferenzierung auf, wobei für einige zentrale Bereiche grundsätzliche Einigkeit besteht. Vor allem die Beteiligung berühmter Personen oder anderer besonderer Beteiligter, sowie regional- und zeittypisches sind für alle Archive wichtige Indikatoren. Stärker und schwächer ausdifferenzierte Modelle erfassen dabei vermutlich viele derselben Akten in verschiedenen Kategorien. Bedeutende Fälle für die freie Meinungsäußerung werden beispielsweise wahrscheinlich eine gewisse Medienresonanz erfahren, als zeittypisch auffallen oder ähnliches. Insgesamt sind die Modelle – von wenigen Besonderheiten abgesehen – als recht ähnlich zu bewerten.

Zu erwähnen ist allerdings, dass besonders Baden-Württemberg derzeit sein Modell überholt und hier auf in Überarbeitung befindliche Dokumente zurückgegriffen werden musste. Für Bayern lag nur die zuletzt 2013 aktualisierte Aussonderungsbekanntmachung Justiz vor – die in der praktischen Arbeit ebenfalls überholt ist.³³ Bayern stützt sich bisher vielmehr auf interne Schreiben der Generaldirektion an die Sprengelarchive, die im Zuge laufender Bewertungen häufig angepasst wurden. Das Modell befindet sich damit sozusagen eher in den Akten des Staatsarchivs.³⁴ Derzeit werden sowohl Modell als auch Methodik überarbeitet.

³⁰ So etwa für das Hessische Landesarchiv, siehe Gniffke, Masterplan, S. 3.

³¹ Arbeitskreis „Archivische Bewertung“ im VdA, Evaluierung, K 1 bis K 4, K 10.

³² Arbeitskreis „Archivische Bewertung“ im VdA, Evaluierung, S. 18-20, Zitat S. 20.

³³ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024, Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

³⁴ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

Hinter den Merkmalen für Archivwürdigkeit stehen idealerweise klar definierte Überlieferungsziele. Diese inhaltlichen Ziele sollten so formuliert werden, dass sie weiter eingegrenzt bzw. konkretisiert werden können. Je besser ein Ziel definiert ist, umso einfacher kann es in die Bewertungsentscheidungen einfließen.³⁵

Für die Zivilprozessakten an sich bestehen nun aber Unterschiede zwischen den Landesarchiven.

Hessen und NRW formulieren ‚nur‘ Ziele für die Gesamtüberlieferung der Justiz in ihren Handreichungen: das Abbilden des Handelns der Justizverwaltung, die Abbildung gesellschaftlicher Entwicklungen und zeit- und regionaltypischer Phänomene oder Unterlagen mit bleibendem Wert für Wissenschaft, Gesetzgebung, Regierung, Verwaltung, Justiz.³⁶ Baden-Württemberg beschreitet in einer Handreichung bezüglich der Zivilprozessakten den gleichen Ansatz wie NRW und bezieht sich auf Unterlagen mit historischem Wert in vielerlei Hinsicht.³⁷ Bayern hat hingegen ein recht genaues, übergeordnetes und nur auf die Zivilprozessakten bezogenes Ziel: die Dokumentation der Alltagsgeschichte.³⁸

Die vorgestellten Modelle scheinen die Evaluationskriterien gut zu erfüllen: die Modelle beschreiben Strategien und Bewertungsziele (K 1 und K 2), mit leichten inhaltlichen Unterschieden und teils auf die Justizüberlieferung allgemein bezogen, teils auf die Zivilprozessakten im Besonderen. Die Modelle sind gut ausgearbeitet, ausgerichtet sind sie bisher vor allem noch auf die Identifikation besonderer Fälle, weniger auf ‚normale‘ Fälle. Letzteres, und damit auch der Evidenzwert, wird eher nebenbei mit abgedeckt, am augenscheinlichsten in der Samplebildung nur im Sachgebiet der Wohneigentumsverfahren in Hessen.³⁹

Die Modelle definieren grundsätzlich auch recht gut die Methoden und Instrumente (Einzelanforderungen und Kennzeichnung durch die Gerichte), die konsequent angewendet werden sollen (K 10). Methoden und Instrumente sollen angemessen sein sowie verständlich, anschaulich beschrieben, nachvollziehbar und plausibel sein (K 11, K 12). Die Modelle beschreiben auch die Aufgaben des Bestandsbildners (K 15) – die in vielen Fällen eine Vorauswahl anhand der Merkmale im Modell treffen.⁴⁰

³⁵ Wendenburg, Prüfen, S. 44-45, Zitat S. 44. Wendenburg bezieht sich auf K 35 aus dem Evaluierungskatalog in Arbeitskreis „Archivische Bewertung“ im VdA, Evaluierung.

³⁶ Vgl. Haberkorn/Herge/Laube/Maaß/Petter/Zilien/Gniffke, Teilbewertungsmodell, S. 6, sowie Landesarchiv NRW, Richtlinien, S. 8.

³⁷ Landesarchiv Baden-Württemberg, Handreichung Archivwürdige Zivilakten.

³⁸ Dieses Ziel ist in der Aussonderungsbekanntmachung Justiz allerdings nicht zu finden, sondern wurde nur im Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024 mitgeteilt.

³⁹ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024. Ähnlich auch Interview Gillner (Landesarchiv NRW) am 15.01.2024.

⁴⁰ Arbeitskreis „Archivische Bewertung“ im VdA, Evaluierung, siehe die jeweiligen Unterpunkte.

Die Gerichte sollen das Modell verstehen und umsetzen. Probleme damit wurden im vorigen Kapitel schon angedeutet. Der Wunsch des Archivs nach Kennzeichnung ist nicht nur zusätzliche Arbeit. Die Modelle geben den Gerichten leider auch vor allem unbestimmte Begriffe an die Hand, die Unschärfen in die Vorgehensweise (K 11, K 12) einbringen. Diese relative Willkür in der Überlieferung (subjektiv interpretiertes ‚Besonderes‘) beeinträchtigt dann aus Archivperspektive auch den Wert der gebildeten Überlieferung etwa für die Alltags- und Regionalgeschichte.⁴¹ Die Stahlschmidt-Arbeitsgruppe konstatierte bereits 1999, dass die gegebenen Kriterienkataloge sich nicht bewährt hätten und schlug einfache, „mechanisch anwendbare Kriterien“ vor.⁴² Zwar folgten die Archive dem nicht, wie die Besprechung der Modelle gezeigt hat.⁴³ Aber auch heute sind sie sich der eher nur annäherungsweisen Umsetzung des Modells bewusst – es liefert Orientierung, bleibt aber abstrakt.⁴⁴ Das nächste Kapitel beschäftigt sich mit dieser ‚manuellen‘ Methodik, der Kennzeichnung von Akten durch die Gerichte und der Autopsie durch die Archivare sowie deren Stellenwert in der bisherigen Methodik der verschiedenen Archive.

3.2 ‚Manuelle‘ Methoden

Dieses Kapitel beschäftigt sich nun mit der Rolle ‚manueller‘ Methoden in der Praxis: Kennzeichnung von Akten durch die Gerichte in ihren Fachverfahren, Autopsie der Archivare ohne komplexere technische Hilfsmittel. Der Stellenwert der Kennzeichnung durch die Gerichte unterscheidet sich zwischen den Archiven, vor allem abhängig davon, wie stark datengetriebene, maschinelle Methoden bereits Einzug gehalten haben. Alle interviewten Archive sind derzeit in der Entwicklung datengetriebenen-maschineller Methoden; manuelle Methoden verlieren damit langsam an Bedeutung. Besonders in Baden-Württemberg werden maschinelle Methoden schon länger auf Zivilprozessakten angewendet. Die Besprechung manueller Methoden ist jedoch wichtig zum Verständnis der Motivation für den Umstieg auf maschinelle.

Ein wesentliches Merkmal der bisherigen ‚manuellen‘ Methoden ist die Beschränkung auf Vorauswahlen durch die Gerichte. Dabei verzichten die Archive auf die Anbieten aller Akten und erhalten stattdessen Listen, in die nur die vom Gericht gekennzeichneten Akten eingetragen sind. Ob vollständige Aussonderungslisten oder Übersichten über die Gesamtzahl der Prozesse auf Anforderung zu Verfügung gestellt werden könnten, ist dabei nicht bekannt. Auch Mitteilungen über die

⁴¹ So besonders Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

⁴² Stahlschmidt (Hrsg.), Empfehlungen, S. 9-10.

⁴³ Mit wenigen Ausnahmen wie den Wohneigentumsverfahren in Hessen (die ersten 5 pro Jahr) oder Streitwerten über 500.000 Euro in NRW.

⁴⁴ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

Gesamtvernichtungsmengen erhalten die Archive nicht, sodass sie über die Zivilprozessakten abseits der Auswahl durch die Gerichte keine Informationen haben.⁴⁵ Die übergebene Liste wird dann der Autopsie unterzogen, also eine Einzelentscheidung über die Archivwürdigkeit getroffen.

Das Landesarchiv Hessen konstatierte etwa, dass es früher „recht großzügig“ gewesen sei, wegen dem oft ungünstigen Verhältnis von Aktenumfang und -inhalt nun aber weniger Verfahren „in Gänze archivieren“ wolle.⁴⁶ Dabei müsse man sich jedoch weiter nach Einschätzungen der Richter sowie Hinweisen durch die Pressestellen der Landgerichte richten. Auch in NRW war das oben skizzierte Verfahren bisher Standard, mit einer Autopsie der von den Gerichten gekennzeichneten Fällen und solchen, die zum BGH kamen oder in Pressemitteilungen erwähnt wurden.⁴⁷

Da die Archive von den Gerichten abhängig sind, wiegt umso schwerer, dass eine große Bandbreite in der Vorgehensweise der Gerichte vorliegt. Wie viel (gar nicht, wenig, zu viel) gekennzeichnet wird, unterscheidet sich von Gericht zu Gericht; manche Gerichte übersenden Datenbankauszüge, manche erstellen manuell Listen.⁴⁸ Allgemein wird jedoch extrem wenig gekennzeichnet; Erinnerungen und Ermutigungen hatten bisher nicht den gewünschten Effekt. Einzelne, bekannte Aktenzeichen fügen die Archivare daher auch selbst noch hinzu.⁴⁹

Eine weitere Möglichkeit der zielgerichteten Reduktion ohne maschinelle Methoden ist die Auswahlarchivierung von nur bestimmten Provenienzen. Eine gewisse Prominenz erlangte hier das baden-württembergische Modell von 1999, das die Konzentration auf vier Gerichte vorsah. Die Städte wurden dabei nach Charakteristika wie Großstadt, Mittelstadt, ländliche Stadt und industrieller Prägung, Universitätsstadt und ländlicher Prägung ausgewählt.⁵⁰ Im Interview wurde dieses Modell jedoch als veraltet beschrieben;⁵¹ die Konzentration auf wenige Gerichte wird allgemein auch nicht mehr verfolgt. Hessen etwa rechnet bei einem größeren Bundesland mit mehr Akten, was aber nicht zu einer größeren Übernahmemenge führen muss. Die Zahl besonderer Einzelfälle korreliert nicht so stark mit der Gesamtzahl der Akten und auch für die Abdeckung des Evidenzwertes reiche eine gewisse Zahl. Da die Überlieferungsbildung möglichst breit aus dem ganzen Land erfolgen solle, sollte auch keine Stelle ausgeschlossen werden, zumal die Zahl der Gerichte nicht so hoch sei. Die Übernahmen könnten dann mit

⁴⁵ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024. Bezüglich der Angewiesenheit auf die Gerichte genauso Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

⁴⁶ Haberkorn/Herge/Laube/Maaß/Petter/Zilien/Gniffke, Teilbewertungsmodell, S. 27.

⁴⁷ Interview Gillner (Landesarchiv NRW) am 15.01.2024.

⁴⁸ Interview Gillner (Landesarchiv NRW) am 15.01.2024.

⁴⁹ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

⁵⁰ Landesarchiv Baden-Württemberg, Erweitertes Auswahlmodell, S. 3 und 6.

⁵¹ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

Registraturanalysen kontrastiert werden.⁵² Die Auswahlarchivierung zielt stattdessen einerseits auf bestimmte Kammern bestimmter Landgerichte, die besonders interessante Sachverhalte oder solche bezüglich Aktiengesetz, Wertpapierbereinigung (v.a. Aktenzeichen OH) verhandeln. Diese Vorgehensweise passt gut zur Ausrichtung der Modelle auf besondere Einzelfälle. Andererseits verfolgt das Landesarchiv NRW noch den Ansatz, zeit- und regionaltypische Fälle von einer Auswahl von acht ausgewählten Landgerichten zu übernehmen.⁵³ Diese Formen der Auswahlarchivierung reduzieren die Übernahmemenge auf einfache Weise, für die Bewertung von Massenakten normalerweise ein wichtiger Faktor. Dies war jedoch aufgrund der geringen Anzahl von den Gerichten gekennzeichneten Akten bisher nicht von Bedeutung.⁵⁴

Die abstrakten Archivwürdigkeitskriterien und der Mangel performanter Instrumente zum Umgang mit der großen Masse der Zivilprozessakten führten zum ‚Verlass‘ auf die Gerichte in der Überlieferungsbildung. Diese Angewiesenheit darauf, dass die Gerichte praktisch eine (Vor-)Bewertung vornehmen, stellte und stellt weiterhin das Hauptproblem ‚manueller‘ Methoden dar.

Die ehemalige baden-württembergische Idee, sich auf Gerichte in vier Städten zu beschränken, führte etwa nicht zum Erfolg. Das Modell sah die Abgabe aller herausragenden Fälle bei allen Gerichten vor und bei bestimmten Aktengruppen in jedem Fall die Übernahme der Akten eines bestimmten Arbeitsgerichts. In der Praxis der letzten ca. zwei Jahrzehnte führte das jedoch dazu, dass die anderen Gerichte keinerlei Akten abgaben. Die Begründung war dabei, dass sie keine herausragenden Fälle hätten. Dadurch wurden praktisch nur noch Akten dieses einen Amtsgerichts übernommen, wodurch eine fachlich angemessene Abbildung der Justiz des Landes nicht erreicht wurde.⁵⁵

Befürworter manueller Methoden sehen in ihnen noch immer den Weg zur Identifikation wichtiger Einzelfälle, während eine Samplebildung und deren Repräsentativität aufgrund von Justizstatistiken weniger wichtig werde. Auch die Bewertungsmodelle zielten ja auf die Ermittlung dieser besonderen Akten.⁵⁶ Das Aufkommen maschineller Methoden erlaubt nun, die Rolle der Kennzeichnung durch

⁵² Interview Gniffke (Landesarchiv Hessen) am 12.01.2024. Als Gegenbeispiel wurde die Polizei genannt, bei der wegen ihrer starken Verästelung nur von bestimmten Stellen übernommen wird.

⁵³ Interview Gillner (Landesarchiv NRW) am 15.01.2024. Siehe Landesarchiv NRW, Richtlinien, S. 46-47.

⁵⁴ Interview Gillner (Landesarchiv NRW) am 15.01.2024. Die Vorgabe von 1% maximaler Gesamtübernahmemenge, für die das Landesarchiv NRW bekannt ist, spielte daher bei den Zivilprozessakten keine Rolle.

⁵⁵ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024. Siehe auch Brenneisen, Black box, S. 8 sowie Haberkorn/Herge/Laube/Maaß/Petter/Zilien/Gniffke, Teilbewertungsmodell, S. 25. Diese drei zusammengenommen zeigen, dass bereits seit den 1970er Jahren die Problematik nicht ausreichender Kennzeichnung bekannt war.

⁵⁶ Brenneisen, Black box, S. 10.

die Gerichte neu zu definieren und in die besagte Richtung (Einzelfallidentifikation) weiterzuentwickeln. Bisher half sie, fehlende Expertise der Archivare in rechtlichen Themen auszugleichen und war notwendig für den Umgang mit den enormen Mengen an Zivilprozessakten.⁵⁷ Die Expertise des Gerichts und insbesondere die inhaltliche Vertrautheit des Richters mit ‚seinem‘ Fall, sind schwer zu ersetzen. Teils wird dies mit Datenbankabgleichen etwa zur Identifikation von Personen versucht (siehe nächstes Kapitel). Ein besonderer Fall könnte einerseits durch die Involvierung bekannter Personen identifiziert werden, andererseits bestreiten auch bekannte Persönlichkeiten wenig interessante Fälle. Insofern begründet ein Abgleich eher eine Annahme der Archivwürdigkeit als ihr tatsächliches Vorliegen. Die Metadaten eines Falls können ggf. auch einen falschen Eindruck vermitteln, etwa ein hoher Streitwert im Mietrecht.⁵⁸ Hier ist die Expertise des Gerichts wertvoll, tatsächlich besondere Fälle zu ermitteln; besonders wenn keine Quellen wie eine Landesrechtsprechungsdatenbank oder die JURIS-Datenbank vorliegen, mit denen aus rechtswissenschaftlicher Sicht bedeutende Fälle über die Aktenzeichen identifiziert werden können. Allgemein herrscht die Erwartung vor, dass die Zahl und Wichtigkeit der Kennzeichnung durch die Gerichte weiter abnehmen wird.⁵⁹

Insbesondere für Archive und Methoden, die nicht auf Datenbankabgleiche zurückgreifen (wollen), bleibt die Kennzeichnung durch die Gerichte relevant. Unter den interviewten Archiven war dies vor allem das Bayerische Staatsarchiv. Dieses will mit den Zivilprozessakten Alltags- und Regionalgeschichte überliefern – die bisherige Methodik, besondere Fälle durch die Gerichte kennzeichnen zu lassen, ist dafür weniger passend. Auch Datenbankabgleiche, die in die gleiche Richtung zielen, verfolgt das Staatsarchiv aus diesem Grunde nicht. Das Fehlen besonderer Fälle sei kein Problem, da das Überlieferungsziel explizit nicht auf das ‚Besondere‘ ziele.⁶⁰

Rückgebunden an die Evaluationskriterien des VdA⁶¹ ergibt sich damit folgendes Bild. Eine Mengensteuerung (K 37) wird durch die ‚manuelle‘ Methode nicht erreicht, ist aber auch nicht nötig, da die Gerichte bereits sehr wenige Akten kennzeichnen. Die praktische Umsetzbarkeit der Bewertungsentscheidungen aus dem Bewertungsmodell (K 33) ist eher negativ zu bewerten. Zwar kennt niemand die Fälle so gut wie die bearbeitenden Richter, sie assoziieren diese aber scheinbar nicht ausreichend mit den

⁵⁷ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

⁵⁸ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024. Im Mietrecht wird etwa der Streitwert eines Mieters in einem Hochhaus auf alle anderen Mietobjekte des Hochhauses hochgerechnet und kann dann schnell in die Millionen gehen.

⁵⁹ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

⁶⁰ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

⁶¹ Arbeitskreis Archivische Bewertung im VdA, Evaluierung. Siehe die jeweiligen Kriterien am angegebenen Ort.

inhaltlichen Kriterien der Bewertungsmodelle oder übergehen die Kennzeichnung aus anderen, nicht bekannten Gründen. Infolgedessen werden kaum Akten gekennzeichnet. Dadurch werden vermutlich auch die inhaltlichen Überlieferungsziele (K 35) schlechter erreicht. Besondere Fälle werden wahrscheinlich zu selten/nicht gekennzeichnet; typische/normale Fälle werden womöglich nicht in ausreichender Zahl gekennzeichnet. Bewährt haben sich die manuellen Methoden dadurch eher nicht (K 26); angemessen (K 11) sind sie angesichts der schlechten Umsetzung und der mangelhaften Ergebnisse sowie den neuen Möglichkeiten maschineller Methoden zumindest bei ihrer alleinigen Anwendung nicht mehr.

3.3 Maschinelle Methoden

Die maschinelle Methode basierte bisher im wesentlichen auf der Datenbanksprache SQL, die derzeit das Haupt-Abfragewerkzeug ist.⁶² Das Archiv erhält dabei die Metadaten zu allen Akten, wählt die relevanten Zivilprozessakten mithilfe von Selektoren (in Form von komplexen SQL-Abfragen) aus und erstellt dann eine präzise Anforderungsliste. Damit wird nicht mehr das Gericht aufgefordert, Fälle zu kennzeichnen und dem Archiv zu schicken, sondern das Archiv fordert präzise bestimmte Aktenzeichen an. Ein Selektor kann dabei auch der Abgleich mit einer Datenbank sein, mit Aktenzeichen, Namen und Geburtsdaten.⁶³

So wie die manuelle Methode auf die Kennzeichnung durch die Gerichte angewiesen ist, ist die maschinelle auf gute Metadaten angewiesen. Die Datenqualität bestimmt wesentlich über das erzielbare Ergebnis, was einen engen Behördenkontakt bedingt und eine menschliche Nachkontrolle der Abfrageergebnisse.⁶⁴ Schlechte Metadaten wurden auch in den Interviews für diese Arbeit noch als Grund angeführt, weswegen sich die Entwicklung bzw. Anwendung derartiger Methoden und ihrer Werkzeuge noch nicht lohne. Hier befindet man sich noch in einer Experimentier-Phase und will maschinelle Methoden anwenden, sobald bessere Metadaten zur Verfügung stehen.⁶⁵ Sobald all diese Metadaten aber vorliegen, sehen sich die Archive aber auch in der Pflicht, mit diesen Daten zu arbeiten und ihre Modelle und Methoden anzupassen.⁶⁶ Aber auch dort wo bereits mit maschinellen Methoden gearbeitet wird, ist die Metadaten-Situation noch

⁶² Vielleicht am besten verkörpert wird diese Methode vom 2016 entwickelten Selesta, einer „datenbankbasierten Bewertungs- und Erschließungsunterstützung via Webbrowser“ (Zitat in Koch/Naumann/Rees/Riek/Schnell/Ziwes, Bewertungsautomat, S. 173), die auf SQL basiert. Selesta wurde mittlerweile abgelöst von neuen Verfahren, die jedoch im Grundzug gleich funktionieren und SQL nutzen. Siehe auch Naumann, Neues, S. 63-64.

⁶³ Siehe Naumann/Koch, Bewertungsautomat, S. 40-44, für eine Beschreibung der Funktionsweise

⁶⁴ Koch/Naumann/Rees/Riek/Schnell/Ziwes, Bewertungsautomat, S. 176-177. Siehe zu den Problemen mit Datenbereitstellung, Datenexport usw. auch das Kapitel zur Justiz.

⁶⁵ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

⁶⁶ Interview Gillner (Landesarchiv NRW) am 15.01.2024.

verbesserungsbedürftig. Oftmals gibt es für die Archive interessante Felder/Daten nicht, oder ihr Export benötigt technische Anpassungen, die nicht finanzierbar sind. Am Ende steht ein Kompromiss zwischen dem, was das Archiv an Daten wünscht und dem, was die Justiz bereitstellen kann.⁶⁷

Die bisher vorherrschende, SQL-basierte Methode zielt auf eine deutlich verbesserte Umsetzung der bisherigen Bewertungsmodelle. Zu den inhaltlichen Kriterien passende Akten sollen auf Basis der Metadaten mit komplexen Abfragen bzw. Datenbankabgleichen gefunden werden. Verbesserungen ergeben sich hier aus besseren Abfragewerkzeugen und besseren Metadaten, die eine genauere Identifikation der archivwürdigen Akten ermöglichen. Einen anderen Weg geht nun das Bayerische Staatsarchiv, begründet in dessen für Zivilprozessakten konkreterem und ‚einfacherem‘ Überlieferungsziel, das auf das ‚Normale‘ zielt statt auf das Besondere. Die Datengrundlage ist dieselbe wie bei den SQL-Verfahren, das Verfahren basiert aber auf einer Programmiersprache für statistisches Computing („R“), mit der bessere Samples erstellt werden können.⁶⁸ Das Verfahren zielt aber bisher nur auf die Amtsgerichte.

Im Folgenden werden nun die zwei Teile der Bewertung mit maschinellen Methoden vorgestellt: der Datenbankabgleich und die Samplebildung und ihre Umsetzung in SQL sowie R.⁶⁹

Datenbankabgleiche sollen den Archiven ermöglichen, selbst etwa über die Wichtigkeit von Personen zu entscheiden oder die rechtswissenschaftliche Bedeutung eines Verfahrens zu erkennen, ohne dabei auf das jeweilige Gericht angewiesen zu sein.⁷⁰ Es bestehen grundsätzlich drei Arten von interessanten Datenbanken: Personendatenbanken, andere externe Datenbanken sowie eigene, vom Archiv erstellte Datenbanken. Personenabfragen sind bei Zivilprozessakten schwierig, da keine Geburtsdaten in den Metadaten erfasst werden. Die Zivilprozessordnung sieht deren Erfassung nicht vor, sie sind für eine genaue Identifizierung von Personen aber elementar. Problematisch ist etwa auch, dass in den Zivilprozessakten-Metadaten die Parteien in einem Datenbankfeld stehen, das für eine maschinelle Verarbeitung geteilt werden müsste. Damit war eine personenbasierte Abfrage bisher nicht möglich. Andere externe Datenbanken können aus dem Bereich der Justiz genutzt werden, so etwa die JURIS-Datenbank. Bei dieser muss allerdings immer ein aktueller Stand vorgehalten werden. Eine weitere Option ist die Rechtsprechungsdatenbank des eigenen Landes

⁶⁷ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

⁶⁸ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024. Das Bayerische Staatsarchiv nutzt dafür „R“ (<https://www.r-project.org/>, 28.03.2024) und verarbeitet damit die typischen Datenexporte aus juristischen Fachverfahren.

⁶⁹ SQL und R werden hier stellvertretend für vergleichbare Instrumente der IT angeführt, weil sie in der deutschen Archivwelt bzw. den interviewten Archiven vorherrschen.

⁷⁰ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

(sofern vorhanden). Zusätzlich kann das Archiv eigene, interne Datenbanken erstellen, in die es manuell interessante Aktenzeichen einträgt.⁷¹ Die Erstellung eigener Datenbanken ist dabei ein Hybrid aus manuellen und maschinellen Methoden – entscheidend ist dabei, dass der Eintrag von der Bewertung zeitlich entkoppelt werden kann und die Aussonderung nur noch mit der Datenbank abgeglichen wird, im Gegensatz zum Richter, der die Akte während seiner Arbeit daran kennzeichnet.

Das Landesarchiv Baden-Württemberg stützt sich bisher aufgrund der beschriebenen Probleme bisher nur auf die JURIS-Datenbank und die eigene Datenbank für besondere Fälle. Das Landesarchiv NRW nutzt bisher die JURIS-Datenbank, von der bisher allerdings nur ein alter Stand vorliegt.⁷² Hier wird derzeit auch ein neues Tool entwickelt, das auf die GND zugreifen und Personen identifizieren soll. Die Strafprozessakte, die auch als erste mit maschinellen Methoden bewertet wurde, wird schon mit mehreren Datenbanken bewertet.⁷³

Das Bayerische Staatsarchiv will hingegen in seiner Vorgehensweise keine Datenbanken nutzen. Es sieht keine Möglichkeit zur Personenidentifikation und setzt darauf, dass die Kennzeichnung durch die Gerichte die Sonderfälle bzw. bekannten Persönlichkeiten abdeckt; zumal wegen der geringen Streitwerte beim Amtsgericht keine besondere Relevanz gesehen wird. Das Staatsarchiv ist außerdem von der Qualität der JURIS-Datenbank nicht überzeugt; ein Eintrag dort sei kein Qualitätsmerkmal und führte zu sehr hohen Übernahmequoten. Auf eine Landesrechtsprechungsdatenbank kann das Staatsarchiv nicht zurückgreifen.⁷⁴ Datenbankabgleiche haben also eher für SQL-basierte Verfahren zur Identifikation von Einzelfällen weiter hohe Relevanz. Problematiken bestehen hier noch durch zu unpräzise oder fehlende Metadaten und teilweise das Fehlen geeigneter Datenbanken.

Ergänzt werden sie durch Samples, die in der Komplexität ihrer Zusammenstellung stark variieren können. Immer dabei sind die von den Gerichten gekennzeichneten Akten sowie selbst vorgemerkte Aktenzeichen.

Eine einfachere Samplebildung nimmt bisher etwa das noch weitgehend ‚manuell‘ arbeitende Hessische Landesarchiv vor, das die ersten fünf Wohneigentumsverfahren pro Jahrgang übernimmt – und nur diese. Es geht dabei vor allem um den Evidenzwert

⁷¹ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024. Siehe für die Diskussion dieser Problemstellungen auch Brenneisen, Black box, S. 14-26. Brenneisen kommt besonders für Personen zum Schluss, dass künftig vielleicht der Akteninhalt ausgelesen und daraus ein Geburtsdatum ergänzt werden kann. Darauf wird auch in späteren Kapiteln noch eingegangen.

⁷² Für die beiden Archive: Interview Koch am 04.01.2024 sowie Interview Gillner (Landesarchiv NRW) am 15.01.2024.

⁷³ Interview Gillner (Landesarchiv NRW) am 15.01.2024. Das Hessische Landesarchiv arbeitet bspw. Mit Metadaten der Strafprozessakten (Staatsanwaltschaften, Fachverfahren MESTA) und experimentiert hier noch mit SQL und der Formulierung von Abfragen.

⁷⁴ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

dieser Akten,⁷⁵ warum aber ausschließlich diese Verfahrensart im Modell erwähnt wird, konnte nicht geklärt werden. Dieses Vorgehen wird damit begründet, dass diese Verfahren bei der Auswahl ‚besonderer‘ Fälle typischerweise nie ausgewählt werden. Komplexere Samplebildungen werden durch die schlechten Metadaten erheblich erschwert (siehe das Kapitel zur Justiz).⁷⁶

Das LABW arbeitet von den interviewten Archiven bisher am stärksten maschinell. Die Arbeit mit prozentualen Mengen und Zufallszahlen ist in SQL derzeit auch nur mit hohem Aufwand programmierbar, weshalb bisher davon abgesehen wurde. Stattdessen orientiert sich das LABW bei der Samplebildung am Streitwert (der recht zuverlässig ausgefüllt wird) und übernimmt bei den Amtsgerichten fünf Fälle, bei den Landgerichten 30 Fälle absteigend nach Streitwert. Das Landesarchiv NRW nutzt ebenfalls den Streitwert, aber in der Weise, dass ein sehr hoher Streitwert (über 500.000 Euro) als Indiz für die Archivwürdigkeit genommen wird.⁷⁷ Hinzugenommen werden in Baden-Württemberg pro Sachgebiet die zwei Fälle mit der längsten Laufzeit, und alle Fälle müssen mit Urteil oder gerichtlichem Vergleich abgeschlossen worden sein. Dadurch werden etwa 30 Akten pro Amtsgericht und 50 Akten pro Landgericht übernommen. Aufgrund der starken Varianz in der Größe der Gerichte ist dies fachlich angreifbar; momentan stehen aber fixe Zahlen in den Abfragen, da anderes technisch nicht möglich ist. Dies reduziert nebenbei die Übernahmemengen so erheblich, dass Mengenfragen in der Zivilprozessakten-Bewertung bisher keine Rolle spielen. Das LABW verfügt damit über das komplexeste Verfahren für die Zivilprozessakten, das im Vergleich mit Strafprozessakten trotzdem noch deutlich simpler ist (vier Abfragen gegenüber etwa einem Dutzend). Insgesamt werden so bei den Zivilprozessakten einige Beispiele übernommen, aber eher nur eine Ersatzüberlieferung erreicht.⁷⁸ Im Sample wird damit von allem ‚Normalen‘ etwas ‚Gutes‘ (umfangreiches, aussagekräftiges) abgebildet. Um eine statistische Repräsentativität geht es in den bis hierhin beschriebenen Ansätzen mit SQL nicht. Die Situation ist besser als beim Verlass auf die ‚manuelle‘ Methode, hohen theoretischen Ansprüchen kann man aber bisher nicht gerecht werden.⁷⁹

Die Repräsentativität war in den 2000er Jahren ein großes Thema, ausgelöst durch Matthias Buchholz' Studie zur Bewertung von Personalakten. Er kritisierte „die [nicht repräsentative] bewußte Auswahl des Typischen und des Besonderen“⁸⁰ und warb für

⁷⁵ Siehe Heck/Stabel/Laube/Rödel, Handreichung, S. 8 sowie Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

⁷⁶ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

⁷⁷ Landesarchiv NRW, Abschlussbericht, S. 76.

⁷⁸ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

⁷⁹ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

⁸⁰ Buchholz, Überlieferungsbildung, S. 255.

eine echte Repräsentativität. Es könnten nicht alle Forschungsfragen an die Unterlagen antizipiert werden und welche Akten/Informationen für deren Bearbeitung notwendig wären. Die „Verteilung der untersuchungsrelevanten Merkmale in der Grundgesamtheit [ist] unbekannt“, weshalb Repräsentativität nur über Zufallsstichproben realisiert werden könnte.⁸¹

Der bayerische Ansatz versucht nun, eine gute Repräsentativität zu erreichen. Dafür muss die Größe der Stichprobe stimmen (vereinfacht: je größer die Grundgesamtheit der Verfahren desto besser) und eine echte Zufallsstichprobe gezogen werden. Der Vorteil ist dabei, dass der prozentuale Auswahlsatz bei wachsender Größe der Grundgesamtheit sinkt.⁸² Das Bayerische Staatsarchiv will sich in seinem neuen Verfahren auf die Daten eines Jahrgangs aller bayerischen Amtsgerichte stützen, der bei Entwicklung der Methode ca. 126.000 Akten umfasste. Unvollständige Daten oder Fehler, ähnlich denen in den Daten der anderen Bundesländer, werden als nicht übermäßig problematisch beurteilt. Die jährliche Ausspielung der Daten wird von der Justiz zwar für nicht ganz einfach gehalten, ist aber möglich. Die Datengrundlage wird zunächst auf die Fälle reduziert, in denen tatsächlich eine Entscheidung erging, weil für das Überlieferungsziel jeder Fall einen Sachverhalt, eine Entscheidung und einen gewissen Umfang haben soll. Der Kern ist dann die Stichprobe. Prozentmengen seien nicht zielführend, daher wird mit absoluten Zahlen gearbeitet. Mit einer Stichprobe von 400 Fällen aus 10.000 Fällen wird bereits ein 95-prozentiges Konfidenzintervall erreicht, allerdings bei ganz einfachen binären Gesichtspunkten einer reinen Zufallsstichprobe.⁸³ Die Sachgebiete werden nun in fünf Gruppen eingeteilt und die Auswahl danach stratifiziert. Dadurch wird verhindert, dass Sachgebiete mit sehr wenigen Fällen nicht in die Stichprobe gelangen. Das erhöht die Wahrscheinlichkeit der Ziehung, beeinträchtigt aber nicht die Repräsentativität – für diese reicht es, wenn klar ist, welche Auswahlwahrscheinlichkeit der einzelne Fall hatte. Die Repräsentativität ist damit das Ziel und der Gewinn des neuen Verfahrens. Künftiger Forschung wird dadurch der Schluss vom Sample auf die Gesamtheit erlaubt; bei einem nach der ‚alten‘ manuellen/SQL-Methode ausgewählten Fall sei die Übertragbarkeit auf die Allgemeinheit der Fälle schwächer gegeben. Letztlich ist die bayerische Methode auch einfacher, da sie keine komplexen

⁸¹ Buchholz, Archivische Überlieferungsbildung, S. 211. Für den statistischen Hintergrund der Stichprobenauswahl siehe S. 215-228. Nicht repräsentative Verfahren können ihre eigenen Vorteile haben. Strebt das Archiv aber Repräsentativität an, sollte eine reine Zufallsauswahl vorgenommen werden und das Sample nicht mehr verfälscht wird, siehe dafür S. 232-233. Siehe auch a.a.O., S. 210-235.

⁸² Buchholz, Methoden, S. 24-27, Zitat S. 25.

⁸³ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024. Die Stichprobe wird dabei von R gezogen, das einen Zufallszahlengenerator enthält. Das Hauptstaatsarchiv arbeitete bei der Entwicklung dieser Methode mit Statistikern von der Ludwig-Maximilian-Universität München zusammen.

mathematischen Terme benötigt, wie sie in Buchholz' Publikationen vorkamen; stattdessen reicht eine vom Werkzeug automatisch gezogene relativ kleine Stichprobe von 400 Fällen.⁸⁴

Damit ergibt sich eine völlige Divergenz der genutzten maschinellen Ansätze. Bayern hat ein Überlieferungsziel (für die Amtsgerichte!) formuliert, das ohne das ‚Besondere‘ auskommt; dementsprechend sind die Anforderungen an die Metadaten niedriger, Datenbanken werden nicht benötigt. Wie das für die Landgerichte ausgestaltet wird, ist noch offen; möglicherweise wird dort das ‚Besondere‘ aber eine wichtigere Rolle spielen und dessen Auswahl könnte wieder komplexere Verfahren bedingen. Bei den Amtsgerichten ist die repräsentative Stichprobe aber mit einem relativ einfachen Tool zu erreichen. Allerdings ist Repräsentativität auch kein Selbstzweck und sollte nicht als „dogmatischer Anspruch aufgefasst werden“.⁸⁵ Ihre Notwendigkeit richtet sich nach der bisherigen Überlieferung und den Überlieferungszielen.

Ein Vergleich der beiden Ansätze (SQL-exemplarisch und R-repräsentativ) muss ihre Unterschiede besonders akzeptieren. Die Evaluationskriterien des VdA helfen hier nämlich nur bedingt. Kriterium K 26 spricht von der Bildung repräsentativer oder exemplarischer Samples, wofür die Verfahren möglicherweise „unsachgemäß eingesetzt oder verzahnt [sind] und damit unzureichende Ergebnisse liefern.“⁸⁶ Die verschiedenen Ziele der Ansätze sind eben verschieden und auch ein unsachgemäßer Einsatz oder Verzahnung waren in den Interviews nicht festzustellen. Die Anwender des exemplarischen Ansatzes verwiesen auf die schlechten Metadaten (insbesondere für Zivilprozessakten, weniger für Strafprozessakten!) und Limits der technischen Möglichkeiten von SQL. Wären diese besser, könnten sie sich ebenfalls in Richtung Repräsentativität bewegen. Der Grund für ‚unzureichende Ergebnisse‘ (K 26) liegt eben nicht im falschen Einsatz der Werkzeuge, sondern mehr in externen Faktoren – vor allem der Datenerstellung und Bereitstellung durch die Justiz. An diesem Punkt kann vielleicht ein direkter Vergleich gezogen werden, auch zu den manuellen Methoden. Das Kriterium ist dabei nicht nur die gebildete Überlieferung (exemplarisch oder repräsentativ) sondern die Kontrolle darüber. So wie maschinelle Methoden den Archivaren mehr Kontrolle über den Prozess und damit auch über die gebildete Überlieferung bieten als manuelle Methoden, bietet der repräsentative Ansatz – bei begrenzteren Zielen – mehr Kontrolle als der exemplarische. Dieses Ergebnis ist so zu verstehen wie Buchholz' zur Repräsentativität: auch wenn man sie befürwortet, sollte sie nicht als dogmatisches, einziges Ziel verstanden werden. Es bleibt damit spannend,

⁸⁴ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

⁸⁵ Buchholz, Methoden, S. 26.

⁸⁶ Arbeitskreis Archivische Bewertung im VdA, Evaluierung, K 26.

wie sich die Ziele, Rahmenbedingungen und Ergebnisse für die beiden Ansätze entwickeln.

3.4 Management von Modellen und Methoden

Das vorige Kapitel stellte fest, dass zwei Ansätze in der modernen, maschinellen Zivilprozessakten-Bewertung bestehen. Da verschiedene Überlieferungsziele und damit auch gebildete Überlieferungen legitim sind, ist ein wichtiges Kriterium die Kontrolle über den Prozess der Überlieferungsbildung. Mit dem Management des Prozesses, zu dem dessen Kontrolle und Evaluation gehören, beschäftigt sich dieses Kapitel. Besonders wichtig ist dabei der Kontakt mit der Justiz, sowie die interne Organisation des Archivs.

Bewertungsmodelle sollen grundsätzlich kontinuierlich evaluiert und fortgeschrieben werden, wobei die Justiz beispielsweise in Hessen sehr hoch priorisiert wird.⁸⁷ Wie in den vorigen Kapiteln beschrieben, waren die Archive bei Vorherrschen der manuellen Methode extrem auf Mithilfe durch die Gerichte angewiesen. Tagungen mit den Justizarchivpflegern der Gerichte sollten diese die Überlieferungsziele vermitteln und sie zur Kennzeichnung von Akten zu animieren. Dies war jedoch wenig effektiv.⁸⁸ Bemängelt wurde in den Interviews auch, das Justizarchivpfleger schlicht keine beliebte Aufgabe sei. Die Person würde dazu bestellt, ohne über ihre Aufgaben aufgeklärt zu werden. Auch hinsichtlich der neuen maschinellen Methode bestehe großer Beratungsbedarf. Am besten wäre, wenn die Gerichte miteinander kommunizieren und sich gegenseitig helfen würden,⁸⁹ denn teils fehlten die Kenntnisse zur korrekten Bedienung der eigenen Fachverfahren.⁹⁰ All dies führte bisher zu schlechten Metadaten, uneinheitlichen Anbieterslisten und erschwerte dadurch die Bewertung.⁹¹ Stattdessen besteht nun die Hoffnung, dass durch die Kooperation mit zentralen IT-Stellen etwa im Justizministerium die Bereitstellung verbessert werden kann.⁹²

Solange die ‚manuelle‘ Methode vorherrschte, bestand auch ein geringerer Anspruch an die Überlieferungsbildung, wodurch auch die Evaluation des Modells weniger dringlich erschien.⁹³ Die Archive mussten mit dem auskommen mussten, was die Gerichte ihnen übermittelten; dadurch erhielten die Archive aber auch nur zum Modell passende Akten,

⁸⁷ Haberkorn/Herge/Laube/Maaß/Petter/Zilien/Gniffke, Teilbewertungsmodell, S. 9.

⁸⁸ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

⁸⁹ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

⁹⁰ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

⁹¹ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

⁹² Interview Gillner am 15.01.2024: Das Landesarchiv NRW setzt auf die Kooperation der eigenen Grundsatzabteilung mit den zuständigen Stellen im Justizministerium. Interview Unger am 27.02.2024: auch in Bayern soll die zentrale Justiz-Service-Stelle beim Justizministerium künftig den Gesamtdatensatz bereitstellen. Interview Gniffke am 12.01.2024: hier besteht enger Kontakt mit der IT-Stelle der hessischen Justiz.

⁹³ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

wodurch sich kein Impuls zur Evaluation und Änderung des Modells ergab. Auch hinsichtlich der geringen Abgabemengen bestand keine Notwendigkeit, Modell und Methoden anzupassen.⁹⁴

Mit der steigenden Verbreitung und Komplexität maschineller Methoden wird eine intensivere Evaluierung notwendig. Die Archive nutzen dafür nun Evaluationstreffen der beteiligten Archivare, bei denen die Anpassung von Modell und Methode besprochen wird. Verantwortung für das Modell hat dabei ein Modellpfleger (NRW), für die Methodik eine Arbeitsgruppe (Baden-Württemberg). Besonders wichtig sind Informationen über Veränderungen in den Datenbanken oder neue Ideen der praktisch arbeitenden Kollegen. Auch hier gilt, dass die Bewertungsmodelle dauerhafter sind, während die Methodik nur noch ein halbes Jahr aktuell bleiben kann. Die Arbeitsgruppe prüft dann, wie viel der neuen Ideen und theoretischen Ziele in die Praxis umgesetzt werden kann.⁹⁵

Das skizzierte, derzeitige Vorgehen entspricht den Evaluationskriterien des VdA sehr gut. Die Modelle und Methoden werden in Gruppen evaluiert und angepasst (K 39) und sind nicht von einzelnen Personen abhängig. Sie reagieren auch sensibel auf alle Veränderungen, die „Fehler und Lücken in den Steuerungsinstrumenten“ produzieren könnten (K 40).⁹⁶ Für die Zivilprozessakten ist das wichtigste ‚Steuerungsinstrument‘ praktisch die Justiz: sie erstellt die Metadaten, sie muss sie den Archiven zur Verfügung stellen. Daher ist der Fokus ihre übergeordneten IT-Stellen sinnvoll. Denn hier sind in naher Zukunft starke Veränderungen zu erwarten, die für bessere Metadaten sorgen können.

4. Zukunftserwartungen

4.1 Landesarchive: Standardisierung, Vereinheitlichung, technische Innovation

Die bisherigen Kapitel erörtern den gegenwärtigen Stand der Zivilprozessakten-Bewertung. Es wurde aber auch darauf hingewiesen, dass sie sich in einer Umbruchsituation befindet und alle interviewten Archive ihre Modelle und Methoden überarbeiten.⁹⁷ Diese Überarbeitung wird natürlich beeinflusst von Zukunftserwartungen, die sich auf Veränderungen in der Justiz, innerhalb oder auch zwischen Archiven, und in technischer Hinsicht ergeben.

⁹⁴ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

⁹⁵ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024 sowie Interview Gillner (Landesarchiv NRW) am 15.01.2024.

⁹⁶ Arbeitskreis Archivische Bewertung im VdA, Evaluierung, K 39 und K 40.

⁹⁷ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

Ein Dauerthema waren und bleiben die Metadaten und die Hoffnung, dass sich bei besseren Metadaten und Methoden die Zivilprozessakten-Bewertung der schon besser funktionierenden Strafprozessaktenbewertung annähert.⁹⁸ Jede Generation von Archivaren müsse danach streben, ihre Datenbasis zu verbessern und anzupassen. Ein gewünschtes Metadatum für die Zukunft wäre etwa der Aktenumfang, der auch in der E-Akte nachgehalten wird und ein weiterer Indikator für interessante Verfahren sei.⁹⁹

Mit der E-Akte verknüpfen sich hingegen keine besonderen Erwartungen, unter anderem, da sie sehr eng an der Papierakte orientiert wurde, schlicht PDF-Dateien enthält und arm an Metadaten ist.¹⁰⁰ Solange die Zivilprozessordnung sich nicht verändert, ändern sich auch die grundsätzlich vorhandenen Metadaten nicht. Eher noch besteht die Befürchtung, dass die Veränderung mit Einführung der E-Akte nicht sofort in die ‚richtige‘ Richtung für das Archiv führt, etwa wenn der Kenntnisstand des Personals hinterher hängt.¹⁰¹

Mehr Hoffnungen richten sich auf den XML-basierten Austauschstandard XJustiz. Dessen Hauptcharakteristika sind eine bundesweite Einheitlichkeit, das Ziel einer einfachen Verschiebbarkeit von Daten „durch einfachen Mausklick“ und ein XML-Schema für Metadaten, an dem sich Softwareentwickler orientieren können.¹⁰² XJustiz hat auch ein erweitertes Schema für Zivilprozessakten. Der erwartete Vorteil besteht darin, dass keine individuellen Lösungen mehr für die Aussonderung entwickelt werden müssen und vor allem der „strukturierte Metadaten- und Aktentransfer“ garantiert ist.¹⁰³ Genau hier bestanden bisher die Probleme, die besonders den exemplarischen Ansatz erschweren.

Gerade das Landesarchiv NRW erwartet etwa, dass der Wechsel zur E-Akte und der XJustiz-Standard zur Anbietung aller Akten führen und die Archive zur Überarbeitung ihrer Modelle und Methoden zwingen werden. Justizseitig würden die Prozesse dann weitgehend automatisiert ablaufen. Alle Gerichte würden gleich nach diesem Standard arbeiten und eine Bewertung über die Daten aller Gerichte (wie in Bayern) möglich.¹⁰⁴ Hinzu kommt in fernerer Zukunft das GeFa, das Gemeinsame Fachverfahren für die Justiz, mit dem sich Hoffnungen auch mehr und auch mehr Metadaten verbinden.¹⁰⁵ Es soll ein Fachverfahren für „die gesamte deutsche Justiz“ sein, „den Datenaustausch mit

⁹⁸ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

⁹⁹ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

¹⁰⁰ Interview Koch am 04.01.2024 und genauso Interview Gillner (Landesarchiv NRW) am 15.01.2024.

¹⁰¹ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

¹⁰² Website XJustiz, https://xjustiz.justiz.de/projektskizze/index.php#ausgangspunkt_ziele (25.03.2024).

¹⁰³ Schweizer, E-Akte, S. 126.

¹⁰⁴ Interview Gillner (Landesarchiv NRW) am 15.01.2024.

¹⁰⁵ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

externen Stellen unterstützen“ und schlägt „die Brücke zwischen den elektronischen Akten- und Textsystemen“.¹⁰⁶

Durch XJustiz und GeFA würden bisher zentrale Probleme der Zivilprozessakten-Bewertung – unzureichende Kennzeichnungen, Bereitstellungsprobleme, schlechte Metadatenbasis – für die Archive gelöst. Die richtigen Ansprechpartner der Zukunft sind dadurch die, auf die die Archive schon jetzt zielen: übergeordnete IT-Stellen.¹⁰⁷

Kritischere Stimmen erwarten von XJustiz nur geringe Veränderungen; sollten hier etwa Codierungen verwendet werden (wie in den Strafprozessakten) und die Archive keine Kataloge mit ihren Auflösungen erhalten, bestehe eher die Gefahr eines Rückschritts.¹⁰⁸

XJustiz und GeFa verweisen auf Standardisierungstendenzen in der Justiz; dem könnten auch Vereinheitlichungstendenzen in den Archiven gegenüber stehen, auch aufgrund der technischen Entwicklung. Schon jetzt wird etwa im Landesarchiv Baden-Württemberg dank maschineller Methoden mit gemeinsam formulierten Abfragen die Bewertung viel einheitlicher vorgenommen als früher.¹⁰⁹ Wenn bald alle Archive massenhaft XJustiz-Datensätze bekommen, wird der Bedarf nach einer einheitlichen Lösung akut sein.¹¹⁰

Dank einer Standardisierung bei den Gerichten und vollständigen Abgaben könnten mithilfe einheitlicher Daten Bewertungsentscheidungen dann über alle Gerichte hinweg getroffen werden.¹¹¹ Die Archive lernen dabei voneinander und das Landesarchiv Baden-Württemberg stellt seine Version des exemplarischen Ansatz mit SQL vor, den das Landesarchiv Hessen adaptieren will. Wesentliche Hürden sind dabei noch Wissen und Multiplikatoren im eigenen Archiv.¹¹² Eine maximale interne Vereinheitlichung der Bewertungspraxis findet hingegen im Bayerischen Staatsarchiv statt; dort wird auch überlegt, inwiefern der statistisch-repräsentative Ansatz auf andere Massenakten ausgeweitet werden kann.¹¹³ Vereinheitlichung und Standardisierung sind aber auch im Kontakt mit der Justiz relevant. Stellt die Justiz einheitliche Metadaten bereit, sollte auch das Archiv standardisierter und objektiver arbeiten. Für ein transparentes Handeln sind etwa gemeinsame Standards eine große Chance.¹¹⁴

Die Arbeit hat allerdings auch gezeigt, dass durchaus unterschiedliche Ansätze bestehen, die sich auf unterschiedliche Technik stützen. Inwiefern technische Neuerungen relevant

¹⁰⁶ Website GeFa, <https://gefa-justiz.de/was-ist-gefa> (16.02.2024).

¹⁰⁷ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024. Laut Gniffke sei es vor allem dort sinnvoll die Gerichte zu involvieren, wo es um die Modalitäten der Übernahme geht, und dafür sei die IT-Stelle der hessischen Justiz der beste Ansprechpartner.

¹⁰⁸ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

¹⁰⁹ Interview Koch (Landesarchiv Baden-Württemberg) am 04.01.2024.

¹¹⁰ Interview Gillner (Landesarchiv NRW) am 15.01.2024.

¹¹¹ Interview Gillner (Landesarchiv NRW) am 15.01.2024. Der bayerische Ansatz zielt genau darauf, ist aber noch nicht reguläre Praxis.

¹¹² Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

¹¹³ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

¹¹⁴ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

oder adaptierbar sind, bleibt eine schwierige Frage. Die Anpassung der eigenen Prozesse an die Arbeit mit XJustiz-Datensätzen wird wohl unverzichtbar sein. Das Landesarchiv NRW hat kürzlich ein Werkzeug hierfür entwickelt, das XJustiz-Datensätze aufnimmt, für die Bewertung bearbeitet, und am Ende eine XJustiz-Nachricht mit der Bewertungsentscheidung an das Gericht sendet.¹¹⁵

KI wird bisher mehr als Unterstützungswerkzeug gesehen. Sie könnte beispielsweise SQL-Kenntnisse weniger notwendig machen, indem sie menschlich produzierte Abfragen in SQL konvertiert oder gegebenenfalls bei der Bereinigung von Metadaten eingesetzt werden könnte.¹¹⁶ Inwiefern KI direkt in die Analyse der Metadaten und möglicherweise Volltexten einbezogen wird, muss sich zeigen. Für die Metadaten wäre es besser, wenn in den Datenbanken schlicht mehr Pflichtfelder eingerichtet würden, etwa für das Geburtsdatum. KI könnte perspektivisch vielleicht eher helfen, mithilfe des Volltextes Bezüge herzustellen oder noch etwas ‚Besonderes‘ zu finden.¹¹⁷ Allgemein bestehen aber noch wenig Vorstellungen, wie KI (und Volltexte) genutzt werden sollten, da die Metadaten eigentlich alle Informationen in kompakter Weise bereitstellen. Besonders für die C-Verfahren bei den Amtsgerichten besteht auch die Frage, welches ‚Besondere‘ dort noch zu finden sein könnte. Dementsprechend wird die Nutzung von KI und Volltexten als relevanter für andere Aktengruppen angesehen.¹¹⁸

4.2 Internationale Forschung: technische Innovation

Aufgrund der Kürze der Arbeit und dem enormen Tempo der technischen Entwicklung kann dieses Kapitel nur eine Momentaufnahme bieten. Interviewt wurden für das Kapitel Jenny Bunn (The National Archives in Großbritannien (TNA)) und Basma Makhlouf-Shabou (Universität Genf).

Neue technische Instrumente müssen von den Archivaren immer so adaptiert werden, dass sie das ‚sense making‘ mithilfe der angebotenen Daten ermöglichen. Hierbei geht es dann um einen besonderen ‚archival sense‘, er sollte aber Inhalt und Hintergründe der in den Dokumenten beschriebenen Sachverhalte einigermaßen akkurat wiedergeben.¹¹⁹ Ein Beispiel wäre etwa die Rekonstruktion von Beziehungen in einer Universität anhand der angebotenen Dokumente. Die Frage ist dabei immer, wie die Verarbeitungsgeschwindigkeit des Computers mit dem menschlichen ‚sense making‘ zusammengebracht werden kann.¹²⁰ Die National Archives prüften seit den 1990er

¹¹⁵ Interview Gillner (Landesarchiv NRW) am 15.01.2024.

¹¹⁶ Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

¹¹⁷ Interview Gillner (Landesarchiv NRW) am 15.01.2024 sowie Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

¹¹⁸ Interview Unger (Bayerisches Hauptstaatsarchiv) am 27.02.2024.

¹¹⁹ Bunn, Frames, S. 72, 74.

¹²⁰ Bunn, Frames, S. 74-76.

Jahren in mehreren Projekten verschiedene Arten moderner Werkzeuge, auch für Unterlagen der Justiz. Den Archivaren und der Justiz stellen sich ähnliche Herausforderungen: immer muss aus einer großen Menge Dokumenten/Informationen das relevante extrahiert werden.¹²¹ 2015 wurden etwa eDiscovery-Tools untersucht, die große, unstrukturierte Datenmengen auf technische Eigenschaften sowie sensible Inhalte (v.a. personenbezogene Daten) durchsuchen können. Zwar zielte die Untersuchung damals auf die Automatisierung der Schwärzung sensibler Inhalte; sie zu finden, Dokumente danach zu gruppieren usw. könnte aber auch für die Bewertung interessant sein.¹²²

Die zwei hauptsächlichen Themen neuerer Forschungen sind die Extraktion von Metadaten und dann die Verbindung der Ergebnisse mit Datenbanken. TNA übernimmt beispielsweise die Veröffentlichung von Urteilen der höchsten Gerichte. Dabei werden nur die Urteile, nicht die gesamte übernommene Akte, von einem Parser zusammengefasst und Basisinformationen in XML extrahiert. Er sorgt damit für eine besser strukturierte Version der gegebenen Informationen,¹²³ die mit den gegebenen Metadaten verknüpft werden könnte. In den Zivilprozessakten könnte die Klageschrift hierfür das aussagekräftigste Dokument sein. Andererseits muss hier kein Limit gesetzt werden; auch die ganze Akte könnte durchsucht und aus dem Text zusätzliche Metadaten extrahiert werden, die bisher uneindeutige Metadaten ergänzen und ihnen Kontext verleihen, der etwa für eine eindeutige Personenidentifikation nützlich sein kann. Nach Erfahrungen der TNA funktioniert die Kombination von Named Entity Extraction (Metadaten) und Named Entity Linking (Verknüpfung Metadatum und Datensatz in einer Datenbank) aber noch nicht ausreichend. Die Technik verknüpfte zwar gleiche Namen, die aber zu unterschiedlichen Personen gehörten. Das Werkzeug konnte nicht intelligent arbeiten, d.h. keine zusätzlichen Informationen einbeziehen, da es die Unterlagen nicht ‚liest‘. Wenn hier Regeln codiert werden könnten („finde den Namen XY im Zusammenhang mit diesen vier Wörtern und nicht mit diesem Wort“), könnte das Ergebnis verbessert werden. Um zu funktionieren, müssten solche Regeln recht genau sein. In Anbetracht der Individualität von z.B. Personen würde das jedoch sehr komplex.¹²⁴ Deutlich besser wären zusätzliche Metadaten: mithilfe von Geburtsdaten sowohl in Akte und Datenbank könnte die Verknüpfung wahrscheinlich wirklich automatisiert erfolgen. Beim jetzigen Stand bleibt damit auch beim Einbezug

¹²¹ Bunn, AI. Siehe auch The National Archives, Using AI.

¹²² The National Archives, Application, S. 15-21.

¹²³ Interview mit Jenny Bunn für The National Archives (UK), 23.02.2024.

¹²⁴ Interview mit Jenny Bunn für The National Archives (UK), 23.02.2024. Es wäre fast für jede Person, oder möglicherweise Personengruppe eine eigene Regel notwendig.

etwa der GND menschliche Intervention notwendig, um die notwendige Intelligenz bereitzustellen.

Makhlouf-Shabou war im Gespräch vor allem die Unterscheidung zwischen Automatisierung und KI wichtig. In einem Projekt mit dem Staatsarchiv Neuchâtel prüfte sie mithilfe von Data Mining, wie stark bestimmte Kriterien für die Bewertung (sowohl formal wie auch inhaltlich) automatisiert geprüft werden können. Die untersuchten Unterlagen umfassten sowohl unstrukturierte wie auch strukturierte Daten. Etwa 60% der Metriken waren voll automatisierbar, viele weitere teilweise.¹²⁵ Neben diesen automatischen Ansatz tritt dann die KI. Beides sollte verwendet werden und dadurch eine doppelte Prüfung ermöglicht, die die Verzerrungen/Biases beider Ansätze aufdeckt.

Ein wichtiger Faktor für die Arbeit mit KI ist laut Makhlouf-Shabou auch, dass KI eben keine extrem große Menge an Metadaten benötigt, sondern nur Sets von unbewerteten und bewerteten Daten. Das KI-Modell wird dann zusätzlich mit Texten, Dokumenten und Informationen gefüttert, etwa über die Funktionen der bewerteten Behörde, legaler und sonstiger regulatorischer Kontexte und weiteres – also praktisch alles, was auch ein menschlicher Archivar in die Bewertung mit einbezieht. Dadurch wird das Training des Modells verbessert.¹²⁶ Ein Problem hierbei ist, dass die bisherige manuelle, aber teils auch noch die maschinelle Bewertungspraxis von den Praktikern in Landes- und Staatsarchiven als recht mangelbehaftet dargestellt wurde. Die vorhandenen, bewerteten Daten bieten damit nicht unbedingt eine passende Grundlage, von der KI-Modelle lernen sollten.

Für den ‚klassischen‘ Fall der Personenidentifikation mit unvollständigen Informationen bestehen ebenfalls weitere Probleme. Wenn die KI nicht nur Daten extrahiert (Johann Müller), sondern selbst hinzufügt (Johann Müller, geb. in X am XX.XX.XXXX), entstehen Authentizitätsprobleme. Fehler und Halluzinationen sind typisch für KI, ihre Ergebnisse sollten daher nur Vorschläge sein und menschlicher Intelligenz zu klaren, besseren Entscheidungen verhelfen.¹²⁷

Daher sollte eher die Verbesserung der Metadaten durch Extraktion im Vordergrund als die mit Unsicherheiten behaftete Verknüpfung. Wenn diese Metadaten gegeben sind, ist die Verknüpfung viel einfacher und genauer. Zusätzliche Metadaten (im Fachverfahren oder im Volltext) sollten auf verschiedenen Wegen gesucht werden. Eine Möglichkeit für den Volltext wären etwa Parser und Microservices, die mit relativ einfachen

¹²⁵ Interview mit Basma Makhlouf Shabou, 26.02.2024.

¹²⁶ Interview mit Basma Makhlouf Shabou, 26.02.2024.

¹²⁷ Interview mit Basma Makhlouf Shabou, 26.02.2024.

Prinzipien zusätzliche Metadaten generieren könnten.¹²⁸ Beispielsweise im Volltext von Verträgen in der Akte Geburtsdaten zu finden,¹²⁹ würde allerdings auch zu größerer Heterogenität der Datenbasis führen. Für manche Akten wären damit (viel) mehr Metadaten vorhanden als für andere. Möglicherweise könnten aber gerade diese zusätzlichen Daten eine umfangreichere Akte mit mehr Dokumenten und mehr Inhalt anzeigen.¹³⁰ Microservices könnten auch das Textmaterial untersuchen, verwendete Gesetze/Paragraphen ermitteln und ausgeben¹³¹ oder besonders häufige Wörter finden und bei diesen einen Abgleich mit der GND und GND-Oberbegriffen ermöglichen. Die Analysemöglichkeiten sind vielfältig. Ob solche Methoden zu besseren oder stärker verzerrten Bewertungsergebnissen führen, wäre damit eine interessante Frage für weitergehende Forschung. Für den exemplarischen Ansatz das ‚Besondere‘ zu finden, wäre aber sicherlich leichter. Im vorigen Kapitel wurden als Ansatzpunkte für mehr/bessere Metadaten etwa die Kommunikation mit IT-Stellen der Justiz genannt oder die Schaffung von Pflichtfeldern in den Fachverfahren. Eine ‚pessimistischere‘ Perspektive ist hingegen, dass nur eine Änderung der ZPO positive Konsequenzen haben werde,¹³² wie etwa die Eintragung von Geburtsdaten zur eindeutigen Identifikation von Personen. Insgesamt würde durch eindeutigere Metadaten die Verknüpfungsleistung deutlich erleichtert und die dort zu erwartende Fehlerrate stark reduziert. Ob dabei Data Mining, KI oder andere Technologien zum Einsatz kommen, ist zweitrangig. Die Option zur Nutzung von KI/Data Mining/Data Science zu haben, ist aber immer positiv zu bewerten.

Die Adoption solcher Technologien/Methoden ist jedoch schwierig. Ein grundsätzliches ‚Problem‘ ist dabei, dass die meisten Archivare eine geisteswissenschaftliche Ausbildung haben und nicht im Sinne von Daten und Data Science denken. Alles digitale Material als Daten verstanden werden, die mit den neuen Werkzeugen anders analysiert werden können. Dafür braucht es auch neue Fähigkeiten im IT-Bereich bei

¹²⁸ Interview Nicola Welch (TNA) am 20.03.2024. Laut Welch sind die Parser komplexer, weil sie an bestimmten Orten nach Informationen suchen und diese aufnehmen. Andere Microservices könnten sich für simple Abfragen auch schlicht auf Regex (regular expressions) stützen. Siehe zum Projekt *Welch, Nicki*, Find Case Law: Reflecting on our first year of publishing judgments, part one, <https://blog.nationalarchives.gov.uk/find-case-law-reflecting-on-our-first-year-of-publishing-judgments-part-one/> (17.03.2024).

¹²⁹ Brenneisen, Black Box, S. 24.

¹³⁰ Gutachten, Dokumentationen und ähnliche Dokumente, die möglicherweise mehr und zusätzliche Daten enthalten, können Indikatoren für Archivwürdigkeit sein, siehe z.B. Brenneisen, Black box, S. 7 oder auch Interview Koch am 04.01.2024, die den Umfang als ein interessantes Metadatum benannte. Andererseits konstatiert das Landesarchiv Hessen, dass sich in Zivilprozessakten oft viele nicht archivwürdige Dokumente befinden und nur Schriftwechsel, Beweisstücke, Dokumentationen und Urteile interessant sind. Die Frage wäre dann, ob die KI mehr Nutzen aus den wertvollen Stücken zieht oder sich das Extraktionsergebnis durch die vielen nutzlosen Dokumente verschlechtert.

¹³¹ Interview Nicola Welch (TNA) am 20.03.2024.

¹³² Interview Gniffke (Landesarchiv Hessen) am 12.01.2024.

den Archivaren,¹³³ um selbst Analysen anzustellen und um eine gewisse Souveränität im Umgang mit IT-Fachleuten zu erreichen. Allgemein sollten die Archivare ‚up to date‘ sein und die IT/KI-Spezialisten die zuständigen Personen für Datenhaltung, Vertraulichkeitssicherung und Zugänglichkeit über die Auswahl der richtigen Werkzeuge informieren.¹³⁴

Insgesamt könnten neue Wege der Datenextraktion und Verknüpfung (darunter KI) dank besserer Datenbasis insbesondere die Ergebnisse des exemplarischen Ansatzes verbessern. Dem gegenüber steht jedoch der Aufwand, den der Einsatz neuer Methoden und Technologien mit sich bringt. Die Zivilprozessakten eignen sich durch ihre eher schlechte Datenbasis (vorläufig) schlechter für den erstmaligen Einsatz neuer Technologien. Andere Aktengruppen mit besseren Metadaten, größerer Standardisierung und bisher besseren Bewertungsergebnissen bieten wahrscheinlich eine bessere Grundlage für die Erarbeitung von Methoden und (KI-)Modellen, die dann vielleicht für die Zivilprozessakten adaptiert werden können. Insofern könnten vielleicht eher wieder die Strafprozessakten Pionier sein, wie sie es auch bei Einführung der ersten maschinellen Methoden waren.¹³⁵

5. Ergebnisse

Die interviewten Archive haben nach dem VdA-Kriterienkatalog zur Evaluierung von Bewertungsmodellen eigentlich gute Voraussetzungen, für die Zivilprozessakten-Bewertung – aus einer ‚formalen‘ Perspektive. Sie verfügen über gute Modelle, Kontakte zu den Bestandsbildnern, über Methoden und Prozesse zum Umgang mit den Zivilprozessakten.

In der praktischen Arbeit ergeben sich aber zahlreiche Schwierigkeiten. Viele davon können – über Umwege – letztlich wieder auf die Basis der Bewertung von (digitalen) Massenakten zurückgeführt werden: die Metadaten.

Am Anfang stehen dabei die unterschiedlichen Arbeitsweisen von Archivaren und der Justiz: erstere wünschen viele und genaue Metadaten, während die Justiz auch wegen der enormen Spannbreite der Sachverhalte in Zivilprozessen eher mit groben Kategorien arbeitet. Abseits davon sind die Metadaten leider auch oft genug schlecht strukturiert, uneinheitlich oder unvollständig.

Diese eher unzureichenden Metadaten trafen dann bisher auf Bewertungsmodelle, die vor allem auf die gezielte Auswahl besonderer Einzelfälle zielten und die Übernahme

¹³³ Interview Bunn am 23.02.2024.

¹³⁴ Interview Makhoul Shabou

¹³⁵ Siehe für eine kurze Zusammenfassung der Entwicklungen bei den Strafprozessakten vom Anfang bis zu Selesta (2018): Naumann/Koch, Bewertungsautomat.

zeit- und regionaltypischer Fälle zielten. Aufgrund der Masse an Akten und dem Mangel an Metadaten und mächtigen Analyse- und Auswahlwerkzeugen, waren die Archive jedoch gezwungen, sich auf die Gerichte zu ‚verlassen‘. Effektiv wurde ihnen ein Großteil der Bewertungsarbeit anvertraut.

Die gebildete Überlieferung zeigte aber, dass die Gerichte diese Aufgabe nicht ausfüllen können. Es wurden kaum Fälle für die Abgabe an das Archiv gekennzeichnet. Die Autopsie wurde auf diese Fälle beschränkt. Das führte letztlich nur zu einer Ersatzüberlieferung einiger weniger exemplarischer Verfahren nach einem archivarischen ‚Bauchgefühl‘. Diese ‚manuellen‘ Methoden bedurften daher dringend einer Überarbeitung/Veränderung.

Die ersten ‚maschinellen‘ Methoden kamen in den 1990er Jahren auf, wurden zunächst für die mit präziseren Metadaten ausgestatteten Strafprozessakten verwendet und stetig weiterentwickelt. Auch bei den Zivilprozessakten sind die Metadaten der wichtigste Faktor, doch sind sie deutlich schlechter. Bei ihrer Bereitstellung aus den Fachverfahren muss das Archiv sich nach den Möglichkeiten der Justiz richten und oft Kompromisse eingehen.

Aufgrund der neuen Möglichkeiten werden die Bewertungsmodelle von den Archiven intensiv überarbeitet. Modell und Methode sind dabei interdependent, die derzeit zwei Ansätze maschineller Methoden entsprechen zwei verschiedenen Ausrichtungen der Modelle. Der erste, ‚exemplarische‘ Ansatz korrespondiert mit den bisherigen Modellen und zielt auf die Auswahl besonderer Akten mithilfe von Datenbankabgleichen und der Bildung eines exemplarischen Samples nach Kriterien wie besonders hohen Streitwerten. Genutzt wird dafür typischerweise SQL, mit dem komplexe Abfragen formuliert werden, das aber beispielsweise für Zufallsziehungen schlechter geeignet ist. Datenbankabgleiche sind vor allem wegen der schlechten Metadaten teilweise noch schwierig, zumal passende Datenbanken ausgewählt werden müssen. Der zweite, ‚repräsentative‘ Ansatz zielt nur auf ein repräsentatives Sample, das mit statistischen Instrumenten gebildet wird – er kommt ohne Datenbanken aus. Eine Stratifizierung nach Sachgebieten sichert, dass die Stichprobe auch kleine Sachgebiete enthält. Der exemplarische Ansatz leidet damit stärker unter technischen Problemen als der repräsentative, das aber eben auch wegen seines deutlich höheren Anspruchs aus dem Bewertungsmodell. Der repräsentative Ansatz folgt deshalb auch auf ein redefiniertes Bewertungsmodell, das die Abbildung von Alltagsgeschichte zum Ziel hat, und daher auf Repräsentativität zielt, auf den Schluss vom Einzelfall auf den typischen Fall.

Die Möglichkeiten komplexerer maschineller Werkzeuge machen die deutlich häufigere Evaluation der eigenen Modelle und Methoden sinnvoll. Aufgrund der zentralen

Bedeutung der bereitgestellten Metadaten ist der Kontakt zur Justiz weiter essenziell, aber in anderer Weise als früher. Die Wichtigkeit der ‚manuellen‘ Kennzeichnung durch einzelne Richter nimmt dank besserer maschineller Analyse- und Auswahlwerkzeuge ab. Sie sind künftig nur noch für solche Fälle wichtig, in denen wirklich eine intime Kenntnis des Falls, nicht nur der Metadaten, zu dessen Beurteilung notwendig ist. Wichtiger werden dafür die IT-Stellen der Justiz, die großen Einfluss auf eine bessere und einheitlichere Bereitstellung der Metadaten haben können.

Denn hierauf richten sich die Zukunftserwartungen der Archive: die Digitalisierung in der Justiz mit dem Austauschstandard XJustiz in der näheren und dem Gemeinsamen Fachverfahren für die Justiz (GeFa) in der fernerer Zukunft. Sie ermöglichen idealerweise die weitgehend automatisierte Anbietung aller Akten nach einem einheitlichen Standard. Damit würden zentrale Probleme der Archive gelöst. Auf Seiten der Archive ist möglicherweise ebenfalls eine Vereinheitlichung der Bewertungspraxis zu erwarten. Sie erfolgte bisher vor allem zwischen den verschiedenen Abteilungen desselben Landes-/Staatsarchivs. In technischer Hinsicht werden künftig alle Archive Interesse an Auswahlwerkzeugen haben, die mit XJustiz-Datensätzen arbeiten können. Möglicherweise nähern sich dadurch auch die Auswahlmethoden an, da viele Archive Interesse haben, voneinander zu lernen und in der Praxis eingesetzte Lösungen für sich zu adaptieren. Andererseits könnten sich auch zwei ‚Blöcke‘ bilden, etwa wenn der repräsentative Ansatz sich ebenfalls als attraktiv für Archive erweist, deren Modelle bisher eher zum exemplarischen Ansatz passen.

Nicht nur in Sachen Bereitstellung und Metadaten-Standardisierung, sondern auch im Bereich komplexerer Auswahlwerkzeuge schreitet die Technik fort. Die Bandbreite der Werkzeuge umfasst einfache Microservices (etwa mit Regular Expressions), Parser, Methoden der Data Science und die Verwendung von KI. KI-Modelle können etwa bereits mit allen Dokumenten trainiert werden, die auch ein menschlicher Archivar einbeziehen würde. Die oben genannten Werkzeuge benötigen allerdings die Volltexte digitaler Akten. Nach den Ergebnissen dieser Arbeit scheinen sich die Technologien bisher noch eher für die Extraktion zusätzlicher Metadaten anzubieten. Für eine Verknüpfung mit Datenbanken reichen die gegebenen/extrahierten Metadaten aber noch nicht aus. Außerdem dürften in den meisten Archiven schlicht Wissen und Kompetenzen fehlen, um derartige Technologien in die etablierten Methoden zu integrieren. Da bessere Metadaten schon allein dank der Trends in der Justiz zu erwarten sind, werden sich die etablierten Methoden wohl auch ohne anderweitige technische Ergänzungen verbessern. Möglicherweise bieten sich mit besseren Metadaten ausgestattete

Aktengruppen (etwa die Strafprozessakte) aber besser für Tests neuer Werkzeuge und Technologien an.

Die verschiedenen Komponenten der Zivilprozessakten-Bewertung waren in unterschiedlichen Geschwindigkeiten vorangeschritten, gleichen sich aber nun wieder an: die Bereitstellung der Metadaten zieht nach gegenüber den maschinellen Methoden. Inhaltliche Probleme mit den Metadaten bleiben jedoch teilweise; ob die Archive diese mit neuer Technologie lösen können, ist noch offen.

Technik und Methoden entwickeln sich zur Zeit rapide. ‚Dauerhaftere‘ Lektionen aus dieser Transferarbeit sind daher abstrakter zu formulieren. Die Hauptsache ist dabei das von Jenny Bunn beschriebene Denken in Systemen. Die Zivilprozessakten-Bewertung ist geprägt von starken Interdependenzen: das Archiv formuliert Bewertungsmodelle, wählt danach Methoden aus, die wiederum von den bereitgestellten Metadaten abhängen. Gleichzeitig muss es sich um bessere Metadaten bemühen, sonst erzeugen auch komplexe Methoden keine gute Überlieferung, da sie die Bewertungsmodelle dann unzureichend operationalisieren können. Bewertungsmodelle müssen sich an den Methoden orientieren, sonst führen sie nicht zu guter, tatsächlicher Überlieferung. Alle Teile müssen aufeinander abgestimmt sein, um eine gute Überlieferung zu erzeugen. Künftige Veränderungen und Fortschritte können sowohl aus dem Archiv kommen als auch aus der Justiz, wobei die Archive immer involviert sein sollten. Allerdings hat der repräsentative Ansatz viel geringere Anforderungen an das Bewertungsmodell und die Metadaten. Er könnte gerade wegen seiner Einfachheit attraktiv sein, auch wenn beide Ansätze valide sind. Entscheidend ist in beiden Fällen die Koordination der verschiedenen ‚System‘-Komponenten, die eine andauernde Herausforderung bleiben wird. Dafür können die Archive dank neuer Technik ihren inhaltlichen Ansprüchen besser gerecht werden und haben deutlich mehr Kontrolle über den Auswahlprozess. Die vom System genutzten/erzeugten Metadaten sollten dann auch für die Erschließung nachgenutzt werden. Insgesamt verändert sich die Rolle der Archivare: in Richtung eines Managers von Bewertungssystemen.

6. Zusammenfassung

Die Arbeit widmet sich den Zivilprozessakten, die durch ihre schlechten Metadaten bisher eine besondere Herausforderung unter den Unterlagen der Justiz sind. Die Archive mussten sich lange auf die Kennzeichnung von Akten durch die Richter verlassen, was zu unbefriedigenden Ergebnissen führte. ‚Maschinelle‘ Methoden erlauben dagegen die eigene, genauere Analyse und Auswahl archivwürdiger Akten. Die Arbeit untersucht mithilfe von Interviews den Entwicklungsstand und verschiedene Ansätze der maschinellen Methoden und bezieht mithilfe zweier weiterer Interviews auch Perspektiven der Forschung auf technische Neuerungen ein. Im Ergebnis zeigt sich, dass neue Gegebenheiten und technische Möglichkeiten eine hohe Veränderungsgeschwindigkeit bedingen. Zentral ist die Kontrolle über den Auswahlprozess, was auch das Aufgabenspektrum der Archivarinnen und Archivare verändert. Sie müssen hier zu Managern komplexer Systeme aus einer Vielzahl aufeinander abzustimmender Komponenten werden.

7. Anhang

7.1 Tabelle Archivwürdigkeit bei Zivilprozessakten

Die folgende Tabelle stellt Merkmale für Archivwürdigkeit in Modellen verschiedener Bundesländer dar. Sie ist leicht vereinfacht, um Gemeinsamkeiten und Unterschiede zu illustrieren.

Registerzeichen C, O, OH

Inhaltliche Kriterien	BY	NW	HS	BW	Erläuterungen, Besonderheiten
Medienresonanz/ Pressemitteilungen der Gerichte		X	X	X	
Prominente Personen	X	X	X	X	
BGH-Entscheidung		X	X		
Besondere Beteiligte (BRD, Land, Vereine, Kirchen, Parteien, Gewerkschaften)	X	X	X	X	BW: v.a. auch betr. deren Besitz- und Rechtsverhältnisse; BY: v.a. auch betr. Innere Angelegenheiten von Vereinen, Körperschaften, Stiftungen
Hohe Streitwerte		X			NW: ab 500.000 Euro
Regional- oder zeittypische Sachverhalte	X	X	X	X	NW: 8 ausgewählte LG; HS: nur allgemein gehalten; BW: Akten ,landesgeschichtlicher Bedeutung'; BY: Akten mit Hinweisen auf Volkssitte, Brauchtum, zeittypische Sachverhalte
Bedeutung für Wissenschaft, Kunst etc. sowie berechnete Belange Betroffener, Urheberrechte, Patente, freie Meinungsäußerung, Verbote	X			X	BY: besonders auch betr. Streit über Namensrechte, Erbrechte, Familienstiftungen
Nachbarrechte, dingliche Rechte, Dienstbarkeiten etc.	X			X	BY: besonders auch betr. Eigentum an Grundstücken, Rechte und Dienstbarkeiten, Privilegien
Karten, Abbildungen Skizzen wichtiger Gebäude				X	
Bedeutung für die Justiz, Rechtsgeschichte	X		X	X	BY: u.a. wenn Landesrecht von vor 1900 angewendet wird
Einzelne Sachgebiete o.ä.		X	X		HS: 5 Wohneigentumsverfahren pro Jahr; NW: v.a. Auswahl einzelner Kammern mit besonderem Profil
Auswahlmodus					
Einzelanforderung durch die Archive		X	X		BW, BY: nicht gesondert erwähnt
Durch die Justiz gekennzeichnete Akten	X	X	X	X	
Titelsammlungen		X	X		NW, HS: v.a. von bestimmten Kammern der LG, die besondere Sachgebiete verhandeln; NW: Titelsammlungen aus drei AG; BW, BY: nicht gesondert erwähnt

Tabelle 2: Merkmale für Archivwürdigkeit¹³⁶

¹³⁶ Vom Autor erstellt. Besondere Aktentypen der Registerzeichen C, O, OH, die in der Tabelle nicht wiedergegeben werden: Wiedergutmachungssachen, Aktiengesetzsachen, Wertpapierbereinigungssachen, Patentrechtssachen für NW, HS. Basis für die Tabelle sind: Bayerisches Staatsministerium der Justiz, Aussonderungsbekanntmachung

7.2 Literaturverzeichnis

Interviews:

Jenny Bunn, Interview mit Jenny Bunn für The National Archives (UK), 23.02.2024.

Bastian Gillner, Interview für das Landesarchiv Nordrhein-Westfalen am 15.01.2024.

David Gniffke, Interview für das Hessische Landesarchiv am 12.01.2024.

Elke Koch, Interview für das Landesarchiv Baden-Württemberg am 04.01.2024 2024.

Basma Makhlouf Shabou, Interview mit Basma Makhlouf Shabou, 26.02.2024.

Michael Unger, Interview für das Bayerische Staatsarchiv am 27.02.2024.

Internetquellen:

Website Gemeinsames Fachverfahren für die Justiz, <https://gefa-justiz.de/was-ist-gefa> (16.02.2024).

Website XJustiz, https://xjustiz.justiz.de/projektskizze/index.php#ausgangspunkt_ziele (25.03.2024).

Graue Literatur (auch digital):

Bayerisches Staatsministerium der Justiz, **Aussonderungsbekanntmachung Justiz** vom 27. April 1994 (JMBL. S. 71), die zuletzt durch Bekanntmachung vom 27. Mai 2013 (JMBL. S. 44) geändert wurde, <https://www.gesetze-bayern.de/Content/Document/BayVwV154213/tru> (28.02.2024).

Gniffke, David, **Masterplan** Bewertungs- und Übernahmestrategie durch Bewertungsmodelle. Version 3.0, 2021, https://landesarchiv.hessen.de/sites/landesarchiv.hessen.de/files/2022-11/masterplan_bewertungsmodelle_3.0.pdf (25.03.2024).

Haberkorn, Eva/Herge, Ina/Laube, Marina/Maaß, Rainer/Petter, Dirk/Zilien, Johann/Gniffke, David, **Teilbewertungsmodell** für die hessische Justiz. Landgerichte Amtsgerichte Staatsanwaltschaften. Genehmigt durch den Präsidenten des Hessischen Landesarchivs am 29.05.2019, zuletzt aktualisiert von David Gniffke am 01.08.2020, 2020, https://landesarchiv.hessen.de/sites/landesarchiv.hessen.de/files/2022-11/teilmodell_amtsgerichte_staatsanwaltschaften_landgerichte.pdf.

Heck, Andrea/Stabel, Florian/Laube, Marina/Rödel, Eva, **Handreichung** zur Aktenaussonderung und Aktenanbietung für Amtsgerichte. Stand 08.04.2020, 2020.

Hessisches Statistisches Landesamt, Tätigkeit der **Zivilgerichte** in Hessen 2014 bis 2022, <https://statistik.hessen.de/unsere-zahlen/rechtspflege> (22.03.2024).

Landesarchiv Baden-Württemberg, **Erweitertes Auswahlmodell** bei Massenakten der Justiz. Stand: 14.11.2022, https://www.landesarchiv-bw.de/sixcms/media.php/120/bewertung_massenakten_justiz_2022_11_14.pdf (02.11.2023).

Justiz; Landesarchiv Baden-Württemberg, Archivwürdige Zivilakten; Landesarchiv Baden-Württemberg, Erweitertes Auswahlmodell; Hessisches Landesarchiv, Handreichung; Haberkorn/Herge/Laube/Maaß/Petter/Zilien/Gniffke, Teilbewertungsmodell; Landesarchiv NRW, Abschlussbericht; Landesarchiv NRW, Richtlinien.

Landesarchiv Baden-Württemberg, **Archivwürdige Zivilakten**, 2023, https://www.landearchiv-bw.de/sixcms/media.php/120/Behoerden_Auswahl_Zivilakten.pdf (25.03.2024).

Landesarchiv NRW, **Richtlinien** zur Anbietung und Archivierung von Unterlagen der Justiz. Erarbeitet von der Projektgruppe „Archivierungsmodell Justiz“ des Landesarchivs Nordrhein-Westfalen Düsseldorf 2008. Version 1.3 (18.11.2020), https://www.archive.nrw.de/sites/default/files/media/files/Behoerdeninfo-Justiz-v1.3-November2020_0.pdf.

Landesarchiv NRW, **Abschlussbericht** der Projektgruppe Archivierungsmodell Justiz Düsseldorf, 2008. Version 1.4 (18.11.2020), 2020, <https://www.archive.nrw.de/sites/default/files/media/files/Archivierungsmodell-Personalverwaltung-v1.2-Oktober2020.pdf> (25.03.2024).

Welch, Nicki, **Find Case Law**: Reflecting on our first year of publishing judgments, part one, <https://blog.nationalarchives.gov.uk/find-case-law-reflecting-on-our-first-year-of-publishing-judgments-part-one/> (17.03.2024).

Sonstige Literatur:

Arbeitskreis "Archivische Bewertung" im Verband deutscher Archivarinnen und Archivare e. V. (VdA) (Hrsg.), Evaluierung von Bewertungsdokumenten. Beiträge zur archivischen Überlieferungsbildung. Stuttgart 2018.

Arbeitskreis "Archivische Bewertung" im Verband deutscher Archivarinnen und Archivare e. V. (VdA), **Kriterien** zur Evaluierung von Bewertungsdokumenten. Eine Handreichung zur archivischen Überlieferungsbildung, in: Arbeitskreis "Archivische Bewertung" im Verband deutscher Archivarinnen und Archivare e. V. (VdA) (Hrsg.), Evaluierung von Bewertungsdokumenten. Beiträge zur archivischen Überlieferungsbildung. Stuttgart 2018, 18–33.

Irmgard Christa Becker/Thomas Henne/Niklas Konzen/Robert Meier/Kai Naumann/Karsten Uhde (Hrsg.), E-Government und digitale Archivierung. Beiträge zum 23. Archivwissenschaftlichen Kolloquium der Archivschule Marburg. (Veröffentlichungen der Archivschule Marburg, Hochschule für Archivwissenschaft, Nr. 67). Marburg 2021.

Susanne Brenneisen, Throwing light into the **black box**? Alternative Optionen für die Bewertung von Zivilprozessakten im Vorfeld der Einführung der elektronischen Gerichtsakte in Baden-Württemberg. Transferarbeit im Rahmen der Laufbahnprüfung für den Höheren Archivdienst an der Archivschule Marburg (51. Wissenschaftlicher Lehrgang). Marburg 2018.

Caroline Brown (Hrsg.), Archival futures. London 2018.

Matthias Buchholz, Archivische **Überlieferungsbildung** im Spiegel von Bewertungsdiskussion und Repräsentativität. Eine Fallstudie am Beispiel von Sozialhilfeakten der oberbergischen Gemeinde Lindlar. Zugl.: Berlin, Humboldt-Univ., Diss., 2001 u.d.T.: Buchholz, Matthias: Überlieferungsbildung bei massenhaft gleichförmigen Einzelfallakten im Spannungsverhältnis von Bewertungsdiskussion, Repräsentativität und Nutzungsperspektive. (Archivhefte / Landschaftsverband Rheinland, Rheinisches Archiv- und Museumsamt, Archivberatungsstelle, Bd. 35). Köln 2001.

Matthias Buchholz, **Archivische Überlieferungsbildung** im Spiegel von Bewertungsdiskussion und Repräsentativität. (Archivhefte, Bd. 35). 2. Aufl. Köln 2011.

- Matthias Buchholz, Statistische **Methoden** als Werkzeug der Überlieferungsbildung bei "Massenakten", in: Brandenburgische Archive, 2014.
- Jenny Bunn, **Frames** and the future of archival processing, in: Caroline Brown (Hrsg.), Archival futures. London 2018, 65–77.
- Jenny Bunn, **AI** for Appraisal and Selection: A personal reflection, in: Arbido, 2023, o.S., <https://arbido.ch/de/ausgaben-artikel/2023/archiv-der-zukunft/ai-for-appraisal-and-selection-a-personal-reflection> (28.02.2024).
- Gregor Gehrke, **Zivilprozessakten** der Amts- und Landgerichte ab 1879, in: Jens Heckl (Hrsg.), Unbekannte Quellen: "Massenakten" des 20. Jahrhunderts. Untersuchungen seriellen Schriftguts aus normierten Verwaltungsverfahren. Band 2. (Veröffentlichungen des Landesarchivs Nordrhein-Westfalen, Bd. 43) Düsseldorf 2012, 142–151.
- Jens Heckl (Hrsg.), Unbekannte Quellen: "Massenakten" des 20. Jahrhunderts. Untersuchungen seriellen Schriftguts aus normierten Verwaltungsverfahren. Band 2. (Veröffentlichungen des Landesarchivs Nordrhein-Westfalen, Bd. 43). Düsseldorf 2012.
- Elke Koch/Kai Naumann/Jochen Rees/Annette Riek/Sabine Schnell/Franz-Josef Ziwes, **Bewertungsautomat** statt Autopsie: Sind jetzt zehntausend Akten in zehn Sekunden bewertet?, in: ARCHIVAR 70, 2017, 173–177.
- National Archives, The **application** of technology-assisted review to born-digital records transfer, Inquiries and beyond 2016.
- National Archives, **Using AI** for Digital Records Selection in Government. Guidance for records managers based on an evaluation of current marketplace solutions 2021.
- Kai Naumann, **Neues** vom Bewertungsautomaten. Workshop über Selesta in Stuttgart und Ludwigsburg, in: ARCHIVAR 73, 2020, 63–64.
- Kai Naumann/Elke Koch, **Bewertungsautomat** statt Autopsie: Neue Möglichkeiten und ihre Konsequenzen am Beispiel der Strafjustiz, in: Verband deutscher Archivarinnen und Archivare e.V. (Hrsg.), Massenakten - Massendaten. Rationalisierung und Automatisierung im Archiv. (Tagungsdokumentationen, Bd. 22) Fulda 2018, 35-52.
- Verena Schweizer, **E-Akte** - XJustiz - Fachverfahren. Entwicklung eines Aussonderungsworkflows für die E-Akte Justiz in Baden-Württemberg, in: Irmgard Christa Becker/Thomas Henne/Niklas Konzen/Robert Meier/Kai Naumann/Karsten Uhde (Hrsg.), E-Government und digitale Archivierung. Beiträge zum 23. Archivwissenschaftlichen Kolloquium der Archivschule Marburg. (Veröffentlichungen der Archivschule Marburg, Hochschule für Archivwissenschaft Nr. 67) Marburg 2021, 119–128.
- Rainer Stahlschmidt (Hrsg.), **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, Bd. 2). Düsseldorf 1999.
- Verband deutscher Archivarinnen und Archivare e.V. (Hrsg.), Massenakten - Massendaten. Rationalisierung und Automatisierung im Archiv. (Tagungsdokumentationen, Bd. 22). Fulda 2018
- Andrea Wendenburg, **Prüfen!** Prüfen! Prüfen! Sind inhaltliche Ziele in der Überlieferungsbildung erreichbar? Oder: Drum prüfe, wer sich ewig bindet!, in: Arbeitskreis "Archivische Bewertung" im Verband deutscher Archivarinnen und Archivare e. V. (VdA) (Hrsg.), Evaluierung von Bewertungsdokumenten. Beiträge zur archivischen Überlieferungsbildung. Stuttgart 2018, 44–47.

Geoffrey Yeo, Can we keep everything? The **future** of appraisal in a world of digital profusion, in: Caroline Brown (Hrsg.), *Archival futures*. London 2018.

Franz-Josef Ziwes, Bewertung zwischen Fingerspitzengefühl und e-Skills. Strategien zur Bewältigung einer archivischen Kernaufgabe, in: Katharina Ernst/Peter Müller (Hrsg.), *Aktuelle Fragen der Überlieferungsbildung*. Vorträge des 79. Südwestdeutschen Archivtags am 16. und 17. Mai 2019 in Ludwigsburg. Stuttgart 2020.