

Hanjo Hamann, Marisa Nest *

Justizforschung und Digitalisierung - Neue Wege zur Erschließung der Personalgeschichte des Bundesgerichtshofs im Internet

JurPC Web-Dok. 57/2018, Abs. 1 - 87

I. Die Namen der Justiz	Abs. 1
II. ... vom Papier ins pdf-Format	Abs. 2
1. Monographien (Festschriften)	Abs. 3
2. Amtliches Verkündungsblatt (BAnz)	Abs. 4
3. Zeitschriftenbeiträge (NJW)	Abs. 5
4. Gerichtsakten (Bundesgerichtshof)	Abs. 6
5. Entscheidungssammlung (BGHZ)	Abs. 7
6. Zwischenergebnis	Abs. 8
III. ... konvertiert in Tabellendokumente	Abs. 9
1. NJW und BGHZ	Abs. 10
2. Festschrift und Geschäftsverteilungen	Abs. 11
IV. ... in ein Linked Open Data-Repository	Abs. 12
1. Auswahl eines Datenrepositoriums	Abs. 13
2. Datenmodell von Wikidata	Abs. 14
3. Import der vorhandenen Daten	Abs. 15
4. Resultat und Nutzungspotential	Abs. 16
5. Künftige Aktualisierungen	Abs. 17
V. Zusammenfassung und Ausblick	Abs. 18

Abs. 19

Abs. 20

I. Die Namen der Justiz...

Die traditionelle europäische Vorstellung vom kodifizierten Gesetzesrecht, das durch Richter lediglich ausgesprochen oder punktuell konkretisiert wird, lässt wenig Spielraum für die Berücksichtigung individueller Richterpersönlichkeiten. Wohl auch deshalb sind Richterbiographien in Deutschland kaum zugänglich und jedenfalls weniger gut erforscht als in anderen Rechtsordnungen.[1] Dabei lässt sich gerade in Deutschland eine Hinwendung zur richterrechtlichen Methodik beobachten,[2] die dazu ermutigt, sich mit der Personalgeschichte der Justiz differenzierter als bisher auseinanderzusetzen.

Abs. 21

Solche Justizforschung steht indes vor der Schwierigkeit, die Personalien und Lebensdaten der Richter aus den verschiedenen Jahrzehnten der Bundesrepublik überhaupt zusammenzutragen. Sogar für die obersten Bundesgerichte ließ sich im Internet bisher allenfalls ein Bruchteil der Richterpersonalien abrufen.[3] Hier konnte das zwischen September 2016 und März 2017 durchgeführte Editionsprojekt „Die Namen der Justiz“ Abhilfe schaffen. Über seine Ergebnisse, die unter www.Richter-im-Internet.de dokumentiert und für die Öffentlichkeit zugänglich gemacht sind, wurde bereits andernorts ausführlich berichtet:[4] Es stellte erstmals die Senatsbesetzungen der obersten Bundesgerichte sowie des Bundespatent- und verfassungsgerichts im Internet zur Verfügung – für den Bundesgerichtshof sind nunmehr sämtliche Richterpersonalien seit seiner Gründung erstmals im Internet abrufbar. Dafür mussten ganz verschiedene Quellen erschlossen und neue Wege der digitalen Verfügbarmachung beschritten werden, die der vorliegende Text nun am Beispiel des Bundesgerichtshofs erläutert.[5]

Abs. 22

Er versteht sich als Anregung und Arbeitshilfe zur Nutzung digitaler Arbeitsprozesse in der rechtsgeschichtlichen Forschung – einer Methodeninnovation, die andere Disziplinen derzeit unter dem Stichwort digital humanities verhandeln. Dies trägt dazu bei, die durch Rechtswissenschaftler immer öfter eingeforderte[6] freie(re) Verfügbarkeit amtlicher Justizdaten zu verbessern, und erleichtert zugleich die Erforschung des Rechtssystems in anderen Fachdisziplinen, die dafür auf offen verfügbare Forschungsdaten (open data) angewiesen sind.[7] Dass die Mitwirkung von Rechtswissenschaftler(inne)n an solchen Bestrebungen als zeitgemäß und gesellschaftlich sinnvoll wahrgenommen wird, belegen nicht zuletzt die Finanzierung des vorliegenden Projekt durch das dotierte Fellowship „Freies Wissen“ des Stifterverbands für die Deutsche Wissenschaft[8] und die Anerkennung der aus dem Projekt entstandenen Informationsplattform www.Richter-im-Internet.de durch Bundespräsident Steinmeier als einer von 17 „Ausgezeichneten Orten“ der deutschen Wissenschaft 2017.[9]

Abs. 23

Auf welchen Wegen nun wurde dieser Ort erreicht?

Abs. 24

Abs. 25

II. ... vom Papier ins pdf-Format...

Zunächst mussten die bislang auf Papier gebannten Richterpersonalien ausfindig gemacht und digitalisiert werden. Dafür wurden verschiedene Quellen herangezogen:

Abs. 26

Abs. 27

1. Monographien (Festschriften)

Einen Anlaufpunkt bildeten die beiden Festschriften, die beim Bundesgerichtshof anlässlich des 25. und 50. Gründungsjubiläums herausgegeben wurden (im Folgenden FS25 bzw. FS50).[10] Sie enthalten einen „Statistischen Teil“ (FS25) bzw. „Personalien und Geschäftsstatistiken“ (FS50) mit einer fortlaufenden Tabelle aller Richter und ihrer jeweiligen Lebens- und Amtsdaten.

Abs. 28

Beide Tabellenwerke sind im Aufbau identisch, das spätere aktualisiert und erweitert lediglich das frühere. Ein manueller Stichprobenabgleich beider Tabellen ergab zwar einige Widersprüche, diese konnten aber zügig ausgeräumt werden.[11]

Nach Klärung dieser Widersprüche wurde die „Personalien“-Tabelle aus Anhang I.A der FS50 (S. 787–832) als Datenquelle ausgewählt. Da die Urfassung dieser Tabelle und die ihr zugrundeliegenden Daten weder beim Verlag noch bei den Herausgebern erhalten geblieben sind, musste sie manuell aus der Druckfassung digitalisiert werden. Dazu wurden die Papierseiten eingescannt und mittels optischer Texterkennung (OCR) in bearbeitbaren Text für den nächsten Bearbeitungsschritt (IV.) umgewandelt.

Abs. 29

Abs. 30

2. Amtliches Verkündungsblatt (BANz)

Da in der FS50-Tabelle zwar Lebensdaten der Richter(innen) und die Rahmendaten ihrer Amtszeit erfasst waren, nicht aber ihre Zugehörigkeit zu einem bestimmten Senat, musste als weitere Quelle auf die sog. Geschäftsverteilungspläne zurückgegriffen werden, in denen für jeden Senat regelmäßig (meist zu Jahresbeginn[12]) seine Zuständigkeiten und Personalbesetzung festgelegt werden. Diese Pläne wurden seit dem Zweiten Weltkrieg jährlich im amtlichen Verkündungsblatt „Bundesanzeiger“ veröffentlicht. Die genauen Seitenangaben lassen sich aus einer öffentlich einsehbaren Fundstellenliste beim Bundesgerichtshof entnehmen.[13] Für die Jahre ab 2013 waren die Geschäftsverteilungspläne bereits im Internet abrufbar,[14] mussten also nicht erst digitalisiert werden. Sie konnten als pdf-Dokumente abgerufen und ohne separate Texterkennung als Grundstock für den Datenbestand genutzt werden. Dadurch lagen die ersten sechs Geschäftsverteilungspläne des BGH (2013–2018) bereits im pdf-Format vor.

Abs. 31

Für die weiter zurückliegenden Geschäftsverteilungspläne kam dem Digitalisierungsvorhaben eine Änderung im Jahr 1967 zugute: Seit jenem Jahr veröffentlichten die obersten Gerichtshöfe des Bundes ihre Geschäftsverteilungspläne gesammelt in einem jährlichen Sonderheft (Beilage) des Bundesanzeigers, das sich aufgrund seines A4-Formats gut für die Erfassung mit handelsüblichen Scannern eignet.

Abs. 32

Es stellte sich allerdings heraus, dass die Beilagenhefte in keiner Bibliothek[15] als lose Hefte verfügbar waren, sondern fest eingebunden in das damals im Format einer Tageszeitung publizierte Verkündungsblatt. Diese gebundenen Bestände konnten schon wegen der starken Mittelwölbung nicht mit der für die spätere Texterkennung erforderlichen Qualität digitalisiert werden. Die von manchen Bibliotheken zusätzlich vorgehaltenen Mikrofiches stellten sich als überwiegend unscharf oder verrauscht heraus. Deshalb musste letztlich der Verlag des Verkündungsblattes ersucht werden, mit seinen hauseigenen Archivexemplaren der Beilagen auszuweichen. Diese wurden aufgrund eines Werkvertrags ins pdf-Format digitalisiert und sogleich mit einer Texterkennung (OCR) behandelt. Da der Digitalisierungsauftrag insgesamt 1.336 Seiten umfasste, waren vereinzelte Fehler unvermeidlich und mussten in einer zweiten Runde durch erneute Digitalisierung von zwölf zuvor fehlerhaft digitalisierten Seiten berichtigt werden. Zudem war die Beilage für das Jahr 1976 im Verlagsarchiv nicht mehr auffindbar und wurde nachträglich aus Bibliotheksbeständen der Universität Mannheim digitalisiert.[16]

Abs. 33

Für die vor 1967 veröffentlichten Geschäftsverteilungspläne erwies sich die Digitalisierung als prohibitiv schwierig: Diese Dokumente waren im großformatigen Hauptteil des Bundesanzeigers abgedruckt worden, auf inzwischen stark gealtertem und brüchigen Zeitungspapier, eingebunden in Folianten, die für handelsübliche Digitalisierungsgeräte zu groß und wegen ihrer Mittelwölbung und mangelhaften Bindung nicht in der nötigen Qualität digitalisierbar waren. Einen anschaulichen Eindruck hiervon vermitteln die beiden Geschäftsverteilungspläne von 1951 und 1952,[17] die aus einem solchen gebundenen Bibliotheksexemplar abfotografiert werden mussten, weil sie anderweitig nicht erhältlich waren.

Abs. 34

Damit konnten aus dem Verkündungsblatt zunächst 47 Geschäftsverteilungspläne für die Jahre 1951/52 und 1967–2012 erfasst und ins pdf-Format digitalisiert werden. Abs. 35

Abs. 36

3. Zeitschriftenbeiträge (NJW)

Da die Geschäftsverteilungspläne für 1951 und 1952 lediglich die Zuständigkeiten der Senate dokumentierten, nicht aber deren Personalbesetzung, und weil für 1950 laut BGH-Fundstellenliste kein Geschäftsverteilungsplan amtlich veröffentlicht wurde, musste für die Anfangsjahre des Gerichts zusätzlich auf eine Fachzeitschrift zurückgegriffen werden: In den Jahrgängen 1950–1952 der Neuen Juristischen Wochenschrift (NJW) fanden sich sowohl ein Bericht eines der ersten BGH-Anwälte über die Geschäftsverteilung bei Gründung des Bundesgerichtshofes.[18] als auch alle Neubesetzungen der BGH-Anfangsjahre in der Rubrik „Mitteilungen“.[19] Diese wurden ebenfalls digitalisiert und für die weitere Bearbeitung genutzt. Abs. 37

Abs. 38

4. Gerichtsakten (Bundesgerichtshof)

Ein wahrer Datenschatz zur Ergänzung und Überprüfung der bis dahin gefundenen Datenquellen ließ sich beim Bundesgerichtshof selbst heben: Nach Beginn des hier skizzierten Editionsprojekts hatte der BGH im Oktober 2016 auch hausintern eine Digitalisierung seiner in Papierakten erhaltenen Geschäftsverteilungspläne durchgeführt. Sie sollte zwar nur der eigenen Bestandssicherung dienen („Eine Veröffentlichung der alten Geschäftspläne auf der Website des BGH ist nicht geplant. Intern werden weiterhin die Papierexemplare genutzt.“[20]), doch wurden ihre Ergebnisse auf Nachfrage bereitwillig zur Verfügung gestellt. Anhand des übermittelten Datenträgers mit pdf-Dokumenten konnte nun auch die Bestandslücke von 1953 bis 1966 mit Digitalisaten der noch auf Schreibmaschine getippten Geschäftsverteilungspläne aus dem Aktenbestand des BGH geschlossen werden. Abs. 39

Abs. 40

5. Entscheidungssammlung (BGHZ)

Zudem hatten die Mitarbeiter des BGH aus den Entscheidungssammlungen der Anfangsjahre, für die keine Senatsbesetzungen dokumentiert worden waren (1950–1952), eine Liste von Richtern mit ihrer jeweiligen Beteiligung an verschiedenen Senaten rekonstruiert, die ebenfalls im pdf-Format übermittelt wurden und damit für die weitere Bearbeitung genutzt werden konnten. Abs. 41

Abs. 42

6. Zwischenergebnis

Aus den genannten Quellen ließen sich also Dokumente digitalisieren, die für jeden BGH-Richter im Zeitraum 1950–2000 Lebens- und Amtsdaten sowie seine Senatszugehörigkeiten ermitteln ließen. Für die Jahre seit 2001, die nach dem Erscheinen der FS50 lagen, ließen sich aus den Geschäftsverteilungsplänen immerhin noch Namen, persönliche Daten (insb. Geschlecht und akademische Grade) und Senatszugehörigkeiten digitalisieren. Damit lagen zuletzt pdf-Dokumente im Gesamtumfang von 1.085 Seiten vor.[21] Abs. 43

Abs. 44

Der Großteil der so digitalisierten Dokumente – nämlich die als amtliche Werke nach § 5 UrhG gemeinfreien Geschäftsverteilungspläne – wurde sodann auf dem eigens eingerichteten Internetportal www.Richter-im-Internet.de öffentlich zugänglich gemacht und zugleich als Forschungsdaten im Repositorium ²Dok des Fachinformationsdienstes für internationale und interdisziplinäre Rechtsforschung archiviert (intr2dok.vifa-recht.de). Dort wurde jedem Geschäftsverteilungsplan zugleich ein persistenter, d.h. global eindeutiger und unveränderlicher,

Dokumenten-Identifikator (DOI) auf Grundlage des Uploaddatums zugewiesen,[22] der eine dauerhafte Zitierbarkeit des Dokuments gewährleistet.

Abs. 45

III. ... konvertiert in Tabellendokumente ...

Sodann wurden mit Unterstützung wissenschaftlicher Hilfskräfte alle pdf-Digitalisate zur weiteren Bearbeitung in Tabellendokumente übertragen.

Abs. 46

Abs. 47

1. NJW und BGHZ

Zunächst wurden die in den NJW-Mitteilungen abgedruckten Richterpersonalien von Hand in eine Tabelle eingetragen. Daraus entstand ein Datensatz von 59 Bundesrichtern und zwei Bundesrichterrinnen (Gerda Krüger-Nieland und Elisabeth Krumme), die bis 1951 bestellt worden waren.

Abs. 48

Auch die aus der Entscheidungssammlung (BGHZ) gesammelten Richterpersonalien wurden aus dem pdf-Dokument in ein Tabellenformat überführt, das danach 301 Datensätze für die Jahre 1950–1952 enthielt, also im Durchschnitt 11,1 Namen pro Senat und Jahr. Insgesamt wurden 80 verschiedene Personen erfasst: 29 Straf- und 47 Zivilrichter sowie Herbert Kleinewefers, der 1951 sowohl in zwei Strafsenaten als auch einem Zivilsenat mitgewirkt hatte, und je eine Straf- (Else Koffka) und eine Zivilrichterin (Gerda Krüger-Nieland) sowie Elisabeth Krumme, die 1952 sowohl in zwei Strafsenaten als auch einem Zivilsenat mitgewirkt hatte.[23] Die übrigen Richter waren pro Geschäftsjahr an je bis zu drei verschiedenen Zivilsenaten (vier Richter) oder vier verschiedenen Strafsenaten (neun Richter und Elisabeth Krumme 1950/51) tätig gewesen.

Abs. 49

Im Anschluss daran wurden die beiden erstellten Datensätze zusammengeführt und abgeglichen. Es ergab sich eine Schnittmenge von 57 Richtern und zwei Richterinnen, deren Benennung sowohl in der NJW mitgeteilt als auch in mindestens einer amtlich veröffentlichten Entscheidung dokumentiert wurde. 21 weitere Richter waren nur in den Entscheidungssammlungen, aber nicht in der NJW auffindbar, obwohl mindestens drei davon in beiden Quellen hätten auftauchen sollen.[24] Umgekehrt benannten die NJW-Mitteilungen nur einen Richter, dessen Mitwirkung an keiner amtlich veröffentlichten Entscheidung in den Jahren 1950–1952 belegt ist,[25] sowie einen weiteren, den es wohl nie gab.[26] Damit konnte die aus den Entscheidungssammlungen rekonstruierte Liste als maßgeblich behandelt und für die weitere Bearbeitung zugrunde gelegt werden.

Abs. 50

Abs. 51

2. Festschrift und Geschäftsverteilungen

Die Digitalisate aus der FS50 sowie aus den Geschäftsverteilungsplänen für die Zivilsenate wurden durch Hilfskräfte von Hand in Tabellendokumente übertragen, um aus den pdf-Dokumenten die gemeinfreien Daten zu extrahieren. In diesen Tabellen wurde größter Wert auf historische Authentizität gelegt, weshalb die Übertragung zunächst 1:1 textkonkordant erfolgte und ggf. festgestellte Fehler bzw. Unstimmigkeiten lediglich in einer zusätzlichen Tabellenspalte notiert wurden. Die Tabelle enthält also genau denselben Text wie die zugrundeliegenden pdf-Dokumente, nur überführt in eine explizite Tabellenstruktur und erweitert um die Spalte „editorische Anmerkungen“. Deshalb wurde sie als „Lesefassung“ bezeichnet und kann zum schnelleren Nachschlagen anstelle der Ausgangsdokumente verwendet werden.

Abs. 52

Da im Zuge dieser Datenextraktion sowohl in der FS50-Tabelle[27] als auch in den Geschäftsverteilungsplänen[28] Fehler aufgefallen waren und die Zusammenführung beider Datenbestände anschließend noch drei Widersprüche ergab,[29] wurden diese Unstimmigkeiten jeweils durch Heranziehung zusätzlicher Quellen aufgelöst.[30] Diese zunächst auf dem

Abs. 53

Tabellenblatt „Lesefassung“ als gesonderte „editorische Anmerkungen“ aufgenommenen Fehlerkorrekturen wurden im Fall der Geschäftsverteilungspläne sodann in einem weiteren Tabellenblatt mit den maschinenlesbar kodierten Rohdaten zusammengeführt, so dass dieses Tabellenblatt die erste vollständige und berichtigte Fassung der BGH-Senatsbesetzungen wiedergibt.

Für den gesamten Bearbeitungsschritt wurden zusätzlich zur Arbeitskraft des Projektleiters 70 Hilfskraftstunden benötigt. Als Ergebnis konnten zwei Tabellendokumente verfügbar gemacht werden: „Senatsbesetzungen des BGH(Z) 1950–2018“ (gvp.xls, 1,35 MB) und „Personalien der BGH-Richter 1950–2000“ (fs50.xls, 0,2 MB). Beide Dokumente enthalten jeweils ein Tabellenblatt („Vorwort“) mit Lizenzhinweisen, Zitiervorschlag, Datenbeschreibung und editorischen Hinweisen, sowie zusammen insgesamt drei Tabellenblätter mit den eigentlichen Daten:

Abs. 54

Datei	Tabellenblatt	Zeilen	Spalten	Datensätze	Stand	csv?
gvp.xls	1, Lesefassung	6.077	9	5.362	2018	nein
	2, maschinenlesbar	5.289	12	5.288	2018	208 kB
fs50.xls	BGH-FS50	434	44	432	2000	59 kB

Abs. 55

Tab. 1: Umfang der zum 1.3.2018 unter www.Richter-im-internet.de verfügbaren Tabellenblätter

Die Datei gvp.xls enthält also eine Lesefassung der Geschäftsverteilungspläne (mit Spalten für Jahr und Senat sowie – für jeden Richter – Amtsbezeichnung,[31] etwaige akademische Grade und Namenszusätze, Nach- und Vorname sowie etwaige Anmerkungen[32]) sowie deren maschinenlesbare Fassung (mit sechs binären Indikatorvariablen für Geschlecht, Professorentitel, ehrenamtliche Doktorgrade, andere Doktorgrade sowie die Funktionen Senatsvorsitz und stellvertretender Senatsvorsitz).

Abs. 56

Die weitere Datei fs50.xls enthält die digitalisierte und infolge des Abgleichs mit den Geschäftsverteilungsplänen korrigierte Fassung des Tabellenwerks aus der BGH-Festschrift (FS50), die zugleich um Mehrfachnennungen in den Untertabellen „Präsidenten“, „Senatspräsidenten“ und „Bundesrichter“ bereinigt wurde, also je Richter(in) nur einen Datensatz aufweist.

Abs. 57

Die beiden letztgenannten Tabellenblätter wurden zudem als kommagetrennte Textdateien im UTF8-kodierten (unicode) csv-Format (comma-separated values) zur Verfügung gestellt, um deren plattform- und softwareübergreifende Nutzbarkeit und Formatkompatibilität zu gewährleisten.

Abs. 58

IV. ... in ein Linked Open Data-Repository

Abs. 59

Im letzten Schritt sollten die maschinenlesbaren Daten, die nun im Tabellenformat auf der Projektwebsite verfügbar waren, zusätzlich in ein Datenrepository übertragen werden, das sowohl weitere Schnittstellen zum Abruf der Daten bietet, als auch den Datensatz gemäß Linked Open Data (LOD)-Prinzipien erreichbar macht.[33] LOD bezeichnet strukturierte Daten, die im Internet für jeden zur freien Nutzung, Verbreitung und Weiterverwendung zur Verfügung stehen und darüber hinaus mit anderen Daten verlinkt werden. Ein Vorteil von LOD und insbesondere der Vernetzung der Daten ist, dass zu bereits abgefragten Informationen weitere verwandte Daten erschlossen werden können. So werden nicht nur die Daten sichtbar gemacht, sondern auch die Erzeugung neuen Wissens ermöglicht.

Abs. 60

1. Auswahl eines Datenrepositoriums

Die Datenrepositorien DBpedia (www.dbpedia.org), YAGO (www.yago-knowledge.org) und Wikidata (www.wikidata.org) folgen den LOD-Prinzipien, weisen aber in der Art, wie Daten importiert werden, Unterschiede auf.[34] So beziehen beispielsweise YAGO und DBpedia ihre Daten hauptsächlich aus der Online-Enzyklopädie Wikipedia,, indem die dort bereits strukturiert vorhandenen Informationen extrahiert und anschließend importiert werden; YAGO ergänzt diese Informationen mit Daten aus WordNet[35] und GeoNames.[36]

Abs. 62

Einen anderen Ansatz verfolgt Wikidata. Es werden zwar ebenfalls Daten aus Wikipedia-Artikeln gewonnen, es können aber auch beliebige andere Datensätze importiert werden. Aktuelle Importe reichen von diversen Datensätzen der UNESCO bis hin zu Betriebsstellenverzeichnissen der Deutschen Bahn.[37] Mithin ist der Datenimport bei Wikidata einfacher als bei DBpedia und YAGO, weil nicht zuerst bestehende Wikipedia-Artikel ergänzt oder neue angelegt werden müssen, um Daten importieren zu können. Zudem lässt sich beobachten, dass die Bedeutung von Wikidata innerhalb der Linked Data-Community kontinuierlich wächst. Beispielsweise überführte die Google-Tochter Metaweb Technologies ab Ende 2014 den gesamten Inhalt ihrer semantischen Datenbank Freebase in Wikidata.[38] Auch gesellschaftliche Institutionen und Organisationen wie das britische Staatsfernsehen (BBC) und das Museum für moderne Kunst in New York (MOMA) binden bisweilen Wikidata-Informationen in ihre Angebote ein.[39]

Abs. 63

Abs. 64

2. Datenmodell von Wikidata

Wikidata ist ein freies, strukturiertes Wissensrepositorium, das sowohl von Menschen, als auch von Maschinen bearbeitet und gelesen werden kann. Es hat seinen Ursprung in Wikipedia, der größten Online-Enzyklopädie, und wurde Ende 2012 mit dem Ziel gegründet, die in Wikipedia-Artikeln enthaltenen strukturierten Informationen, zentralisiert zu verwalten, um sie dann Wikipedia und anderen Wikimedia-Projekten zur Verfügung zu stellen. Da die Daten an zentraler Stelle gepflegt und aktualisiert werden, entfällt insbesondere deren gesonderte Aktualisierung in jeder Sprachversion von Wikipedia, was eine enorme Arbeitersparnis bedeutet. Zugleich können andere Internetprojekte und Dienste[40] die in Wikidata enthaltenen Daten z.B. über Schnittstellen (z.B. MediaWiki), Softwarepakete (z.B. pywikibot) und den Wikidata-SPARQL-Endpoint abrufen und frei weiterverwenden.

Abs. 65

Das Datenmodell von Wikidata[41] lässt sich wie folgt zusammenfassen: Auf Wikidata sind Daten als sogenannte Datenobjekte modelliert. Datenobjekte besitzen eine eindeutige Identifizierungsnummer (Id) und bilden sprachunabhängig ein Konzept ab. Ein Beispiel ist das Datenobjekt mit der Id Q26933282.[42] Dieses Objekt wird durch seine Charakteristika bestimmt, die als sog. Aussagen hinterlegt sind. Jedes Objekt kann beliebig viele Charakteristika haben, wobei jede Aussage ihrerseits aus drei Teilen besteht: Dem Datenobjekt, auf das sie sich bezieht, der Eigenschaft, die sie bezeichnet, und deren Wert. Eine solche Aussage lautet beispielsweise, dass das Datenobjekt Q26933282 die Eigenschaft „ist ein(e)“ mit dem Wert „Mensch“ hat. Folglich handelt es sich bei dem von Q26933282 abgebildeten Konzept um einen Menschen.

Abs. 66

Aussagen

Abs. 67

The image shows two Wikidata statement cards. The first card is for the property 'ist ein(e)' (P31) with the value 'Mensch' (Q5). It shows 0 instances and options to edit, add instances, or add values. The second card is for the property 'Geburtsdatum' (P569) with the value '10. Juni 1938'. It shows 1 instance and options to edit or add values.

Abb. 1: Aussagen über das Datenobjekt Q26933282 auf Wikidata (www.wikidata.org/wiki/Q26933282)

Die Eigenschaft „ist ein(e)“ ist wiederum durch eine Identifizierungsnummer (P31) repräsentiert und auch deren Wert besteht aus einer Id (hier Q5 für „Mensch“),[43] wodurch der gesamte Wissensbestand untereinander verknüpft wird.

Abs. 68

Neben diesen sprachunabhängig hinterlegten Informationen, die erst im Zuge ihrer Darstellung in einer bestimmten Sprache ausgegeben werden (Q5 wird dann zu „Mensch“ im Deutschen, zu „human“ im Englischen, usw.), werden einige Informationen auch in verschiedenen Sprachvarianten hinterlegt. So erhält jedes Datenobjekt eine Bezeichnung, Beschreibung und etwaige Bezeichnungsalternativen in beliebig vielen Sprachen. So hat das Datenobjekt Q26933282 in der deutschen Sprache die Bezeichnung „Falk Freiherr von Maltzahn“, die Beschreibung „deutscher Jurist und Richter am Bundesgerichtshof“ und die Bezeichnungsalternative „Falk von Maltzahn“. Erst aus diesen Informationen erfahren wir also, welche konkrete Person der Datensatz Q26933282 repräsentiert. Die Akte bekommt gewissermaßen einen Reiter.

Abs. 69

Falk Freiherr von Maltzahn (Q26933282)...

Abs. 70

The image shows the Wikidata page for 'Falk Freiherr von Maltzahn' (Q26933282). It displays language-dependent information in a table. The table has four columns: 'Sprache', 'Bezeichnung', 'Beschreibung', and 'Auch bekannt als'. The data is as follows:

Sprache	Bezeichnung	Beschreibung	Auch bekannt als
Deutsch	Falk Freiherr von Maltzahn	deutscher Jurist, Richter am Bundesgerichtshof	Falk von Maltzahn
Englisch	Falk von Maltzahn	Keine Beschreibung vorhanden	
Französisch	Keine Bezeichnung vorhanden	Keine Beschreibung vorhanden	
Bairisch	Keine Bezeichnung vorhanden	Keine Beschreibung vorhanden	

Below the table, there is a link 'Alle eingegebenen Sprachen'.

Abb. 2: Sprachabhängige Informationen über das Datenobjekt aus der vorigen Abbildung.

Abs. 71

3. Import der vorhandenen Daten

Bevor die bereits aufbereiteten Daten in das Wikidata-Datenschema übertragen wurden, waren sie zunächst durch Informationen zu ergänzen, die auf Wikipedia bereits unstrukturiert vorlagen. Hierfür wurde die „Liste der Richter am Bundesgerichtshof“ genutzt,[44] die ein Referent in der

Abs. 72

Landesdirektion Sachsen (Benutzer .Manu.) am 22. Juni 2013 angelegt hatte. Nach eigenen Angaben speist sich diese Liste aus dem jährlichen „Handbuch der Justiz“ des Deutschen Richterbunds, und sie enthält auch über das Jahr 2000 (Erscheinungsjahr der FS50) hinaus die Namen und Amtsdaten der an den BGH berufenen Richter. Dadurch konnten 159 Richter(innen) identifiziert werden, die im FS50-Datensatz und den Geschäftsverteilungsplänen der Zivilsenate nicht enthalten waren. Nach Aufnahme dieser Einträge sowie sechs neu bestellter Richter(innen), die in den Geschäftsverteilungsplänen 2017 und 2018 erstmals auftauchten, enthielt der Datensatz nunmehr 597 Personen.

Für jede dieser Personen wurde zunächst überprüft, ob sie bereits durch ein Datenobjekt auf Wikidata repräsentiert sind. Dafür wurden Eigenschaften der Richter(innen) ausgewählt, die sowohl die Person möglichst eindeutig beschreiben als auch in beiden Datensätzen zu finden sein würden. Eine Analyse bereits existierender Richter(innen)-Objekte auf Wikidata ergab, dass die meisten dieser Objekte bereits deutsche Bezeichnungen besitzen – den Namen der Person – sowie Beschreibungen mit den Wortbestandteilen Richter(in) oder Jurist(in). Als Aussagen waren oft das Geburtsdatum, eine Berufsangabe oder die Eigenschaft „Mensch“ hinterlegt (vgl. Tab. 2). Diese Eigenschaften wurden sodann für den Abgleich mit dem hier erzeugten Datensatz verwendet.

Abs. 73

Eigenschaft	Wert	von 50
ist ein	Mensch	50
Geburtsdatum	Geburtsdatum der Person	50
Tätigkeit	Richterin bzw. Richter	48
Beschreibung	enthält Richter(in) oder Jurist(in)	48
Bezeichnung	Name der Person	50

Abs. 74

Tab. 2: Stichprobe von 50 zufällig ausgewählten Richterinnen und Richtern aus dem Datensatz; geprüft wurde, wie oft die genannte Eigenschaften auf Wikidata mit einem Wert belegt sind.

Anhand der so identifizierten Eigenschaften konnte festgestellt werden, dass zu 352 Richter(innen) bereits ein Datenobjekt auf Wikidata existiert. Für die übrigen 245 Richter(innen) wurden neue Datenobjekte auf Wikidata angelegt, die im Folgenden mit Daten ergänzt wurden. Um die zu importierenden Daten richtig auf Wikidata abbilden zu können, mussten diese zunächst modelliert werden (vgl. Abb. 3).

Abs. 75

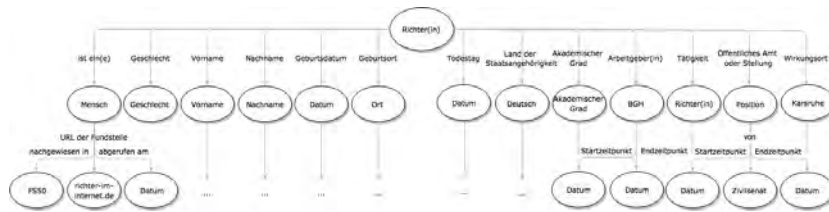
Dafür wurde der Name der Richterin oder des Richters inklusive eines etwaigen Namenszusatzes als Bezeichnung abgebildet, im Fall von Namenszusätzen jedoch zugleich der Name ohne Namenszusatz als Bezeichnungsalternative gesetzt. Zugleich erhielt jedes Datenobjekt eine Beschreibung in acht Sprachen, die sich an die häufigste schon auf Wikidata vorgefundene deutschsprachige Beschreibung anlehnte („deutscher Jurist, Richter[in] am Bundesgerichtshof“) und in die Sprachen Chinesisch, Englisch, Französisch, Italienisch, Niederländisch, Polnisch und Spanisch übersetzt wurde.

Abs. 76

Schließlich wurden als einzelne Aussagen zu jedem Datenobjekt die Eigenschaften Mensch, Geschlecht und Staatsangehörigkeit, Vor- und Nachname sowie akademischer Grad, Geburtsdatum und -ort sowie Todestag, Beruf (Richter/in), Arbeitgeber (BGH) und Ort (Karlsruhe), sowie die Senatszugehörigkeit und die darin ausgeübten Positionen (Vorsitz/stellvertretender Vorsitz/Beisitz) ergänzt. Für Arbeitgeber, Senat und ausgeübte Positionen wurden jeweils Konkretisierungen (sog. Qualifikatoren) erfasst, die Anfangs- und Enddatum der jeweiligen Tätigkeit enthalten. Jede Aussage wurde zugleich mit der jeweiligen Quelle der Daten belegt, also Richter-im-Internet.de (Q32961325) oder der deutschsprachigen

Abs. 77

Wikipedia (Q48183).



Abs. 78

Abb. 3: Concept Map für die Modellierung der Daten

Basierend auf dieser Modellierung der Daten wurde der eigentliche Import unter Einsatz des pywikibot-Frameworks[45] durchgeführt.

Abs. 79

Abs. 80

4. Resultat und Nutzungspotential

Nachdem für zwei Fünftel (245/597) der bisherigen BGH-Richter(innen) neue Datenobjekte in Wikidata angelegt wurden, ergab ein erneuter Abgleich, dass nunmehr alle 597 bekannten Richterinnen und Richter des Bundesgerichtshofs auf Wikidata vertreten und mit einem Datenprofil des o.g. Umfangs dokumentiert sind.

Abs. 81

Die Daten können nun in verschiedenen Formen sowohl von Menschen als auch von Maschinen abgerufen werden.[46] Manuell kann beispielsweise über den Browser direkt auf Datenobjekte zugegriffen werden, wenn die Id des Datenobjektes bekannt ist, oder die Suchfunktion unter www.wikidata.org genutzt werden. Zudem bietet Wikidata einen sog. SPARQL-Endpoint (query.wikidata.org), also eine Schnittstelle, über die die Daten im RDF-Format mittels der graphenbasierten Abfragesprache SPARQL abgefragt werden können. Dies erlaubt es, komplexere Suchmuster zu definieren, als sie mit einer Volltextsuche möglich wären:

Abs. 82

```
#defaultView:Timeline
SELECT ?item ?itemLabel ?startTime ?endTime
WHERE {
  ?item wdt:P106 wd:Q16533 ;
        wdt:P21 wd:Q6581072 ;
        wdt:P569 ?dateOfBirth ;
        p:P108 [ ps:P108 wd:Q687323; pq:P580 ?startTime ; pq:P582 ?endTime ] .

  BIND(ROUND(FLOOR((?startTime - ?dateOfBirth)/365)) AS ?age)
  FILTER(?age < 50 && 1979 < YEAR(?startTime) && YEAR(?startTime) < 1990).
  SERVICE wikibase:label { bd:serviceParam wikibase:language "de". }
}
```

Abs. 83

Abb. 4: Anfrage an den SPARQL-Endpoint (tinyurl.com/y93hjnvh) zur Ermittlung aller Frauen unter 50 Jahren, die in den 1980ern am BGH eingestellt wurden.

Abs. 84

5. Künftige Aktualisierungen

Um die nun verfügbaren Daten auf aktuellem Stand zu halten, müsste jede(r) neu ins Amt berufene Richter(in) des BGH nachgetragen werden – entweder durch manuellen Nachtrag (wikidata.org – „Ein neues Datenobjekt erstellen“), da es sich hierbei um keine umfangreichen neuen Datensätze handelt, oder aber durch einen automatisierten Prozess, der z.B. die

Abs. 85

Wikipedia-Liste auf neue Einträge prüft und diese ggf. zu Wikidata überträgt. Die hierfür erforderlichen Import-Skripte wurden online zur Verfügung gestellt^[47] und können frei weiterverwendet werden.

Da die Daten auch auf Wikidata jederzeit geändert werden können, stellt der einmalige Import der Daten nicht sicher, dass auch zukünftig die Daten in derselben Form vorliegen. So kann es beispielsweise passieren, dass Aussagen gelöscht oder geändert werden. Um dennoch zu gewährleisten, dass die vollständigen Daten auch in Zukunft auf Wikidata zu finden sind, ist es sinnvoll, zyklisch erneute Abgleiche durchzuführen und bei verloren gegangenen Informationen diese noch einmal nachzutragen. Dadurch würde auch auf lange Sicht eine Nachhaltigkeit der Daten geschaffen werden.

Abs. 86

Abs. 87

V. Zusammenfassung und Ausblick

Der vorliegende Beitrag zeigt neue Wege der Digitalisierung und Archivierung rechtsgeschichtlichen Wissens auf und erläutert sie an konkreten Beispielen. Ausgehend von einem Editionsprojekt zu den Geschäftsverteilungsplänen der Bundesgerichte (www.Richter-im-Internet.de) stellt der Beitrag drei verschiedene Digitalisierungsstrategien und ihr jeweiliges Verwendungspotential dar: Von der Digitalisierung ins pdf-Format, die einen einfachen Abruf im Internet und die menschliche Lektüre am Bildschirm erlaubt, über die Konvertierung in Tabellendokumente, die zusätzlich zur Bildschirmlektüre auch den automatisierten Import in Statistiksoftware ermöglicht, bis hin zur Modellierung der Daten als Linked Open Data und ihrem Import in ein offenes Datenrepositorium, das Menschen und Maschinen gleichermaßen zugänglich ist und jede digitale Weiterverarbeitung zulässt. Damit können rechtsgeschichtliche Editionsprojekte an die Informatisierung anderer Geisteswissenschaften (digital humanities) anschließen und den juristischen Nachbardisziplinen zugleich offene Forschungsdaten (open data) zur Verfügung stellen. Die in verschiedenen Disziplinen betriebene Erforschung der deutschen Justiz wird von solcher Methodeninnovation nicht nur hinsichtlich der Personalien des Bundesgerichtshofes dauerhaft profitieren.

Abs. 88

Fußnoten

* Dr. Dr. Hanjo Hamann ist wissenschaftlicher Mitarbeiter am Max-Planck-Institut zur Erforschung von Gemeinschaftsgütern und war Fellow Freies Wissen 2016/17 des Stifterverbands und von Wikimedia Deutschland. Marisa Nest ist Mitarbeiterin der Forschungsgruppe Human-Centered Computing (HCC) unter der Leitung von Claudia Müller-Birn am Institut für Informatik der Freien Universität Berlin.

[1] Neben etlichen Einzelbiographien bspw. Irons, *A People's History of the Supreme Court: The Men and Women Whose Cases and Decisions Have Shaped Our Constitution*, 1999; Toobin, *The Nine. Inside the Secret World of the Supreme Court*, 2007; Dalin, *Jewish Justices of the Supreme Court: From Brandeis to Kagan*, 2017; schon Mitte der 1990er beschrieb Posner diese Literatur als "newish, growing, increasingly popular", *NYU L. Rev.* 70 (1995) 502.

[2] Ausf. Ipsen, *Richterrecht und Verfassung*, 1975; Müller, 'Richterrecht'. *Elemente einer Verfassungstheorie*, 1986; Langenbucher, *Die Entwicklung und Auslegung von Richterrecht*,

1996; Bumke (Hrsg.), Richterrecht zwischen Gesetzesrecht und Rechtsgestaltung, 2012; Effer-Uhe u.a. (Hrsg.), Richterliche Rechtsfortbildung und kodifiziertes Richterrecht. 25. GJZ-Tagung 2014, 2016; Rütters, Die heimliche Revolution vom Rechtsstaat zum Richterstaat, 2016; Payandeh, Judikative Rechtserzeugung: Theorie, Dogmatik und Methodik der Wirkungen von Präjudizien, 2017; schon vor zehn Jahren beendete Haferkamp seinen Bericht über die Geschichte des Bürgerlichen Gesetzbuchs mit den Worten „Das deutsche Zivilrecht der Gegenwart wird daher zunehmend vom Richterrecht geprägt.“ (in: Basedow/Hopt/Zimmermann (Hrsg.), HWB-EuP Bd. 1, Tübingen 2009, hwb-eup2009.mpipriv.de/index.php /B%C3%BCrgerliches_Gesetzbuch).

[3] Die Personalbesetzung im Rahmen der Geschäftsverteilung wird erst seit 2013 systematisch online dokumentiert (s.u. bei Fn. 14), in der wohl umfangreichsten Internetdatenbank Wikidata verfügt im September 2017 weniger als 60 % der BGH-Richter (352/591) über einen Eintrag.

[4] Hamann, Richter im Internet. Editionsbericht zur Digitalisierung der Geschäftsverteilungspläne der deutschen Bundesgerichte seit dem Zweiten Weltkrieg, fhi 21 (2017) 8, www.forhistiur.de/2017-08-hamann.

[5] Damit schließt der Beitrag an den in der vorigen Fn. genannten an und trägt für den Bundesgerichtshof die dort noch als „Ausblick“ gekennzeichneten weiteren Arbeitsschritte nach.

[6] Aus der reichen Literatur z.B. Fuchs, Die Weiterverwendung der gemeinfreien Rechtsdatenbank „juris“, 3.4.2011, delegibus.com/2011,2.pdf; Hamann/Heinicke, Zwei Drittel der Wirklichkeit, SSRN 30.8.2012, dx.doi.org/10.2139/ssrn.2391465; Bruss, Die Verträge zwischen der juris GmbH und der Bundesrepublik Deutschland – Angriff auf die Gemeinfreiheit?, HFR 18 (2013) 16, www.humboldt-forum-recht.de/media/Druckansicht/pdf/2013-03.pdf; Podolski, Juris: Lexxpress gegen BVerfG: Sieg über mächtige Gegner, LTO 1.9.2015, www.lto.de/persistent/a_id/16771; international etwa Hürlimann, Publikation von Urteilen durch Gerichte, sui generis 1 (2014) 82, dx.doi.org/10.21257/sg.8; Mitee, The Right of Public Access to Legal Information: A Proposal for its Universal Recognition as a Human Right, GLJ 18 (2017) 1429, www.germanlawjournal.com/volume-18-no-06.

[7] Vgl. Hönnige/Gschwend, Das Bundesverfassungsgericht im politischen System der BRD – ein unbekanntes Wesen?, PVS 51 (2010), 507, insb. 513, nbn-resolving.de/urn:nbn:de:0168-ssoar-257817.

[8] Gemeinsam mit Wikimedia Deutschland e.V.; dazu schon Hamann, Gerichtsurteile als Menschenwerk. Zum Editionsprojekt „Die Namen der Justiz“, WMde-Blog 23.2.2017, blog.wikimedia.de/2017/02/23/g.

[9] [www.land-der-ideen.de/ausgezeichnete-orte/preistraeger?year=2017&category\[\]=Wissenschaft](http://www.land-der-ideen.de/ausgezeichnete-orte/preistraeger?year=2017&category[]=Wissenschaft); dazu Behrens, Offene Wissenschaft – Wie kann sich Freies Wissen weiter entwickeln? (Interview), WMde-Blog 28.6.2017, blog.wikimedia.de/2017/06/28/o.

[10] Krüger-Nieland (Hrsg.), 25 Jahre Bundesgerichtshof, München 1975; Geiß/Nehm/Brandner (Hrsg.), 50 Jahre Bundesgerichtshof, Köln 2000.

[11] Details in Hamann (Fn. 4) Fn. 25: www.forhistiur.de/2017-08-hamann#notes_n25.

[12] Vgl. Hamann (Fn. 4) Rn. 31: „Alle Senatsbesetzungen datieren mithin auf die erste Januarhälfte (1.–12.1.), mit Ausnahme der Jahre 1958 (19.2.) und 1968 (1.3.).“

[13] Unter www.bundesgerichtshof.de > Das Gericht > Stellung im Gerichtssystem > Rechtliche Grundlagen (www.t1p.de/2clz).

[14] www.bundesanzeiger.de (Suchbegriff „Geschäftsverteilungsplan“, Suchbereich „Amtlicher Teil“). Soweit Bundesgerichte ihre Geschäftsverteilung auch auf der eigenen Website veröffentlichen, wird sie mit Beginn des Folgejahres wieder gelöscht; kein Bundesgericht unterhält ein Online-Archiv, daher ist unklar, wann Geschäftsverteilungspläne erstmals im

Internet abrufbar waren.

[15] Einschließlich der Deutschen Nationalbibliothek (S. S., DNB Benutzung und Bestandsverwaltung, E-Mail vom 5.10.2016).

[16] Für die qualitativ hochwertige Digitalisierung danke ich Christoph Kling und der Universitätsbibliothek Mannheim.

[17] Unter www.richter-im-internet.de/bgh/1951 bzw. www.richter-im-internet.de/bgh/1952.

[18] Möhring, NJW 1950, S. 885–887.

[19] Nach Durchsicht jedes einzelnen NJW-Hefts im Suchzeitraum ergaben sich acht Fundstellen: NJW 1950, 777; 1950, 860; 1950, 901; 1950, 941; 1951, 105; 1951, 145; 1951, 348; 1951, 433.

[20] R. S., BGH Informationsdienste, E-Mail vom 9.1.2017.

[21] 45 Seiten FS50 zzgl. 1.040 Seiten Geschäftsverteilungspläne.

[22] Bspw. DOI 10.17176/20170224-104127 für den BGH-Geschäftsverteilungsplan 2017.

[23] Bei Richter „Schmidt“, der 1952 im V. Zivilsenat auftauchte, ließ sich nicht abschließend klären, ob es sich um Guido Schmidt handelte, der seit 1950 im I. Zivilsenat tätig war, oder um Adolf Schmidt, der 1952 an den 5. Strafsenat berufen wurde.

[24] Emil Lersch, Georg Rietschel und Fritz von Werner waren schon 1950 an amtlich veröffentlichten BGH-Entscheidungen beteiligt, hätten also eigentlich in den NJW-Mitteilungen auftauchen sollen. Von den übrigen 18 Richtern waren fünf seit 1951 in der amtlichen Sammlung nachweisbar, 13 erst ab 1952, als BGH-Besetzungen schon nicht mehr in der NJW mitgeteilt wurden.

[25] Willi Geiger, erstmals im Geschäftsverteilungsplan 1954 als Präsident des III. Zivilsenats belegt.

[26] Bundesrichter „Wolfhart“, zuvor Richter beim OGH Köln (so NJW 1950, 777). Wolfhart war lediglich der Vorname des ebenfalls 1950 berufenen OGH-Richters „W. Werner“ (NJW 1950, 901), der in jenem Jahr im 1., 2., 3., und 4. Strafsenat tätig war.

[27] So war als Austrittsdatum von Rudolf Schmitt der nicht existierende 29.2.1974 angegeben, als Geburtsort von Siegfried Räfte „Spitzkummersdorf/Oberlausitz“ – richtig sind wohl 28.2.1974 und Spitzkunnersdorf. Gerhard Schäfer wurde zum 20.4.2000 sogar für „verstorben“ erklärt, dabei war er noch bis 2002 im Amt und lebt bis heute; das falsch zugeordnete Todesdatum betraf den eine Zeile darunter stehenden Bernd-Arthur Paulusch.

[28] Im Geschäftsverteilungsplan 1988 waren die Senatsbesetzungen mit dem Vorjahresstand („Stand: 1. Januar 1987“) datiert; in den Jahren 2003/04 war der Name des Richters im IX. Zivilsenat und späteren Bundestagsabgeordneten Wolfgang NeškoviÄ als „Neškovi“ abgedruckt, während 1959 dem Richter im V. Zivilsenat „Piepenbrock“) ein überflüssiges Klammerzeichen beigegeben war. Weitere Fehler fielen aufgrund widersprüchlicher Genusbegriffe auf, vgl. Hamann (Fn. 4) Fn. 18.

[29] Die Geschäftsverteilungspläne wiesen im VIII. Zivilsenat 1967–1976 einen Richter (bzw. 1978–1989 Vorsitzenden Richter) namens „Braxmaier“ aus, während die FS50 nur Wolfram „Braxmeier“ verzeichnet; ebenfalls im VIII. Zivilsenat führten die veröffentlichten Geschäftsverteilungspläne 1994/95 den Richter Dr. (Wolfgang) Brunotte, der laut FS50 bereits zum 31.3.1993 in den Ruhestand getreten war; der in den Geschäftsverteilungsplänen von 1981 bis 1995 enthaltene Richter Teplitzky wurde in der FS50 Prof. Dr. Otto „Teplitzki“ geschrieben.

[30] Dazu mit Nachw. Hamann (Fn. 4) Rn. 41–43.

[31] Bsp.: Senatspräsident bzw. Bundesrichter (bis 1972), Vorsitzende Richterin bzw. Richter am Bundesgerichtshof (seit 1972); bei den bis 1960 noch nachgewiesenen „Juristischen Hilfsarbeitern“ auch andere Amtsbezeichnungen.

[32] Als Anmerkungen wurden (dem Namen in Klammern nachgestellt) fast ausschließlich Sonderzuständigkeiten, parallele Zugehörigkeiten zu anderen Senaten, die Position als stellvertretender Senatsvorsitzender und etwaige Übergangsbestimmungen vermerkt.

[33] LOD weist einzelnen Datensätzen eigene Adressen (URIs) zu, die einen Abruf im HTTP-Protokoll ermöglichen. Wird eine URI abfragt, so sollen die bereitgestellten Informationen auch Links zu anderen URIs beinhalten; für die Kodierung und Verlinkung kommt der offene W3C-Standard RDF zum Einsatz.

[34] Näher zum Folgenden Färber u.a., Linked Data Quality of DBpedia, Freebase, OpenCyc, Wikidata, and YAGO, SWJ 2016, Nr. 1465, unter www.semantic-web-journal.net/system/files/swj1465.pdf.

[35] WordNet ist ein Datenrepositorium, das Informationen zu Wörtern der englischen Sprache beinhaltet. Diese Informationen stellen semantische und lexikalische Beziehungen zwischen den Wörtern dar.

[36] GeoNames ist ein Datenrepositorium, das geographische Daten wie z.B. Ortsnamen in unterschiedlichen Sprachen, Längen- und Breitengrade, sowie Geokoordinaten enthält.

[37] Dazu www.wikidata.org/wiki/Wikidata:Data_Import_Hub.

[38] Näher www.wikidata.org/wiki/Wikidata:WikiProject_Freebase.

[39] Vgl. jeweils www.wikidata.org/wiki/Wikidata:Wikidata_for_authority_control.

[40] Vgl. z.B. www.reasonator.info oder tools.wmflabs.org/sqid.

[41] Beschrieben unter www.mediawiki.org/wiki/Wikibase/DataModel.

[42] <https://www.wikidata.org/wiki/Q26933282>

[43] Die zweite Eigenschaft in Abb. 1 dagegen („Geburtsdatum“) enthält ein sog. Literal (Zeichenkette) in Gestalt einer Datumsangabe.

[44] Unter de.wikipedia.org/wiki/Liste_der_Richter_am_Bundesgerichtshof.

[45] Dazu www.mediawiki.org/wiki/Manual:Pywikibot.

[46] Zusammenfassung der Zugriffswege unter www.wikidata.org/wiki/Wikidata:Data_access.

[47] Unter https://github.com/FUB-HCC/wikidata_bot