SSOAR

Open Access Repository
www.ssoar.info

Register zum Repertorium Germanicum: EDV-gestützte Erschließung spätmittelalterlicher Geschichtsquellen

Borchardt, Karl; Ott, Hannelore

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:
Borchardt, K., & Ott, H. (1993). Register zum Repertorium Germanicum: EDV-gestützte Erschließung spätmittelalterlicher Geschichtsquellen. *Historical Social Research*, *18*(3), 120-125. https://doi.org/10.12759/hsr.18.1993.3.120-125

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
https://creativecommons.org/licenses/by/4.0/deed.de

Terms of use:
This document is made available under a CC BY Licence (Attribution). For more Information see:
https://creativecommons.org/licenses/by/4.0

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-51249

Historical Social Research, Vol. 18 — 1993 — No. 3, 120-125

Register zum Repertorium Germanicum: EDV-gestützte Erschließung spätmittelalterlicher Geschichtsquellen

*Karl Borchardt (Würzburg); Hannelore Ott (Tübingen)**

Das Repertorium Germanicum ist ein Verzeichnis der in den päpstlichen Registern und Kameralakten vorkommenden Personen, Kirchen und Orte des Deutschen Reiches, seiner Diözesen und Territorien vom Beginn des Schismas bis zur Reformation, d.h. von 1378 bis 1523. Nach der Öffnung des Vatikanischen Archivs für die wissenschaftliche Forschung 1880/81 begann das kgl. Preußische Historische Institut in Rom 1892 dieses Projekt, das heute das Deutsche Historische Institut in Zusammenarbeit mit der niedersächsischen Archivverwaltung fortführt. Das Repertorium Germanicum benutzt die über 5000 mitunter dickleibigen Register- und Aktenbände des Papstes, der Kardinäle, der Kanzlei, Kammer und Rota, welche im fraglichen Zeitraum zusammen mehr als zwei Millionen Einträge enthalten. Diese Register und Akten übertreffen die in Hunderten von Archiven verstreute Empfängerüberlieferung um ein Mehrfaches und sind zudem im Vatikanischen Archiv, teilweise auch im Staatsarchiv Rom und der Vatikanischen Bibliothek konzentriert. Vom Repertorium Germanicum liegen bisher die Verzeichnisse zu den Akten der Päpste von 1378 bis 1431 und von 1447 bis 1458 vor. In Vorbereitung sind die Verzeichnisse zu Eugen IV. (1431-47), Pius II. (1458-64) und Paul II. (1464-71). Die Materialaufnahme ist bis in die ersten Jahre von Sixtus IV. (1471-84) gediehen. Insgesamt wurden bis jetzt für zwei Drittel des Zeitraums des Gesamtprojekts schätzungsweise eine knappe Viertelmillion Einträge verzettelt.

Der Nutzen des Repertorium Germanicum liegt nicht so sehr in der politischen Korrespondenz der Päpste, die durch Regesten nur unzureichend erschlossen wird und eigentlich einen Vollabdruck erfordert, sondern in der Unzahl prosopographischer Daten, vom Pfründenbesitz über Indulgenzen und Tragaltäre bis zu allerlei Dispensen (uneheliche Geburt, zu niedriges Alter, Nichterfüllung der Residenzpflicht wegen Studium oder Fürstendienst usw.). Vor allem bietet sich eine sozialgeschichtliche Auswertung an, was Herkunft und Karriere der Geistlichkeit in Deutschland sowie der Deutschen im päpstlichen Dienst, aber auch adelige und bürgerliche Bittsteller an der römischen Kurie angeht. Ebenso werden Finanztransaktionen zwischen Deutschland und Rom erhellt. Die einzelnen Einträge sind nach Kirchenrecht und Kanzleibrauch

---

* Protokoll des 55. Kolloquiums über die Anwendung der EDV in den Geisteswissenschaften 4. Juli 1992

Der Kolloquiumsbeitrag von Thomas Kühne und Andreas Gawatz ist als eigenständiger Aufsatz in der HSR veröffentlicht worden, so daß die Wiedergabe des Protokolls an dieser Stelle unterbleibt; vgl. Andreas Gawatz: Die Tübinger Dokumentation der preußischen Landtagswahlen im Kaiserreich (1867-1918), in: HSR, Vol. 17, 63 (1992).

hochgradig formalisiert, so daß durch die Gleichförmigkeit fast der Eindruck serieller Geschichtsquellen entsteht.

Drucklegung und Erschließung des Repertorium Germanicum wurden wesentlich durch den Einsatz von TUSTEP beschleunigt, zu dem sich das Deutsche Historische Institut in Rom auf Initiative des unvergessenen Dr. Hermann Diener (t 1988) zusammen mit dem Max Niemeyer Verlag, dem Zentrum für Datenverarbeitung der Universität Tübingen und der pagina GmbH in Tübingen entschlossen haben. Bei Martin V. (1417-31) lag der Text 1958 vor, das Personenregister erschien 1979, das Ortsregister steht noch aus. Bei Nikolaus V. (1447-55) erschien der Text 1985; mit Hilfe von TUSTEP konnte der Indexband 1989 folgen. Seither werden Text und Indices gemeinsam erstellt, was mehrere Vorteile bringt. Indem man frühzeitig die Zettel in den Computer eingibt, kann das Manuskript schneller zusammengestellt werden. Oft sieht man erst mit Hilfe der Indices, daß zwei Einträge sich auf dieselbe Person oder denselben Ort beziehen, so daß etwa für Pius II. der Umfang von 6455 auf schließlich 5984 Personen und Orte schrumpfte. Unvermeidliche Tippfehler fallen in den Indices leichter auf. Außerdem enthalten nun die Indices, wie schon lange gefordert, nicht bloß die Personen- und Ortsnamen, sondern auch Patrozinien, Orden und religiöse Gemeinschaften, Kalenderdaten, Fundstellen und Sachbetreffe. Das macht die Bände vielseitiger benutzbar.

Bisher wurden in die Indices zu übernehmende Textpassagen jeweils durch acht besondere Steuerzeichen eingeschlossen: $ Personen, + Orte, § Patrozinien, % Orden und religiöse Gemeinschaften, * akademische Grade, " Standesbezeichnungen, & Kalenderdaten der Einträge, &/ in dem Eintrag genannte Vordaten. TUSTEP und die sachlich knappe Regestenformulierung erlauben jedoch, solche Kodierungen künftig auf ein Minimum zu beschränken. Namen beispielsweise sind durch Großbuchstaben, Orte durch nachfolgendes *dioc,* Patrozinien durch voraufgehendes *s., ss., b., bb.* eindeutig gekennzeichnet. So werden Kodierungsfehler vermieden, und außerdem gewinnt der Bearbeiter Zeit für wichtigere, nicht durch EDV zu leistende Arbeitsschritte wie die Gewinnung des Zunamensindex aus den Vornamen. *Wilhelmus, Johannes et Bertoldus fr. et com. in Hennberg* muß nämlich bei den Vornamen dreimal unter *W, J* und *B* stehen sowie bei den Zunamen einmal unter *H.*

Als besonders schwierig erwies sich die Vereinheitlichung der Flexionsformen schon bei gewöhnlichen Wörtern. Ein *vidimus* (beglaubigte Urkundenabschrift) darf nicht unter *videre* (sehen) verschwinden. Der Prior eines Klosters muß von *prior* (früher) unterschieden werden. *Mense* kann der Ablativ von *mensis* (der Monat) oder der Genitiv usw. von *mensa* (der Tisch) sein. Mitunter hilft ein KWIC-Index (key word in context), um etwa *m.* mit den Bedeutungen *mandare, mandatum, marca* und *murus* zu differenzieren. Noch komplizierter wird die Vereinheitlichung bei Namen, denn hier geht es nicht nur um Flexion - *Theoderico* zu *Theodericus,* während *Otto* trotz Endung auf *-o* bleibt -, sondern auch um Varianten. Während *Hinricus, Henricus* und *Arigo* eindeutig unter

Heinricus stehen müssen, wäre z. B. bei *Rutgerus* und *Ratgerus* eine philologische Abklärung notwendig.

Bei so viel Handarbeit, die trotz des TUSTEP-Einsatzes notwendig bleibt, erfüllen die Indices natürlich nicht alle Wünsche. Der Text wird aber auf Band gespeichert, so daß Interessierte für spezielle Anliegen darauf zurückgreifen können. Beispielsweise kann man sich Pfründen mit mehr als 20 Mark Silber Jahreseinkünften heraussuchen lassen oder Personen, bei denen angegeben wird, daß sie jünger als 14 Jahre sind, oder Uneheliche mit einem Priester als Vater und einer unverheirateten Frau als Mutter: *def. nat. (p., s.).* Wünschenswert wäre, daß die bereits gedruckten Bände mit einem Scanner eingelesen werden, damit man ihre Angaben in ähnlicher Weise auswerten kann.

Da der Text lateinisch und abgekürzt ist, wodurch viele Flexionsformen wegfallen, kommt das Repertorium Germanicum TUSTEP entgegen. Eben deshalb lassen sich andererseits hier gemachte Erfahrungen nicht ohne weiteres übertragen. Dennoch treten einige Probleme bei allen spätmittelalterlichen und frühneuzeitlichen Urkunden, Kopial- und Registerbüchern, Lehenbüchern, Bürgerlisten und Rechnungen auf. Wie oft beklagt wird, kapitulieren Bearbeiter von Urkundenbüchern und Regesten in Deutschland häufig vor der ab etwa 1300 einsetzenden Materialfülle; TUSTEP könnte einen Ausweg aus dieser mißlichen Lage eröffnen.

Technische Hintergründe der Arbeit an den Registern zum Repertorium Germanicum

Statt die Programme vorzustellen, die zu den Repertoriumregistern führen, soll hier die Arbeitsstrategie beschrieben werden, die die Erarbeitung solcher Register mit Herausgebern ermöglicht, die nicht am Ort sind und z. T. nicht TUSTEP benutzen.

Mit dieser Strategie mußten zwei Ziele erreicht werden:

a) Die Abfolge der Arbeitsschritte war so zu wählen, daß unvermeidliche Handarbeit nicht ein zweites Mal geleistet werden muß.

b) Die beiden Kooperationspartner sollten die Sicherheit erhalten, daß das Ergebnis der eigenen Arbeit nicht durch den anderen Partner gefährdet ist.

Bei der Erfassung des Repertoriums zum Pontifikat Pauls VI. anfangs der 80er Jahre wurden zur Vorbereitung der Registerarbeit bereits Registerauszeichnungen vorgesehen. Da *Apel Fochs de Fochsberg* im Register an drei Stellen auftauchen sollte: unter *Fochs de Fochsberg, Apel,* unter *Fochsberg, Apel Fochs de,* und schießlich unter *Apel Fochs de Fochsberg,* wurden sogenannte »Drehpunkte« miterfaßt, die jeweils die Anfänge von Registereinträgen bezeichnen sollten, die durch zyklische Permutation aus den vorgefundenen Einträgen erzeugt werden sollten.

Trotz dieser und weiterer Vorkehrungen ergaben sich bei fast allen Arbeitsschritten immer wieder neue Anforderungen, die erst bei der Registerarbeit

selbst auftauchten, wie: Erzeugung von zusätzlichen Registereinträgen; Überführung von flektierten Formen in die Grundform; Sortierung der vorgefundenen graphischen Formen nach davon verschiedenen Ansetzungsformen.

Außerdem wurde klar, daß der für die »technische« Seite Verantwortliche äußerst vorsichtig mit den Daten umgehen mußte, da er nur einen kleinen Teil der sachlichen Anforderungen überblickte und daher den Herausgebern in jeder Phase die Möglichkeit geben mußte, auch seine Verständnisfehler aufzuspüren. So mußten speziell Entscheidungen, die auf Grund falscher oder unvollständiger Regeln getroffen wurden, wieder rückgängig zu machen sein.

Im Laufe der Bearbeitung der Register stellte sich außerdem heraus, daß die bei der Erfassung manuell vorgenommene Kennzeichnung der Registereinträge häufig unvollständig oder falsch war. Als Folge davon haben wir es als notwendig erachtet, zunächst - also zu Beginn der Registerarbeit - die Namen der genannten Personen und Orte an Hand einer unsortierten Liste auf Vollständigkeit zu prüfen und gegebenenfalls zu ergänzen.

Aus der Erfahrung, daß trotz manueller Vorarbeit alle Einträge kontrolliert werden mußten, und aus der Tatsache, daß der Repertorien-Text selbst stark formalisiert ist, ergab sich bald die Überlegung, für die weiteren Bände die Kodierungsarbeit zu sparen und zu versuchen, die Namen der Personen und Orte automatisch aus dem Text zu gewinnen. Dieses Verfahren wurde inzwischen bei der Registerarbeit zu weiteren Bänden des Repertorium erprobt.

Registerarbeit ohne Kodierung im Text

Der erste Schritt besteht jetzt darin, daß zu jedem Abschnitt alle automatisch erkennbaren Einträge erzeugt werden. Diese werden zunächst in der Reihenfolge belassen, in der sie im Text vorkommen. Als weitere Korrekturhilfe wird jedoch eine rein alphabetisch sortierte Liste der vorkommenden Namen mit ausgegeben, in der z. B. Verschreibungen leichter auffallen als im fortlaufenden Text oder in der endgültig vorgesehenen Sortierung.

Die erste Korrektur der Register-Einträge wird dann in der Datei durchgeführt, in der alle Orts- und Personennamen noch in der Reihenfolge stehen, in der sie im Text vorkommen. Vor jedem Eintrag steht eine Kodierung, die das zugehörige Register angibt, z. B: o: (Ort), n: (Name), ?: (Registerzuordnung ungewiß). Der Bearbeiter kontrolliert und korrigiert die Einträge in dieser Datei (bei Verschreibungen zusätzlich in der Text-Datei). Soll ein automatisch erzeugter Eintrag ganz gelöscht werden, so wird der Eintrag nicht aus der Datei entfernt, sondern durch »-:« als Registerkennzeichen markiert.

Nach der ersten Korrektur werden die Registereinträge per Programm auf die Dateien für die verschiedenen Register verteilt.

Für die Personennamen werden in einem zweiten Arbeitsschritt per Programm die flektierten Formen der Vornamen in den Nominativ überführt; gleichzeitig wird für jeden Bestandteil des Namens, der mit einem Groß-

buchstaben beginnt, durch die bereits erwähnte zyklische Permutation ein eigener Registereintrag erzeugt. Bei Personennamen, die eine Ortsbezeichnung enthalten, müssen auch entsprechende Einträge für das Ortsregister erzeugt werden. Diese werden, sofern automatisch erkennbar, mit einer entsprechenden Markierung versehen.

Nun werden die Registereinträge nach Regeln sortiert, die ein Zusammenführen gleicher, aber verschieden geschriebener Namen ermöglichen, wobei also z. B. th wie t behandelt, Doppelbuchstaben wie ein Einzelbuchstabe einsortiert werden. In das dabei entstehende Register wird außerdem eine Liste schon bekannter Zuordnungen von vorgefundenen Namensformen zu zugehörigen Ansetzungsformen in Form von Verweisen einsortiert, z. B.: *Guillermus siehe Wilhelmus.*

Es folgt ein weiterer Korrekturschritt, bei dem neben den übrigen Korrekturen die noch fehlenden Verweise auf Ansetzungsformen ergänzt und überflüssige als solche markiert werden.

Die in diesem so durchkorrigierten Register enthaltenen Verweise werden dazu benutzt, um per Programm jedem vorkommenden Namen seine Ansetzungsform beizugeben, nach der das Material anschließend sortiert wird. Die Verweise von der Textform auf die Ansetzungsform bleiben selbstverständlich erhalten.

Strategien für kontrollierte Register-Arbeit

Um ein kontrolliertes Arbeiten für die einzelnen Arbeitsschritte zu ermöglichen, wurde auf folgende Punkte besonderer Wert gelegt:

1. Für alle Bearbeitungsschritte, die manuelle Eingriffe durch den Herausgeber erfordern, gilt als eiserne Regel, daß nur Korrekturen am Wortlaut direkt in der Datei vorgenommen werden, während für erforderliche Löschungen und für Verschiebungen oder zusätzliche Aufnahme in andere Register nur entsprechende Markierungen in der Datei angebracht werden.

2. Für jeden Bearbeitungsschritt sind automatische Kontrollen vorgesehen. Dies gilt nicht nur für die von Hand vom Bearbeiter durchgeführten Korrekturen, sondern auch für alle Arbeitsschritte, in denen auf Grund von Regeln Veränderungen an den Registereinträgen vorgenommen werden. Die Kontrollen selbst bestehen darin, daß jeweils die (unkorrigierte) ursprüngliche Fassung mit der korrigierten bzw. durch veränderte Regeln erzeugten Fassung automatisch verglichen wird. Während für die Protokollierung der manuellen Korrekturen ein einfaches Vergleichsprotokoll ausreicht, wird zur Kontrolle der über Regeln durchgeführten Änderungen (z. B. der Reduzierung der flektierten Formen auf die Grundform) eine alphabetisch sortierte Liste ausgegeben, die jeweils die ursprüngliche und die über Programm erzeugte Form eines Namens enthält.

3. Bei der Veränderung der Registereinträge selbst müssen manuelle Eingriffe von automatischen Veränderungen unterscheidbar bleiben. Beispiel: Für

die Vorbereitung der Sortierung werden nach *ux.* und etlichen anderen Abkürzungen vor dem nächsten großgeschriebenen Wort automatisch Nichtsortierzeichen eingefügt, weil in der Regel ein bei der Sortierung zu übergehender Vorname folgt. Wo diese Regel nicht greift, wird das Nichtsortierzeichen von Hand nachgetragen. Hierfür wird ein anderes Zeichen gewählt als das vom Programm für den gleichen Zweck vergebene, so daß man auch bei der Kontrolle schnell eine falsche Regel von einer falschen Einzelkorrektur unterscheiden kann.

4. In den Registern müssen neben Korrekturen auch Ergänzungen vorgenommen werden können. Um insbesondere automatische Ergänzungen durchführen zu können, muß man das fertige (vorläufige) Register wieder in einzelne Einträge zurückverwandeln und dann neu sortieren können.

Die hier eingeschlagene Strategie für die Zusammenarbeit zwischen einer mehr technisch orientierten Stelle einerseits und den Fachwissenschaftlern andererseits, die sich auf die sachlichen Fragen beschränken wollen und sollen, hat sich bei den Registern von Orts- und Personen-Namen bereits bestens bewährt. Prinzipiell die gleichen Regeln gelten natürlich auch für die Sachregister, auch wenn der Ausgangspunkt und die einzelnen Schritte der Arbeit sich in vielen Einzelheiten unterscheiden.

Das Findebuch zum mittelhochdeutschen Wortschatz. Probleme der Kompilation von Ausgabenglossaren zu einem Gesamtglossar.

*Kurt Gärtner (Trier); Gerhard Hanrieder (Erlangen)**

I.

Vor kurzem ist ein neuartiges lexikographisches Werk zum Mittelhochdeutschen, der von ca. 1050 bis ca. 1350 reichenden Periode der deutschen Sprache, erschienen, das sogen. >Trierer Findebuch<, das Ergebnis einer Gemeinschaftsarbeit der Trierer Altgermanisten:

> Kurt Gärtner, Christoph Gerhardt, Jürgen Jaehrling, Ralf Plate, Walter **Roll,** Erika Timm. Datenverarbeitung: Gerhard Hanrieder: *Findebuch zum mittelhochdeutschen Wortschatz. Mit einem rückläufigen Index.* Stuttgart: Hirzel 1992.

Mit dem Findebuch soll die desolate Situation auf dem Gebiet der mittelhoch-

---

* Protokoll des 56. Kolloquiums über die Anwendung der EDV in den Geisteswissenschaften 21. November 1992

deutschen Lexikographie für zwei bis drei Jahrzehnte, genauer: bis zum Erscheinen eines neuen Sprachstadienwörterbuchs zum Mittelhochdeutschen, provisorisch überbrückt werden. Zwar sind in den letzten Jahren einige Autoren- und Werkwörterbücher sowie Spezialwörterbücher zu bestimmten Textsorten erschienen oder in Angriff genommen worden. Nicht nur Germanisten, sondern auch Historiker, Theologen und andere Forscher, die sich mit mittelhochdeutschen Texten beschäftigen, sind dennoch immer noch angewiesen auf zwei große, einander ergänzende Wörterbücher, die inzwischen weit über 100 Jahre alt sind:

1. das Mittelhochdeutsche Wörterbuch< von Benecke/Müller/Zarncke [- BMZ] (1854-1866, letzter Nachdruck 1990), und
2. das Mittelhochdeutsche Handwörterbuch< von Matthias Lexer [= Lexer] (1872-1878, letzter Nachdruck 1992).

Von diesen beiden Wörterbüchern ist nur der BMZ ein selbständiges Sprachstadienwörterbuch; der Lexer dagegen ein Handwörterbuch auf der Basis des BMZ, zu dem er auch als Supplement fungiert und schließlich noch als alphabetischer Index, denn die Lemmata des BMZ sind nicht streng alphabetisch, sondern nach Wortstämmen und ihren Wortfamilien angeordnet.

Die beiden alten Wörterbücher basieren vorwiegend auf Quellen aus der Zeit von 1150 bis 1230, insbesondere auf den wenigen in den drei Jahrzehnten um 1200 entstandenen >klassischen< Texten >Iwein<, >Parzival< und >Tristan<, die etwa ein Viertel der Belege stellen. Für das Mittelhochdeutsche in den zeitlichen Grenzen von 1050-1350 waren diese Wörterbücher von Anfang an nicht repräsentativ. Heute sind sie es noch weit weniger, denn die Quellenerschließung ist nach dem Abschluß des Lexer 1878 weitergegangen.

Das Findebuch hat daher die Aufgabe, die beiden Wörterbücher zu ergänzen, indem es die Lemmata der Ausgabenglossare zu den nach 1878 erschienenen Editionen mittelhochdeutscher Texte aus dem Zeitraum von ca. 1050 bis 1350/1400 kompiliert und damit die Informationen liefert, die ein gezieltes Nachschlagen und Auffinden der gesuchten Lemmata in den Einzelglossaren ermöglicht. Rund 90 der über 100 für das Findebuch herangezogenen Einzelglossare stammen aus Editionen von Texten, die vor und nach den drei Jahrzehnten um 1200 entstanden sind. Der Lemmabestand des Findebuchs (rund 35.000) erlaubt damit erstmals einen annähernd repräsentativen Überblick über den gesamten mittelhochdeutschen Wortschatz.

Das Findebuch ist kein selbständiges Wörterbuch, sein eingeschränkter Wert für den Benutzer kommt deshalb unmißverständlich in seinem Titel zum Ausdruck. Die Bezeichnung >Findebuch< war in Analogie zu Roy Wisbeys >finding lists</>Findelisten< gewählt worden, mit denen er das Auffinden aller Verbformen in den von ihm herausgebrachten Indizes und Konkordanzen ermöglichen wollte. Das Findebuch hat einen ganz ähnlichen Zweck; es dient lediglich als Wegweiser zu dem in den Ausgabenglossaren gebuchten Wortschatz und enthält nur soviel Informationen, wie zum raschen und sicheren Auffinden der