

HistVV: Die Leipziger Vorlesungsdatenbank 1814/15–1914

Frank Fischer, Tobias Grave, Carsten Milling, Jörg Ossenkopp

► **To cite this version:**

Frank Fischer, Tobias Grave, Carsten Milling, Jörg Ossenkopp. HistVV: Die Leipziger Vorlesungsdatenbank 1814/15–1914. Ulf Morgenstern; Thomas Riechert. *Catalogus Professorum Lipsiensis – Konzeption, technische Umsetzung und Anwendungen für Professorenkataloge im Semantic Web*, 21, pp.131-135, In press, *Leipziger Beiträge zur Informatik*, 978-3-941608-08-5. hal-01876835

HAL Id: hal-01876835

<https://hal.archives-ouvertes.fr/hal-01876835>

Submitted on 18 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HistVV: Die Leipziger Vorlesungsdatenbank

1814/15-1914

Frank Fischer, Tobias Grave¹; Carsten Milling, Jörg Ossenkopp²

DFG-Projektgruppe „Vorlesungsverzeichnisse als Quellen disziplinär organisierter Wissenschaft. Die Ausdifferenzierung wissenschaftlicher Fächer an der Universität Leipzig 1814/15–1914“

¹Team Leipzig
{fischer|grave}@ub.uni-leipzig.de

² Team Berlin
cmil@hashtable.de, joerg.ossenkopp@gmail.com

Kurzfassung: Der Beitrag beschreibt die praktische Umsetzung eines Projekts, bei dem exemplarisch die Vorlesungsverzeichnisse einer der damals größten deutschen Universitäten für einen bestimmten Zeitraum digital erfasst und semantisch ausgezeichnet worden sind. Für die 200 Semester zwischen 1814/15 und 1914 stehen nun eine Volltextsuche sowie ein vollständiger Dozentenkatalog zur Verfügung. Fachhistoriografische Forschungen erhalten so neue Impulse. Die während des Projekts entstandene Software kann auch für andere Universitäten und Zeiträume Verwendung finden.

Im Februar 2009 wurde ein Portal freigeschaltet, mit dessen Hilfe die Vorlesungsverzeichnisse der Universität Leipzig vom Wintersemester 1814/15 bis zum Sommersemester 1914 nach inhaltlichen und strukturellen Kriterien erschlossen werden können. Mit Hilfe der Deutschen Forschungsgemeinschaft wurde es unter der Leitung von Ulrich Johannes Schneider an der Universitätsbibliothek Leipzig entwickelt und wird dort nun auch weiter betrieben. Erreichbar ist das Portal unter der Adresse <http://histvv.uni-leipzig.de/> – das verwendete Kürzel HistVV weist auf den Zugang zu „HISTorischen VorlesungsVerzeichnissen“ hin und ist auch als Hinweis auf eine mögliche künftige Verwendung der entwickelten Erfassungsstruktur an anderen Universitäten als Second-Level-Domain an den Anfang der Adressausdrucks gesetzt.

Kapitel 4: Weiterführende und thematisch verwandte Projekte

Innerhalb der Website des Projekts kann jedes der erfassten Vorlesungsverzeichnisse im digitalisierten Volltext gelesen und unmittelbar mit dem Original verglichen werden: Durch ein Anklicken der Seitenzahlen erscheinen die Scans der originalen Druckseiten der Verzeichnisse auf dem Bildschirm. So bleibt jedes Textstück immer mit seiner Vorlage vergleichbar.

Für die Recherche steht eine Volltextsuche zur Verfügung, mit der gezielt z. B. die Konjunktur von bestimmten Lehrthemen nachvollzogen werden kann. Die Suchfunktion ist so eingerichtet, dass sie sich auf einen bestimmten zu untersuchenden Zeitraum und auf einzelne Fakultäten einschränken lässt. Abgesehen von der chronologischen Wiedergabe der Lehrbiografien im ausgewählten Zeitsegment helfen die Abfragemöglichkeiten auch bei der Suche nach spezifischeren Antworten. Man ist beispielsweise schnell darüber im Bilde, zu welchem Zeitpunkt in Leipzig erstmalig die Darwin'schen Theorien gelehrt wurden. Die Eingabe des Begriffs „Darwin“ in die Suchmaske genügt. Darwins „Die Entstehung der Arten“ war 1859 erschienen, und bereits im Wintersemester 1864 kündigte Julius Victor Carus, einer der wichtigsten Vermittler der Darwin'schen Theorien in Deutschland, eine entsprechende Veranstaltung an. Danach gab es zehn Jahre lang keine speziell angekündigte Lehre zum Thema. Ab dem Sommersemester 1873 allerdings gehörte Darwin dann zum regelmäßigen Lehrprogramm.

Auf ähnliche Weise ist es möglich zu erfahren, wann sich Nietzsches Leipziger Alma Mater zum ersten Mal explizit mit dessen Philosophie beschäftigt hat. Nietzsche selbst beklagte noch 1888 in seiner Autohagiografie „Ecce homo“ die Wirkungslosigkeit seiner Schriften in Deutschland. Anhand der Vorlesungsverzeichnisse ist erschießbar, dass in Leipzig im Wintersemester 1902, also 14 Jahre nach der Niederschrift des damals noch nicht veröffentlichten „Ecce homo“ und zwei Jahre nach Nietzsches Tod, eine erste Veranstaltung zu seinem Werk angekündigt war, die von dem Philosophen Raoul Richter geleitet wurde.

Wer nun weiterhin wissen will, um wen es sich dabei handelt, kann sich von der angezeigten Veranstaltung zum Stammdatenblatt des Dozenten im Dozentenkatalog weiterleiten lassen, einer eigens für die Vorlesungsdatenbank zusammengestellten Kartei aller 1184 Lehrenden im erfassten Zeitraum. Dabei wurden im Unterschied zum weithin üblichen Vorgehen bei Professorenkatalogen *alle* im Rahmen der Ankündigungen aufgeführten Dozenten berücksichtigt – Hilfslehrer und Stallmeister eingeschlossen, insofern sie mit Lehrangeboten in den Verzeichnissen vertreten sind. Die Individualisierung der lehrenden Personen erfolgte systematisch durch Standardlexika. Unbekanntere Dozenten wurden im Stadtarchiv Leipzig in den Bürgerakten aus dem 19. Jahrhundert recherchiert. Die Stammdaten der Dozenten enthalten neben den Geburts- und Sterbedaten auch Angaben zur Profession und eine Lehrbiografie mit allen in Leipzig angekündigten Veranstaltungen. Außerdem wird zu anderen biografischen Projekten verlinkt, die eine mehr oder weniger große Schnittmenge mit dem Bestand unseres Dozentenkatalogs aufweisen – und zwar in folgender Verteilung (Stand: 12. 2. 2010):

- Personennamendatei (PND): 980
- Allgemeine Deutsche Biographie (ADB): 234
- Neue Deutsche Biographie (NDB): 292
- Biographisch-Bibliographisches Kirchenlexikon von Bautz (BBKL): 73
- deutschsprachige Wikipedia: 527 (plus darin noch nicht vorhandenen Einträgen in anderssprachigen Wikipedias: 12)
- Leipziger Professorenkatalog: 568
- Hallenser Professorenkatalog: 23
- Rostocker Professorenkatalog: 3
- „Biographisches Lexikon hervorragender Ärzte des neunzehnten Jahrhunderts“ von Julius Pagel (1901): 104
- „Sächsische Biografie“ des ISGV: 18
- Personen-Wiki der SLUB: 266

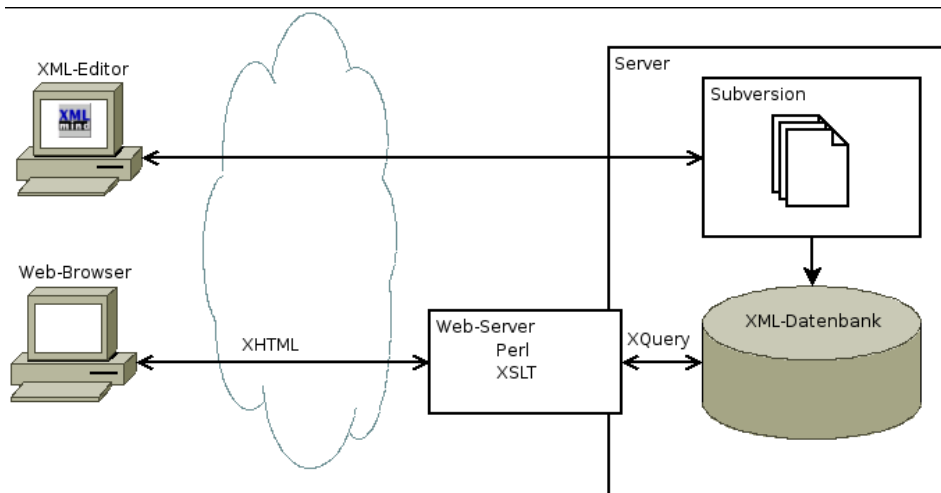
Neben diesen Standardquellen haben wir bei der Recherche der Stammdaten auf das World Biographical Information System (WBIS) des K. G. Saur Verlags zurückgegriffen. Gesondert angegeben haben wir Quellen nur dann, wenn sie auch über diese Schnittstelle nicht zugänglich waren. Das gilt insbesondere für die Bürgerakten und Meldebücher im Leipziger Stadtarchiv. Wenn sich Geburts- und Sterbedaten in den einzelnen Quellen widersprechen, wird das in den Anmerkungen angegeben.

Die Namensansetzung folgt in der Regel der PND (die sich normalerweise an der Namensform orientiert, die für Publikationen benutzt wurde). Es werden stets alle Vornamen angegeben, wobei der Rufname separat ausgezeichnet und nur dessen Anfangsbuchstabe für die systeminterne ID verwendet wird. „Baur, Gustav Adolf Ludwig“ hat also die ID „baur_g“ zugewiesen bekommen. Als Geburts- und Sterbeorte verzeichnet sind die im jeweiligen Jahr gebräuchlichen Ortsnamen, also etwa „Deutsch Eylau“ (heute: Iława), „Königsberg“ (heute: Kaliningrad), „Langenschwalbach“ (heute: Bad Schwalbach) usw. Bei Zweideutigkeiten wird der übliche Zusatz ergänzt (z. B. „Halle (Saale)“, „Freiburg im Breisgau“, „Frankfurt am Main“). Bei kleineren, unbekannteren und inzwischen häufig eingemeindeten Orten wird ggf. die nächste bekanntere Stadt angegeben (z. B. „Bernsdorf bei Zwickau“, „Brüheim bei Gotha“ usw.). Als Berufsbezeichnungen dienen im Idealfall nur die generischen, d. h. aus den Universitätsdisziplinen selbst abgeleiteten Tätigkeiten ohne jede weitere Differenzierung (z. B. „Mediziner“, „Physiker“, „Philologe“ usw.). Derartige Tätigkeitsbezeichnungen wurden als ausreichend erachtet, weil sie innerhalb des Dozentenkatalogs zur schnellen Identifizierung, nicht aber zur weiteren Qualifizierung einer Person beitragen sollen. Ausgegangen wurde bei den Bezeichnungen meist von der entsprechenden Angabe im zugehörigen PND-Datensatz. Erwies sich ein Fall als zweifelhaft, konnte die „Liste der weiten, nor-

Kapitel 4: Weiterführende und thematisch verwandte Projekte

mierten Berufs- und Tätigkeitsbezeichnungen“ der Arbeitsstelle für Standardisierung bei der DNB zu Rate gezogen werden.¹

Die Vorlesungsverzeichnisse wurden für die Eingabe nicht in besondere Einheiten zerlegt, sondern ganzheitlich erfasst, d. h. mit Hilfe der eXtensible Markup Language (XML) semantisch getaggt, bevor sie in eine XML-fähigen Datenbank überführt wurden. Ein wichtiges Kriterium dabei war, die Textgestalt bei der Eingabe nicht verändern und die semantisch bedeutsamen Teile nicht isolieren zu müssen, sondern sie im ursprünglichen Zusammenhang darstellen zu können. Das Schema der Auszeichnungssprache war also in der Lage, die im Verlauf des 19. Jahrhundert deutlich variierende Gestalt der Vorlesungsverzeichnisse ohne weitere Zurichtung zu verarbeiten. Von der Dateneingabe via XML-Editor bis zur Datenausgabe via Browserfenster war die Erstellung des Portals wie folgt organisiert:



Zur Sicherung der einzelnen Arbeitsschritte kam eine Versionsverwaltung zum Einsatz, die sowohl für den Programmcode als auch für die erfassten Daten alle Bearbeitungsstände reproduzierbar macht. Die Abbildung der in XML-Dokumente konvertierten Vorlesungsverzeichnisse auf eine XML-fähige Datenbank erfolgte über die Berkeley DB. Für die Dateneingabe selbst wurde der Editor XMLmind benutzt, der als Personal Edition kostenfrei ist. Die gesamte technische Umsetzung ist mit Open-Source-Software (Apache-Webserver, Berkeley DB XML, Trac) und unter der Berücksichtigung offener Standards (XML, RELAX NG, XQuery) bewältigt worden. Die von Google weiterentwickelte OCR-Software Tesseract ließ sich für die Digitalisierung eines Teils der Vorlesungsverzeichnisse verwenden, insbesondere für Verzeichnisse ab dem Sommersemester 1856, die nicht mehr durchgängig in Fraktur gesetzt waren.

¹ Vgl. <http://www.d-nb.de/standardisierung/pdf/berufeliste.pdf> (Stand: 1. September 2006).

Vorlesungsverzeichnisse sind Gebrauchstexte mit serieller Struktur, bei denen einzelne Passagen semantisch ausgezeichnet werden müssen, um sie maschinenlesbar abzubilden. Nicht alle Besonderheiten des Layouts sind übernommen worden. Gesperrt gedruckte Professorennamen beispielsweise wurden lediglich semantisch getaggt, die Sperrung wird daher im digitalisierten Volltext nicht ohne Weiteres reproduziert. Sperrungen und Antiqua-Passagen innerhalb von Vorlesungstiteln sind jedoch mit Rücksicht auf deren jeweilige Semantik übernommen worden, letztere als Kursivierungen. Durch den immer möglichen Einblick in die gescannten Originaldokumente geht der typografische Ursprungszustand aber ohnehin nicht verloren. Wenigstens bis 1901, dem Jahr der II. Orthographischen Konferenz, gab es keine genormte deutsche Rechtschreibung. In den Vorlesungsverzeichnissen des 19. Jahrhunderts wechseln deshalb verschiedene Schreibungen einander ab („Medicin“ steht neben „Medizin“, „Litteratur“ neben „Literatur“, „Göthe“ neben „Goethe“ usw.). Für die Volltextversionen des Portals wurden keinerlei Vereinheitlichungen vorgenommen. Offensichtliche Fehler im Drucksatz sind aber korrigiert und auch als solche kenntlich gemacht. Darüber hinaus gibt es „[sic!]“-Vermerke bei einigen wenigen Fällen, die ausgehend von den in den Verzeichnissen sonst gebräuchlichen Schreibungen anderes hätten erwarten lassen. Vokabular und Syntax der Verzeichnisse sind wie üblich denkbar verknüpft und auf das Wesentliche reduziert. Die Veranstaltungen werden deshalb mit akademischen und teils heute nicht mehr gebräuchlichen Abkürzungen und Begriffen beschrieben. Für deren Verständnis wurde eine Abkürzungsliste erstellt.

Insgesamt war das 19. Jahrhundert auch für die Leipziger Universität ein Jahrhundert der Umgestaltung. Aus der traditionellen, eigenständigen und bewahrenden Lehrinstitution ist die moderne, budgetierte und experimentierende Forschungsuniversität geworden. Der Gestalt der Vorlesungsverzeichnisse merkt man das deutlich an. Die traditionelle Fakultätszugehörigkeit der Professoren kommt als Ordnungsprinzip mit der systematischen Ordnung der Disziplinen kaum überein. Will man deren Modifikationen und den stetigen Zuwachs des Lehrangebotes erforschen, stellen die gedruckten Verzeichnisse einen beträchtlichen Fundus bereit. Fachhistoriografen haben deshalb seither gern in ihnen geblättert. Die elektronische Erfassung der Vorlesungsverzeichnisse im HistVV-Portal kann nun hoffentlich deren strukturelle Auswertung erleichtern.