

GUNTER VASOLD, Graz

Vom „Allerunderthenigsten Anruffen“ zum analytischen Abrufen

Ein Werkzeug zur kooperativen Erfassung, Verwaltung und Analyse von Untertanensuppliken

A Web-based Environment for Cooperative Research on Supplications filed by Subjects at the Imperial Aulic Council

When it comes to analyzing large numbers of historical records from the 16th and 17th centuries, the approach of using a database seems promising. Implementing such a database requires special measures considering the nature of the data and the way historians construct knowledge about their research topics. This paper introduces a virtual research environment that supports collaborative research on supplications filed by subjects at the Imperial Aulic Council under the reign of Emperor Rudolf II (1576–1612). Special emphasis is given to conceptual problems rather than technical details, which is useful for anyone planning a similar project. After a short description of functional principles, it is discussed how the hermeneutic nature of historical research needs to be reflected in the software development process and how specialized interfaces for typical tasks like revision, annotation, normalization or analyzing the data can be provided.

Keywords: collaborative research – virtual research environment – research data – long time preservation – supplications

Das Forschungsprojekt „Untertanensuppliken am Reichshofrat Kaiser Rudolfs II. (1576–1612)“ untersucht anhand von Suppliken Aspekte der politischen Kommunikation zwischen Untertanen und Obrigkeit. Dazu wurden mehr als 4.000 im Wiener Haus-, Hof- und Staatsarchiv lagernde Suppliken durchgesehen und 1.471 für das Projekt relevante Suppliken im Umfang von 38.520 Aktenseiten in einer Datenbank erfasst. Auf Grundlage dieser Datenbasis konnten Handlungsmuster zwischen Kaiser und Untertanen rekonstruiert und systematisch analysiert werden.¹

Mit der Entwicklung und Betreuung dieser Datenbank samt einer als Frontend zur Datenbank dienenden virtuellen Arbeitsumgebung wurde das an der Universität Graz beheimatete Zentrum für Informationsmodellierung – Austrian Centre for Digital Humanities (ZIM-ACDH) betraut. Dieses Zentrum ist eine seit 2008 existierende, aus einer bereits 1992 gegründeten Vorgängerinstitution namens GewiLab hervorgegangene Einrichtung der Geisteswissenschaftlichen Fakultät. Das Zentrum leistet angewandte Forschung auf dem Gebiet der Informationsverarbeitung in den Geisteswissenschaften mit den Forschungsschwerpunkten Digitale Edition, Collection Management, Datenrepositorien und Langzeitarchivierung von geisteswissenschaftlichen Forschungsdaten. Weitere Schwerpunkte

¹ Vgl. dazu die Beiträge von Ulrich Hausmann und Thomas Schreiber im vorliegenden Band.

liegen in der Anwendung semantischer Technologien und der Entwicklung fachspezifischer Software. Darüber hinaus bietet das Zentrum in der Lehre mehrere akademische Ausbildungsschienen für angehende Digital Humanists an und verfügt über eine eigene, leistungsfähige Serverinfrastruktur.²

Die folgenden Ausführungen beschreiben die Umsetzung des Projekts aus einer konzeptionell-technischen Perspektive. Dabei wird zuerst auf die Anforderungen des konkreten Forschungsprojekts und deren technische Umsetzung eingegangen, darauf aufbauend folgt eine allgemeinere Beschreibung von Herausforderungen und Lösungen für die IT-basierte Unterstützung ähnlicher Forschungsvorhaben.

Anforderungen und Workflows

Ziel der vom Zentrum für Informationsmodellierung übernommenen Aufgabe war es, eine verteilte Arbeitsumgebung im Sinne einer virtuellen Forschungsumgebung zu entwickeln, die es den beteiligten HistorikerInnen erlaubte, kollaborativ Daten aus Untertanensuppliken in einer Datenbank zu erfassen, diese Daten mit Fachwissen anzureichern und schließlich zu analysieren. Die geographische Verteilung auf die beiden Universitätsstandorte (Graz und Eichstätt) sowie die Anforderung, Daten auch direkt im Archiv erfassen zu können, erforderte eine netzwerkbasierte Lösung.

Im Juni 2012 kam es zu einem ersten Gespräch zwischen den FachwissenschaftlerInnen aus dem Projekt und MitarbeiterInnen des Zentrums. Zunächst galt es ein gemeinsames Verständnis des zu lösenden Problems zu erarbeiten und eine gemeinsame Vorstellung hinsichtlich der zu programmierenden Software zu entwi-

ckeln. In solchen Prozessen spielen Fragen der Datenmodellierung, d.h. welche Daten in welcher Form zu erfassen und zu speichern sind, eine zentrale Rolle, aber auch Fragen von Arbeitsabläufen, weil die Software diese möglichst gut unterstützen und sich in diese integrieren soll. In solchen Gesprächen konkretisieren sich erfahrungsgemäß nicht nur die Anforderungen an die zu schaffende Software, sondern es kommt nicht selten auch zu einer Anpassung der geplanten inhaltlichen Herangehensweise, weil sich Möglichkeiten, aber auch Herausforderungen einer computergestützten Methodik konkretisieren. Als hilfreich hat sich hier und in zahlreichen anderen Projekten erwiesen, dass die MitarbeiterInnen des Zentrums neben ihrer technischen Kompetenz auch über einen geisteswissenschaftlichen Hintergrund verfügen.

Diese Konzeptionierungsphase erstreckte sich über mehrere Wochen intensiver Kommunikation. Im Zuge des Entwurfs erster Modelle und Prototypen verbesserte sich das wechselseitige Verständnis der Anforderungen. Auch aus der konkreten Arbeit an den Quellen ergaben sich mehrfach neue oder veränderte Ansprüche, die nicht nur die Datenstruktur betrafen, sondern auch Erfordernisse der Datenbearbeitung, d.h. die Werkzeuge zur Datenmanipulation. Solche, sich über lange Zeiträume hinziehende iterative Entwicklungszyklen sind typisch für geisteswissenschaftliche Projekte und der diesen zugrunde liegenden hermeneutischen Methode. Man kann die den Forschungsprozess unterstützende Software als Teil des hermeneutischen Zirkels betrachten, weil die Software den mit wachsenden Erkenntnissen sich verändernden Erkenntniszielen anzupassen ist.³

Im konkreten Projekt kam es in diesem Sinne fast über den gesamten Projektzeitraum hinweg immer wieder zu Anpassungen der Software. Am Anfang standen jedoch konkret formulier-

²Nähere Informationen zum Zentrum finden sich unter [<http://informationsmodellierung.uni-graz.at/>] (abgerufen am: 19. 3. 2015).

³ Vgl. GADAMER, Zirkel 57–65.

bare Anforderungen, deren Implementierung sich in vier Projektphasen gliedern ließ. Im Fokus der ersten Projektphase stand die Entwicklung der projektinternen Datenbank. Dabei galt es, ein Datenmodell zu entwickeln, in dem alle für die Forschungsfragen der einzelnen Teilprojekte benötigten Daten abgebildet werden konnten, und dieses in einem Datenbankmanagementsystem zu implementieren. Außerdem war die Benutzerschnittstelle für ein möglichst einfaches Erfassen und Verwalten der Daten zu programmieren. In der zweiten Phase sollten Werkzeuge zur Normalisierung und Klassifizierung dieser Daten entwickelt werden. Phase drei war der Programmierung von Such- und Analysewerkzeugen für die Auswertung der erfassten Daten gewidmet. In einer abschließenden vierten Phase ging es darum, parallel zu dieser ausschließlich für projektinterne Zwecke gedachten Software einen funktional vereinfachten öffentlichen Zugang zu den Daten und zu den 20.485 fotografischen Abbildungen der Archivalien zu schaffen.⁴

Datenmodell und Datenbank

Bereits bei der ersten Besprechung wurde deutlich, dass wir es mit hochstrukturierten Daten zu tun bekommen würden. In der Informatik wird zwischen unstrukturierten, semistrukturierten und hochstrukturierten Daten unterschieden.⁵ Unstrukturierte Daten sind beispielsweise fortlaufende Texte, die keinerlei explizite Strukturinformation beinhalten. Implizit können solche Informationen vorhanden sein, sie sind jedoch für Computer nicht oder nur schwer interpretierbar. Während Strukturen wie Zeilen oder

Absätze durch entsprechende Trennzeichen noch einigermaßen einfach erkennbar sind, ist das Identifizieren einer Überschrift oder gar eines Personennamens im Text softwaretechnisch deutlich schwieriger.⁶

Für geisteswissenschaftliche Forschungsvorhaben wird häufig ein semistrukturierendes Datenmodell gewählt, weil damit Texte auf eine flexible Weise angereichert und erschlossen werden können. Dabei werden syntaktische, grammatische oder semantische Strukturen explizit und direkt im Text gekennzeichnet, was deren Identifizierung und Auswertung durch ein Computerprogramm erheblich erleichtert. Konkret werden dabei in den fortlaufenden Text Steuerzeichen in Form von Start- und Stoppsymbolen eingestreut, die etwa eine Überschrift, eine Orts- oder Personennennung oder Elemente des Urkundenformulars als solche kennzeichnen. Bereits in den 1980er Jahren hat ein international besetztes Konsortium namens *Text Encoding Initiative* (TEI) damit begonnen, eine primär für geisteswissenschaftliche Zwecke genutzte Auszeichnungssprache zu entwickeln und zu pflegen, die heute eine maßgebliche Rolle in diesem Bereich spielt.⁷

Für das hier beschriebene Projekt war diese Herangehensweise jedoch nicht geeignet, weil angesichts der fast 40.000 Aktenseiten eine Volltexterfassung nicht zu leisten war,⁸ und somit

⁴ Der Unterschied zu den oben genannten 38.520 Aktenseiten ergibt sich daraus, dass die meisten Fotos Doppelseiten (d.h. den aufgeschlagenen Akt) abbilden.

⁵ Vgl. etwa KAZAKOS, SCHMIDT, TOMCZYK, Datenbanken 26f.; VOSSEN, Datenmodelle 337.

⁶ Sie ist jedoch grundsätzlich, wenn eine gewisse Fehlerrate akzeptabel ist, machbar. Für inhaltliche Strukturen hat die Computerlinguistik dazu Werkzeuge wie das *Natural Language Processing* entwickelt. Vgl. dazu besonders in Hinblick auf historische Sprachquellen PIOTROWSKI, *Natural Language Processing*.

⁷ Vgl. TEI P5 Guidelines [<http://www.tei-c.org/Guidelines/P5/>] (2014 / abgerufen am: 19. 3. 2015).

⁸ Die manuelle Transkription tausender Aktenseiten ist ein sehr zeitaufwändiges Unterfangen. Mit der Technik der *Handwriting Text Recognition* (HTR) wird jedoch in absehbarer Zeit ein Werkzeug verfügbar sein, das diesen Aufwand erheblich reduzieren wird. Vgl. dazu MÜHLBERGER, *Volltexterkennung*; aus tech-

auch keine fortlaufenden Texte verfügbar waren. Stattdessen wurden durch BearbeiterInnen die für die Erkenntnisinteressen des Projekts relevanten Informationen aus den Akten extrahiert und hoch strukturiert, d.h. stark in einzelne Datenfelder segmentiert, in einer relationalen Datenbank gespeichert. Der Hauptgrund für diese Entscheidung und gegen Alternativen, wie etwa Dokumentendatenbanken oder Triple Stores, lag im hohen Standardisierungsgrad und der erprobten Technologie relationaler Datenbanken, insbesondere in Hinsicht auf die Prinzipien der Atomarität, Konsistenz, Isolation und Dauerhaftigkeit, der referentiellen Integrität, aber auch im Vorhandensein einer entsprechenden Infrastruktur am Zentrum, durch die ohne wesentlichen Mehraufwand Monitoring, Replikation und Datensicherung gewährleistet werden konnten. Wesentlich war auch die Performanz und Flexibilität hinsichtlich der Datenspeicherung und des *Data Retrieval* sowie die Verfügbarkeit eingeführter Programmierschnittstellen.

Da historische Daten oft heterogen, unscharf, unvollständig oder ambivalent sind, stellt deren Abbildung in ein striktes relationales Schema besondere Herausforderungen sowohl an die Datenmodellierung als auch die Abstraktionsfähigkeit der DatenerfasserInnen. Zahlreiche Probleme der Abbildung solcher Daten in eine Datenbankstruktur wurden erst im Zuge der Datenerfassung erkannt und bedingten mehrfach die Anpassung des Modells. Während zu Beginn ein noch relativ einfacher Datenbankentwurf mit weniger als 20 Tabellen stand, erhöhte sich diese Zahl während der darauf folgenden Diskussions- und Erprobungsphase auf mehr als 60 Tabellen.

nischer Perspektive die Beiträge in Institute of Electrical and Electronics Engineers, Handwriting Recognition.

Datenerfassung und Datenkorrektur

Die unterschiedlichen Phasen der Arbeit an und mit den Daten, nämlich die Erfassung der aus den Quellen extrahierten Primärdaten, deren Korrektur, die Normalisierung und Klassifizierung der Daten im Sinne einer Erschließung sowie die Datenanalyse, sind in der virtuellen Arbeitsumgebung durch jeweils eigene Bereiche repräsentiert, wobei die Erfassung und Korrektur aus praktischen Gründen in einem gemeinsamen Bereich durchgeführt wurden.

Die Datenerfassung erfolgte direkt an den Quellen im Archiv bzw. anhand eigens für das Projekt angefertigter Fotografien der Akten. Da diese Arbeiten gleichzeitig durch mehrere Personen erfolgten, musste sichergestellt werden, dass alle Beteiligten auf denselben, jederzeit aktuellen Datenbestand zugreifen konnten. Dies war wichtig, weil bereits erfasste Daten wie etwa Orte, Personen, Bestandsangaben usw. nur einmal erfasst werden sollten, um Redundanzen zu vermeiden und den Erfassungsvorgang zu beschleunigen. Gleichzeitig musste dafür gesorgt werden, dass ein exklusives Arbeiten an einem Datensatz auch über längere Zeiträume möglich war, ohne dadurch die Datenkonsistenz zu gefährden.

Die Eingabe erfolgte über HTML-Formulare, war also über jeden modernen Webbrowser möglich. Da eine beträchtliche Menge an Daten in einer komplexen Struktur erfasst wurden, die ein einzelnes Formular extrem unübersichtlich gemacht hätten, wählten wir einen modularen Ansatz: Ausgehend von einem auf eine Bildschirmseite passenden Hauptformular konnte eine Reihe von Subformularen geöffnet werden, die sich als verschiebbare Fenster über das Hauptformular legten. Das war überall dort besonders wichtig, wo mehr als ein zusätzlicher Teildatensatz eingegeben werden konnte, etwa wenn für ein Verfahren mehrere Supplikanten zu erfassen waren, oder sich das Verfahren in

mehrere Verfahrensschritte gliedern ließ. War ein Teildatensatz, etwa eine bestimmte Person oder ein Ort, bereits vorhanden, so konnten diese Daten durch einen einzigen Klick in das gerade bearbeitete Verfahren übernommen werden.

Bei der Entwicklung des Datenmodells und dessen Repräsentation im Erfassungsformular wurde an zahlreichen Stellen darauf Rücksicht genommen, dass die zu erfassenden Daten nicht immer eindeutig und vollständig vorlagen. Zusätzlich wurde die Möglichkeit geschaffen, einzelne Angaben als unsicher oder erschlossen zu kennzeichnen. Für die Erfassung von Datumsangaben wurde darauf geachtet, dass unscharfe Datumsangaben auch so gespeichert wurden, dass diese bei der Datenanalyse maschinell verarbeitet werden konnten. Dies wurde über Intervalldatierungen gelöst, die von BearbeiterInnen während der Erschließungsphase eingegeben wurden. Ergänzend dazu wurden automatisch aus den Datumsangaben generierte *Fuzzy Set Logic* Daten gespeichert.⁹ Zu dokumen-

⁹ *Fuzzy Set Logic* bezeichnet eine von Lofti A. Zadeh entwickelte unscharfe Logik, die mit Angaben wie „wahrscheinlich“, „vermutlich“ oder „eher nicht“ umgehen kann, indem entsprechende Zugehörigkeitsfunktionen definiert werden. Ein unscharfes Datum lässt sich beispielsweise als Trapez darstellen, das durch 4 Eckpunkte definiert ist. Die X-Achse entspricht dabei der Zeit, die Y-Achse einem Wert zwischen 0 (sicher nicht) und 1 (sicher). Der Y-Wert eines bestimmten Zeitpunkts (auf der X-Achse) bestimmt dann den Grad der Zugehörigkeit zur Datierung. Dieser zweite, automatisch generierte Ansatz zur Speicherung von Datumsangaben sollte dazu genutzt werden, die Brauchbarkeit dieser Methode gegen die durch menschliche Bearbeiter festgelegten Datierungen zu validieren. Allerdings stellte sich rasch heraus, dass wegen der unerwartet stark normalisierten Erfassung der quellennahen Datumsangaben die Validität einer solchen Evaluierung zweifelhaft war. Teile des dafür entwickelten Programmcodes wurden später dazu verwendet, bei der Erschließung der Daten den BearbeiterInnen Intervalldatierungen vorzuschlagen.

tarischen Zwecken wurde außerdem für jeden Datensatz der jeweilige Bearbeitungsstand protokolliert und zusätzlich wann und durch wen er gespeichert wurde.

Korrekturarbeiten erfolgten in den Formularen für die Datenerfassung. Um eine möglichst effektive Korrektur zu ermöglichen, konnten alle (Teil-)Datensätze mit den jeweils wichtigsten Feldern als filterbare Listen ausgegeben werden. Dadurch war es möglich, fehlerhafte Datensätze effizient zu identifizieren und aus den Listen heraus direkt zur Bearbeitung zu öffnen. Zusätzlich wurden Detailansichten realisiert, die jeweils alle Datenfelder eines Datensatzes in übersichtlicher Form darstellen, und durch die geblättert werden konnte. War Korrekturbedarf gegeben, ließ sich der betroffene Datensatz auch hier direkt durch einen Klick im Bearbeitungsmodus öffnen. Dieses Prinzip ist nicht nur auf ganze Verfahren anwendbar, sondern auch für einzelne Teildatensätze wie Orte, Personen, Verfahrensschritte, Personengruppen oder Kategorien. Dadurch lässt sich mühsames Navigieren durch größere Datensätze vermeiden. Dieser Ansatz hat sich in der Praxis als sehr nützlich und zeitsparend erwiesen.

Die Erfahrung aus früheren Projekten hat gezeigt, dass Datenkorrekturen bessere Ergebnisse bringen, wenn die Kontrollarbeiten nicht nur an den ursprünglichen Daten am Bildschirm durchgeführt werden. Die Möglichkeit, bestimmte Daten aus ihrem ursprünglichen Kontext herauszunehmen und gesammelt nebeneinander in Form sortierter Listen darzustellen, schärft ebenso den Blick für Fehler, wie die Kontrolle der Daten auf Papier. Deshalb bietet das System sowohl die Möglichkeit, solche Listen zu generieren, als auch des Exports der Detailansichten und Listen als PDF und im Excel-Format.

Normalisierung und Klassifizierung

Neben bzw. nach der eigentlichen Datenerfassung kam der Normalisierung und Klassifizierung der Daten große Bedeutung zu. Dabei wurde besonders darauf geachtet, dass die Resultate dieser Interpretationsleistungen die zuvor erfassten Daten so wenig als möglich veränderten, damit diese für eine allfällige Neuinterpretation verfügbar bleiben. Die Anreicherung der Daten erfolgte daher auch nicht über die für die Datenerfassung geschaffenen Formulare, sondern über speziell dafür programmierte tabellarische Darstellungen. Dabei wurde für jedes zu normalisierende oder zu klassifizierende Feld der Beleg (z.B. ein Name, ein Ort, eine Berufs- oder Standesbezeichnung oder eine Datumsangabe) unmittelbar neben einem zu befüllenden Eingabefeld dargestellt. Falls die in der Tabelle angezeigte Information nicht ausreichte, um eine Klassifizierung oder Normalisierung durchzuführen, konnte der vollständige Datensatz über einen Hyperlink eingesehen werden.

Die Tabellen ließen sich spaltenweise sortieren und nach Mustern filtern, sodass BearbeiterInnen zahlreiche Möglichkeiten hatten, identische oder ähnliche Daten zu identifizieren, nebeneinander darzustellen und damit sicherzustellen, dass konsistente Resultate erzielt wurden. Dabei wurden auf bereits vorhandenen Daten basierende Vorschläge gemacht, die direkt übernommen werden konnten, was nicht nur die Bearbeitungsgeschwindigkeit erhöhte, sondern auch die Datenkonsistenz verbesserte. Zu Kontroll- und Korrekturzwecken war es auch hier möglich, diese Tabellen im Excel-Format oder als PDF zu exportieren.

Such- und Analysemöglichkeiten

Ein komplexes Datenmodell erfordert auch komplexe Werkzeuge für die Filterung, Analyse, Darstellung und Weiterverarbeitung der Daten.

Zentral dafür sind die Suchmöglichkeiten. In der Datenbank erstreckt sich die Suche über mehr als 50 Datenfelder. Die üblicherweise als statisches Formular realisierten Suchmöglichkeiten hätten hier dazu geführt, dass ein über mehrere Bildschirmseiten reichendes und entsprechend unübersichtliches Formular entstanden wäre. Deshalb wählten wir einen anderen Weg in Form eines dynamischen, mit der Komplexität der Suchanfrage mitwachsenden Formulars. Dabei trifft der oder die Suchende aus einer Liste aller durchsuchbaren Feldnamen eine Auswahl (z.B. „Konfession“). Dadurch wird eine Formularzeile generiert, in der beliebig viele, durch „oder“ verknüpfte, zu suchende Werte für dieses Feld eingetragen bzw. ausgewählt werden können. Dieser Vorgang, also die Auswahl eines zu durchsuchenden Feldes mit der anschließenden Festlegung beliebig vieler zu suchender Werte, kann so oft wiederholt werden wie nötig, wobei jeweils eine zusätzliche Zeile im Suchformular entsteht. Die einzelnen Zeilen müssen durch logische Operatoren miteinander verknüpft werden. Das Suchergebnis wird dabei jeweils sofort an die gewählten Felder und Werte angepasst, sodass der oder die Suchende unmittelbar einen Eindruck von den Auswirkungen der Suche auf das Resultat hat.

Suchen nach Supplikanten [Hilfe](#)

Herrschaft: Bayern, Herzogtum oder Österreich / Böhme (148 Treffer)

und

Adressat: Kaiser oder Kaiserliche Hofkam (82 Treffer)

und nicht

Konfession: römisch-katholisch (80 Treffer)

Bedingung hinzufügen

Auf diese einfache Art ist es möglich, mit wenigen Mausklicks auch komplexe Suchabfragen zu formulieren wie etwa „zeige mir alle weiblichen Supplikanten lutherischen Glaubens aus Nürnberg, die älter als 50 Jahre waren, in deren Supplik in der Narratio ein bestimmtes Wort oder eine bestimmte Phrase vorkommt und deren Supplik von bestimmten kaiserlichen Ämtern behandelt wurde“.

Je nach Erkenntnisinteresse liefern diese Suchmöglichkeiten unterschiedliche Tabellen, deren Zeilen jeweils einen Supplikanten, ein Verfahren oder einen Verfahrensschritt beschreiben. Die tabellarischen Suchergebnisse lassen sich nach den einzelnen Spalten sortieren und weiter filtern, wodurch den BenutzerInnen ein interaktives Erkenntniswerkzeug zur Verfügung steht. Auch die Suchergebnisse lassen sich in verschiedene Formate exportieren, etwa um sie auszudrucken oder mit einem Statistikprogramm weiterzuverarbeiten. Jede Suche wird eindeutig über einen *Uniform Resource Identifier* (URI) referenziert. Dadurch können Suchabfragen als Lesezeichen gespeichert oder per E-Mail versendet werden. Beim erneuten Aufruf einer solchen Adresse werden nicht nur die Suchergebnisse reproduziert, sondern auch das Suchformular wiederhergestellt, wodurch Suchabfragen nachträglich verändert oder erweitert werden können.

Reicht die im tabellarischen Suchergebnis gebotene Information zur Beantwortung einer inhaltlichen Frage nicht aus, kann direkt aus dem Suchergebnis heraus der vollständige Datensatz zum jeweiligen Verfahren, Verfahrensschritt oder Supplikanten angefordert werden. Da diese Bereiche auch untereinander verknüpft sind, wird für die BenutzerInnen sofort ersichtlich, ob und welche anderen Verfahren durch einen Supplikanten angestrengt wurden, welche Verfahrensschritte für ein Verfahren durchgeführt wurden, welche Personen beteiligt waren usw.

Öffentlicher Zugang zur Datenbank

Die Datenbank enthält Daten, die als Grundlage für die inhaltliche Arbeit in den einzelnen Teilprojekten des Forschungsprojekts „Untertanensuppliken am Reichshofrat Kaiser Rudolfs II.

(1576–1612)“ dienen. Sie sind somit digital vorliegende Forschungsdaten.¹⁰ Da die Struktur dieser Daten relativ komplex ist und die Datenbank auch nicht für die Öffentlichkeit bestimmte Daten wie interne Arbeitsvermerke enthält, wurde eine zusätzliche, vereinfachte und öffentlich verfügbare Zugangsmöglichkeit zur Datenbank geschaffen.¹¹

Dieser öffentliche Zugang zur Datenbank bietet Recherche- und Suchmöglichkeiten nach Verfahren, Supplikanten und Akten. Der Zugriff auf die Daten erfolgt wahlweise entweder registerartig über sortier- und filterbare tabellarische Listen oder über ein Suchformular. Die Akten und damit die faksimilierten Seiten der Faszikel können zusätzlich noch über die Archivtektonik abgerufen werden. Diese drei Bereiche können je nach Erkenntnisinteresse gezielt verwendet werden, integrieren aber jeweils den Gesamtdatenbestand. Dies betrifft sowohl die Suchmöglichkeiten – die Suche nach Suppliken erlaubt auch Supplikanten-bezogene Suchbedingungen – als auch die Ausgabe der Suchergebnisse. Über Verlinkungen kann man im Suchergebnis jederzeit von einer Supplik zu den beteiligten Supplikanten wechseln, von einem Akt zur Supplik oder von einem gefundenen Supplikanten zu allen von dieser Person eingereichten Suppliken.

Für das Forschungsprojekt wurden durch einen Dienstleister im Archiv fotografische Faksimiles aller für das Projekt relevanten Akten erstellt. Diese Bilder waren vor allem als Hilfsmittel für die Erfassungs- und Korrekturarbeiten der Daten gedacht; eine Veröffentlichung der Bilder

¹⁰ Dem Thema Forschungsdaten wird seit einiger Zeit auch im geisteswissenschaftlichen Bereich vermehrt Aufmerksamkeit gewidmet. Zu Forschungsdaten allgemein vgl. LUDWIG, ENKE, Leitfaden; zu den Besonderheiten geisteswissenschaftlicher Forschungsdaten vgl. PEMPE, Geisteswissenschaften; SAHLE, KRONENWETT, Daten.

¹¹ [<http://www.gewi.uni-graz.at/suppliken/>] (abgerufen am: 19. 3. 2015).

war zunächst nicht beabsichtigt. Erst im Laufe des Projekts wurde die Möglichkeit diskutiert, diese Bilder im *World Wide Web* bereitzustellen. Es konnte mit dem Österreichischen Staatsarchiv eine entsprechende Vereinbarung getroffen werden, die es erlaubte, diese mehr als 20.000 Fotografien von fast 40.000 Aktenseiten auf den Servern des Zentrums für Informationsmodellierung bereitzustellen.

Damit diese Bilder möglichst allgemein verwendbar bleiben, wurden sie nicht direkt in die Datenbank integriert, sondern in Form von digitalen Objekten im Repositorium des Zentrums publiziert.¹² Dazu wurde ein Programm entwickelt, das für jeden Akt die wesentlichen Metadaten aus der Projektdatenbank extrahiert und daraus ein Dokument im METS-Format generiert, in dem auch alle diesen Akt repräsentierenden Bilder referenziert werden.¹³ Diese METS-Dateien wurden zusammen mit allen darin referenzierten Bildern ins Repositorium importiert. Beim Zugriff auf einen Akt wird dieser zusammen mit den Metadaten in einem Viewer dargestellt, der es erlaubt, durch den Akt zu blättern, in ein Bild zu zoomen oder direkt eine bestimmte Seite anzusteuern. Selbstverständlich enthält jeder Akt auch einen Hyperlink, der zum entsprechenden Verfahren in der öffentlich zugänglichen Datenbank führt. Auch umgekehrt können die Bilder eines jeden Akts direkt aus der Datenbank heraus angezeigt werden.

¹² Dieses unter dem Akronym GAMS (Geisteswissenschaftliches Asset Management System) bekannte Repositorium, das am Zentrum für Informationsmodellierung entwickelt und international von mehreren wissenschaftlichen Einrichtungen verwendet wird, findet sich unter der Adresse [<http://gams.unigraz.at/>] (abgerufen am: 15. 5. 2015). Vgl. STIGLER, Asset Management.

¹³ Vgl. Digital Library Federation, METS; ENDERS, Metadata Encoding.

Technische Lösung

Der Vollständigkeit halber soll hier noch kurz die hinter der Arbeitsumgebung stehende technische Lösung beschrieben werden. Die Benutzerschnittstelle wurde in HTML und JavaScript realisiert und ist somit in jedem aktuellen Webbrowser verwendbar. Die Kommunikation mit dem Server erfolgte über HTTP unter Anwendung des REST-Paradigmas.¹⁴ Die serverseitige Logikebene wurde über die WSGI-Schnittstelle und das Flask-Webframework¹⁵ implementiert, wobei die Arbeit durch eine am Zentrum für Informationsmodellierung entwickelte Programmbibliothek namens ZIMCRUD erheblich erleichtert wurde. Als Persistenzlayer fungiert eine PostgreSQL-Datenbank. Die Entscheidung für dieses Datenbankprodukt wurde auf Basis der oben beschriebenen hochgradigen Strukturierung der Daten getroffen. Rückblickend lässt sich festhalten, dass ein Documentstore als Datenbank vermutlich ebenso geeignet gewesen wäre, wegen der Möglichkeit flexiblerer Schemata angesichts der häufigen Nachbesserungen an der Datenstruktur vermutlich sogar die Entwicklungszeit (jedoch auf Kosten der Datenkonsistenz) verkürzt hätte. Wesentlich für die Auswahl aller beteiligten Softwarekomponenten war deren Verfügbarkeit unter einer *Open Source* Lizenz, die nicht nur die freie Verwendung und Weitergabe erlaubt, sondern auch Zugriff auf den Quelltext der Programme ermöglicht.

Verallgemeinerungen

Den Abschluss sollen ein paar allgemeine, also nicht nur für dieses Projekt spezifische Überlegungen zur Überführung historischer Quellen

¹⁴ *Representational State Transfer* beschreibt eine auf HTTP basierte verteilte Architektur. Vgl. TILKOV, REST.

¹⁵ [<http://flask.pocoo.org/>] (abgerufen am: 19. 3. 2015).

in eine digitale Repräsentation und deren Nutzung mithilfe computergestützter Werkzeuge und Methoden bilden.

Die Verwendung von Computern in der Geschichtswissenschaft hat eine lange Tradition.¹⁶ Bereits früh wurde von ExpertInnen verstanden, dass hier die vor allem für die Wirtschaft entwickelten Herangehensweisen und Werkzeuge nur bedingt geeignet waren. Vor allem in den 1980er Jahren wurde damit begonnen, die theoretischen Grundlagen einer historischen Fachinformatik zu erarbeiten.¹⁷ Diese theoriegeleiteten Ansätze traten in den Jahren der großen Digitalisierungsprojekte des letzten Jahrzehnts etwas in den Hintergrund, wurden aber zuletzt, als es darum ging, die wachsende Menge digital vorliegender Daten sinnvoll zu nutzen, wieder stärker beachtet.

Dabei kommt der Schnittstelle zwischen Geschichtswissenschaft und Informatik eine wichtige Bedeutung zu; nicht zuletzt weil hier die strikt logisch und formal bestimmte Informatik, also eine Naturwissenschaft, auf eine weitgehend hermeneutisch bestimmte Geisteswissenschaft trifft. Das Potential für wechselseitiges Miss- oder sogar Unverständnis ist hier groß, eine Erfahrung, die in zahlreichen Projekten gemacht werden musste. Als ein Indiz für das nicht unproblematische Verhältnis dieser beiden Disziplinen zueinander darf auch gewertet werden, dass während der letzten Jahre international zahlreiche Zentren für Digitale Geisteswissenschaften gegründet wurden, die eine Vermittlerfunktion zwischen den beiden Disziplinen einnehmen.¹⁸

¹⁶ Als frühe Publikation etwa GUNDLACH, LÜCKERATH, Datenverarbeitung.

¹⁷ Hier sind im deutschsprachigen Raum vor allem die Arbeiten von Manfred Thaller und Ingo H. Kropač zu nennen. Vgl. KROPAČ, Konzeption; THALLER, Exaktheit; DERS., Datenverarbeitung.

¹⁸ Einen Überblick bietet [<http://www.dhcenternet.org/>] (abgerufen am: 19. 3. 2015).

In einem geschichtswissenschaftlichen Forschungsprojekt, das computergestützte Werkzeuge erfolgreich nutzen will, ist also besonders darauf zu achten, dass zwischen den beteiligten FachwissenschaftlerInnen und InformatikerInnen ein gegenseitiges Verständnis der beabsichtigten Herangehensweise entsteht. Dies ist ein wechselseitiges Herantasten: Die InformatikerInnen müssen nicht nur das zu lösende Problem an sich verstehen, sondern auch den Weg, wie in den Geisteswissenschaften Wissen generiert wird. Umgekehrt müssen erfahrungsgemäß HistorikerInnen häufig erst darauf hingewiesen werden, welche zusätzlichen Möglichkeiten sich aus der Anwendung computergestützter Methoden ergeben.

Vor einigen Jahren begann die Diskussion über sog. Virtuelle Forschungsumgebungen (*Virtual Research Environment*, VRE). Die Idee einer integrierten Software, die Forscher in ihrer Forschungstätigkeit, beginnend bei der Literaturrecherche, der Datenerfassung, der Kommunikation bis hin zum Schreiben und Publizieren der Ergebnisse unterstützt, erwies sich als zu ambitioniert. Zu verschiedenartig sind die Erfordernisse, zu unterschiedlich die Workflows, um diese sinnvoll in eine gemeinsame Software integrieren zu können.¹⁹ Inzwischen zeichnet sich ein Trend zur Spezialisierung ab. Nicht mehr die eine, allgemein und für alles verwendbare Lösung wird angestrebt, sondern disziplin- oder sogar teildisziplinabhängige Lösungen. Dass das weitgehende Fehlen solcher Umgebungen durchaus als Problem wahrgenommen wird, zeigt die hohe Zahl an entsprechenden Einreichungen im Rahmen von *Horizon2020*: 95

¹⁹ Vgl. DOMBROWSKI, Project Bamboo. Überblicksartig zum Thema: HUEGI, SCHNEIDER, Forschungsinfrastrukturen.

Proposals für virtuelle Forschungsumgebungen gingen im Bereich E-Infrastructures ein.²⁰

Solange noch keine brauchbaren, flexibel verwendbaren Forschungsumgebungen existieren, die bei der Modellierung, Erfassung und Analyse historischer Daten unterstützen, bleiben zwei Ansätze für die Realisierung solcher Projekte: die Verwendung von Standardsoftware oder die Entwicklung neuer bzw. Adaption bestehender Software. Wenn es darum geht, Daten strukturiert zu erfassen, stehen zahlreiche Möglichkeiten zur Verfügung. Die Wahl hängt stark von den zu erfassenden Daten und davon ab, wie mit den Daten später gearbeitet werden soll. Im geisteswissenschaftlichen, insbesondere im historischen Bereich hat sich gezeigt, dass der Einsatz von Softwareprodukten „von der Stange“ problematisch ist. Der Historiker und Informatiker Manfred Thaller hat bereits 1988 an seinem „Preußenbeispiel“ gezeigt, dass die Verarbeitung historischer Information alles andere als trivial ist, weil diese Information in der Regel heterogen, unvollständig, unscharf und mehrdeutig ist. Noch dazu hängt die korrekte Interpretation solcher Daten stark von sich verändernden Kontexten wie politischen und administrativen Grenzen oder Währungs- und Maßsystemen ab.²¹ Als Konsequenz hat Thaller eine eigene Datenbanksoftware entwickelt, die Unterstützung für genau diese Probleme bot,²² deren Weiterentwicklung jedoch leider vor etwa zehn Jahren stagnierte. Die Probleme blieben aber dieselben und sind nach wie vor mit Standardsoftware nur unzureichend lösbar.

Die Verwendung von Standardsoftware geht daher immer mit einem Kompromiss bezüglich der Datenqualität, genauer gesagt mit oft unzureichenden Abbildungsfunktionen der Daten,

sowie mit Abstrichen hinsichtlich der Arbeitsmöglichkeiten einher. Daher ist Individualsoftware, wie sie für das Supplikenprojekt entwickelt wurde, meist der bessere, wenngleich kostenintensivere Weg, weil die Software an die individuellen Bedürfnisse des Projekts angepasst wird. Es wäre wünschenswert, wenn im Lauf der Zeit aus solchen Projekten eine Masse an wiederverwendbaren Softwarekomponenten entstehen würde, wodurch der Entwicklungsaufwand für ähnliche Projekte reduziert werden könnte. Erfahrungsgemäß scheitert dieser Anspruch aber häufig daran, dass aus Zeit- und Kostengründen darauf verzichtet wird, die Software entsprechend modularisiert und von der konkreten Problemstellung abstrahiert zu entwickeln.

Letztendlich muss für jedes Projekt hinsichtlich der eingesetzten Software individuell zwischen benötigter Anpassung und Funktionalität und dafür einzusetzenden Ressourcen entschieden werden. Das ist eine Entscheidung, die erfahrungsgemäß erst im Dialog zwischen FachwissenschaftlerInnen und IT-ExpertInnen (idealerweise mit einem geisteswissenschaftlichen Hintergrund) getroffen werden kann. Ebenso muss bereits früh überlegt werden, was mit Daten und Software nach Beendigung des Projekts geschehen soll. Daten und Software können immer nur in einer spezifischen technischen Umgebung genutzt werden. Daher muss dafür gesorgt sein, dass zumindest die Daten, idealerweise aber auch die Software, auch nach Projektende an veränderte technische Umgebungen angepasst werden und so nutzbar bleiben. Diese Tätigkeit wird als Datenkuratierung (*data curation*) bezeichnet und ist unabdingbar, wenn Daten und Software längerfristig nutzbar bleiben sollen. Leider gibt es zurzeit noch kaum institutionelle Einrichtungen, die diese Aufgaben für geisteswissenschaftliche Daten übernehmen. Allerdings hat sich im Rahmen der wissenschaftlichen Vereinigung „Digital Humanities im deutschsprachigen Raum“ vor etwas mehr

²⁰ [<http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/calls/h2020-einfra-2015-1.html>] (2015 / abgerufen am: 19. 3. 2015).

²¹ THALLER, Datenverarbeitung 57–83.

²² DERS., Automation 40–65.

als einem Jahr eine international besetzte Arbeitsgruppe zu diesem Thema gebildet, die entsprechende Vorschläge liefern soll.²³

Die Wahl der Software beeinflusst die Struktur und in weiterer Folge das Format der Datenspeicherung. Die Erfahrungen des Zentrums für Informationsmodellierung und der dort tätigen Personen aus rund 20 Jahren als Projektpartner und Datenhoster für geisteswissenschaftliche Forschungsprojekte zeigen deutlich, dass die langfristige Verfügbarkeit solcher elektronischer Ressourcen wesentlich davon bestimmt wird, dass bereits in der Konzeptionierungsphase hinsichtlich Auswahl von Software und Datenformaten auf diesen Aspekt Rücksicht genommen wird. Grundsätzlich muss zwischen Langzeitarchivierung und Langzeitverfügbarkeit unterschieden werden. Unter Langzeitarchivierung versteht man das langfristige sichere Aufbewahren von Daten. Das beinhaltet nicht nur eine reine Bitstream-Preservation, sondern inkludiert auch, dass diese, bevorzugt in einem textbasierten und dokumentierten Format vorliegenden Daten durch Metadaten erschlossen sind. Langzeitverfügbarkeit hingegen zielt nicht nur auf die Erhaltung der Daten, sondern ebenso auf die Erhaltung der Präsentation dieser Daten ab, wozu zunehmend auch interaktive Elemente gehören. Es geht also darum, nicht nur die Daten längerfristig zu bewahren, sondern auch die Software, die diese Daten darstellt, durchsuchbar und analysierbar macht.

Diesen Fragen wird unter dem Begriff Forschungsdatenmanagement seit einigen Jahren verstärkt Aufmerksamkeit gewidmet. Die Idee dahinter ist, dass diese oft mit erheblichem Aufwand generierten Daten aufbewahrt und idealerweise öffentlich zugänglich gemacht werden sollten. Der ursprüngliche Grund dafür war die vor allem in den Naturwissenschaften

extrem wichtige Reproduzierbarkeit von Ergebnissen, die die Verfügbarkeit der diesen zugrunde liegenden Forschungsdaten bedingt. Ein weiterer Grund ergibt sich aus der potentiellen Wiederverwendbarkeit und somit Nachnutzbarkeit von Daten für andere Forschungsfragen. Diskutiert werden in diesem Zusammenhang vor allem standardisierte Datenformate, Nachweis der Provenienz und der Regeln, nach denen diese Daten erfasst wurden, aber auch Schnittstellen zu den Daten und rechtliche Probleme der Nachnutzung. Auch dies sind Fragen, die bereits in der Konzeptionierungsphase eines Projekts bedacht werden sollten. Je früher man sich über Nutzungs- und Nachnutzungsszenarien klar wird, desto geringer sind die Auswirkungen hinsichtlich der nachträglichen Anpassung von Datenformaten und Softwarelösungen.

Zusammenfassung

Abschließend kann festgehalten werden, dass sich die Umsetzung der Arbeitsumgebung hinsichtlich der sich daraus ergebenden Forschungsmöglichkeiten als erfolgreich erwiesen hat. Der Einsatz von Datenbanken ermöglicht neue Fragestellungen an das Material. Dabei darf jedoch der Aufwand der Datenerfassung nicht unterschätzt werden. Einmal mehr hat sich gezeigt, dass ein solches Projekt nur dann erfolgreich umgesetzt werden kann, wenn beide Seiten, nämlich die HistorikerInnen und die für die technische Umsetzung Verantwortlichen im Vorfeld so lange kommunizieren, bis beide Seiten eine gemeinsame Sprache und einen Konsens bezüglich der zu schaffenden Möglichkeiten gefunden haben. Ebenso wichtig ist, dass diese Verbindung auch während der Arbeit an den Daten nicht verloren geht, weil eine hermeneutische Herangehensweise nicht nur sich verändernde Sichten auf das Problem, sondern damit Hand in Hand gehend auch veränderte Ansprüche an die Software mit sich bringt.

²³ [<http://www.dig-hum.de/arbeitsgruppedatenzentren/>] (abgerufen am: 19. 3. 2015).

Ein auf der digitalen Erfassung von historischen Daten basierendes Projekt hat jedoch nicht nur technische Probleme im Umgang mit den dabei entstehenden Daten zu lösen, sondern sollte auch Überlegungen auf strategischer Ebene anstellen. Dabei geht es um Fragen der Bereitstellung, d.h. Veröffentlichung von Daten und somit um deren potentielle Nachnutzung. Sehr wichtig in diesem Zusammenhang ist das Problem der Kuratierung als Voraussetzung für Langzeitarchivierung und Langzeitverfügbarkeit. Bereits zu Beginn des Projekts sollten Überlegungen angestellt werden, was nach Projektende mit den Daten und der Software geschehen soll, die im Rahmen des Projekts entstanden sind. Wird eine längerfristige Verfügbarkeit ins Auge gefasst, so sollte idealerweise bereits in der Konzeption des Projekts mit der Institution, die später für die Aufbewahrung der Daten verantwortlich ist, Kontakt aufgenommen werden, weil die Entscheidung für eine bestimmte Software oder ein bestimmtes Datenformat den Aufwand für eine längerfristige Bereitstellung massiv beeinflussen kann. Ideal ist, wenn wie im vorliegenden Projekt die Softwareentwicklung von derselben Institution geleistet oder zumindest überblickt wird, die später auch die Verantwortung für die Aufbewahrung der Daten übernimmt.

Korrespondenz:

Mag. Gunter Vasold
Zentrum für Informationsmodellierung – Austrian
Centre for Digital Humanities (ZIM-ACDH)
Elisabethstraße 59/III, 8010 Graz
gunter.vasold@uni-graz.at

Abkürzungen:

Siehe das allgemeine Abkürzungsverzeichnis:
[<http://www.rechtsgeschichte.at/files/abk.pdf>]

Literatur:

- Digital Library Federation (Hg.), METS – Metadata Encoding and Transmission Standard: Primer and Reference Manual, Vers. 1.6 Revised
[<http://www.loc.gov/standards/mets/METSPRimerRevised.pdf>] (2010 / abgerufen am: 19. 3. 2015).
- Quinn DOMBROWSKI, What ever happened to Project Bamboo?, in: *Literary and Linguistic Computing* 29 (2014) 326–339.
- Markus ENDERS, Metadata Encoding and Transmission Standard – Einführung und Nutzungsmöglichkeiten, in: Heike NEUROTH u.a. (Hgg.), *Nestor-Handbuch. Eine kleine Enzyklopädie der digitalen Langzeitarchivierung* (Boizenburg 2009) Kapitel 6, 3–8.
- Hans-Georg GADAMER, Vom Zirkel des Verstehens, in: DERS., *Gesammelte Werke*, Bd. 2: Hermeneutik. Wahrheit und Methode, Bd. 2 (Tübingen 1999) 57–65.
- Rolf GUNDLACH, Carl August LÜCKERATH, *Historische Wissenschaften und elektronische Datenverarbeitung* (Frankfurt am Main 1976).
- Jasmin HUEGL, René SCHNEIDER, *Digitale Forschungsinfrastrukturen für die Geistes- und Geschichtswissenschaften*
[http://www.infoclio.ch/sites/default/files/standard_page/studie_forschungsinfrastrukturen_small.pdf] (2013 / abgerufen am: 1. 6. 2015).
- Institute of Electrical and Electronics Engineers (Hg.), *14th International Conference on Frontiers in Handwriting Recognition (ICFHR 2014)* (Piscataway, NJ 2014).
- Wassilios KAZAKOS, Andreas SCHMIDT, Peter TOMCZYK, *Datenbanken und XML* (Berlin–Heidelberg 2002).

- Ingo H. KROPAČ, Zur Konzeption von Informationssystemen in der Geschichtswissenschaft, in: Karl KASER, Karl STOCKER (Hgg.), *Clios Rache. Neue Aspekte strukturgeschichtlicher und theoriegeleiteter Geschichtsforschung in Österreich* (Wien-Köln-Weimar 1992) 87–129.
- Jens LUDWIG, Harry ENKE, *Leitfaden zum Forschungsdatenmanagement* (Glückstadt 2013).
- Günter MÜHLBERGER, Die automatisierte Volltexterkennung historischer Handschriften als gemeinsame Aufgabe von Archiven, Geistes- und Computerwissenschaftlern, in: Irmgard Ch. BECKER, Stephanie OERTEL (Hgg.), *Digitalisierung im Archiv. Neue Wege der Bereitstellung des Archivguts. Beiträge zum 18. Archivwissenschaftlichen Kolloquium der Archivschule Marburg* (= Veröffentlichungen der Archivschule Marburg, Hochschule für Archivwissenschaft 60, Marburg 2015) 87–116.
- Wolfgang PEMPE, Geisteswissenschaften, in: Heike NEUROTH u.a. (Hgg.), *Langzeitarchivierung von Forschungsdaten* (Boizenburg 2012) 137–159.
- Michael PIOTROWSKI, *Natural Language Processing for Historical Texts* (San Rafael, CA 2012).
- Patrick SAHLE, Simone KRONENWETT, *Jenseits der Daten. Überlegungen zu Datenzentren für die Geisteswissenschaften am Beispiel des Kölner „Data Center for the Humanities“*, in: *Libreas* 23 (2013) 76–96.
- Hubert STIGLER, *Warum Asset Management?*, in: *Zeitschrift für Hochschulentwicklung* 1 (2006) 63–74.
- Manfred THALLER, *Automation on Parnassus. Clio – A Databank Oriented System for Historians*, in: *Historical Social Research* 5 (1980) 40–65.
- DERS., *Gibt es eine fachspezifische Datenverarbeitung in den historischen Wissenschaften?*, in: Karl Heinrich KAUFHOLD, Jürgen SCHNEIDER (Hgg.), *Geschichtswissenschaft und elektronische Datenverarbeitung* (= Beiträge zur Wirtschafts- und Sozialgeschichte 36, Wiesbaden 1988) 45–83.
- DERS., *Ungefähre Exaktheit: Theoretische Grundlagen und praktische Möglichkeiten einer Formulierung historischer Quellen als Produkte „unscharfer Systeme“*, in: Herta NAGL-DOCEKAL, Franz WIMMER (Hgg.), *Neue Ansätze in der Geschichtswissenschaft* (Wien 1984) 77–100.
- Stefan TILKOV, *REST und HTTP. Einsatz der Architektur des Web für Integrationsszenarien* (Heidelberg 2011).
- Gunter VASOLD, *Edition à la carte? Usability, Interfacing und Datenmigration für webbasierte Informationssysteme*, in: Klaus VAN EICKELS, Ruth WEICHSELBAUMER, Ingrid BENNEWITZ (Hgg.), *Mediaevistik und Neue Medien* (Ostfildern 2004) 261–279.
- Gottfried VOSSEN, *Datenmodelle, Datenbanksprachen und Datenbankmanagementsysteme* (München 2008).

