SSOAR

Open Access Repository
www.ssoar.info

Von der Quelle zum Datensatz: Vorschläge zum Einsatz datenbankorientierter Verfahren in der universitären Lehre

Kropac, Ingo H.

Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

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

Empfohlene Zitierung / Suggested Citation:
Kropac, I. H. (1986). Von der Quelle zum Datensatz: Vorschläge zum Einsatz datenbankorientierter Verfahren in der universitären Lehre. In M. Thaller (Hrsg.), *Datenbanken und Datenverwaltungssysteme als Werkzeuge historischer Forschung* (S. 199-216). St. Katharinen: Scripta Mercaturae Verl. https://nbn-resolving.org/urn:nbn:de:0168-ssoar-341419

Nutzungsbedingungen:
Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.
Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:
This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.
By using this particular document, you accept the above-stated conditions of use.

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

INGO H. KROPAC

Von der Quelle zum Datensatz. Vorschläge zum Einsatz datenbankorientierter Verfahren in der universitären Lehre.

1. Vorbemerkung

Während die grundlegenden Fragen, wie quantifizierende Geschichte an den Universitäten zu lehren wäre, in den späten Siebzigerjahren in der Bundesrepublik Deutschland und auch in Österreich ausführlich diskutiert und zum Gutteil beantwortet wurden, [1] fehlen bislang Konzepte zum Einsatz der quellennahen oder quellenorientierten Datenverarbeitung [2] in der historisch/akademischen Ausbildung. Vorteile und Nutzen dieses zumindest auf dem Gebiet der Geschichtsforschung relativ jungen Zweiges der EDV-Anwendung stehen heute wohl außer Zweifel [3];so schien es sowohl aufgrund mehrerer in Graz laufender Projekte [4] als auch in Hinblick auf die Verbes-

---

[1] Zu grundlegenden Problemen der Didaktik von quantifizierenden Methoden vgl. *Thomas Schuler:* Quantifizierung und Geschichtsstudium, in: HSR 17, 1981,S. 61f. - *Heinrich Best und Wilhelm H. Schröder:* Basiscurriculum für eine quantitative historische Sozialforschung. Vorschläge für eine Einführungsveranstaltung am Beispiel des ZENTRUM-Herbstseminars, in: HSR 17, 1981, S. 3-50. - *Henning Bauer:* Der Einsatz archivierter Daten in der Lehre der Historischen Sozialforschung, in: HSR 24, 1982, S. 63-72.
Davon zeugen auch die im deutschen Sprachraum erschienenen Lehrbücher, wie *Norbert Ohler:* Quantitative Methoden für Historiker, München, 1980., oder *Manfred Thaller:* Numerische Datenverarbeitung für Historiker, Wien / Köln, 1982 (= Materialien zur Historischen Sozialwissenschaft. 1.).

[2] Vgl. *Konrad H. Jarausch, Gerhard Arminger, Manfred Thaller:* Quantitative Methoden in der Geschichtswissenschaft. Eine Einführung in die Forschung, Datenverarbeitung und Statistik, Darmstadt, 1985, S. 65-69.

[3] Eine Beschreibung der wichtigsten sechs Punkte findet sich bei *Konrad H. Jarausch, Gerhard Arminger, Manfred Thaller:* Quantitative Methoden in der Geschichtswissenschaft. Eine Einführung in die Forschung, Datenverarbeitung und Statistik, Darmstadt, 1985, S. 68.

[4] Neben dem in diesem Band vorgestellten Unternehmen zur Migrationsgeschichte steht in Graz das Projekt "Untersuchung der Einsatzmöglichkeiten der EDV für historische Faktendokumentation im Bereich der historischen Grundwissenschaften (HIFADOK)" in Bearbeitung, das sich dzt. in der Hauptsache mit maschinengestützter Edition und dem Aufbau einer prosopographischen Datenbank

serung und Vorbereitung statistischer Analysen wünschenswert, eine Einführung auf diesem Gebiet in das Lehrangebot aufzunehmen. Eine Möglichkeit der Vermittlung solcher quellennaher und datenbankorientierter Verfahren, wie sie in Graz verwirklicht wurde [5], soll im folgenden diskutiert werden.

Bereits seit 1980 werden am Institut für Geschichte Lehrveranstaltungen angeboten, die eine Einführung in die Probleme der Quantifizierung zum Inhalt haben [6]. Diese von St. Karner betreuten Übungen betreffen besonders drei große Sachgebiete:

- Vorstellung des EDV-Zentrums und Erlernen der wichtigsten Betriebssystem- und Editorbefehle der UNIVAC 1100.
- Einführung in die Quantifizierung und Diskussion theoretischer und methodischer Fragen.
- Erste Arbeiten mit dem statistischen Programmsystem BMDP.

Die Studierenden, welche die Mühe auf sich nehmen, ein arbeitsintensives Freifach zu besuchen, betreten damit in zweifacher Hinsicht Neuland: Zum einen müssen sie den Umgang mit dem Rechner bzw. seiner Peripherie erlernen, zum anderen werden sie mit einem ihnen noch unbekannten Methodenkanon konfrontiert [7]. Die Absolventen dieser Lehrveranstaltungen hatten jedoch bisher auf universitärem Boden keine Möglichkeit, nach der Einführung auch weiter ausgebildet zu werden, weder in rein arbeitstechnischer Hinsicht (EDV-Einsatz), noch in methodischer (Quantifizierung). Nur für eine Minderzahl bot sich die Gelegenheit, innerhalb eines Forschungsprojektes von Fachkräften betreut und unterwiesen zu werden und/oder den

---

bis 1250 beschäftigt und am Forschungsinstitut für Historische Grundwissenschaften unter der Leitung von Univ.-Prof. Dr. *Friedrich Hausmann* durchgeführt wird.

[5] Das vorliegende Papier beschäftigt sich mit dem Konzept und ersten Erfahrungen über eine Lehrveranstaltung, die unter der gemeinsamen Leitung von Dr. *Hermine Losch* und des Verfassers im Sommersemester 1985 neben der obenerwähnten Einführung erstmals abgehalten wird; ihr Titel lautet "EDV und Geschichtswissenschaft II: Von der Quelle zum Datensatz."

[6] Vgl. *Stefan Karner:* Erfahrungen mit Quantifizierung und EDV in der Lehre, In: Bericht über den 16. Öst. Historikertag in Krems / Donau (= Veröff. d. Verbandes Öst. Geschichtsvereine 25), o.O., 1985, S. 684-686.

[7] Zur Situation der Lehre an den österreichischen Universitäten in bezug auf Quantifizierung vgl. *Gerhard Botz:* Quantifizierende Geschichte als Gegenstand (inter-)universitärer Weiterbildung, In: Bericht über den 16. Öst. Historikertag in Krems / Donau (= Veröff. d. Verbandes Öst. Geschichtsvereine 25), o.O., 1985, S. 667-675.

QUANTKURS in Salzburg zu besuchen, dessen Bedeutung für die Quantifizierung in Österreich gar nicht hoch genug veranschlagt werden kann [8].

Im Anschluß an die Sitzung des Arbeitskreises "Quantifizierung in der österreichischen Geschichtsforschung" auf dem österreichischen Historikertag 1984, in der diverse didaktische Konzepte der Quantifizierung und EDV-Anwendung diskutiert wurden, erfolgte am Ludwig-Boltzmann-Institut für Historische Sozialwissenschaft eine integrative Formulierung jener Grundsätze, die in der Folge auch Basis für die einschlägigen Grazer Lehrveranstaltungen wurden.

2. Zielsetzungen

Um die Vermittlung eines einheitlichen Basiswissens über EDV-Anwendung und Quantifizierung zu ermöglichen, sollten universitäre Lehrveranstaltungen demnach folgende Schwerpunkte zum Inhalt haben [9]:

- Einführung in die methodologischen Voraussetzungen
- Quellenkritik
- theorie- und hypothesengeleitetes Entwickeln von Fragestellungen
- Einführung in die (zumindest deskriptive) Statistik
- Einführung in das lokale EDV-System
- Einführung in statistische Programmpakete (SPSS, SPSS-X, BMDP)
- Einführung in ein historisches Datenbanksystem (CLIO)
- Interpretation und Diskussion von Ergebnissen quantifizierenden Arbeitens.

Diese Bereiche entsprechen den Erfahrungen aus dem QUANTKURS und sollen derzeit an den Universitäten Wien, Salzburg und Graz realisiert werden.

Die bereits angesprochene äußerst unbefriedigende Situation an der Grazer Universität im Bereich der EDV-Anwendung und Quantifizierung führte im Sommer 1984 zu Überlegungen, auf welche Weise hier Abhilfe gefunden werden könnte. Aufgrund der obengenannten Schwerpunkte und der Forderung, auch quellennahe und so eher qualitative Datenverarbeitung zu vermitteln, wurde eine Zielvorstellung entwickelt, die in mehreren, aufeinander aufbauenden Lehrveranstaltungen verwirklicht werden soll:

---

[8] Vgl. *Gerhard Botz:* Quantkurs am Scheideweg, in: Historical Social Research (HSR) 29, 1984, S. 86-95.

[9] Überlegungen zur Integration der Ausbildung in Quantifizierung und EDV-Anwendung für Historiker in den regulären Studiengang österreichischer Universitäten, hrsg.v. Ludwig-Boltzmann-Institut für Historische Sozialwissenschaft, 1984 (unveröffentlichtes Papier).

Vorerst gilt es, grundlegendes Wissen über die EDV als Arbeitstechnik (ob man sie nun als Grund-, Hilfs- oder Nachbarwissenschaft sehen will, spielt dabei keine Rolle) zu vermitteln und deren Anwendungsmöglichkeiten kurz vorzustellen. Damit auf das engste verknüpft ist die Kenntnis der theoretischen, methodischen und konzeptuellen Voraussetzungen quellennaher und methodenorientierter Datenverarbeitung.

Weiters soll der Historiker mit den für ihn spezifischen Programmsystemen umzugehen lernen, nämlich mit dem fachspezifischen und datenbankorientierten Programmsysten CLIO sowie dem Statistikpaket SPSS. Dies soll ihm selbständiges Arbeiten am Computer ermöglichen oder dazu anregen.

Studierende sollen demnach in die Lage versetzt werden, aufgrund ihrer universitären Ausbildung kleinere Projekte von der Seminararbeit bis zur Dissertation selbsttätig zu planen und durchzuführen, und zwar sowohl arbeitstechnisch am Rechner als auch methodisch, z.B. quantifizierend. Daneben sollten sie ohne grössere Probleme kompliziertere Verfahren in beiderlei Hinsicht erarbeiten können, entweder durch Literaturstudium und/oder durch weiterführende Lehrangebote, wie sie etwa die Fortgeschrittenen-Gruppen des QUANTKURSES darstellen [10].

Ausgangssituation

Das didaktische Konzept einer solcherart vordefinierten Lehrveranstaltung ist natürlich von zahlreichen Imponderabilien abhängig, von denen die Ausstattung und Kapazität des lokalen Rechenzentrums, die Zusammensetzung der Zielgruppe, also der Studierenden, sowie die rechtlichen Grundlagen wie Bewilligung, Anrechenbarkeit im Studienplan etc. herauszustreichen sind. Obwohl hier ein Vergleich mit anderen Universitäten kaum möglich ist, trugen diese lokal bedingten Einflüsse wesentlich zur Konzeption der Lehrveranstaltung bei und bedürfen demnach auch einer Erläuterung. Vor allem sind es jedoch die methodisch/konzeptuellen Vorüberlegungen, die den Rahmen für die Verwirklichung der vorgegebenen Ziele abstecken.

Rechenzentrum

Wie die meisten auf naturwissenschaftliche, also rein numerische Anwendungen spezialisierten Universitätsrechenzentren bietet auch das "EDV-Zentrum der Karl-Franzens-Universität Graz" dem Historiker vorerst nicht

[10] Zu weiteren Möglichkeiten, wie etwa Sommerschulen, vgl. *Gerhard Botz:* Quantifizierende Geschichte als Gegenstand (inter-)universitärer Weiterbildung, In: Bericht über den 16. Öst. Historikertag in Krems / Donau (= Veröff. d. Verbandes Öst. Geschichtsvereine 25), o.O., 1985, S. 671.

allzu viel. Die Grazer Universität hat neben anderen Institutionen Anteile an der Rechenkapazität der Großrechenanlage UNIVAC 1100/81 (SPERRY) und besitzt als eigenen Bereichsrechner eine VAX 11/750 (DEC). Beide Systeme sind jedoch durch ihre spezifische Ausstattung an Hardwarekomponenten und Software für die Anliegen eines Historikers keineswegs ideal [11] . Während auf der UNIVAC zwar die Programmsysteme CLIO, SPSS und SPSS-X laufen, erweist sich der zeilenorientierte Editor als äußerst unkomfortabel und das Betriebssystem unterstützt weder den deutschen Zeichensatz noch die für die maschinelle Edition oft unentbehrlichen Sonderzeichen. Zudem ist auch der am EDV-Zentrum installierte Drucker der UNIVAC nicht in der Lage, Groß- und Kleinschreibung wiederzugeben. Im Gegensatz dazu zeichnet sich die VAX durch einen sehr komfortablen Full-Screen-Editor aus, unterstützt seit jüngstem auch den deutschen Zeichensatz, und auf den zur Verfügung stehenden Druckern lassen sich - neben der eigentlich selbstverständlichen Groß- und Kleinschreibung - auch Umlaute und verschiedenste Sonderzeichen darstellen. Leider ist sie ausschließlich mit naturwissenschaftlich orientierter Software ausgestattet, was Anwendung auf nichtnumerischem Gebiet - mit Ausnahme des Editors - derzeit verhindert; allerdings ist die VAX als Remote-Station an die UNIVAC angeschlossen.

Ein weiteres Problem stellen die zugänglichen Kapazitäten der überlasteten und zum Teil nicht mehr ganz neuen Systeme dar. Abgesehen davon, daß die Abarbeitung größerer Jobs sehr lange dauert, sind komplexere Aufgaben innerhalb der bestehenden Limits nur in Einzelfällen zu verwirklichen. Auch der einzige Terminalraum, der mit 15 Terminals (!) ausgestattet ist, verhindert die Abhaltung von Lehrveranstaltungen untertags.

3.2. Zielgruppe

Die heterogene Zusammensetzung jener Gruppe von Studenten und Graduierten, welche sich im Sommersemester zusammengefunden hat(te), bestätigte nur die Erwartungen. Der unterschiedliche Ausbildungsstand, nicht nur in bezug auf EDV-Einsatz oder Quantifizierung, sondern auch was etwa paläographische Kenntnisse oder Quellenkritik betrifft, erschwert und verlangsamt das gemeinsame Arbeiten enorm. Was für den Einen absolutes Neuland darstellt, führt beim Anderen zur Langeweile; allerdings bringt dieser Unterschied in den Vorkenntnissen paradoxerweise wieder einen gewissen

[11] Über die Novissima auf dem Gebiet der Hard- und Softwarekomponenten berichtet das vom EDV-Zentrum herausgegebene Periodikum "Memory" (hier besonders 6, 1984 und 7,1985).

Ausgleich: so benötigt der Teilnehmer mit EDV-Erfahrung etwa die Hilfe eines EDV-unerfahrenen Kollegen, um die Quelle richtig lesen zu können.

Auch der Erwartungshorizont der potentiellen Besucher einer solchen Lehrveranstaltung ist schwer abzuschätzen. Während besonders Anfänger eine äußerst diffuse Vorstellung von EDV und Quantifizierung entwickeln, haben jene Kollegen, die in Projekten arbeiten oder den QUANTKURS absolviert haben, sehr konkrete und spezifische Erwartungen. Besonders erfreulich ist die Teilnahme von Studierenden und Graduierten verwandter Fachbereiche wie etwa der Volkskunde.

3.3. Studienplan

Noch heute ist die EDV "von einem manchmal lockenden, manchmal abschreckenden, stets aber sehr geheimnisvollen und manchmal geradezu magischen Nebel unvollständiger und ungenauer Information umgeben" [12]. Aus diesem Grund ist es wohl nicht immer möglich, in den Gremien die erhoffte Aufgeschlossenheit zu finden, wenn es um die Einrichtung von einschlägigen Lehrveranstaltungen und deren Anrechenbarkeit geht. Immerhin ist es in Graz gelungen, im Sommersemester 1985 drei solcher Kollegs anzubieten, die als Wahlfach für den Studienzweig Geschichte laut Studienplan angerechnet werden [13]. Obwohl von einer Verankerung im Studienplan noch keine Rede sein kann, scheint doch das Angebot an EDV-Lehrgängen für Historiker mittelfristig gesichert.

3.4. Methodische/konzeptuelle Vorüberlegungen

Es ist hier nicht der Ort, die Vor- und Nachteile datenbankorientierter Verfahren im allgemeinen oder eines DBMS im speziellen zu diskutieren. Es

---

[12] *Manfred Thaller:* Descriptor. Probleme der Entwicklung eines Programmsystems zur computerunterstützten Auswertung von Bildquellen, In: Europäische Sachkultur des Mittelalters, Wien, 1980. (= Veröffentlichungen des Instituts für mittelalterliche Realienkunde Österreichs.4.), S. 167.

[13] Karner: EDV und Geschichtswissenschaft I: Einführung in Probleme der Quantifizierung in der Geschichtswissenschaft (mit Einführung in einfache EDV-Programme), UE, 2st.

Kropac/Losch: EDV und Geschichtswissenschaft II: Von der Quelle zum Datensatz. VL, 2st.

Thaller: EDV und Geschichtswissenschaft III: Quellenkritik und Quellenkunde rechnergestützter historischer Forschung. VL, 2st.

soll nur kurz skizziert werden, warum diesen Verfahren ein so hoher Stellenwert eingeräumt wird. Wie bereits aus dem Titel der Lehrveranstaltung hervorgeht, soll nicht die Quantifizierung als Methode, sondern der EDV-Einsatz als Arbeitstechnik im Mittelpunkt der Vorlesung stehen. Selbstverständlich muß dabei auch das Hauptanwendungsgebiet, die quantifizierenden Methoden, ausführlich behandelt werden, allerdings sollen auch nützliche Anwendungen ohne Anspruch auf methodische Innovation - wie beispielsweise die maschinelle Verwaltung umfangreicher Bibliographien - Erwähnung finden. Die oft betonte Trennung und ebensooft stillschweigend vorgenommene Gleichsetzung von EDV-Einsatz und Quantifizierung erscheint als problematisch; es ist besonders bei Einführungen wichtig zu demonstrieren, daß eine Arbeitstechnik keine Methode impliziert und umgekehrt. Dieser Lehrveranstaltung liegt daher das ausdrückliche Bekenntnis zur Methodenpluralität zugrunde [14]: der Einsatz des Rechners soll die Anwendung mehrerer Methoden erleichtern oder erst ermöglichen, so beispielsweise die rechnergestützte, aber methodisch traditionelle Hermeneutik neben der Quantifizierung.

Daher sollte in einem ersten Teil die quellenorientierte Datenverarbeitung behandelt werden, was datenbankorientiertes Arbeiten sowohl mit fortlaufenden Texten als auch mit formalisierten Quellen(teilen) bedeutet. Diese Phase soll einerseits die hermeneutische Interpretation unterstützen, andererseits als explorativer Vorschaltprozeß für die Kodierung dienen, um die Quelle besser "kennenzulernen", und das Entwickeln statistisch zu testender Hypothesen zu begünstigen [15]. Die quellenorientierte Datenverarbeitung bildet somit Gegenstück und/oder Vorstufe zur methodenorientierten, mit Hilfe derer die erstellten Hypothesen durch diverse statistische Verfahren getestet werden. Letztere soll demnach in einer zweiten Stufe behandelt werden, besonders, da ein gewisses statistisches Grundwissen unbedingte Voraussetzung ist.

---

[14] Zu diesen Problemen neuerdings *Konrad H. Jarausch, Gerhard Arminger, Manfred Thaller:* Quantitative Methoden in der Geschichtswissenschaft. Eine Einführung in die Forschung, Datenverarbeitung und Statistik, Darmstadt, 1985, bes. S.1-6, 203-206. - Vgl. auch *Manfred Thaller:* Zur Formalisierbarkeit hermeneutischen Verstehens in der Historie, In: Mentalitäten und Lebensverhältnisse, Rudolf Vierhaus zum 60. Geburtstag, Göttingen, 1982, S. 439-454.

[15] Vgl. *Manfred Thaller:* Datenbankorientierte Verfahren in der Geschichtswissenschaft, In: Bericht über den 15. österreichischen Historikertag in Salzburg, Salzburg, 1984. (= Veröffentlichungen des Verbandes Österreichischer Geschichtsvereine. 23.), S.433 f.

4. Alternativen zur Verwirklichung

Von J.Ehmer wurden drei grundlegende Alternativen genannt [16], die bei der Konzeption einschlägiger Lehrveranstaltungen zu berücksichtigen sind:

Fixes Format - freies Format
Kodierte Eingabe - quellennahe Eingabe
Datenfileorientierung - Datenbankorientierung

Diese Alternativen wurden bereits andernorts ausführlich diskutiert [17]. Hier fiel die Entscheidung für eine quellennahe Eingabe im freien Format, die Arbeitsweise soll datenbankorientiert erfolgen. Neben den obengenannten Vorgaben sprachen in der Hauptsache zwei Gründe dafür:

Die Teilnehmer sollen nach Absolvierung der Vorlesung in der Lage sein, beliebige Quellen auch komplexerer Struktur zu formalisieren und maschinenlesbar zu machen, sowie auch mit fortlaufenden Texten arbeiten zu können.

Das derzeit einzige Programmpaket, welches Forderungen dieser Art erfüllen kann, ohne zur Programmierung auf der Ebene Höherer Programmiersprachen zu zwingen, ist das von M.Thaller in Göttingen entwickelte System namens CLIO [18]. Aufgrund diverser Projekte konnte CLIO in Graz implementiert werden und ist dankenswerterweise auch für Dritte benutzbar.

Neben diesen drei Wahlmöglichkeiten ergeben sich bei der Konzeption dieser Vorlesung weitere Optionen, teils aus der speziellen lokalen Situation, teils aus grundlegenden Überlegungen resultierend.

Als erstes stellte sich die Frage, welcher Rechner nun tatsächlich zu benützen sei. Die in Punkt 3.1 angesprochene Verbindung der beiden Grazer

---

[16] *Josef Ehmer:* Erfahrungen mit EDV-gestützten Projekten als Grundlage für die Konzeption einer Lehrveranstaltung "Einführung in die EDV für Historiker", In: Bericht über den 16. Öst. Historikertag in Krems / Donau (= Veröff. d. Verbandes Öst. Geschichtsvereine 25), o.O., 1985, S. 676-683.

[17] *Josef Ehmer:* Erfahrungen mit EDV-gestützten Projekten als Grundlage für die Konzeption einer Lehrveranstaltung "Einführung in die EDV für Historiker", In: Bericht über den 16. Öst. Historikertag in Krems / Donau (= Veröff. d. Verbandes Öst. Geschichtsvereine 25), o.O., 1985, S. 680f. - Vgl. auch *Donald E. Ginter, Peter Grogono, Frederick A. Bode:* A Review of Optimal Input Methods: Fixed Field, Free field, and the Edited Text, in: Historical Methods Newsletter 10, 1977, S. 166-176.

[18] Vgl. *Manfred Thaller:* CLIO. Einführung und Systemüberblick, Göttingen, 1983 (Preprint). Neben dieser Einführung veröffentlichte M.Thaller weitere, hier nicht angeführte Dokumentationen sowie eine größere Anzahl von Studien zu Konzeption und technischer Realisierung von CLIO.

Rechner macht es möglich, die VAX wegen ihrer moderneren I/O-Komponenten als "Arbeitsplatzrechner" und die UNIVAC als "Hintergrundsystem" zu benutzen. Für den Benutzer ergeben sich daraus nur zwei wesentliche Unterschiede: Batch-Jobs werden statt mit "@START" mit dem Befehl "REMOTE" gestartet, Outputs liegen als Printfile (Filetyp .lis) in der Default-Directory der VAX und können nach Bedarf ausgedruckt werden. Die HELP-Facilities in Betriebssystem und Editor der VAX und der im Vergleich mit der UNIVAC-Steuersprache wesentlich einfachere Aufbau der Digital Command Language kommen den Anfängern besonders entgegen. Die der UNIVAC zu übergebenden Batch-Run-Karten werden nur vorgegeben, die UNIVAC also im Prinzip als "Blackbox" behandelt. Das bedeutet, daß die angeführte Alternative de facto durch die Benützung beider Rechner umgangen wird.

Ähnlich verhält es sich mit der Frage, ob - speziell für statistische Verfahren - ein bestehender Datensatz zu verwenden oder ein neuer anzulegen sei [19]. Das Konzept M. Thallers [20], bei mehreren QUANTKURSen bestens bewährt und hier übernommen, behandelt die Errichtung einer SPSS-Systemdatei erst *nach* der Systemfilebeschreibung und einfacheren statistischen Auswertungen. Dieses Vorgehen bedingt einen bestehenden, bereits getesteten Datensatz, der Variable verschiedener Meßniveaus enthält und die Anwendung verschiedener statistischer Routinen erlaubt; ein solcher stand bei der Planung der Lehrveranstaltung bereits zur Verfügung [21]. Dagegen spricht der Umstand, daß die Studierendenden oft langwierigen Weg der Datenaufbereitung bis zum statistisch auswertbaren File kennenlernen und an allen Schritten beteiligt werden sollen. Auch die in Punkt 3.4 angestellten Vorüberlegungen scheinen dem Thallerschen Konzept zu widersprechen. Einen Ausweg aus diesem Dilemma würde die Verwendung derselben Quelle (oder Quellengattung) für die quellenorientierte Datenverarbeitung bringen, wie sie für die Erstellung des bestehenden (statistisch auswertungsbereiten) Datensatzes verwendet wurde. Um sinnlose Doppelarbeit zu vermeiden, könnte u.U. ein anderer Zeitraum gewählt werden, immer unter der Voraussetzung, daß es sich um identische Strukturen handelt. Daraus ergibt sich auch eine Entscheidung in diesem Bereich - es wäre günstig, beide Wege gleichzeitig zu beschrei-

[19] Dazu vgl. *Thomas Schuler:* Quantifizierung und Geschichtsstudium, In: HSR 17, 1981,S. 61f.

[20] *Manfred Thaller:* Numerische Datenverarbeitung für Historiker. Eine praxisorientierte Einführung in die quantitative Arbeitsmethode und in SPSS, Wien-Köln, 1982 (= Materialien zur Historischen Sozialwissenschaft. 1.).

[21] Herrn Dr. *Gerhard Jaritz*, Institut für mittelalterliche alienkunde Österreichs, sei für die Weitergabe dieses Datensatzes herzlich gedankt.

ten. Weiters wurde es notwendig zu entscheiden, ob quellenorientiert oder methodenorientiert, ob streng quantifizierend oder methodisch eher neutral vorzugehen sei. Die Beantwortung dieser Fragen ergibt sich zwingend aus Punkt 3.4: es soll sowohl quellen- als auch methodenorientiert und nach mindestens zwei geschichtswissenschaftlichen Methoden gearbeitet werden, ohne einseitige Wertung.

5. Inhalt der Lehrveranstaltung

5.1. Übersicht

Einführungsphase:

- *Grundwissen über den Computer (Hard/Software) und dessen Peripherie*
- *Anwendungsmöglichkeiten der EDV in der Geschichtswissenschaft*
- *Probleme datentechnischer Verfahren in der Historie - Theorien, Methoden, Konzepte*

1. inhaltliche Phase - quellenorientierte Datenverarbeitung

- *Quelle und Quellenkritik; Informationsgehalt und -hierarchie der Quelle*
- *CLIO-Format: Eingabekonventionen, Datenstruktur, Plausibilität und Praktikabilität*
- *Erfassung der Quelle am Bildschirmterminal*
- *CLIO-Systemfilegenerierung (Befehle und deren Elementarmodi); Auswertung des "Prüflaufprotokolls"; wenn notwendig, dann Korrektur der Rohdaten und erneute Systemfilegenerierung*
- *Information Retrieval: Formulierung inhaltlicher Fragen, Übersetzung in die CLIO-Retrieval-Sprachkomponenten;*
- *"Linguistische Datenverarbeitung" im weitesten Sinne: Erstellung von Wortzählungen und eines Wörterbuchs*

2. inhaltliche Phase - methodenorientierte Datenverarbeitung (Quantifizierung)

- *Einführung in die Quantifizierung (Wiederholung)*
- *Statistisches Grundwissen (z.B. Meßniveau, Stichprobe / Grundgesamtheit etc.)*
- *Vorstellung des Programmpakets SPSS*
- *SPSS-Systemfilebeschreibung*
- *einfache Auswertungen aus dem Gebiet der deskriptiven Statistik*
- *maschinelles Kodieren mit Hilfe der entsprechenden CLIO-Routinen*
- *SPSS-Systemfilegenerierung.*

5.2. Praktische Durchführung

Beide inhaltlichen Phasen beschäftigen sich mit derselben Quelle, einem Rechnungsbuch der Chorfrauen von Klosterneuburg, 1495-1508 [22]. Diese Quelle eignet sich durch ihre leicht leserliche Schrift und ihre klare Struktur (Abb. 1) besonders, um dieses Konzept für die eingangs erwähnte heterogen zusammengesetzte Zielgruppe zu realisieren. Die Teilnehmer sollen lernen, den oft langwierigen Weg von der Originalquelle zum (analytischen) Datensatz mit allen seinen Problemen und Tücken selbst zu beschreiten.

In der ersten Phase werden die Jahre 1495/96 des Rechnungsbuchs als Xerokopie vorgelegt und quellenkritisch betrachtet. In der Folge wird die Quelle in zweifacher Hinsicht zur Eingabe aufbereitet: einmal formalisiert in einer hierarchischen Struktur für Retrieval-Programme und maschinelle Kodierung, zum anderen als fortlaufender Text gleichsam als Vorbereitung zur Edition, um "Textanalysen" in Form von Wortzählungen, KWICs und KWOCs vornehmen zu können. Bei der hierarchischen, datenbankorientierten Behandlung der Quelle gilt es, deren Informationsgehalt und -struktur logisch und möglichst unverkürzt abzubilden. Das bedeutet eine Differenzierung der einzelnen Informationen nach mehreren Richtungen. Es ist zu prüfen, ob die Information unmittelbar aus der Quelle übernommen werden kann wie hier etwa der Wochentag oder die Angabe des Preises, oder ob sie erst mittelbar aus anderen Angaben zu erschließen ist, wie beispielsweise das Kalenderdatum oder die Fastenregelung. Weiters bedarf es einer Zuweisung der Informationen an die verschiedenen hierarchischen Ebenen, wobei anfangs die Struktur der Quelle möglichst direkt in jene des Datenmodells übernommen werden sollte. Daraus ergibt sich für die vorgestellte Quelle die Struktur von Schema 1 (in der Notation von CLIO).

Die im Quellenbeispiel (Abb. 1) gezeigte Woche würde unter Zugrundelegung des Datenmodells aus Schema 1 aussehen wie Schema 2.

Die Bearbeitung der Quelle als fortlaufender Text erscheint vom Datenmodell her wesentlich einfacher, muß doch lediglich der Text erfaßt werden und der Ort in der Quelle, wo jener zu finden ist (auf die seltenen Textvarianten und alternative Betrachtungsweisen wurde vorerst verzichtet). Wir erhalten dadurch Schema 3.

Dieselbe Woche, die oben bereits formalisiert dargestellt wurde, müßte in der Eingabe dann wie Schema 4 aussehen.

---

[22] Niederösterreichisches Landesarchiv, HS. 424. - Dazu vgl. *Gerhard Jaritz:* Aderlaß und Schröpfen im Chorfrauenstift Klosterneuburg (1445 -1533), In: Jahrbuch des Stiftes Klosterneuburg N.F. 9, 1975, S.78 f.

```
D(okument)$QUELLE/ZITAT
⇑
⇑≪ W(ochentag)$NAME/DATUM/FASTEINREGELUNG/
⇑     ⇑ FESTREGELUNG
⇑     ⇑
⇑     ⇑≪ E(intragung)$NUMMER/ANZAHL%EINHEIT/OBJEKT/
⇑          RECHTSFORM/PREIS/ZIELGRUPPE/HERKUNFT
⇑
⇑≪ S(umme)$GESAMTAUSGABEN
```

Schema 1

Durch einfache Manipulationen läßt sich in der Folge aus einer Dokumentendatenbank eine Wortdatenbank errichten.

Parallel zu Modellbildung und Datenerfassung ist es bereits notwendig, das Wie und Warum der einzelnen Entscheidungen festzuhalten und zu dokumentieren; zum einen muß jeder Mitarbeiter vom anderen wissen, wie er zu seinen Eingabedaten gekommen ist, zum anderen ist eine solche Dokumentation condicio sine qua non für eine spätere Verwendung des Datensatzes.

Aufgrund der geschilderten Situation am lokalen Rechenzentrum sind die Studierenden angehalten, nach einem Einführungsabend an den Terminals alle Manipulationen in Zusammenhang mit der Datenerfassung selbständig durchzuführen. Lediglich die Kollationierung der Quelle und damit verbunden die Vorgabe der Korrekturen werden von den Leitern der Lehrveranstaltung vorgenommen. Dieser Zwang zur selbständigen Arbeit erfordert natürlich eine ständige begleitende Betreuung der Teilnehmer, die ihre Probleme hauptsächlich in den Lehrveranstaltungsblöcken, aber auch außerhalb der regulären Vorlesungszeiten mit den Vortragenden und untereinander besprechen sollen. Auf diese Weise wird der Notwendigkeit entsprochen, möglichst wenig Gruppenarbeit am Rechenzentrum zu absolvieren. Die Erarbeitung der Aufgaben und Programme, sowie deren Analyse und Interpretation erfolgt ausschließlich am Institut; lediglich drei Arbeitsabende am Rechenzentrum sollten ausreichen, den Umgang mit Hardware und Software zu erklären bzw. zu vertiefen. Das bedeutet zwar die Trennung von theoretischer und praktischer Arbeit, bereitet aber andererseits auf den Alltag der EDV-Anwender vor, auftretende Probleme ohne unmittelbare Beratung aus eigener Kraft lösen zu müssen. Um die Eingabearbeit und in der Folge auch

```
D$RECHNUNGSBUCH KLOSTERNEUBURG 1495-96/NÖLA HS. 424 FOL. 18'
W$SONNTAG%p(ri)ma po(st) Trini(ta)te/1496.6.5/N/F
E$1/14/HÜHNER/K/3S. 6D.
E$2/1/SCHWEINEBRATEN/K/70D.
E$3//SEMMELN/K/23D./GÄSTE
E$4/3/GEKRÖSE/K/12D.
E$5/0.5/KALB/K/3S. 20D.
W$MONTAG/1496.6.6/N/W
E$1//SCHWEINEFLEISCH/K/32D.
E$2/2/EUTER/K/3S. 20D.
E$3/2/SPANFERKEL/K/32D.
E$4/2/HÜHNER/K/15D./KRANKE
W$DIENSTAG/1496.6.7/N/W
E$1//SELCHFLEISCH/E
E$2/0.5/SCHAFBAUCH/K/36D.
E$3/1%VIERTEL/KALBFLEISCH/K/60D.
W$MITTWOCH/1496.6.8/N/W
E$1//MILCH/K/12D.
E$2/1%ACHTEL/MILCHRAHM/K/32D.
E$3//KREBSE/K/12D.
E$4//EIER/K/53D.
W$DONNERSTAG/1496.6.9/N/W
E$1/0.5/KASTRAUN/K/38D.
E$2/7/HÜHNER/G
E$3/1/KALBSBRUST/K/66D.!
E$3/1%VIERTEL/KALBFLEISCH/K/66D.!
W$FREITAG/1496.6.10/N/W
E$1//KREBSE/K/68D.
E$2//MILCH/K/12D.
E$3//ZWIEBEL/K/8D.
W$SAMSTAG/1496.6.11/N/W
E$1//EIER/K/77D.
E$2//MILCHRAHM/K/46D.
S$4TL. 36D.
```

Schema 2

```
D(okument)$FOLIERUNG
↑
↑≪$ZEILE/TEXT
```
Schema 3

die Erstellung von Befehls- und Runkarten zu erleichtern, arbeiten jeweils zwei Studierende an einer Aufgabenstellung.

In der zweiten Phase wird das Jahr 1497 des obengenannten Rechnungsbuchs bearbeitet, das als fertiger - am QUANTKURS getesteter und bestens bewährter - Datensatz vorliegt. Da der Inhalt der Quelle innerhalb der beiden Jahre kaum differiert, kann durch die vorausgehende Bearbeitung der Jahre 1495/96 in der ersten Phase ein verbessertes inhaltliches Verständnis jener statistischen Verfahren, denen der Datensatz unterworfen wird, sowie der folgenden Kodierungsarbeit erreicht werden. Die maschinelle Kodierung der eingegebenen Daten, die Übersetzung der Kodes und die Errichtung eines SPSS-Systemfiles sollen den Grundkurs abschließen.

6. Zukunftsperspektiven

6.1. Ein Grazer Curriculum?

Wie eben durch die Bezeichnung Grundkurs angeklungen ist, soll auf Wunsch der Studierenden eine Fortsetzung angeboten werden, die allerdings organisatorisch und konzeptuell wesentliche Abweichungen vom bisher erörterten Modell enthalten muß. Eine weiterführende Lehrveranstaltung sollte verstärkt inhaltlich orientiert sein, d.h. vermehrt auf das Entwickeln von Hypothesen und Fragestellungen abzielen und die im Grundkurs erlernten Routinen anhand inhaltlicher Schwerpunkte vertiefen. Weiters sollte es den Studierenden ermöglicht werden, eigene Quellen, Fragestellungen und Probleme einzubringen und zu diskutieren. Die Organisation als Blockvorlesung soll auch in der Fortsetzung aufrecht bleiben.

Die in Punkt 3.3 vorgestellte Abfolge der Lehrveranstaltungen soll in der nächsten Zeit beibehalten werden. Stufe I ermöglicht ein erstes Hineintasten in die EDV und Quantifizierung, ist für Anfänger konzipiert und wird jedes Semester gelesen. Die hier diskutierte Stufe II soll sich über zwei Semester erstrecken und bietet zwar auch die Einstiegsmöglichkeit für Anfänger, geht aber wesentlich intensiver vor und benötigt einen nicht geringen Arbeitsaufwand. Stufe III vermittelt als reine Vorlesung theoretisches Wissen auf hohem Niveau; je nach Verfügbarkeit geeigneter Lektoren sollte diese Vorlesung doch

```
D$18'
$1/Do(m)i(ni)ca p(ri)ma po(st) Trini(ta)te
$2/Item umb 14 huendel 3 ß 6 d
$3/Item umb 1 sweinenpraten 70 d
$4/Item umb seml den gesten 23 d
$5/Item umb 3 kröß 12 d
$6/Item umb 0.5 kalb 3 ß 20 d
$7/Mantag
$8/Item umb sweinesfleisch 32 d
$9/Item umb 2 au(o)tter 40 d
$10/Item umb 2 spensaw 60 d
$11/Item umb 2 huendel den kranken 15 d
$12/Eritag
$13/Item geselichtsfleisch hab ich gehabt
$14/Item umb 0.5 schofpauch 36 d
$15/Item umb 1 viertl kalbfleisch 60 d
$16/Mitichen
$17/Item umb milich 12 d
$18/Item umb 1 achtel milichra(e)n 32 d
$19/Item umb kreussn 48 d
$20/Item umb a:yr 53 d
$21/Phintztag
$22/Item umb 1 kastraun 38 d
$23/Item 7 huendel hat man uns geschenkt
$24/Item umb 1 kelbreineprust und ain viertel kalbfleisch 66 d
$25/Freytag
$26/Item umb kreussn 68 d
$27/Item umb milich 12 d
$28/Item umb zwival 8 d
$29/Sambstag
$30/Item umb a:yr 77 d
$31/Item umb milichräm 46 d
$32/Summa facit 4 t 36 d
```

Schema 4

zumindest jedes zweite oder dritte Semester mit wechselnden Schwerpunkten abgehalten werden. Ob das derzeitige Angebot den Anforderungen, Historiker in "neuen" Methoden und Arbeitstechniken grundlegend auszubilden,

gerecht werden und sich dadurch auch in Graz etablieren kann, wird erst die weitere Zukunft zeigen. Mit einer Integration in den Studienplan kann kurzfristig jedoch kaum gerechnet werden; die Ausbildung in dieser Richtung beruht auf Freiwilligkeit.

6.2. Lehrpaket

In der geschilderten Weise soll sukzessive das gesamte Rechnungsbuch aufgearbeitet und als überschaubare "Mini-Datenbank" für den Lehrbetrieb adaptiert werden. In ein solcherart erstelltes Lehrpaket müßte neben den zur Diskussion gestellten Intentionen der Vortragenden aber auch die Rezeption der Inhalte durch die Studierenden eingehen. Dies soll die Protokollierung jedes Vorlesungsblockes durch dessen Teilnehmer und intensive Diskussion innerhalb der Gruppe gewährleisten. Die auf dem Rechner erstellten Protokolle werden nach Abschluß des Kurses von den Vortragenden redigiert und sollen zusammen mit dem Vorbereitungsmanuskript unter Einarbeitung der wichtigsten einführenden Literatur ein Skriptum bilden, das den Studenten in der Folge zur Verfügung gestellt werden soll. Dieses Skriptum mit der edierten Quelle im Anhang muß auch Beispiele für Befehlskarten und Outputs der speziellen Programme enthalten. Daraus könnte sich nach weiterer Diskussion und den sich daran knüpfenden Umstellungen ein Lehrbuch entwickeln, das seinerseits wieder die Basis eines Lehrpakets bildet. Ein vollständiges Lehrpaket müßte demnach folgende Komponenten enthalten:

1. Lehrbuch mit eingearbeiteten Beispielen und Literaturangaben; edierte Originalquelle.
2. Kommentar zum Lehrbuch für Vortragende, in dem das grundlegende Konzept transparent und ev. Vorschläge für die Realisation gemacht werden,
3. Magnetband mit Rohdaten (Datensatz), Systemdateien (Datenbanken), Befehlskarten und Outputs analog zu den im Lehrbuch enthaltenen Beispielen, sowie der maschinell edierten Quelle.

Die Konzipierung und Einrichtung eines Lehrpakets für andere Großrechner als die UNIVAC 1100/xx oder Mikrocomputer wird vor allem mit den zukünftigen Strategien von CLIO abhängen; eine portable Version des Programmsystems ist in Entwicklung.

Abschließend sei angemerkt, daß die unterbreiteten Vorschläge keinesfalls für sich in Anspruch nehmen wollen, endgültige Modelle und programmatische Zielsetzungen zu präsentieren. Eine Möglichkeit von sicherlich vielen soll hier zur Diskussion gestellt werden, auf welche Weise Benützung und

Aufbau von Datenbanken die universitäre Lehre, aber auch die Forschung erleichtern und fördern könnten. Alle Überlegungen zur Lehrveranstaltung selbst gelten natürlich auch für ein eventuell daraus resultierendes Lehrpaket; Sinnhaftigkeit und Verbesserungsmöglichkeiten am Konzept und dessen Realisation gilt es in der Folge zu prüfen.

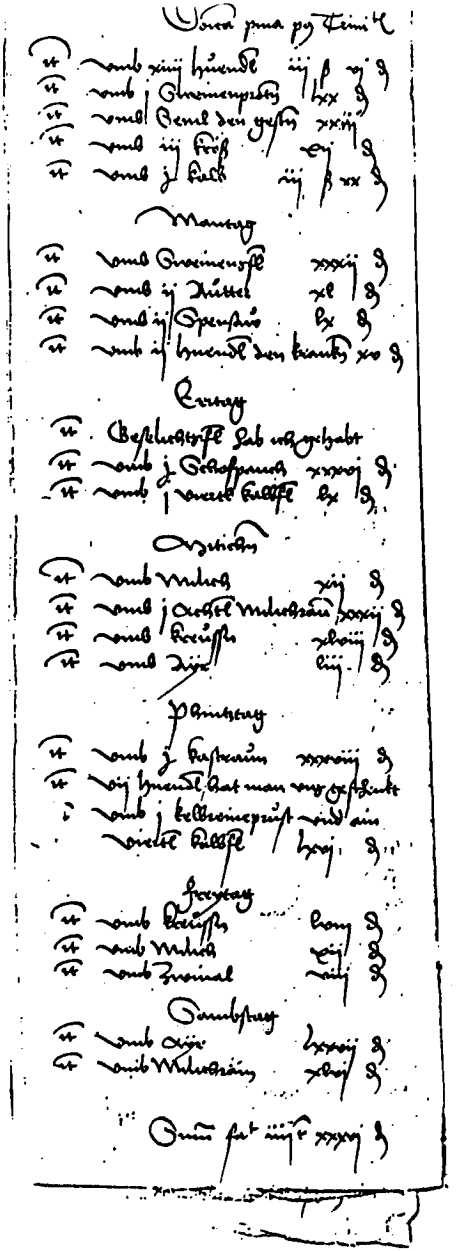

Abbildung 1