

SGML-ähnliche Textauszeichnung in der TUSTEP-Umgebung

Bader, Winfried

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

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

Empfohlene Zitierung / Suggested Citation:

Bader, W. (1993). SGML-ähnliche Textauszeichnung in der TUSTEP-Umgebung. *Historical Social Research*, 18(1), 92-112. <https://doi.org/10.12759/hsr.18.1993.1.92-112>

Nutzungsbedingungen:

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

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:
<https://creativecommons.org/licenses/by/4.0>

SGML-ähnliche Textauszeichnung in der TUSTEP-Umgebung

*Winfried Bader**

Abstract: Working with computer becomes real data processing in opposite to simply working with a typewriter, when the user does not face only the printout but faces the data to be processed for several purposes: quantitative analysis, grammatical analysis, statistical approaches, retrieval for information. The success of this purposes needs not only powerful tools but a good method of data management. The paper gives a practical introduction into the generic text markup. Generic text markup means here markup with regard to the content of a text and not only to the typography in opposite to the sepcific markup for specific formatters. In a first part the paper describes the idea of generic markup. It stresses the advantages of generic markup in the daily work. The second part shows the four fundamental steps in working with generic markup. This steps are explained with regard to SGML. The third part gives examples. They present TUSTEP as a tool for working with SGML-like tags. TUSTEP-programs are demonstrated for parsing texts with generic markup. This examples include not only different layout programs, but also analytical work with tagged texts.

Die Arbeit mit Texten im Computer hebt sich dann von der elektronischen Schreibmaschine ab und beginnt zur Datenverarbeitung zu werden, wenn im Mittelpunkt des Interesses nicht mehr allein der Ausdruck auf Papier steht, sondern auch der Zugriff auf die Textdaten für Fragen der Analyse, Statistik und inhaltlichen Auswertung bei der Erfassung und Speicherung des Textes berücksichtigt wird. Eine erfolgreiche Weiterverarbeitung von Textdaten setzt neben geeigneten Werkzeugen eine geeignete Methode der Datenerfassung voraus. Oft wird das Bedürfnis nach Weiterverarbeitung erst spät erkannt, dem bei

* Address all communications to Winfried Bader, Universität Tübingen - Zentrum für Datenverarbeitung, Brunnenstraße 27, D-7400 Tübingen; e-mail: bader@mailserv.zdv.uni-tuebingen.de.

einer Textauszeichnung unter ausschließlicher Rücksicht auf die eine gedruckte Form meist nicht mehr entsprochen werden kann.

Der vorliegende Beitrag ist abseits langwieriger Handbücher eine praktische Einführung in die Idee und Anwendung der sachlich orientierten Textauszeichnung, die die Grundlage für eine Textdatenverarbeitung bieten. Sachlich orientierte Textauszeichnung (*generic markup*) steht dabei als Oppositionsbegriff zu den auf ein spezielles Formatier-Programm bezogenen spezifischen Steueranweisungen (*specific markup*). Mit diesem Verfahren entsteht bei den Daten ein fließender Übergang von Textlayout, Volltextzugriff, strukturierte Daten bis hin zu einer Datenbank. Die Beispiele im dritten Teil des Artikels verdeutlichen das Spektrum und ermuntern auch den Laien zur Anwendung von sachlichen Textauszeichnungen. Ausführlich wird das Programmpaket TUSTEP (1) als Werkzeug für den Umgang mit dieser Art von Textdaten für Druckausgabe und Weiterverarbeitung dargestellt. Die Idee der sachlichen Textauszeichnung ist auch in SGML (2) und den Vorschlägen der TEI (Text Encoding Initiative) verwirklicht. Dadurch ist das Verfahren mit Hilfe der vorgeschlagene Richtlinien für den internationalen Datenaustausch geeignet.

1. Vorteile der sachlich orientierten Textauszeichnungen

Ausgangspunkt der folgenden Überlegungen sind Texte, deren primärer Zweck es ist, ausgedruckt zu werden, ein Gebiet, auf dem die elektronische Schreibmaschine eigentlich genügen würde. Um einen im Computer gespeicherten Text in schöner Form auf Papier zu drucken, ist es notwendig, im Text Auszeichnungen anzubringen, die entweder dem anschließenden Formatier-Programm sagen, welche Typographie an der betreffenden Stelle gewünscht ist (so bei *batch formatter* wie TUSTEP-FORMATIERE, TUSTEP-SATZ, TeX oder dem UNIX-Befehl *nroff*), oder die sofort auf dem Bildschirm das gewünschte Ergebnis liefern (so bei WYSIWYG-Systemen - *interactive formatter* - wie WORD und WordPerfect). In beiden Fällen ist der elektronisch gespeicherte Text das Gegenstück der Papierversion. Beide enthalten dieselben Informationen.

Demgegenüber verleiht die Verwendung von sachlich orientierten Auszeichnungen dem gespeicherten Text eine neue Dimension. Durch freigewählte eindeutige Zeichenfolgen wird der Text ausschließlich nach sachlichen Gesichtspunkten strukturiert. Damit ist er mehr als nur das Gegenstück der gedruckten Form: die sachlichen Strukturen können für unterschiedliche Formen der elektronische Weiterverarbeitung genutzt werden (Indexerstellung, Retrieval, Ablegen in Datenbanken, Plausibilitätstests etc.). Die Umsetzung der Struktur in eine typographische Form, die durch ein Programm geschieht, ist nur eine mögliche Weiterverarbeitung unter anderen.

Die Verwendung von sachlich orientierten Auszeichnungen (*generic markup*) hat gegenüber der Verwendung von spezifischen Steueranweisungen für das jeweilige Formatier-Programm (*specific markup, formatting instructions*) folgende Vorteile:

- Die sachliche Struktur und die äußere gedruckte Form sind getrennt. Man kann jedes der beiden Gebiete dem Fachmann überlassen, die sachliche Strukturierung dem wissenschaftlichen Autor oder Autorin, das Layout dem Setzer oder Setzerin. Beide Arbeitsschritte sind zeitlich voneinander getrennt, so daß man nicht schon bei der ersten Texterfassung an die typographische Gestaltung denken muß.
- Bei der Texterfassung kann man sich ganz auf den Inhalt und seine Struktur konzentrieren.
- Sachliche Auszeichnung ist in der Regel differenzierter als typographische Auszeichnung. Man kann damit in den Daten auch typographisch gleich behandelte Elemente unterscheiden, z. B. die einzelnen Bestandteile in einem bibliographischen Eintrag, die - bis auf den Autorennamen - in der Regel alle in gleicher Schrift gedruckt werden.
- Sachlich Gleiches wird immer gleich ausgezeichnet und kann damit bei der Weiterverarbeitung immer als solches erkannt werden, um es jeweils typographisch gleich auszugeben oder in anderer Weise weiterzuverarbeiten. Spezifische Steueranweisungen lassen dagegen selbst bei typographischer Gleichheit mehrere Möglichkeiten der Auszeichnung zu.
- Sachlich verschieden Ausgezeichnetes (differenziert für die Weiterverarbeitung) kann dennoch typographisch gleich behandelt werden.
- Die Hierarchie der Überschriften (wichtig für die inhaltliche Struktur des Textes) läßt sich leicht einhalten und für die automatische Erstellung eines Inhaltsverzeichnisses eindeutig auffinden.
- Sachliche Auszeichnungen sind oft kürzer als notwendige Kombinationen von Steueranweisungen.
- Sachliche Auszeichnungen sind kompatibel. Sie sind nicht auf ein Ausgabeprogramm festgelegt. Stehen entsprechende Konvertierungs-Programme zur Verfügung, kann dieselbe Datei z. B. mit TeX, WORD, WordPerfect, TUSTEP-FORMATIERE und TUSTEP-SATZ ausgegeben werden.
- Sachliche Auszeichnungen haben auch bei Verwendung nur eines Formatier-Programms einen Vorteil: man muß nicht schon zu Beginn der Arbeit den typographischen Stil festlegen.
- Kommentare, Bearbeitungshinweise und Anmerkungen unterschiedlichster Art kann man in den Text mit aufnehmen und damit weiterarbeiten, ohne sich gleich Gedanken machen zu müssen, ob diese Informationen auf Papier gedruckt werden sollen oder nicht.

Sachliche Auszeichnungen sind nicht immer notwendig, auch die elektronische Schreibmaschine hat ihre Berechtigung. Bei kurzen Texten, Texten von geringer Lebensdauer und begrenztem Zweck (wie z. B. ein Brief, eine Ankündi-

gung, ein Thesenpapier, ein Handout) kann es eine Arbeitersparnis bedeuten, direkt typographische Auszeichnungen zu verwenden. Sachliche Auszeichnungen lohnen sich aber, wenn eine Struktur für die Weiterverarbeitung gebraucht wird (Serienbriefe) oder mit anderem Inhalt wiederverwendet wird (wiederkehrende Ankündigungen). Betrachtet man die oben aufgezählten Vorteile, so sieht man daran, daß das Verfahren der sachlich orientierten Textdatenauszeichnung auch und vorzüglich auf Textdaten anzuwenden ist, bei denen das Ausdrucken und das Layout nur eine untergeordnete, wenn auch vorhandene Rolle spielen, und der Hauptzweck das strukturierte Abspeichern von Informationen ist, die später ausgewertet werden sollen.

Unbedingt verwendet werden sollten sachliche Auszeichnungen für folgende Zwecke:

- für Texte, die in verschiedenen Formatier-Systemen verwendet werden.
- für Texte, die ganz oder teilweise in Datenbanken eingefügt werden.
- für Texte, die voraussichtlich mehrere Revisionen und Überarbeitung erfahren werden.
- für Texte, die verschiedenen - auch noch unvorhergesehenen - Zwecken dienen werden.
- für Texte, die durch ein Team von Mitarbeiterinnen erfaßt und bearbeitet werden.
- für Texte, die über eine lange Zeit verwendet werden. **Die sachliche Auszeichnungen garantieren nicht nur ein Überleben bei neuen Rechnern und neuen Versionen des Ausgabesystems, sondern schaffen auch Kontinuität beim Wechsel von Mitarbeiterinnen, denen die inhaltliche Struktur leichter vermittelt werden kann als bestimmte typographische Erscheinungen.**
- für Texte, die sehr lang sind, so daß die innere Konsistenz nur über sachliche Auszeichnungen gewahrt werden kann.
- für Texte, die streng strukturiert sind, wie z. B. **der Artikel eines Wörterbuchs.**
- für Texte, die mit zusätzlichen Informationen (Registerkennzeichnungen, Grammatikanalysen, Verweisen, Übersetzungen etc.) angereichert werden.
- für Textdaten, die primär zum Abspeichern von Informationen für späteres Retrieval und erst sekundär zum Ausdrucken dienen.

2. Vier Schritte zur praktischen Anwendung von sachlich orientierten Textauszeichnungen

Die Verwendung von freigewählten, projektspezifischen sachlichen Textauszeichnungen, erfordert vier grundlegende Arbeitsschritte und Programmelemente. Dies gilt sowohl für die Realisierung der Textauszeichnung nach den SGML-Richtlinien wie auch für die Verwendung von sachlichen Textauszeichnungen in der TUSTEP-Umgebung wie sie unten in den Beispielen vorgestellt werden.

2.1 Auflisten der benötigten Auszeichnungen

Der erste Schritt zur Verwendung sachlicher Auszeichnungen ist eine Analyse der Struktur des zu erfassenden Textes, aus der eine Zusammenstellung der benötigten Auszeichnungen resultiert: Welche Hierarchie von Überschriften gibt es? Werden verschiedene Stufen von Abschnitten benötigt? Welche besondere Informationen müssen hinsichtlich Registererstellung und Weiterverarbeitung (u. U. mit typographischer Differenzierung) ausgezeichnet werden (Autoren, Sachwörter, Schlüsselwörter, Literaturangaben etc.)? Welche verschiedenen Textteile gibt es (Zitate, Nebenbemerkungen, Merksätze, Zusammenfassung, Gedichtzeilen etc.)? Welche Art von Anmerkungen gibt es (übliche wissenschaftliche Fußnoten, Sacherklärungen, textkritische Apparate in verschiedenen Variationen)? Werden Kürzel für häufig vorkommende Ausdrücke, Sonderzeichen und Querverweise gebraucht? Sind Kommentare vorzusehen?

Die aufgrund dieser Analyse benötigten Auszeichnungen werden in einer Liste (Datei) zusammengestellt und ihre sachliche Bedeutung festgehalten. Gegebenenfalls werden Bedingungen angegeben, wo eine solche Auszeichnung erlaubt ist, welche Kombinationen notwendig, welche verboten sind. Z. B. muß jede Anfangsauszeichnung durch eine entsprechende Endauszeichnung abgeschlossen sein. Schachtelungen von Auszeichnungen sind nur bedingt erlaubt: ein Autorennamen innerhalb eines Sachworts ist unmöglich, eine textkritische Anmerkung gehört nicht in eine Sacherklärung, Kommentare dürfen keinerlei Auszeichnungen enthalten etc. Umgekehrt kann z. B. eine Gedichtzeile nur innerhalb eines Literaturzitats vorkommen. (3) Für die Hierarchie der Überschriften kann eine feste Abfolge vereinbart werden, z. B. daß der Überschrift einer bestimmten Stufe nur eine gleichwertige, eine untergeordnete oder eine unmittelbar übergeordnete vorausgehen kann (also nicht Überschrift 1.1.1 nach 1. ohne 7.7).

Diese Liste wird zunächst sicherlich nur grob entworfen und im weiteren Verlauf des Projektes ergänzt. (4) Die Erarbeitung der Liste ist Team-Arbeit. Hier müssen die Bedürfnisse aller an einem Projekt Beteiligten berücksichtigt werden: Autorin, Herausgeberin (FachwissenschaftlerIn, TextkritikerIn), Datenbankmanagerin, Layouterin. Bei der Erarbeitung der Liste muß man im Auge behalten, wer die Auszeichnungen eintragen muß, wer sie weiterverarbeitet, und dabei u. U. eine Trennung der Auszeichnungen in verschiedene Klassen vornehmen (siehe unten 2.2).

In SGML entspricht dieser Liste die *Document Type Definition*. Die DTD verlangt - noch mehr als die vorgeschlagene Liste -, genaue Bedingungen für den Aufbau des Textes anzugeben: die Abfolge der einzelnen Elemente (Überschriften, Abschnitte, Zitate etc.) und der Inhalt der einzelnen Elemente (ein Kapitel besteht aus Unterüberschriften, Abschnitten, Zitaten und Anmerkungen; ein Abschnitt besteht aus Text und Anmerkungen; Text sind Buchstaben

und Ziffern etc.). Der Zweck der DTD ist die automatische Weiterverarbeitung, der Zweck der Liste ist Gedächtnisstütze fürs Textauszeichnen und Programmieren. Beide verfolgen dasselbe Ziel, der Unterschied der Mittel bedingt die höheren formalen Anforderungen an die DTD.

Ist jede Fachkraft im Team mit der Liste einverstanden, so kann man sich sicher sein, daß später die Schritte 3 und 4 auf jeden Fall durchführbar sind und das gewünschte Ergebnis erreichen.

2.2 Auszeichnen des Textes

Der zweite Schritt ist das Auszeichnen des Textes. Dieser zentrale Schritt fällt oft zusammen mit der Texterfassung selbst. Bei großen Projekten wird er sich - trotz Einsatz von OCR-Lesern, Rechtschreibprüfern und halbautomatische Unterstützung bei der Auszeichnung - über Jahre hinziehen und von vielen Mitarbeiterinnen geleistet werden.

In SGML wird bei den Auszeichnungen zwischen verschiedenen Arten (*tag, entity, reference*) unterschieden. Auch bei den selbsterfundenen Auszeichnungen lohnt es sich, auf qualitativ verschiedene Arten von Auszeichnungen zu achten und diese u. U. in ihrer Form zu unterscheiden und sie ggf. verschiedenen Arbeitsschritten zuzuordnen. Zu unterscheiden sind z. B. die strukturell notwendigen Auszeichnungen (*tags*), die auch Einfluß auf die Typographie haben und die Grundstruktur der Daten wiedergeben, Kürzel für lange Wörter oder Sonderzeichen (*entities*), Kürzel zum Einfügen von Teilen aus anderen Texten (*references*), Querverweise innerhalb des Textes (*references*) und Registerinträge. Außerdem können verschiedene Analyseergebnisse über den Text gelegt werden (Quantitäten, Versmaße, Zuordnungen zu Themenbereichen etc.).

Eine sachliche Auszeichnung in diesem Sinne ist auch die Strukturierung von Daten ähnlich einer Datenbank, wie es z. B. für eine Bibliographie angebracht ist. Hier dienen die Auszeichnungen erst in letzter Linie dem Ziel, die Typographie des Ausdrucks zu gestalten.

2.2.1 Sinnvolle Zeichenkombinationen für sachliche Auszeichnungen

Eine praktische Frage ist, welche Zeichenkombination man als Auszeichnungen in den Daten verwendet. Die Zeichenkombinationen für die Sachliche Auszeichnungen können frei gewählt werden. Sie müssen lediglich eindeutig sein und sich vom Text und den übrigen Daten unterscheiden. Die folgenden Gesichtspunkte sind bei der Festlegung der Zeichenkombinationen zu berücksichtigen:

- Mnemotechnische Gesichtspunkte sind bei der Wahl der Zeichenfolgen wichtig.

- Für Ungeübte ist wichtig, daß sich die Auszeichnungen im Editor vom eigentlichen Text optisch gut unterscheiden. Sie sollen auffallen, aber beim Lesen des Textes nicht stören.
- Sachliche Auszeichnungen sollten so kurz wie möglich sein.
- Die für sachliche Auszeichnungen verwendeten Zeichen sollten auf der Tastatur günstig liegen.

Eine bewährte Form ist es, einheitlich zwei mnemotechnisch gewählte Zeichen zwischen spitze Klammern zu setzen: <xy>. Dies geht allerdings nur, wenn im Text selbst keine spitzen Klammern vorkommen oder wenn man für diese im Text eine Ersatzcodierung wählt. (5) Ist diese Einheitlichkeit nicht notwendig, so kann eine beliebige - für häufig gebrauchte natürlich möglichst geringe - Anzahl von Zeichen in spitzen Klammern verwendet werden, wie es dem Vorschlag von SGML entspricht. SGML hat die durchaus zu empfehlende Konvention, daß bei paarig vorkommenden Auszeichnungen die Enderkennung aus denselben Zeichen besteht wie die Anfangserkennung, jedoch durch </ anstelle von < eingeleitet wird (z. B. <la> ... </la>).

Die spitzen Klammern erfüllen allerdings nicht das Kriterium, daß sie auf der (PC-)Tastatur günstig liegen, sie haben jedoch den Vorteil, daß sie im Text gut zu erkennen sind. Eine Auszeichnung der Form &xy, wie sie für Makros in TUSTEP-SATZ vorgesehen sind, sind etwas leichter zu schreiben, im Text aber schwieriger zu erkennen.

Für strukturierte Daten, die in Feldern erfaßt werden, wie z. B. bibliographische Daten oder grammatische Analysen zu je einem Wort, ist als Textauszeichnung empfehlenswert ein Sonderzeichen (z. B. @, &, * etc.) gefolgt von einem mnemonisch gewählten Buchstaben und einem Blank.

2.3 Weiterverarbeitung des ausgezeichneten Textes (parsing)

Der nächste Schritt ist die Bearbeitung des ausgezeichneten Textes mit Programmen. In der TUSTEP-Umgebung ist dies in der Regel das Feld des Kommandos KOPIERE.

Ein erster, wichtiger Punkt wird die Plausibilitätsprüfung der Auszeichnungen sein, entsprechend den in der Liste festgelegten Bedingungen (erlaubte Auszeichnungen, Paarigkeit, legale Schachtelungen, korrekte Abfolge etc.)

Für die Weiterverarbeitung zum Drucken ist unverzichtbar ein Programm zur Umsetzung der sachlichen Auszeichnungen (*generic markup*) in die spezifischen Steueranweisungen (*specific markup*) für das anschließend verwendete Formatier-Programm. Hier erfolgt die Umsetzung und Weichenstellung für WORD, TeX, TUSTEP-SATZ oder TUSTEP-FORMATIERE. Selbstverständlich kann ein und derselbe Text durch verschiedene Programme für verschiedene Formatierer aufbereitet werden, kann derselbe Text durch verschiedene Programme in verschiedenem typographischem Stil aufbereitet werden, können aus demselben Text verschiedene Teile für die Ausgabe berücksichtigt werden

(Weglassen der textkritischen Anmerkungen, Auswählen einer Sprache bei mehrsprachigen Texten, Ausgabe der Kommentare etc.)

Der große Vorteil der sachlich orientierten Auszeichnungen ist aber, daß außer diesen Programmen zur Vorbereitung der Druckausgabe, was für viele Daten nur ein sekundärer Einsatz ist, in diesem Schritt auch Programme verwendet werden können, die die Daten analytisch weiterverarbeiten, für andere Zwecke aufbereiten (Einspeisen in eine Datenbank) oder auswerten.

Bei SGML ist dieser Schritt die Aufgabe von Parsern, die - so die Theorie - selbständig durch andere Programme erzeugt werden, indem sie die Document Type Definition interpretieren und auswerten. In jedem Fall muß das Parser-Programm an die in der Liste bzw. in der DTD festgelegten Bedingungen angepaßt sein.

In den einfachen Fällen wird es genügen, die Auszeichnungen 1:1 in die spezifische Steueranweisungen des Formatier-Programms auszutauschen. In der TUSTEP-Umgebung genügt hier die einfache Verwendung des Austauschparameters X innerhalb des Kommandos FORMATIERE. Bei komplizierteren Strukturen wird eine solche 1:1-Umsetzung nicht genügen, so daß man die spezifischen Steueranweisungen durch algorithmische Verarbeitung erzeugen muß.

2.4 Formatieren und Ausgeben

Durch Schritt 3 liegt der Text in einer für ein formatier-Programm spezifischen Form vor. Im vierten Schritt wird das entsprechende Formatier-Programm aufgerufen, das - je nach Programm - noch mehr oder weniger großen Einfluß auf die Typographie nehmen kann (Satzspiegel, Kopftext etc.). Daraufhin erfolgt das Ausdrucken auf Papier.

3. Beispiele

Die Beispiele versuchen das Spektrum des Einsatzes der sachlichen Auszeichnungen abzudecken. Die vorgestellten (einfachen) Programme entstammen der TUSTEP-Umgebung und beziehen sich auf die Aufbereitung der Daten zum Druck. Für die anderweitige Weiterverarbeitung sollen Hinweise genügen.

3.1 Auszeichnung eines Fließtextes

Die Auszeichnungen des folgenden Beispiels streben keine Vollständigkeit an, sondern sind als Anregung gedacht. Die anschließende Umsetzung bezieht sich ausschließlich auf die TUSTEP-Umgebung. Sie zeigt innerhalb von TUSTEP vor allem die Offenheit der Auszeichnungen für FORMATIERE und SATZ.

3.1.1 *Schritt 1: Die Liste der sachlichen Auszeichnungen*

<1a>	Überschrift 1. Ordnung, Anfang
<1e>	Überschrift 1. Ordnung, Ende
<2a>	Überschrift 2. Ordnung, Anfang
<2e>	Überschrift 2. Ordnung, Ende
<3a>	Überschrift 3. Ordnung, Anfang
<3e>	Überschrift 3. Ordnung, Ende
<ga>	Gedicht, Anfang
<gs>	Gedicht-Strophenanfang
<gz>	Gedicht-Versanfang
<ge>	Gedicht, Ende
<za>	Zitat, Anfang
<ze>	Zitat, Ende
<ea>	Exkurs-Anfang
<ez>	Zeilenwechsel im Exkurs
<ee>	Exkurs-Ende
<ab>	Neuer Abschnitt
<aa>	Autorenname, Anfang
<ae>	Autorenname, Ende
<sa>	Sacherklärung, Anfang
<se>	Sacherklärung, Ende

Bemerkungen:

Im Hinblick auf die Typographie ist wichtig zu wissen, daß die Auszeichnungen für die Überschriften die Leerzeilen davor und dahinter enthalten, daß also nach einer Überschrift keine Auszeichnung für Abschnitt-Anfang zu stehen braucht.

3.1.2 *Schritt 2: Auszeichnung des Textes*

In der folgenden Form wird der Text im Computer (für die Programmbeispiele in einer Datei namens *text1*) erfaßt, ausgezeichnet und aufbewahrt. Diese Daten sind die Grundlage für alle weiteren Bearbeitungsschritten, die noch vorgesehen sind.

```
<1a>Wilhelm Busch - Leben und Werk<1e>
<2a>1. Biographie<2e>
<aa>Wilhelm Busch<ae> wurde am 15 . 4 .1982 in
Wiedenstadt
<sa>Wiedenstadt ist ein Dorf westlich von
```

Hannover.<se>
geboren. <aa>Busch<ae> sagt über seinen Vater:
<za>'Mein Vater war Krämer, klein, kraus, rührig,
mäßig und gewissenhaft; stets besorgt, nie zärtlich,
zum Spaß geneigt, aber ernst gegen Dummheiten. <ze>
<2a>2. Werk<2e>
1865 entsteht das Werk 'Max und Moritz' .
<ab>In der fruchtbarsten Zeit von 1872 bis 1884
entstehen u. a. 'Die fromme Helene' und die
'Knopp-Trilogie'. Daraus der Anfang:
<ga>Sokrates, der alte Greis,
<gz>Sagte oft in tiefen Sorgen:
<gz>'Ach wieviel ist doch verborgen,
<gz>Was man immer noch nicht weiß. '
<gs>Und so ist es. - Doch indessen
<gz>Darf man eines nicht vergessen:
<gz>Eines weiß man doch hinieden,
<gz>Nämlich, wenn man unzufrieden.<ge>

3.1.3 Schritt 3: Umsetzung der sachlichen Auszeichnungen

Schritt 3 ist die maschinelle Bearbeitung der Daten aufgrund der sachlichen Auszeichnungen. Als erstes muß hier ein Programm stehen, das die Daten auf ihre Konsistenz prüft. In der TUSTEP-Umgebung lassen sich mit dem Kommando KOPIERE solche Programme schreiben. (6)

Was hier vorgestellt werden soll, ist die Umsetzung der sachlichen Auszeichnungen in die spezifischen Auszeichnung für FORMATIERE- oder SATZ-Steueranweisungen. Sie erfolgt in einem eigenen, dem Kommando SATZ bzw. FORMATIERE vorgeschalteten Programm (Kommando KOPIERE), das durch die (hier nicht gebrauchten) Möglichkeiten der algorithmischen Programmierung auch die Umsetzung von komplizierten Strukturen erlaubt. (7)

Das Beispiel ist nur eine mögliche typographische Form. Der Vorteil der sachlichen Textauszeichnung ist ja gerade, daß man auf die typographische Form nicht festgelegt ist, daß man also für die Ausgabe auch einen ganz anderen Stil (z. B. mit zentrierten Überschriften) wählen könnte. Die Umsetzung versucht, auch typographische Feinheiten zu berücksichtigen, die allerdings in dem kurzen Text nicht zur Geltung kommen. Die sachlichen Auszeichnungen sind dabei allein wegen der Länge und Komplexität der notwendigen Steueranweisungen im Vorteil.

a) Umsetzung für FORMATIERE

#kopiere, text1, zwischen.fo, loeschen=+, parameter=*

Überschriften

```
xx      !<<1a>>!$$$0 &v0 @f @s- @r- __ $5 #f+#s+#3+!  
xx      !<<1e>>!#?- @s+ @r+ $2 !  
xx      !<<2a>>!$$$8 &v0 @s- @r- $2 #3+!  
xx      !<<2e>>!#3- @s+ @r+ $1 !  
xx      !<<3a>>!$$$5 $1 @r- @s- #f+!  
xx      !<<3e>>!#f- @r+ @s+ $0!
```

Gedicht

```
xx      !<<ga>>!$0 &v1 &:3&;6!  
xx      !<<gs>>!$1 &:3&;6!  
xx      !<<gz>>!$$$ &:3&;6!  
xx      !<<ge>>!$0!
```

Zitat

```
xx      !<<za>>!&l5 &r5 $0 &v1 !  
xx      !<<ze>>!&l0 &r0 &v0 $0 !
```

Exkurs

```
xx      !<<ea>>!&l5 &r5 $0 &v1 !  
xx      !<<ez>>!$!  
xx      !<<ee>>!&l0 &r0 &v0 $0 !
```

Abschnitt

```
xx      !<<ab>>!$ !
```

Autorenname

```
xx      !<<aa>>!!  
xx      !<<ae>>!!
```

Sacherklärung

```
xx      !<<sa>>!@f+!  
xx      !<<se>>!@f-!
```

*eof

Bemerkungen:

Bei den Überschriften wird grundsätzlich Silbentrennung und Randausgleich ausgeschaltet, damit lange Überschriften, die sich über mehrere Zeilen erstrecken, nicht wie ein kurzer Abschnitt aussehen.

Zitat (gedacht ist an längere Zitate) und allgemeiner Exkurs werden typographisch gleich behandelt. Ihre Unterscheidung dient in diesem Beispiel nur potentiellen anderen Verarbeitungsschritten.

Die Autorennamen sollen im gedruckten Text nicht besonders gekennzeichnet sein. Die sachliche Auszeichnung bleibt daher bei der Umsetzung in spezifische Steueranweisung unberücksichtigt, d. h. sie wird eliminiert. Die Auszeichnung dient in diesem Fall z. B. der Erstellung eines Autorenregisters.

Die Daten in der Datei *zwischen.fo* sehen folgendermaßen aus:

```
$$$0 &v0 @f @s- @r- __ $5 #f+#s+Wilhelm Busch - Leben
  und Werk#?- @s+ @r+ $3
$$$8 &v0 @s- @r- $3 #3+1. Biographie#3- @s+ @r+ $2
Wilhelm Busch wurde am 15.4.1982 in
Wiedenstadt
@f+Wiedenstadt ist ein Dorf westlich von
. . .
```

b) Umsetzung für SATZ

Eine Umsetzung desselben Textes für das Formatieren mit dem Kommando SATZ könnte folgendermaßen aussehen.

```
#kopiere, text1, zwischen.sa, loeschen=+, parameter=*
```

Überschriften

```
xx      !<<1a>>! &&&!
xx      !<<1e>>!&&&{!
xx      !<<2a>>! &&!
xx      !<<2e>>!&&{!
xx      !<<3a>>! &!
xx      !<<3e>>!&{!
```

Gedicht

```
xx      !<<ga>>!&1& &=010&=-020!
xx      !<<gs>>!$$$/$$$
xx      !<<gz>>!$ &=010&=-020!
xx      &=010&=-020!
xx      !<<ge>>!&1&{
```

Zitat

```
xx      !<<za>>! &2&!
xx      !<<ze>>!&2&{!
```

Exkurs

```
xx      !<<ea>>! $$!
xx      !<<ez>>! $$$!
xx      !<<ee>>! $$${!
```

Abschnitt

```
xx      !<<ab>>! $!
```

Autorenname

```
xx      !<<aa>>!#k+!
xx      !<<ae>>!#k-!
```

Sacherklärung

```
xx      !<<sa>>!@f+!
xx      !<<se>>!@f-!
```

```
*eof
```

3.1.4 Schritt 4: Umbruchberechnung und Druckausgabe

Das durch Schritt 3 produzierte Ergebnis in der Datei *zwischen.fo* ist auf ein spezifisches Formatier-Programm festgelegt, das TUSTEP-Kommando FORMATIERE bzw. SATZ. Diese Kommandos verarbeiten die jeweiligen spezifischen Steueranweisungen, berechnen daraus den Umbruch und machen die Daten fertig zur Druckausgabe.

a) Umbruch mit dem Kommando FORMATIERE

Die Gestalt des Textes wird hauptsächlich durch die Auszeichnungen bestimmt. Im Kommando FORMATIERE selbst wird linker und rechter Rand, die Zahl der Zeilen je Seite und der Fuß- und Kopf text festgelegt werden.

```
#formatiere, zwischen.fo, mo=hpii, lo=+, par=*  
dr          1 15 60  
kt          !Beispiel @/ xxx!  
*eof
```

Das Ergebnis von FORMATIERE ergibt den in Abbildung 1 wiedergegebenen Ausdruck.

Wilhelm Busch - Leben und Werk

1. Biographie

Wilhelm Busch wurde am 15.4.1822 in Wiedenstadt¹ geboren. Busch sagt über seinen Vater:

'Mein Vater war Krämer, klein, kraus, rührig, mäßig und gewissenhaft; stets besorgt, nie zärtlich, zum Spaß geneigt, aber ernst gegen Dummheiten.'

2. Werk

1865 entsteht das Werk 'Max und Moritz'.

In der fruchtbarsten Zeit von 1872 bis 1884 entstehen u. a. 'Die fromme Helene' und die 'Knopp-Trilogie'. Daraus der Anfang:

Sokrates, der alte Greis,
Sagte oft in tiefen Sorgen:
'Ach wieviel ist doch verborgen,
Was man immer noch nicht weiß.'

Und so ist es. - Doch indessen
Darf man eines nicht vergessen:
Eines weiß man doch hinieden,
Nämlich, wenn man unzufrieden.

¹ Wiedenstadt ist ein Dorf westlich von Hannover.

b) Umbruch mit dem Kommando SATZ

Das Ergebnis *zwischen,sa* aus obiger Umsetzung kann mit dem Kommando SATZ verarbeitet werden.

Die endgültige Gestaltung der Überschriften, des Gedichtes, des Zitates und des Exkurses wird über Parameter vorgenommen. Die Fußnoten sind mit dem Standard-Makro *FUNO zu extrahieren und zunächst getrennt zu Setzen.

Ausgedruckt erscheint der Text in ähnlicher typographischer Gestaltung wie die Umsetzung mit FORMATIERE. Die Qualität der Schrift und die Möglichkeiten der Typographie entsprechen der Erscheinung des vorliegenden Artikels, der - wie die ganze Zeitschrift - mit TUSTEP gesetzt ist.

3.2 Auszeichnung von kritischen Apparaten

Bei kritischen Apparaten ist der Einsatz von sachlichen Auszeichnungen aus zwei Gründen sinnvoll. Der erste praktische Grund ist, daß die Wiedergabe von verschiedenen Apparaten mit Verweis auf den Zeilenzähler mit TUSTEP nur im Satzprogramm möglich ist. Die Verwendung sachlicher Auszeichnung erlaubt eine Umsetzung der Apparate auch mit FORMATIERE, so daß Erfassung und Probeausdruck für die Korrektur auch auf dem PC (8) möglich ist. Der zweite Grund liegt darin, daß man mit sachlichen Auszeichnungen Varianten kategorisieren und auszeichnen kann, um später vollständig auf alle Varianten zuzugreifen, (9) und unabhängig davon die Möglichkeit hat, für den Druck daraus je nach Zweck nur bestimmte Varianten auszuwählen.

3.2.1 *Darstellung eines kritischen Apparates für Korrekturzwecke*

Als Auszeichnung für kritische Apparate empfehlen sich Auszeichnungstripel. Im Text wird der Anfang und das Ende der Stelle ausgezeichnet, zu der es eine Variante gibt; direkt anschließend wird die Variante selbst angeführt, die durch eine Auszeichnung abgeschlossen wird. (10)

Beispiel

<tt> Textkritische Stelle im Text, Anfang
<ta> Textkritische Anmerkung, Anfang
<te> Textkritische Anmerkung, Ende

Ein damit ausgezeichnete Text könnte so aussehen:

Und Simson liebte eine Frau im Tal Sorek; ihr Name war
<tt>Delila.<ta>Mult Mss von LXX: Dalida<te> Die Für-
sten der Philister kamen zu ihr und sagten: Versuch ihn
zu verführen und finde heraus, wie <tt>wir<ta>LXX:

du<te> ihn überwältigen <tt>können<ta>LXX:
kannst<te>.

Beim Ausdrucken für Korrekturzwecke mit FORMATIERE ist es sinnvoll, nicht eine Annäherung an die endgültige typographische Form (11) anzustreben, sondern davon zu abstrahieren und auch im Ausdruck vor allem die inhaltlichen Informationen (ohne die unübersichtlichen Auszeichnungen) darzustellen. Die Apparateinträge werden an Ort und Stelle im laufenden Text ausgegeben, jedoch am rechten Rand, der vorher entsprechend groß festgelegt wurde, freigestellt. Man kann damit auf dem Papier Text und Apparat klar unterscheiden. Die Zuordnung erfolgt ohne Verweisziffern durch das Nebeneinanderstehen. Die Steueranweisungen für diese Umsetzung des Apparats sehen folgendermaßen aus. (12)

```

xx      !<tt>!#/+!
xx      !<ta>!#/- &v1 &r0 &a38 !
xx      !<te>!&10 &r25 $0 &v0 !
    
```

Diese Steueranweisung ergeben das in Abbildung 2 wiedergegebene Druckbild.

Und Simson liebte eine Frau im Tal	
Sorek; ihr Name war Delila.	Mult Mss von LXX: Da-
	lida
Die Fürsten der Philister kamen zu	
ihr und sagten: Versuch ihn zu ver-	
führen und finde heraus, wie wir	LXX: du
ihn überwältigen können	LXX: kannst

3.2.2 Verschiedene Auswahl des Apparats

Im obigen Beispiel wurde im Haupttext die Stelle, zu der es eine Variante gibt, durch Kursive gekennzeichnet und die Variante als Apparat wiedergegeben. Die Kennzeichnung im Haupttext kann dabei natürlich jederzeit weggelassen werden.

Wenn nicht die philologische Detailarbeit, sondern die Fachinformation des Textes interessiert, bieten die sachlichen Auszeichnungen die Möglichkeit, die Variante ganz wegzulassen und nur den Haupttext (mit oder ohne Kursive) wiederzugeben. Die Umsetzung in SATZ-Steueranweisungen sieht dafür folgendermaßen aus:

```
xx      !«tt»i#/+!  
xx      !«ta>>!#/-&x!  
xx      !<<te>>!&x{!
```

Das Beispiel hätte gedruckt die folgende Form:

... ihr Name war *De lila*. Die Fürsten der ...

Umgekehrt kann es sinnvoll sein, den Text der Variante - ohne die Sigle - im Text wiederzugeben und den Haupttext wegzulassen. Damit erhält man fortlaufend ausschließlich den Text der Variante:

```
xx      !«tt»!&x!  
xx      !«ta»<></: !&x{!  
xx      !<<te>>!!
```

Ausgedruckt sieht das folgendermaßen aus:

Und Simson liebte eine Frau im Tal Sorek; ihr Name war Dalida. Die Fürsten der Philister kamen zu ihr und sagten: Versuch ihn zu verführen und finde heraus, wie du ihn überwältigen kannst.

Soll dieses Verfahren für verschiedene Textzeugen durchgeführt werden, so müssen - ggf. durch etwas kompliziertere Auswahlverfahren - die Siglen der Textzeugen bei der Auswahl berücksichtigt werden. Dazu ist es sinnvoll, wenn die Siglen - wie hier - eindeutig zu erkennen sind.

Verwendet man für verschiedene Arten von textkritischen Varianten (13) verschiedene Auszeichnungen, so können auch diese in ähnlicher Weise unterschieden werden.

Eine weitere sinnvolle Möglichkeit für große Projekte ist es, die Varianten in der skizzierten Weise möglichst vollständig zu erfassen und eine bestimmte Behandlung der Varianten generell vorzugeben, z. B. daß die Variante prinzipiell unberücksichtigt bleibt und nur der Haupttext abgedruckt wird. Von dieser generellen Regelung kann dann im Einzelfall durch geringfügiges Verändern der Auszeichnung abgewichen werden, z. B. die Verwendung von <f*> anstelle von <it> könnte bedeuten, die Variante soll als Apparat gedruckt werden; <*r> könnte bedeuten, die Variante soll in den Text und der eigentlich Text in den Apparat. (14)

3.3 Sekundäre Verwendung eines Textes

In manchen Fällen kommt es vor, daß ein Text in Teilen für andere Zwecke verwendet werden soll. Z. B. wenn in einem Editionsbericht verschiedene (auch sich überschneidende) Passagen des Editionstextes zitiert werden sollen, man dafür aber den Editionstext mit seinen letzten Korrekturen verwenden will. Hier kann es sinnvoll sein, nicht sofort den Editionstext in den Editionsbericht zu kopieren, sondern im Bericht und im Text Markierungen anzubringen, die den Text entsprechend zuordnen und später - nach der letzten Korrektur des

Editionstextes - per Programm die entsprechenden Passagen zu übernehmen. Auszeichnungen für eine solche Auswahl im Text könnten folgende Form haben:

<a123> Anfang Auswahl Nr. 123

<e123> Ende Auswahl Nr. 123

Diese Auszeichnungen haben eine ganz andere Qualität als die bisher vorgestellten. Sie bauen zwischen Editionsbericht und -text eine Art Hypertext-Struktur auf. Sie haben mit dem eigentlichen Text und seiner Struktur nichts zu tun, sondern zeigen eine bestimmte Verwendungsweise an. Die neuen Auszeichnungen stören die bisherige Verwendung des Textes nicht. Man muß lediglich in den Programmen berücksichtigen, daß diese Auszeichnungen übergangen werden. Für die Realisierung der eigentlichen Auswahl durch ein Programm gibt es viele Möglichkeiten. (15)

3.4 Verarbeitung strukturierter Daten

3.4.1 Bibliographie

Ein typisches Beispiel für datenbankartig strukturierte Daten ist eine Bibliographie, bei der jeder Eintrag verschiedene Felder enthält. (16) Es ist klar, daß der gezielte Zugriff auf die einzelnen Felder für Sortierung und Retrieval notwendig sein muß und durch die strukturierte Abspeicherung auch möglich ist. Dies wurde in HSR bereits ausführlich dargestellt. (17) Aus dem hier gesagten ist aber auch deutlich geworden, daß diese strukturierte Form der Abspeicherung kein Hindernis für einen schönen Ausdruck in der gewohnten Form einer gedruckten Literaturliste sein muß. (18)

Beispiel

Das Beispiel der Bibliographie aus HSR 14 läßt sich mit den dort besprochenen Programmen in folgende Form bringen:

&a Autorin
&t Titel
&z Zeitschrift
&b Band
&o Ort
&j Jahr
&s Signatur
&k Kommentar

Ausgedruckt als Literaturliste sollen diese Daten folgende Form haben:

Autorin, Titel, Ort Jahr.

bzw.

Autor, Titel, in: Zeitschrift Band (Jahr) .

Signatur und Kommentar fallen also weg. Band wird nur berücksichtigt, wenn es sich um eine Zeitschrift handelt; das Jahr kommt dann in Klammern. Der Ort wird nur genannt, wenn es keine Zeitschrift ist. Durch diese Bedingungen ist eine 1:1-Umsetzung der Auszeichnungen nicht möglich. Das folgende Programm ist ein Beispiel für eine algorithmische Umsetzung der Auszeichnungen.

```
#ko,bib, zwischen.fo, lo=+, par=*
```

```
Verteilen zu verschiedener Bearbeitung
```

```
ken      !&a! &t! &z<%! &b<%! &o! &j!<%!  
ksp      12 22 32 41 51 61 0  
spw      8 8 8 0 0 72  
sp2      3 3 3 3 3 3 3
```

```
Autor: Zeilenwechsel, abgeschlos-  
sen durch Komma
```

```
erg      1  ! !#f-, !  
xx       1  !&a<> ! $0 $c; 5 #f+!
```

```
Titel: abgeschlossen durch Komma
```

```
erg      2  ! !, !  
xx       2  i&t!!
```

```
Zeitschrift: falls vorhanden eingelei-  
tet durch 'in: '
```

```
xx       3  !&z!in: !  
wss      3  wsl
```

```
Band: nur ausgeben, wenn Zeitschrift
```

```
ws-      4  wsl  
xx       4  !&b!!
```

```
Ort: nicht ausgeben bei Zeitschrift
```

```
ws +     5  wsl  
xx       5  !&o!!
```

```
Jahr: abgeschlossen durch Punkt
```

```
ws+      6  wsl  
erg      6  ! ! . !  
xx       6  l&j ! !
```

```
Jahr: bei Zeitschrift in Klammern,
```

```
erg      7  ! ! ) . !  
xx       7  !&j<> ! ( !  
wsl      7  wsl
```

```
*eof
```

3.4.2 Analytisch angereicherte Texte

Ein anderes Beispiel für datenbankartig strukturierte Daten ist die Anreicherung eines Textes mit grammatikalischen Informationen in einem Format wie es bei der Westminster Hebrew Morphological Database (19) verwendet wird. Pro Wort gibt es eine Zeile mit folgenden Informationen: Stellenangabe, Blank als eindeutige Textauszeichnung, objektsprachlicher Text, Blank, Lemma, das Zeichen @, Analyse. (20) Ohne weiteres kann auch hier der objektsprachliche Text isoliert und am Stück ausgedruckt werden. Für Analysen sind die Kennungen eindeutig, um jeweils auf bestimmte Felder zugreifen zu können und die gewünschte Information zu erhalten. (21)

Anmerkungen:

- (1) TUSTEP (TUebinger System für TEXTverarbeitungs-Programme) ist eine Entwicklung der Abteilung Literarische und Dokumentarische Datenverarbeitung am Zentrum für Datenverarbeitung der Universität Tübingen. — TUSTEP gibt es als PC-Version unter MS-DOS und DR-DOS (IBM-PCs und kompatible). Auf Großrechnern und Workstations läuft TUSTEP derzeit unter den Betriebssystemen MVS (IBM), VMS (DEC VAX), BS2000 (SIEMENS) und UNIX (CONVEX, SunOS, AIX). Weitere Implementierungen unter UNIX sind geplant.—Für Hochschulen gibt es für TUSTEP Campus-Lizenzen. Nähere Informationen über den Vertrieb von TUSTEP und den Bezug des Handbuchs sind bei Prof. Dr. Ott (e-mail: TUSTEP@zdv.uni-tuebingen.de) am Zentrum für Datenverarbeitung der Universität Tübingen zu erhalten. Für die Beratung in Fragen zu TUSTEP stehen die Mitarbeiter der Abteilung »Literarische und Dokumentarische Datenverarbeitung« am ZDV in Tübingen zur Verfügung.
- (2) **Standard** Generalized Markup Language (SGML), ISO-Norm 8879 von 1986. Die komplette ISO-Norm von SGML mit einem Kommentar findet sich leicht zugänglich in C. F. **GOLDFARB**, The SGML Handbook, Oxford 1990, ISBN 0-19-853737-9. Eine Einführung in SGML bietet: E. van **HERWIJNEN**, Practical SGML, Dordrecht 1990, ISBN 0-7923-0635-X. - Programme um SGML-Dokumente zu verarbeiten finden sich **als public domain Software auf** verschiedenen **ftp-servern**, z.B. sgml.ex.ac.uk, rs3.hrz.th-darmstadt.de
- (3) Besonders bei **stark strukturierten Texten wie z. B. Wörterbuchartikel** gewährleisten das Festlegen solcher Bedingungen und die Vorgabe der Reihenfolge der einzelnen Auszeichnungen die Konsistenz des gesamten Werkes.

- (4) Zu beachten ist dabei, daß die Offenheit zur Ergänzung gegeben ist, ebenso die Offenheit zur Simplifizierung; eine nachträgliche Änderung oder Differenzierung aber ist mit Aufwand verbunden.
- (5) Die Beschränkung auf zwei Zeichen bietet in der TUSTEP-Umgebung außer der geringen Schreiarbeit die Möglichkeit der direkten Umsetzung in SATZ-Makros der Form $\langle E.x \rangle$. Denn für die direkte Verwendung dieser Makros im Kommando SATZ sind hinter & nur zwei Zeichen zulässig.
- (6) Die Vorstellung dieser Möglichkeit von TUSTEP führt im Rahmen dieses Artikels zu weit.
- (7) Das vorliegende Beispiel benötigt lediglich eine lil-Umsetzung, die auch direkt beim Kommando FORMATIERE vorgenommen werden könnte.
- (8) Unter MS-DOS und DR-DOS wird das TUSTEP-Kommando SATZ nicht unterstützt.
- (9) Um z. B. den Text eines bestimmten Textzeugs auszudrucken.
- (10) Das Ende der Stelle im Text und der Anfang der Variante fallen also zusammen.
- (11) Apparat am Fuß des Textes. Mehrere Kurze Apparateinträge in einer Zeile. Verweis in den Text über die Zeilennummer.
- (12) Im Kommando FORMATIERE muß der rechte Rand für den Text dazu - mit Hilfe der Spezifikation VORSPANN und der Steueranweisung &r25 - auf 25 Zeichen festgelegt sein. Es wird von einer Spaltenbreite von 60 Zeichen ausgegangen.
- (13) Anderer Textzeuge; handschriftlicher Zusatz; Rechtschreibkorrektur etc.
- (14) Dises Verfahren wurde in dem Projekt »Das Grundgesetz - Dokumentation seiner Entstehung« der Forschungsstelle für Zeitgeschichte des Verfassungsrecht angewendet und in dem Artikel von Ulrich BACHMANN - Kirsten NEMITZ, Das Grundgesetz - Dokumentation seiner Entstehung. Projektkonfiguration und EDV-Einsatz für die nach Artikel gegliederte Edition, in: Nachrichten für Dokumentation. Zeitschrift für Informationswissenschaft und -praxis 43 (1992), 317-325 ausführlich vorgestellt.
- (15) In dem obengenannten Forschungsprojekt wurde auch dieses Verfahren verwendet, siehe hierzu Winfried BADER, The German Basic Law. Documentation on its Genesis in order by Articles, in: Association for History & Computing, Seventh International Congress, Bologna 29/VHI - 2/IX/1992. Abstracts, Bologna 1992, 31-32.
- (16) Autor, Titel, Jahr, Verlag, Sigle, Kommentar etc.
- (17) Winfried BADER, Das Sortieren mit TUSTEP: Dargestellt am Beispiel einer Bibliographie, in: HSR No. 51, Vol. 14 (1989), 139-160.
- (18) Mit dem TUSTEP-Kommando FORMULAR-AUFBEREITE ist auch die Ausgabe in Form von Bibliothekskärtchen möglich.
- (19) Herausgeber dieser vollständigen morphologischen Analyse der Hebräischen Bibel ist Alan GROVES, Philadelphia.

- (20) Die Textauszeichnung durch die beiden Blanks entspricht zwar nicht den Regeln für Eindeutigkeit gemäß SGML, doch ist es in diesem Fall hinreichend.
- (21) Eine ähnlich **strukturierte** grammatische Analyse ist zusammen mit einer ausführlichen Beschreibung von dazugehörigen TUSTEP-Programmen dargestellt in Winfried **BADER**, Data base functions in TUSTEP, in: Association Internationale Bible et Informatique (AIBI) en collaboration avec l'Eberhard-Karls-Universität de Tübingen, Bible et Informatique: Interprétation, Herméneutique, Compétence informatique / Bibel und Informatik: Interpretation und Hermeneutika. Actes du Troisième Colloque International / Akten des dritten Internationalen Kolloquiums Tübingen, 26-30 August 1991, (Travaux de Linguistique Quantitative 49; zugleich DEBORA 7), Paris, Genève 1992, 449-470.