

Dritter Tübinger TEI-Workshop im ZDV

Vom 26. bis 28. Februar 1997 fand im ZDV der dritte TEI-Workshop zum Thema *SGML-konforme Textauszeichnung nach den Richtlinien der Text Encoding Initiative (TEI)* statt. Wie bei den ersten beiden Workshops (November 1995 und März 1996) referierten Michael Sperberg-McQueen (University of Illinois at Chicago) und Winfried Bader (jetzt: Deutsche Bibelgesellschaft, Stuttgart). 42 Teilnehmerinnen und Teilnehmer aus dem In- und Ausland kamen nach Tübingen (FU Berlin, Bielefeld, Köln, Mainz, IDS Mannheim, Münster, Trier und Tübingen; Italien, Malta, Österreich, Slowenien). Im Unterschied zu den beiden ersten Workshops gehörten alle Teilnehmerinnen und Teilnehmer dem Wissenschafts- und Hochschulbereich an; Angehörige von Verlagen oder von anderen Firmen nahmen nicht teil.

Die Interessen reichten von der Anlage großer elektronischer Archive über die technische Aufbereitung bis zur ästhetischen Gestaltung elektronischer Ausgaben. Es wurde deutlich, wie sehr sich der Stand der Vorkenntnisse gegenüber den vorangegangenen Veranstaltungen verändert hatte. Alle Anwesenden hatten bereits Erfahrungen im Umgang mit dem Internet gesammelt und verfügten außerdem über grundlegende HTML-Kenntnisse. Hierdurch bedingt, bestanden andere Erwartungen: Die Verarbeitungsmöglichkeiten sollten nicht ausschließlich auf ein gedrucktes Endprodukt hin ausgerichtet sein; vorrangig war eher die Frage nach der Schnittstelle zwischen gedruckter und elektronischer Publikation sowie nach ästhetisch anspruchsvollen WWW-Präsentationsmöglichkeiten.

Die Arbeitsschwerpunkte lagen im Bereich von Urkunden und historischen Quellen, Repertorien, Anlage von übers WWW zugänglichen Textdatenbanken, Editionsprojekte aus dem Bereich der Orientalistik und Ägyptologie mit den entsprechenden Zeichensatz-Anforderungen, Transkription von gesprochener Sprache, grammatische Analysen sowie bibliothekarische Fragen und Themenstellungen aus dem Verwaltungsbereich.

Nach einer Einführung in die Grundlagen von SGML wurden die TEI-Richtlinien vorgestellt. Grundsätze der Dokumentenanalyse und der Auszeichnung von Texten wurden erörtert und die Auszeichnung an eigenen Textbeispielen ausprobiert. Zum Programm gehörten dann weiterführende Vorträge zu SGML und TEI über komplexe Analysen und Merkmalsstruk-

turen sowie Schreibsystemvereinbarungen. Eingeschlossen war die Vorstellung von SGML-Tools sowie von Verarbeitungsmöglichkeiten. Es bestand außerdem die Möglichkeit, im PC-Pool selbst mit den Werkzeugen umzugehen. Im wesentlichen entsprachen damit die Themen des Workshops den beiden ersten Veranstaltungen (vgl. dazu BI 95/11+12; BI 96/3+4).

XML: eXtensible Markup Language

Als Neuerung wurde XML präsentiert. XML ist ein vereinfachter Dialekt von SGML, der im November 1996 als *draft*-Dokument fertiggestellt wurde (siehe <http://www.textuality.com/sgml-erb/WD-xml-lang.html>). Das Ziel ist, einerseits, den Zugang zur Erstellung von SGML-Dokumenten zu erleichtern, andererseits, SGML-Dokumente im WWW anbieten und verarbeiten zu können. XML ist kompatibel zu SGML, d. h. daß jedes XML-Dokument gleichzeitig ein vollwertiges SGML-Dokument gemäß dem ISO-Standard 8879 ist.

Während HTML oder TEI SGML-Anwendungen darstellen, handelt es sich bei XML um ein Derivat. XML ist eine Untermenge von SGML mit einer strengeren, unflexibleren, aber wesentlich einfacheren Struktur. Beispielsweise besteht in SGML die Möglichkeit, durch entsprechende Angaben in der DTD die Endmarkierung zu minimieren. Wird in der DTD angegeben

`<!ELEMENT TEI.2 – O (teiHeader, text)>`, so bedeuten die Angaben nach dem *generic identifier* 'TEI.2', daß eine Anfangsmarkierung des Elements gesetzt werden muß (–), eine Endmarkierung aber auch fehlen darf (Großbuchstabe 'O').

In XML gibt es diese Möglichkeit nicht: Ein Element, das mit einer Anfangsmarkierung beginnt, muß durch eine Endmarkierung geschlossen werden – außer es handelt sich um ein leeres Element. Die Syntax kann nicht modifiziert werden (*concrete reference syntax*). Diese strenge Syntax bedingt, daß in XML im Grunde keine DTD eigens zum Dokument dazugeliefert werden muß. Die DTD geht aus dem Dokument selbst hervor.

Neuer Parser nsgmls

Zusätzlich zu den bisher verwendeten Tools (siehe BI-Artikel) wurde das neue SGML-

Parser-Paket SP-1.1.1 von James Clark vorgestellt. Dazu gehört nsgmls, die neue Version des Parsers sgmls, der sich durch eine wesentlich höhere Geschwindigkeit auszeichnet. Weitere Informationen unter:

<http://www.sil.org/sgml/publicSW.html#parsers>

Resümee und Ausblick

Die elektronische Publikation gewinnt in der Diskussion um geisteswissenschaftliche Editionen zunehmend an Gewicht. Bei der Anlage neuer Projekte wird sie von vornherein in die Überlegungen einbezogen. Als Forum für den

Austausch zwischen verschiedenen Fächergruppen wurde der Tübinger Workshop ein weiteres Mal reichlich genutzt. In der abschließenden Diskussion wurde mehrfach der Wunsch nach einem weiterführenden Kurs deutlich – nicht ohne daß sehr unterschiedliche Leitvorstellungen dabei formuliert worden wären. Eine Folgeveranstaltung wird für 1998 geplant. Einzelheiten und Termine werden rechtzeitig im WWW unter dem URL

<http://www.uni-tuebingen.de/zdv/Termine>
angekündigt.

Annegret Fiebig
fiebig@zdv.uni-tuebingen.de