

Die Text Encoding Initiative (TEI) und SGML: Internationale Richtlinien zur Codierung von elektronischen Texten

Das TEI Advisory Board traf sich im Sommer 1993 in Chicago, um die bisherige Arbeit der TEI zu besprechen. Dr. Bader (ZDV) nahm als Vertreter der Association Internationale Bible et Informatique (AIBI) daran teil.

Die Text Encoding Initiative, seit 1987 eine gemeinsame Arbeitsgruppe der Association for Computers and the Humanities (ACH), der Association for Literary and Linguistic Computing (ALLC) und der Association for Computational Linguistics (ACL), hat Vorschläge zur Codierung von elektronischen Texten erarbeitet. Ziel dabei ist, einen maschinenunabhängigen Datenaustausch zu garantieren und die vielfältige Verwendung (multiuse) von Texten sicherzustellen. Die Codierungsvorschläge der TEI sind eine Anwendung von SGML (Standard Generalized Markup Language, ISO-Norm 8879 von 1986), die seit 1969 auch mit industrieller Beteiligung entwickelt wurde. »Markup language« bedeutet, daß es sich um eine Metasprache handelt, die es erlaubt, die Struktur von Texten durch »tagging« zu beschreiben. Die Charakteristik von SGML liegt darin, daß sie *deskriptiv* und nicht prozedural ist, daß sie das Konzept der *Dokument-Typen* kennt und daß sie *maschinenunabhängig* ist. Die Anwendung von SGML auf einen Text ist ein interpretativer Vorgang mit dem Ziel, die Eigenschaften eines Textes für den Computer explizit zu machen, die ein menschlicher Leser implizit erkennt.

SGML stellt für eine solche Textcodierung nur die formalen Mechanismen zur Verfügung. Die TEI füllt diese Mechanismen inhaltlich für die Bedürfnisse der Geisteswissenschaften. Sagt SGML formal, daß ein Dokument aus Elementen besteht, so schlägt TEI die inhaltliche Anwendung vor, daß ein Roman aus Kapiteln, ein Gedicht aus Strophen und ein

Gesetz aus Paragraphen besteht. Die TEI-Richtlinien enthalten Codierungsvorschläge für sehr vielfältige Dokument-Typen: Prosa, Poesie, Dramen, gesprochene Texte, Briefe, Lexika etc. Außerdem sind vielfältige analytische Codierungen vorgesehen: kritischer Apparat, grammatikalische Beschreibungen, Querverweise und Hyperlinks, die »underground tunnels« zwischen Texten, die die Essenz vieler geisteswissenschaftlicher Forschung bilden.

Die Resultate der TEI liegen als Guidelines vor. Diese wurden den Vertretern zahlreicher Organisationen auf dem Gebiet Computers and the Humanities wiederholt zur Diskussion vorgelegt, darunter auch Prof. Dr. Ott und Dr. Bader vom ZDV. Das Advisory Board Meeting hat die bisher erarbeiteten Guidelines »TEI P2« (proposal 2) verabschiedet (siehe die Schlußvereinbarung unten), die sukzessive unter dem Namen »TEI P3« (proposal 3) als vorläufig endgültige Version erscheinen.

Die buchstabengetreue Anwendung der Guidelines, die zum internationalen Standard werden sollen, für eigene Projekte ist wegen ihres Umfangs (1300 Seiten) und der z. Zt. noch unzureichenden Software-Unterstützung sicherlich schwierig. Jedoch sollte jedes mit der elektronischen Erfassung von Texten beschäftigte Projekt diese Guidelines zur Kenntnis nehmen, da hier das für den wissenschaftlichen Umgang mit Textdaten zu empfehlende Konzept der sachlichen Auszeichnung durch eine erfahrene internationale Arbeitsgruppe unter Anwendung einer festgelegten Norm realisiert ist.

Die Guidelines sind über ftp-server (*listserv@uicvm.uic.edu* oder *sgml.ex.ac.uk*) als ASCII- oder PostScript-File zugänglich. Eine gedruckte Version liegt bei W. Bader, (Zi. 125 im ZDV) zur Einsicht bereit.

Closing Statement of the TEI Advisory Board Meeting (Chicago, 29 June 1993)

The TEI Advisory Board, including the individuals below acting as representatives of the participating organizations indicated, has considered the work of the TEI to date as represented by the 6 June 1993 draft of TEI P2.

We approve of the directions in which this effort has proceeded, and the progress which has been made in reaching the TEI's goals in accord with the Memorandum of Understanding among Participating Organizations and the Poughkeepsie Principles [=Schlußstatement

der Vassar Planning Conference 1987 in Poughkeepsie, New York].

We endorse the TEI's plans for continuation of this effort and for the development, dissemination, and implementation of the TEI Guidelines, and look forward to continuing and enhancing cooperation between the organizations we represent and the TEI.

(Unterzeichnet von den Vertretern der beteiligten Organisationen)