

## EDV und Quellenforschung: ITUG-Jahrestagung 2003 in Burgos/Spanien

Die Jahrestagung 2003 der ITUG (International TUSTEP User Group) fand vom 4. bis zum 10. Oktober in Peñaranda de Duero in der Nähe von Burgos (Spanien) statt. Über 30 Anwender aus Deutschland, der Schweiz und Spanien trafen sich zum Erfahrungs- und Gedankenaustausch über den EDV-Einsatz bei ihren mit TUSTEP bearbeiteten Editions- und Erschließungsprojekten. Neben der Information über die mit der Version 2004 von TUSTEP verfügbaren neuen Leistungen und der Formulierung zahlreicher Wünsche an die Weiterentwicklung von TUSTEP wurden in insgesamt 14 Referaten mit TUSTEP bearbeitete Projekte vorgestellt und Anwendungsprobleme diskutiert. Da die Referate der ITUG-Jahrestagungen Gegenstände betreffen, die über den Kreis der unmittelbaren Teilnehmer hinaus Interesse finden, ist es mittlerweile Tradition geworden, die Beiträge in einem Sammelband zu veröffentlichen: Ein Sammelband, der die Beiträge dieser Jahrestagung allgemein zugänglich machen soll, wird an der Universität Burgos vorbereitet.

Das Generalthema der Jahrestagung wurde von Nicolas Castrillo mit einem Beitrag über Ferdinand I. (dessen 500. Geburtstag in Spanien im Jahr 2003 gefeiert wurde) konkretisiert. Durch quantifizierende Untersuchungen der Sprache einschlägiger Urkunden lassen sich zum einen Entwicklungen in der Hofkanzlei feststellen, zum anderen wird es durch diese Untersuchungen möglich, weiteres Quellenmaterial zu qualifizieren. Die Quellenmaterialien, die in den weiteren hier vorgestellten Forschungsprojekten und Arbeitsvorhaben bearbeitet werden, stammen aus den unterschiedlichsten Epochen und Entstehungszusammenhängen. Neben Wörterbüchern und Goethebriefen sind mittelalterliche Urkunden und Bibelkommentare zu nennen, weiter Inkunabeln, »der Krünitz« (eine Enzyklopädie des 18. und 19. Jahrhunderts), die Werke Goethes und Leibniz', zoologische Literatur des 18. Jahrhunderts (Carl von Linné), der Jerusalemer Talmud sowie Grabinschriften auf jüdischen Friedhöfen. Ungeachtet dieser Vielfalt lassen sich zahlreiche gemeinsame Problemstellungen beschreiben. Gemeinsam ist allen Vorhaben zunächst das allgemeine Ziel, gedrucktes bzw. epigraphisches Quellenmaterial zu digitalisieren (soweit nicht schon, wie im Fall von Leibniz, im Rahmen von bereits vorliegenden Editionen geschehen), zu erschließen und im digitalen Medium einem

größeren Nutzerkreis zugänglich zu machen. Eine derartige Erschließung und Publikation setzt zunächst voraus, dass eine geeignete Strukturdefinition die zu digitalisierenden Daten angemessen beschreibt und den Inhalt in einer Weise erschließt, die den besonderen Möglichkeiten des digitalen Mediums entspricht. Dieser Schritt, für den einschlägige Standards (z. B. TEI – Text Encoding Initiative) zur Verfügung stehen, wurde in den Digitalisierungsprojekten (Repertorium Biblicum, altdeutsche Originalurkunden, Krünitz, Early Zoological Literature) beschreibenden Referaten thematisiert und problematisiert. Die Erschließung für einen größeren Nutzerkreis, dessen Interessen nicht von vornherein bekannt sind, gelingt nur dann, wenn die in digitalisierten Texten vorliegenden Informationen auch tatsächlich aufgefunden werden können. Die Verwendung einer standardisierten oder eigenen DTD (Document Type Definition) garantiert – neben der Weiterverarbeitung und der Präsentation digitaler Quellen mit aktuellen Werkzeugen – eine Erschließung, die die Binnenstruktur eines Dokumentes berücksichtigt und »feldbezogenes« Suchen ermöglicht. Sollen indessen Informationen innerhalb von Strukturen, d. h. innerhalb eines Feldes bzw. Containers, in einem nicht weiter gegliederten Volltext identifiziert werden, reichen die im Vorfeld der Datenaufnahme möglichen Entscheidungen und Maßnahmen nicht aus. Die Suche im »mikroskopischen« Bereich bedarf spezieller Strategien und Hilfsmittel. Demonstriert wurden zum einen Programmkonstrukte, die eine effiziente Suche nach Kollokationen ermöglichen; zum anderen wurde diskutiert, inwieweit eine derartige Suche durch Indizes unterstützt werden kann, die den Inhalt unter Anwendung definierter Regeln (Lemmatisierung, inhaltliche Zusammenhänge) erschließen. Ein anspruchsvoller Sonderfall des *information retrieval* ist die Identifikation von Mishna-Zitaten in Talmud-Texten. Da hier nicht immer wörtlich und vollständig zitiert wird, kann bloßes *pattern matching* in diesem Fall nicht zum Ziel führen. Damit Zitate identifiziert werden können, müssen zitierte Partien zunächst permutiert werden. Die mit dem resultierenden Cluster erzielten Suchergebnisse können dann quantifiziert und bewertet werden.

Prof. Dr. Nicolás Castrillo, der an der Universität Burgos die lateinische Philologie vertritt, hatte die Tagung bereits zum zweiten

Mal ausgerichtet. Für die großen organisatorischen Leistungen und für die überwältigende Gastfreundschaft sei ihm, der Universität Bur-

gos und der Provinzialregierung Castilla y León auch an dieser Stelle herzlich gedankt.

Einige der Referate stehen mit Projekten in Zusammenhang, die über eine eigene Webpräsenz verfügen. Die betreffenden URLs sind im folgenden aufgelistet (und können hier direkt angeklickt werden):

Hebräische Grabsteinepigraphik, Steinheim-Institut, Univ. Duisburg-Essen  
<http://sti1.uni-duisburg.de/projekte/epigraphie/hamburg/intro.xml>

Leibniz-Forschungsstelle, Univ. Münster  
<http://www.uni-muenster.de/Leibniz/>

Krünitz, Univ. Trier  
<http://www.kruenitz.uni-trier.de>

Stiftung Weimarer Klassik  
[http://ora-web.swkk.de:7777/swk-db/db\\_all.html](http://ora-web.swkk.de:7777/swk-db/db_all.html)  
<http://www.swkk.de>

Early Zoological Literature Online, SUB Göttingen  
<http://www.sub.uni-goettingen.de/projekte/ezoolo.html>

Corpus der altdeutschen Originalurkunden, Univ. Trier  
<http://germa83.uni-trier.de/Corpus/>

Digitalisierung des *Repertorium Biblicum Medii Aevi*, Univ. Trier  
<http://www.repbib.uni-trier.de>

Inkunabel-Katalog deutscher Bibliotheken (INKA), UB Tübingen  
[www.ub.uni-tuebingen.de/pro/kata/inkun.php](http://www.ub.uni-tuebingen.de/pro/kata/inkun.php)

Goethe-Wörterbuch, Berlin-Brandenburgische Akademie der Wissenschaften (BBAW)  
[www.bbaw.de/forschung/gwb/](http://www.bbaw.de/forschung/gwb/)

*Matthias Kopp*  
[kopp@zdv.uni-tuebingen.de](mailto:kopp@zdv.uni-tuebingen.de)