

# TextGrid – Eine Geschichte

von Werner Wegstein, Andrea Rapp und Fotis Jannidis

## Vorbemerkung

In guter philologischer Tradition gliedern wir ‚TextGrid – Eine Geschichte‘ in ‚Die Vorgeschichte‘, ‚Das Projekt und seine Phasen‘ und den visionären Ausblick ‚Perspektiven‘. Mit der Dreiteilung ergibt sich die Verantwortung für die verschiedenen Teile von selbst. Da einige der frühen Erfahrungen ihren Ursprung an der Universität Würzburg haben, fällt dieser Teil auf mich, das ‚Projekt und seine Phasen‘ auf Andrea Rapp, weil sie schon an den Anfängen in Göttingen beteiligt war und anschließend alle Projektphasen von TextGrid mitgestaltet hat. Fotis Jannidis ist für den visionären Ausblick prädestiniert, denn er hat im Jahr 2007, auf der TEI-Jahrestagung *TEI@20: 20 Years of Supporting the Digital Humanities* in College Park, Maryland, U.S., schon einmal in die Glaskugel geblickt, Visionäres gesehen und in Worte gefasst, die man nicht so schnell vergisst, z. B. die verlassenen XML-Bäume in den TEI-Kodierungswäldern.<sup>1</sup> Eine Vision, wie es mit TextGrid nach zehn Jahren als Virtueller Forschungsumgebung weitergeht, wird ihm daher leichtfallen.

*Werner Wegstein*

*Ein Mann hat eine Erfahrung gemacht,  
jetzt sucht er die Geschichte seiner Erfahrung ...*

Max Frisch<sup>2</sup>

## Die Vorgeschichte

Der Anstoß geht von Prof. Dr. Wolfgang Riedel aus, seinerzeit Dekan der damaligen Philosophischen Fakultät II der Universität Würzburg. Im April

---

<sup>1</sup> Keynote presentations 1.3 *TEI in a crystal ball* (<http://www.tei-c.org/Vault/Members-Meetings/2007/program.html#jannidis>)

<sup>2</sup> Das aufgegebene Thema erinnert mich an Gedanken des Erzählers im Anfangsbereich des Romans ‚Mein Name sei Gantenbein‘ von Max Frisch (1964), S. 8 und S. 11 – daher mein Motto.

2003 lädt er für einen Workshop am 11. Juli 2003 nach Würzburg ein zum Thema: ‚Zukunft von TUSTEP und Perspektiven philologischer EDV‘.

Anlass ist die ungewisse Zukunft von TUSTEP. Wird es (in Tübingen oder woanders) weiterentwickelt? Wird es auf dem bis dato erreichten Entwicklungsstand ‚eingefroren‘ und in der gegebenen Form (und wie lange dann) weiterverwendet? Gibt es Alternativen zu TUSTEP? Die Würzburger Germanistik, die seit langem und intensiv mit TUSTEP arbeitet, betrachtet diese Ungewissheiten mit einiger Sorge. Es erscheint uns ratsam, die einschlägig kompetenten und engagierten Kollegen aus dem Fach einmal zu einem Meinungsaustausch und Brainstorming zusammenzurufen.<sup>3</sup>

Es folgt eine Liste von Fragen, über die diskutiert werden sollte, sowie die Namen bereits Eingeladener und ein Dank vorab für weitere Einladungsvorschläge.

Der Workshop am 11. Juli 2003 ist ein Erfolg. 33 Wissenschaftler (v. a. aus Würzburg, Tübingen, Trier) und der Schweiz (Zürich) nehmen teil, und diskutieren die skizzierten Fragen ausführlich und weitestgehend einmütig. Als Ergebnis ist festzuhalten:

1. Die Verdienste von TUSTEP sind unstrittig: die stolzen Bilanzen der Würzburger Forschergruppen *Prosa des deutschen Mittelalters* und *Das Bild des Krieges im Wandel vom späten Mittelalter zur frühen Neuzeit* sowie des Sonderforschungsbereichs 226 *Wissensorganisierende und wissensvermittelnde Literatur im Mittelalter* wären ohne TUSTEP nicht möglich gewesen. Die Erfahrungen daraus sind von Anfang an auch in die Lehre eingeflossen und haben die beiden Aufbaustudiengänge in Würzburg, ‚Linguistische Informations- und Textverarbeitung‘ und ‚EDV-Philologie‘ beeinflusst. Für die Organisation der Weiterentwicklung in Form einer Evolution wird ein Netzwerk empfohlen. Die leitende Frage bleibt: Wie wird philologische Software weiterentwickelt?
2. Die Diskussion am 11.07.03 zeigt auch, dass bislang nirgends erwogen worden ist, organisatorisch und technisch eine Lösung zu suchen. Deshalb erfährt der Vorschlag, an der Universität Würzburg ein Kompetenzzentrum ‚EDV-Philologie‘ einzurichten, Zustimmung von allen TeilnehmerInnen des Workshops. In Würzburg sind sowohl das Know-how in Forschung und Lehre als auch das Bedürfnis nach der Weiterentwicklung philologischer Software vorhanden.

---

<sup>3</sup> Der Abdruck des Briefftextes erfolgt mit Zustimmung des Verfassers.

Als Konsequenz des Arbeitstreffens beantragt das Institut für deutsche Philologie im Juli 2003 die Einrichtung eines Kompetenzzentrums für EDV-Philologie am Institut für deutsche Philologie in Verbindung mit einer entsprechenden Professur.

An dem Wochenende um den 7./8. November 2003 kann das TEI Members Meeting<sup>4</sup> in dem nicht allzu weit entfernten Nancy genutzt werden, seit 1990 bestehende Kontakte zur TEI-Community zu vertiefen und in Nancy weitere Ansichten zu den in Würzburg diskutierten Fragen einzuholen.

Zum 1. Dezember 2003 richtet die Universität Würzburg das ‚Kompetenzzentrum für EDV-Philologie‘ am Institut für deutsche Philologie ein und beauftragt mich mit der Vertretung einer Professur für EDV-Philologie.

Zum 19. März 2004 lädt nunmehr das Kompetenzzentrum für EDV-Philologie zum nächsten <philtag>-Arbeitstreffen ein. Nach dem Erfolg des Workshops im Juli 2003 bekommt dieses Arbeitstreffen damit die Nummer 2. (Die Nomenklatur schwankt eine Weile je nach Sprache der Veranstalter und Teilnehmer zwischen Workshop und Arbeitstreffen.) Es melden sich 20 Teilnehmende an.<sup>5</sup> Zwölf TeilnehmerInnen registrieren sich für die geplante Projektgruppe <philtag>.

Das Programm besteht aus sechs Vortragsblöcken und einer Diskussionsrunde. Eingebunden sind Informatik-Studierende aus Darmstadt (Jannidis) und Würzburg (Seipel), sowie Informatikstudierende aus Würzburg mit Nebenfach Linguistik (Wegstein).

---

4 <http://www.tei-c.org/Membership/Meetings/2003/mm22.xml>

5 Neben den elf Referenten und ihren Betreuern sind dies: Hans-Werner Bartz, Deutsches Wörterbuch Projekt, Univ. Trier; Dr. Thomas Burch, Kompetenzzentrum Elektronisches Publizieren, Univ. Trier; Conrad Drescher, Informatik, Univ. Würzburg; Dipl.-Inf. Marbod Hopfner, Informatik, Univ. Tübingen; Dr. Stephan Moser, Kompetenzzentrum EDV-Philologie, Univ. Würzburg; Martin Omasreiter, Informatik, Univ. Würzburg; Dr. Klaus Prätor, Berlin-Brandenburgische Akademie der Wissenschaften, Berlin; Dr. Andrea Rapp, SUB Göttingen; Dr. Peter Stahl, Kompetenzzentrum EDV-Philologie, Univ. Würzburg.

*Die Vorträge:*

1. Michael Stolz (Basel): Computergestütztes Kollationieren. Ein Werkstattbericht aus dem Basler Parzival-Projekt
2. Dietmar Seipel / Michael Wenger (Informatik, Würzburg): Textvergleich
3. Fotis Jannidis/Alexander Dotor/Thomas Ries (TU Darmstadt): Automatisches Tagging von Strukturen (Inhaltsverzeichnisse, Register, ...)
4. Werner Wegstein / Thomas Hemmerich / Florian Thienel (Universität Würzburg):
  - a. Archivierung von TUSTEP-Dateien im XML-Format (Hemmerich)
  - b. Editor's workbench: Kollationierungsarbeitsplatz mit XMetal und Grafik (Thienel)
5. Hans-Walter Gabler (Universität München): Werkzeuge für die Edition von Virginia Woolfs Roman ‚To the Lighthouse‘
6. Michael Nedo (Wittgenstein-Archiv, Cambridge): EDV-Werkzeuge für die Wittgenstein-Edition
7. Diskussionsrunde über Procedere und Präzisierung der Rahmenbedingungen
  1. Java +/- Eclipse
  2. Votum für Plattformunabhängigkeit und Open Source. Diskussion über die Nachteile von Eclipse; angeregt wird eine Untersuchung der Abhängigkeit und die Suche nach Alternativen zu Eclipse.
  3. XML-Dateiformat
  4. Kodierungsminimalia
  5. Einbindung anderer Module
  6. Integration weiterer Programmier- und Skriptsprachen (Prolog soll geprüft werden.)
  7. Test-Rahmen
  8. Dokumentationsstandards
  9. Stand der Arbeit in Würzburg

Zum Abschluss wird festgehalten, dass Termine im März aus verschiedensten Gründen ungünstig sind, und es wird angeregt, die Arbeitstreffen in den Herbst zu verlegen.

7.10.2004: TextGrid-Besprechung in Würzburg Zimmer 4E13:

Wenn ich mich recht entsinne, haben wir in Würzburg damals über die Zusammensetzung der Antragsgruppe diskutiert und versucht, eine gute Balance zu finden. Die Informatik im Kernbereich ist mit DAASI International GmbH (z. B. zur Anbindung an die Middleware) und mit der Saphor GmbH, die noch etwas später dazukommt, vertreten. Aber auch die FachwissenschaftlerInnen kommen für bestimmte Aufgaben nicht ohne InformatikerInnen in ihrem Teilprojekt aus. Die richtigen dafür zu finden, wird nicht einfach. Die Fachwissenschaft (Germanistik) selbst ist in guter Breite vertreten; die Mediävistik mit Andrea Rapp, die Neuere Deutsche Literaturwissenschaft mit Fotis Jannidis, die Sprachwissenschaft mit jemandem wie mir, der auch ein verkappter Mediävist sein könnte – wie viele Sprachwissenschaftler mit einem Schwerpunkt in der Sprachgeschichte. Rückblickend zeigt sich aber: Keiner von uns war von Dauer und ist heute noch dort, wo er war, als wir begannen. Unserem gemeinsam gestarteten Projekt hat das aus meiner Sicht nicht geschadet. Immerhin ist die SUB immer noch in Göttingen und das IDS in Mannheim, *stabilitas loci* – fast schon etwas Benediktinisches in all dem Wandel, aber für Langzeitprojekte aus meiner Sicht eine wichtige Grundlage für den administrativen Kern.

Außerdem bringen unsere Institutionen, die auf Dauer angelegt sind, Spezielles ein: die SUB natürlich die Metadaten, Katalogstrukturen und alles, was mit Buch im weitesten Sinn zu tun hat, und das IDS, die Institution, die auf Sprache (vor allem, aber nicht nur auf Gegenwartssprache) spezialisiert ist, bringt z. B. auch die Korpuslinguistik, die Lexikografie und weitere interessante Bereiche mit ein. Außerdem wurden 1971 und 1973 am IDS die ersten beiden Tagungen mit dem Titel ‚Maschinelle Verarbeitung altdeutscher Texte‘ veranstaltet, in Deutschland der Einstieg in die Textanalyse von Sprache mithilfe der Datenverarbeitung (vgl. Lenders/Moser 1978a; 1978b).

Das Einzige, dessen ich mir in meiner Erinnerung ganz sicher bin, ist die Tatsache, dass ich nach dieser Besprechung Heike Neuroth und Joachim Migl vom Hubland-Campus der Universität in Würzburg unter Nutzung aller Schleichwege und Abkürzungen so schnell (und ohne Strafzettel) an den Würzburger Bahnhof bringen konnte, dass sie den Zug nach Göttingen noch erreicht haben, obwohl wir bis zur letzten Minute zu diskutieren hatten.

Die Verlegung der <philtag>-workshops in den Herbst wird noch 2004 vollzogen. Als Frucht der TEI-Tagung in Nancy können Lou Burnard und Laurent Romary als Referenten gewonnen werden.

8./9.10.2004: philtag 3 in Würzburg TEI Workshop, 8.–9. Oktober 2004 mit Lou Burnard und Laurent Romary:<sup>6</sup>

Alle Details sind auf der Website festgehalten: die Workshop-Ziele, eine Leseliste einschlägiger Literatur, mit der man sich zuvor beschäftigen sollte, das Programm und die TeilnehmerInnen. Interessant erscheint im Rückblick die dichte Kreuzung von Aktivitäten: Donnerstag trifft sich die Würzburger philtag-Initiative nach dem Mittagessen und vor Lou Burnards Vortrag über XAIRA im Toscana-Saal der Würzburger Residenz und am Samstag wird der Spätnachmittag genutzt für ein Treffen des Technical Committee of the German Working Group on Philosophical Editions. Im Rundbrief Nr. 20 & newsletter 12 der Arbeitsgemeinschaft philosophischer Editionen der Deutschen Gesellschaft für Philosophie e.V. vom Dezember 2004, S. 3–4, findet sich dazu ein ausführlicher Bericht von Klaus Prätör.

Am 21.10.2004 wurde der Antrag beim BMBF eingereicht. Im Sommer 2005 waren noch Rückfragen von Gutachtern zu beantworten.

14./15.10.2005: philtag 4 in Würzburg mit Lou Burnard und Sebastian Rahtz (Oxford University):<sup>7</sup>

Alle Details sind auf der Website festgehalten: die Workshop-Ziele, eine Leseliste einschlägiger Literatur, mit der man sich zuvor beschäftigen sollte und das Programm. Die TeilnehmerInnenliste ist leider (schon in Würzburg) verloren gegangen. Einer der Gründe dafür mag der technische GAU gewesen sein, der in Würzburg am 15.10.2005 zuschlug. Im Laufe des Vormittags brachen im Abstand von 30 Minuten nacheinander beide Beamer, die zur Verfügung standen, zusammen. Die Glühbirnen versagten. Zusätzliche Reservebirnen hatten wir damals nicht.

25.10.2005: Das Bundesministerium für Bildung und Forschung bewilligt unseren Antrag. Freude macht sich breit. Jetzt kann es losgehen.

*Werner Wegstein*

---

6 <http://www.tei-c.org/Talks/2004/Wuerzburg/programme.html>

7 <http://projects.oucs.ox.ac.uk/teiweb/Talks/2005/Wuerzburg/>

## Das Projekt und seine Phasen

„Die Geschichte, plur. ut nom. sing. von dem Zeitworte geschehen.

- 1) Was geschehen ist, eine geschehene Sache, so wohl in weiterer Bedeutung, eine jede, so wohl thätige als leidentliche Veränderung, welche einem Dinge widerfähret, als auch in engerer und gewöhnlicherer, von verschiedenen mit einander verbundenen Veränderungen, welche zusammen genommen ein gewisses Ganze[s!] ausmachen (...). Eine wahre Geschichte, im Gegensatze der erdichteten. In engerer Bedeutung führet nur die erstere den Nahmen der Geschichte. (...)
- 2) Die Erzählung solcher Geschichte oder geschehenen Begebenheiten; die Historie. (...)<sup>8</sup>

So wird ‚Geschichte‘ in Adelungs Wörterbuch erläutert, das hier nach der Version im *Trierer Wörterbuchnetz* zitiert ist, die auf den vom Bundesministerium für Bildung und Forschung gekauften und seitdem frei verfügbaren Daten der *Digitalen Bibliothek* von [www.zeno.org](http://www.zeno.org) basiert. Sowohl das Wörterbuchnetz als auch die *Digitale Bibliothek* sind Teil der mehr als zehn Jahre umfassenden TextGrid-Geschichte, die hier erzählt wird. Zehn Jahre bedeuten nicht nur in der digitalen Welt eine Zeit intensiven Wandels; dies spiegelt der vorliegende Band sehr deutlich. Daher ist es nicht die Aufgabe unseres Beitrags, die TextGrid-Geschichte so zu erzählen, dass ein „gewisses Ganzes“ entsteht; dies leisten vielmehr neben den zahlreichen TextGrid-Publikationen und -Berichten vor allem die TextGrid-Produkte und die Erfahrungen der EntwicklerInnen NutzerInnen. Unser Blick auf die Geschichte soll vielmehr ein ganz persönlicher und auf die „nicht herausgegebenen“ Rahmenbedingungen, Meilensteine, Schlüsselmomente und Veränderungen sein, die die Geschichte von TextGrid ausmachen.

---

8 Johann Christoph Adelung: *Grammatisch-kritisches Wörterbuch der Hochdeutschen Mundart mit beständiger Vergleichung der übrigen Mundarten, besonders aber der oberdeutschen*. Zweyte, vermehrte und verbesserte Ausgabe. Leipzig 1793–1801, s. v. ‚Geschichte‘, Bd. 2, Sp. 605 f. [http://woerterbuchnetz.de/cgi-bin/WBNetz/wbgui\\_py?sigle=Adelung&lemid=DG01581](http://woerterbuchnetz.de/cgi-bin/WBNetz/wbgui_py?sigle=Adelung&lemid=DG01581); Imagedigitalisat: <http://www.zeno.org/Adelung-1793/K/adelung-1793-02-0605>; <http://www.zeno.org/Adelung-1793/K/adelung-1793-02-0606>.

Das TextGrid-Konsortium bildete eine außergewöhnliche, aber im Rückblick sehr gut funktionierende Mischung: Die beteiligten Philologinnen mit unterschiedlicher Schwerpunktsetzung (s. o.) vertrauten sich in der Leitung des Projekts einer erfahrenen Bibliothek an, hinzu kamen zwei Unternehmen mit technologischer, aber auch philologischer Expertise. In den folgenden Phasen wurde dieser Kreis erweitert durch weitere fachwissenschaftliche Ausdifferenzierung wie Klassische Philologie, Judaistik oder Musikwissenschaften sowie durch Rechenzentren als zentrale Infrastrukturpartner.<sup>9</sup> Obwohl die Initiativgruppe vordergründig einen heterogenen Eindruck machte, gab es für einen großen Teil der Projektpartner einen gemeinsamen Hintergrund, sie hatten bereits eine gemeinsame Geschichte – nicht zuletzt im TUSTEP-Kontext, in der sie Textwissenschaften mit dem Einsatz und der Entwicklung digitaler Technologie verbunden hatten. Einer der Schlüsselmomente des Projektes war sicherlich die erste Antragsbesprechung des Konsortiums in Raum 4E13 in Würzburg am 7.10.2004 (s. o.), bei der die Kerngruppe – deren genaue Zusammensetzung bis dahin nicht allen bekannt war – zusammentraf und überrascht feststellte, dass alle sich in unterschiedlichen Konstellationen kannten und bereits in irgendeiner Form zusammengearbeitet hatten. Offenheit und ein aus Erfahrungen resultierendes Grundverständnis für die jeweils anderen Forschungskulturen waren also vorhanden.

Dennoch erforderte die Zusammenarbeit im Projekt eine Phase, in der eine gemeinsame, die Disziplinen überbrückende Sprache und eine spezifische gemeinsame Projektkultur entwickelt werden mussten. Auf dieser Basis gelang es, Kontinuität mit Flexibilität und Entwicklungsfähigkeit zu verbinden und über den für Projekte sehr langen Zeitraum von zehn Jahren vertrauensvoll zusammenzuarbeiten. Überhaupt scheint mir diese spezifische Projektkultur – nicht nur Aspekte aus den Geisteswissenschaften und der Informatik zusammenfügend, sondern ein neues Drittes bildend – durchaus kennzeichnend für die Digital Humanities zu sein und damit etwas Verbindendes zu stiften in diesem Disziplinbündel, das – kaum definiert – sich bereits entfaltet und ausdifferenziert. Auch aus diesem Blickwinkel scheint TextGrid ein Digital-Humanities-Modellprojekt zu sein, von dem mehr bleiben wird als eine Software-Entwicklung, da es ein wesentlicher Motor der Community-Entwicklung und Kulturbildung ist.

---

9 Vgl. dazu auch die Beiträge von Joachim Veit, Philipp Wieder und anderen in diesem Band.



Auch die Förderer im Bundesministerium für Bildung und Forschung schenkten der Gruppe Vertrauen und ließen sich auf das Experiment ein, den Geisteswissenschaften Zugang zu avanciertester Technologie zu ermöglichen. Die Bereitschaft zu referatsübergreifender Kommunikation und Zusammenarbeit kann wohl ebenfalls kaum hoch genug eingeschätzt werden.

Technologische Rahmenbedingungen wurden zunächst von der durch die Weitsicht Elmar Mittlers<sup>10</sup> ermöglichten Einbindung in den D-Grid-Verbund gesetzt, was einerseits Folgen für technologische Grundlagen und Architekturentscheidungen, andererseits aber auch für die Kommunikation und für die Vermittlung der TextGrid-Idee in die geisteswissenschaftlichen Communitys hatte. Sie wurde häufig als „zu naturwissenschaftlich“ empfunden und war in den Geisteswissenschaften nicht immer einfach zu erläutern. Mittlerweile sind jedoch viele Aspekte der Grid-Philosophie in allen Wissenschaftszweigen angekommen: digitale Forschungsinfrastrukturen, netzbasiertes und kollaboratives Arbeiten, Datenoffenheit, Nachnutzbarkeit usw. Zugleich konnten die Mitglieder des TextGrid-Konsortiums vom Austausch mit den anderen natur- und ingenieurwissenschaftlichen Grid-Communitys profitieren, wie überhaupt der disziplinübergreifende Blick auf den Technologiebedarf bereichernd für alle Beteiligten war.<sup>11</sup> Ein weiterer Meilenstein für die TextGrid-Initiatoren war daher sicherlich die Zwischenbegutachtung des D-Grid-Verbunds am 1. und 2. März 2007 in Berlin im Konrad-Zuse-Zentrum, bei der wir den Eindruck gewonnen hatten, „angekommen zu sein“ – und das in doppelter Perspektive: Die geisteswissenschaftlichen Angebote, Fragestellungen und Bedarfe stießen nicht nur auf Verständnis und Interesse bei den übrigen Communitys, sondern wurden als Bereicherung erkannt, während wir unsererseits die naturwissenschaftlichen Infrastruktur-Perspektiven immer besser integrieren konnten.

Bereits in den ersten vorsichtigen Projektideen und Konzepten spielten frei verfügbare, vor allem qualitätvolle digitale Inhalte eine wesentliche Rolle. Qualität definierte sich nach den Kriterien von inhaltlicher Qualität (aktuelle und wissenschaftlich akzeptierte kritische Referenzangaben), Erfassungsqualität (Erfassungsgenauigkeit, Fehlerrate, Vollständigkeit, Zitierfähigkeit), Datenformat und Aufbereitungsqualität (Standards, XML/TEI, Metadaten) sowie Archivierbarkeit. Allerdings standen diese Dinge in der

---

10 Siehe dazu auch sein Interview in diesem Band.

11 Diese Aspekte wurden systematisch weiterverfolgt und dokumentiert im Projekt *WissGrid – Grid für die Wissenschaft*; s. [www.wissgrid.de](http://www.wissgrid.de) (26.01.2015).

ersten Projektphase nicht im Fokus, da diese Phase sich auf die Entwicklung einer Workbench für textwissenschaftliche (insbesondere editionswissenschaftliche) Arbeits- und Forschungsprozesse konzentrierte. Gleich zu Beginn wurden jedoch die Wörterbücher des Trierer Wörterbuchnetzes eingebunden.<sup>12</sup> Die Wörterbuchdaten wurden zu eng mit anderen Werkzeugen verknüpften Services (vgl. Büdenbender/Leuk 2009).

Die Inhalte rückten während der Projektlaufzeit immer stärker ins Blickfeld und es wurde deutlich, dass der zuverlässige und langfristige Zugang dazu ein entscheidendes Kriterium für die Weiterentwicklung und die Akzeptanz der Digital Humanities sein würde. Die Community artikulierte ganz dezidiert den entsprechenden Bedarf nach dem TextGrid Repository. Ein besonderer Glücksfall und wiederum wichtiger Meilenstein war daher der Kauf der Daten der *Digitalen Bibliothek*, die seitdem in Würzburg weiter aufbereitet werden und für Forschung und Lehre frei zur Verfügung stehen.<sup>13</sup> Erst wenn genügend standardisierte und offene Daten, die Geisteswissenschaftlerinnen und Geisteswissenschaftler für ihre Forschung benötigen, zur Weiterverarbeitung und Nachnutzung zur Verfügung stehen, werden Digital-Humanities-Methoden in der Mitte der Fächer ankommen. Gerade auch in der Lehre ist ein solcher freier Zugang zu verlässlichen Daten, mit denen man Studierenden auch Qualitätsaspekte nahe bringen kann, von immenser Bedeutung.

Als glücklichen Umstand sehe ich schließlich, dass die Entwicklung von TextGrid parallel mit der Entwicklung und der Etablierung der Digital Humanities erfolgte und dass es uns ermöglicht wurde, TextGrid immer stärker als Forschungsprojekt zu profilieren, wenn auch der Infrastrukturaspekt nie aus den Augen geriet. Auf diese Weise kann – so hoffen die Initiatoren – der Anspruch verfolgt werden, als Forschungsinfrastruktur „wesentliche Beiträge zum wissenschaftlichen Erkenntnisgewinn“ zu liefern, zu „Inkubatoren für neue und innovative wissenschaftliche Fragestellungen“ zu werden und „ganz neuartige Möglichkeiten der forschenden Erschließung von Bibliotheks-, Archiv- und Sammlungsbeständen“ zu bieten sowie „eine Kultur des internationalen und zunehmend auch interdisziplinären Austauschs“ zu fördern.<sup>14</sup>

*Andrea Rapp*

---

12 [www.woerterbuchnetz.de](http://www.woerterbuchnetz.de) (26.01.2015).

13 Siehe hierzu den Beitrag von Katrin Betz im vorliegenden Band.

14 Empfehlungen des Wissenschaftsrates zu Forschungsinfrastrukturen in den Geistes- und Sozialwissenschaften, 28.1.2011, Drs. 10465-11, S. 7: <http://www.wissenschaftsrat.de/download/archiv/10465-11.pdf>.

## Die Zukunft TextGrids

Das Projekt TextGrid wird 2015 nach neunjähriger Förderung abgeschlossen sein. Wie steht es um seine Zukunft? Auf diese Frage ist keine einfache Antwort möglich. Anders als bei vielen anderen Projekten haben sich die Verantwortlichen engagiert, ihr Projekt über dessen Förderlaufzeit hinaus weiter zu betreiben. Dazu haben sie sich in einem Verein zusammengeschlossen, dessen vornehmliche Aufgabe die „persistente Sicherstellung des fachwissenschaftlich nachhaltigen Gebrauchs der Angebote der TextGrid-VFU in einer heterogenen Community wissenschaftlicher Nutzerinnen und Nutzer“<sup>15</sup> ist. Und gerade Institutionen wie das Institut für Deutsche Sprache und die Staats- und Universitätsbibliothek Göttingen, die plausibel Nachhaltigkeit versprechen können, sind dabei in führender Position engagiert. Zugleich ist TextGrid ein Modul in dem europäischen Infrastruktur-Projekt DARIAH-EU, das zurzeit als DARIAH-DE in Deutschland wiederum mit Förderung des BMBF umgesetzt wird.<sup>16</sup>

Das klingt alles sehr gut, aber es wird meines Erachtens nichts an der Tatsache ändern, dass TextGrid in einigen, wahrscheinlich wenigen Jahren verschwunden sein wird. Und das ist gut so. Lassen Sie mich erläutern, wie ich das meine. Schon seit einiger Zeit ist es klar geworden, dass DH-Institutionen wie TEI oder eben auch TextGrid aus mehreren Komponenten bestehen: zum einen aus der Software, und hier kann man im Fall von TextGrid noch zwischen dem Repository, in dem die Daten liegen und mit einem Persistent Identifier versehen sind, der API und den Clients unterscheiden. Zum anderen handelt es sich um eine Community von Geisteswissenschaftlern, Vertretern von Bibliotheken und Rechenzentren sowie Informatikern. Und schließlich handelt es sich um Standards und Best Practices, die im Laufe des Projekts entwickelt oder adaptiert wurden. Wenn ich nun sage, dass TextGrid in einigen Jahren verschwunden sein wird, dann bezieht sich das vor allem auf die Community und die Clients. Die Community wird sich, gerade weil TextGrid Teil einer umfassenderen Infrastruktur wird, ändern und daran adaptieren. Für die Clients muss ich ein wenig weiter ausholen.

Zu den wichtigsten Lehren, die zumindest ich aus meinen Erfahrungen mit der Entwicklung von TextGrid ziehen kann, gehört die Einsicht, dass

---

15 <http://www.textgrid-verein.de/>

16 DARIAH-DE (<https://de.dariah.eu/>) ist ein Teil der gesamteuropäischen Initiative (<https://www.dariah.eu/>).

digitale Infrastrukturen einem besonderen Entwicklungsdruck unterliegen – und das gilt natürlich umso mehr, wenn die angezielte Nutzergruppe selbst sich in einem so grundlegenden Wandel befindet, wie das in den Digitalen Geisteswissenschaften der Fall ist. Das Design einer Infrastruktur muss sich unter diesen Umständen auf die Extrapolation von Entwicklungstendenzen in eine doch weitgehend ungewisse Zukunft stützen. Und nicht nur das Verhalten der Anwender unterliegt Wandlungsprozessen, auch die Techniken zur Lösung von bestimmten Problemen ändern sich. In den Jahren, in denen wir an TextGrid gearbeitet haben, hat sich etwa der Browser als Arbeitsplattform für Anwendungen im Bereich des PC etabliert – und dann ist auch diese Technik wieder unwichtiger geworden, da im Rahmen der mobilen Geräte eine neue Kultur stark spezialisierter Programme Mode wurde.

Ein Grundproblem von Software im Bereich der Digitalen Geisteswissenschaften scheint gleich zu bleiben: Wie kann man die weniger technikaffinen Anwender komplexe Arbeiten ausführen und sich dabei ganz auf ihre fachwissenschaftliche Fragestellung konzentrieren lassen, ohne sie in ein Korsett zu zwingen, das der jeweiligen Fragestellung erst mühsam angepasst werden muss? Doch mögliche Antworten auf diese Frage sind von den größeren Paradigmen im Umgang mit Computern in all ihren Formen abhängig. Anders gesagt: Eine gelungene Lösung zu finden angesichts von sich ändernden Anwendererwartungen und -kompetenzen, von Techniken und Standardlösungen, die stets im Fluss sind, kann nur das Ergebnis eines längeren Forschungsprozesses sein, der nie endgültig abgeschlossen sein wird. Die Rolle der Fachwissenschaftler wird sich hierbei wohl vor allem darauf beschränken, die wissenschaftlichen Anforderungen zu formulieren, die Prototypen auf ihre Funktionalität zu testen und die Institutionen, die Dauer versprechen, dabei zu unterstützen, in die richtige Richtung weiterzuentwickeln.

Diese Teile von TextGrid werden sich also auflösen, werden teilweise eingehen in andere Strukturen, werden teilweise obsolet werden und werden evtl. zum Teil Grundlage für eigenständige neue Entwicklungen sein. Dass bei all der hierbei notwendigen Flexibilität und Beweglichkeit einige Dinge doch sicher und verlässlich bleiben – nicht zuletzt die Daten der Nutzer, die publizierten Daten und ihre URLs, kurzum das Repository –, dafür stehen die Betreiber von TextGrid.

Es ist sehr zu hoffen, dass TextGrid dabei auch zum Anlass wird, die Förder- und Unterhaltsstruktur in Deutschland zu überdenken. Zurzeit sehen sich alle einschlägigen Fördereinrichtungen in Deutschland nicht in der Lage, dauerhafte Infrastrukturen nach dem Auslaufen der Projektförderung zu fi-

nanzieren. Das Problem ist allen Beteiligten bekannt, aber die politischen und finanziellen Herausforderungen, die damit verbunden sind, erlauben bislang keine einfache, klare Lösung. Die Zeit, in der Fachwissenschaftler versuchen konnten, ihre Bedürfnisse im Rahmen von Projekten durch den Aufbau von ‚Infrastrukturprojekten‘ zu befriedigen, scheint auf jeden Fall vorbei zu sein. Die Notwendigkeit, solche Projekte nachhaltig zu gestalten, erfordert die Kooperation mit Institutionen, die – durchaus mit sehr guten Gründen – ganz anders arbeiten, als es der ungeduldige Fachwissenschaftler tut. Für diesen Prozess, der einerseits die Einsicht des Wissenschaftlers braucht, welche Werkzeuge fehlen, und der andererseits diesen übergreifende Stabilität abverlangt, gibt es einige Modelle, z. B. die Open-Source-Entwicklung oder die Webservice-Architekturen, aber hier ist der richtige Weg, Kreativität und stabile Organisation zu verbinden, wohl noch nicht gefunden.

Aber natürlich ist TextGrid nicht nur ein Forschungsprojekt in Sachen Virtuelle Forschungsinfrastruktur in den Geisteswissenschaften – TextGrid ist auch Infrastruktur, deren Nutzer in diesem Kontext das Arbeiten mit digitalen Werkzeugen lernen und weiterentwickeln. Ihre Anforderungen, die sich aus der Auseinandersetzung mit TextGrid und den Forschungsplänen für ihr Material ergeben, werden ihre Umgangsweise und Ansprüche an alle folgenden Werkzeuge prägen. Auch in dieser Form wird TextGrid also weiterleben.

*Fotis Jannidis*

## **Literaturverzeichnis**

- Büdenbender, Stefan; Leuk, Michael (2009): Daten als Dienste: Wörterbücher als Erschließungsinstrumente in der virtuellen Arbeitsumgebung „TextGrid“. In: *it – Information Technology* (Themenheft „Informatik in den Geisteswissenschaften“) 51 (4): 191–196.
- Lenders, Wilfried; Moser, Hugo (Hrsg.) (1978 a): *Maschinelle Verarbeitung altdeutscher Texte. 1 Beiträge zum Symposium Mannheim, 11./12. Juni 1971*. Berlin: E. Schmidt.
- Lenders, Wilfried; Moser, Hugo (Hrsg.) (1978 b): *Maschinelle Verarbeitung altdeutscher Texte. 2 Beiträge zum Symposium Mannheim, 15./16. Juni 1973*. Berlin: E. Schmidt.