

I11-1#Interviewer: Gut. Als erstes würde ich gerne wissen, wie sie ihre EDV-philologische Erfahrung beschreiben würden und ihre Aktivitäten in dem Bereich bisher. Befragte/r: Na ja, ich bin hier im Rechenzentrum gewissermaßen die Schnittstelle für den EDV-Einsatz in den Geisteswissenschaften und würde sagen, dass ich ex officio einiges an Erfahrungen mitbringe im Umgang mit solchen Themen – ganz unbescheiden. **I:** Ja, natürlich, also gut, das war jetzt öfter bei den Interviews so. Das sind natürlich viele Interviewpartner, die damit... natürlich viel Erfahrung mitbringen. Können sie da noch ein bisschen genauer darauf eingehen, was sie genau in dem Bereich machen? **B:** So buchstäblich alles, was in irgendeiner Weise mit Sprache, Geschichte, Geschichte der Sprache, Philologie, Linguistik... grundsätzlich überall, auf allen Gebieten würde ich sagen, in denen Text jeglicher Form zum Objekt wissenschaftlichen Arbeitens gemacht wird. Dort werde ich um Rat angegangen. Dort bringe ich auch gewisse Kompetenz mit. [Telefon klingelt] Entschuldigung. Ich gehe mal kurz dran. [kurzes Telefonat] So, jetzt können wir weitermachen. Also wir werden gleich nochmal unterbrochen. **I:** Wir können auch solange noch warten, also das ist... **B:** Ich habe keine Ahnung, wann die kommt. Kann durchaus sein, dass die dann erst, was weiß ich, in einer Stunde oder so aufkreuzt. **I:** Ach so. Nein, so lange wird es, glaube ich, nicht dauern. Also, gut. **B:** Also, um nur zu sagen: es betrifft, also die Einschlägigkeit der Arbeiten, das kann eine Urkundenverwaltung sein. Das kann die Verwaltung von Gerichtsurteilen sein. Das kann, also jede Form von Text-Datenbankverwaltung ist damit gemeint. Also Bibliographieverwaltung genauso, wie z.B. das Quantifizieren oder Analysieren von Texten. Textvergleichen, also wenn es darum geht, Textzeugen miteinander zu vergleichen oder Textausgaben, um eine Textgenese nachzuvollziehen. Das gehört ganz genauso dazu, wie eine kritische Edition usw. usw. usw. Buchstäblich alles, was in irgendeiner Weise, wie ich schon sagte, die allgemeine Formulierung dafür - es ist alle Arbeitsbereiche, in denen Text zum Gegenstand wissenschaftlichen Fragens gemacht wird. **I:** Und sie geben ja auch Kurse, also TUSTEP-Kurse zum Beispiel und ähnliches. Gut.

I11-2#Ja, dann würde ich sie jetzt gern nach ihrem Gesamteindruck des TextGrid-Lab in der Version, wie es im Workshop getestet wurde, fragen. Was fanden sie gut? Was fanden sie schlecht? Einfach mal übergreifend formuliert, vor dem Hintergrund ihres Erfahrungsschatzes in dem Bereich. **B:** Also da, ich weiß gar nicht genau, wo anfangen, nicht? Also zunächst einmal der Raum – fand ich nicht besonders glücklich gewählt. Aber vermutlich war der gar nicht gewählt, sondern einfach durch unsere Raumvorgabe, so, so... **I:**

Also mir geht es jetzt um das TextGrid-Lab. Was war da ihr Gesamteindruck? B: Ja, gut. Also ein Problem bestand unter anderem darin, dass ich das, was vorgetragen worden ist, teilweise nicht verstanden habe. Der Platz an dem ich saß, ne? Also ich... Stefan Büdenbender hat das wirklich, soweit ich das mitbekommen habe, wirklich sehr gut gemacht. Nur ich habe dann auch mehrfach darauf hingewiesen, dass ich meine Probleme hätte, es zu verstehen. Habe mich dann in der zweiten Hälfte des Kurses ganz nach vorne gesetzt, wo ich das besser verstanden habe. Also manches ist einfach an mir vorbeigegangen, wobei ich hinzufügen muss, dass ich eine gewisse, eine bestimmte Form von Schwerhörigkeit habe. Ich hab vor zehn Jahren einen Hörsturz gehabt. Habe seitdem, bin seitdem von Tinnitus geplagt und habe einfach Filterprobleme. Also das sind die äußeren Lagerungsbedingungen für... Also da – ich bin überzeugt davon, dass manches, was für mich wichtig gewesen wäre und auch für mein Urteil wichtig gewesen wäre, vielleicht einfach nicht mitgekriegt habe. Ne? Das muss man mit in Rechnung stellen. Ja. Ansonsten – ich bin mit sehr hohen Erwartungen reingegangen und insofern ist dann auch mein, mein Gesamteindruck gemischt. Gemischt insofern, als ich, also es sind halt, das haben sie ja, wissen sie ja auch, sie sind ja an der vordersten Front dort tätig. Es gibt zahlreiche Unfertigkeiten. Ne? Also ich, ich hatte den... Also das, was präpariert worden ist, was als Musterdaten geliefert worden ist, das, was vorgestellt worden ist, das hat funktioniert. Aber ich bin überzeugt davon, dass wesentlich mehr und tiefere, sowohl in die Breite als auch in die Tiefe, wesentlich mehr Perspektiven für die Zukunft noch damit verbunden sind. Das war nicht mehr als nur ein erstes hineinschnuppern in den Kosmos der Möglichkeiten, die sich da auftun.

I11-3#I: Ja. Was ist ihnen besonders negativ aufgefallen? Waren das diese Kinderkrankheiten? Oder mal, wenn man die mal ausklammert... Also das, was funktioniert hat, ist ihnen da was besonders negativ aufgefallen, oder...? B: Irgendetwas, was besonders negativ gewesen wäre, nicht. Also, ich hab natürlich auch von vorneherein mit bestimmten Kinderkrankheiten gerechnet. Nur die Vielfalt von Dingen – also wenn man irgendetwas gemacht hat, was nicht genau eine Schritt für Schritt ..., was nicht genau der Schritt für Schritt Anleitung folgte, wenn man irgendetwas nebenher gemacht hat, dann lief man innerhalb kürzester Frist irgendwie gegen die Wand. Nicht? Also, es sind unterm Strich, jeder einzelne Punkt, den ich erwähnen könnte, ist nur eine Kleinigkeit und er lässt sich garantiert auch ausmerzen, aber nach einer Weile hat es ganz einfach genervt. Um nur ein Beispiel zu nennen, nicht, an der Stelle z.B., wo man z.B. eine XML-Auszeichnung einfügen kann. Ich markiere erst – also jeder XML-Editor kann das – ich markiere ein Wort. Ich sage dann, ich will ein bestimmtes Tag einfügen und das muss natürlich um den markierten

Bereich herum eingefügt werden. Das hat es aber nicht – der hat es vorne eingefügt. Nicht? Und nicht drumherum. Also man musste anschließend noch einiges an manueller Arbeit investieren, um das hinzubekommen. Ich zweifle nicht im Geringsten daran, dass sich das ganz leicht und schnell ausmerzen lässt, aber solche und ähnliche Dinge gab es für meinen Geschmack zu viele. Sorry!

I11-4#I: Sie brauchen nicht mir gegenüber ‚Sorry‘ zu sagen. Ich evaluiere. Gut. Ist Ihnen etwas besonders positiv aufgefallen in dem Stadium, in dem das jetzt war? B: Also die Oberfläche scheint mir recht anspruchsvoll und durchdacht zu sein, wobei ich, wobei hier ganz genauso gilt: Es wird sich im Verlaufe der praktischen Arbeit mit dem TextGrid-Lab noch zeigen. Da wird noch, da wird noch einiges zu tun sein. Aber nein, insgesamt hat es mir eigentlich – also ich, die Oberfläche finde ich gelungen. Ja, also ich könnte, ja. **I: Auch was die optische Gestaltung angeht. Oder das Look & Feel sozusagen.** B: Ja, also nach der... Also das Look & Feel. Es ist nicht so, dass man es einfach nur aufrufen konnte und sofort damit klarzukommen. Aber mit..., nach..., man ist ja... Für die ersten Schritte ist man sozusagen bei der Hand genommen worden und es hat eigentlich ganz gut geklappt. Sobald man sich mal auf die Philosophie eingelassen hat, glaube ich, kommt man gut zurecht. **I: Hatten sie mit Eclipse vorher auch schon gearbeitet?** B: Nein, noch nie. Also ich war mal mittelbar damit befasst, als mal jemand, ein Kunsthistoriker hier, sich mit Eclipse beschäftigt hat, aber das war sehr marginaler Natur, dieser Kontakt.

I11-5#I: Ja. Ja wie finden sie denn die... Die Eclipse-Oberfläche bringt ja eine bestimmte Art der Navigation mit sich, der Organisation. Ja, was haben sie davon gehalten? Sind sie damit gut zurechtgekommen oder weniger gut? B: Gut zurechtgekommen, ja. Also wie angemessen das ist, das wird sich erst im fernerer Verlauf noch zeigen. Nicht? Also es sind ja jede Menge Perspektiven für die Zukunft damit verbunden. Nicht? Aber da wird man abwarten müssen, glaube ich, vor einem endgültigen Urteil. Derzeit hat es mir eigentlich ganz gut gefallen. Ja. Also ich glaube schon, dass man, dass man mit einer relativ kurzen Eingewöhnungszeit – und damit meine ich weniger als einen Tag – also, ein EDV-erfahrener Anwender wird damit gut klarkommen. **I: Es bringt ja einige Besonderheiten mit sich, z.B., dass es verschiedene Wege zu ein und demselben Ziel ermöglicht – von der Navigation her. Finden sie das gut, oder finden sie das eher verwirrend?** B: Nein, finde ich gut. Je breiter die Möglichkeiten sind, je mehr Wege nach Rom führen, desto besser. Ne? Also ich bin der Meinung, dass man so wenig wie möglich gegängelt werden sollte. Dass man die Möglichkeit haben sollte, den Weg zu wählen, der einem am einsichtigsten, am transparentesten erscheint. Nicht? Also je mehr Möglichkeiten, desto besser, meine ich. Da

wird auch... Gerade in dieser Hinsicht, glaube ich, wird auch noch viel zu tun sein. Also für die weitere Verarbeitung, wenn dann mal später z.B. ein Indexierungsprogramm eingebaut wird, nicht? Da wird man einfach zahlreiche anwählbare Optionen vorsehen müssen, um zu sagen: Ich will es genau auf diese Art und Weise haben, ne? Also der eine will vielleicht rückläufig sortiert haben, der andere vorläufig. Der andere sagt: Ich will gerne , naja z.B. nach den russischen oder nach den französischen oder nach den italienischen nationalen Konventionen sortiert haben und solche Dinge. **I: Für sie spielt also dann die Individualisierbarkeit des Systems eine wichtige Rolle?** B: Eine fundamentale, ganz wesentliche... **I: Also, auch im Sinne, die eigenen Ansichten zu wählen und auch über Schnittstellen das Ganze selber ergänzen zu können.** B: Also das ist sozusagen die Quintessenz all meiner Beschäftigung mit diesem Thema. Es gibt keine fertiggestrickten automatisierten, vollständig automatisierten Abläufe, die man für die wissenschaftliche Bearbeitung von Texten vorsehen kann. Es ist immer so, dass die Vielfalt wissenschaftlichen Fragens nicht durch irgendwelche Automatismen vorzuantizipieren sind. Nicht? Man wird einfach, also zahlreiche spezifische, projektspezifische Erfordernisse in irgendeiner Weise vorsehen müssen. Also der Benutzer muss die Möglichkeit haben, das, was für ihn und für seine spezifische Art zu arbeiten wesentlich ist, anzuwählen. Nicht? Das ist von ganz, ich würde sogar sagen von entscheidender Bedeutung. Zu glauben, dass man irgendeinen Generalweg sozusagen, eine – wie soll ich sagen, die eierlegende Wollmilchsau sozusagen auf Knopfdruck haben könnte – das ist eine Illusion. Und das wird nicht funktionieren. Da bin ich felsenfest überzeugt.

I11-6#I: Wenn wir jetzt schon perspektivisch uns darüber, damit auseinandersetzen, würde ich gerne zu einem Punkt kommen, der jetzt weniger direkt mit dem Lab in der Version, wie sie getestet wurde, zu tun hat, sondern zu eHumanities allgemein. Zu dem Begriff, was damit verbunden wird..., also einer digitalen Infrastruktur in den Geisteswissenschaften. Welche Potentiale sehen sie da? Welche Bedeutung messen sie dieser, dieser Richtung bei? B: Ich glaube, ich verstehe die Frage nicht. **I: Ja. Also ich will jetzt allgemein zu eHumanities ihre Meinung wissen, gerne hören, welche Potentiale sie damit verbinden, welche Chancen oder Risiken sie da sehen, welche Bedeutung sie diesem..., dieser Richtung zumessen.** B: Also meine persönliche Einschätzung ist eine andere, als das, was ich, sozusagen... Na, ich fang andersherum an. Also ich würde meinen: Da sehe ich ein riesiges Potential für die Geisteswissenschaften, ihre Arbeitsperspektiven zu verbessern, zu verbreitern – im komplementären Sinne. Das bedeutet nicht, dass die traditionellen Methoden und die Methodenlehre und auch die Methodologie abgelöst werden

sollen durch eine neue, sondern es sollen einfach die neuen Möglichkeiten, die man hat, die man ganz zweifellos hat, und zwar in reichem Maße hat, die soll man nutzen lernen und einsetzen lernen, für seine eigenen Vorhaben fruchtbar machen lernen. Da sehe ich ein riesiges Potential. Aber das sehe ich schon seit 20 Jahren. Und ich sehe, dass auf der anderen Seite, auf der Seite der Klientel, so gut wie kaum Gebrauch davon gemacht wird. Also ich sehe, dass von den 100% Volumen, das als Perspektiven oder als Möglichkeiten, überhaupt als Chancen zu Verfügung steht, dass da, ich würde sagen vielleicht 5% genutzt werden. Die Leute sehen die Nützlichkeit, wenn man mit ihnen darüber spricht, und erläutert ihnen: Schau mal, es wäre doch vielleicht interessant für dich, wenn du Folgendes machen würdest... Nicht? Früher bin ich durch die Gänge der Universität gelaufen und hab versucht, mir Klientel zu schaffen. Ich hab versucht, die Leute gezielt anzusprechen. Sie waren immer bereit, sich mit mir auseinanderzusetzen. Sie haben das eingesehen, das ist was ganz Nützliches. Aber sie haben es nicht gemacht, sie haben sich nicht darauf eingelassen.

I11-7#I: Und was glauben sie, woran das liegt? B: An der zunehmenden Verschulung des Studiums. Ich glaube, dass in dieser Hinsicht auch durch die BA-/MA-Studiengänge ein schwerer Fehler gemacht wird. Die Studiengänge sind zu stark modularisiert, linearisiert, diszipliniert. Die Leute haben kaum Möglichkeiten, nach links und rechts zu schnuppern, um sich..., um sich selbst neue Perspektiven zu erschließen, einfach mal zu gucken.... Man konnte sich... Früher gab es eine Menge Freiraum, wo man sich mit Dingen beschäftigen konnte, die nicht mit unmittelbarer Nützlichkeit für den künftigen beruflichen Werdegang, vielleicht auch nur für den Studienverlauf waren. Früher konnte man das. Heute kann man das nicht mehr...

I11-8#I: Ja, fehlt die Zeit, wenn man fest... B: Die Zeit, unter anderem die Zeit. Ja genau. Ich sehe das auch. Also, hier unser Sonderforschungsbereich, dem von den Geldgebern, sprich von der DFG ins Pflichtenheft geschrieben wurde, dass die anstehenden Arbeiten eben auch zur Ergebnissicherung mit Hilfe der EDV bewältigt werden sollen. Von diesen Auflagen wird kaum Gebrauch gemacht. Nicht? Wir haben hier, ich nenne jetzt keine Namen, z.B. in den Altertumswissenschaften: Die Leute sind voll und ganz damit zufrieden, wenn sie ihre nichtlateinischen Schriften, also sprich Griechisch, Hebräisch, vielleicht auch noch ein bisschen Aramäisch oder Koptisch oder was auch immer, wenn sie das auf dem Bildschirm sehen. So als wenn die ganze Verarbeitung solcher Texte im Ausdrucken bestünde. Das Gegenteil ist richtig. Das Gegenteil ist richtig. Wenn man sich mit ihnen darüber unterhält, also der, an den ich jetzt gerade denke und der jetzt übrigens gerade auch einen Ruf bekommen hat als Professor, der hat einfach gesagt: Ich kann absolut nicht erkennen, worin da der Wert für mich bestehen soll, das mechanisch, automatisch zu verarbeiten. Dass seine

Texte, die er geschrieben hat, in fünf, vielleicht sogar schon in zwei Jahren nicht mehr gelesen werden können, einfach weil es die WinWord-Version und den Makroaufsatz, den er dafür benötigt, nicht mehr gibt. Das interessiert ihn nicht. Das ist für ihn von untergeordneter Bedeutung. Ich halte das für kurzsichtig. Entschuldigung. Nicht? Aber, das kann ich, das habe ich ihm auch so gesagt, dass ich das für kurzsichtig halte, aber... Was mich dabei stört ist auch die Naivität, mit der man es auf der anderen Seite zu tun hat. Die Leute sind der festen Überzeugung... Sie kommen zu mir, z.B. da muss ein Text digitalisiert werden. Also zunächst einmal geht es nur um den Text. Nicht? Und die Leute kommen tatsächlich zu mir, das sind also die EDV-unerfahrenen, und sind der festen Überzeugung, man bräuchte nur ein Buch auf den Scanner drauflegen und unten fällt eine CD raus und da ist der Text drauf, so wie bei einem Kopierer vielleicht. Das ist natürlich lächerlich. Aber der von dem ich jetzt rede, war z.B. der Meinung, man müsste eine Prosopographie der Römischen Republik machen. Das heißt einfach nur, die entsprechenden Bücher, in denen tabellarische Übersichten über bekannte Personen der römischen Republik drin sind, also die Magistrate etwa aufgezählt sind, Konsuln, aber auch viele andere. Die bräuchte man nur einscannen und dann in irgendeiner Weise dafür sorgen, dass das in eine Datenbank reinkommt. Das geht bestimmt irgendwie automatisch. Nicht? Da muss man nur irgendwie den richtigen Knopf finden dafür. Und dann anschließend kann man damit Analysen machen. Was das genau für Analysen sind und... Egal. Aber zunächst mal einfach ne Datenbank daraus machen. Das hat er mitgekriegt, dass Datenbanken was Wichtiges sind. Aber solche naiven Vorstellungen...

I11-8#I: Was denken sie, woran liegt das? Ich meine, bei dem Klientel liegt es ja nicht an der Modularisierung des Studiums, beispielsweise. Die haben ja vielleicht in der Lehre noch damit zu tun, aber jetzt in der Forschung, oder in ihrer, bei ihrer Arbeit, da muss es ja einen anderen Grund dafür geben als Modularisierung des Studiums. Wo liegen da die Hürden, dass man sich weniger... B: Ich kann es ihnen nicht sagen. Also um halt irgendetwas Anspruchvolles mit Hilfe der EDV machen zu können, muss man Arbeitsaufwand investieren, mindestens Lernaufwand. Man muss sich ja mit einer völlig neuen Technologie, mit einer völlig neuen Art des Arbeitens auseinandersetzen. Man muss... Da müssen Texte kodiert werden und zwar nicht nur auf Teufel komm heraus, sondern mit einem ganz bestimmten Auswertungs- und Analyseziel kodiert werden. Ich kann ja nur das an Informationen herausbekommen, aus einem z.B. XML-kodierten Text, was ich vorher an Informationen, an Kodierungsinformationen reingesteckt habe. Nicht? Dafür muss eine Vorverarbeitung stattfinden. Nicht? Also ich muss Gehirnschmalz investieren. Und das scheuen die Leute, würde ich sagen.

I11-9#I: Also wäre es dann ein wichtiger Punkt, die Erlernbarkeit zu verbessern. Wenn irgendwie transportiert werden könnte, dass der Aufwand, um mit so etwas arbeiten zu können, mittlerweile geringer ist, als es vielleicht vorher, vor ein paar Jahren noch der Fall war. Sehen sie dann höhere Chancen? B: Könnte sein, ja. Ich weiß nur nicht, wie man die Erlernbarkeit verbessern kann. Aber vielleicht haben sie ja ein paar gute Ideen. Es wäre schön, wenn das so wäre. Ich sehe halt immer wieder, also wirklich eine nahezu erschreckende Naivität. Die Leute sind der festen Überzeugung: Ja, wir machen jetzt EDV. Nicht? Und dazu, wir haben auch genügend Projektmittel bekommen. Wir brauchen nur eine entsprechend leistungsfähige EDV-Infrastruktur, d.h. ein möglichst leistungsfähiger Rechner muss angeschafft werden. Und natürlich auch noch Peripherie: Drucker, Scanner usw. Das muss alles noch mit dazu und, natürlich auch noch ein ganz teures Programm, am besten auch noch fünfstelliger Bereich. Ein teures Programm ist gleichzeitig ein gutes Programm. Und dann lege ich los. Nicht? Dann wird aber alles, was ich habe, wird von der EDV automatisch erledigt. Und wenn ich den Leuten in solchen Projektsitzungen erkläre: Nein, ihr könnt EDV nur dann sinnvoll einsetzen, wenn ihr euch vorher verarbeitend Mühe macht und euer Gehirn anstrengt. Dann werde ich angeguckt, als wenn ich nicht recht bei Trost wäre. Und das ist etwas Deprimierendes. Das kann ich ihnen versichern, das wird ihnen übrigens auch mit ihrer Klientel mit dem TextGrid-Lab noch passieren. Warten sie mal drauf, bis sich noch eine breitere Öffentlichkeit darauf einlässt, wird es sehr viel unsachliche Kritik genau aus dieser Ecke geben. **I: Das kann gut sein. Deshalb bin ich auch daran interessiert, jetzt zu erfahren, was in der Community an Verbesserungsvorschlägen oder eventuell Lösungsmöglichkeiten dieser Hürden gedacht wird oder in Betracht kommt. Wenn sie sagen, da herrscht eine große Naivität. Dem kann man ja eigentlich nur begegnen, wenn man die Hürden absenkt. Also dass man praktisch auf Augenhöhe mit dem Benutzer geht. Also es soweit vielleicht auch veranschaulicht, dass da eine höhere Nachvollziehbarkeit vorliegt.** B: Ja, aber es fängt schon auf einem ganz primitiven Level an. Und ich glaube nicht, dass man das..., dass es eine Lösung dafür gibt, diese Einstiegshürde zu senken. Das Problem besteht in der Kodierung. Nicht? Also wir reden ja davon. Es geht ja nicht darum, irgendwie in eine Tastenkombination zu drücken und dann erscheint das Richtige, zum Beispiel ein geschwänztes z, oder ein langes s, oder sonst irgendetwas auf dem Bildschirm, sondern ich will eine Repräsentation dieses seltsamen Zeichens einfügen. Und es ist völlig egal, ob das jetzt auf dem Bildschirm erscheint oder nicht. Hauptsache: Ich habe die semantisch richtige Information in kodierter Form eingetragen. Aber das ist eine Hürde, die für viele nicht zu überschreiten ist. Also ich kenne jemanden. Ich will wiederum keine Namen

nennen. Aber, jemand, der eigentlich, sagen wir mal, ein Pionier auf diesem Gebiet hätte sein können. Nicht? Der weigert sich, eine Kodierung einzugeben. Dann kann ich ja nicht mehr sehen, was das ist. Verstehen Sie? Nicht? Also das ist... **I: What you see is what you get. Unterscheidung dieser Ebenen...Und, ja.**

I11-10#Aber mal aus einer anderen Perspektive. Wenn es jetzt nicht darum geht, dass jemand kodieren muss, sondern dass er vielleicht einfach nur rezipiert in dem Sinne, dass er recherchiert. Also in Datenbankbeständen suchen kann, evtl. in dem Wörterbuchservice, der in dem TextGrid-Lab zur Verfügung stand. Das kann ja auch eine Form der Nutzung von..., im Sinne von eHumanities sein. Wenn jemand nicht editorisch tätig ist zum Beispiel. Dann würde diese Hürde ja beispielsweise ausfallen, der Kodierung. Wenn er nicht als jemand tätig ist, der da einspeist, zunächst einmal. B: Gut. Ok. Und dann? I: Oder? Wäre das dann vielleicht ein erster Zugang? Eher über die Nutzung, nicht als jemand, der Daten einspeist, sondern der das lediglich nutzt, um zu recherchieren, Datenbestände zu durchsuchen... B: Aber wenn er suchen will, z. B. in irgendetwas, was in nichtlateinischer Form kodiert im Text enthalten ist. Wenn er nicht darüber weiß, was an Kodierungen drin ist, dann kann er doch auch nicht suchen. I: Das kommt drauf an – auf die Suche an. Oder? B: Ja also, wenn das, wonach er sucht, keine Entsprechung im lateinischen Alphabet hat. Was haben wir denn dann? Wie sucht man so z.B. nach chinesischen Schriftzeichen oder nach japanischen oder arabischen? I: Gut. Das ist jetzt auch ein ziemlich spezieller Bereich. Aber es wäre dann... Also, wenn man etwas gut indexiert und gut recherchierbar macht, dann müsste das eigentlich auch so durchsuchbar sein, dass derjenige nicht unbedingt sich mit den speziellen Codes auskennt. Oder? B: Wenn das abgenommen wird – ja. Aber es gibt doch auch immer wieder... Also, also jemand der sich ... Also ein Hebraist beispielsweise, nicht? Vielleicht gibt es da irgendeine Möglichkeit über eine Transliteration zum Ziel zu kommen, nicht? Aber wenn er dann nach einer ganz bestimmten Erscheinungsform eines bestimmten, ein bestimmtes Diakritikum sucht, nicht? Da wird er schon seine Probleme haben, könnte ich mir denken. Also ich, ich stelle es mir nicht so einfach vor, ihm alle Arbeit auf diesem Gebiet abzunehmen. Ich finde es mal interessant, auf diesem Gebiet zu suchen, ob es geht. Es könnte sein, dass dabei etwas rausschaut. Aber bis ich es nicht gesehen habe, glaube ich es nicht.

I11-11#I: Gut. Gut. Dann wird der nächste Schritt vielleicht nicht ganz leicht fallen. Aber nichtsdestotrotz würde ich das gerne mal durchspielen, das Szenario. Gehen wir mal davon aus: Es gibt einige Nutzer, die diese digitale Infrastruktur, die beabsichtigt ist, auch wirklich nutzen wollen. Die also diese Hürde überwunden haben. Es wird

vielleicht eine kleine Community sein, aber mal angenommen: Es gibt so eine. Dann befinden die sich ja in einer Infrastruktur, die man als Vernetzung von Inhalten, aber auch von Arbeitsprozessen ansehen kann. B: Vernetzung von Inhalten, meinen sie also die Texte. Nicht? I: Zum Beispiel Texte, aber auch Wörter, wie in dem Wörterbuchservice – das ist ja eher ein Wortnetz. Also es gibt Wortnetze, Textnetze, aber auch vernetzte Projekte. Also die Projektverwaltung wurde ja auch getestet bei diesem Workshop. Das soll ja auch zu Verknüpfungen führen, bzw. bringt das modulare Konzept ja mit sich, dass man auch Tools irgendwann anschließen kann bzw. auch kommunizieren kann mit anderen und so – also in vielfacher Weise vernetzt. Grob unterschieden jetzt in Inhalte und Arbeitsprozesse. Welche Potentiale sehen sie in den Kollaborationsmöglichkeiten über diese Netzwerke? Möglichkeiten der Zusammenarbeit? B: Gigantische. Also mein Problem war in den 80er Jahren noch zum Beispiel... Also ich... Ein altes Hobby von mir besteht in der quantifizierenden Textanalyse. Nicht? Ich habe einfach geguckt: Was kann ich herausfinden über einen Text, über den Modus Scribendi des Autors in genau diesem Text. Was kann ich als autorspezifisches Kriterium herausarbeiten? Nicht? Und ich habe dabei auch ziemlich viele interessante Dinge gefunden. Mein größtes Problem bestand einfach nicht zu sehr im Programmieren von analytischen Zugriffen auf die Daten, sondern mein Problem bestand darin, an Texte heranzukommen bzw. an verlässliche Texte. Auf diesem Gebiet hat z.B. das Projekt Gutenberg unheimlich viel geleistet. Aber da sehe ich einen gigantischen Bedarf. Nicht? Also, die eHumanities, überhaupt die Gemeinschaft, die Community der eHumanities, muss an der Stelle zusammenwachsen und einfach zunächst einmal eine Datengrundlage, aber natürlich dann anschließend auch einen methodologischen Pool schaffen, um einen fundierten Zugriff auf Texte aller und jedweder Art zu ermöglichen.

I11-12#I: Und da wären wir jetzt auch bei den vernetzten Inhalten. Das ergibt ja dann eine riesige Masse an Daten. Was halten sie da für wichtig, in der Erschließung bzw. im Umgang mit diesen Datenmengen? B: Ja, von fundamentaler Bedeutung ist es, gute Ausgaben zu schaffen. Also nicht... Also bei Projekt Gutenberg, das ich vorhin erwähnt habe, das stellt ein riesiges Reservoir an Texten zur Verfügung. Aber diese Texte sind von höchst unterschiedlicher Qualität. Also das heißt aus editorischer Sicht, aus wissenschaftlicher Sicht, sind diese Texte – ich würde sagen – zu mehr als 80% unbrauchbar. Also man greift auf irgendetwas zu. Man will sich z.B. mit der französischen oder der spanischen oder italienischen Version etwa von Moby Dick beschäftigen und Texte... die Texte findet man. Aber aus editorischer Sicht ist das meistens unter aller Sau. Entschuldigung. Nicht? Und das ist ein umso schwierigeres Problem, je weiter der Autor dieser Texte von uns zeitlich entfernt

ist. Beim Zugriff auf antike Texte, da... Die Situation ist derzeit mehr als enttäuschend. Und da wird viel viel Arbeit zu investieren sein. Also es kommt nicht darauf an, dass ein editorisch guter und solide gemachter Text zur Verfügung steht, von einem bestimmten Autor und von einem bestimmten Text, sondern es kommt darauf an, dass eine möglichst breite Basis geschaffen wird. Dass also aus verschiedener editorischer Perspektive verlässlich gearbeitet wird und da eben ein entsprechendes Reservoir, ein riesiger Pool von Texten gebildet wird. **I: Ja, da wäre jetzt die Qualität der einzelnen Datensätze der Punkt. Was mich eben noch interessieren würde: Bei dieser Masse an Daten, kommt es ja auch auf die Handhabbarkeit an. Also gehen wir mal davon aus, dass das ein Kriterium ist. Die Qualität der einzelnen Datensätze. Also...** B: Also Entschuldigung. Als ich von der Qualität sprach, meinte ich natürlich auch, dass die editorische Erschließung, sprich, dass Varianten beispielsweise, vielleicht sogar Kommentare, inhaltliche Kommentare, entsprechend ausgezeichnet in diesen Texten eingebettet sind, so dass ich darauf zugreifen kann. Dass ich z.B. bei einer philologischen Krux, die ich nicht als solche erkennen kann, dass ich da einen Hinweis bekomme: An der Stelle haben wir es mit einer Schwärzung im Papier zu tun. Das, was ich hier hingeschrieben habe, als Wort, das ist eine Vermutung von mir. Ich vermute, dass an der Stelle folgendes Wort steht. Aber es kommt auch noch, kommen auch noch folgende andere Möglichkeiten noch mit in Betracht. Also Textkritik, textkritischer Apparat. Nicht? Muss mit rein. Unbedingt. **I: Was dann auch die Recherchefunktion, natürlich, verbessern würde.** B: Ganz wesentlich sogar. **I: Also über Erschließung und Indexierung hin zu einer Übersicht über die großen Datenmengen.** B: Also das ist das was ich vorhin mit Qualität gemeint habe. Also nicht irgendeinen Text ... Vor – Wie lange ist das her? – vor eineinhalb Jahren waren, war die ITUG-Tagung in Zürich. Und da hat Fotis Jannidis einen Vortrag gehalten und hat, also natürlich ein guter Vortrag, es war halt eine dreiviertel Stunde, was kann man an der Stelle schon richtig bringen? Aber da war davon die Rede, dass irgendein Text aus irgendeiner Quelle genommen wird. Und dass der, der TextGrid-Lab benutzt, gar nicht wissen muss, aus welcher Quelle der Text kommt. Nicht? Und der wird dann in irgendeinen Verarbeitungsmechanismus reingesteckt, dieser Text. Also der Text kann zergliedert werden, in ein Wortformenverzeichnis zum Beispiel. Oder er wird formatiertechnisch aufbereitet und solche Dinge. Und das ist in dieser Verkürzung einfach zu naiv. Nicht? Ich muss wissen aus welcher Quelle der Text kommt. Ich muss wissen nach welchen Qualitätskriterien er erarbeitet worden ist. Und ich brauche jede Menge periphere Informationen dazu. Es gilt umso mehr, je weiter eben der Autor von uns zeitlich entfernt ist. Nicht?

I11-13#I: Gut. Sehen sie da eine Gefahr bei großen Datenmengen und wenn... Also selbst wenn sie einigermaßen gut erschlossen sind, dass man sich da verliert? Also in dieser Vernetzung von zahlreichen Inhalten nicht sich orientieren kann? B: Das sehe ich nicht. Nein, ganz im Gegenteil. Also wenn sich jemand aus wissenschaftlicher Perspektive mit irgendeinem Text beschäftigen will, dann muss er in der Lage sein: a) den Text ausfindig zu machen. Er muss in der Lage sein, sich selbst, seine eigenen Qualitätsanforderungen zu bestimmen und zu sagen: Ich habe für mich verschiedene Versionen davon und ich entscheide mich bewusst für diese hier, weil die mir eben nach meinen Qualitätsanforderungen, die ich habe, die ich an den Text richte, der richtige, der gute Text zu sein scheint. Wenn er dazu nicht in der Lage ist, dann hat er in der Wissenschaft nichts verloren, finde ich. Kleinen Moment... [Telefonat]. So. **I: Gut, ja, ich habe jetzt eher gemeint, wenn jemand mit dem System arbeitet und auf große Datenmengen zugreift ... Klar, wenn er gezielt sucht und das alles soweit gut erschlossen ist, dann dürfte er ja auch fündig werden. Aber jetzt einmal abgesehen von der gezielten Suche gibt es ja auch andere Möglichkeiten, sich damit zu beschäftigen und deshalb...** B: Sprechen sie jetzt von irgendeiner Struktur für den Datenpool? **I: Ja, genau. Also, inwiefern, wenn wir von Netzwerkstrukturen ausgehen, inwiefern halten sie es da für notwendig... Oder wie würden sie sagen, wäre es am besten, zu erschließen, für den Nutzer, dass er sich da gut zurechtfinden kann.** B: Also unbedingt eine Systematik anlegen. Nicht? Wobei die Systematik einfach wächst und gedeiht im Verhältnis und entlang der Daten, die in das System eingespeist werden. Also, ich weiß nicht. Zunächst einmal würde ich versuchen, so eine grobe literaturhistorische Systematik einzurichten – also zuerst einmal Antike, Mittelalter, neuzeitliche Literatur – und dann nach Sprachen, nach Sprachräumen geordnet. Aber ohne Systematik kommt so ein riesiger Datenpool nicht aus. Ich hätte es von daher für eine Selbstverständlichkeit gehalten, dass eine Struktur appliziert werden muss. Also ohne dass man jetzt mit einer festen Struktur von vorneherein daran geht, sondern dass man einfach sich mal... Also nach einer Weile, nachdem die ersten Daten eingespeist worden sind, sich anhand dessen, was sich drin befindet, überlegt: Wie gliedere ich das denn? Also wie schaffe ich dem Benutzer einen möglichst leichten Zugang von oben, von einem übergeordneten Gesichtspunkt aus auf die Daten.

I11-14#I: Ja, also: Halten sie es dann auch sinnvoll, das für den Nutzer sichtbar zu machen, weil man ist ja zunächst einmal in dem System drin und hat vielleicht keinen genauen, keine genaue Vorstellung davon, was eigentlich dahintersteht. Oder? B: Ja, ja. Also die Systematik muss für ihn natürlich nachvollziehbar sein, sonst hat die Systematik

keinen Sinn. Nicht? Das ist also... Das würde ich für apriorisch halten. Das ist ganz fundamental. Also wenn eine Systematik da ist, muss sie für den Benutzer auch erschließbar sein. Also ich würde nicht den Versuch machen, diese Systematik zu übertreiben, also zu feingliedrig zu machen. Das kann man später noch tun. Nicht? Nur aber zur ersten Orientierung. Also Benutzer kommen typischerweise mit drei bis fünf Hierarchiestufen ganz problemlos zurecht. Ja, daran würde ich mich orientieren.

I11-15#I: Und ansonsten auch die Möglichkeit, selber Strukturen zu schaffen, in Erschließungsprozessen, oder nicht? B: Nein, auf keinen Fall. [Unterbrechung, Scheinausgabe an eine Studentin]. Also was eine Systematik angeht. Es widerspricht dem Grundgedanken der Systematik, dass viele Köche Zugriff haben. Viele Köche verderben immer den Brei. Man muss die Möglichkeit vorsehen, dem Benutzer die Möglichkeit einzuräumen, Vorschläge zu machen. Natürlich. Aber auf keinen Fall darf er selbst an der Systematik herumfummeln. Systematik findet in einem einzigen Kopf statt. Also da muss es einen Administratoren bzw. eine administrative Ebene geben. Und die beschäftigt sich mit solchen Vorschlägen und baut sie entweder ein oder verwirft sie. Nicht? Aber auf keinen Fall den Benutzern die Möglichkeit geben. Das wäre... Also ich hielte das für einen schweren Fehler, auch für einen Verstoß schon gegen den Systemgedanken.

I11-16#I: Vor dem Hintergrund: Wie, finden sie, ist im Moment die Darstellung von Rechercheergebnissen im Lab gelungen? Haben sie mal mit der Recherchefunktion gearbeitet? B: Ich weiß es nicht mehr. **I: Mit dem Wörterbuchservice?** B: An welcher Stelle war denn das? Mit dem Wörterbuchservice habe ich gearbeitet, aber nur zum Lemmatisieren. Ich habe also nicht... Was hätte ich denn z.B. suchen können? **I: Es gibt eine allgemeine Recherchemaske. Also einfach einen Suchbegriff eingeben und sehen, welche, zu welchen Daten der führt.** B: Also nach Wörtern suchen. **I: Ja, genau. Und nach Suchbegriffen...** B: Nein, das habe ich nicht gemacht. Deswegen kann ich mich auch nicht dazu äußern.

I11-17#I: Ok. Gut. Dann abschließend: Wo sehen sie die wichtigsten Fortschritte jetzt bei TextGrid bzw. das größte Manko noch? Auch im Vergleich zu anderen EDV-philologischen Anwendungen, die sie kennen. Können sie das zusammenfassen? B: Ja, also was die... Also ich habe vorhin schon einmal das Wort gebraucht von der Vielfalt historischen Fragens. Das wird man in einer solchen Oberfläche nur sehr schwer abfangen können. Also was ich noch ein bisschen vermisse, war ganz einfach die Feingliedrigkeit des Zugriffs unter verschiedenen fachspezifischen Gesichtspunkten. Also z.B.: Wir haben ja diesen Text von Goethe beispielsweise gehabt. Nicht? Und in diesem Text von Goethe kamen

sehr viele historische Schreibweisen vor. Und die hat natürlich der Lemmatisierer, der sich an dem Wörterbuch des Instituts für deutsche Sprache orientiert hat, der konnte damit nichts anfangen bzw. hat sogar falsche Informationen geliefert. An einer Stelle war Herz mit tz geschrieben. Das hat er nicht... Das hat er tatsächlich lemmatisiert, aber auf Hertz und zwar auf die Schwingungen. Wissen sie? Der... elektromagnetische Schwingungen. Also er hat es nicht auf das richtige Wort... Wie sollte er auch, ne? Also ich wollte nur sagen: Es wird sehr schwer sein, allen Aspekten wissenschaftlichen Zugriffs, in einer solchen Oberfläche gerecht zu werden. Aber darin sehe ich auch die besondere Herausforderung drin. Und das ist gleichzeitig auch... Das ist der zweite Teil ihrer Frage, das beachtliche Potential dafür. Sie werden an dem gemessen werden. Oder bzw. das TextGrid-Lab wird an dem gemessen werden. Also es wird immer, zumindest in den ersten Jahren ganz besonders, wird sich sehr viel Kritik einstellen. Die wird praktisch unvermeidlich sein. Einfach weil, weil es sehr lange dauern wird, bis man allen möglichen Gesichtspunkten gerecht wird. Aber wie gesagt, das ist auch eine besondere Chance. Ich würde mich dieser Chance stellen und würde die Herausforderung annehmen. Einfach mal gucken, was in zwei, drei Jahren sein wird. Nur es erfordert bis... Der Weg dorthin ist sehr steinig.