



Abschlussbericht

Öffentliche Fassung

TextGrid – Institutionalisierung einer Virtuellen Forschungsumgebung in den Geisteswissenschaften

Version: Version 1.0 vom 30.11.2015	Dokumentenstatus: Final
Verantwortlicher Partner: SUB Göttingen	Verfügbarkeit: Öffentlich

Projekt: TextGrid	Förderkennzeichen: 01UG1203A-F
Vorhabenbezeichnung: TextGrid – Institutionalisierung einer Virtuellen Forschungsumgebung in den Geisteswissenschaften	
Laufzeit des Vorhabens: Juni 2012 bis Mai 2015	

Förderkennzeichen der Verbundpartner:

1. **01UG1203A** Niedersächsische Staats- und Universitätsbibliothek Göttingen
2. **01UG1203B** Julius-Maximilians-Universität Würzburg
3. **01UG1203C** Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen
4. **01UG1203D** Technische Universität Darmstadt
5. **01UG1203E** Institut für Deutsche Sprache Mannheim
6. **01UG1203F** Max-Planck-Institut für Wissenschaftsgeschichte Berlin

Autoren: TextGrid Konsortium (alle Partner)

Partner	MitarbeiterInnen
Berlin-Brandenburgische Akademie der Wissenschaften, Berlin	Krauth, Wolf-Hagen Neumann, Gerald Roeder, Torsten Zinsmeister, Elke
DAASI International GmbH, Tübingen	Gietz, Peter Haase, Martin Funk, Stefan E.
Hochschule Worms, Zentrum für Technologietransfer und Telekommunikation (ZTT) ¹	Küster, Marc Wilhelm Selig, Thomas Stuck, Christian Stuck, Simon
Georg-August-Universität Göttingen, Niedersächsische Staats- und Universitätsbibliothek Göttingen	Blümm, Mirjam Brodhun, Maximilian Funk, Stefan E. Grupe, Nadja Hausner, Roman Neuroth, Heike Söring, Sibylle Veentjer, Ubbo
Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen	Berber, Fatih Wieder, Philipp
Institut für Deutsche Sprache, Mannheim	Fiedler, Norman Witt, Andreas
Julius-Maximilians-Universität Würzburg, Institut für Deutsche Philologie	Betz, Katrin Jannidis, Fotis Vitt, Thorsten
Max-Planck-Institut für Wissenschaftsgeschichte, Berlin	Rieger, Simone Thoden, Klaus Wintergrün, Dirk
Technische Universität Berlin, Zentrum Technik und Gesellschaft	Kaden, Ben Palfner, Sonja
Technische Universität Darmstadt, Institut für Sprach- und Literaturwissenschaft	Leuk, Michael Rapp, Andrea Schmid, Oliver Vanscheidt, Philipp

¹ Zum Antragszeitraum Fachhochschule Worms

Inhaltsverzeichnis

1. Aufgabenstellung.....	4
2. Voraussetzungen, unter denen das Vorhaben durchgeführt wurde	5
3. Planung und Ablauf des Vorhabens	6
3.1 Projektüberblick	6
3.2 Arbeitspakete und Arbeitsgruppen.....	7
3.3 Projektergänzungen	7
4. Wissenschaftlicher und technischer Stand, an den angeknüpft wurde.....	9
4.1 Wissenschaftlicher Stand	9
4.2 Technischer Stand	10
5. Zusammenarbeit mit anderen Stellen.....	14
5.1 Projekte stellen sich vor	14
5.2 DARIAH-DE.....	14
5.3 SADE	15
5.4 Digilib & IIF	15
5.5 TUSTEP.....	16
5.6 Johann Friedrich Blumenbach - Online	16
5.7 Die DH-Landschaft	16
6. Verwendung der Zuwendung und des Ergebnisses im Einzelnen, mit Gegenüberstellung der vorgegebenen Ziele	17
6.1 AP-übergreifend: TextGrid 1.0., TextGrid 2.0 und TextGrid 3.0.....	17
6.2 AP 1: Betriebsmodell	19
6.2.1 AP 1.1 Rechtsform, Kostenkategorien.....	20
6.2.2 AP 1.2 Verein, Geschäftsstelle.....	25
6.2.3 AP 1.3 (Technische) Integration und zukünftiges Weiterentwicklungspotential, Zukunftsfähigkeit.....	27

6.3 AP 2: Prozessorganisation	29
6.3.1 AP 2.1 Umgang mit großen Nutzerzahlen und heterogenen Nutzergruppen sowie Nutzerverwaltung (Praxis und Theorie) und AP 2.2 Verteilte / Zentrale Entwicklungen.....	29
6.3.2 AP 2.3: Finanzierungsmodelle und -wege, verantwortet von ProWiss.....	32
6.4 AP 3: Nutzer-Kommunikation.....	34
6.4.1 AP 3.1 Dokumentation, Nutzer-Information.....	34
6.4.2 AP 3.2 eHumanities und TextGrid in universitären Curricula.....	35
6.4.3 AP 3.3 Workshops, Schulungen, Nutzer-Interaktion.....	37
6.5 AP 4: Betrieb und Softwarepflege Repository.....	42
6.5.1 AP 4.1 Humanities Data Centre	42
6.5.2 AP 4.2 Performanz, Skalierbarkeit, Lasttests.....	43
6.5.3 AP 4.3 Qualitätsmanagement (SUB).....	45
6.5.4 AP 4.4 Ingest Datensammlung, Publikation, Präsentation.....	46
6.5.5 AP 4.5 Technischer Regelbetrieb, Architektur, innovatives Evolutionspotential.....	46
6.6 AP 5: Betrieb und Softwarepflege Laboratory	47
6.6.1 AP 5.1 Sicherheit, Bugfixing, Technische Dokumentation	47
6.6.2 AP 5.2 Funktionale Tests	48
6.6.3 AP 5.3 Usability.....	48
6.6.4 AP 5.4 Koordination von Weiterentwicklungen und Satellitenprojekten (UWÜ)	50
6.6.5 AP 5.5: Weiterentwicklung der Digitalen Bibliothek.....	53
6.7 AP 6: Integriertes Monitoring und Qualitätssicherung	54
6.8 H. AP 7: Projektmanagement und Öffentlichkeitsarbeit (Koordination durch AP-Leitung/SUB)	58
6.8.1 AP 7.1 Projektmanagement i.e.S und Koordinierung potentieller Kooperationen.....	58
6.8.2 AP 7.2 Abstimmung mit den Fachgesellschaften	59
6.8.3 AP 7.3 Marketing und Öffentlichkeitsarbeit.....	60
6.8.4 Finalisierung des Nachhaltigkeitskonzepts und Koordination	62
7. Die wichtigsten Positionen des zahlenmäßigen Nachweises	64

8. Die Notwendigkeit und Angemessenheit der geleisteten Arbeit	66
9. Der voraussichtliche Nutzen	67
10. Während der Durchführung eines Vorhabens dem ZE bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen	69
11. Bereits erfolgte oder geplante Veröffentlichungen des Ergebnisses nach Nr. 6	70
11.1 Bereits erfolgte Veröffentlichungen.....	70
11.2 Geplante Veröffentlichungen.....	72
Annex	73
I. Partnerliste	73
II. Arbeitspakete und Deliverables	74

1. Aufgabenstellung

Digitale Werkzeuge etablieren sich zunehmend auch in den Geisteswissenschaften. Virtuelle Arbeitsumgebungen unterstützen den gesamten Forschungsprozess des Recherchierens, Analysierens, Publizierens und Archivierens und erlauben erstmals den Zugriff auch auf große Mengen wissenschaftlicher (Primär-)Daten. Seit 2006 baut TextGrid eine Virtuelle Forschungsumgebung für die text- und quellenbasierten Geisteswissenschaften auf.

Mit seinen beiden in der ersten Projektphase (2006-2009) entwickelten Hauptkomponenten TextGridRep und TextGridLab konnte TextGrid in der zweiten Projektphase (2009-2012) eine Infrastruktur zur Verfügung stellen, die sowohl Storage für Datenmanagement und Datenspeicherung bzw. -archivierung (TextGridRep) als auch verschiedene Basiswerkzeuge und -dienste (TextGridLab) in einer stabilen Produktivversion bot. Die Schaffung einer breiten Nutzerbasis als Voraussetzung für fachspezifische Weiterentwicklungen der Forschungsumgebung und ihrer Komponenten stand bereits in der zweiten Förderphase im Fokus und wurde in der dritten Projektlaufzeit weiterverfolgt. Schwerpunkt der dritten Förderphase (2012-2015) war die Etablierung eines nachhaltigen Dauerbetriebs: Software, Archiv und damit das gesamte Angebot sollten in Zusammenarbeit mit AnwenderInnen, FachwissenschaftlerInnen und Fachgesellschaften auf eine langfristige Nutzerbasis gestellt werden. Denn nur wenn Dienste und Werkzeuge, Archiv und Software in einer sicheren, stabilen Umgebung langfristig verfügbar sind und den verschiedenen spezifischen fachwissenschaftlichen Anforderungen und Bedürfnissen nicht nur entsprechen, sondern auch die Möglichkeit zu projektspezifischen Anpassungen und Weiterentwicklungen bieten, wird die geisteswissenschaftliche Community dem Angebot einer Virtuellen Forschungsumgebung langfristige Akzeptanz und das notwendige Vertrauen entgegenbringen.

Die Virtuelle Forschungsumgebung (VFU) TextGrid ist bundesweit Vorreiter und Beispielfall für die Entwicklung eines nachhaltigen Finanzierungs- und Betriebsmodells, das alle Facetten der Institutionalisierung berücksichtigt und dessen Ergebnisse auf andere Infrastrukturen übertragen bzw. mit diesen gebündelt werden (z.B. DARIAH-DE, CLARIN). TextGrid hat belastbare Kalkulationen und Organisationsmodelle zum dauerhaften Betrieb einer VFU vorgelegt und einen Verein gegründet, der die Institutionalisierungsbemühungen zusammenführt und rechtlich-organisatorisch absichert. Auf europäischer Ebene ist hier auf das „Legal Framework for a European Research Infrastructure Consortium (ERIC)“ zu verweisen, mit dem das Infrastrukturprojekt DARIAH-EU seine Institutionalisierung vorantreibt. Die Entwicklung eines nachhaltigen Finanzierungsmodells ist noch nicht abgeschlossen, da schon aufgrund der wachsenden Breite und Diversität von Anforderungen in den Geisteswissenschaften eine Finanzierung nicht „aus einer Hand“ realisierbar ist. Vielmehr wird ein Modell konzipiert, das von unterschiedlichen Finanzierungsstufen, -modalitäten und -dynamiken ausgeht und in dem institutionelle Grundsicherung, verteilte Eigenanteile, projektbezogene Kostenerstattung, Erhebung von Gebühren, Nutzung von Overheads und weitere Finanzierungsquellen flexibel zusammenspielen. Die Realisierung eines solchen Modells ist zeitintensiv, da mit vielen verschiedenen Stakeholdern und Akteuren ein mögliches finanzielles Engagement auf diesem Gebiet besprochen und abgestimmt werden muss.

2. Voraussetzungen unter denen das Vorhaben durchgeführt wurde

Die Geistes- und Sozialwissenschaften sind – nicht nur im engeren Bereich der Digital Humanities – auf den Zugang zu digitaler Forschungsinfrastruktur angewiesen. Deren Entwicklung ist derzeit überwiegend projektförmig realisiert, obwohl Infrastruktur verlässlich, nachhaltig und dauerhaft zur Verfügung stehen muss, um Lehre und Forschung auf dem geforderten Qualitätsniveau langfristig sicherzustellen. Nicht zuletzt ergibt sich für Forschende durch einfachen und stabilen Zugang zu digitaler Infrastruktur ein Standortvorteil im nationalen und internationalen Wettbewerb um Forschungsmittel und qualifizierten Nachwuchs. In seiner dritten Förderphase konnte das Projekt sowohl auf der technischen Ebene als auch in strukturell-organisatorischer und fachwissenschaftlicher Hinsicht auf die Voraussetzungen und erzielten Ergebnisse der ersten und zweiten Förderphase aufbauen. Diese dienten als Grundlage für eine kontinuierliche Fortführung des langfristigen Projektziels, eine nachhaltige e-Humanities-Infrastruktur für die Geisteswissenschaften aufzubauen und in der Community zu etablieren.

3. Planung und Ablauf des Vorhabens

Die kooperative und offene Vorgehensweise aller Partner, die bereits die erste und zweite Projektphase etabliert hatte, konnte in der dritten Förderphase fortgesetzt werden. Sie zeichnete sich durch intensive Kommunikation mit internen wie externen Partnern sowie potentiellen und aktiven NutzerInnen aus. Präsentationen auf relevanten nationalen wie internationalen Konferenzen, die Einbindung auch externer Expertinnen und Experten in projektinterne Debatten und Entscheidungsprozesse wie auch die Kooperation mit externen Projekten hatten Transparenz und die Aufmerksamkeit seitens der Community zur Folge. Der TextGrid / DARIAH-DE DH Summit im März 2015 verzeichnete etwa 280 TeilnehmerInnen aus dem In- und Ausland. Das Angebot der regelmäßig stattfindenden TextGrid-Nutzertreffen nahmen pro Treffen durchschnittlich etwa 50 Personen wahr, bei der TextGrid Grand Tour im Rahmen des DH-Summits nahmen sogar gut 70 Personen teil. Bei Projektende verzeichnete TextGrid mehr als 1.200 registrierte NutzerInnen.

3.1 Projektüberblick

Das TextGrid-Konsortium bestand in der dritten Projektphase aus elf Partnern (siehe Annex A), die in allen sieben Arbeitspaketen sehr eng zusammengearbeitet haben.

Der Kommunikation diente auch in dieser Projektphase eine Reihe paralleler, komplementärer Kanäle:

- Ein Projekt-Wiki (interner Bereich der Projekt-Webseite), enthaltend alle wichtigen Dokumente, das detailliert Ablauf und Entscheidungen des Projekts archiviert.
- Verschiedene Mailinglisten für die jeweiligen Arbeitspakete, Arbeitsgruppen und das Gesamtprojekt zur zielgerichteten Kommunikation auch mit externen Partnern. Alle TextGrid-Mailinglisten werden über die Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG) gehostet.
- Regelmäßige persönliche Treffen des gesamten Konsortiums sowie für Arbeitsgruppen und Task Forces nach Bedarf.
- Ein- oder mehrtägige „Programmiersprints“ für die technische Entwicklungsarbeit
- Video- bzw. Telefonkonferenzen für den kontinuierlichen Austausch u.a. der Steering Group (monatlich) und der AG Architektur (monatlich). Für Videokonferenzen stand das Videokonferenzsystem des DFN-Vereins zur Verfügung. Für Telefonkonferenzen hat sich im Laufe des Projektes „Skype“ als Mittel der Wahl durchgesetzt, das auch kurzfristige Termine und Absprachen ermöglicht.
- Zur Außendarstellung dienten vor allem die TextGrid-Website und der DHd-Blog - Digital Humanities im deutschsprachigen Raum (zusammen mit DARIAH-DE). Die monatlich über 1.100 Besucher auf der TextGrid-Website (insgesamt seit Dezember 2012 knapp 36.000) und die bis Mai 2015 202 registrierten AutorInnen und 613 Artikel beim DHd-Blog unterstreichen das öffentliche Interesse an der Virtuellen Forschungsumgebung.

3.2 Arbeitspakete und Arbeitsgruppen

Neben der Arbeit in den Arbeitspaketen wurden thematische Arbeitsgruppen (AGs) und Task Forces (TFs) gebildet, die arbeitspaketübergreifend agierten und damit die Kohärenz des Gesamtprojekts sicherstellten.

- AG Architektur – Konzeption der grundlegenden TextGrid-Architektur und TextGrid-Infrastruktur
- AG Textformate – Austauschtextformat
- AG Tests – koordiniert, delegiert und kontrolliert Tests aller Art
- AG Evaluation neuer Anforderungen
- TF Metadaten – Metadatenschema
- TF Service Level Agreements
- TF Usability, User-Szenarien
- TF Usability, TextGridRep.de – Erweiterung, Verbesserung und Anpassung des Repository-Browsers (<http://www.textgridrep.de>)
- TF TextGrid / DARIAH-DE AK Nachhaltigkeit – Finanzierung des Übergangs in den Dauerbetrieb
- TF TextGrid / DARIAH-DE AK Architektur – Zusammenführung bzw. Migration der technischen Infrastruktur beider Projekte

3.3 Projektergänzungen

Unter der Prämisse, TextGrid in die (inter)national entstehende e-Humanities-Infrastruktur einzubetten und zu vernetzen, hat das Projekt bereits in seiner ersten und zweiten Phase zahlreiche Kontakte und Kooperationen geknüpft und etabliert, die den Projektverlauf positiv beeinflussten.

So wurde die Zusammenarbeit mit dem ebenfalls vom BMBF geförderten Projekt DARIAH-DE, das eine digitale Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften aufbaut, so weit intensiviert, dass zum Projektende von TextGrid eine Migration von zahlreichen wichtigen technologischen Kernkomponenten wie etwa die Nutzerverwaltung und Repository-Technologien nach DARIAH-DE erfolgte und so die dauerhafte und langfristige Nutzung der Angebote von TextGrid auf neuestem technologischen Stand auch weiterhin gewährleistet ist.

TextGrid ist zudem stetig in Gesprächen mit verschiedenen Forschungsprojekten und -verbänden, um an bestehende Entwicklungen anzuknüpfen und diese fortzuführen. Aktuell werden 16 laufende Projekte im TextGrid-Umfeld verzeichnet, darunter die genetisch-kritische und kommentierte "Hybrid-Edition von Theodor Fontanes Notizbüchern", die Erstellung von "Textdatenbank und Wörterbuch des Klassischen Maya (IDIOM)" sowie das Projekt "Bibliothek der Neologie".

SADE (Scalable Architecture for Digital Editions), ein eXist-basiertes Werkzeug, welches viele Komponenten umfasst, die zum Entwickeln und Darstellen einer digitalen Edition besonders geeignet sind, wird seit 2012 unter anderem von der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW), der Österreichischen Akademie der Wissenschaften (ÖAW; im Rahmen des Projektes

CLARIN: Common Language Resources and Technology Infrastructure), dem Max-Planck-Institut für Wissenschaftsgeschichte (MPI-WG), dem Cologne Center for eHumanities (CCeH) und der Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB; im Rahmen von TextGrid) weiterentwickelt. Zur Integration in das TextGrid Laboratory existiert ein Plugin, das NutzerInnen über den Marketplace in ihre Labinstallation integrieren können.

Im Bereich Öffentlichkeitsarbeit und Community-Bildung hat TextGrid in Fortsetzung zu dem 2012 veröffentlichten Imagefilm "Virtuelle Forschungswelten: Neue Wege in den Geisteswissenschaften" 2014 gemeinsam mit DARIAH-DE den Imagefilm "Digitale Wissensräume: Virtuelle Forschungslabore der Geisteswissenschaften" produziert.² Der Film nimmt die in der Praxis angekommenen DH-Methoden, -Tools und -Technologien in aktuellen Projektzusammenhängen sowie in Forschung und Lehre in den Blick. Darüber hinaus wurden zusammen mit DARIAH-DE drei „Erklärfilme“³ zu den Themen „Kollaborativ arbeiten in digitalen Forschungsinfrastrukturen“⁴, „Datensicherheit für Forschungsdaten in den Geisteswissenschaften“⁵ und „Digitales Edieren mit TextGrid“⁶ sowie ein Imagefilm zum DH Summit 2015⁷ in Berlin produziert. Zu sehen sind die Filme und Clips auf dem seit 2014 von TextGrid ebenfalls zusammen mit DARIAH-DE betriebenen DHd-Kanal, der nach gut einem Jahr im Mai 2015 bereits etwa 10.500 Abrufe sowie 130 AbonnentInnen verzeichnete.⁸ Der von TextGrid, DARIAH-DE und dem Max-Planck-Institut für Wissenschaftsgeschichte 2012 eingerichtete DHd-Blog – Digital Humanities im deutschsprachigen Raum⁹ wird von der Community ebenfalls sehr gut angenommen und umfasst im Mai 2015 über 600 Posts von über 100 AutorInnen.

² <https://www.youtube.com/watch?v=tMBq7dlNuLg>

³ Die Kurzfilme (Clips) erklären – jeweils in deutscher und englischer Sprache - in Legetechnik einfach und verständlich die komplexen Zusammenhänge.

⁴ <https://www.youtube.com/watch?v=AWa661KfQGO>

⁵ <https://www.youtube.com/watch?v=jJgrZ4AaVE>

⁶ <https://www.youtube.com/watch?v=jJgrZ4AaVE>

⁷ https://www.youtube.com/watch?v=P2NTLs2IV48&list=PL_aqAE00s6w0hYiKe7MopvpzrW_BKx61U

⁸ <https://www.youtube.com/user/dhdkanal>

⁹ <http://www.dhd-blog.org>

4. Wissenschaftlicher und technischer Stand, an den angeknüpft wurde

4.1 Wissenschaftlicher Stand

Auch am Ende der dritten Projektphase gilt, was bereits zu Beginn des Projekts die nationale wie internationale geisteswissenschaftliche Forschungslandschaft im Blick auf computergestützte Anwendungen, Methoden und Technologien charakterisierte¹⁰: Akzeptanz und Nutzung in den Textwissenschaften weisen, bedingt durch ein bisweilen ambivalentes Verhältnis zum Medium Computer sowie zu den Neuen Technologien insgesamt generell noch immer ein uneinheitliches Bild auf, das vor allem durch lokale Installationen mit jeweils projektbezogenen Applikationen, in weitaus geringerem Maße aber durch Systeme kollaborativer IT-gestützter Methoden geprägt ist. Eine tragfähige übergreifende Infrastruktur, die einen Zugriff auf verteilte Daten- und Computing-Ressourcen auf der Grundlage verbindlicher Standards und Schnittstellen für Digitalisate, Formate und Werkzeuge ermöglichen würde, bleibt ein Desiderat, wo Einzellösungen die mangelnde Netzwerk- und Gridfähigkeit, die fehlende Unterstützung von TEI und Unicode oder auch eine mitunter schwerfällige Usability befördern. Als geisteswissenschaftliche Anwendergruppe mit langer EDV-Tradition kann die Computerlinguistik, insbesondere mit solchen Arbeiten, die sich mit dem Semantic Web und dem automatischen Clustering von Textressourcen befassen, hier indessen bereits vielversprechende Ansätze vorweisen.

Unter dem Begriff des *Distant Reading* lassen sich zudem zunehmende Bestrebungen auch in den Literaturwissenschaften subsumieren, die sehr große Textmengen bzw. Korpora (Big Data) zur Grundlage von IT-gestützten literaturwissenschaftlichen Analyseverfahren machen. Vielbeachtete Argumentationen von den Vordenkern der Digital Humanities, Gregory Crane¹¹, John Unsworth¹² u.a., haben diese neuen Forschungsansätze, Literatur „zu rechnen“¹³, flankiert und einen Paradigmenwechsel statuiert. Ein weiteres zentrales Anwendungsgebiet der Digital Humanities ist zudem das der digitalen Editionen, die entsprechend einen Großteil der mit TextGrid bearbeiteten Forschungsgegenstände ausmachen.

Grundsätzlich lassen sich Techniken, Werkzeuge und Technologien, die bereits in der Informatik oder den Informationswissenschaften entwickelt sind, nicht ohne weiteres und in unveränderter Form in geisteswissenschaftliche Anwendungen transferieren. Innovative Techniken, Werkzeuge und Technologien lassen sich jedoch dort entwickeln, wo sie in Kooperation zwischen Informationswissenschaft einerseits und Fachwissenschaft andererseits entstehen. Solche Kollaborationen bergen nicht nur das Potential, die so genannten eHumanities in ihrem *state of the art* voranzubringen, sondern resultieren ebenfalls in Prozessen der Wechselwirkung, die wiederum auch die Informationswissenschaften befördern. Während zahlreiche GeisteswissenschaftlerInnen

¹⁰ Für die folgenden Ausführungen vgl. TextGrid II-Abschlussbericht

¹¹ Gregory Crane: *What do you do with a million books?*. In: D-Lib Magazine 12/3, 2006 (<http://www.dlib.org/dlib/march06/crane/03crane.html>)

¹² Tanya Clement u.a.: *How Not to Read a Million Books* (<http://people.brandeis.edu/~unsworth/hownot2read.html>)

¹³ Gerhard Lauer: *Literatur rechnen: Lektüre im Computerzeitalter*. In: Frankfurter Allgemeine Zeitung v. 26.08.2009 (<http://www.faz.net/aktuell/feuilleton/geisteswissenschaften/literatur-rechnen-lektuere-im-computerzeitalter-1840973.html>)

jedoch einzelne digitale Werkzeuge, darunter v.a. digitale Wörterbücher, für ihre Forschungsarbeit nutzen, blieb indessen die eigentliche Methodik tradierter geisteswissenschaftlicher Forschungsarbeit bislang noch weitgehend unberührt von digitalen Angeboten und deren Potential; entsprechend fanden digitale Werkzeuge und Dienste in der methodologischen Diskussion und Reflexion bislang nur vereinzelt Niederschlag. In jüngerer und jüngster Zeit zeichnet sich hier jedoch ein Wandel ab. Zusätzlich zu den bereits etablierten und international agierenden Zentren wie dem Centre for Computing in the Humanities, King's College London, und dem Center for Digital Humanities an der Universität Trier lässt sich auf nationaler wie internationaler Ebene ein Zuwachs an ausgewiesenen Digital Humanities-Zentren verzeichnen, so u.a. am University College London und der Royal Irish Academy sowie an den bundesdeutschen Universitäten Darmstadt, Gießen, Göttingen, Hamburg, Köln, Würzburg und weiteren. Aktivitäten auf europäischer und pan-europäischer Ebene verweisen auf die zentrale Rolle, die den Digital Humanities zunehmend nicht nur in Forschung und Lehre, sondern auch als wichtigem Wettbewerbs-Faktor des jeweiligen Forschungsstandortes zugewiesen wird. Dafür steht im europäischen Raum v.a. die ESFRI Roadmap der Europäischen Kommission¹⁴, die Forschungsinfrastrukturen als zentrale strategische Instrumente im Rahmen der europäischen Wissenschaftslandschaft verortet. Auf Bundesebene haben zentrale Förderorgane wie das Bundesministerium für Bildung und Forschung (BMBF) und die Deutsche Forschungsgemeinschaft (DFG) in jüngerer Zeit ebenfalls den Einsatz digitaler Werkzeuge und/oder Virtueller Forschungsumgebungen zum Gegenstand geisteswissenschaftlicher Forschungsausschreibungen gemacht, darunter v.a. die eHumanities-Förderlinie 24 des BMBF (2011/12), der ebenfalls vom BMBF geförderte eHumanities Innovation Award (2012)¹⁵ sowie verschiedene Förderlinien im Rahmen von Horizon2020¹⁶.

4.2 Technischer Stand

Zu Beginn der dritten Förderphase bietet TextGrid eine produktive virtuelle Forschungsumgebung für die Geisteswissenschaften an. Hauptkomponenten dieser Umgebung sind das TextGrid Repository (kurz TextGridRep) und das TextGrid Laboratory (kurz TextGridLab). Ersteres, auch TextGrid Middleware genannt, stellt die Infrastruktur da, die weitestgehend unsichtbar für die NutzerInnen die Funktionen der Forschungsumgebung realisiert. Letzteres ist ein auf Eclipse¹⁷ basierender Client, der die genannten Funktionen nutzerfreundlich über eine grafische Oberfläche bereitstellt. Die folgenden Abschnitte skizzieren den technischen Stand von TextGridRep und TextGridLab, an den angeknüpft wurde.¹⁸

TextGrid Repository

Die Entwicklungsarbeiten der TextGridRep Middleware wurden bis zu Beginn der dritten Förderphase komplett im TextGrid-Konsortium umgesetzt; die gewonnene Expertise konnte entsprechend in die neue Projektphase überführt werden. Es lag gut dokumentierter Quellcode vor; eine umfassende

¹⁴ http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri

¹⁵ <http://www.e-humanities.net/events/2012-ehum-award-call.html>

¹⁶ <http://ec.europa.eu/programmes/horizon2020/en/h2020-section/european-research-infrastructures-including-e-infrastructures>

¹⁷ <https://eclipse.org/>

¹⁸ Zu Beginn der dritten Förderphase lagen das TextGridRep und TextGridLab jeweils in der Version 2.0 vor.

Dokumentation des Gesamtsystems und der zu Grunde liegenden Architektur war allerdings nicht vollumfänglich vorhanden.

Die Middleware besteht aus den Hautkomponenten TG-auth, TG-crud und TG-search. Sie stellen die grundlegenden Funktionen bereit, die für den Betrieb eines Repositorys wie dem TextGridRep benötigt werden, und bilden somit das Rückgrat der Virtuellen Forschungsumgebung für die Geisteswissenschaften.

TG-auth

TG-auth ist die zentrale Infrastrukturkomponente, die die notwendigen Funktionen für föderierte Authentifizierung und Autorisierung sowie Zugriffskontrolle zur Verfügung stellt. TG-auth integriert folgende technische Teilkomponenten:

- Ein auf OpenLDAP¹⁹, einer Implementierung des Standards LDAP, basierender Authentifizierungsserver, in dem TextGrid-Benutzer-Accounts angelegt und verwaltet werden können. Dies gilt für die BenutzerInnen, die sich nicht über die DFN-AAI-Föderation mit ihrem lokalen Campus-Account anmelden können.
- OpenRBAC²⁰ ist eine Implementierung des Standards RBAC (Role Based Access Control), über die die Zugriffskontrolle auf alle TextGrid-Objekte festgelegt werden kann. Diese Komponente verwendet wiederum einen OpenLDAP-Server als Datenbackend für Rollen, Zugriffsrechte etc.
- TGextra ist ein TextGrid-spezifischer Web Service, der die OpenRBAC- Funktionalität erweitert.
- Shibboleth²¹ ist eine Implementierung des Standards SAML (Security Assertions Markup Language), wodurch sich sowohl BenutzerInnen im oben erwähnten LDAP-Server als auch BenutzerInnen, deren Heimatorganisation im Rahmen der DFN-AAI an TextGrid-Dienste eine eindeutige ID schicken, bei TextGrid anmelden können.
- WebAuthN ist eine Web-Anwendung für Browser-basierte Authentifizierung, die LDAP- und SAML-basierte Authentifizierung (auch über die Metaföderation eduGAIN) ermöglicht.

TG-crud

TG-crud (crud steht für "create, retrieve, update, delete" und bezieht sich ursprünglich auf Datenbankoperationen) ist in TextGrid die Komponente, die sich um das Management von Objekten kümmert, die erzeugt, gelesen, aktualisiert oder gelöscht werden. TG-crud nimmt dazu die Anfragen der NutzerInnen an, die zumeist das TextGridLab nutzen und bekommt dabei Daten und Metadaten geliefert. Es verteilt diese Daten dann – je nach Art der Daten – mit unterschiedlichen Methoden auf die unterschiedlichen Datensourcen, die in die Middleware integriert sind. Dazu existieren folgende Instanzen:

- Index-Datenbank Elasticsearch (ermöglicht die Suche über die Inhalte des TextGridRep)
- Triple-Store Sesame, RDF-Datenbank (speichert Beziehungen zwischen TextGrid-Objekten)

¹⁹ <http://www.openldap.org/>

²⁰ http://www.openrbac.de/en_startup.xml

²¹ <http://shibboleth.net/>

- TG-auth, TGextra-crud (Rechteverwaltung der TextGrid-Objekte, realisiert mit OpenRBAC und OpenLDAP)
- das StorNext Filesystem (greift per JavaGAT auf das TextGrid Speicher-Backend zu)
- TG-noid (erzeugt die TextGrid URIs und verwaltet internes und externes Locking der einzelnen Objekte)
- XML-Datenbank eXist (wird über SADE für Publikation projektbasierter Portale genutzt)

Je nach Objekttyp werden noch weitere Aktionen über so genannte Adaptoren angestoßen, z.B. die Extraktion von EXIF-Metadaten aus Bildobjekten und die Übernahme in die TextGrid-Metadaten.

TG-search

TG-search ist ein Kerndienst von TextGrid, der die Suche über alle im Repository enthaltenen Daten ermöglicht. Er erlaubt die Volltextsuche über alle (von NutzerInnen lesbaren) TEI- und XML-Daten sowie die Suche in den TextGrid-Metadaten. Des Weiteren stellt TG-search Navigationsinformationen (z.B. für den Navigator im TextGridLab) und Revisionsinformationen zur Verfügung. TG-search greift für die Suche nach Meta- und Struktur-(Volltext-)Daten auf die ElasticSearch- und die RDF-Datenbank zu.

Für die Navigations- und die Revisionsinformationen sowie für das Filtern nach den letzten Revisionen eines Objektes nutzt TG-search Sesame als RDF-Speicher. Anfragen werden mittels der Abfragesprache SPARQL gestellt und die Ergebnisse werden für die Suche ausgewertet.

Für den nicht öffentlichen Bereich des Repositorys werden TG-search-Treffer mit TG-auth abgeglichen. Auf diesem Weg wird sichergestellt, dass der Nutzer nur Suchtreffer von Objekten angezeigt bekommt, für die er die entsprechenden Rechte hat. Im öffentlichen Bereich des Repositorys wird TG-auth nicht angefragt.

TextGrid Laboratory

Das TextGrid Laboratory (TextGridLab) bildet die Benutzerschnittstelle zum dynamischen Teil des TextGridRep und vereint damit die Werkzeuge zum Erstellen und Bearbeiten von Daten in TextGrid.

TextGrid hat in der letzten Projektphase die stabile Version 2.0 (Mai 2012) veröffentlicht. Diese basiert auf der *Eclipse Rich Client Platform* in Version 3.7. In der Basisversion, die als vollständiger Download von der TextGrid-Homepage erhältlich ist, beinhaltet die Version die Gesamtinfrastruktur, Tools, die für den Betrieb der VRE zentral sind, einzelne fachwissenschaftliche Leuchtturmprojekte (Text-Bild-Linkeditor) bzw. Tools, die bereits produktionsreife Stabilität erlangt hatten (Wörterbuchnetz).

- TextGridLab-Grundkomponenten und Navigator
- Projekt-, User- und Rechtemanagement
- Suche in Lab- und Rep-Inhalten
- Metadateneditor, incl. Integration von PND und Getty-Thesaurus
- XML-Editor
- Text-Bild-Linkeditor

- Wörterbuchnetz-Suche
- Import und Export

Weitere Komponenten können hinzuiinstalliert werden und integrieren sich so in das TextGridLab. Das betrifft zum einen etwa fachspezifische, im Rahmen von TextGrid entwickelte Tools:

- Workflowtool
- Noteneditor
- Kollationierer
- Glosseneditor
- Text-Text-Linkeditor
- Integration Digilib
- Wörterbuch-Linkeditor
- COSMAS-, LEXUS- und ANNEX-Integration
- Web-Preview bzw. Publikation mittels SADE

Zum anderen gilt das aber auch für Tools, die ohne spezifische Anpassung an TextGrid für Eclipse entwickelt wurden, etwa den kommerziellen XML-Editor oXygen oder aus dem Eclipse-Umfeld stammende Editoren für XML-Schema- und CSS-Dateien.

5. Zusammenarbeit mit anderen Stellen

Ein intensiver Austausch mit nationalen und internationalen Digital Humanities-Projekten und -Initiativen ist für TextGrid von Beginn an ein entscheidender Faktor bei der Umsetzung eines der Hauptziele, der bundes- und europaweiten Vernetzung in der DH-Landschaft. Diese Anbindung hat TextGrid durch Kooperationen, Projektpartnerschaften und gemeinsame Entwicklungen auf inhaltlicher, strategischer und/oder technischer Ebene erreicht. Die strukturelle und semantische Interoperabilität der TextGrid-Komponenten spielte dabei eine wesentliche Rolle. Bereits in den ersten beiden Projektphasen initiierte Kooperationen wurden weiter ausgebaut.

5.1 Projekte stellen sich vor

Editionsvorhaben und Forschungsverbünde, die TextGrid für ihre Arbeit nutzen, stellen ihr Digital Humanities-Forschungsvorhaben auf der TextGrid-Website vor.²² Potentielle Nutzerinnen und Nutzer, aber auch solche, die bereits mit der Virtuellen Forschungsumgebung arbeiten, haben so die Möglichkeit, zu spezifischen Themen direkt mit den jeweiligen Ansprechpartnern in Kontakt zu treten.

- [Carl Louis Bargheer: Fiedellieder plus](#)²³
- [eCodicology](#)²⁴
- [Georg Greflinger – Digitale Archiv-Edition](#)²⁵
- [Hybrid-Edition von Theodor Fontanes Notizbüchern](#)²⁶
- [Johann Friedrich Blumenbach – online](#)²⁷
- [RiR – Relationen im Raum](#)²⁸
- [Savigny-Bang-Briefwechsel](#)²⁹
- [SlaVaComp](#)³⁰
- [Virtuelles Skriptorium St. Matthias](#)³¹

5.2 DARIAH-DE

Mit dem Projekt DARIAH-DE, das eine digitale Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften aufbaut und der deutsche Beitrag des EU-Projektes DARIAH-EU ist, wurde im Verlauf der dritten Förderphase auf strategischer, struktureller und technischer Ebene die Zusammenarbeit weiter intensiviert, um Synergien zu nutzen. Viele DARIAH-DE Dienste bauen auf

²² <https://textgrid.de/kooperationsprojekte>

²³ <https://textgrid.de/bargheer>

²⁴ <https://textgrid.de/ecodicology>

²⁵ <http://greflinger.hypotheses.org/>

²⁶ <https://textgrid.de/fontane>

²⁷ <https://textgrid.de/blumenbach>

²⁸ <https://textgrid.de/rir>

²⁹ <https://textgrid.de/savigny-bang>

³⁰ <https://textgrid.de/slavacomp>

³¹ <https://textgrid.de/st-mathias>

Diensten bzw. Erkenntnissen aus TextGrid auf, umgekehrt nutzt TextGrid verschiedene Basis-IT-Komponenten von DARIAH-DE. Die bis zum Projektende von TextGrid erfolgte Migration zahlreicher technologischer Kernkomponenten wie etwa der Nutzerverwaltung und Repository-Technologien in DARIAH-DE gewährleistet die dauerhafte und langfristige Nutzung der Angebote von TextGrid auf neuestem technologischen Stand. In Kooperation mit DARIAH-DE werden außerdem auch nach Projektende von TextGrid weiterhin Nutzertreffen und Workshops durchgeführt werden, die den NutzerInnen die Angebote beider Forschungsumgebungen bzw. -infrastrukturen im Zusammenspiel nahebringen.

Durch das Zusammenwachsen mit DARIAH-DE ist zudem eine Vielzahl von Kooperationspartnern, die Komponenten beider Infrastrukturen nutzen und dadurch auch zu einem weiteren „Zusammenwachsen“ der Communitys beitragen, hinzugekommen.

TextGrid und DARIAH-DE kooperieren aktuell mit 80 Forschungsprojekten (Stand 01.10.2015), davon: mit 62,5% in laufenden Vorhaben (50 Projekte), mit 15% in Vorhaben, die sich derzeit in Begutachtung befinden und bei denen im Frühjahr / Sommer 2016 mit einer Entscheidung zu rechnen ist (12 Projekte), mit 12,5% in Vorhaben, die sich im Antragsstadium bzw. kurz vor der Einreichung befinden (10 Projekte) und mit 10% in abgeschlossenen Vorhaben (8 Projekte), von denen Projektergebnisse über DARIAH-DE und TextGrid gesichert werden. Die Projekte werden von unterschiedlichen Förderinstitutionen betreut, ein ermutigendes Zeichen sowohl für die Offenheit und Innovationskraft der Forschungsinfrastrukturen, ihre Rolle als „Inkubator“ für Spitzenforschung als auch für ihre Nachhaltigkeitsperspektive.

5.3 SADE

Die "Skalierbare Architektur für digitale Editionen", kurz SADE, ist ein Programmpaket zum Entwickeln und Darstellen einer digitalen Edition. Der Fokus liegt dabei auf der Präsentation von XML-Dokumenten, die den Richtlinien der Text Encoding Initiative (TEI) folgen. Die Entwicklung der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW) wird seit 2012 kooperativ v.a. von der BBAW, der Österreichischen Akademie der Wissenschaften (ÖAW; im Rahmen des Projektes CLARIN), dem Max-Planck-Institut für Wissenschaftsgeschichte (MPI-WG), dem Cologne Center for eHumanities (CCeH) und der Niedersächsischen Staats- und Universitätsbibliothek Göttingen (SUB; im Rahmen von TextGrid) weiter entwickelt. Mittels SADE wurde eine Publikationsplattform geschaffen, die es den mit TextGrid arbeitenden Projekten ermöglicht, eigene Webseiten mit ihren erarbeiteten Daten zu füllen und diese auch selbst gestalten zu können.³²

5.4 Digilib & IIF

Auch nach dem Abschluss des von der Mellon Foundation geförderten Projekts "TextGrid Shared Canvas" im März 2014, das die SUB Göttingen zusammen mit der Stanford University durchgeführt hat, beteiligt sich TextGrid an der Weiterentwicklung der Software.³³ TextGrid setzt IIF ein (vgl. z.B.

³² vgl. R 4.4.2 SADE - Dokumentation und Referenzinstanz (http://textgrid.de/fileadmin/TextGrid/reports/TG-R-4_4_2-SADE_Dokumentation_und_Referenzinstanz.pdf)

³³ vgl. Vortrag: DHd Tagung in Graz 2015 (<http://gams.uni-graz.at/o:dhd2015.v.003>)

die Visualisierung des Graphic Recording auf der DH-Summit 2015 Webseite³⁴) und bietet es (auf Nachfrage) seinen Kooperationsprojekten an, z.B.

- Projekt Fontanes Notizbücher (Produktivschaltung Dez. 2015)
- Projekt Virtuelles Scriptorium St. Mathias³⁵

Die Besonderheit bei TextGrid ist, dass es einzelnen Editionsprojekten möglich ist, über den zentralen TextGrid-Bilderservice (digilib) eigene Scans bereitzustellen, auch ohne Bindung an eine Bibliothek oder sonstige Institution.

5.5 TUSTEP

TUSTEP ist ein Werkzeug zur wissenschaftlichen Bearbeitung von Textdaten (auch solchen in nicht-lateinischen Schriften), wie es vor allem in den Geisteswissenschaften benötigt wird. Im Vorprojekt wurde grundsätzlich gezeigt, wie sich TUSTEP in TextGrid integrieren lässt. Zwar wurde diese Integration in TextGrid III technisch nicht weiter verfolgt, allerdings wurden die neuen Entwicklungen in TUSTEP, insbesondere die neue XML-basierte Notation für TUSTEP-Befehle (TXSTEP) von Seiten des Projekts verfolgt und mit Diskussionsbeiträgen begleitet. Grundsätzlich erleichtert die XML-Notation die Integration. Es wäre ein lohnenswertes Folgeprojekt, diese Punkte aufzugreifen.

5.6 Johann Friedrich Blumenbach - Online

In einem Teilprojekt wurde mit dem Akademievorhaben "Blumenbach online" exemplarisch ein Satellitenprojekt eingebunden, wobei es insbesondere um die Integration von neuen Werkzeugen in TextGrid und ihre Verwendung in den jeweiligen Arbeitskontexten ging. Das Kollationierungstool CollateX wurde im ersten Projektjahr an die besonderen Anforderungen bezüglich des synoptischen Vergleichs strukturell ähnlicher Texte (XML-Dokumente; zwei oder mehr) angepasst, insbesondere hinsichtlich der Anzeige/Ausgabe von Gemeinsamkeiten und Differenzen auf struktureller Ebene auch bei Texten mit nur wenigen textlichen Übereinstimmungen. Es wurde ein CollateX-Wrap-Service entwickelt, der als Bindeglied zwischen dem TextGridRep und dem CollateX-Service fungiert, sowie eine Erweiterung des CollateX-Moduls für den TextGrid-Repository-Browser.

5.7 Die DH-Landschaft

Die Landschaft der Digital Humanities Projekte in Deutschland ist sehr lebendig, wie die aktuelle Raum-Zeit-Visualisierung der bisher erfassten DH-Projekte auf dem DARIAH-DE Portal zeigt.³⁶ Hier finden sich auch zahlreiche Vorhaben wieder, die mit TextGrid arbeiten.

³⁴ <https://de.dariah.eu/dhsummit2015>

³⁵ <http://textgridlab.org/1.0/iiif/m2/?json=561f7f1de4b01190df3bcd7a>

³⁶ <https://de.dariah.eu/projekte>

6. Verwendung der Zuwendung und des Ergebnisses im Einzelnen, mit Gegenüberstellung der vorgegebenen Ziele

6.1 AP-übergreifend: TextGrid 1.0., TextGrid 2.0 und TextGrid 3.0

Ein Schwerpunkt der dritten Förderphase von TextGrid war die Verstetigung und Optimierung der TextGrid-Architektur. In diesem Zusammenhang wurde eine umfassende und leicht verständliche Dokumentation aller TextGrid Services (auch in ihrem Zusammenspiel) erstellt. Außerdem sind Services wie die Suchfunktion optimiert und auch neue Funktionalitäten entwickelt worden. Die neu gegründete TaskForce Usability TextGridRep.de hat ein Konzept für die Verbesserung der Usability der TextGrid Repository Webseite erstellt und umgesetzt.

Performanz durch Software-Wechsel

Bei der Suchfunktion bestand Anfang 2013 ein Performanzproblem. Die Suche war an die XML-Datenbank eXist angeschlossen und mit der Datenmenge des Repositorys, vor allem bei zahlreichen gleichzeitigen Schreibzugriffen, überlastet. Aus diesem Grund wurden Alternativen zu eXist getestet und dabei die beiden Search-Anwendungen Solr und Elasticsearch untersucht. Aufgrund der guten Ergebnisse fiel die Wahl auf Elasticsearch. Seit November 2013 läuft das TextGridRep mit Elasticsearch produktiv. Es konnten durch diesen Wechsel auch die Ausfälle des TextGridReps deutlich verringert werden.

Export und Darstellung

Die ursprünglich in eXist integrierte Aufbereitung der in TextGrid gespeicherten TEI-Dokumente wurde in eine eigene Komponente, den *Aggregator*, verlagert. Diese Komponente unterstützt neben der Aufbereitung ganzer Sammlungen (etwa Gedichtbände) auch den Export in Formate wie EPUB zur Darstellung auf E-Book-Readern und Tablets oder TEI-Corpus und ZIP zum massenweisen Export, auch in automatisch umgewandelter Form (etwa nach plain-text).

Server-Architektur

Die Server-Architektur wurde vereinfacht und optimiert, indem ein Wechsel der bis dahin auf mehreren (virtuellen) Servern verteilten Infrastruktur auf einen Server durchgeführt wurde. Ein weiterer Server dient als Backup für die Hochverfügbarkeit (High Availability – HA) des Repositorys. Diese Server sind im ESX-Bereich der GWDG angesiedelt (textgrid-esx1.gwdg.de und textgrid-esx2.gwdg.de). Zusätzlich wurden zwei Testserver in der GWDG ComputeCloud erstellt, um die Produktiv- und die HA-Architektur mit zwei Servern testen zu können.

Für die einfachere und schnellere Konfiguration dieser Server-Architektur wurden Puppet-Module erstellt, mit denen die Server automatisch eingerichtet und aktualisiert werden können. Diese Module werden derzeit (im Rahmen von DARIAH-DE) weiter entwickelt und auf die Anforderungen der Hochverfügbarkeit angepasst.

Usability

Die Usability wurde weiter optimiert, insbesondere die Bedienung der Webpräsenz des TextGrid Repositorys. Die von der TaskForce Usability TextGridRep.de ausgearbeiteten Optimierungsvorschläge wurden im Rahmen des Wechsels der Portal-Software von Typo3 in ein

Liferay-CMS und dem in diesem Zusammenhang vorgenommenen Redesign entsprechend umgesetzt. Neben der intuitiveren Bedienung trägt vor allem eine umfangreiche Dokumentation aller TextGrid-Services zu einer verbesserten Usability bei. Auch die TextGrid-Homepage ist überarbeitet worden. Analog zum Repository wurde das Design angepasst. Eine Umstrukturierung des Menüs ermöglicht eine noch leichtere Navigation.

Integration technischer Komponenten in DARIAH-DE

Im Rahmen der Integration von TextGrid in die DARIAH-DE Infrastruktur wurde eine Zusammenlegung der Nutzeraccounts durchgeführt, die Nutzung beider Services ist nun mit einem DARIAH-DE Account möglich. Die gesamte Autorisierungs- und Authentifizierungs-Infrastruktur (AAI) von TextGrid wurde in die DARIAH-DE AAI integriert. Zudem ist die Möglichkeit, sich auch ohne einen manuell angelegten TextGrid/DARIAH-DE Account im TextGrid Laboratory zu authentifizieren, deutlich verbessert worden. Dies erfolgte durch zwei Maßnahmen: a) die bisherige Authentifizierung über die deutsche Hochschul föderation DFN-AAI wurde auf die europäische Meta-Föderation eduGAIN ausgedehnt. Zugleich wurde b) die Initiative zum „GÉANT Data Protection Code of Conduct for Service Providers“ ratifiziert, die es ermöglicht, dass die von TextGrid benötigten personenbezogenen Daten von den Organisationen skalierbar bereitgestellt werden können. TextGrid nutzt hierbei die Service Privacy Policy der DARIAH-DE Dienste nach.

Durch den Umzug von TextGrid-Homepage und -Repository von einem TYPO 3 Content Management System zur Portal-Software Liferay, mit der auch die DARIAH-DE Webpräsenz umgesetzt ist, lässt sich die gemeinsame Pflege und Weiterentwicklung beider Webpräsenzen noch einfacher umsetzen. Auch auf Systemadministrations- und Entwicklerebene wurde die Migration vollzogen. So werden die gleichen Cloud-Umgebungen genutzt, das gleiche Monitoring sowie die gleichen Entwicklungstools.

Neue Funktionalitäten

Im Zuge der Zusammenarbeit mit den Projekten “Textdatenbank und Wörterbuch des Klassischen Maya”, “Bibliothek der Neologie”, “Johann Friedrich Blumenbach – online” und DARIAH-DE wurde eine RDF-Eingabemaske entwickelt, um Objekt-Metadaten in RDF/XML zu erfassen und in einen Triplestore zu speichern. Diese Eingabemaske benötigt nur ein RDF-Schema für die Metadaten, da diese sich von Projekt zu Projekt unterscheiden können. Die Maske kann daher in generischer Weise ohne weitere Anpassungen genutzt werden, sobald ein RDF-Metadaten-Schema vorhanden ist.

Für das TextGrid Repository wurde eine OAI-PMH-Schnittstelle implementiert. Mit ihr können die Metadaten zu allen Editionen des TextGridRep heruntergeladen werden. Die Schnittstelle implementiert momentan alle erforderlichen Funktionen des OAI-PMH-Protokolls. Im Rahmen des Projektes DARIAH-DE, das ebenfalls diesen Service für das DARIAH-DE Repositorium nutzt, wird diese Schnittstelle noch um weitere Funktionalitäten, wie z.B. Resumption Tokens, erweitert. Momentan können die Metadaten der Objekte des TextGridRep nur als Ganzes heruntergeladen werden. Die Resumption Tokens werden einen schrittweisen Zugriff auf die Daten ermöglichen.

Für eine Möglichkeit, die erstellten TEI-Dokumente (und auch die Metadatenobjekte) neben dem Repository in einem projekteigenen Portal zu veröffentlichen, wurde SADE (Scalable Architecture for Digital Editions) mitsamt eines SADE-Backends, das unter anderem eine eXist-Datenbank nutzt, in das TextGridLab integriert. So können z.B. auch noch vorläufige, in Arbeit befindliche Projektdaten aus dem TextGridLab der Öffentlichkeit präsentiert werden, ohne sie persistent publizieren zu müssen.

Durch diese Variante werden also keine Funktionalitäten der Langzeitarchivierung, wie z.B. die Vergabe von PIDs, zur Verfügung gestellt. Die Nutzung des SADE-Frameworks sowie die technische Umsetzung der Integration in TextGrid ist in Report 4.4.2 *SADE Dokumentation und Referenzinstanz*³⁷ sowie in Milestone 4.4.1 *Update SADE (Dokumentation und Referenzinstanz)*³⁸ dokumentiert.

TextGridLab

Der Fokus bei der Weiterentwicklung des TextGridLab lag auf der Behebung von Bugs, der Anpassung an aktuelle Entwicklungen, der Verbesserung der Usability und der Implementierung kleinerer Features, die von Satellitenprojekten benötigt wurden. In diesem Zuge wurden während der Projektlaufzeit acht getestete und für stabil befundene neue Releases des TextGridLab zur Verfügung gestellt: Zunächst Wartungsreleases zu der aus der vorherigen Projektphase übernommenen Version 2.0, später dann neue Releases aus dem Hauptentwicklungszweig und schließlich das umfangreiche Abschlussrelease TextGridLab 3.0 auf der Basis einer neuen Eclipse-Version. Die neuen Versionen stehen dabei nicht nur zum Download bereit, sondern werden den Benutzern auch direkt aus dem TextGridLab zum automatischen Update angeboten.

Daneben wurde laufend eine (ungetestete) Nightly-Version mit dem aktuellsten Entwicklungsstand zur Verfügung gestellt, um so rasch auf das Feedback einiger *early adopter* reagieren zu können.

Eclipse, das die technische Grundlage des TextGridLab bildet, ist bereits seit einigen Jahren unter dem Label *e4* mit einer Reimplementierung auf der Basis neuer Technologien befasst. Diese Entwicklung war in den vorherigen Projektphasen jedoch noch nicht weit und stabil genug für eine Integration mit dem TextGridLab. Mit dem Eclipse-Release *Kepler* im Jahr 2013 jedoch wurde das erste Mal auf die Veröffentlichung einer Version der 3.x-Versionslinie verzichtet, sodass zur Erhaltung der Zukunftsfähigkeit des TextGridLab eine Migration auf die 4.x-Versionslinie unausweichlich war. Dies ist mit TextGridLab 3.0 nach einigen Anstrengungen aufgrund von Inkompatibilitäten und Verhaltensänderungen gelungen.

Die Codeorganisation und -dokumentation wurde ebenfalls hin zu einer stärkeren Öffnung angepasst. Dies schließt die Migration des Codes vom zentralistischen Versionskontrollsystem *Subversion* zum verteilten und damit für Kollaborationen geeigneteren *git* ebenso ein wie die Einrichtung von Dokumentations- und Quellpaketen, die es Entwicklern ermöglicht, für das TextGridLab zu entwickeln, ohne dessen Quellcode komplett auszuchecken. Daneben erleichtert die neue Codeorganisation das Update von einzelnen Teilen des TextGridLab.

Um Editionsprojekten die Einschätzung zu erleichtern, wie ihre Daten nach der Publikation aussehen werden, wurde eine Vorschaukomponente implementiert und über den Marketplace verteilt. Auch der Vorschaumodus des XML-Editors wurde vor diesem Hintergrund überarbeitet.

6.2 AP 1: Betriebsmodell

Ziel dieses Arbeitspakets war es, Rahmenbedingungen für einen nachhaltigen Betrieb der VFU TextGrid zu schaffen, die einen juristisch fundierten und wirtschaftlich abgesicherten Betrieb der Arbeitsumgebung ermöglichen. Generalisierbare Lösungen für die sich dabei stellenden Probleme wurden von den an AP 1 maßgeblich beteiligten Partnern, dem IDS in Mannheim und der Universität

³⁷ http://textgrid.de/fileadmin/TextGrid/reports/TG-R-4_4_2-SADE_Dokumentation_und_Referenzinstanz.pdf

³⁸ https://textgrid.de/documents/10644/18815/M441_SADE+Dokumentation+und+Referenzinstanz.pdf/8d37e70c-f955-4077-9df1-aa78035a34cc

Würzburg, erarbeitet. Die zentralen Aufgaben dieses Arbeitspakets bestanden im Betrieb einer Geschäftsstelle (AP 1.1) für den zu Beginn der Förderphase zu gründenden eingetragenen und gemeinnützigen Verein „TextGrid – Verein zum nachhaltigen Betrieb einer digitalen Forschungsinfrastruktur in den Geisteswissenschaften e.V.“, in der Erarbeitung eines Vorschlags für eine tragfähige Rechtsform, die für einen Dauerbetrieb von TextGrid die besten rechtlich-organisatorischen Rahmenbedingungen schafft sowie in der Kalkulation der Kosten, die bei einem kontinuierlichen und nachhaltigen Betrieb der Virtuellen Forschungsumgebung TextGrid zu erwarten sind (AP 1.2). TextGrid kann technisch gesehen nur dann nachhaltig arbeiten, wenn Entwicklungsperspektiven für die Werkzeuge vom jetzigen Stand der Technik zu neuen sich abzeichnenden Web-Standards exploriert werden und durch Machbarkeitsstudien und Blaupausen vorbereitet werden. Diesem Aspekt der Nachhaltigkeit an der Schnittstelle von organisatorischem zu technischem Betrieb widmete sich das Unterarbeitspaket 1.3.

6.2.1 AP 1.1 Rechtsform, Kostenkategorien

Basierend auf den Arbeiten im AP 3 „Strukturelle und organisatorische Nachhaltigkeit“ der zweiten Förderphase von TextGrid wurde für die Übergangsphase in den Dauerbetrieb zu Beginn der dritten Förderphase die Gründung eines Vereins vorgeschlagen. Die in diesem Zusammenhang erarbeiteten Statuten und die rechtliche Reglementierung für den technischen Betrieb wurden am Anfang der dritten Projektphase implementiert (vgl. AP 1.2). Diese Maßnahme hatte zum Ziel, dem Betrieb zunächst eine andernorts bereits bewährte, flexible, demokratisch legitimierte Struktur und eine rechtliche Grundlage als gemeinnützige Organisation zu geben. Die gewählte Rechtsform soll den technischen Betrieb in seiner Erprobung in der Übergangsphase begleiten und in der Praxisanwendung ihrerseits auf ihre Akzeptanz beim Nutzer getestet werden. AP 1.1 hat in der Folge die Erfahrungen und Prozesse der täglichen Vereinsarbeit dokumentiert und ausgewertet sowie die Tragfähigkeit dieser Rechtsform und die Akzeptanzkriterien bei potentiellen Nutzern evaluiert. Mögliche Änderungsbedarfe an den organisatorischen Rahmenbedingungen wurden zum Abschluss des Arbeitsprozesses zusammengetragen und mögliche Lösungswege aufgezeigt. Ziel war die Institutionalisierung von TextGrid, die den Gegebenheiten des technischen Betriebs gerecht wird und den Forderungen nach Rechtssicherheit und demokratischer Einflussnahme seitens der NutzerInnen Rechnung trägt. Konkret gilt es zu prüfen, ob die jetzt gewählte Rechtsform des Vereins verstetigt oder eine Alternative gefunden werden muss. Darüber hinaus wurden insbesondere detaillierte Kostenkalkulationen kontinuierlich erstellt, um die bisher noch nicht kalkulierbaren Aufwände einer VFU im Umfang von TextGrid erstmals verlässlich einschätzen zu können. AP 1.1 hat daher anfallende Kosten des täglichen Betriebs erfasst und zum Ende der Förderphase empirisch gestützte Kostenkalkulationen und Kostenmodelle für eine nachhaltige Finanzierung bereitgestellt.

Erfahrungen aus dem organisatorischen Betrieb der Rechtsform

Virtuelle Forschungsumgebungen wie TextGrid versuchen mit ihrem Leistungsportfolio die aktuellen Strömungen der geistes-, kultur- und sozialwissenschaftlichen Forschung aufzugreifen, zu adressieren und voranzutreiben. So heterogen die Anforderungen, Kapazitäten und Lösungen der beteiligten Disziplinen ausfallen, so redundant gestalteten sich oft die Arbeiten an vergleichbaren Problemstellungen. Nach dem Auslaufen einer meist auf Projektbasis erfolgenden Finanzierung, mit dem Weggang verantwortlicher MitarbeiterInnen, mit der Einstellung des Supports für proprietäre Bearbeitungsplattformen oder dem Veralten von digitalen, oftmals nicht allgemein akzeptierten

Standards folgenden Formaten stellt sich recht schnell die Frage nach der Nachhaltigkeit solcher Unterfangen. Das Projekt TextGrid hat zum Ziel, dieses Spannungsfeld zu überbrücken, indem es als VFU nicht nur mit einer für unterschiedliche Nutzeranforderungen und neue technische Entwicklungen offenen und flexiblen technischen Plattform sowie durch möglichst starke fachwissenschaftliche und curriculare Verankerung zur Verstetigung des Erreichten beizutragen, sondern auch ein stabiles und den Anforderungen angepasstes organisatorisches Gerüst zur Verfügung stellt.

Um dem anzustrebenden Organisationsmodell einen rechtlich stabilen Rahmen zu geben, hat sich TextGrid in seiner zweiten Förderphase entschlossen, einen eingetragenen Verein (e.V.) ins Leben zu rufen, der das Projekt im Laufe seiner dritten Förderphase testweise begleiten soll, um zu Projektende als Option der Verstetigung durch Übernahme sämtlicher bisher dem Projektkonsortium zufallenden Aufgaben zur Verfügung zu stehen. Diese Entscheidung basierte auf festen Kriterien, anhand derer mögliche Rechtsformen auf Kompatibilität mit den für eine VFU vermuteten Bedarfen geprüft werden konnten. Jedes Kriterium erhielt eine eigene Wertigkeit mit einer genauen Definition der für TextGrid relevanten und auf das Projekt abgestimmten Anforderungen:

- Gemeinnützigkeit: Herstellung der Förderfähigkeit,
- Haftung: Limitierung von Haftungsansprüchen auf die gewählte Körperschaft,
- Mitgliedschaft: Geringe Hürden, Transparenz, Internationalität, Abbildbarkeit von bisherigen Projektstrukturen, Repräsentativität von Fach-Communities,
- Gründungsaufwände/ -kosten: Beschränkung von Kosten und administrativer Belastung in der Gründung,
- Betriebsaufwände/ -kosten: Beschränkung von Kosten und administrativer Belastung im Betrieb.

Die zentrale Frage im organisatorischen Nachhaltigkeitskonzept von TextGrid war sicherlich, seinen zahlreichen NutzerInnen selbstverständliche Planungssicherheit zu gewährleisten und die so für die Forschung unabdingbare Infrastruktur uneingeschränkt und nachhaltig zur Verfügung zu stellen. Eine künftige Verstetigungsambitionen stützende Mitgliedschaft in dem zu gründenden Verein sollte als starkes Signal an Förderer und Fach-Communities gesandt werden, die VFU auch perspektivisch nutzen und fortentwickeln zu wollen. Zu diesem Zweck wurde der Verein testweise konzipiert, um TextGrid in seiner organisatorischen Stabilität unter Bedingungen des Echtbetriebs operativ zu erproben. Bis zur endgültigen Festlegung der organisatorischen Struktur besteht die vornehmliche Aufgabe des Vereins darin, den dauerhaften Betrieb von TextGrid als digital vernetzter Forschungsinfrastruktur in den Geisteswissenschaften zunächst zu begleiten und zu unterstützen, um die Frage nach seiner Befähigung zur endgültigen Ablösung der Projektstrukturen beantworten zu können.

Aufgrund der temporär beschränkten Auslegung des Projekts und seiner Organe wird es nötig sein, Fähigkeiten und Aufgaben des Konsortiums sukzessive an den Verein oder die jeweils gewählte Nachfolgestruktur zu übertragen, um die von TextGrid entwickelten Produkte auch nach Ende der Förderung zukunftsfähig und nutzerorientiert zu erhalten, bis ein endgültiger Organisationsrahmen für die Virtuelle Forschungsumgebung gefunden wird. Wichtige Kernaspekte sind hierbei insbesondere:

- Nutzerschulungen
- Bedarfserhebung und Technology-Watch
- Beteiligung an Drittmittelvorhaben
- Wissenschaftskommunikation
- Öffentlichkeitsarbeit
- Wissenschaftspolitik

Der TextGrid e.V. konnte mit dieser Zielsetzung bereits 17 akademische Institutionen bzw. individuelle VertreterInnen aus dem Bereich der eHumanities ansprechen und als MitgliederInnen gewinnen, die sämtlich als einflussreiche Akteure die digital arbeitenden Geisteswissenschaften im deutschsprachigen Raum gestalten. Das Spektrum der MitgliederInnen erstreckt sich hierbei von Einzelpersonen, die als ausgewiesene Expertinnen und Experten auf diesem Gebiet die Arbeit des Vereins bereichern, über Universitäten, außeruniversitäre Forschungseinrichtungen und Akademien. Besonders erfreulich ist, dass die Einschlägigkeit des Vorhabens auch über den nationalen Rahmen hinweg erkannt wurde.

Die vor der Gründung festgelegten Kriterien als Eckpunkte für den Erfolg der Vereinsarbeit konnten sämtlich positiv abgearbeitet werden. Der Verein, der seine Gemeinnützigkeit nachweisen konnte und in das Vereinsregister eingetragen wurde, operiert derzeit als Kommunikationsplattform für in den Digital Humanities engagierte Wissenschaftsorganisationen. Um dem Projekt vollumfänglich nachzufolgen zu können, wird er oder jede andere Organisationsstruktur neben seinen Disseminierungsaktivitäten zur zentralen Koordinationsinstanz aller Arbeiten rund um TextGrid und muss als solche Bedarfe hinsichtlich technischer Funktionalität und Weiterentwicklung, Finanzierung und fachwissenschaftlicher Verankerung aufgreifen. Er wird sich dabei nicht nur den FachwissenschaftlerInnen als Ansprechpartner anbieten, sondern auch die Interessen der NutzerInnen der Virtuellen Forschungsumgebung TextGrid gegenüber den Förderern vertreten müssen. Bedarfe könnten sich somit über die Schaffung zusätzlicher funktionaler Einheiten ergeben, die, Satzungsänderungen vermeidend, keine hoheitlichen Aufgaben wahrnehmen werden, sondern ExpertInnen zu jeweils anfallenden Fachgebieten (Management, Entwicklung und technischer Betrieb, Fachwissenschaften) zusammenführt und die einzelnen Aspekte des Nachhaltigkeitskonzepts von TextGrid abbilden.

Zwar hat der Verein im Rahmen seiner Satzung die Möglichkeit verankert, Mitgliedsbeiträge zu erheben, die momentan im Sinne der fachwissenschaftlichen Akzeptanz moderat zum Vereinsvermögen beitragen, doch wird der Verein die mit dem Dauerbetrieb verbundenen Kosten nur zu einem vergleichsweise geringen Teil aus diesen Einnahmen bestreiten können. Die Mitgliedschaftsbeiträge werden sich – insbesondere in der Startphase – in einer Größenordnung bewegen, die sowohl Institutionen mit geringem Budget als auch Einzelpersonen eine aktive Mitgliedschaft erlaubt. Für juristische Personen (Institutionen) wird derzeit ein Mitgliedsbeitrag von 250 Euro pro Jahr erhoben. Für natürliche Personen bleibt er mit 50 Euro deutlich darunter. Rechnet man dies hoch, wird der Verein selbst bei Aufnahme weiterer Institutionen und einer größeren Zahl einzelner NutzerInnen mit Einnahmen nur im unteren fünfstelligen Bereich rechnen können. Wenngleich TextGrid für den Dauerbetrieb demnach zusätzlicher Finanzierungsquellen bedarf, um einen vollfinanzierten Betrieb zu stemmen, scheint es für die Bewältigung zumindest einiger der

genannten Unterstützungsaufgaben hinzureichen. Der vollständige technische Betrieb kommt hingegen ohne weitere Förderung und stabile Ertragsmodelle nicht aus, die sich auf eine breite Basis heterogener Finanzierungsquellen aus Bund und Ländern stützt. Wichtig scheint hierbei die synergetische Nutzung existierender Strukturen in anderen Verbänden – zu denken wäre hier vornehmlich an DARIAH, perspektivisch vielleicht auch an CLARIN – zu sein, um Redundanzen zu vermeiden und eine sichere Einbettung in weitaus stabilere und breiter gefächerte Strukturen, als sie ein Verein bietet, zu ermöglichen.

Kostenkategorien

Die großen Geldgeber des Bundes und der Länder nehmen das Kriterium der Verstetigung beantragter Forschungsvorhaben zusehends in die Formalitäten ihrer Ausschreibungen auf – eine Mindestvorhaltdauer von Daten und Werkzeugen von 10 Jahren gilt bei der Deutschen Forschungsgemeinschaft. Die Frage nach der Finanzierbarkeit der Aufbewahrung und Pflege immer größer werdender digitaler Datenarchive stellt sich auch jenseits gewöhnlicher Projektlaufzeiten insbesondere kleineren Disziplinen immer mehr. Es wird hierbei zusehends deutlicher, dass eine kurzfristige, initial ausgelegte Finanzierung nicht mehr ausreicht, um die auf lange Zeit anfallende Aufgabe der Bestanderhaltung zur Verhinderung von Erkenntnisverlust und Vermeidung kostspieliger, redundanter Neuerhebungen zu stemmen. Der Übergang von einer Einzelförderung in eine institutionelle Förderung ist damit unumgänglich.

Bisher erfolgte die Finanzierung des Aufbaus und – inzwischen in der dritten Förderphase – des testweisen Echtbetriebs von TextGrid über eine zeitlich befristete projektgestützte Förderung. Um eine mittel- bis langfristig stabile (als Richtwert werden zunächst 10 Jahre angenommen) Finanzierung zu realisieren, sollte eine jede Einzelförderung in eine institutionelle Förderung übergehen. Die wesentlichen Akteure, denen eine Finanzierung der VFU perspektivisch zukommen wird, sind hierbei an der Aufschlüsselung der beim Betrieb entstehenden Kosten interessiert, sodass für die weitere Entwicklung eines Finanzierungsmodells die Kostenermittlung im Rahmen eines Kostenmodells erfolgen muss. Der Versuch, dies für eine VFU in den genannten Fachdisziplinen zu unternehmen, stellt eine erste Annäherung dar, zumal es einige Unwägbarkeiten zu berücksichtigen gilt.

Zwar existieren zahlreiche Untersuchungen zu Kostenermittlungen rund um den Umgang mit Forschungsdaten, doch konzentrieren sich diese fast ausschließlich auf den Themenkomplex des Lebenszyklus von Forschungsdaten. Für die auf eine VFU bezogenen Kalkulationen wurde eine induktive Methode angewandt, um von einzelnen Arbeitsschritten zu einer Gesamtkostenaufstellung zu gelangen. Zur Kalkulation der Aufwände, die beim Umgang mit Forschungsdaten im Rahmen einer solchen VFU entstehen, ist somit der entscheidende Schritt, die Faktoren zu identifizieren, die Kosten verursachen. Es wurden also sämtliche Dienste und Werkzeuge, die in einen Dauerbetrieb überführt werden sollen, aufgelistet und die momentan für ihre Inbetriebhaltung notwendigen, durch Monitoring der jeweils zuständigen Stellen errechneten Geldmittel ermittelt.

(Weiter-)Entwicklungskosten wurden hierbei folglich nicht berücksichtigt. Die Kosten selbst untergliedern sich in Personal- (berechnet nach den von der DFG zugrunde gelegten Mittelwerten und zugehörigen Personenmonaten) und Sachmitteln (berechnet nach den im Projekt derzeit notwendigen Aufwänden). Für die Untersuchung wurde folgende Nomenklatur angewendet:

- “Kostenkategorie” bezeichnet den Serviceaspekt einer VFU (Rechnerbetrieb und LZA; Tools, Archiv, Nutzerbetreuung sowie Organisation und Verwaltung).
- “Kostenkomponente” bezeichnet die in der jeweiligen Kategorie anfallenden Dienste und Tools.
- “Einzelposten” bezeichnet die zur Erhaltung der Funktionalität notwendigen Arbeitsschritte.

Grundsätzlich trifft auch bei der Erarbeitung eines Kostenmodells für eine VFU die Unterscheidung zwischen Initial-/Entwicklungskosten und Folge-/Betriebskosten zu. Für das nun für TextGrid zu entwickelnde Modell stehen jedoch die Betriebskosten einer funktionsfähigen VFU im Fokus, wurden die vorgelagerten Kosten des Aufbaus doch bereits in den projektbasierten drei Förderphasen aufgebracht. Für ein späteres Finanzierungskonzept ist demnach zunächst mit einer institutionellen Finanzierung zu rechnen. Die im Laufe des Betriebs anfallenden Anpassungen der bestehenden Architektur zur Erfüllung von Nutzerwünschen und das Aufgreifen neuester technischer Möglichkeiten soll ein solches Modell nicht berücksichtigen. TextGrid hat sich vielmehr entschieden, solche Expansionen der VFU künftig über zusätzliche Projektmittel vorzunehmen. Eine erweiterte Funktionalität und geänderte Service Level Agreements befreien TextGrid allerdings nicht von der Notwendigkeit der Anpassung seines Kostenmodells für den Betrieb. Geänderte Anforderungen ziehen andere Kosten nach sich, sodass das vorliegende Modell lediglich den Istzustand abzubilden vermag.

Bei den durchgeführten Berechnungen überraschte es kaum, dass die Aufwände für MitarbeiterInnen der VFU den Löwenanteil der Gesamtkosten ausmachen: 85% aller Mittel des laufenden Betriebs fließen in Personalkosten. Bei einem Gesamtarbeitsumfang von 12 FTEs pro Jahr entfallen im Rahmen der Kostenverteilung nach Kostenarten (vgl. Abbildung 1) je ein Drittel der aufzubringenden Mittel auf die Langzeitarchivierung der Forschungsdaten (Pflege und Support für das TextGridRep) sowie Pflege und Support der Werkzeuge und Dienste (TextGridLab), während sich das letzte Drittel einerseits auf Nutzerschulungen- und -betreuung (Fachwissenschaftliche Nachhaltigkeit) zu fast einem Viertel der Gesamtkosten und andererseits den Erhalt der Basisinfrastruktur (Server, Sicherheit, Storage, Persistent Identifiers etc.) bzw. den administrativen Betrieb (Projektmanagement, Öffentlichkeitsarbeit, Rechtsform, Dienstreisen etc.) verteilt.

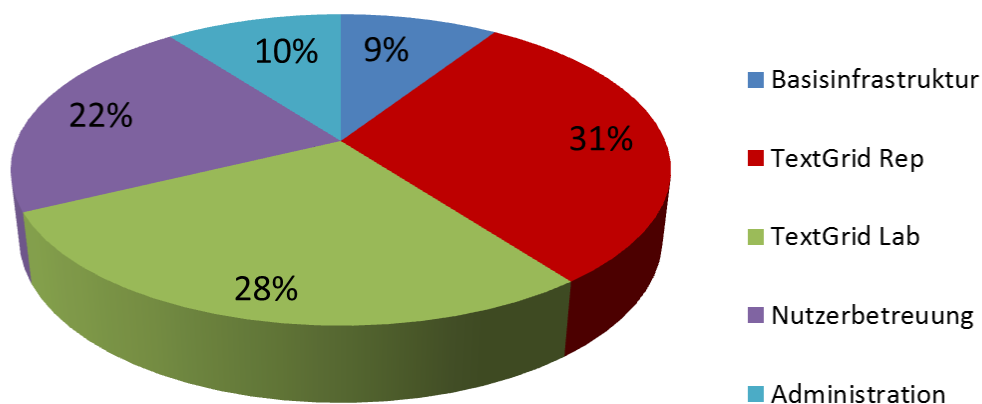


Abbildung 1: Kostenverteilung in TextGrid (Stand Dezember 2013)

Zusammengenommen setzt TextGrid auf ein modulares Betriebskonzept mit autonomen Servicekomponenten. Um den Bedarf der wissenschaftlichen Community sinnvoll und vollständig abdecken zu können, empfiehlt sich ein auf eine gemeinsame organisatorische Basis sowie einer hinreichend ausbaufähigen, zuverlässigen aber kostengünstigen Hardwareausstattung gestützter Dauerbetrieb im Zusammenspiel aus einem Anwendungslabor für Forschungsdaten, einem Langzeitarchiv und beständiger wissenschaftlicher Begleitung des/der Nutzers/Nutzerin. Diese Modularisierung entspricht der Philosophie des Projekts TextGrid hinsichtlich technischer, fachwissenschaftlicher und organisatorischer Nachhaltigkeit einer VFU. TextGrid bietet damit eine flexible und offene Struktur, die, den nachhaltigen Bestand aller Komponenten vorausgesetzt, nicht nur den Anforderungen von NutzerInnen aus dem Bereich der Digital Humanities an ein modernes Forschungsdatenmanagement entgegenkommt, sondern den WissenschaftlerInnen auch eine organisatorisch stabile Umgebung vorhält und sie mit den teils komplexen Herausforderungen im Umgang mit Forschungsdaten nicht allein lässt.

6.2.2 AP 1.2 Verein, Geschäftsstelle

Der TextGrid-Verein benötigt eine Geschäftsstelle, die alle den Verein betreffenden organisatorischen Tätigkeiten übernimmt. Im Unterschied zu dem in AP 7 konzentrierten Projektmanagement beschränkt sich die Geschäftsstelle ausschließlich auf die unmittelbar den Verein selbst betreffenden organisatorischen Aufgaben, arbeitet dabei aber eng mit dem Projektmanagement zusammen. Die Geschäftsstelle untersteht satzungsgemäß dem Vorstand des Vereins und unterstützt diesen bei der Erfüllung seiner Aufgaben. Zu den wesentlichen Tätigkeitsfeldern der Geschäftsstelle gehören insbesondere die administrative Vereinsarbeit, die Bearbeitung rechtlicher Fragestellungen, Hilfestellungen bei der Einberufung der Mitgliederversammlungen und die Unterstützung des Vorstandes bei der Abfassung der Jahresberichte etc.

Vereinsgründung

Die Gründung des Vereins vollzog sich in folgenden Schritten:

- 06.08.2012: Notarielle Anmeldung des Vereins durch den Vorstand (L. Eichinger, A. Rapp in Vertretung);
- 18.12.2012: Antrag auf Anerkennung der Gemeinnützigkeit des Vereins; Einleitung des Eintragsverfahrens beim Amtsgericht Mannheim;
- 03.01.2013: Vorläufige Anerkennung der Gemeinnützigkeit des Vereins durch das Finanzamt Mannheim mit Auflagen zur Satzungsänderung,
- 26.06.2013: Abschluss des Eintragsverfahrens am Amtsgericht Mannheim durch Vorlage von Vollmachten zur Unterschriftsberechtigung sämtlicher Gründer und Einholung von Vertretungsnachweisen aller Vollmachtgeber,
- 17.07.2013: Notarielle Eintragung des TextGrid-Vereins durch den Vorstand unter der Registernummer VR 700598 (L. Eichinger, A. Rapp in Vertretung). Die für die Eintragung des Vereins anfallenden Gebühren wurden aus den Mitteln des IDS beglichen.

- Die zur Ankerkennung der Gemeinnützigkeit notwendigen Änderungen der Vereinssatzung (insb. Mittelverwendung bei Vereinsauflösung) wurden am 20.05.2014 notariell unter Anwesenheit zweier vertretungsberechtigter Vorstände durchgeführt. Die Gemeinnützigkeit wurde dem Verein bereits am 11.02.2014 zuerkannt, die Zuteilung einer Steuernummer erfolgte am 02.06.2014.

Vereinsarbeit

- Der TextGrid-Verein hat im Berichtszeitraum fünf Mitgliederversammlungen und eine Sitzung des Vorstandes durchgeführt:
- 21.09.2012: Sitzung des TextGrid-Vorstandes mit der Entscheidung über den Aufnahmeantrag der Berlin-Brandenburgischen Akademie der Wissenschaften als Vereinsmitglied;
- 23.10.2012: 1. Mitgliederversammlung mit ersten Absprachen zur weiteren Vorgehensweise im Verein;
- 07.06.2013: 2. Mitgliederversammlung des Vereins mit Bericht über den Stand des Gründungsverfahrens und Diskussion zur Arbeitsweise des Vereins. Hierbei wurden die Eckpunkte der Vereinsarbeit skizziert. Auftrag und Arbeitsweise des Vereins leiten sich demnach aus der Satzung ab (Koordination des Fachwissenschaftler, Betrieb u. Pflege, Nutzerbetreuung, Öffentlichkeitsarbeit, Qualitätssicherung). Hierbei hat sich die Umsetzung dieser Ziele am künftigen Finanzierungsmodell auszurichten. In den nächsten Schritten wird ein Fahrplan zur Übergabe von Kompetenzen vom Projekt auf den Verein erarbeitet.
- 10.07.2013: Konzeption und Freischaltung einer Vereinswebseite auf der eigens erworbenen Domain <http://www.textgrid-verein.de/> unter Finanzierung einer Lizenz zur Nachnutzung des Corporate Designs von TextGrid. Inhalt: Vereinszweck, Mitglieder, Gremien, Satzung, Links, Kontaktpersonen.
- 25.11.2013: 3. Mitgliederversammlung des Vereins mit Bericht über den Stand des Gründungsverfahrens und Beschluss der zur Anerkennung der Gemeinnützigkeit notwendigen Satzungsänderungen;
- 16.12.2014: 4. Mitgliederversammlung mit der Diskussion zur Finanzlage des Vereins, der Entlastung des Vorstandes sowie Konkretisierung möglicher Finanzierungskonzepte und des künftigen Aufgabenspektrums des Vereins. Somit steht der Verein als Institutionalisierungsoption nach Projektende bereit.
- 16.04.2015: 5. Mitgliederversammlung mit der Wahl eines neuen Vorstandes und der Bestellung der Geschäftsstelle für die weitere Mandatszeit.

Finanzierung des Vereins

Zur Finanzierung des Vereins entrichten alle Mitglieder die in der Satzung bestimmten Mitgliedsbeiträge. Die Rechnungen werden jeweils am Ende des Kalenderjahres an die Mitglieder versandt. Zum Abschluss des Kalenderjahres 2014 erfolgte erstmals ein Kassenbericht, im Zuge dessen der Vorstand entlastet wurde. Die Diskussion innerhalb der Mitglieder ergab, dass die auf diese Weise erzeugten Einnahmen für flankierende Maßnahmen zum eigentlichen technischen

Betrieb der VFU bereitgestellt werden sollen. Mit Wirkung zum 01.03.2014 hat der Verein ein eigenes Konto mit den notwendigen Vollmachten bei der Commerzbank AG eingerichtet.

6.2.3 AP 1.3 (Technische) Integration und zukünftiges Weiterentwicklungspotential, Zukunftsfähigkeit

Responsive Design für textgridrep.de

Die Benutzung von Smartphones und Tablets gerade zum Rezipieren von Inhalten hat in den letzten Jahren stark zugenommen. Für das Web-Frontend zum TextGridRep wurde deshalb das Design mit *Responsive-Design*-Techniken so weiterentwickelt, dass auf Smartphones und Tablets automatisch eine an die spezifischen Bildschirmformate angepasste Darstellung gewählt wird.

E-Book-Export

Es wurde ein Tool entwickelt, mit dem TEI-Inhalte aus TextGrid als E-Books im EPUB-Format exportiert und so etwa auf E-Book-Readern, Tablets und Smartphones gelesen werden können. Der Export ist dabei auf einer beliebigen Ebene des Repositorys möglich, es können also z.B. ganze Sammlungen exportiert werden. Das Tool lief zunächst auf einem Testserver und wurde Anfang 2013 nach einer Überarbeitung der formatierenden Stylesheets in die Produktivversion des Portals übernommen.

Webversion TextGridLab

Das TextGridLab bietet als *Rich Client* zwar ein leistungsfähiges und erweiterbares User Interface – und z.B. ein webbasierter XML-Editor mit dem Leistungsumfang von etwa oXygen existiert auch heute nicht –, Webtechniken haben allerdings in ihrer Leistungsfähigkeit deutlich zugenommen und bieten den Vorteil eines installationsfreien, plattformübergreifenden Zugangs. In diesem AP wurden deshalb Möglichkeiten evaluiert, eine Weboberfläche für den dynamischen Teil von TextGrid zu entwickeln. Dabei sollte insbesondere die Komplexitätsreduktion der Oberfläche eine Rolle spielen – die mit der Flexibilität und Anpassbarkeit der Oberfläche des TextGridLab einhergehende Komplexität erfordert einen gewissen Lernaufwand und wird insbesondere von gelegentlichen NutzerInnen bisweilen kritisiert.

Wir haben deshalb begonnen, einen experimentellen webbasierten Zugang auch zum dynamischen Bereich von TextGrid zu entwickeln. Dies geschah in Form einer *single page webapp*, die direkt mit den regulären Schnittstellen des TextGrid-Repository arbeitet und wenig lokale Installation erfordert. Damit kann sie als Grundlage / Steinbruch für künftige projektspezifische Entwicklungen dienen.

In diesem Rahmen wurden (bedarfsorientiert) JavaScript-Zugänge für die Kernutilities TG-auth*, TG-crud und TG-search entwickelt und dabei auf ein populäres MVC-Framework, *backbone.js*, zurückgegriffen. Navigations- und Anzeigemöglichkeiten wurden ebenso wie ein einfacher Editor integriert und dabei eine objekttypenabhängige Pluginschnittstelle implementiert. Objekte können zu Sammlungen zusammengestellt und gemeinsam exportiert, aber die Sammlungen auch als Aggregationen wieder in TextGrid gespeichert werden. Für Projekte wurde die Verwaltung der Benutzerrollen und -rechte umgesetzt.

Die Implementierung ist effizient und reaktiv, hat aber auch die Schwierigkeiten und Grenzen dieses Ansatzes gezeigt – unter anderem einen recht hohen Aufwand für die Reimplementierung von in den Java-basierten Bibliotheken bereits vorhandener oder angelegter Grundfunktionalität in JavaScript,

Zusatzaufwand für die Bedienung bestimmter TextGrid-APIs, die teilweise nicht auf die Bedürfnisse von JavaScript-basierten Webapps ausgerichtet sind, und das für diesen Use Case noch nicht endgültig gelöste Problem der Einbindung der Authentifizierung.

Eine Auswertung dieses Ansatzes wird in R 1.3.1 *Machbarkeitsstudien zu neuen technischen Entwicklungen für TextGridLab*³⁹ vorgenommen.

Zukunftsfähigkeit TextGridLab

TextGridLab 1.0 und 2.0 beruhen auf der Eclipse-Versionsreihe 3.x. Eclipse verfolgt bereits seit einiger Zeit eine neue, architektonisch umgebaute Versionsreihe 4.x, und mit dem aktuellen Eclipse 4.3 (Kepler) ist es erstmals zu einem simultaneous release ohne eine 3.x-Version gekommen.

Für das TextGridLab wurde die Migration auf die Eclipse-Version 4.3 negativ evaluiert, da in Verbindung mit dieser Eclipse-Version eine Reihe von Bugs auftraten, die nicht endgültig diagnostiziert und gelöst werden konnten. Ein technologisches Upgrade des z.Zt. auf Eclipse 3.7 basierenden TextGridLab bis Projektende ist dennoch unausweichlich, auch um der technologischen Weiterentwicklung der Plattformen Rechnung zu tragen. Eine Portierung auf die zum Projektende aktuelle Version *Eclipse 4.4 (Luna)* konnte erfolgreich umgesetzt werden, so dass die für den Fall des Scheiterns vorbereitete Migration auf den letzten Stand der Eclipse-3-Reihe nicht zur Anwendung kommen musste.

Mehr Interaktivität im Lab mit ECF

Für das TextGridLab wurde testhalber das Eclipse Communication Framework (ECF) integriert und ein Konzept für eine produktive Integration erstellt. Mögliche Szenarien damit:

- Kollaboratives Arbeiten: An einem TextGrid-Dokument können zwei BenutzerInnen gleichzeitig arbeiten, die Änderungen werden dabei sofort synchronisiert und jeder sieht die Markierungen des jeweils anderen.
- Über einen in die Lab-Oberfläche integrierten Chat können z.B. MitarbeiterInnen am selben Projekt miteinander kommunizieren.
- Eine Integration von Präsenz- und Chatdiensten mit der TextGrid-Benutzerverwaltung könnte die Kollaboration von an mehreren Standorten an einem Projekt arbeitenden Benutzerinnen und Benutzern erleichtern.
- Neben Editoren können auch weitere Bildelemente der eigenen Arbeitsumgebung beim Partner eingeblendet werden.

Massenexport aus dem Repository

Für die Entwicklung analytischer/Big-Data-Verfahren ist es üblich, Korpora zusammenzustellen, die dann exportiert und mit unterschiedlichen, oft von den Forscherinnen und Forschern selbst entwickelten lokalen Werkzeugen weiterverarbeitet werden. Um diese Anforderungen zu bedienen, wurde die Exportschnittstelle weiterentwickelt, u.a. um Umfang und Dateinamen der exportierten Daten näher zu bestimmen, die zu exportierenden Daten in ein von BenutzerInnen gewünschtes Zielformat zu transformieren, und den Umgang mit Metadaten auszuwählen. Außerdem wurde ein

³⁹ <https://textgrid.de/documents/10644/18815/R+1.3.1+und+M+1.3.2.pdf>

Ansatz zum Export aller Ergebnisse einer Suchanfrage (z.B. „alle Gedichte in deutscher Sprache“) implementiert. Dies wurde etwa vom Metricalizer-Projekt für die Einbindung aller Gedichte des Repositories in deren Werkzeug zur Analyse von Gedichten genutzt (z.Zt. in einer Testversion).

Umgang mit großen, strukturierten TEI-Dokumenten

Für das TextGridRep wurde ein Konzept zur Darstellung, Repräsentation und Suche von Teilen großer TEI-basierter Inhalte mitentwickelt und die Darstellungskomponente wurde implementiert. Dies betrifft beispielsweise die Enzyklopädien der Digitalen Bibliothek.

6.3 AP 2: Prozessorganisation

Das AP hatte das Ziel, Prozesse zur Verwaltung großer heterogener Nutzergruppen zu etablieren.

Dabei spielten mehrere Aspekte eine Rolle: Zum einen die Frage, wie vorhandene NutzerInnen betreut und deren wachsendes Expertenwissen in die wachsende TextGrid-Community integriert werden kann, und zum anderen, wie neue Nutzer- und Entwicklergruppen gewonnen werden können (Community Building). Dazu gehörte die Entwicklung eines Beratungskonzepts für unterschiedliche Nutzergruppen anhand von exemplarischen Workflows aus der wissenschaftlichen Nutzung.

Der zweite Aspekt betraf die Dokumentation, die Entwicklung von Richtlinien und Qualitätsstandards für Plugins bzw. ergänzende Tools für TextGridLab und TextGridRep, um künftigen EntwicklerInnen und NutzerInnen selbsterklärende Vorlagen zu liefern (Marketspace für die Toolentwicklung).

Schließlich war der dritte Aspekt die Entwicklung und Etablierung eines Finanzierungsmodells für die Verstetigung des bisher in TextGrid erreichten Standards.

Für die Umsetzung der genannten Ziele sind drei Unterarbeitspakete eingerichtet worden:

- AP 2.1: Umgang mit großen Nutzerzahlen und heterogenen Nutzergruppen sowie Nutzerverwaltung (Praxis und Theorie), verantwortet von der TU Darmstadt
- AP 2.2: Verteilte / Zentrale Entwicklungen, verantwortet vom MPIWG
- und AP 2.3: Finanzierungsmodelle und -wege, verantwortet von ProWiss

Das AP hat sich bis zum Erreichen dieser Ziele mehrere Meilensteine gesetzt und die laufenden Ergebnisse unter anderem mit öffentlich zugänglichen Reports dokumentiert.

6.3.1 AP 2.1 Umgang mit großen Nutzerzahlen und heterogenen Nutzergruppen sowie Nutzerverwaltung (Praxis und Theorie) und AP 2.2 Verteilte / Zentrale Entwicklungen

M 2.1.1 (Laufende) Kartierung der heterogenen Nutzercommunity

Zur Kartierung der heterogenen Nutzercommunity wurden die Nutzerdaten der TextGrid-Accounts gemäß geltender Datenschutzbestimmungen anonymisiert ausgewertet und graphisch aufbereitet. Auf Basis der entwickelten Daten wurden zahlreiche Abbildungen erstellt, die die quantitative Entwicklung der TextGrid-Accounts seit 2009 sowie deren geographische Verteilung innerhalb Deutschlands und weltweit graphisch darlegen. Mit Hilfe der Daten konnte ein stetiges Wachstum von im Mittel ungefähr 20 neuen Accounts pro Monat über den gesamten Beobachtungszeitraum hinweg nachgewiesen werden. Auch konnten durch die Daten Wissenschaftsstandorte mit großem

Interesse an TextGrid identifiziert werden, was insbesondere mit Blick auf die Forschungseinrichtungen außerhalb des TextGrid-Konsortiums aufschlussreich ist. Des Weiteren wurden die Daten aufbereitet, um eine Visualisierung der zeitabhängigen Entwicklung der Accounts mit dem DARIAH-DE Geobrowser zu ermöglichen. Eine Analyse der fachlichen Zusammensetzung der Nutzercommunity von TextGrid war aufgrund der starken Überschneidungen bei der fachlichen Zuordnung nur eingeschränkt möglich. Da es bei den Teilcommunitys häufig Überlappungsbereiche gibt, ist in vielen Fällen eine eindeutige Zuordnung der NutzerInnen zu genau einer Teilcommunity nicht möglich.

M 2.1.2 Einrichtung eines Nutzerforums und dauerhafte Pflege

Nach ausführlicher Diskussion über die Wahl der am besten geeigneten Forumslösung für TextGrid fiel im Frühjahr 2013 die Entscheidung, das TextGrid-Forum in das zum damaligen Zeitpunkt neu zur Verfügung stehende DARIAH-DE Liferay-Portal einzubinden. Die Integrationslösung sollte auch die enge Zusammenarbeit zwischen den beiden Projekten unterstreichen und ermöglicht dabei Synergieeffekte in den Bereichen Betrieb und Nutzerverwaltung. Nach einer kurzen Testphase im Juni 2013 und der Behebung noch vorhandener technischer Probleme wurde das Nutzerforum im Herbst 2013 den TextGrid-NutzerInnen zugänglich gemacht. Vorgestellt wurde die Kommunikationsplattform bereits am 21. Juni 2013 im Rahmen des 2. TextGrid-Nutzertreffens in Mainz; bei späteren Veranstaltungen, insbesondere den weiteren Nutzertreffen, wurde das Forum wiederholt beworben. Das nutzerbetriebene Forum wird auch im Report R 3.3.1 *Erfahrungen aus den Nutzerschulungen, Empfehlungen für die weitere Ausgestaltung des Konzepts*⁴⁰ im Zusammenhang mit den weiteren Formen der Wissensvermittlung beschrieben.

Die Erwartungen an einen hohen Nutzungsgrad und ein entsprechendes Aktivitätsniveau dieses Kommunikations- und Supportkanals haben sich bis Ende der Förderphase im Mai 2015 trotz intensiver Bewerbung nicht erfüllt. Das Forum soll daher nicht länger angeboten werden, um einen möglichen negativen Effekt auf die Außendarstellung aufgrund mangelnder Aktivität zu vermeiden.

M 2.2.1 Entwicklerworkshop mit externen Projektpartnern in Kooperation mit AP 5.4

- Erster Entwicklerworkshop „Technische Anforderungen von Nutzern an TextGrid“
- Dieser Entwicklerworkshop fand am am 18. und 19. April 2013 am MPIWG statt und hatte das Ziel, Anforderungen für weitere Entwicklungen von TextGrid herauszuarbeiten.
- Anhand von Best-Practice-Beispielen diskutierten aktive TextGrid-NutzerInnen und -EntwicklerInnen sowie NutzerInnen anderer relevanter Infrastrukturen Anforderungen, die an eine Infrastruktur und Arbeitsumgebung wie TextGrid anlegt werden. Dabei wurden vor allem übergreifende Fragen der Zielgruppenrelevanz, der Langzeitarchivierung und Lizenzierung, des Publizieren aus TextGrid heraus sowie Stabilitäts- und Sicherheitsaspekte aus Sicht der NutzerInnen und EntwicklerInnen angesprochen. In diesem Zusammenhang wurde die Zusammenarbeit zwischen den KollegInnen der Technischen Universität Darmstadt und des Max-Planck-Instituts für Wissenschaftsgeschichte intensiviert. Die Themen und Diskussionsergebnisse des Workshops dienten dabei der Vorbereitung von Nutzertreffen, darüber hinaus flossen die Rückmeldungen der TextGrid-NutzerInnen aus der

⁴⁰ https://textgrid.de/fileadmin/TextGrid/reports/TextGrid_Report_3_3_1_Final.pdf

Nachbereitung vorangegangener Treffen in die Diskussion des Workshops als Feedback und Bedarfsanalyse ein.

- Zweiter Entwicklerworkshop „Usability in Forschungsinfrastrukturen in den digitalen Geisteswissenschaften“
 - Der zweite Entwicklerworkshop fand am 22. und 23. September 2014 am DAI statt und wurde in Zusammenarbeit mit Cluster 1 von DARIAH-DE organisiert und durchgeführt. Der Workshop führte EntwicklerInnen und AnwenderInnen von verschiedenen Forschungsinfrastrukturen und Best-Practice-Projekten aus dem geisteswissenschaftlichen Bereich – größtenteils nicht aus dem Umfeld von TextGrid und DARIAH-DE – zusammen. Die Teilnehmer diskutierten Anforderungen und Optimierungsmöglichkeiten von Usability insbesondere in neugeschaffenen digitalen Kommunikations- und Arbeitsumgebungen.
 - Die WorkshopteilnehmerInnen stimmten darin überein, dass umfassende und grundlegende Forschung zum Thema Usability in Forschungsinfrastrukturen bzw. generell im Digital-Humanities-Bereich kaum durchgeführt werde und Usability-Überlegungen und -Strategien zu wenig Beachtung fänden.
 - Ferner wurde festgestellt, dass zu wenig über die NutzerInnen und ihr Nutzungsverhalten von Digital Humanities-Werkzeugen und -infrastrukturen bekannt sei, obwohl Forschungsförderung, Projektmanagement und Fragen der Verbreitung, Nachhaltigkeit sowie Nachnutzbarkeit auf den Erwartungen und dem Verhalten der NutzerInnen aufbauen (sollen).
 - Ein Ergebnis des Workshops ist ein Katalog von Empfehlungen, die den Bereich der Usability-Forschung im Rahmen der Infrastrukturentwicklung in den digitalen Geisteswissenschaften besser etabliert und berücksichtigt sehen wollen (siehe auch R 5.3.3 *Usability in Forschungsinfrastrukturen in den digitalen Geisteswissenschaften*⁴¹)
- Dritter Entwicklerworkshop „Annotieren ↔ Publizieren“
 - Der dritte Entwicklerworkshop fand am 20. November 2014 am MPIWG statt.
 - Im Rahmen dieses eintägigen Workshops kamen 15 EntwicklerInnen und NutzerInnen von DH-Werkzeugen und Forschungsumgebungen aus sieben Wissenschaftseinrichtungen im MPIWG zusammen, um anhand von Best-Practice-Beispielen die Konzeption, Gestaltung und Einsatzmöglichkeiten von Tools zum Annotieren von Text- und Bilddateien sowie Publikationswerkzeugen zu diskutieren. Das TextGrid-eigene Publikationstool SADE zeigte sich entlang der Diskussion als mächtiges Werkzeug, das den Anforderungen der Wissenschaftlergemeinschaft bereits entspricht. Zur Einbindung weiterer externer Annotationstools in das TextGridLab und zur Optimierung vorhandener TextGrid-Werkzeuge wurden für das erste Quartal 2015 zusätzliche Entwicklersprints vereinbart.

⁴¹ https://textgrid.de/fileadmin/TextGrid/reports/TextGrid_Report5_3_3_final.pdf

M 2.2.2 Bereitstellung des Market Space / Marketplace

Dieser Meilenstein beschreibt die Implementation des „Marketplace“, dem Informationssystem für Plugins in TextGrid. Da das Konzept des Marketplace auch die Integration von Kommunikationsinstrumenten und Interaktionswerkzeugen vorsieht, wurde das bereits bestehende Angebot dieser Werkzeuge gesichtet und daraus Empfehlungen für eine bessere Gestaltung des Informationsangebots herausgearbeitet. Zusätzlich dazu wurde eine neue Katalog-Software, die die Verbindung des TextGridLab mit der Dokumentation der externen Plugins herstellt, entwickelt und in das bestehende Informationsangebot integriert. Die Präsentation des Marketplaces erfolgte auf dem Konsortialtreffen im Dezember 2014.

R 2.2.2 (Laufende) Dokumentation von externen Plugins⁴²

Dieser Report ging einher mit der Bereitstellung des Market Space/ Marketplace.

Darin wird die Neuimplementierung des TextGrid-Marketplaces beschrieben, der hauptsächlich zwei Aspekte behandelt: zum einen der Marketplace als eine erweiterte Informations-Plattform, die über eine reine Dokumentation der externen Plugins hinausgeht und über sämtliche Belange von TextGrid informieren will und die bestehende Dokumentation und Interaktionsmöglichkeiten bündelt; zum anderen die Dokumentation der externen Plugins für TextGrid, für die Strukturen geschaffen werden, wie die Beiträge von TextGrid-NutzerInnen in Form von funktionalitätserweiternden Plugins reibungslos in das bestehende Angebot eingefügt werden können und wie das schon bestehende Angebot dokumentiert werden kann, um Doppelentwicklungen zu verhindern. Dazu stellt der Bericht Dokumentationskriterien und Methoden zur Qualitätssicherung vor, die in Zukunft – auch in Bezug auf die Verschmelzung mit DARIAH-DE – implementiert werden sollten.

6.3.2 AP 2.3: Finanzierungsmodelle und -wege, verantwortet von ProWiss

Innerhalb dieses Unterpakets gab es keine Meilensteine/ Berichte. Das Aufgabenfeld dieses Unterpakets umfasste die Punkte „Kostenfaktoren, Kostenmodelle, Organisationsstrukturen“ und „Nutzerkommunikation, Monitoring, Qualitätssicherung, Projektmanagement, Öffentlichkeitsarbeit“:

Kostenfaktoren, Kostenmodelle, Organisationsstrukturen

- Mitwirkung an der Erarbeitung unterschiedlicher Kostenfaktoren und -kategorien unter Federführung des Instituts für Deutsche Sprache (IDS) in Mannheim sowie Ausarbeitung und Diskussion des Kostenmodells mit Differenzierung nach Infrastruktur, Betrieb/Software-Pflege TextGridLab und TextGridRep, Nutzerbetreuung und laufendem organisatorischem Aufwand: Die betreffenden Konzepte liegen vor.
- Weiterentwicklung, Prüfung und Diskussion möglicher Finanzierungsvarianten (institutionelle Grundfinanzierung, Eigenanteile, Gebühren, Mitgliedsbeiträge, Drittmittel, Projektpauschalen etc.) und entsprechender Gestaltungsoptionen für die Trägerschaft (z.B. TextGrid-Verein, Verbund von Hochschulen und/oder Akademien, Leibniz-Institut) sowie Differenzierung nach lang-, mittel- und kurzfristigem Finanzierungsbedarf: Entsprechende Konzepte liegen vor.

⁴² https://textgrid.de/fileadmin/TextGrid/reports/TextGrid_Report2_2_2.pdf

- Mitwirkung an der Erarbeitung der Satzung des TextGrid-Vereins sowie Teilnahme an den Mitgliederversammlungen mit Berichterstattung zur institutionellen Verstetigung von TextGrid: Die Satzung ist erarbeitet, der Verein gegründet und als gemeinnützig anerkannt. Mehrere Mitgliederversammlungen haben bereits stattgefunden.
- Entwicklung und Diskussion eines sog. „Prämien- und Partnermodells“ als Angebot für Länder mit besonders DH-aktiven wissenschaftlichen Einrichtungen zur Beteiligung an der Grundfinanzierung im Zusammenwirken mit dem Bund: Das Modell wird derzeit im Detail ausgearbeitet und die Grundlage für die weiteren Sondierungsgespräche zur Verstetigung von TextGrid bilden.
- Mitwirkung an der Erarbeitung eines Positionspapiers zum Verhältnis TextGrid/DARIAH-DE und von Szenarien zur nachhaltigen Zusammenführung beider Vorhaben sowie Mitwirkung im Arbeitskreis „Nachhaltigkeit / Finanzierungskonzept“ von TextGrid und DARIAH-DE: Ein erster Entwurf befindet sich derzeit in der Überarbeitung.
- Mitwirkung an der Erarbeitung und Fortschreibung des Dossiers „Nachhaltigkeit digitaler Forschungsinfrastrukturen und Forschungsumgebungen“ als Diskussionsgrundlage für Sondierungsgespräche mit dem BMBF, einzelnen Landesministerien sowie weiteren Akteuren und Stakeholdern im Bereich Digital Humanities: Das Dossier wird kontinuierlich fortgeschrieben.
- Mitwirkung an der konzeptionellen und organisatorischen Planung, Vorbereitung, Durchführung und Moderation sowie Nachbereitung eines Arbeitstreffens zum Thema „Virtuelle Forschungsumgebungen und ihre Verstetigung – Herausforderungen, Möglichkeiten und Perspektiven für Digital Humanities“, gemeinsam veranstaltet von TextGrid und dem BMBF am 10. Juli 2013 in Berlin unter Beteiligung maßgeblicher Akteure und Stakeholder aus Wissenschaft, Wissenschaftsorganisation und -förderung sowie -politik: Einen zusammenfassenden Bericht enthält das o.a. Dossier „Nachhaltigkeit digitaler Forschungsinfrastrukturen und Forschungsumgebungen“.
- Planung, Vorbereitung, Durchführung sowie Nachbereitung von Sondierungsgesprächen mit den Landesministerien von Sachsen (21. Oktober 2013 und 13. März 2014 in Dresden), Niedersachsen (15. November 2013 in Hannover), Berlin (17. April 2014) und Baden-Württemberg (3. Dezember 2014 in Stuttgart) sowie mit dem BMBF (zuletzt am 19. August 2014 in Bonn), gemeinsam mit Prof. Dr. Heike Neuroth und Prof. Dr. Andrea Rapp sowie Vorbereitung weiterer diesbezüglicher Gespräche: Zusammenfassende Berichte über die Gespräche enthält das o.a. Nachhaltigkeitsdossier.

Zusätzlich dazu fand eine Beratung und Kommentierung der zu Nutzerkommunikation, Monitoring und Qualitätssicherung vorgelegten Konzepte und Präsentationen im Rahmen verschiedener Konsortialtreffen und Telefonkonferenzen statt. Projektmanagement und Öffentlichkeitsarbeit von TextGrid wurden im Zusammenwirken mit der Projektleitung beratend und unterstützend begleitet sowie Moderationsaufgaben bei verschiedenen Veranstaltungen zu TextGrid und DARIAH-DE wahrgenommen.

Fazit

Die gestellten Aufgaben wurden wie geplant bearbeitet. Maßgebliche Weichenstellungen zur Verstetigung von TextGrid und seiner Zusammenführung mit DARIAH-DE wurden vorgenommen. Das betraf insbesondere die weitere Konkretisierung und Diskussion des Angebots an die Länder zur Beteiligung an der institutionellen Grundfinanzierung von TextGrid (als integraler Teil von DARIAH-DE), das in den zuletzt geführten Sondierungsgesprächen auf erfreuliches Interesse stieß und als aussichtsreiches Konzept nachhaltiger (Grund-)Finanzierung die Diskussionsbasis für die weiteren Gespräche mit maßgeblichen Akteuren und Stakeholdern bilden wird. In der kostenneutralen Verlängerung der Förderung von TextGrid bis Ende 2015 wurde die angestrebte Verstetigung von TextGrid und seine Zusammenführung mit DARIAH-DE als breit aufgestellte Virtuelle Forschungsumgebung für die Geisteswissenschaften weiter vorangetrieben. Einen wichtigen Schritt auf diesem Wege stellte die zweite Stakeholder-Konferenz am 19.10.2015 dar.

Insgesamt bestätigt sich der Eindruck eines zunehmenden Interesses seitens der geisteswissenschaftlichen Community an Digital Humanities und damit auch an einer stabilen, breitgefächerten Verfügbarkeit von Virtuellen Forschungsinfrastrukturen und -umgebungen, insbesondere hinsichtlich der Langzeitarchivierung und langfristigen Verfügbarkeit von Forschungsdaten und -resultaten. Dafür spricht nicht nur das sich deutlich erweiternde Stellenangebot im Bereich der Digital Humanities, sondern auch die zunehmende Resonanz auf unterschiedliche Veranstaltungen und Veranstaltungsangebote auf diesem Gebiet – auch über die engeren Fachgemeinschaften hinaus. Dabei steht zu erwarten, dass nicht zuletzt die anstehende Zusammenführung von TextGrid und DARIAH-DE diese Tendenzen befördern und das Bewusstsein für die sich den Geisteswissenschaften damit eröffnenden neuen Möglichkeiten und Perspektiven sowohl in der scientific community selbst als auch bei den Wissenschaftsorganisationen und Mittelgebern weiter wachsen wird. Um bei den beteiligten Stakeholdern und Akteuren auch die Bereitschaft zu nachhaltigem finanziellem Engagement zu erhöhen, bedarf es indes auch nach Ende der BMBF-Förderung kontinuierlicher, intensiver Überzeugungsarbeit.

6.4 AP 3: Nutzer-Kommunikation

6.4.1 AP 3.1 Dokumentation, Nutzer-Information

Das primäre Ziel des Arbeitspakets war die Etablierung von TextGrid im geistes- und kulturwissenschaftlichen Arbeitsalltag durch die Betreuung der Nutzer-Community. Zur Erreichung des Ziels wurden die NutzerInnen einerseits in Schulungsmaßnahmen wie Workshops und Nutzertreffen bei der Arbeit mit der virtuellen Forschungsumgebung betreut und es wurden ihnen andererseits Hilfsmittel wie das Nutzerhandbuch und (Video-)Tutorials zur Verfügung gestellt. Des Weiteren wurde die Einbindung von TextGrid in universitären Lehrveranstaltungen vorangetrieben.

R 3.1.1 Einrichtung von Sandboxes, Entwicklung von Use Cases, Zusammenstellung von Demo-Materialien⁴³

Bei der Planung und Umsetzung neuer Text-Tutorials (mit zahlreichen Screenshots) und Video-Tutorials wurden zunächst die inhaltlichen Schwerpunkte diskutiert. Zu den wichtigsten Werkzeugen und Arbeitsschritten im TextGridLab wurden im Laufe des Jahres 2013 detaillierte Text-Tutorials

⁴³ https://textgrid.de/documents/10644/18815/TextGrid_Report_3_1_1_Final.pdf

ausformuliert und im Zuge der Erstellung der Video-Tutorials noch einmal überarbeitet. Im Herbst 2013 wurden dreizehn Tutorials zum Einstieg in TextGrid, zur Objekt-Verwaltung und zum Text-Bild-Link-Editor in Textform mit zahlreichen Screenshots und als druckbare pdf-Dateien veröffentlicht. Des Weiteren wurden bis Anfang 2014 zwölf Tutorials auch mit der Software Camtasia im Video-Format sowie als englische Text- und Druck-Version auf der TextGrid-Website zur Verfügung gestellt. Darüber hinaus wurden die Videos im Laufe des Jahres 2014 auf dem DHd-Kanal bei YouTube eingestellt.

Darüber hinaus wurde ein Template für TextGrid-Tutorials entwickelt, das gleichzeitig als Anleitung für die Erstellung von Tutorials für Software im Bereich Digital Humanities dient. Dabei wurden die Anforderungen kooperierender Projekte berücksichtigt, so dass das „Meta-Tutorial“ beispielsweise auch für DARIAH-DE nachgenutzt werden kann.

In seiner Masterarbeit „Methoden und Praktiken digitalen Edierens: Die ‚Darmstädter Musteredition‘ des ‚Meßtraktats‘ aus der Handschrift Trier Stb. 852/1311 4° in der virtuellen Forschungsumgebung TextGrid“ skizzierte Florian Enders die wichtigsten Arbeitsabläufe beim Erstellen einer digitalen Edition. Die Arbeit ist auf die Auszeichnung von Texten in XML und den Funktionsumfang des Text-Bild-Link-Editors fokussiert. Darüber hinaus wurden im Rahmen von Nutzertreffen und Schulungsmaßnahmen Aufgabenblätter mit Workflows entwickelt, die die Arbeitsabläufe aus dem realen Forschungsalltag in TextGrid nachbilden und durch ihren didaktischen Aufbau für weitere Schulungen und Workshops nachgenutzt und weiterentwickelt werden können.

R 3.1.2 Dokumentation für Anwender⁴⁴

Das Handbuch in Form des Wikis wurde im Zusammenhang mit der Integration nachgereicherter Module des TextGridLab überarbeitet. Im Handbuch, das im öffentlichen Bereich des TextGrid-Wikis zur Verfügung steht, wurden anschließend während der gesamten Projektlaufzeit kontinuierlich Korrekturen vorgenommen, die Rückmeldungen der NutzerInnen spielen dabei eine besonders wichtige Rolle.

6.4.2 AP 3.2 eHumanities und TextGrid in universitären Curricula

An der Technischen Universität Darmstadt hat Dr. Oliver Schmid in Zusammenarbeit mit der DARIAH-DE Kollegin Ruth Reiche im Wintersemester 2012/2013 eine gemeinsame Lehrveranstaltung angeboten. Unter dem Namen „Digitale Infrastrukturen in den Geisteswissenschaften“ wurde ein Seminar bzw. eine Übung angeboten, in dessen Rahmen die Studierenden Tools des TextGridLab sowie DARIAH-DE Demonstratoren unter (inter)aktiver Teilnahme des Auditoriums vorführten. Einzelne TeilnehmerInnen erarbeiteten zu Themen im TextGrid-Kontext bzw. DARIAH-DE-Kontext Hausarbeiten, darunter eine sehr gute Arbeit mit dem Thema „Analyse einer Benutzerfreundlichkeitsstudie zur Software TextGrid“, deren Ergebnisse bereits Eingang in die Überlegungen zur Verbesserung der Usability des TextGridLab gefunden haben. Im Sommersemester wurde die Veranstaltung „Text in den Digital Humanities“ von Michael Bender und Prof. Dr. Andrea Rapp angeboten, in der textlinguistische Theorie-Diskussionen mit praktischer Arbeit mit TextGrid verbunden wurden. Praktische Aspekte, die in studentischen Projekten u.a. mit dem TextGridRepository und den Voyant-Tools exploriert werden, standen im Seminar „Textanalyse in den Digital Humanities“ von Michael Bender (WS 2013/2014) im Mittelpunkt, während im Seminar

⁴⁴ https://textgrid.de/documents/10644/18815/TextGrid_Report_3_1_2_Final.pdf

„Digital Humanities und Computerphilologie“ von Prof. Dr. Andrea Rapp digitale Forschungsinfrastrukturen und digitale Methoden in einem gesamtwissenschaftlichen Kontext diskutiert wurden. Die Seminare bedienen nicht nur Module im Master Linguistic and Literary Computing, sondern können auch im Joint BA (Germanistik) und in den verschiedenen Lehramtsstudiengängen Deutsch belegt werden. Auf diese Weise werden die Kenntnisse über und die Arbeitsweisen mit den neuen Technologien in den Fächern verankert. An der TU Darmstadt wird kontinuierlich in jedem Semester mindestens ein Seminar aus dem TextGrid-Kontext angeboten, digitale Forschungsinfrastrukturen sind darüber hinaus in der nach der Re-Akkreditierung des Master LLC (WS 2013/2014) neu eingeführten Pflichtvorlesung „Einführung in die Digital Humanities“ integriert. Auch im neu akkreditierten, im WS 2015/16 startenden BA bzw. Joint BA „Digital Philology“ werden Aspekte digitaler Infrastrukturen nun bereits auf Bachelor-Niveau in der Lehre vermittelt.

Im Wintersemester 2012/13 hat Dr. Oliver Schmid erstmals im Rahmen des Seminars „Digitales Publizieren und Edieren in der Geschichtswissenschaft und den Literaturwissenschaften“ von Prof. Dr. Elmar Mittler Studierenden der Universität Göttingen das Projekt TextGrid vorgestellt und sie in interaktiven Übungen bei ihren ersten Schritten in der Virtuellen Forschungsumgebung begleitet. Der Beitrag erfreute sich einer solch positiven Resonanz sowohl seitens Prof. Mittler als auch der Studierenden, dass am Ende dieser Sitzung eine Fortsetzung im Folgesemester vereinbart wurde. In den folgenden Semestern wurde die Zusammenarbeit in der computergestützten Lehre mit jeweils ein bis zwei Terminen pro Semester fortgesetzt, bis zum Ende der Projektlaufzeit. In den insgesamt acht Veranstaltungen wurden die Schulungsmaßnahmen sukzessive auf das sich konkretisierende Ziel einer digitalen Edition des Werks „Bellifortis“ von Conrad Kyeser zugeschnitten, einhergehend mit einer ausführlichen Beratung zu Fragen des digitalen Edierens in TextGrid und der Publikation der Ergebnisse nach der Fertigstellung der Edition. Das Ergebnis wurde im Rahmen des siebten Nutzertreffens „Archivieren und Publizieren“ im Mai 2015 in Göttingen präsentiert.

Im Rahmen der Übung „Office, Internet und virtuelle Forschungsumgebungen – Arbeitshilfsmittel der digitalen Geschichtswissenschaften“ von Dr. Holger Müller stellte Philipp Vanscheidt im Juni 2014 TextGrid als virtuelle Forschungsumgebung für historische Forschungen vor. Der Schwerpunkt lag hierbei auf der Quellenedition von der Handschriftenbeschreibung bis zur Veröffentlichung. Der Kontakt mit Dr. Müller hatte sich beim dritten TextGrid-Nutzertreffen ergeben und führte neben einem Vortrag bei der interdisziplinären Tagung „Gespräche zur keltologischen Forschung“ auch zu einem Publikationsangebot (siehe Kapitel 11.2 Geplante Veröffentlichungen).

Im Rahmen des Seminars „Digital Edieren in der Neueren Deutschen Literatur“ von Dr. Anne Baillot wird an der Universität Stuttgart bei einem philologischen Projekt TextGrid eingesetzt. Um den Studierenden den Zugang zu dem Programm und seinen Prinzipien zu erleichtern, wurde im Dezember 2014 von Philipp Vanscheidt eine Unterrichtseinheit vor Ort geleitet. Das Projekt wird weiterhin im persönlichen Kontakt mit Lehrenden und Studierenden der Stuttgarter Universität betreut.

Der Einsatz von TextGrid in Seminaren wird im Sommersemester 2015 und darüber hinaus fortgesetzt. Im Kontext des Forschungsprojekts „Kulturen der Intimität“⁴⁵ werden an der TU Darmstadt (A. Rapp), der Universität Koblenz-Landau (E. Wyss) und der Universität Münster (A. Dammel) komplementäre Seminare zur Briefkorpus-Erschließung, zur linguistischen Analyse sowie zur historischen Pragmatik durchgeführt. Die dezentrale kollaborative Bearbeitung der Briefe erfolgt über das TextGridLab, mit dem die Studierenden der drei Standorte eine gemeinsame Arbeitsplattform, Werkzeuge und unkomplizierten Zugang zu den Materialien erhalten. Von besonderer Relevanz ist hier u.a. auch die Sicherheit der Daten und der kontrollierte Zugang der Nutzenden, da es sich bei den Forschungsobjekten um zeitgenössische Privatbriefe handelt, bei denen der Schutz der Persönlichkeitsrechte zu beachten ist.

6.4.3 AP 3.3 Workshops, Schulungen, Nutzer-Interaktion

Die sechs Nutzertreffen waren als Milestones fest im Projektplan verankert, zudem wurden in den vergangenen drei Jahren über 20 Schulungsmaßnahmen durchgeführt. Dazu zählen neben einführenden Grundlagenschulungen und Workshops für Fortgeschrittene auch Präsentationen ohne direkten Schulungscharakter, die von Institutionen und Projekten gezielt angefragt worden waren. Darüber hinaus wurde TextGrid auch im Rahmen von Tagungen und Konferenzen aus Nutzersicht präsentiert.

M 3.3.1-3.3.6 Nutzertreffen

Das Konzept, Nutzertreffen als eine Kombination von Vorträgen, Schulungen und Workshops sowie dem TextGrid-Café als Kommunikationsplattform durchzuführen, wurde aus der zweiten Förderphase von TextGrid übernommen und auf Basis der Erfahrungen der folgenden Veranstaltungen kontinuierlich angepasst. Da das erste Nutzertreffen bereits in der zweiten Förderphase veranstaltet worden war, wurde die Nummerierung der Nutzertreffen in der dritten Förderphase fortgeführt. Die Nutzertreffen M 3.3.1-6 werden daher als das zweite bis siebte Nutzertreffen bezeichnet. Alle Nutzertreffen sind im Bereich „Nutzertreffen“ der TextGrid-Website dokumentiert.

M 3.3.1 Nutzertreffen (Edieren mit TextGrid)

Das zweite TextGrid-Nutzertreffen fand am 21. und 22. Juni 2013 an der Akademie der Wissenschaften und der Literatur in Mainz statt. Für das Thema „Erstellung einer digitalen Edition mit TextGrid“ interessierten sich über 40 TeilnehmerInnen, darunter auch Gäste aus Österreich, der Schweiz und Norwegen. Im Rahmen der Veranstaltung wurde sechs Projekten die Gelegenheit geboten, ihre Arbeit mit TextGrid vorzustellen, darunter zwei Masterarbeiten. In Workshops wurde die Verwaltung von TextGrid-Projekten und die Arbeit mit dem XML- und dem Noten-Editor vorgeführt und Probleme im Umgang mit diesen Werkzeugen diskutiert. Zahlreiche Anregungen für die kommenden Nutzertreffen wurden in die Planung der nächsten Veranstaltungen einbezogen.

⁴⁵ <https://liebesbriefarchiv.wordpress.com/>

M 3.3.2 Nutzertreffen (Bilder in TextGrid)

Das dritte TextGrid-Nutzertreffen fand am 20. und 21. Februar 2014 an der Technischen Universität Darmstadt statt. Der Fokus der Veranstaltung lag erstens auf den Werkzeugen, um „Bilder in TextGrid“ zu bearbeiten und zweitens bei der Möglichkeit, die in der VRE entwickelten Daten außerhalb von TextGrid nutzen zu können, beispielsweise durch die Konvertierung der im Text-Bild-Link-Editor entwickelten Verknüpfungsinformationen für interaktive Grafiken auf Webseiten. Dabei wurden dem etwa 50-köpfigen Publikum neben aktuellen Projekten auch Projektideen präsentiert, deren offene Fragen in den anschließenden Workshops mittels praktischer Umsetzung geklärt werden konnten. Durch das Nutzertreffen konnte mit der Amerikanistik, vertreten durch Prof. Dr. Priewe (Universität Stuttgart), eine neue Fachdisziplin für die TextGrid-Community gewonnen werden.

M 3.3.3 Nutzertreffen (Anpassungen und Erweiterungen)

Das vierte TextGrid-Nutzertreffen wurde am 4. und 5. August 2014 in Zusammenarbeit mit dem Projekt SlaVaComp an der Universität Freiburg veranstaltet. Trotz des Termins in den Sommerferien konnte das Thema „Anpassungen und Erweiterungen“ ungefähr 30 Interessierte an die Albert-Ludwigs-Universität locken. Die Veranstaltung war geprägt von den Erfahrungen bei der Anpassung der virtuellen Forschungsumgebung an die eigenen Bedürfnisse, die verschiedene Projekte in ihren Präsentationen mit den Anwesenden teilten. In zwei interaktiven Workshops wurden die Möglichkeiten zur Einbindung von XML-Datenbanken in TextGrid vorgestellt sowie die Erstellung synoptischer Ansichten mit XSLT vorgeführt.

M 3.3.4 Nutzertreffen (Norm- und Metadaten)

Das fünfte TextGrid-Nutzertreffen mit dem thematischen Schwerpunkt „Norm- und Metadaten“ fand mit etwa 40 TeilnehmerInnen am 25. und 26. November 2014 in Zusammenarbeit mit dem Salomon-Ludwig-Steinheim Institut in Essen statt. Am ersten Tag bildeten zahlreiche Vorträge zu Metadaten im Allgemeinen sowie bibliographischen Daten und Personendaten im Speziellen den Schwerpunkt. Am zweiten Tag konnten die TeilnehmerInnen in Workshops zu kontrollierten Vokabularen und Entitäten in TEI selbst Erfahrung bei der Arbeit mit Metadaten in TextGrid sammeln.

M 3.3.5 Nutzertreffen (TextGrid Grand Tour)

Das sechste Nutzertreffen wich in seiner Form von den bisherigen Nutzertreffen ab und gab als „TextGrid Grand Tour“ am 5. März 2015 als Begleitveranstaltung zum DH Summit 2015 im Harnack-Haus der Max-Planck-Gesellschaft in Berlin einen Überblick über die Vielfalt der virtuellen Forschungsumgebung. Zu Beginn wurde in einem Rückblick von Prof. Dr. Werner Wegstein ein Einblick in die Geschichte des Projektes gegeben, es folgten ein Überblick über die verschiedenen Komponenten und Erweiterungen der Software und das digitale Edieren mit TextGrid. Auf fünf „Inseln“ konnten die BesucherInnen je nach Interesse die verschiedenen Möglichkeiten der Forschungsumgebung mit MitarbeiterInnen in halbstündigen Exkursionen ausprobieren und sich dabei von Thema zu Thema, von Insel zu Insel fortbewegen.

M 3.3.6 Nutzertreffen (Archivieren und Publizieren)

Das siebte Nutzertreffen fand als gemeinsame Veranstaltung mit dem Lizenzierungs-Workshop von DARIAH-DE am 12. und 13. Mai 2015 in Zusammenarbeit mit der SUB Göttingen statt. Die beiden Kernthemen „Archivieren und Publizieren“ lockten ungefähr 40 TeilnehmerInnen in die Historische Sternwarte Göttingen. Neben dem TextGrid-Repository als einer der tragenden Säulen von TextGrid standen das Ingest-Tool Kolibri sowie das Publikationswerkzeug SADE im Mittelpunkt.

R 3.3.1 Erfahrungen aus den Nutzerschulungen, Empfehlungen für die weitere Ausgestaltung des Konzepts

Der Report über die „Erfahrungen aus den Nutzerschulungen, Empfehlungen für die weitere Ausgestaltung des Konzepts“ wurde im Mai 2014 veröffentlicht. In Bezug auf die Schulungen und Workshops wurden die Rahmenbedingungen beleuchtet, deren einflussreichster Faktor die Homogenität bzw. Heterogenität der Teilnehmergruppe ist, sowie die technischen Voraussetzungen für die erfolgreiche Durchführung einer Schulungsmaßnahme. Dabei bezieht sich die mögliche Heterogenität des Teilnehmerfeldes hauptsächlich auf die sehr breit gefächerte fachliche Zusammensetzung und die teilweise großen Unterschiede in Bezug auf IT-Kenntnisse und Erfahrungen mit der Arbeit im digitalen Umfeld. Des Weiteren wurden Überlegungen zur Vorbereitung, zum Ablauf und zu den Inhalten von Schulungen und Workshops angestellt. Den zweiten Schwerpunkt des Reports bilden die Analyse der Nutzertreffen und ihre Entwicklung im Verlauf der vergangenen drei Jahre. Zuletzt wurden weitere Formen der Wissensvermittlung, die in TextGrid zum Tragen kommen, beschrieben, wie beispielsweise die Tutorials und die Dokumentation. Der vollständige Bericht ist im Bereich „Materialien > Anträge und Berichte“ auf der TextGrid-Website veröffentlicht. Im Frühjahr 2015 wurde der Report noch einmal überarbeitet, insbesondere wurden die neu gewonnenen Erfahrungen aus dem sechsten Nutzertreffen („Grand Tour“) in den Erfahrungsbericht aufgenommen.

In ihrem Vortrag „TextGrid für Sprachwissenschaftler“ stellten Michael Bender und Philipp Vanscheidt bei der 22. Tagung der Computerlinguistik-Studierenden (TaCoS) am 1. Juni 2012 die am Trierer Kompetenzzentrum für Erschließungs- und Publikationsverfahren in den Geisteswissenschaften, am Institut für Deutsche Sprache in Mannheim sowie am Max-Planck-Institut für Psycholinguistik in Nijmegen entwickelten und in TextGrid integrierten Komponenten vor. Schwerpunkte waren hierbei das Wörterbuchnetz, der Lemmatisierer Morphisto für die deutsche Sprache, das korpusanalytische Werkzeug Cosmas und die Videoannotationssoftware Annex.

In interaktiven Schulungen am 16. November 2012 beim Projekt „eScience – Forschungsnetzwerk Sachsen“ an der Technischen Universität Dresden und am 21. November 2012 am Centre Virtuel de la Connaissance de l'Europe in Sanem (Luxemburg) wurde das TextGridLab interessierten potentiellen NutzerInnen im In- und Ausland präsentiert. Bei einem Vortrag im Rahmen eines Arbeitstreffens an der Max-Weber-Stiftung in Bonn konnte sich am 29. und 30. November 2012 TextGrid im Vergleich mit anderen DH-Projekten gut präsentieren.

Am 18. Januar 2013 haben Philipp Vanscheidt, Celia Krause und Florian Enders auf der Tagung „Digitale Rekonstruktionen mittelalterlicher Bibliotheken“ in Trier einen Beitrag zum Edieren in TextGrid präsentiert. Unterstützt von Stefan Funk (SUB Göttingen) wurde TextGrid am 20. Februar 2013 an der Universitätsbibliothek Mainz mit den Schwerpunkten Infrastruktur und Nutzungsszenarien vorgestellt.

Eine umfassende Einführung in die Arbeit mit TextGrid wurde an der Universität Stuttgart im Rahmen der Workshop-Reihe „Techniken der Digital Humanities in der Diskussion“ des Instituts für Literaturwissenschaft gegeben. Am 26. und 27. April 2013 hat Philipp Vanscheidt die TeilnehmerInnen die Arbeit mit TEI P5 näher gebracht. Am 28. und 29. Juli 2013 hat er dann mit Dr. Oliver Schmid eine zweitägige TextGrid-Schulung durchgeführt, in der fast die komplette Bandbreite der in TextGrid zur Verfügung gestellten Werkzeuge vorgeführt wurde.

Im Rahmen der einwöchigen Edirom-Summer-School des Musikwissenschaftlichen Seminars der Hochschule Detmold fand am 25. und 26. September 2013 an der Universität Paderborn eine TextGrid-Schulung statt. Der Beitrag war eingebettet in Schulungen zur Textkodierung in TEI und musikwissenschaftliche Workshops. Am Ende der Summer-School wurde eine Fortsetzung der Zusammenarbeit für kommende Summer Schools in Paderborn/Detmold vereinbart.

Von 7. bis 12. Oktober 2013 leitete Philipp Vanscheidt gemeinsam mit Dr. Marianna Czisnik ein Seminar am Maritime Museum Greenwich (Großbritannien). In der Veranstaltung mit dem Titel „Letters to Nelson from the Iberian Peninsula (1803-1805)“ wurden Briefe britischer Funktionsträger von der iberischen Halbinsel an Admiral Nelson in XML transkribiert. Das Resultat wird als kritische Ausgabe mit Bildern und Texten online erscheinen. Bei der Erstellung wird TextGrid genutzt, die Studierenden wurden entsprechend in den Gebrauch der Virtuellen Forschungsumgebung eingeführt.

Bei der Tagung „Möglichkeiten der automatischen Manuskriptanalyse“ am 24. und 25. Februar 2014 an der Universität Trier wurde das Projekt eCodicology vorgestellt. In diesem Rahmen kam auch der Masseningest mit dem Importwerkzeug KoLibRi zur Sprache, mit dem das Projekt große Mengen an Digitalisaten mittelalterlicher Handschriften in TextGrid speichert und verwaltet.

Am 30. April 2014 wurde in einer ganztägigen Veranstaltung TextGrid an der Karl-Franzens-Universität Graz vorgestellt. Eingeladen hatte eine mediävistische Forschergruppe um Dr. Karin Kranich und Helmut Klug, die das sogenannte Tegernseer Wirtschaftsbuch ediert und zum Teil schon einer Einführung beim Treffen der AG Germanistische Edition in Bern beigewohnt hatte. An der Veranstaltung in Graz nahmen etwa 20 WissenschaftlerInnen teil, neben einer Einführung in die Virtuelle Forschungsumgebung wurden außerdem verschiedene editorische Arbeitsschritte von der Transkription bis zur Publikation in der Umgebung praktisch erprobt.

In einer zweitägigen Veranstaltung am 5. und 6. Mai 2014 wurde an der Europa-Universität Viadrina in Frankfurt (Oder) das von der Deutschen Forschungsgemeinschaft geförderte Projekt „Multimediales Archiv jüdischer Autorinnen und Autoren in Berlin 1933-1945“ betreut. Das Projekt hat 2014 seine Arbeit aufgenommen und zwei Vertreter hatten bereits dem dritten TextGrid-Nutzertreffen in Darmstadt beigewohnt. Neben einer Einführung in die Umgebung TextGrid wurden mit einer etwa 20-köpfigen Teilnehmergruppe editorische Transkriptionen nach TEI erklärt und geübt sowie insbesondere die Möglichkeiten zur Erstellung von Personen- und Ortsregistern vorgestellt und diskutiert. Das Projekt wurde auch beim fünften TextGrid-Nutzertreffen in Essen mit einem Vortrag vorgestellt.

Bei der interdisziplinären Tagung „Gespräche zur keltologischen Forschung“ am 9. und 10. Mai 2014 an der Universität Stuttgart wurde TextGrid von Dr. Oliver Schmid und Celia Krause unter dem Titel „TextGrid in der Keltologie – Die virtuelle Forschungsumgebung als digitale Fundgrube für Archäologen“ vor ungefähr 15 KeltologInnen vorgestellt. Der inhaltliche Schwerpunkt der

Präsentation wurde dabei auf die bildwissenschaftlichen Aspekte des heterogenen Forschungsbereichs gelegt. Die wichtigsten Erkenntnisse aus dem Vortrag und den Rückmeldungen in der anschließenden Diskussion wurden im Beitrag „TextGrid in der Keltologie. Die virtuelle Forschungsumgebung als digitale Fundgrube für Archäologen und Epigraphiker.“ für den Tagungsband aus der Reihe „Keltologische Kontroversen“ zusammengefasst (siehe Kapitel 11.2 Geplante Veröffentlichungen).

An der Universitätsbibliothek Heidelberg versammelten sich am 10. Juni 2014 etwa 25 Interessierte verschiedener Einrichtungen der Universität zu einer Einführung in die Arbeit mit der Virtuellen Forschungsumgebung. Anwesend waren neben MitarbeiterInnen der Universitätsbibliothek vor allem Mitglieder der Junior Research Group „Digital Humanities and Digital Cultural Heritage“. Der Mehrheit der TeilnehmerInnen war TextGrid schon bekannt, einige zählten bereits zum Kreis der aktiven NutzerInnen. Die Anwesenden nutzten bei der anschließenden Diskussion die Gelegenheit, die Möglichkeiten einer gemeinsamen Nutzung der VRE an der Universität Heidelberg zu eruieren.

Beim DFG-Rundgespräch „Digitalisierung in der Literaturwissenschaft“ am 26. und 27. Juni in Berlin, veranstaltet vom DFG-Fachkolleg 105 „Literaturwissenschaften“ sollten Kriterien und Standards für die Begutachtung Digitaler Editionen diskutiert werden. Andrea Rapp stellt dort Philosophie und Möglichkeiten von TextGrid als Werkzeug für Digitale Editionen vor.

Am Robert Bosch Institut für Geschichte der Medizin in Stuttgart wurde am 15. September 2014 in die Software und in XML gemäß TEI eingeführt. Konkret soll ein Editionsprojekt, das bisher vorrangig auf Microsoft Word basiert, in die für digitale Editionen übliche Auszeichnungssprache XML überführt werden. Entsprechend wurden auch verschiedene Konvertierungsverfahren (Microsoft XML, Open Office, TUSTEP, Konverter, XSLT, beliebige Programmiersprache) diskutiert. Zwei Teilnehmer der Schulung reisten auch zum fünften Nutzertreffen nach Essen an. Das Projekt wurde beim siebten Nutzertreffen in einem Vortrag vorgestellt.

Das Forschungsnetzwerk „DiXit – Digital Scholarly Editions Initial Training Network“ veranstaltete vom 14. bis 19. September 2014 an der Universität Graz eine Konferenz zu „XML/TEI for Digital Scholarly Editions“, bei der „Standards, tools and software“ im Mittelpunkt standen. TextGrid wurde dort in einem 90-minütigen Vortrag als Forschungsumgebung für XML-basierte Editionen vorgestellt und mit rund 70 interessierten TeilnehmerInnen präsentiert und diskutiert.

Am 7. und 8. Oktober 2014 wurde an der Akademie der Wissenschaften und der Literatur in Mainz SchülerInnen im Rahmen der Veranstaltung „Digital Humanities Academy – Digitale Geisteswissenschaften als Berufsperspektive und Karrierechance“, einer Veranstaltung im Rahmen des „Wissenschaftsjahres Digitale Gesellschaft“ ein Einblick in die geisteswissenschaftliche Arbeit im Bereich der Digital Humanities gegeben. An beiden Tagen entschieden sich jeweils ungefähr 30 SchülerInnen für den Beitrag „Hiphop, Märchen, Mittelalter“ des DH-Teams der Technischen Universität Darmstadt, in dem die Arbeit mit TextGrid im Rahmen der Vorstellung von geisteswissenschaftlichen Methoden und Werkzeugen präsentiert wurde.

Für das DFG-Projekt „Von Kanaan nach Gibraltar und zurück – die Phönizier im Mittelmeerraum“ wurde am 22. Mai 2015 eine Schulung an der Technischen Universität Darmstadt durchgeführt. Bereits im Vorfeld bestand ein enger Kontakt zu Dr. Marion Bolder-Boos, die maßgeblich an der Entwicklung des Forschungsnetzwerks beteiligt ist. Durch diese Zusammenarbeit konnte auch bei ersten VertreterInnen der archäologischen Community das Interesse für TextGrid geweckt werden.

Die Nutzerinnen und Nutzer, die per E-Mail oder telefonisch mit den Support-MitarbeiterInnen Kontakt aufgenommen haben, wurden während der Projektlaufzeit kontinuierlich betreut. Die Anfragen, die über die Support-Mailingliste oder über direkten Kontakt an den Support herangetragen wurden, wurden so zeitnah und verständlich wie möglich beantwortet, bei technischen Fragen wurden im Zweifelsfall die EntwicklerInnen konsultiert.

6.5 AP 4: Betrieb und Softwarepflege Repository

Ziel des Arbeitspakets "Betrieb und Softwarepflege Repository" war der Betrieb des fachwissenschaftlichen Langzeitarchivs für geisteswissenschaftliche Forschungsdaten und die Pflege der dazu benötigten Software. Dazu zählen neben TextGridRep auch die dem TextGridLab als Schnittstelle zum Rep dienenden Middleware-Dienste. Dabei war die Hauptaufgabe des Arbeitspakets die Verbesserung des bestehenden TextGridRep (siehe hierzu auch Stand der Technik, S. 10 ff.) sowie dessen Zukunftssicherheit, um langfristig für höhere Nutzerzahlen und einen Produktivbetrieb mit hoher Performanz und Skalierbarkeit gewappnet zu sein. Zudem war es notwendig, für eine Evaluation der umgesetzten Maßnahmen regelmäßige Lasttests durchzuführen. Um den Produktivbetrieb effizienter zu gestalten, mussten diverse technische Prozesse innerhalb von TextGrid optimiert werden. Dazu zählt beispielsweise die Bereitstellung großer Datensammlungen in TextGrid (Ingest), die zu Beginn der Förderphase noch einen hohen manuellen Aufwand erforderten. So hat Arbeitspaket 4 diesen Prozess im Detail dokumentiert und weiter automatisiert sowie die projektspezifische Präsentation dieser Daten umgesetzt.

Die folgenden Abschnitte beschreiben die Ergebnisse des Arbeitspakets ausgerichtet an dessen Task-Struktur.

6.5.1 AP 4.1 Humanities Data Centre

Ziel des Arbeitspakets 4.1 war die Koordination einer engen Zusammenarbeit mit dem Projekt „Humanities Data Centre“ (HDC), das vom Niedersächsischen Ministerium für Wissenschaft und Kultur, Niedersachsen-Vorab, gefördert wird. Ursprünglich wurde mit einem Beginn des Projekts in 2012 oder 2013 gerechnet, am Ende startete das HDC am 01.05.2014. Obwohl damit eine längere, abgestimmte Zusammenarbeit zwischen den beiden Projekten nicht möglich war, wurden doch aufgrund verstärkter Bemühungen und der Tatsache, dass einige TextGrid-Partner in beiden Projekten mitarbeiten, substantielle Ergebnisse erzielt.

Das Projekt Humanities Data Centre erarbeitet in einer Designphase (2014 - 2016) die konzeptionellen Grundlagen für den Aufbau und den Betrieb eines Forschungsdatenzentrums für die Geisteswissenschaften. Dazu gehören (i) die Definition des Dienstleistungsangebots des zukünftigen Forschungsdatenzentrums, (ii) die Konzeption der zur Umsetzung des Angebots erforderlichen Prozesse und ihrer technischen Implementierung sowie (iii) der Entwurf eines stabilen organisatorischen Rahmens. Neben den konzeptionellen Aufgaben werden zudem die Prototypen für den Ingest sowie den Zugriff auf die Forschungsdaten entwickelt. An die Designphase sollen sich eine Aufbauphase und eine initiale Betriebsphase von einem Jahr anschließen.

Aus TextGrid haben insbesondere die Entwicklungen um das TextGrid/DARIAH-DE Repository ihren Weg in das HDC gefunden (siehe hierzu auch AP-übergreifende Ergebnisse, S. 17ff.). Hier flossen bereits während der Entwicklung die Anforderungen aus dem HDC ein, so dass die entstandene

Codebasis funktional geeignet ist, um als Prototyp für die Speicherung von Forschungsdaten zu fungieren.

6.5.2 AP 4.2 Performanz, Skalierbarkeit, Lasttests

Mit den Ergebnissen der zweiten Projektphase lag für das TextGrid Repository eine produktive und leistungsfähige Infrastruktur vor. Allerdings ist es bei einer solchen essentiell, die existierende Architektur und Umsetzung in regelmäßigen Abständen mit den Kundenwünschen, Nutzerzahlen, Dienstgütedaten und aktuellen technischen Entwicklungen abzugleichen und zu bewerten. Nur so kann gewährleistet werden, dass die Software den aktuellen Anforderungen genügt, weiterentwickelt werden kann und wartbar bleibt. Bei TextGrid waren es insbesondere die steigenden Nutzerzahlen und der Wunsch nach einem möglichst ausfallsicheren Betrieb der Infrastruktur, die eine Neubewertung notwendig machten. Um diese zu realisieren, wurden zunächst für die Kernkomponenten des TextGridRep jeweils die folgenden drei Schritte durchgeführt:

1. Identifikation von Lücken (die so genannte *gap analysis*)
2. Bestimmung von Maßnahmen zur Schließung der Lücken (Optimierung)
3. Bewertung der Machbarkeit der einzelnen Maßnahmen

Die entsprechenden Resultate sind ausführlich im Projektbericht R 4.5.1 *Architekturskizze inklusive Identifikation der Lücken und der Optimierungsmöglichkeiten*⁴⁶ dokumentiert.

Basierend auf der Machbarkeitsbewertung wurde die Umsetzung einer Reihe von Maßnahmen beschlossen und eine entsprechende Roadmap erarbeitet. Hierbei wurde insbesondere Wert auf die verwendeten Technologien gelegt und abgekündigte oder nicht mehr adäquat unterstützte Produkte ersetzt. Flankiert wurden diese Maßnahmen durch kontinuierliche Tests der Softwarekomponenten und der TextGrid Middleware. Nur so konnten die EntwicklerInnen belastbares Feedback erhalten und die Entwicklungen auf die Verbesserung der Performanz hin prüfen.

In einem nächsten Schritt wurde basierend auf der aktualisierten Infrastruktur ein Hochverfügbarkeitskonzept für TextGrid erarbeitet. Dazu wurden in einem ersten Schritt alle Komponenten von TextGrid sowie die verwendete Software auf ihre Cluster-Fähigkeit geprüft. Dabei galt es zu ermitteln, ob die einzelnen Teile der TextGrid Middleware parallel auf mehrere Servern so laufen können, dass ein Server im Fehlerfall die Aufgaben des anderen übernehmen kann. Das Hochverfügbarkeitskonzept wurde schrittweise umgesetzt und in Betrieb genommen.

Einsatz von Elasticsearch zur Suche in TextGrid

TG-search ist ein Kerndienst von TextGrid, welcher die Suche über alle Daten erlaubt (siehe hierzu auch Kapitel 4.2 Technischer Stand). So ist die Volltextsuche über alle TEI-Daten ebenso möglich wie die Suche nach Metadaten. Vor Beginn der dritten Förderphase griff TG-search dazu auf eine eXist-db Datenbank⁴⁷ zu. Diese Lösung verursachte häufiger Probleme, beispielsweise bei der Re-Indizierung oder Restaurierung des Datenbestandes; zudem wurde die Performanz der Suche als nicht ausreichend angesehen. Eine Auswertung verfügbarer Alternativen im Zuge der oben genannten *gap*

⁴⁶ https://textgrid.de/fileadmin/user_upload/TextGrid_R451_Architekturskizze.pdf

⁴⁷ <http://exist-db.org/exist/apps/homepage/index.html>

analysis führte zu der Entscheidung, das Open Source-Produkt ElasticSearch⁴⁸ zu verwenden. Bei ElasticSearch handelt es sich um eine verteilte Such- und Analyseplattform für große Datenbestände, die die Anforderungen von TextGrid an eine Suchtechnologie erfüllt hat. Es ermöglicht komplexe Suchanfragen mit Facettierung ebenso wie die Volltextsuche in einer Vielzahl von Dateiformaten. ElasticSearch kann flexibel skaliert werden, so dass sich auch TG-search in das Hochverfügbarkeitskonzept von TextGrid integrieren ließ.⁴⁹ Durch den Einsatz von Technologien wie ElasticSearch, die einen großen Entwicklerstamm haben und im Enterprise-Umfeld eingesetzt werden, wird nicht nur die Performanz des Gesamtsystems verbessert, sondern auch die Zukunftssicherheit von TextGrid erhöht.

Umsetzung des Hochverfügbarkeitskonzepts⁵⁰

Das in TextGrid erarbeitete und umgesetzte Hochverfügbarkeitskonzept basiert auf einer Architektur, die die vorab genannten Änderungen am TextGrid Repository berücksichtigt und die auf den ausfallsicheren Betrieb der Forschungsumgebung ausgerichtet ist.

Das Grundkonzept ist in Abbildung 2 skizziert, wobei generell zwei Aspekte getrennt betrachtet werden: Daten und Dienste. Die unpublizierten und publizierten Daten werden von den TextGrid-NutzerInnen erzeugt, bearbeitet und schließlich veröffentlicht. Für diese Aufgaben werden die Nutzerdaten dauerhaft in einer Speicherinfrastruktur abgelegt. Um eine hohe Verfügbarkeit zu gewährleisten und die Daten vor Verlust zu schützen, ist es notwendig, dass beim Wegfall von Teilen der unterliegenden Speicherinfrastruktur weiterhin der Datenzugriff möglich ist. Dies wird durch die Replikation der Daten mittels Policies erreicht. So werden für gewöhnlich alle Daten an zwei geographisch unterschiedlichen Standorten gespeichert, um ihre Verfügbarkeit zum Beispiel auch im Falle eines Stromausfalles an einem Standort zu gewährleisten.

⁴⁸ <http://www.elasticsearch.org/>

⁴⁹ vgl. P. Wieder: *Hochverfügbarkeit und Performanz: Beiträge zum langfristigen Betrieb von TextGrid*. In: *TextGrid: Von der Community – für die Community*, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 183-190.

⁵⁰ vgl. hierzu auch: P. Wieder: *Hochverfügbarkeit und Performanz: Beiträge zum langfristigen Betrieb von TextGrid*. In: *TextGrid: Von der Community – für die Community*, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 183-190.

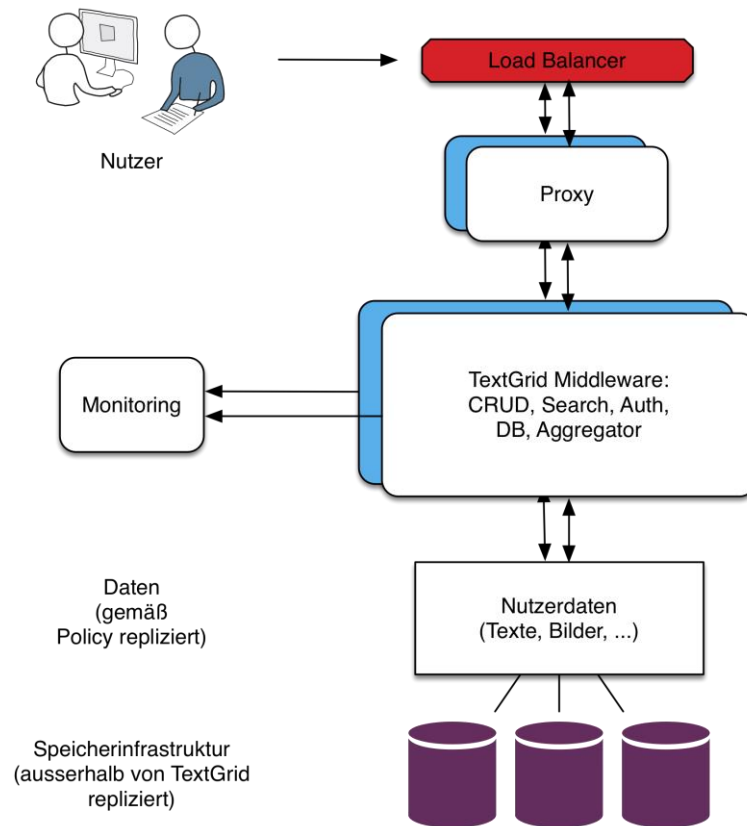


Abbildung 2: Grundkonzept der Hochverfügbarkeit

Auf Dienstebene wird eine höhere Verfügbarkeit durch den Betrieb von mehreren Instanzen der TextGrid-Middleware erreicht. Dazu werden über einen Load-Balancer Anfragen der NutzerInnen an verschiedene Proxies verteilt, die diese dann zur Verarbeitung an die jeweilige Instanz der TextGrid-Middleware weiterreichen. Auf diese Weise kann das System den Ausfall einer Instanz kompensieren, indem der Load-Balancer in einem solchen Fall Anfragen nur an die funktionstüchtige Middleware weitergibt. Ergänzt wird die TextGrid-Infrastruktur durch ein Monitoring-System, das in die Überwachung bei den jeweiligen Rechenzentren integriert ist und es erlaubt, auf den Ausfall einzelner Instanzen zeitnah zu reagieren.

Zusätzlich zu den Infrastrukturmaßnahmen erforderte der parallele Betrieb mehrerer Instanzen der TextGrid-Middleware auch Softwareentwicklungen. So musste auf verschiedenen Ebenen der TextGrid-Middleware, wie beispielsweise TG-crud oder TG-auth, sichergestellt werden, dass die notwendigen Daten synchron gehalten werden und eine Instanz den Ausfall einer anderen kompensieren kann. Hierzu waren Änderungen an den entsprechenden Komponenten notwendig.

6.5.3 AP 4.3 Qualitätsmanagement (SUB)

Ziel dieses Arbeitspakets war die Erstellung eines Handbuchs, das die verschiedenen Aspekte des Qualitätsmanagements dokumentiert und als Anleitung für einen Regelbetrieb der TextGrid-Middleware dient. Das Qualitätsmanagement kann unterteilt werden in Administration, Technische Dokumentation, Bugfixing und Monitoring.

Die Pflege aller Komponenten wurde dahingehend vereinfacht, dass die Installation und Konfiguration des Servers automatisch über Puppet abgewickelt wird. Dadurch ist gewährleistet, dass

sich die Server gleichen und unter gleichen Bedingungen getestet werden können. Die Konfiguration per Puppet läuft derzeit auf den Testservern.

Für den technischen Anteil des TextGrid-Rep wird eine technische Dokumentation mit einer allgemeinen Beschreibung aller Services und darunter liegenden Technologien angefertigt. Des Weiteren werden die Application Programming Interfaces (API) dokumentiert. Diese Dokumentation wird weiterhin in dem durch das Projekt DARIAH-DE gehostete Wiki zur Verfügung gestellt.

Für die Pflege der Software wurde Umzüge des Systems zur Quellcode-Versionskontrolle durchgeführt. In erster Linie betrifft dies einen Umzug aller Quellcodes von SVN nach git, das in das Chili-System der GWDG integriert wurde. Des Weiteren wurde auch der frühere Inhalt des Jira-Systems zu Chili umgezogen. Somit sind auch die offenen Arbeitsschritte und Fehler-Behebungen (Bugfixing) integriert und an einer Stelle zusammen mit den Quellcodes verfügbar.

Das aufgebaute umfangreiche Monitoring mittels Icinga/Nagios wurde in das DARIAH-DE Monitoring überführt. Über diese Anwendungen werden die Administratoren über Probleme und Ausfälle von Services und Servern informiert, um entsprechend zu handeln.

6.5.4 AP 4.4 Ingest Datensammlung, Publikation, Präsentation

Die SUB leistete inhaltlichen sowie auch technischen Support für den Import von neuen Daten in das TextGrid-Rep, die Dokumentation wurde erweitert und Fehler korrigiert. Der Dienst TG-publish wurde erweitert sowie verbessert, und im Zuge der Zusammenlegung des Basicodes mit dem des DARIAH-DE Repositoriums (DH-publish) generalisiert, so kann nun z.B. das PID-Modul von beiden Services genutzt werden. Weiterhin wurden Anpassungen für den Ingest von Wörterbüchern implementiert.

Für die Präsentation der Daten wurde das SADE Framework in das Lab integriert. Durch dieses Framework können Bild- und TEI-Dokumente sowie Metadatenobjekte auch ohne das TextGrid Repository in projekteigenen Portalen veröffentlicht werden (in diesem Fall werden allerdings keine Mechanismen der Langezeitarchivierung, wie zum Beispiel die Vergabe eines PID, zur Verfügung gestellt). Die Veröffentlichung im Rep und über SADE ist ebenso möglich.

6.5.5 AP 4.5 Technischer Regelbetrieb, Architektur, innovatives Evolutionspotential

Der Betrieb des TextGrid Repository wurde, zusätzlich zu den Verbesserungen bezüglich Performanz und Skalierung, vor allem durch die Umstrukturierung der Server-Architektur optimiert. Während vor der dritten Förderphase die verschiedenen Dienste auf einer stark verteilten Server-Struktur liefen, wurden im Zuge der Konsolidierung und Umsetzung der Hochverfügbarkeit alle Dienste auf einen leistungsfähigen Server im VMware ESX-Bereich der GWDG umgezogen. Zusätzlich wurde das System geklont, so dass ein redundantes System für den hochverfügbaren Betrieb zur Verfügung steht. Durch die Integration des Systems in die Überwachung und Prozesse des Rechenzentrums sind nun für den Betrieb von TextGrid auch die notwendigen operativen Voraussetzungen für hohe Verfügbarkeit und verlässlichen Regelbetrieb gegeben.

Die Änderung des Produktivsystems bedingte auch eine entsprechende Umstrukturierung des Testsystems, um gezielt die Funktionen der Hochverfügbarkeit testen zu können. Dazu wurden Test-Server in der GWDG Compute Cloud bereitgestellt, die von allen TextGrid-EntwicklerInnen genutzt

werden können. Für die speziellen Bedürfnisse der Hochverfügbarkeit des RDF-Triplestores wurde ein dritter Test-Server benötigt, der ebenfalls in der GWGD Compute Cloud angelegt wurde.

Zusätzlich zu den betrieblichen Aspekten sowie den genannten Verbesserungen im Bereich der Middleware war es TextGrid ein Anliegen, das Potential des Repository innovativ über den Einsatzbereich in TextGrid zu erweitern und damit neue NutzerInnen zu gewinnen. Hier wurde in Zusammenarbeit mit DARIAH-DE eine gemeinsame Codebasis für ein erweiterbares DARIAH-DE Repositorium⁵¹ geschaffen. Dies bündelt Ressourcen und schafft neue Möglichkeiten, die bereits getätigten Investitionen nachzunutzen.

6.6 AP 5: Betrieb und Softwarepflege Laboratory

6.6.1 AP 5.1 Sicherheit, Bugfixing, Technische Dokumentation

Nachdem sich der Entwickler Michael Leuk, der in der zweiten Förderphase des Projekts für die Betreuung einzelner Komponenten des TextGridLab verantwortlich gewesen war, in den Quellcode und die Architektur des Gesamt-Labs eingearbeitet hatte, stand das Bugfixing in den verschiedenen Komponenten des Labs im Fokus seiner Arbeit. In diesem Bereich gab es mehrere Schwerpunkte, so stand auch in der dritten Förderphase die Verbesserung der Nutzerfreundlichkeit im Vordergrund. Dabei hat sich bei der Suche nach Bugs im TextGridLab auch weiterhin die wichtige Rolle des Feedbacks aus der Nutzercommunity gezeigt. So wurden beispielsweise Fehler in der Menüführung behoben, die im Rahmen eines Workshops entdeckt wurden. Auch im Bereich der Objektverwaltung wurden mehrere Bugs behoben, die beispielsweise beim Kopieren von Objekten und bei der Bearbeitung von Aggregationen auftraten. Teilweise konnten Probleme auch zunächst nur für bestimmte Betriebssysteme gelöst werden, beispielsweise waren nicht funktionierende Schaltflächen auf der Willkommen-Seite zunächst nur für Windows und MacOS repariert worden, während das Problem für Linux-Rechner erst zu einem späteren Zeitpunkt gelöst werden konnte. Während zahlreiche kleinere Bugreports auch kurzfristig bearbeitet werden konnten, mussten einige Probleme, für die eine sehr komplexe Lösung erforderlich war, über einen längeren Zeitraum hinweg bearbeitet werden, wie beispielsweise bei der Anpassung der Fenstergröße des Metadaten-Editors.

Des Weiteren stand beim Bugfixing die Stabilität der Werkzeuge und Dienste im Mittelpunkt. Insbesondere wenn die Funktionalität des TextGridLab vorübergehend eingeschränkt war, wurde die Behebung der dafür verantwortlichen Fehler mit höchster Priorität vorangetrieben. So musste nach einem serverseitigen Problem der Dienst, der für die Kommunikation des TextGridLab mit dem Trierer Wörterbuchnetz verantwortlich ist, komplett neu aufgesetzt und angepasst werden, um die fehlerfreie und zuverlässige Funktion der Wörterbuchsuche wieder sicherstellen zu können.

Eine weitere zeitintensive Aufgabe war die Umstellung des TextGridLab auf die Eclipse-Version 4.3 („Kepler“), viel Zeit beanspruchte dabei die Fehlersuche und die Behebung der Probleme einzelner Werkzeuge, insbesondere des Navigators. Da die Umstellung über viele Monate hinweg nicht vollständig fehlerfrei durchgeführt werden konnte, wurde als neues Ziel die Anpassung des TextGridLab auf die aktuelle Eclipse-Version 4.4 („Luna“) für die TextGridLab-Version 3.0 in Angriff genommen, die im Frühjahr 2015 umgesetzt werden konnte.

⁵¹<https://dev2.dariah.eu/wiki/download/attachments/14651583/M%204.3.2.1-DARIAH-Repositorium-Prototyp-final.pdf?version=1&modificationDate=1430220062670&api=v2>

M 5.1.1 Technische Dokumentation

Nach der Bündelung und Reorganisation der relevanten Seiten und Dokumente im Wiki erfolgte zunächst die Strukturierung der technischen Dokumentation. Die neue Struktur erleichtert den Überblick über die aktuell vorhandene technische Dokumentation bezüglich Architektur, Frontend, Backend, Middleware, Repositorium, Development, High Availability, Testing, Monitoring und Administration. Aufgrund einer stärkeren Verschränkung von Ressourcen von TextGrid und DARIAH-DE wurde eine umfangreiche Überarbeitung der Dokumentation durchgeführt. Nach der abschließenden Zuordnung verbleibender Dokumentationsaufgaben konnte die technische Dokumentation im Frühjahr 2015 fertiggestellt werden.

6.6.2 AP 5.2 Funktionale Tests

Automatisierte Tests des Gesamtsystems aufgrund typischer Arbeitsschritte von geisteswissenschaftlichen Anwendern erlauben es, bei Programmveränderungen und Anpassungen auftauchende Probleme schnell zu erkennen und zu lokalisieren. Schon in TextGrid II hat die Hochschule Worms eine umfangreiche Testumgebung für diese Art der automatisierten funktionalen Tests aufgebaut. In einem Labor der Hochschule sind auf allen Arbeitsplätzen zusätzliche Festplattenpartitionen mit der Testumgebung Sikuli eingerichtet, sodass Tests mit bis zu 21 parallel arbeitenden Rechnern durchgeführt werden können. Ein kompletter Durchlauf der aktuell existierenden funktionalen Tests dauert etwa 70 Minuten. Zu Projektbeginn deckten die funktionalen Tests allerdings erst etwa 10% des Funktionsumfangs des TextGrid Laboratory ab, so dass im Rahmen des Projekts der aktuellen Projektphase die Testabdeckung systematisch ausgebaut wurde. Gleichzeitig wurden die Tests für zwei Betriebssysteme (Windows und Mac OS X) umgesetzt und während der Laufzeit an diverse Betriebssystem-Updates angepasst.

Es wurden somit weitere funktionale Testpläne, basierend auf den Anforderungen der NutzerInnen, erstellt. Durch Use Case Beschreibungen wurden die Testpläne an die jeweiligen Anforderungen angepasst und teilweise auch mit Lasttests verzahnt. Besonders kritische Funktionalitäten wurden dabei besonders intensiv getestet. Gleichzeitig hat die Hochschule Worms im Produktivbetrieb anfallenden Metriken zur Programmfunktionalität analysiert, um wesentliche Fehlerszenarien zu erkennen.

6.6.3 AP 5.3 Usability

M/R 5.3.1 Usability, Zusammenfassung der bisherigen Nutzerszenarien und Identifikation der Testgruppen⁵²

Dieser Bericht befasste sich zunächst mit einer praxistauglichen Definition von Usability und wie Usability im Rahmen der dritten Förderphase systematisch untersucht werden kann. Zusätzlich wurde ein Überblick über die Usability-Studien aus den vorherigen Förderphasen sowie über die Studien begleitender Projekte gegeben.

Für das weitere Vorgehen wurde als Arbeitsdefinition für Usability in TextGrid formuliert, dass untersucht werden sollte, wie das System in der wissenschaftlichen Praxis eingesetzt wird. Auf Grundlage dessen wurden im zweiten Abschnitt des Berichts Vorschläge für zusätzliche denkbare

⁵² Vgl. https://textgrid.de/fileadmin/TextGrid/reports/TextGrid_R531_Usability.pdf

Nutzungsszenarien gemacht, insbesondere vor dem Hintergrund, die Zahl der aktiven NutzerInnen zu erhöhen und zusätzliche Wissenschaftsdisziplinen abzudecken.

Der dritte Abschnitt gibt einen Überblick über Verfahren, wie die Anwenderfreundlichkeit vom System ermittelt werden kann, auch wenn keine dezidierten Geräte zur Verfügung stehen.

R 5.3.2 Usability: Anforderungen an TextGrid aus der Arbeit mit den Testgruppen⁵³

Im zweiten Bericht dieses Unter-Arbeitspaketes wurden die Anforderungen an TextGrid aus Sicht ausgewählter Nutzergruppen weiter spezifiziert. Hierzu wurden relevante Testgruppen – insbesondere aus der textbasierten historischen Forschung – identifiziert und deren Arbeitsabläufe analysiert, um zu sehen, wie diese auf die vorhandenen Funktionalitäten von TextGrid abgebildet werden können.

Näher untersucht wurden hier einerseits Projekte, die die Nutzung von TextGrid fest in der Projektplanung verankert haben, andererseits zwei Fallbeispiele aus dem wissenschaftshistorischen Kontext, in denen der mögliche Einsatz von TextGrid von der technisch-inhaltlichen Projektplanung über die textuelle Bearbeitung der Quellen bis zur Publikation durchgespielt wurde sowie Optimierungsmöglichkeiten des Systems dokumentiert wurden.

Ein weiterer Aspekt zur Ermittlung der Usability lag auf der Untersuchung der Zufriedenheit der NutzerInnen von TextGrid. In enger Zusammenarbeit mit anderen Arbeitspaketen wurden weitere Bedarfe evaluiert.

Der Bericht dokumentiert die Auswahl der Testgruppen, die Überlegungen zu typischen Anwendungsszenarien, die Anforderungen aus den untersuchten Modellprojekten und die sich daraus ergebenden Empfehlungen für die Weiterentwicklung von TextGrid.

M/R 5.3.3 Usability: Abschluss des Fallbeispiels und sich daraus ergebende Anforderungsanalyse⁵⁴

Der letzte Bericht in dem Arbeitspaket fasst die Diskussionen des 2. Entwicklerworkshops zusammen, der im September 2014 von AP 2.2 in Zusammenarbeit mit AP 5 und Cluster 1 von DARIAH-DE organisiert und durchgeführt wurde. Das Thema Usability wurde hier noch einmal breiter aufgefasst, als es in diesem Arbeitspaket ursprünglich vorgesehen war, durch die weitere Beschäftigung im Rahmen von DARIAH-DE bot es sich aber an, das Thema noch von einem höheren Standpunkt aus zu beleuchten.

Der Bericht gibt die Struktur des Workshops wieder, indem er die Impulsreferate zu den verschiedenen Themen („Usability in Forschungsinfrastrukturen der digitalen Geisteswissenschaften“, „Praktische Beispiele und strategische Überlegungen: Projekte und Toolentwicklungen“, „Methoden und Verfahren“) zusammenfasst und die wichtigsten Punkte der anschließenden Diskussionen protokolliert. Die WorkshopteilnehmerInnen stimmen darin überein, dass umfassende und grundlegende Forschungen zum Thema Usability in Forschungsinfrastrukturen bzw. generell im Digital Humanities-Bereich kaum durchgeführt werden und Usability-Ideen und -

⁵³ Vgl. https://textgrid.de/fileadmin/TextGrid/reports/TextGrid_Report_532-Final.pdf

⁵⁴ https://textgrid.de/fileadmin/TextGrid/reports/TextGrid_Report5_3_3_final.pdf

Strategien zu wenig Beachtung finden.⁵⁵ Andererseits wird allgemein beklagt, dass zu wenig über die NutzerInnen und das Nutzungsverhalten von Digital Humanities-Werkzeugen und -Infrastrukturen bekannt sei, obwohl Forschungsförderung, Projektmanagement und Fragen der Verbreitung, Nachhaltigkeit sowie Nachnutzbarkeit auf den Erwartungen und dem Verhalten der NutzerInnen aufbauen (sollen).

Den Abschluss des Berichts bildet ein Katalog von Empfehlungen, wie die Erforschung von Usability und die Umsetzung ihrer Konzepte, im Rahmen der digitalen Geisteswissenschaften und damit assoziierten Infrastrukturprojekten.

6.6.4 AP 5.4 Koordination von Weiterentwicklungen und Satellitenprojekten (UWÜ)

TextGridLab-Bau

In der vorherigen Projektphase ist das TextGridLab von einer eher monolithischen Software zu einem modularen System aus Komponenten geworden, von denen ein Teil (core) die Grund-API, zusammen mit einer Reihe von Komponenten (XML-Editor, Text-Bild-Linkeditor, Hilfe, Basiskomponenten – praktisch das „Hauptprogramm“) der zum Download und Update zur Verfügung gestellten Applikation bildet, in die weitere, z.B. fach- oder projektspezifische Tools etwa über den *TextGridLab Marketplace* nachinstalliert werden können.

Im Rahmen von AP5.4 wurden die Releases des TextGridLab geplant, koordiniert und durchgeführt – insgesamt wurden im Rahmen der Laufzeit acht neue Versionen des Gesamtlabs an die EndnutzerInnen in allen Projekten verteilt, wobei i.A. automatische Updates oder Download einer neuen Version zur Wahl standen. Daneben steht jederzeit die *Nightly*-Version mit dem aktuellen, ungetesteten Entwicklungsstand zur Verfügung.

Während der Projektlaufzeit wurde dabei anhand der Erfahrungen das Releasemodell umgestellt: Während zu Anfang mehr oder weniger reine Bugfix-Releases zur stabilen Version 2.0 veröffentlicht wurden, wurde 2014 aufgrund der relativ stabilen Hauptentwicklungslinie auf Releases, die von der Hauptentwicklungslinie abzweigen und in einer Release-Phase auf Stabilität getestet wurden, umgestellt. Dieses Prozedere hat den Vorteil, dass neben klassischen Bugfixes sowohl kleinere Änderungen (Weiterentwicklungen der Übersetzungen etwa) als auch neue Features in die releaste Version übernommen werden und dass durch die größere Nähe zwischen Release und Entwicklungsversion Entwicklung, Release und Bugfixing einfacher werden.

2015 wurde, nach längerer Entwicklung in Nebenzweigen, dann die aktuelle Weiterentwicklung des Lab mit einer auf die aktuellen 4.x-Versionen von Eclipse portierten Version zusammengeführt und das Ergebnis als TextGridLab 3.0 releast. Damit basiert das TextGridLab zum Projektende auf einer aktuellen Eclipse-Version.

Um das TextGridLab für Entwicklungen in Satellitenprojekten weiter zu öffnen und auch die interne Entwicklung weiter zu vereinfachen, wurde eine *Reorganisation der TextGridLab-Buildpakete*

⁵⁵ vgl. hierzu: M. Bender: *Forschungsumgebungen in den Digital Humanities: Nutzerbedarf, Wissenstransfer, Textualität*. Reihe: Sprache und Wissen (SuW) 22. Berlin, Boston: de Gruyter. (erscheint voraussichtlich im März/April 2016)

durchgeführt. Unter anderem wurde die Struktur zu einer einheitlicheren Unterstützung von Maven, Eclipse und m2eclipse verändert, Source Packages gebaut und publiziert und eine Entwicklerdokumentation geschrieben. Damit ist es nun intern wie für externe Projekte möglich, Plug-ins für das TextGridLab zu entwickeln, ohne den Lab-Sourcecode auszuchecken; wer existierende Tools weiterentwickelt, muss zudem nur diejenigen Komponenten auschecken, die verändert werden sollen.

- Die Verwaltung des TextGrid-Sourcecode wurde von einem durch einen Partner selbst gehosteten Subversion zu dem von der GWDG angebotenen Git migriert. Git bietet als modernes verteiltes Versionskontrollsystem deutlich bessere Möglichkeiten zur Kooperation mit lose verbundenen EntwicklerInnen, eine erhöhte Datensicherheit sowie eine breitere Unterstützung für experimentelle Zweige – aufgrund einer anderen Repositorystruktur waren jedoch eine Reihe von Umbauten insbesondere im Buildverfahren für das TextGridLab notwendig. Im AP wurde das Migrationsverfahren konzipiert und für eine Reihe von Komponenten durchgeführt, zudem wurden Build- und Releaseverfahren für das Lab angepasst. Recht aufwändig war dabei auch die Rekonstruktion von Code und Versionsgeschichte aus dem bei einem Systemfehler beschädigten Subversion-Repository.
- Für die Webdarstellung auf textgridrep.de wurde auf Anforderung von Drittprojekten ein Konzept entwickelt und implementiert, mit dem bestimmte projektspezifische Darstellungsformen (aus dokumentarischer Beschreibung in den TEI-Dokumenten) auch in die HTML-Darstellung übernommen werden können. Dabei war entsprechend ein Kompromiss zwischen der Übernahme projektspezifischer Renderings einerseits und einer relativ einheitlichen Darstellung auf textgridrep.de andererseits zu finden. Zur Verwendung kommen die in der TEI standardisierten Methoden zur formalen Beschreibung der Darstellung in Originaldokumenten (rendition, tagUsage, style), soweit als formale Sprache hierfür CSS verwendet wird.
- Zur besseren Unterstützung des Publikationsprozesses wurde eine TextGridLab-Komponente implementiert, die die spätere Darstellung auf textgridrep.de auch für noch nicht veröffentlichte Dokumente als Vorschau simuliert.

Ebenfalls in einem Satellitenprojekt intensiv benutzt wird die Preview-Seite des XML-Editors. Diese zeigt das HTML-Rendering des aktuell bearbeiteten XML-Dokuments anhand einer Transformation direkt im Lab auf Wunsch sogar noch vor dem Speichern. Das verwendete Stylesheet kann pro Dokument definiert werden, mit einer konfigurierbaren Standardtransformation (zunächst wurde die Grundeinstellung der beliebten TEI-Stylesheets von Sebastian Rahtz verwendet, mittlerweile die Simulation der Webdarstellung auf textgridrep.de).

Aggregator

Eine im Rahmen von AP1.3 und AP5.4 entwickelte Komponente ist der *Aggregator (Arbeitstitel)*. Dieser Service dient dazu, die TextGrid-spezifische *Aggregationsstruktur* zur Repräsentation aller Arten von zusammengesetzten Daten abzulaufen und in Formen zu exportieren, die für Dritte leichter nutzbar sind. Das beinhaltet u.a.:

- *TEI-corpus-Export*: Diese Zusammenfassung von TEI-codierten Einzelobjekten im standardisierten Corpusformat TEI-corpus ermöglicht später eine gemeinsame Weiterverarbeitung der aggregierten Objekte z.B. für eine Konvertierung oder Analyse.
- Zum Offline-Lesen größerer Texte auf E-Book-Readern, Tablets etc. wurde ein automatischer Export als EPUB realisiert.
- Ein ZIP-Export incl. Roundtripping zum Re-Import evtl. überarbeiteter Versionen über das Lab wurde ergänzt.

Zudem wurde die Software so überarbeitet, dass sich künftig weitere Exportformate oder zusätzliche Features für die vorhandenen leichter andocken lassen. Daneben wurde eine Unterstützung für Caches und Proxies integriert.

Die Komponente ist mittlerweile im Produktivbetrieb auf <http://textgridrep.de/> im Einsatz und übernimmt dort neben diversen Exportmöglichkeiten insbesondere auch die Webdarstellung der TEI-Dokumente im Repository.

Sonstiges

Im Rahmen eines gemeinsamen Programmiersprints wurde der im Exzellenzinitiativen-Projekt *Brain / Concept / Writing* an der RWTH Aachen entwickelte Multitouch-Tisch an TextGrid angebunden.⁵⁶ Dazu wurde u.a. ein experimenteller Basis-PDF-Export-Service für TextGrid-Objekte entwickelt, da PDFs auf dem Tisch unmittelbar dargestellt werden können.

Das technische Monitoring wurde weiterentwickelt, um Auswirkungen und Prioritäten eines Problems besser zu berücksichtigen. Das Monitoring wurde zudem von der eigenen, selbstgehosteten Lösung aus der vorherigen Projektphase zu DARIAH-DE migriert. Dabei wurde ein Konzept für die Pflege von projektspezifischem Monitoring in DARIAH-DE mitentwickelt.

In Kooperation mit dem DFG-geförderten Projekt „Hybrid-Edition von Goethes Faust“ wurde ein Konzept für die Veröffentlichung der XML-Daten via TextGrid incl. Anbindung der Faust-spezifischen Webanwendung entwickelt und die Aufbereitung der Daten und ihrer Metadaten für die Veröffentlichung vorbereitet.

Im Kontext des Zentrums für digitale Edition und quantitative Analyse „Kallimachos“ wurde eine Python-basierte Anwendung zur programmatischen Extraktion von Metadaten aus der TextGrid-Suche entwickelt, die vorhandene Metadaten mit in TextGrid vorhandenen Daten abgleicht und passende TextGrid-Adressen zu den entsprechenden Dokumenten ermittelt.

Eine Weiterentwicklung des Aggregators erlaubt den Export auch großer Datenmengen, etwa aller Ergebnisse einer Suchanfrage wie »alle Gedichte in TextGrid« (mit über 50000 TEI-Dokumenten). Dies wurde beispielsweise benutzt, um die Gedichte in Klemens Bobenhausens Analysetool *Metricalizer* zu integrieren (z.Zt. noch in einer Testversion).

⁵⁶ Vgl. hierzu Christian Cherek, Stefan Funk, Roman Hausner, Ubbo Veentjer, Thorsten Vitt: *Kombiniere: Analog und Digital – Virtuelle Forschungsumgebungen auf dem Schreibtisch*. Beitrag zum DHD-Blog (<http://dhd-blog.org/?p=992>)

6.6.5 AP 5.5: Weiterentwicklung der Digitalen Bibliothek

Literatur

In der vorherigen Projektphase wurden die Daten des Literaturordners – des größten und aus literaturwissenschaftlicher Sicht interessantesten Teils der Digitalen Bibliothek – aus dem von Zeno verwendeten, historisch gewachsenen und auf die Webdarstellung optimierten XML-Format in semantisches TEI-Markup zusammen mit TextGrid-Metadaten konvertiert und im TextGridRep veröffentlicht. Dafür wurde ein komplexer Workflow von XSLT-Transformationen entwickelt.

Diese in der ersten Version veröffentlichten Daten des Literaturordners der digitalen Bibliothek enthielten eine Reihe von Fehlern, die unter anderem auf fehlerhafte Heuristiken im Transformationsworkflow zurückzuführen waren. Um diese Problematik künftig zu begrenzen, wurden eine Reihe von Validierungswerkzeugen entwickelt, die nach Überarbeitungen des Workflows automatisch laufen und vor gegebenenfalls auftretenden Problemen warnen.

Der Transformationsworkflow wurde aus diesen Erfahrungen heraus überarbeitet, deutlich vereinfacht und klarer dokumentiert. Die Überarbeitung zielt auch auf die Nachnutzbarkeit: Um diese Transformationsroutine auf die übrigen zeno-Daten übertragen zu können, sind Anpassungen an die jeweilige Datenstruktur nötig. Es wurde deshalb daran gearbeitet, die vorhandene Transformationsroutine modular zu gestalten, so dass die nötigen Anpassungen effizienter vorgenommen werden können. Der überarbeitete Literaturworkflow wird darüber hinaus als Basis für die neben Enzyklopädien und Literatur noch offenen Bereiche (Philosophie etc.) dienen. Da diese Bereiche wesentlich einfacher strukturiert sind, ist der Ablauf für diese Bereiche (1) Übernahme des Literatur-Workflows, (2) Entfernung der Literatur-Spezifika, (3) Werkauszeichnung für den Bereich und (4) Anpassung an die jeweiligen Ideosynkrasien des zu bearbeitenden Bereichs.

Bis zum Projektende wurde eine korrigierte Version der Literaturdaten fertiggestellt, die in den nächsten Monaten dann ins TextGrid-Repository eingespielt werden kann.

Philosophie

Auf der Basis der o.g. Modularisierung wurde ebenfalls ein Workflow für den Philosophiebereich der Digitalen Bibliothek entwickelt, der gewissermaßen als Benchmark für den Modularisierungsansatz dient. Die Philosophiedaten wurden zum Projektende ebenfalls fertiggestellt und werden zunächst zur Sichtung in eine Testinstanz, in den folgenden Monaten dann in die Produktivinstanz des TextGrid-Repository eingespielt.

Daneben wurden eine Reihe von Entwicklungstools gebaut bzw. überarbeitet, u.a. eine GUI, mit der für einen Bereich manuell ausgezeichnet werden konnte, auf welcher Hierarchieebene ein Werk (aus philologischer Sicht) vorliegt – diese Information ist in den Ausgangsdaten nicht vorhanden, bildet jedoch die Grundlage für die Zerlegung entsprechend dem TextGrid-Metadatenmodell.

Enzyklopädien

Die Sammlung der digitalen Bibliothek enthält 22 Enzyklopädien, die ebenfalls in ein einheitliches TEI-Encoding überführt worden sind.

Die Textsorte unterscheidet sich recht deutlich von den übrigen Inhalten der Digitalen Bibliothek, weshalb von den für die anderen Textsorten entwickelten Stylesheets nur ein Teil übernommen werden konnte. Hier ist im Grunde jede Enzyklopädie nur ein Werk, das jedoch aus einer Vielzahl von

Einzelartikeln besteht. Für die Zerlegung wurde deshalb entschieden, als Kompromiss zwischen Dateigröße und Dateianzahl auf eine Zerlegung in Buchstabenstrecken (jeweils separat für Hauptteil und Anhang/Nachtragsbände) zurückzugreifen, und eine Unterstützung für einzelne Artikel in Suche und Webanzeige im TextGridRep (im Rahmen generischerer Mechanismen) zu implementieren.

Die Strukturierung und Auszeichnung der Originaldateien der Enzyklopädien erwies sich während der Bearbeitung als sehr uneinheitlich. Deshalb wurde ein stark modularisierter Workflow entworfen: Für die meisten Schritte gibt es jeweils eine für jede Enzyklopädie individuelle Steuer-Transformation, die auf gemeinsame Schritte für die einheitlich codierten Komponenten zurückgreift. Trotz Modularisierung sind hier 140 Einzelstylesheets entstanden und entsprechend erwies sich die Entwicklung des Enzyklopädiensworkflows als höchst zeitaufwändig. Der Enzyklopädiensworkflow konnte 2014 soweit abgeschlossen werden, dass nun (auch nach entsprechenden Anpassungen am Repository für eine sinnvolle Behandlung der Textsorte mit einzelnen Enzyklopädieartikeln, die in Dateien für die Buchstabenstrecken vorliegen) die Daten probenhalber eingespielt und auf letzte Fehler korrigiert werden können.

Folgende Schritte wurden bei der Workflowgestaltung umgesetzt:

- Manuelles Kodieren der Enzyklopädien-teiHeader
- Überführen der von editura erstellten Grundstruktur (z.B. Aufteilung einer Enzyklopädie in Nachtragsbände, Appendix und Lexikonteil) in eine TEI-teiCorpus-Struktur. Hier musste für jede der Enzyklopädien ein eigenes Mapping erstellt werden.
- Transformieren der Enzyklopädien nach TEI.
- Maschinelles Splitten der Enzyklopädien nach Buchstabengrenzen, manuelle Korrekturen der Ergebnisse.
- Mapping der gesplitteten Daten auf das TextGrid-Metadatenmodell
- Auflösung von Verweisen zwischen zwischen Artikeln als Links
- Generierung von Metadaten und Links für die in den Texten eingebundenen Grafiken

Toolweiterentwicklung und -migration:

Einige Tools konnten bis Projektende nicht wie geplant weiterentwickelt und in die DARIAH-DE Infrastruktur migriert werden. Konkret sind davon insbesondere der Text-Bild-Link-Editor (TBLE) und der Text-Text-Link-Editor (TTLE) betroffen, zwei zentrale Werkzeuge von TextGrid, die eine starke Nutzung seitens der Community erfahren. Um die Toolweiterentwicklung und -migration in den Produktivbetrieb abzuschließen, wurde eine kostenneutrale Verlängerung über sieben Monate von Juni-Dezember 2015 bewilligt, finanziert aus Restmitteln des IDS Mannheim über Unterauftrag DAASI GmbH.

6.7 AP 6: Integriertes Monitoring und Qualitätssicherung

Das Ziel des Arbeitspakets war die Entwicklung und Implementierung eines Integrierten Monitorings in TextGrid. Die Aufgabe eines solchen Monitorings ist eine betriebsbegleitende und multidimensionale Beobachtung und Analyse von in einer Virtuellen Forschungsumgebung (VFU) ablaufenden Prozessen. Es soll dabei als Analysewerkzeug vor allem da eingesetzt werden, wo

Prozesse nicht automatisiert überwacht werden und / oder wo ein akuter Analysebedarf festgestellt wird. Im Gegensatz zu einem technischen Monitoring, das Verfügbarkeit und Performanz technischer Prozesse auswertet und dokumentiert, beobachtet und analysiert dieses Monitoring qualitativ die Wechselbeziehungen und Kommunikationsprozesse zwischen den NutzerInnen, der Technik und den BetreiberInnen der VFU. Das IM setzt an der Schnittstelle zwischen den sozialen, den technischen und den organisatorischen Aspekten der Forschungsinfrastruktur an. Seine Funktion ist es, ausgehend von den Mensch-Technik- und Mensch-Mensch-Interaktionen, die zweckmäßige und erfolgreiche Nutzbarkeit der VFU abzusichern bzw. diese Absicherung zu unterstützen. Bei Bedarf regt es steuernde, also gestaltende Eingriffe durch die Betreiber an. Es ist daher Teil einer Qualitätssicherung innerhalb der VFU.

Folgende Definition wurde für das Integrierte Monitoring vorgeschlagen:

Das Integrierte Monitoring ist ein Verfahren, um alle in einer Virtuellen Forschungsumgebung auftretenden Prozesse mit dem Ziel einer höheren Prozessrationalität analytisch zu begleiten, zu überprüfen und zu dokumentieren, um möglichen negativen Entwicklungen durch steuerndes Eingreifen frühzeitig zu begegnen und somit die Erfolgchancen einer Virtuellen Forschungsumgebung zu erhöhen.

R 6.1.1 und R 6.1.2 Dimensionen des integrierten Monitoring einer digitalen Forschungsinfrastruktur und Konzept Beurteilungsmethodik, Beurteilungskategorien und Governance-Werkzeuge⁵⁷

In einem ersten Schritt wurde die Ausgangsidee hinsichtlich der Dimensionen für ein solches Monitoring Beurteilungsmethodik, der Beurteilungskategorien und möglichen Steuerungsmaßnahmen zu einem Konzept ausdifferenziert. Die Ergebnisse liegen in den Reports 6.1.1 und 6.1.2 vor. Diese wurden als Ausgangspapier für die Folgeschritte sowie als Dokumentation des Konzepts des Monitorings zusammengeführt. Der Bericht wurde im Sommer 2013 vorgelegt und auf dem Konsortialtreffen in Worms (11./12.11.2013) diskutiert. Das Konzept des Monitorings wurde weiterhin auf der DHd-Tagung 2014 in Passau einem externen Publikum präsentiert und zur Diskussion gestellt. Im nächsten Schritt galt es, das Monitoring auf eine im vorliegenden Kontext und die gegebenen Bedingungen praktikable Anwendungsform hin zu konkretisieren. Das Ergebnis der Diskussion auf dem Konsortialtreffen war die Entscheidung einer pragmatischeren Fokussierung des Monitorings auf naheliegende und bereits etablierte Anwendungskontexte. In der Folge wurden u.a. Interviews mit dem Ziel einer weiteren Bedarfsanalyse mit AnwenderInnen und ZielgruppenvertreterInnen geführt, wobei die Interviews zugleich als Testlauf für eines der konkreten Erhebungsverfahren des Monitorings evaluiert wurden. Der dabei ermittelte Hauptbedarf liegt bei der Vermittlung von Nutzungskompetenzen und der Organisation von Anpassungen der technischen Lösungen an die Anforderungen der (potentiell) TextGrid nutzenden Projekte. Das Verfahren des Telefoninterviews erwies sich als für das Monitoring sehr geeignet. Aufbauend auf diesen Erkenntnissen wurden Nutzerinteraktionen, also Schulungen und Nutzertreffen als Hauptanknüpfungspunkt für die Konkretisierung des Monitorings definiert. Zudem zeigte sich eine

⁵⁷ https://textgrid.de/documents/10644/18815/TextGrid_R6.1.1undR6.1.2_integriertes_Monitoring_und_Governance.Werkzeuge.pdf/58fc9971-c4d0-4103-b855-e31dd29b938e

naheliegende Verbindung dieser Bedarfe mit dem AP 5.3 (Usability), mit dem in der Folge eine verstärkte Zusammenarbeit erfolgte.

Aktivitäten des Monitorings im TextGrid-Kontext

Parallel wurde das Monitoring in zwei organisationalen Anwendungsfällen aktiv. So wurde ein Arbeitspapier zur wechselseitigen Konstituierung von DARIAH-DE und TextGrid entworfen und vorgelegt. Darin wurden mögliche Synergien zwischen den beiden Projekten differenziert untersucht, wobei das AP besonders die Position von TextGrid erarbeitete. Aus dieser Position heraus wurde eine Zusammenführung von TextGrid und DARIAH-DE aus drei Gründen empfohlen: Erstens ist TextGrid eines der umfänglichsten und zugleich eines der frühesten dezidiert im Bereich der Digitalen Geisteswissenschaften angesiedelten und damit auch bekanntesten Projekte. Zweitens bietet TextGrid einen von vielen Projekten genutzten Produkt- und einen Servicerahmen, ist also fest in der geisteswissenschaftlichen Forschungslandschaft verankert. Drittens ist es aus Sicht der Nachhaltigkeit geboten, Kompetenz, Expertise und Entwicklungsmöglichkeiten sowohl von TextGrid wie auch von DARIAH-DE zu bündeln.

Weiterhin wurde auf dem TextGrid-Konsortialtreffen am 11./12.11.2013 in Worms in Kooperation mit der Konsortialleitung ein World-Café zu den Perspektiven der VertreterInnen des Konsortiums auf TextGrid durchgeführt. Das World-Café dient als Verfahren zur Erfassung und Analyse heterogener Einstellungen, Perspektiven und Deutungen innerhalb eines Projektzusammenhangs. Dabei wird ein offener Kommunikationsraum geschaffen, der alle Positionen gleichberechtigt zulässt, erfasst und expliziert. Die Ergebnisse werden zugänglich gemacht, zusammengefasst und für einen weiteren konstruktiven Dialog mit dem Ziel der Verständigung und gegebenenfalls Anpassung der Projektziele bzw. der Einleitung von Gestaltungsschritten aufbereitet und vermittelt. Das Verfahren wurde im Rahmen von TextGrid erfolgreich auf dem Konsortialtreffen in Worms angewandt. Die Ergebnisse wurden auf dem Konsortialtreffen am 30.06.2014 in Mannheim präsentiert und in einem Report an die Konsortialleitung kommuniziert.

Am 22. und 23. September 2014 führte das AP 2 in Kooperation mit DARIAH-DE einen Entwicklerworkshop zum Thema Usability durch. Das Integrierte Monitoring begleitete die Veranstaltung mit einer ausführlichen Vor-Ort-Dokumentation, die zu einer Publikation führte (Kaden, Rieger, 2015) Dabei wurde auch das Verfahren einer rückkopplungsgerichteten Erkenntnisgewinnung während des Workshops erfolgreich durchgeführt. Auf einer inhaltlichen Ebene wurden in diesem Zusammenhang Eckpunkte für ein Usability-Monitoring erarbeitet.

Eine gute Usability ist für die Akzeptanz und damit auch den Erfolg von Angeboten wie TextGrid entscheidend. Das Usability-Monitoring versucht, eine entsprechende Sensibilität bei möglichen weiteren Entwicklungsschritten von TextGrid abzusichern und auf die Umsetzung eines passenden Usability Engineerings hinzuwirken. Es kann bei Bedarf und je nach Ressourcenlage auch eigene Usability-Tests anregen, koordinieren bzw. durchführen.

M 6.1.3 Workshop mit Projektpartnern zum erarbeiteten Konzept des integrierten Monitoring

Am 21.11.2014 führte das AP 6.1 im MPIWG in Berlin einen Workshop zur *Implementierung des Integrierten Monitorings* in die Nutzungs- und Nutzerkommunikationsstrukturen von TextGrid durch. Dieser Workshop hatte das Ziel, Implementierungspotentiale anhand von drei Use Cases zu spezifizieren. Entsprechend war der Gegenstand des Workshops ein exemplarisches Framing des

Monitorings auf die Anwendungsfälle (a) Nutzerschulungen, (b) Nutzertreffen, (c) Nutzerforum. Dabei ging es vor allem um die Frage, wie sich das Konzept des Integrierten Monitorings in einem realistischen Anwendungskontext vor dem Hintergrund überschaubarer Ressourcen in einen Dauerbetrieb von TextGrid implementieren lässt. Weiterhin wurde konkretisiert, welche Prozesse mit welchen Verfahren welche Art von Wissen über das System VFU TextGrid ermittelbar machen. Die Erkenntnisse und weitere während des Workshops ermittelte Anwendungsfälle wurden ausgewertet und flossen in die Erstellung eines Dokumentationssystems ein, welches den Kern der Implementierung des Monitorings bildet.

M 6.1.4 Implementierung des integrierten Monitoring, Report Integriertes Monitoring einer digitalen Forschungsinfrastruktur

Die Implementierung basiert darauf, dass das Verfahren umfassend als Anwendungslösung dokumentiert zusammen mit dem Dokumentationssystem für die Monitoring-Prozesse zunächst im TextGrid-Confluence-Wiki abgelegt wird. Der Ansatz des Monitorings ist so generisch gehalten, dass er prinzipiell auch in anderen Zusammenhängen, beispielsweise auf DARIAH-DE Anwendung finden kann. Die Dokumentation umfasst ein navigierbares Ablagesystem für Beschreibungen einzelner Monitoringprozesse, das es ermöglicht, sowohl auf den Erkenntnisstand des Monitorings als auch auf Best-Practice-Fälle der durch das Monitoring durchgeführten Prozessbeobachtung zurückgegriffen werden kann.

R 6.2.1 Report Ergebnisse der Nutzerbefragung⁵⁸

Die Aufgabe des AP 6.2 war es, Bedürfnisse verschiedener Nutzergruppen zu ermitteln, wobei der Hauptfokus auf Editionsprojekten des Akademienprogramms lag. Diese bilden als Langzeitvorhaben im Bereich der geisteswissenschaftlichen Grundlagenforschung eine große und besonders stabile Gruppe von potentiellen NutzerInnen, die an kollaborativer Arbeit, hybriden Publikationsmöglichkeiten und der langfristigen Nutzbarkeit ihrer Forschungsdaten durch die wissenschaftliche Gemeinschaft interessiert ist.

In drei aufeinander folgenden Durchgängen wurden anhand eines Fragebogenleitfadens offene Interviews mit MitarbeiterInnen aus Vorhaben verschiedener Akademien, universitärer und außeruniversitärer Forschungsvorhaben sowie mit EinzelwissenschaftlerInnen geführt. Die Ergebnisse der Befragung wurden ausführlich im Bericht R 6.2.1 *Ergebnisse der Benutzerbefragung* publiziert. Es stellte sich heraus, dass TextGrid zu diesem Zeitpunkt bei den FachwissenschaftlerInnen nur oberflächlich bekannt war und in der Anwendung keine größere Verbreitung gefunden hatte. Daher wurden zum einen verschiedene TextGrid-Leistungen selbst experimentell getestet und zum anderen ein Mittel gesucht, TextGrid innerhalb der Akademien bekannter zu machen. Als vielversprechender Weg zeichnete sich der Kontakt zu den IT-Verantwortlichen der Akademien ab, die in der Regel als FachberaterInnen den Forschungsvorhaben bei der Antragstellung und in der laufenden Projektarbeit in informationstechnologischen Belangen zur Seite stehen, und deren Empfehlungen für gewöhnlich befolgt werden.

Im Rahmen eines an der Bayerischen Akademie der Wissenschaften in München durchgeführten Workshops der AG Digitale Forschung der Union der Akademien konnte TextGrid am 05.11.2013 den

⁵⁸ <https://textgrid.de/documents/10644/18815/R+6.2.1+Ergebnisse+der+Nutzerbefragung.pdf/c8670f61-c00b-42b6-b98c-6e67414a40af>

IT-Verantwortlichen der Akademien präsentiert werden. Für die weitere Zusammenarbeit und Planung von Veranstaltungen in den einzelnen Akademien wurde zudem eine Liste der AnsprechpartnerInnen in den Akademien erstellt.

6.8 H. AP 7: Projektmanagement und Öffentlichkeitsarbeit (Koordination durch AP-Leitung/SUB)

Ziel des Arbeitspakets war die Unterstützung der Projektpartner, um sicherzustellen, dass das Projekt alle Ziele innerhalb des vorgegebenen Zeitrahmens und des Budgetrahmens erreicht. Dazu gehörte die fachliche wie auch die formale und administrative Leitung des Projektes. Weitere Schwerpunkte dieses Arbeitspakets waren die koordinierte Öffentlichkeitsarbeit für das Verbundprojekt sowie die Vernetzung mit anderen Digital Humanities-Initiativen und Projekten (z.B. CLARIN, DARIAH-DE). Im Zuge der Etablierung einer nachhaltigen Institution kam zudem der Abstimmung mit den Wissenschaftsorganisationen (Fachgesellschaften etc.) eine Schlüsselrolle zu. Darüber hinaus fiel die Koordination der projektinternen Arbeitsgruppen und Task Forces sowie die Vorstellung von TextGrid im Rahmen von Schulungen, Nutzertreffen und verschiedenen nationalen und internationalen Fachtagungen (z.B. auf der DHd-Tagung vom 25.-28.03.2014 in Passau, der DH-Konferenz vom 8.-11.07.2014 in Lausanne) in diesen Bereich.

6.8.1 AP 7.1 Projektmanagement i.e.S und Koordinierung potentieller Kooperationen

Zentrale Aufgaben im Rahmen der Gesamtleitung des Projekts waren die Abstimmung der einzelnen Arbeitspakete untereinander und deren Verknüpfung miteinander sowie die Koordination und Verknüpfung von Ergebnissen. Zum administrativen und organisatorischen Projektmanagement gehörten die finanzielle Verwaltung der zentral eingesetzten Mittel ebenso wie die Sicherstellung der Einhaltung von Meilensteinen und der termingerechten Abgabe von Ergebnisberichten. Das Projektmanagement hatte ebenfalls die Verantwortung für den reibungslosen Ablauf der internen und externen Kommunikation. Es wurde z.B. eine Dokumentation der drei Projektphasen im Confluence-Wiki und fortlaufende Nutzerstatistiken erstellt. Presse- und Fachveröffentlichungen wurden regelmäßig publiziert und wie die Projekt-Ergebnisse auf der TextGrid-Homepage dokumentiert.

Die Koordinierung potentieller fachlicher wie auch strategischer Kooperationen bildete einen weiteren Schwerpunkt dieses APs. So wurde ein stetiges Zusammenwachsen mit DARIAH-DE forciert (TaskForce TextGrid/DARIAH-DE AK Architektur und AK Nachhaltigkeit/Finanzierungskonzept). Im Sommer 2014 wurde die TextGrid Nutzerverwaltung in die DARIAH-DE Nutzerverwaltung integriert. Von den aktuell etwa insgesamt 2.385 existierenden DARIAH-DE Accounts sind rund 1.760 Accounts TextGrid zuzuordnen.

Bereits in der zweiten Förderphase konnte TextGrid zahlreiche Kooperationen mit Forschungsverbänden konzipieren und gestalten, die in der dritten Förderphase in die konkrete Umsetzung gingen. Dazu gehören u.a. die Projekte, die im Rahmen des BMBF-Calls für eHumanities als förderungswürdig eingestuft wurden, aber auch strategische Kooperationen wie die Anbindung von SADE (Scalable Architecture for Digital Editions). Kooperationen im Rahmen gemeinsamer Forschungsvorhaben mit anderen Forschungsverbänden bestehen z.B. mit den Vorhaben „Hybrid-Edition von Theodor Fontanes Notizbüchern“ der Fontane-Arbeitsstelle an der Georg-August-

Universität Göttingen, dem Sammlungs- und Editionsprojekt „Johann Friedrich Blumenbach – online“ der Akademie der Wissenschaften zu Göttingen, der „Bibliothek der Neologie“ an der Universität Münster und dem Projekt „Textdatenbank und Wörterbuch des klassischen Maya“ der Universität Bonn. Darüber hinaus erfolgte eine Beratung mehrerer Forschungsvorhaben im Rahmen der Konzeption von Förderanträgen (DFG, Akademie der Wissenschaften) und dem Einsatz von TextGrid.

6.8.2 AP 7.2 Abstimmung mit den Fachgesellschaften

Zentrale Aufgabe war die koordinierte Vermittlung von Ergebnissen in die Fach-Communitys durch intensiven Austausch mit den für das Projekt und dessen Nutzer-Community relevanten Fachgesellschaften. Einen weiteren Schwerpunkt bildete die enge Vernetzung mit eHumanities-Zentren sowie Forschungsinfrastrukturinitiativen und -projekten im nationalen und internationalen Kontext (z.B. DARIAH), um Aufbau und Weiterentwicklung einer übergreifenden eHumanities-Forschungsinfrastruktur weiter voranzutreiben. Zur Unterstützung der anderen Arbeitspakete und zur Ansprache der einzelnen Zielgruppen wurden regelmäßige Experten- und Strategiegelgespräche geplant und durchgeführt. Es erfolgten außerdem zahlreiche Gespräche mit Fachgesellschaften wie Digital Humanities-Zentren hinsichtlich der Verstetigung Virtueller Forschungsumgebungen und im Hinblick auf mögliche Finanzierungskonzepte.

R 7.2.1 Konzept / Organisation Experten-, Strategiegelgespräche und TextGrid Summit⁵⁹

Konzeption, Organisation und Durchführung des von TextGrid zusammen mit DARIAH-DE ausgerichteten Digital Humanities Summit im März 2015 im Harnack-Haus in Berlin:

Die Konferenz richtete sich in erster Linie an die vom BMBF geförderten 24 DH-Projektverbünde, DH-Zentren und Nachwuchsgruppen sowie weitere assoziierte Projektpartner von TextGrid und DARIAH-DE. Die zweitägige Hauptkonferenz am 3. und 4. März 2015 bot zahlreiche Vorträge namhafter VertreterInnen aus Politik und Wissenschaft zu verschiedenen Teilbereichen der digitalen Geisteswissenschaften, inklusive Poster-Slam und DARIAH-DE DH-Award-Verleihung. Im Rahmen der TextGrid Grand Tour (sechstes TextGrid-Nutzertreffen) am 5. März 2015 konnten die verschiedenen Möglichkeiten der Forschungsumgebung mit MitarbeiterInnen in Exkursionen ausprobiert werden. Der Summit wurde im Nachhinein umfassend auf der entsprechenden Veranstaltungsseite im DARIAH-DE Portal unter www.dhsummit2015.de dokumentiert (Veröffentlichung des Grafic Recording, einer filmischen Dokumentation inklusive Imagefilm, der vorgestellten Poster und gehaltenen Präsentationen).

R 7.2.2 Kooperation mit Fachgesellschaften⁶⁰

Die kontinuierliche Re-Evaluation der Projektziele und ihrer Umsetzung, aber auch die potentielle Erschließung neuer Zielgruppen erfolgte in Abstimmung mit den für das Projekt und dessen Nutzercommunity relevanten Fachgesellschaften, wie etwa dem Mediävistenverband, der Deutschen Gesellschaft für Sprachwissenschaft oder dem deutschen Germanistenverband.

Über bestehende Kooperationen hinaus ist TextGrid stetig im Gespräch mit anderen Forschungsprojekten, -verbänden und -institutionen, um an bestehende Entwicklungen anzuknüpfen

⁵⁹ https://textgrid.de/documents/10644/18815/R+7.2.1_Experten-+und+Strategiegelgespr%C3%A4che.pdf/052fd53e-c8fa-4a82-8600-b09feba0fe99

⁶⁰ https://textgrid.de/fileadmin/TextGrid/reports/TG_R_7.2.2_-_Kooperation_mit_Fachgesellschaften.pdf

und sie fortzuführen. Dazu zählen u.a. DARIAH-DE, CLARIN, TEXTvire, das King's College London sowie die Stanford University und die North Eastern University (beide USA). Als Forschungsinfrastrukturprojekt für die Geistes- und Kulturwissenschaften weist DARIAH-DE eine große Übereinstimmung mit TextGrid in Bezug auf Zielgruppe und Vermittlungsstrategie auf, weswegen eine Zusammenarbeit auf diesem Gebiet sinnvoll erscheint und viele Synergieeffekte verspricht. Um die Kommunikation mit den Fachgesellschaften aufzunehmen, hat DARIAH-DE ein eigenes Stakeholdergremium „Fachgesellschaften“ etabliert, das vom Verband „Digital Humanities im deutschsprachigen Raum“, vertreten durch die Universität Hamburg, koordiniert wird. Es soll sich aus VertreterInnen von Fachgesellschaften zusammensetzen und jährlich zu einem persönlich Treffen zusammenkommen. Im Rahmen des DH Summits im März 2015 fand das erste Treffen des Stakeholdergremiums „Fachgesellschaften“ statt. Hierbei wurden u.a. die Fragen thematisiert, welche Bedeutung digitale Forschungsprozesse für einzelne Disziplinen haben, ob und wie dies auf curriculare Entwicklung Einfluss nimmt und welche Anforderungen einzelne Fachdisziplinen an Digitale Forschungsumgebungen und Forschungsinfrastrukturen haben. Eine weitere Maßnahme stellt die Präsenz auf den Tagungen der Fachgesellschaften dar, auf denen TextGrid beispielsweise durch Projektvorstellungen, Posterpräsentationen und Demo-Sessions vertreten ist. So wurden beim Historikertag 2014 in Göttingen FachwissenschaftlerInnen aktiv zu ihren Anforderungen und Bedürfnissen befragt.

Auch durch die regelmäßige Organisation und Durchführung von Workshops und Konferenzen erhöhte sich die Aufmerksamkeit seitens relevanter Forschungsprojekte. In Zusammenarbeit mit dem EU-Projekt EHRI (European Holocaust Research Infrastructure) und DARIAH-EU haben DARIAH-DE und TextGrid am 09. Juli 2013 im Jüdischen Museum Berlin eine internationale Konferenz zum Thema „*Public History of the Holocaust: Historical Research in the Digital Age*“ veranstaltet, die sich vorwiegend an HistorikerInnen und InformationswissenschaftlerInnen richtete. Der Fokus lag auf den Möglichkeiten und Grenzen des Internets als öffentlichem Raum für die Holocaust-Forschung.

6.8.3 AP 7.3 Marketing und Öffentlichkeitsarbeit

Die koordinierte Öffentlichkeitsarbeit und eine professionelle Marketingstrategie waren Kernaufgaben, bei denen es darum ging, einen möglichst hohen Grad an Akzeptanz bei den Fach-Communitys als Zielgruppe und den politischen Entscheidungsträgern zu erreichen. Eine den gesamten Projektverlauf umfassende PR- und Marketingstrategie und die daraus folgenden konkreten Umsetzungen gründeten die Basis und unterstützten damit maßgeblich die Aktivitäten in AP 3 (Community-Bildung) sowie den Aufbau eines nationalen eHumanities-Netzwerkes.

Die koordinierte Vermittlung von Ergebnissen in die Fach-Communitys erfolgte mittels verschiedener geeigneter PR-Materialien und Social-Media-Strategien. Ereignisorientierte Aktivitäten waren z.B. die in 2014 erfolgte Produktion des DH-Films „Digitale Wissensräume: Virtuelle Labore der Geisteswissenschaften“ von TextGrid und DARIAH-DE ebenso wie die Einrichtung eines übergreifenden Digital Humanities-Blogs im deutschsprachigen Raum, ferner die Etablierung eines Bug-Reporting- bzw. Support-Systems, mit dem unerwartete Softwarefehler sowie Feature Requests rund um TextGridLab und TextGridRep über JIRA an TextGrid gemeldet werden können (seit Herbst 2013). Diese Aktivitäten machen die erzielten Arbeitsfortschritte des Verbundvorhabens innerhalb und außerhalb der Zielgruppe sichtbar.

M 7.3.1 Webauftritt TextGrid

Die Präsenz von TextGrid im World Wide Web war im Rahmen der Akquise neuer Nutzergruppen für den Auf- und Ausbau der Community von zentraler Bedeutung. Die TextGrid-Webseite⁶¹ stellt als externe Kommunikationsplattform die Schnittstelle zur Öffentlichkeit dar. Es stehen Informationen über Software und Dienste, Anwendungen und Nutzungsszenarien selbst im Vordergrund. Die Startseite bietet mit Laboratory, Repository und Community/Dokumentation direkten Zugriff auf die „drei Säulen von TextGrid“. Die Webseite wurde laufend an die jeweiligen Anforderungen angepasst, aktualisiert und überarbeitet. Im Frühjahr 2015 wurde die Webseite von der TYPO3-Instanz auf eine Liferay-Instanz umgezogen. In diesem Zusammenhang erfolgte eine strukturelle und inhaltliche Überarbeitung von TextGrid-Homepage und -Repository inklusive Redesign und grafischer Überarbeitung des TextGrid Laboratorys.

R 7.3.1 Öffentlichkeitsarbeit (laufend): Fortlaufende Dokumentation aller ÖA-Maßnahmen⁶²

Die koordinierte Öffentlichkeitsarbeit dient als zentrales Instrument, Fachcommunitys gezielt über die Angebote der Virtuellen Forschungsumgebung TextGrid zu informieren und damit sowohl den nachhaltigen Betrieb als auch die Akzeptanz und Nutzung von TextGrid zu unterstützen und weiter auszubauen. Dazu wurde eine Reihe von Maßnahmen eingesetzt, z.B. die Konzeption und Koordination des TextGrid-Sammelbandes zum Projektabschluss, der im Rahmen einer Anthologie das Projekt TextGrid aus verschiedenen Perspektiven beleuchtet. Imagefilme, Explainity-Videoclips (zu den Themen „Datensicherheit für Forschungsdaten in den Geisteswissenschaften“, „Kollaborativ arbeiten in digitalen Forschungsinfrastrukturen“ und „Digitales Edieren in einer Virtuellen Forschungsumgebung“), Mitschnitte von Tagungen und Vorträgen sowie die TextGrid Video-Tutorials stehen im 2014 eingerichteten DHd-Kanal auf YouTube zur Verfügung. Verschiedene Projekt-Werbematerialien wurden für den erwünschten Wiedererkennungseffekt und die Etablierung von TextGrid als „Marke“ erstellt, im Rahmen des DH-Summits 2015 z.B. DIN A5 Blöcke, Kugelschreiber und Taschen mit TextGrid und DARIAH-DE Logo plus URL, darüber hinaus z.B. Roll-Ups und zwei neue TextGrid-Flyer, die auf Konferenzen und im Rahmen von Vorträgen und Schulungen über TextGrid informieren.

R 7.3.2 Social-Media Konzept, Umsetzung & Betreuung⁶³: Dokumentation aller Nutzungsformen von Social-Media innerhalb des Projektes

Social Media-Formate und deren Akzeptanz und Potential in der Wissenschaftscommunity im Allgemeinen sowie in der TextGrid-Community im Besonderen stehen im Vordergrund dieses Reports. TextGrid nutzt verschiedene Kommunikationswerkzeuge wie den DHd-Blog, ein öffentliches Wiki und öffentliche Mailinglisten, Twitter sowie den DHd-Kanal auf YouTube. Der gemeinsam von TextGrid, DARIAH-DE und dem Max-Planck-Institut für Wissenschaftsgeschichte initiierte DHd-Blog - Digital Humanities im deutschsprachigen Raum⁶⁴ informiert über aktuelle Themen, Entwicklungen, Dienste, Projekte und Veranstaltungen sowie über Förder- und Stellenausschreibungen in Deutschland, Österreich und der Schweiz und lädt zur Diskussion ein. Für den Blog wurde ein eigenes

⁶¹ <http://www.textgrid.de>

⁶² https://textgrid.de/documents/10644/18815/TextGrid_R7+3+1_%C3%96ffentlichkeitsarbeit+Vol3_final.pdf/916fdeba-2681-45e4-80ec-247dc198a3b6

⁶³ https://textgrid.de/fileadmin/TextGrid/reports/TextGrid_R7_3_2_Social_Media_2_0_final.pdf

⁶⁴ <http://dhd-blog.org/>

Logo und Design erstellt. Bei Projektende umfasste der Blog über 200 registrierte AutorInnen und über 600 Artikel. Ein öffentliches Wiki stellt Informationen zur Arbeit mit TextGrid (Laboratory, Repository, Architektur, technische Infrastruktur) zur Verfügung⁶⁵. Das Wiki umfasst ein User Manual sowie eine technische Dokumentation, die die Architektur, das TextGridLab als Frontend, das TextGridRep und Services als Backend dokumentiert.

Die öffentlichen Mailinglisten textgrid-extern@gwdg.de, textgrid-support@gwdg.de, textgrid-user@gwdg.de und textgrid-dev@gwdg.de stehen ForscherInnen, EntwicklerInnen und IT-SpezialistInnen, die mit TextGrid arbeiten oder ein Forschungsvorhaben planen, für Fragen oder zum aktiven Austausch mit anderen AnwenderInnen zur Verfügung.

Anlässlich des ersten TextGrid-Nutzertreffens im Februar 2012 hat TextGrid einen Twitter-Account⁶⁶ eingerichtet, um im Format des Microblogging von Tagungen und Workshops sowie zu aktuellen Themen rund um TextGrid, die beteiligten Forschungsvorhaben und -verbünde, aber auch zu relevanten Themen der Digital Humanities informieren zu können. Das Profil weist per Mai 2015 511 Follower auf, die Zahl der AbonnentInnen wächst jedoch auch nach Projektende beständig weiter.

R 7.3.3 Marketingkonzept & Umsetzung⁶⁷: Dokumentation aller Marketing-Maßnahmen

Das Disseminationskonzept wurde gemeinsam mit DARIAH-DE entwickelt und nimmt die zwischen TextGrid und DARIAH-DE existierenden Wechselbeziehungen zum Anlass, diese Vernetzungen weiter zu intensivieren und Synergien zu nutzen. In der Wechselbeziehung von DARIAH-DE und TextGrid wirkt DARIAH-DE als digitale Forschungsinfrastruktur als Fundament für digitale Geistes- und Kulturwissenschaften. TextGrid als konkrete, für die textbasierte Quellen- und Editionsarbeit optimierte Anwendung kann hierauf aufsetzen. Durch die gemeinsame Marketingstrategie wurden Synergien noch stärker genutzt und bereits entwickelte Angebote (nach Möglichkeit auch Forschungsdaten) noch stärker nachgenutzt, um Parallelentwicklungen zu vermeiden sowie Lessons-Learned und Best-Practice-Erfahrungen in kommende Entwicklungen einzubeziehen.

Ziel der Marketingmaßnahmen war es, sowohl die breitere Fachöffentlichkeit für das Thema der digitalen Geistes- und Kulturwissenschaften zu sensibilisieren und den nachhaltigen Ausbau der digitalen Forschungsumgebungen TextGrid und DARIAH-DE öffentlichkeitswirksam zu begleiten als auch auf die spezifischen Anforderungen der verschiedenen Nutzergruppen einzugehen und somit eine stärkere Wirkung von DARIAH-DE und TextGrid in die entsprechenden Fach-Communitys hinein zu erzielen. Außerdem wurde eine enge Verzahnung der Marketingaktivitäten von DARIAH-DE und TextGrid angestrebt, um Synergieeffekte zu nutzen.

6.8.4 Finalisierung des Nachhaltigkeitskonzepts und Koordination

Um die Finanzierung und damit den Weiterbetrieb von TextGrid über das Projektende im Mai 2015 hinaus sicher zu stellen, steht TextGrid im Gespräch mit verschiedenen Stakeholdern, wie dem Bundesministerium für Bildung und Forschung, mehreren Länderministerien und VertreterInnen der Hochschulrektorenkonferenz sowie des Wissenschaftsrates. Um die nachhaltige Finanzierungsstrategie zu vollenden und dauerhaft umzusetzen, wurde eine kostenneutrale Verlängerung über sieben Monate von Juni-Dezember 2015 bewilligt. Finanziert wird die

⁶⁵ <https://dev2.dariah.eu/wiki/display/TextGrid/Main+Page>

⁶⁶ <https://twitter.com/TextGrid> bzw. @TextGrid

⁶⁷ https://textgrid.de/fileadmin/user_upload/TG_R_7_3_3_-_Marketingkonzept_final.pdf

kostenneutrale Verlängerung aus Restmitteln der SUB Göttingen über einen Unterauftrag mit Prof. Axel Horstman (ProWiss), der bislang den Kontakt zu den Stakeholdern koordiniert und die Gespräche maßgeblich begleitet hat.

Für die Durchführung eines Nachhaltigkeitstreffens mit Bund- und LändervertreterInnen im Oktober 2015 im Harnack-Haus in Berlin, wurden ebenfalls Gelder aus Restmitteln der SUB Göttingen bewilligt.

Damit die noch zu leistenden Abschlussarbeiten sowie restlichen Migrationsprozesse ausreichend koordiniert und unterstützt werden können, wurde außerdem die kostenneutrale Verlängerung einer Koordinationsstelle bewilligt.

7. Die wichtigsten Positionen des zahlenmäßigen Nachweises

Georg-August-Universität Göttingen, Niedersächsische Staats- und Universitätsbibliothek Göttingen mit den Unterauftragnehmern:

- **Berlin-Brandenburgische Akademie der Wissenschaften Berlin**
- **Technische Universität Berlin**
- Der abschließende zahlenmäßige Nachweis wird nach Ende der kostenneutralen Verlängerung (01.06.2015-31.12.2015) nachgereicht

Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen mit den Unterauftragnehmern:

- **DAASI International GmbH, Tübingen**
- **Hochschule Worms, Zentrum für Technologietransfer und Telekommunikation (ZTT)**
- Bewilligte Summe Insgesamt: 433.396,00
- Ausgaben Insgesamt: 349.225,23
- Rest Insgesamt: 84.170,77
- Mittelpositionen wurden nicht überschritten
- Umwidmungen sind nicht erfolgt

Institut für Deutsche Sprache, Mannheim

- Der abschließende zahlenmäßige Nachweis wird nach Ende der kostenneutralen Verlängerung (01.06.2015-31.12.2015) nachgereicht

Julius-Maximilians-Universität Würzburg, Institut für Deutsche Philologie

- Bewilligte Summe Insgesamt: 298.752,00 € (incl. PP)
- Ausgaben Insgesamt: 260.183,53 €
- Rest Insgesamt: 38.568,47 €
- Mittelpositionen wurden nicht überschritten
- Umwidmungen sind nicht erfolgt

Max-Planck-Institut für Wissenschaftsgeschichte, Berlin

- Bewilligte Summe Insgesamt: 199.278,00 €
- Ausgaben Insgesamt: 249.783,44 €
- Rest Insgesamt: -50.505,44 €
- Mittelposition 0812 wurde um 51925,48 Euro überschritten. Unser ursprünglich für dieses Projekt mit einer halben Stelle eingesetzte wissenschaftlicher Mitarbeiter, Herr Urzua, musste im Laufe des Projektes (am 1.3.2013) aus fachlichen Gründen auf eine andere Stelle wechseln, da seine Kenntnisse dort dringend gebraucht wurden und zugleich für TextGrid zur

Erreichung der Projektziele eine Mitarbeiterin mit langjähriger Erfahrung im Bereich VREs benötigt wurde. Diese Stelle musste deshalb kurzfristig neu besetzt werden. Hierzu konnten wir die wissenschaftliche Mitarbeiterin Frau Rieger mit Erfahrung in der Organisation virtueller Forschungsumgebungen gewinnen. Da diese auf Grund ihrer Vorerfahrungen in eine höhere Besoldungsgruppe als beantragt eingeordnet wurde, ist das entsprechende Defizit entstanden. Die bereits in 2013 und 2014 angefallenen Mehrkosten wurden durch Eigenmittel der MPG gedeckt. In Bezug auf das Jahr 2015 finden z.Zt. noch Abstimmungen, in die der Projektträger einbezogen ist, statt.

- Der Posten 0843 wurde um 5191,94 Euro überschritten. Dieses ergibt sich rechnerisch durch die prozentuale Ansetzung des Betrages, wurde jedoch aus Eigenmitteln der MPG gedeckt.
- In Position 0846 besteht eine Überschuss von 6611,98 Euro.
- Umwidmungen sind nicht erfolgt

Technische Universität Darmstadt, Institut für Sprach- und Literaturwissenschaft

- Bewilligte Summe Insgesamt: 511.521,00 €
- Ausgaben Insgesamt: 502.169,17 €
- Rest Insgesamt: 9.351,83 €
- Mittelposition 0812 wurde um 1.594,49 € überschritten, Mittelposition 0817 um 10.810,93 €. Beide Überschreitungen sind durch Änderungen in den Tarifverträgen (rückwirkende Tarifierhöhungen sowie Einmalzahlungen) während der Projektlaufzeit bedingt. Mittelposition 0822 wurde um 4.086,21 € überschritten. Grund dafür waren die zu Projektende vermehrt anfallenden Arbeiten der Dokumentation, die durch Hilfskräfte unterstützt werden konnten sowie Anfragen für weitere Schulungen, die mit Unterstützung der Hilfskräfte ebenfalls durchgeführt werden konnten. Alle Überschreitungen konnten durch andere Mittelpositionen ausgeglichen werden.
- Umwidmungen sind nicht erfolgt

8. Die Notwendigkeit und Angemessenheit der geleisteten Arbeit

Der in der Vorhabenbeschreibung vorgestellte Arbeitsplan zum Erreichen des Gesamtziels wurde erfolgreich umgesetzt. Die dort formulierten Aufgaben wurden in sieben Arbeitspaketen, deren Teilziele in den jeweiligen Meilensteinen und Reports formuliert wurden, bearbeitet und in jährlich vorgelegten Zwischenberichten dargelegt. Der Verlauf der Arbeit im Projekt folgte der im Antrag formulierten Planung.

9. Der voraussichtliche Nutzen

Waren im Rahmen der ersten beiden Projektphasen die Voraussetzungen für eine geisteswissenschaftliche Virtuelle Forschungsumgebung geschaffen und grundlegende Entwicklungsarbeit hinsichtlich der Tools und Dienste erbracht worden, so sollte TextGrid III das grundsätzliche Ziel umsetzen, diese Forschungsumgebung in einen nachhaltigen Betrieb zu überführen und auf eine breite Nutzerbasis zu stellen. Zum Ende der ersten

Förderphase waren folgende zentrale Aspekte identifiziert:

- **Produktionsreife:** Die vorhandenen Werkzeuge und die Infrastruktur von TextGrid sollten sich nach den Beta-Tests in der ersten Projektphase im Produktionsbetrieb unter Volllast bewähren und entsprechend modifiziert werden.
- **Breitere Basis:** Um TextGrid auf eine breitere Anwender-Basis zu stellen, mussten die vorhandenen Werkzeuge an die Bedürfnisse weiterer Communities angepasst werden.
- **Geschäftsmodell:** Das vorliegende Geschäftsmodell-Konzept sollte umgesetzt und ggf. optimiert werden.
- **Nachhaltigkeit von Diensten und Daten:** Die nachhaltige Existenz und Auffindbarkeit der Dienste von TextGrid sowie die langfristige Verfügbarkeit der in TextGrid erstellten Daten sollte gewährleistet werden.
- **Community:** Eine große, international vernetzte Nutzergemeinde aus FachwissenschaftlerInnen, Institutionen und EntwicklerInnen ist unabdingbare Voraussetzung für den Fortbestand und die Weiterentwicklung von TextGrid. Für die langfristige Pflege der TextGrid-Produkte sollte um das Projekt eine breite fachwissenschaftliche Community aufgebaut werden.

Diese zentralen Aufgabenbereiche, in der zweiten Projektphase umgesetzt, formulierten jeweils eine wissenschaftliche, technische bzw. Nutzungs- bzw. Verwertungsperspektive. In die eigentliche Verwertungsphase trat TextGrid noch im Laufe der zweiten Förderphase, nämlich mit Erreichen der Produktionsreife im Juli 2011 ein. Seither wird TextGrid in Projekten sowie in der Lehre genutzt, wird mit KooperationspartnerInnen gemeinsam weiterentwickelt, ist in strategischen oder inhaltlichen Allianzen aktiv sowie Gegenstand wissenschaftlicher Begleitforschung. Rund 60 z.T. langfristig (über 10 Jahre) angelegte Forschungsverbünde nutzen heute die Virtuelle Forschungsumgebung TextGrid aktiv für die Erstellung wissenschaftlicher Editionen, für die Analyse ihrer Forschungsdaten, für die Entwicklung projektspezifischer Analyse- sowie Visualisierungswerkzeuge und schließlich für die langfristige digitale Archivierung und Bereitstellung der Forschungsdaten zur weltweiten Nachnutzung durch die Scientific Community.

Konkrete Nutzungsanfragen vonseiten verschiedener bei der Akademie der Wissenschaften und der Deutschen Forschungsgemeinschaft eingereichten bzw. einzureichenden Forschungsvorhaben wurden und werden begleitet von detaillierten Konzeptions- bzw. Kooperationsgesprächen. Allein im Sommer 2012 (Juni – August) erreichten TextGrid sechs konkrete Nutzungs- bzw. Kooperationsanfragen, darunter – auf internationaler Ebene – auch die Einladung seitens der Mellon Foundation, U.S.A., zur Einreichung eines gemeinsamen Antrags mit der Stanford University. Auf die dezidierte Nutzungsempfehlung von TextGrid im Rahmen des BMBF-e-Humanities-Calls sind weitere

zwölf Forschungsvorhaben zurückzuführen, die per 2012 sowie überwiegend mit Beginn 2013 ihre Arbeit aufgenommen haben. Im Mai 2014 starteten zwei weitere Langfristvorhaben: das auf 15 Jahre angelegte, von der Akademie der Wissenschaften und der Künste NRW geförderte Vorhaben "Textdatenbank und Wörterbuch des Klassischen Maya" sowie die an der Universität Münster/Seminar für Kirchengeschichte situierte "Bibliothek der Neologie", die über einen Zeitraum von 12 Jahren von der DFG gefördert wird. Weitere vier (Editions-)Vorhaben befinden sich derzeit im Begutachtungsverfahren der DFG, drei weitere Vorhaben in der Antragskonzeption (Einreichung: Sommer 2015-Frühjahr 2016).

Diese zahlreichen konkreten Nutzungsmodelle und Nutzungsanfragen, nicht zuletzt aber auch das große Interesse seitens der Fachwissenschaftlerinnen und Fachwissenschaftler, von Entwicklungsprojekten und weiteren Akteuren der eHumanities an den von TextGrid erstmals im Februar 2012 ausgerichteteten Nutzertreffen verweisen auf den hohen Bedarf an computergestützter, interdisziplinärer und fachübergreifender Forschung, wie TextGrid sie ermöglicht. Dabei wird der kontinuierliche und enge Austausch zwischen AnbieterInnen und NutzerInnen dezidiert nachgefragt. Auf technischer Ebene werden dabei zukünftig hinsichtlich Support- und Beratungs-, aber auch Kooperationsszenarien v.a. die Evaluation neuer Werkzeuge auf ihre Interoperabilität mit dem TextGridLab, die Integration von Software-Tools als Webservices im Lab, sowie die Konfiguration komplexer Workflows relevant.

Ein nachhaltiges Organisations- und Betriebsmodell der Virtuellen Forschungsumgebung bildet die Voraussetzung für das fortgesetzte Vertrauen der fachwissenschaftlichen Community, ihre Forschungsvorhaben mit TextGrid zu realisieren. Die Verstetigung von TextGrid, in der zweiten Projektphase mit der Gründungsversammlung des Vereins „TextGrid – Verein zum nachhaltigen Betrieb einer digitalen Forschungsinfrastruktur in den Geisteswissenschaften e.V.“ begonnen, war deshalb zentraler Gegenstand der dritten Förderphase.

10. Während der Durchführung eines Vorhabens dem ZE bekannt gewordener Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

DARIAH-DE

Mit dem Projekt DARIAH-DE, digitale Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften, wurde im Verlauf der dritten Förderphase die Integration verschiedener TextGrid-Komponenten in die Forschungsinfrastruktur vorangetrieben und eine Zusammenarbeit auf strategischer, struktureller und technischer Ebene vorbereitet.

Digilib & IIF

TextGrid beteiligt sich an der Weiterentwicklung der Software IIF⁶⁸, die Editionsprojekten ermöglicht, über den zentralen TextGrid-Bilderservice (digilib) eigene Scans bereitzustellen.

EDIARUM

Die von der Berlin-Brandenburgischen Akademie der Wissenschaften entwickelte digitale Arbeitsumgebung ediarum⁶⁹, ein u.a. auf oxygen basierendes Framework für Editionen, ermöglicht Editionsprojekten, ihre Ergebnisse in TEI-XML zu bearbeiten, zu speichern und zu präsentieren. Dieses Framework kann ins TextGridLab integriert werden.

EUDAT

Das europäische Projekt EUDAT hat sich auf die Fahnen geschrieben, eine Infrastruktur für das gemeinschaftliche Arbeiten mit Forschungsdaten zu entwickeln. Diese basiert im Kern auf einem integrierten Satz von Diensten, um Daten zu verwalten, zu teilen und für Berechnungen zu nutzen. Der hier gewählte Ansatz ist unabhängig von der Forschungsdisziplin. Der sogenannte B2SHARE Dienst, dessen Zweck das Speichern und Teilen von Forschungsdaten ist, erfüllt dabei eine Funktion, die in Teilen mit dem TextGrid/DARIAH Repository vergleichbar ist, wenn auch auf einer weniger spezifischen Ebene. B2SHARE basiert dabei selbst auf Invenio, einer populären Repository-Lösung. B2SHARE/Invenio ist hier nur als Beispiel für eine Reihe ähnlicher Produkte gewählt, die sich aktuell in der Entwicklung befinden bzw. im Forschungsumfeld genutzt werden.

SADE

TextGrid ist aktiv an der Weiterentwicklung der "Skalierbaren Architektur für digitale Editionen" (SADE) beteiligt, einem Programmpaket zum Entwickeln und Darstellen digitaler Editionen.⁷⁰

TUSTEP

Neue Entwicklungen in TUSTEP, insbesondere die neue XML-basierte Notation für TUSTEP-Befehle (TXSTEP) wurden von Seiten des Projekts verfolgt und mit Diskussionsbeiträgen begleitet.

⁶⁸ vgl. Vortrag: DHd Tagung in Graz 2015 (<http://gams.uni-graz.at/o:dhd2015.v.003>)

⁶⁹ <http://www.bbaw.de/telota/software/ediarum>

⁷⁰ vgl. R 4.4.2 SADE - Dokumentation und Referenzinstanz (http://textgrid.de/fileadmin/TextGrid/reports/TG-R-4_4_2-SADE_Dokumentation_und_Referenzinstanz.pdf)

11. Bereits erfolgte oder geplante Veröffentlichungen des Ergebnisses nach Nr. 6

11.1 Bereits erfolgte Veröffentlichungen

M. Bender, C. Krause, A. Rapp, O. Schmid und Ph. Vanscheidt: *TextGrid – eine virtuelle Forschungsumgebung für die Geisteswissenschaften*. In: *Digitale Medien – Werkzeuge für exzellente Forschung und Lehre*, hrsg. von G. Csanyi, F. Reichl und A. Steiner, Münster 2012, S. 124-125.

K. Betz: *Ein virtuelles Bücherregal: Die Digitale Bibliothek im TextGrid Repository*. In: *TextGrid: Von der Community – für die Community*, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 229-238.

M. Blümm, S. E. Funk und S. Söring: *Die Infrastruktur-Angebote von DARIAH-DE und TextGrid*. In: *Information. Wissenschaft & Praxis* 2015, 66(5–6): S. 304–312.

R. Casties und D. Wintergrün: *Bilder als Quelle in TextGrid*. In: *TextGrid: Von der Community – für die Community*, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 153-163.

F. Enders, C. Krause, R. Stotzka, D. Tonne und Ph. Vanscheidt: *Nach der Digitalisierung. Zur computergestützten Erschließung mittelalterlicher Handschriften*. In: *Digitale Rekonstruktionen mittelalterlicher Bibliotheken (Trierer Beiträge zu den Historischen Kulturwissenschaften, Band 12)*, hrsg. von S. Philippi und Ph. Vanscheidt, Wiesbaden 2014, S. 81-98.

F. Enders: *Methoden und Praktiken digitalen Edierens: Die ‚Darmstädter Musteredition‘ des ‚Meßtraktats‘ aus der Handschrift Trier Stb. 852/1311 4° in der virtuellen Forschungsumgebung TextGrid*. Masterarbeit an der Technischen Universität Darmstadt, 01.09.2014.

S. E. Funk, W. Pempe: *Vom Konzept zur Umsetzung - Einblicke in die Entstehung des TextGrid Repository*. In: *TextGrid: Von der Community – für die Community*, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 191-200.

S. E. Funk, U. Veentjer und T. Vitt: *Digitale Werkzeuge in den digitalen Geisteswissenschaften 277. Die Virtuelle Forschungsumgebung TextGrid - Status quo und neue Entwicklungen*. In: *Evolution der Informationsinfrastruktur Kooperation zwischen Bibliothek und Wissenschaft*, hrsg. v. H. Neuroth, N. Lossau, A. Rapp, Glückstadt, 2013, S. 277-300.

M. Göbel: *Das TextGrid Laboratory: Zehn Jahre Software-Entwicklung*. In: *TextGrid: Von der Community – für die Community*, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 251-258.

P. Gietz: *Authentifizierung und Autorisierung im Kontext Virtueller Forschungsumgebungen in den Geisteswissenschaften*. In: *TextGrid: Von der Community – für die Community*, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 77-90

N. Grupe, M. Kerzel: *TextGrid in Zahlen. NutzerInnen, MitarbeiterInnen, Finanzen*. In: *TextGrid: Von der Community – für die Community*, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 271-285.

- M. Hedges, H. Neuroth, K. M. Smith, T. Blanke, L. Romary, M. Küster, M. Illingworth: *TextGrid, TEXTvire, and DARIAH: Sustainability of Infrastructures for Textual Scholarship*. In: Journal of the Text Encoding Initiative (Online), Issue 5 (2013).
- M. de la Iglesia, N. Moretto, M. Brodhun: *Metadaten, LOD und der Mehrwert standardisierter und vernetzter Daten*. In: TextGrid: Von der Community – für die Community, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 91-102.
- B. Kaden, S. Rieger: *Usability in Forschungsinfrastrukturen für die Geisteswissenschaften. Erfahrungen aus TextGrid III*. In: TextGrid: Von der Community – für die Community, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 63-76
- M. Kerzel, M. Reich, H. Weber: *Die Edition „Johann Friedrich Blumenbach - online“ der Akademie der Wissenschaften zu Göttingen*. In: Evolution der Informationsinfrastruktur Kooperation zwischen Bibliothek und Wissenschaft, hrsg. v. Heike Neuroth, Norbert Lossau, Andrea Rapp, Glückstadt, 2013, S.107-136
- Th. Kollatz und S. Schmunk: *Datenvisualisierung. Geo-Browser und DigiVoy*. In: TextGrid: Von der Community – für die Community, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 165-180.
- H. Neuroth, A. Rapp, S. Söring (Hrsg.): *TextGrid: Von der Community - für die Community. Eine Virtuelle Forschungsumgebung für die Geisteswissenschaften*. Glückstadt 2015.
- H. Neuroth: *Das vielschichtige und komplexe Thema Nachhaltigkeit. Ein Interview mit H. Neuroth*. In: TextGrid: Von der Community – für die Community, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 241-249.
- G. Radecke, M. Göbel, S. Söring: *Theodor Fontanes Notizbücher. Genetisch-kritische und kommentierte Hybrid-Edition, erstellt mit der Virtuellen Forschungsumgebung TextGrid*. In: Evolution der Informationsinfrastruktur Kooperation zwischen Bibliothek und Wissenschaft, hrsg. v. Heike Neuroth, Norbert Lossau, Andrea Rapp, Glückstadt, 2013, S.85-196.
- O. Schmid, Ph. Vanscheidt, C. Krause, A. Rapp und S. Rettelbach: *Der Weg in die TextGrid-Community. Schulungen, Nutzertreffen, Support*. In: TextGrid: Von der Community – für die Community, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 287-299.
- S. Söring und M. Blümm: *Grenzgänge. An der Schnittstelle von Geisteswissenschaft, Informatik und Informationswissenschaft*. In: TextGrid: Von der Community – für die Community, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 159-270.
- U. Veentjer und M. W. Küster: *Publizieren mit TextGrid. SADE und XML-Print*. In: TextGrid: Von der Community – für die Community, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 211-227.
- W. Wegstein, A. Rapp und F. Jannidis: *TextGrid - Eine Geschichte*. In: TextGrid: Von der Community – für die Community, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 23-35.
- P. Wieder: *Hochverfügbarkeit und Performanz: Beiträge zum langfristigen Betrieb von TextGrid*. In: TextGrid: Von der Community – für die Community, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 183-190.

A. Witt, L. Eichinger und N. Fiedler: *Virtuelle Forschungsumgebungen nachhaltig organisieren - ein Erfahrungsbericht*. In: *TextGrid: Von der Community – für die Community*, hrsg. von H. Neuroth, A. Rapp und S. Söring, Glückstadt 2015, S. 301-314.

11.2 Geplante Veröffentlichungen

C. Krause und O. Schmid: *TextGrid in der Keltologie. Die virtuelle Forschungsumgebung als digitale Fundgrube für Archäologen und Epigraphiker*. Erscheint in: *Keltologische Kontroversen III*, hrsg. von H. A. Müller, 2016.

M. Bender: *Forschungsumgebungen in den Digital Humanities. Nutzerbedarf - Wissenstransfer - Textualität*. Reihe: Sprache und Wissen 22. Berlin, Boston: de Gruyter, 2016.

Annex

I. Partnerliste

- Berlin-Brandenburgische Akademie der Wissenschaften (BBAW) Berlin
Arbeitsschwerpunkt: Nutzerakzeptanz inkl. Befragungen (AP 6.2)
- DAASI International GmbH, Tübingen
Arbeitsschwerpunkt: Betrieb und Softwarepflege Repository: Performanz/Skalierbarkeit (AP 4.2); Exemplarische Durchführung einer Anpassung für ein Satellitenprojekt (AP 5.4)
- Hochschule Worms, Zentrum für Technologietransfer und Telekommunikation (ZTT)
Arbeitsschwerpunkt: Lasttests (AP 4.2), funktionale Tests (AP 5.2)
- Georg-August-Universität Göttingen, Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB) (Projektleitung)
Arbeitsschwerpunkt: Betrieb und Softwarepflege (TextGrid Repository, AP 4), Konsortialleitung und Öffentlichkeitsarbeit (AP 7)
- Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen (GWDG)
Arbeitsschwerpunkt: TextGrid Regelbetrieb und Qualitätsmanagement (AP 4.2, 4.3, 4.5)
- Institut für Deutsche Sprache (IDS), Mannheim
Arbeitsschwerpunkt: Betriebsmodell (AP 1)
- Julius-Maximilians-Universität Würzburg, Institut für deutsche Philologie
Arbeitsschwerpunkt: Betrieb und Softwarepflege Laboratory (AP 5)
- Max-Planck-Institut für Wissenschaftsgeschichte (MPI WG), Berlin
Arbeitsschwerpunkt: Usability-Untersuchungen mit besonderem Schwerpunkt auf der fachwissenschaftlichen Anwendung (AP 5.3), Prozessorganisation des Übergangs von Prototypen in den Dauerbetrieb (AP 2)
- Technische Universität Berlin, Zentrum Technik und Gesellschaft (ZTG)
Arbeitsschwerpunkt: Integriertes Monitoring und Qualitätssicherung (AP 6)
- Technische Universität Darmstadt – Institut für Sprach- und Literaturwissenschaft
Arbeitsschwerpunkt: Nutzer-Kommunikation (AP 3); Dokumentation (3.1), Coaching und Schulungsmaßnahmen (AP 3.3), Lab Support (AP 5.1)

II. Arbeitspakete und Deliverables

Alle Berichte zu den Deliverables sind auf der TextGrid-Homepage unter <https://textgrid.de/antrage-und-berichte> veröffentlicht und abrufbar.

AP 1 Betriebsmodell

Bezeichnung	Beschreibung
R 1.1.1	Identifikation von Kostenfaktoren und Erarbeitung von Kostenmodellen
R 1.1.2	Erfahrungsbericht Rechtsform: Praxisbewährung und Nutzeranforderungen
R 1.1.3	Finalisierung Kosten- und Finanzierungsmodelle
M 1.2.1	Geschäftsstelle Verein
R 1.3.1	Machbarkeitsstudien zu neuen technischen Entwicklungen für TextGridLab
M 1.3.2	Blaupause zur Integration zukunftsfähiger Entwicklungen

AP 2 Prozessorganisation

Bezeichnung	Beschreibung
M 2.1.2	Einrichtung eines Nutzerforums und dauerhafte Pflege
M 2.2.1	Entwicklerworkshop mit externen Projektpartnern in Kooperation mit AP 5.4
M 2.2.2	Bereitstellung des Market Space
R 2.2.2	(Laufende) Dokumentation von externen Plugins

AP 3 Nutzer-Kommunikation

Bezeichnung	Beschreibung
R 3.1.1	Einrichtung von Sandboxes, Entwicklung von use cases, Zusammenstellung von Demo-Materialien
R 3.1.2	Dokumentation für Anwender
R 3.3.1	Report über die Erfahrungen aus den Nutzerschulungen, Empfehlungen für die weitere Ausgestaltung des Konzepts
M 3.3.1	Nutzertreffen mit je spezifischen Schwerpunkten

AP 4 Betrieb und Softwarepflege Repository

Bezeichnung	Beschreibung
M 4.4.1	Update SADE (Dokumentation und Referenzinstanz)
R 4.4.1	Dokumentation/Leitfaden Import (technisch/inhaltlich)
R 4.4.2	SADE Dokumentation, Referenzinstanz
M 4.5.1	Architekturskizze eines produktiven Regelbetriebs
R 4.5.1	Architekturskizze inklusive Identifikation der Lücken und der Optimierungsmöglichkeiten
R 4.5.2	Update der in Monat 6 erstellten Architekturskizze, Potenzial der Optimierungen, erste Umsetzungen

AP 5 Betrieb und Softwarepflege Laboratory

Bezeichnung	Beschreibung
M 5.1.1	Technische Dokumentation
R 5.2.1	Laufende Funktionale Tests
M 5.2.1	Laufende funktionale Tests
R 5.3.1	Usability, Zusammenfassung der bisherigen Nutzerszenarien und Identifikation der Testgruppen
M 5.3.1	Usability, Zusammenfassung der bisherigen Nutzerszenarien und Identifikation der Testgruppen
M 5.4.1	Konzepte und Workflows zur Integration von Tools in das TextGridLab
R 5.5.1	Tiefenerschließung der Wörterbücher und Enzyklopädien
M 5.5.1	Vollständige Veröffentlichung der aufbereiteten Daten
R 5.5.2	Usability: Anforderungen an TextGrid aus der Arbeit mit den Testgruppen
R 5.5.3	Usability: Abschluss des Fallbeispiels und sich daraus ergebene Anforderungsanalyse
M 5.5.3	Usability, Abschluss des Fallbeispiels und sich daraus ergebende Anforderungsanalyse

AP 6 Integriertes Monitoring

Bezeichnung	Beschreibung
R 6.1.1	Report zur Analyse „Dimensionen des integrierten Monitoring einer digitalen Forschungsinfrastruktur“
R 6.1.2	Konzept Beurteilungsmethodik, Beurteilungskategorien und Governance Werkzeuge
M 6.1.3	Workshop mit Projektpartnern zum erarbeiteten Konzept des integrierten Monitoring
M 6.1.4	Implementierung des integrierten Monitoring, Report „Integriertes Monitoring einer digitalen Forschungsinfrastruktur“
R 6.1.5	Abschließender Bericht zur Implementierung inklusive Evaluierungen
R 6.2.1	Report „Ergebnisse der Nutzerbefragung“

AP 7 Projektmanagement und Öffentlichkeitsarbeit

Bezeichnung	Beschreibung
R 7.2.1	Konzept / Organisation Experten-, Strategiegespräche und TextGrid Summit etc.
R 7.2.2	Kooperation mit Fachgesellschaften
R 7.3.1	Öffentlichkeitsarbeit
M 7.3.1	Webauftritt TextGrid
R 7.3.2	Social-Media Konzept, Umsetzung &

	Betreuung
R 7.3.3	Marketingkonzept & Umsetzung