

Sprachressourcen in der Lehre – Erfahrungen, Einsatzszenarien, Nutzerwünsche

1 Einleitung

Im zweiten Teil dieses Themenheftes steht eine andere Konstellation im Mittelpunkt. Hier geht es vor allem um die Frage, wie die forschungsbezogene Nutzung von Sprachressourcen (Daten, Werkzeuge, Services) in der akademischen Lehre vermittelt werden kann und welche Rahmenbedingungen den Einsatz von Sprachressourcen in der Lehre erleichtern können.

Vor dem Hintergrund vielfältiger aktueller Bestrebungen, die Möglichkeiten der digitalen Geisteswissenschaften durch die Schaffung von Forschungsinfrastrukturen zu erweitern, stellt sich neben den anvisierten forschungsbezogenen Nutzungsszenarien digitaler Ressourcen auch die Frage, wie entsprechende Inhalte und Arbeitsweisen in die akademische Lehre integriert werden können. Vielerorts werden Sprachressourcen in Forschungsprojekten erarbeitet und genutzt. Neben der langfristigen Sicherung dieser Ressourcen in technischen Infrastrukturen ist deren nachhaltige Nutzung und Pflege im Interesse sowohl der Nutzer als auch der Anbieter. Im Sinne des Humboldt'schen Ideals der Einheit von Forschung und Lehre ist dabei insbesondere der Transfer von der Forschung in die Lehre und zurück von Interesse.

Den Lernenden bzw. den Studierenden erschließen sich die Möglichkeiten der Nutzung von Sprachressourcen und Werkzeugen, und sie werden auf lange Sicht zu deren Nutzern, die wiederum den Fortbestand der Ressourcen sichern. Eine Schlüsselrolle zwischen den Anbietern und den zukünftigen Nutzerinnen und Nutzern kommt dabei den Dozentinnen und Dozenten zu, die die Technik und den forschungsmethodischen Umgang mit den Ressourcen vermitteln. Neben der Arbeit mit akademischen Lehrwerken spielt dabei auch die unmittelbare Nutzung von Sprachressourcen eine zunehmend stärkere Rolle, sei es direkt in den Lehrveranstaltungen oder im Rahmen von Hausarbeiten oder studentischen Projekten. Der Einsatz von Sprachressourcen stellt dabei in mehrfacher Hinsicht eine Herausforderung dar, weist jedoch auch Parallelen zu anderen Bereichen auf, in denen computerbasierte Verfahren in Forschung und Lehre genutzt werden.

Vor diesem Hintergrund wurde im Rahmen des vom BMBF geförderten Projekts D-SPIN¹ (Deutsche Sprachressourcen-Infrastruktur) ein Workshop zu „Sprachressourcen in der Lehre“ veranstaltet. Aus einigen Vorträgen des Workshops sind nun Beiträge für dieses Themenheft hervorgegangen. Wir bedanken uns ausdrücklich bei den Autorinnen und Autoren², welche sich bereit erklärt haben, die jeweiligen Workshopbeiträge zu Artikeln für dieses Themenheft aufzubereiten.

¹ D-SPIN: Deutsche Sprachressourcen-Infrastruktur: <http://d-spin.org>

² Auf die doppelte Nennung von Personenbezeichnung in maskuliner und femininer Form wird im weiteren Verlauf des Artikels zugunsten der Lesbarkeit stellenweise verzichtet. Männliche Personen-, Berufs- und Rollenbezeichnungen schließen dabei entsprechende weibliche Bezeichnung mit ein.

Die entstandenen Beiträge sind in diesem Teil des Bandes so angeordnet, dass zwei überblicksartige Artikel als Rahmen für drei Beiträge zu spezielleren Themen dienen (Sprachressourcen in Szenarien der theoretischen und der Computerlinguistik, für die historische Semantik und in der Lehrerbildung).

In diesem dem zweiten Hefteil vorangestellten Artikel stellen wir zum einen die weiteren Beiträge zu diesem Themenschwerpunkt vor und präsentieren zum anderen noch einmal die Gesamtperspektive auf den zugrundeliegenden Workshop.

2 Vorstellung der Beiträge in diesem Teil des Themenheftes

Im ersten Beitrag, *Chancen und Probleme der Nutzung von Korpora, Taggern und anderen Sprachressourcen in Seminaren*, geht Heike Zinsmeister von didaktischen Überlegungen aus und diskutiert zunächst die etablierte Bühnen-Metaphorik zum Einsatz von Sprachressourcen in der Lehre, welche – so stellt die Autorin fest – in der Praxis oftmals um die didaktisch problematische Kategorie des „Schattenspiels“ erweitert werden muss. Indem sie anschließend existierende Evaluationskriterien für sprachverarbeitende Software an Korpusressourcen und den Lehrkontext adaptiert, entwickelt die Autorin eine umfangreiche Liste von Entscheidungskriterien für den Einsatz von Sprachressourcen in der Lehre. Kriterien wie z.B. „Sozialform“, „Technisches Knowhow“, „Lizenzierung“ und „Operabilität“ sind zu Dimensionen wie „Nutzer“ und „Verfügbarkeit“ gruppiert und in Form von Fragen, die ihre möglichen Ausprägungen spezifizieren, weiter ausgeführt. Im anschließenden Diskussionsteil werden diese möglichen Ausprägungen anhand von Beispielen des Einsatzes von konkreten Ressourcen in Lehrheiten von computerlinguistischen Seminaren der Autorin ausführlich illustriert, wobei auch unterschiedliche Ausprägungen für Studenten eines computerlinguistischen Studiengangs vs. eines sprachwissenschaftlichen Studiengangs eine Rolle spielen.

In dem Beitrag *Digitale Korpora in der Lehre – Anwendungsbeispiel aus der Theoretischen Linguistik und der Computerlinguistik* leitet Stefanie Dipper Anforderungen an Sprachressourcen aus vier Fallbeispielen, d.h. Lehrsituationen in den Fächern Theoretische Linguistik und Computerlinguistik her. Dabei arbeitet sie die unterschiedlichen, teilweise konträren Anforderungen dieser beiden Fächer heraus. So plädiert sie für die Bereitstellung von webbasierten Korpus-Abfragetools und Annotationswerkzeugen für die Lehre in der theoretischen Linguistik, in der die Bildung und Überprüfung von linguistischen Theorien (z.B. Regeln der Vorfeldbesetzung im Deutschen) anhand von Korpusbelegen im Vordergrund steht. Hingegen ist die Lehre in der Computerlinguistik, in der es häufig um die Vermittlung von Methoden der Anwendungs- oder Ressourcenentwicklung geht, auf herunterladbare und lokal installierbare Korpora und Werkzeuge angewiesen, um Daten flexibel vor- und weiterverarbeiten und Werkzeuge in eigene spezifische Verarbeitungsketten einbinden zu können (z.B. für eine Anwendung zur *authorship attribution*).

Die beiden Beiträge von Zinsmeister und Dipper verfolgen somit komplementäre Strategien, indem Zinsmeister top-down und deduktiv Kriterien für Sprachressourcen aus einem didaktischen Konzept und den EAGLES-Evaluationskriterien für Software entwickelt und diese dann auf reale Unterrichtssituationen anwendet. Dipper hingegen beginnt bottom-up und induktiv mit der Beschreibung von vier Unterrichtsszenarien, anhand derer sie abstra-

hierend die unterschiedlichen Anforderungen an Sprachressourcen ableitet und schließlich vergleicht.

Im dritten Beitrag mit dem Titel *Der eHumanities Desktop als Werkzeug in der historischen Semantik: Funktionsspektrum und Einsatzszenarien* schlagen Alexander Mehler et al. eine andere Richtung ein. Mit dem „eHumanities Desktop“ stellen Sie ein System vor, welches den integrierten, webbasierten Zugriff auf Sprachressourcen und korpuslinguistische Werkzeuge erlaubt, und illustrieren dessen Einsatzmöglichkeiten für das Forschungsgebiet der Historischen Semantik. Nachdem die Arbeit am „eHumanities Desktop“ vor dem Hintergrund der *Digital Humanities* verortet wird, liefern die Autoren zunächst eine kompakte Darstellung des Forschungsgebietes der Historischen Semantik, bevor sie das Funktionsspektrum des „eHumanities Desktop“ vorstellen und anschließend auf verschiedene Lehrkontexte eingehen, in denen das System bereits genutzt wurde. Die komplett webbasierte Bereitstellung des Systems erleichtert dabei den Einsatz in der fachwissenschaftlichen Forschung und Lehre, bei dem aus Sicht der Anwender ein Forschungsinteresse im Mittelpunkt steht und die technischen Hürden als solche zu minimieren sind. Erste Erfahrungen im Einsatz in der Lehre an der Universität Frankfurt schildert das Autorenteam anhand verschiedener Veranstaltungsformen, hier Seminare und Übungen sowie Studiengruppen und Anwenderworkshops. In der abschließenden Diskussion werden wiederum verschiedene Zielgruppen identifiziert, welche entsprechend unterschiedliche Anforderungen hinsichtlich der Weiterentwicklung des „eHumanities Desktop“ erkennen lassen.

Im vierten Beitrag *Digitale Sprachressourcen in Lehramtsstudiengängen: Kompetenzen – Erfahrungen – Desiderate* stellen Michael Beißwenger und Angelika Storrer ein weitreichendes Einsatzgebiet für Sprachressourcen dar. Vor dem Hintergrund gesamtgesellschaftlicher Entwicklungen wie etwa der generellen Zunahme internetbasierter Kommunikation schildern die Autoren, wie im Bereich der Lehrerbildung sowohl schulische Lehrpläne als auch Curricula für germanistische Lehramtsstudiengänge inzwischen diverse Anknüpfungspunkte enthalten, die die Arbeit mit digitalen Sprachressourcen in beiden Kontexten motivieren, wenn nicht gar erfordern. Die Autoren unterscheiden dabei zwei Typen didaktischer Konzepte zur Einbindung von Sprachressourcen in die Lehrerbildung und stellen die von ihnen entsprechend konzipierten und an der TU Dortmund angebotenen Lehrveranstaltungen vor. Beide Lehrveranstaltungen bzw. Konzepttypen werden detailliert erläutert, durch zugehörige Aufgabenbeispiele illustriert und daraus gewonnene Erfahrungen diskutiert. Die Autoren liefern somit wertvolle Anregungen für die Integration der Arbeit mit Sprachressourcen in sprachbezogene Lehramtsstudiengänge. Der Beitrag wird durch eine Auflistung der genutzten Materialien und eingesetzten Ressourcen abgerundet.

In dem letzten Beitrag des Bandes, *Korpuslinguistik in der linguistischen Lehre: Erfolge und Misserfolge* berichtet Noah Bubenhofer im ersten Teil von seinen Erfahrungen als Dozent an den Universitäten Zürich und Mannheim, an denen er Einführungen in die Korpuslinguistik und sprachwissenschaftliche Seminare durchführte. Zunächst wird das didaktische und inhaltliche Konzept der Einführungsveranstaltungen erläutert. Um das Potenzial von Korpuslinguistik in sprachwissenschaftlichen Seminaren zu illustrieren, werden verschiedene Seminar- und Abschlussarbeiten vorgestellt, in denen Forschungsfragen aus der Semantik, der Textlinguistik und der Diskurs- und Kulturanalyse mit korpuslinguistischen

Methoden angegangen wurden. Im zweiten Teil präsentiert Bubenhofer als Autor und Betreuer einer Online-Einführung in die Korpuslinguistik eine Klassifikation der Bedarfe der Nutzer seines Webangebots anhand der Zugriffsstatistiken. Das Fazit fasst die sich aus den beiden Teilen ergebenden Anforderungen an Werkzeuge und Softwaremodule für die Korpuserstellung und -analyse zusammen, die gewährleistet sein müssen, um Studierende mit philologisch-sprachwissenschaftlichem Hintergrund erfolgreich an das empirische Arbeiten heranzuführen.

Wie bereits erwähnt, haben alle fünf Beiträge in diesem Teil des Themenheftes ihren Ursprung in einem gleichnamigen Workshop, zu welchem wir im Folgenden einen Gesamtüberblick darstellen.

3 Hintergrund: Der D-SPIN Workshop zu „Sprachressourcen in der Lehre“

Im Rahmen des Projekts D-SPIN wurde ein Workshop zu *Sprachressourcen in der Lehre: Erfahrungen, Einsatzszenarien, Nutzerwünsche* von den Mitherausgebern Frank Binder, Henning Lobin und Harald Längen organisiert. Dieser fand am 18.1.2011 an der Berlin-Brandenburgischen Akademie der Wissenschaften statt. An dem Workshop nahmen ca. 30 Wissenschaftlerinnen und Wissenschaftler teil, zum einen Universitätsdozenten, die in ihrer Lehre Sprachressourcen einsetzen und zum anderen Mitarbeiter von Einrichtungen, die Sprachressourcen bereit stellen. In acht Vorträgen (siehe Ablaufschema in Abbildung 1) stellten Lehrende ihre Erfahrungen mit dem Einsatz von Sprachressourcen in der universitären Lehre in Studiengängen wie Computerlinguistik, Allgemeine Sprachwissenschaft, Geschichtswissenschaft und Literaturwissenschaft/Philologien dar und formulierten Anforderungen und Wünsche an die Bereitsteller von Ressourcen für die Zukunft

Im Vorfeld des Workshops waren bereits eine Liste von Leitfragen zum Thema bekannt gemacht worden (siehe Abbildung 2). In mehreren Diskussions-Slots ergab sich somit ein anregender Austausch zwischen Lehrenden und Ressourcen-Providern. Im Folgenden liefern wir zunächst eine Gesamtperspektive auf die im Workshop vertretenen Positionen, bevor wir einige Antwortvorschläge zu den Leitfragen wiedergeben.

3.1 Zusammenfassung der Aussagen des Workshops

Insgesamt wurde deutlich, dass Sprachressourcen (Daten, Werkzeuge und Services) in der akademischen Lehre in sehr verschiedenartigen Szenarien und Kontexten eingesetzt werden, da sie in unterschiedlichen curricularen Zusammenhängen eine Rolle spielen. Gleichzeitig sind einige übergeordnete Tendenzen erkennbar, welche wir hier mit einer kurzen Synopsis darstellen möchten. Im Anschluss an diese Darstellung sind einige Antwortvorschläge zu den Leitfragen des Workshops zusammengetragen, so wie sie im Kontext des Workshops sichtbar wurden. Materialien zu den Vorträgen der im Folgenden namentlich aufgeführten Vortragenden sind auf der Webseite zum Workshop verlinkt.³

³ Workshop „Sprachressourcen in der Lehre“:
<http://www.uni-giessen.de/~g91254/dspin-workshop-lehre/>

Ablauf des Workshops:

- Begrüßung und Eröffnung
- 1. Heike Zinsmeister (Universität Konstanz): Chancen und Probleme der Nutzung von Korpora, Taggern und anderen Sprachressourcen in sprachwissenschaftlichen Seminaren – ein Erfahrungsbericht
- 2. Noah Bubenhofer (IDS Mannheim): Korpuslinguistik in der linguistischen Lehre: Erfolge und Misserfolge
- 3. Werner Wegstein (Julius-Maximilians-Universität Würzburg): Sprachressourcen in der Lehre: Erfahrungen aus dem Fachgebiet 'Deutsche Sprachwissenschaft'
- 4. Alexander Mehler (Goethe-Universität Frankfurt): eHumanities Desktop und historische Semantik
- Moderierte Diskussion (35 min)
- 5. Cristina Vertan (Universität Hamburg): Sprachressourcen in der Lehre: Erfahrungen aus der historischen Korpuslinguistik
- 6. Stefanie Dipper (Ruhr-Universität Bochum): Elektronische Korpora in der Lehre - Anwendungsbeispiele aus der theoretischen und der Computerlinguistik
- 7. Michael Beißwenger / Angelika Storrer (Technische Universität Dortmund): Digitale Sprachressourcen in den Lehramtsstudiengängen: Kompetenzen - Erfahrungen – Desiderate
- 8. Sabine Bartsch (Technische Universität Darmstadt): Prototypen und Processing Chains: Werkzeuge und Kompetenzen für die linguistische Sprachverarbeitung
- Moderierte Diskussion (45 min)
- Schlussworte, Verabschiedung

Abbildung 1: Vorträge und Ablauf des Workshops „Sprachressourcen in der Lehre“ am 18.01.2011 in Berlin

Betrachtet man zunächst die zugrunde liegenden Daten, also Text- oder Sprachmaterial, so wird deutlich, dass für viele fachwissenschaftliche Fragestellungen – im Unterschied zu methodischen oder verfahrenswissenschaftlichen Fragestellungen – sehr spezifisches Datenmaterial benötigt wird (vgl. Beitrag von Mehler) oder auf renommierte Editionen ganz bestimmter Texte zurückgegriffen werden muss (Wegstein). Große Mengen allgemeinsprachlicher Daten hingegen, wie sie etwa in der theoretischen Linguistik (Dipper), Computerlinguistik (Dipper, Zinsmeister) und Korpuslinguistik (Bubenhofer, Zinsmeister) eine Rolle spielen, dienen einerseits der Entwicklung sprachtechnologischer oder computerlinguistischer Werkzeuge, andererseits vor allem als Vergleichsgrundlage in differenzanalytischen Verfahren. Es besteht also generell Bedarf an standardkonformen und gleichzeitig möglichst flexibel einsetzbaren Referenzressourcen (z.B. Korpora), insbesondere auch für dezidierte Sprachen oder Sprachvarianten. Generell scheint gesprochene Sprache im Rahmen dieses Workshops noch unterrepräsentiert zu sein, ebenso wie multimodale Kombinationen aus Text, Ton, Bild und Video. Es bleibt daher noch zu klären, inwieweit sich die gewonnenen Erkenntnisse aus den Bereichen der primär text-bezogenen Sprachressourcen auch auf diese Bereiche übertragen lassen.

Hinsichtlich der eingesetzten Werkzeuge und Services weist der Workshop in zwei zunächst konträre, aber gleichermaßen interessante Richtungen. Beiden Richtungen ist gemein, dass jeweils Portfolios von spezialisierten Werkzeugen eingesetzt und diese zur Anreicherung

rung von Textmaterial meist zu linguistischen Verarbeitungsketten kombiniert werden (Bartsch, Bubenhofer, Dipper, Mehler, Zinsmeister). Einerseits werden Werkzeuge eingesetzt, die man zwar lokal auf Arbeitsrechnern installieren und administrieren muss (Bartsch, Dipper, Zinsmeister), dafür aber im idealen Falle der Verfügbarkeit als *Open Source*-Software auch selbst modifizieren und neu kombinieren kann (Bartsch). Wenn man im Rahmen von technischen Studiengängen oder Lehrveranstaltungen auch die Entwicklung derartiger Werkzeuge beherrschen will, ist diese Art der Verfügbarkeit bisher die optimale. Auch bietet sich ein solches Ökosystem aus quelloffenen Werkzeugen ideal für interdisziplinäre Team-Arbeit unter Studierenden an (Bartsch, Bubenhofer). Ein anderer Ansatz scheint für technikfernere Fachwissenschaften (wie z.B. Vertan, Wegstein, Dipper) vielversprechender: Webbasierte Anwendungsumgebungen (wie z.B. Mehler), in denen die integrierten Werkzeuge auch ohne technisches Know-How genutzt und darüber hinaus perspektivisch auch Dienste zum kollaborativen Arbeiten sowie zur Sicherung und Versionierung von Daten und Werkzeugen angeboten werden können. Beide Ansätze erachten wir als wertvoll, da einerseits sowohl künftige Werkzeugentwickler gewonnen und trainiert werden müssen, andererseits die Früchte ihrer Arbeit so vielen Anwendern wie möglich zur Verfügung stehen sollen. Inwieweit sich beide Ansätze kombinieren lassen, d.h. inwieweit die webbasierte Arbeitsumgebung gleichermaßen für Nutzer und Entwickler der darin bereitgestellten Werkzeuge interessant ist – etwa indem sie neben intuitiver Bedienung und Kombinierbarkeit auch die Modifikation von Werkzeugen erlaubt und die dafür nötige Verwaltung von referenzierbaren Versionen und Varianten integriert –, bleibt zunächst eine offene Frage.

Neben den bereits angesprochenen Nutzergruppen – textbezogen arbeitende Fachwissenschaftler auf der einen Seite und (zukünftige) Entwickler von computerlinguistischen und sprachtechnologischen Werkzeugen auf der anderen Seite – muss eine dritte Nutzergruppe ebenfalls bedacht werden: Texte und Sprache sind – gemäß ihrer grundlegenden Bedeutung in vielen Berufsfeldern – zentraler Lehrgegenstand in einer Reihe von Schulfächern. Dementsprechend sollten Lehrerinnen und Lehrer im Umgang mit Sprachressourcen geschult, ihre Nutzung in Lehrplänen verankert und die Vermittlung der Nutzungsmöglichkeiten in Curricula für Lehramtsstudiengänge integriert werden (Storror & Beißwenger). Weiterhin ist anzumerken, dass die Integration der Arbeit mit Sprachressourcen in Curricula auch allgemeine berufsrelevante Fähigkeiten schulen kann, etwa in den Bereichen der Erhebung-, Verwaltung und Auswertung verschiedener Formen von Daten. Dies ist insbesondere der Fall, wenn für diese Zwecke auch die Nutzung weit verbreiteter Standardsoftware (z.B. Techniken der Tabellenkalkulation oder Datenbank-Nutzung) oder grundlegende Techniken der Text- und Sprachtechnologie am Praxisbeispiel demonstriert und vermittelt werden.

3.2 Leitfragen und Antworten des Workshops

Der Workshop orientierte sich an einer zuvor den Vortragenden zur Verfügung gestellten Liste von Leitfragen (s. Abb. 2), die im folgenden in kondensierter, nicht abschließender Form auf der Grundlage der Workshop-Beiträge beantwortet werden sollen:

Leitfragen des Workshops:

1. Welche Themen und Sprachressourcen sind für welche Zielgruppen in der Lehre relevant?
2. Wie ist die Nutzung von Sprachressourcen heute in Curricula integriert?
3. Welche Ressourcen werden bevorzugt und gern praktisch eingesetzt? Aus welchen Gründen? In welchen Szenarien? Lassen sich daraus "best practices" hinsichtlich der Bereitstellung von Ressourcen ableiten?
4. Welche Wünsche und Vorschläge zur Verbesserung und Erleichterung des didaktischen Einsatzes von Sprachressourcen gibt es seitens der Lehrenden?
5. Welche möglichen Hürden und Herausforderungen aus Sicht der Anbieter treten dabei auf?
6. Welche Rahmenbedingungen sind nötig, um den Transfer zwischen Forschung und Lehre im Bereich der Sprachressourcen zu fördern?

Abbildung 2: Leitfragen zum Workshop „Sprachressourcen in der Lehre“

1. *Welche Themen und Sprachressourcen sind für welche Zielgruppen in der Lehre relevant?*

Für einen Einblick in die Vielfalt der Themen und Zielgruppen verweisen wir auf die verlinkten Präsentationen auf der Webseite des Workshops. Die im Workshop repräsentierten Zielgruppen erstrecken sich von Lehramtsstudierenden für sprachliche Fächer über Studierende der Philologien, Geschichtswissenschaften und Sprachwissenschaften (inkl. Computerlinguistik, theoretische Linguistik, praktische Linguistik, d.h. z.B. Feldforschung und Korpuslinguistik) bis in Bereiche der Informatik.

2. *Wie ist die Nutzung von Sprachressourcen heute in Curricula integriert?*

Sie ist in den verschiedenen Disziplinen und Wissenschaftsbereichen auf unterschiedliche Weisen integriert. In einigen Bereichen sind „Endnutzer“ an einsatzbereiten Sprachressourcen interessiert, in anderen Bereichen werden der Entwicklungsvorgang und dessen handwerkliche Seite oder dessen theoretische Grundlagen in den Vordergrund gestellt. In wieder anderen Bereichen verschränken sich diese Perspektiven – zum Beispiel in interdisziplinären bzw. fächerübergreifenden studentischen Projekten.

3. *Welche Ressourcen werden bevorzugt und gern praktisch eingesetzt? Aus welchen Gründen? In welchen Szenarien? Lassen sich daraus "best practices" hinsichtlich der Bereitstellung von Ressourcen ableiten? Antworten bezogen auf die beiden grundlegenden Arten von Ressourcen:*

a. Daten

- i. Teilweise ist sehr spezifisches Datenmaterial wie z.B. ganz bestimmte Texte oder auch renommierte Editionen von Texten von Interesse.
- ii. Referenzressourcen werden sowohl für die „Allgemeinsprache“ als auch für dezidierte Sprachen oder Sprachvarianten als Vergleichsgrundlage für differenzanalytische Verfahren benötigt.

- b. Werkzeuge
- i. Portfolios spezialisierter Werkzeuge, verknüpft zu linguistischen Verarbeitungsketten, werden disziplinübergreifend genutzt.
 - ii. Konkret wurde der Wunsch nach einer expliziten Verknüpfung von Annotationswerkzeugen mit Recherche-Tools für die Abfrage (Querying) von annotierten Daten deutlich.
 - iii. Modifizierbarkeit, Anpassbarkeit, Transparenz und Flexibilität sollten gegeben sein, z.B. in einem Ökosystem quelloffener Werkzeuge.
 - iv. Weiterhin sollten diese Werkzeuge als integrierte Komponenten einer nutzerfreundlichen webbasierten Arbeitsumgebung bereitgestellt werden.
 - v. Die Anknüpfung der Werkzeuge an Standardsoftware zur Erhebung, Verwaltung und Auswertung von Daten (z.B. Tabellenkalkulation, Statistik-Software, Datenbanken) bietet neben der Vermittlung text- und sprachtechnologischer Verfahren und Techniken eine breite Anknüpfungsmöglichkeit in Richtung allgemeiner berufsrelevanter Kompetenzbildung.
4. *Welche Wünsche und Vorschläge zur Verbesserung und Erleichterung des didaktischen Einsatzes von Sprachressourcen gibt es seitens der Lehrenden?*
- a. Daten sollten standardkonform in Formaten vorliegen, die von den diversen Werkzeugen ohne eigene Vorarbeit verarbeitet werden können.
 - b. Werkzeuge: An dieser Stelle verweisen wir auf die Vorschläge aus 3.b.ii bis 3.b.iv.
 - c. Mehrmals wurde der Wunsch nach einer besseren Dokumentation von Ressourcen benannt, idealerweise auch im Hinblick auf den Einsatz in Lehrsituationen, z.B. durch Tutorials oder E-Learning-Einheiten oder auch durch eine Sammlung von Referenzen auf Studien, in denen die Ressource schon verwendet wurde.
 - d. Es gibt einen Bedarf an institutionalisierten Dienstleistungen für Ressourcennutzer wie z.B. Help-Desk-Services und die Bildung von Expertennetzwerken.
5. *Welche möglichen Hürden und Herausforderungen aus Sicht der Anbieter treten dabei auf?*
- Im Kontext von D-SPIN sind diese folgenden Herausforderungen sichtbar:
- a. Es braucht Zeit, Geld, Geduld und Engagement sowie ein gutes Erwartungsmanagement.
 - b. Die teils aufwändige Klärung rechtlicher Fragestellungen, gegebenenfalls verbunden mit nötigen Maßnahmen des Interessensausgleichs, sowie bestehende Vereinbarungen mit den „Spendern“ von linguistischem Datenmaterial stehen ambitionierten Wünschen teils schlicht entgegen. Die Änderung von Rahmenbedingungen sind nur in engen Grenzen möglich, und

die Achtung bestehender vertraglicher Vereinbarungen ist für die Vertrauensbildung und Reputation essenziell.

- c. Die technische Unterstützung rechtlich sicherer und ethisch unbedenklicher Bereitstellungsmöglichkeiten von Sprachressourcen ist ein Feld aktueller und künftiger Forschung und Entwicklung – daher verweisen wir hier noch einmal auf Punkt 5.a
6. *Welche Rahmenbedingungen sind nötig, um den Transfer zwischen Forschung und Lehre im Bereich der Sprachressourcen zu fördern?*

Zwei Anregungen traten zutage:

- a. Die Integration der Nutzung von Sprachressourcen in Curricula und Lehrpläne sollte vorangetrieben werden.
- b. Implizit: Technisch-administrative und Community-bildende Arbeit im Bereich der Aufbereitung und Verfügbarmachung von Sprachressourcen sollten weiterhin anerkannt und gefördert werden.

3.3 Training und Ausbildung im Kontext von CLARIN/D-SPIN

Der in diesem Beitrag dargestellte Workshop zu „Sprachressourcen in der Lehre“ ist Teil einer Reihe von Aktivitäten im Bereich Training und Ausbildung im Kontext der Forschungsinfrastrukturprojekte CLARIN⁴/D-SPIN während ihrer „Preparatory Phase“ 2008-2011 (s. ESFRI 2011). Diese reichen von initialen Orientierungaktivitäten über verschiedene durchgeführte Lehrveranstaltungen bis hin zur Bereitstellung von Tutorials.

In einer ersten Phase wurden innerhalb der Projektkonsortien auf nationaler und europäischer Ebene Ideen und Strategien für die Arbeit in den einzelnen Bereichen zusammengetragen und diskutiert, so auch für den Bereich Training und Ausbildung (siehe etwa BINDER & CRISTEA 2009). Strategische Leitfragen zu den Aktivitäten in diesem Bereich wurden in dieser Zeit gesammelt und im Projektverlauf dokumentiert (AHLBORN & BINDER 2010). So wurde deutlich, dass man vielfältige Angebote benötigt und zwischen regionalen, nationalen und internationalen Angeboten sowie zwischen universitären und außercurricularen Angeboten für verschiedene Qualifikationsstufen unterscheiden muss, zusätzlich zur jeweiligen thematischen Schwerpunktsetzung. Weiterhin besteht Bedarf an Training und Ausbildung sowohl auf Seiten der teils stark spezialisierten Nutzer als auch auf Seiten der Entwicklerteams und Anbieter von Sprachressourcen. Ebenfalls in dieser Phase wurde durch den „Sprachressourcen-Gipfel“ in Mannheim 2009 frühzeitig Kontakt zu verschiedenen Fachcommunities aufgebaut und die Vielfalt der vorhandenen Ressourcen und Interessensgebiete verdeutlicht.

In der folgenden Zeit verfolgte das D-SPIN Arbeitspaket 6 „Training und Ausbildung“ vor allem eine nutzerorientierte Strategie. Neben der Integration des Themengebietes in lokale universitäre Curricula geisteswissenschaftlicher Studiengänge war vor allem die Ausrichtung der D-SPIN Sommerschule 2010 ein Höhepunkt in dieser Phase des Projekts (BINDER ET AL. 2010).

⁴ CLARIN: Common Language Resources and Technology Infrastructure: <http://www.clarin.eu>

Für die Zielgruppe der Entwickler und Anbieter von Sprachressourcen wurden durch die Projektpartner parallel weitere Angebote geschaffen. So werden derzeit und auch in Zukunft wiederholt Workshops und Trainingskurse zu technischen Themen angeboten – einmal im Zuge des Überganges zur neuen Projektphase in CLARIN-D, aber auch im Rahmen von CLARA⁵ – einem europäischen „Initial Training Network“, welches thematisch und institutionell eng mit CLARIN verknüpft ist. Für WebLicht, der in D-SPIN gemeinsam von Teams aus Tübingen, Leipzig, Berlin, Stuttgart und Mannheim entwickelten Web-Anwendung für linguistische Verarbeitungsketten (HINRICHS ET AL. 2009, HINRICHS ET AL. 2010), stehen auf der Webseite verschiedene Tutorials unter anderem zur Integration eigener linguistischer WebServices bereit.⁶

Spätestens auf der gemeinsam von CLARIN und DARIAH⁷ ausgerichteten Konferenz „Supporting the Digital Humanities“ (SDH) im Oktober 2010 in Wien wurde deutlich, dass nutzerorientierte Ausbildung und Lehre durch die beteiligten Partner der entstehenden Forschungsinfrastrukturen ein bedeutender Erfolgsfaktor sind und auch zukünftig gemeinsame Anstrengungen erfordern.

4 Diskussion und Ausblick

Mit dem Übergang von der „Preparatory Phase“ zur „Implementation Phase“⁸ des europäischen Projekts CLARIN endet auf nationaler Ebene das Projekt D-SPIN, vor dessen Hintergrund der Workshop zu „Sprachressourcen in der Lehre“ konzipiert wurde, und das unmittelbar anschließende Verbundprojekt CLARIN-D nimmt seine Arbeit auf. Die Fragen nach Möglichkeiten einer besseren Integration der Nutzung von elektronischen Sprachressourcen in die akademische Lehre bleiben damit auf der Agenda.

Mit der Konzeption des Workshops zu „Sprachressourcen in der Lehre“ war es beabsichtigt, den Status Quo zur Nutzung von Sprachressourcen in der Lehre zu erkunden, von dem aus es nun gilt, die nächsten Schritte zu identifizieren und anzugehen. Im Idealfall werden so Lücken im Vorhandenen identifiziert, deren Schließung eine Erleichterung für die Nutzung von Sprachressourcen darstellt. Wenn also etwa etablierte Werkzeuge eine an aktuelle Betriebssysteme angepasste Benutzungsschnittstelle – sei diese nun webbasiert oder nicht – erhalten können, andere Ressourcen unter neuen, weniger restriktiven Lizenzen bereitgestellt werden können oder an verschiedenen Stellen die Zugänglichkeit zu und Interoperabilität zwischen Werkzeugen und Daten verbessert werden kann, dann sind dies zweifelsohne hilfreiche Entwicklungen. Dass derartige Anforderungen insbesondere auch aus den Bereichen der akademischen Lehre kommen können, zeigt deren Relevanz als Einsatzfeld von Sprachressourcen. Gleichfalls bietet das Konzept eines solchen Workshops stets Raum zur Weiterentwicklung, den wir im Folgenden ein wenig skizzieren wollen.

Der Fokus des Workshops und seiner hier veröffentlichten Beiträge liegt stark auf textuellen Ressourcen. Wenngleich die sich dahinter verbergende Vielfalt nicht unterschätzt

⁵ CLARA: Common Language Resources and their Applications – a Marie Curie ITN. <https://clara.uib.no/>

⁶ WebLicht Tutorials: <http://weblight.sfs.uni-tuebingen.de/englisch/tutorials/html/index.html>

⁷ DARIAH: Digital Research Infrastructure for the Arts and Humanities: <http://dariah.eu/>

⁸ Die Bezeichnungen der einzelnen Konstruktionsphasen europäischer Forschungsinfrastrukturen wurden wie in ESFRI (2011) verwendet.

werden sollte, scheint es doch geboten, diesen Fokus auszuweiten oder in ähnlichen Formaten multimodale Ressourcen in den Vordergrund zu stellen. Dabei wären sowohl Kombinationen aus Texten und Bildern zu betrachten als auch der Umgang mit Sammlungen von Audio- und Video-Daten, welche zum Zwecke wissenschaftlicher Analysen üblicherweise textuell transkribiert und anschließend mit weiteren Annotationen versehen werden. Auch für derartige Ressourcen bieten die aktuellen Infrastrukturentwicklungen neue Perspektiven, die es in die akademische (Methoden-)Lehre zu integrieren gilt.

Ebenfalls deutlich wird eine sprachwissenschaftliche Gewichtung der Beiträge im Workshop und in diesem Themenheft, während CLARIN/D-SPIN einen größeren Kreis potentieller Nutzer erreichen will. Über die Zusammenarbeit mit diversen Fachcommunities soll diese Fokussierung im Verlauf der derzeit beginnenden Konstruktionsphase von CLARIN-D ausgeweitet werden. Parallel dazu ist es ebenso wichtig, neben der Ausrichtung auf fachwissenschaftliche Nutzer, auch für Entwickler und Anbieter von Sprachressourcen entsprechenden Themen und Lehrveranstaltungen in technisch orientierte Studiengänge zu integrieren bzw. vorhandene Angebote zu pflegen.

Die Konzeption des Workshops ist weiterhin gezeichnet durch seine Infrastrukturperspektive, aus der die Lehre als ein Einsatzgebiet von Sprachressourcen erscheint. In dieser Konstellation denkt man leicht eher in Begriffen wie „Einsatzszenarien“, als sich etwa über „Lernziele“ Gedanken zu machen. Dabei ist die Didaktisierung eine Herausforderung an sich. Aus Sicht der Hochschullehrerinnen und Hochschullehrer ist der didaktische Nutzen bzw. Mehrwert, welcher mit dem Einsatz von Sprachressourcen verbunden sein kann, nicht als selbstverständlich anzusehen. Erfahrungen aus anderen Bereichen, etwa der softwaregestützten Statistik in empirischen Geistes- und Sozialwissenschaften oder des computergestützten Fremdsprachenlernens, zeigen, dass man im Spannungsfeld zwischen Computereinsatz und didaktischen Zielen gleichermaßen Erfolge und Kontroversen finden kann. Es sollte sich daher lohnen, die Entwicklungen um den Einsatz computerbasierter Sprachressourcen in fachwissenschaftlichen Lehrkontexten auch unter dem Gesichtspunkt der Didaktisierung weiter zu verfolgen. Diesbezüglich ist auch anzumerken, dass im Rahmen des Workshops nicht explizit nach vorhandenen und eingesetzten Lehrwerken – etwa zur Methodenlehre – gefragt wurde, wenngleich natürlich einzelne Beiträge solche Werke benennen. Hier könnte man ansetzen und noch einmal systematisch das derzeitige Angebot betrachten und gegebenenfalls entsprechenden Bedarf ermitteln.

Der Workshop war bewusst auf die deutschsprachige Fachcommunity ausgerichtet. Man kann ihn dadurch einordnen in umfassendere Bestrebungen mit dem Ziel, Inhalte der *Digital Humanities* in die Curricula deutscher Hochschulen einzubinden und diese Entwicklungen sichtbar zu machen (SAHLE 2010). Natürlich gibt es im vorliegenden Themenfeld auch im internationalen Kontext Aktivitäten, durch die der Austausch von Szenarien, Erfahrungsberichten und Empfehlungen zur akademischen Lehre in den *Digital Humanities* innerhalb der Fachcommunities gefördert werden soll (s. etwa HIRSCH 2010 oder TOMASEK & DAVIS 2011 sowie die darin enthaltenen Verweise). Szenarien, die Lehre und Forschung etwa im Bereich kollaborativer Annotation von Korpusdaten eng verzahnen – zum Beispiel realisiert als „'Classroom' production method“ (BAMMAN & CRANE 2010) –, werden erst durch geeignete Korpusinfrastrukturen ermöglicht. Hier geht es also nicht nur darum, bisherige Lehr- und

Lernformen effizient zu unterstützen, sondern auch neue Formate in der Lehre zu ermöglichen. Derartige Weiterentwicklungen zählen sicherlich zu den wichtigsten Impulsen, die geisteswissenschaftliche Forschungsinfrastrukturen den eHumanities geben können.

Dabei wird CLARIN nicht all diese Gebiete aus sich selbst heraus bedienen können. Vielmehr sind einzelne Fachcommunities im Rahmen der Kooperation mit CLARIN-D, DARIAH-DE und ähnlichen Initiativen angehalten, kompatible Lösungen unter Nutzung der entstehenden Infrastruktur selbst zu gestalten und zu erschaffen. Dies schließt auch den Transfer in die Lehre ein, sei es durch erhöhte Sichtbarkeit für einschlägige studentische Projekte, die Erstellung von praktischem Übungsmaterial oder Beiträgen zu Methodenlehrwerken oder durch regelmäßige Workshops und Schulungen für Forschende und Lehrende zu den Einsatzmöglichkeiten und –erfahrungen der neuen elektronischen Werkzeuge in speziellen fachwissenschaftlichen Kontexten.

Danksagungen

Konzeption, Organisation, Durchführung und Nachbereitung des Workshops „Sprachressourcen in der Lehre“ am 18.01.2011 an der Berlin-Brandenburgischen Akademie der Wissenschaften (BBAW) wurden ermöglicht durch die Förderung des Bundesministeriums für Bildung und Forschung (BMBF) im Rahmen des Projekts D-SPIN, ergänzt durch Förderungen des Hessischen Ministerium für Wissenschaft und Kunst (HMWK) im Rahmen des LOEWE-Schwerpunkts „Kulturtechniken und ihre Medialisierung“⁹.

Für das sorgfältige Reviewing der Beiträge in diesem Teil des Themenheftes danken wir Volker Boehlke, Gertrud Faaß, Gerhard Heyer, Marc Kupietz, Lothar Lemnitzer und Elke Teich. Weiterhin danken wir Heike Zinsmeister für Kommentare zu einer früheren Version des Abschnitts „Leitfragen und Antworten des Workshops“. Besonderer Dank gilt allen Vortragenden und Teilnehmenden des Workshops für die rege Beteiligung und die interessanten Diskussionen.

Literatur

- AHLBORN, SVETLANA & FRANK BINDER. (2010). D-SPIN Report R6.1: Training materials and guidelines for their applications. http://weblicht.sfs.uni-tuebingen.de/Reports/D-SPIN_R6.1.pdf
- BAMMAN, DAVID & GREGORY CRANE (2010). Corpus Linguistics, Treebanks and the Reinvention of Philology. In Fähnrich, K.-P. and Franczyk, B (Hrsg.) Proceedings of INFORMATIK 2010: Service Science, September 27 - October 01, 2010, Leipzig, volume 2 of Lecture Notes in Informatics, pages 542-551. GI.
- BINDER, FRANK & DAN CRISTEA (2009). Training and dissemination in CLARIN – the potential to bridge gaps. In: CLARIN Newsletter #6 June 2009. <http://www.clarin.eu/newsletter>
- BINDER, FRANK, SVETLANA AHLBORN, JOST GIPPERT AND HENNING LOBIN (2010) Connecting with User Communities: D-SPIN Summer School 2010. Poster at SDH+NEERI 2010. Vienna, Austria. <http://www.dspin-sommerschule.de/>

⁹ LOEWE-Schwerpunkt „Kulturtechniken und ihre Medialisierung“: <http://www.kulturtechniken.info/>

- ESFRI (EUROPEAN STRATEGY FORUM ON RESEARCH INFRASTRUCTURES) (2011). Strategy Report on Research Infrastructures - Roadmap 2010. Luxembourg: Publications Office of the European Union. <http://dx.doi.org/10.2777/23127>
- HINRICHS, ERHARD, MARIE HINRICHS, THOMAS ZASTROW, GERHARD HEYER, VOLKER BOEHLKE, UWE QUASTHOFF, HELMUT SCHMID, ULRICH HEID, FABIENNE FRITZINGER, ALEXANDER SIEBERT AND JÖRG DIDAKOWSKI (2009). WebLicht - Web-based LRT services for German. In: Elke Teich, Andreas Witt and Peter Frankhauser (Eds.) GSCL Workshop: Linguistic Processing Pipelines – Book of Abstracts (pp. 11-14)
- HINRICHS, MARIE, THOMAS ZASTROW AND ERHARD HINRICHS (2010). WebLicht: Web-based LRT Services in a Distributed eScience Infrastructure. In: Nicoletta Calzolari, Khalid Choukri, Bente Maegaard, Joseph Mariani, Jan Odijk, Stelios Piperidis, Mike Rosner and Daniel Tapias (Eds.) Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC 2010)
- HIRSCH, BRETT D. & MEAGAN TIMNEY (2010). The Importance of Pedagogy: Towards a Companion to Teaching Digital Humanities. Poster at Digital Humanities 2010. London, UK. <http://dh2010.cch.kcl.ac.uk/academic-programme/abstracts/papers/html/ab-853.html>
- SAHLE, PATRICK (2010). Digitale Geisteswissenschaften: Studieren! In: Silke Schomburg, Claus Leggewie, Henning Lobin und Cornelius Puschmann (Hrsg.) Beiträge der Tagung „Digitale Wissenschaft - Stand und Entwicklung digital vernetzter Forschung in Deutschland“. (S. 51-56) http://www.hbz-nrw.de/dokumentencenter/veroeffentlichungen/Tagung_Digitale_Wissenschaft.pdf
- TOMASEK, KATHRYN & REBECCA FROST DAVIS (2011). Integrating Digital Humanities Projects into the Undergraduate Curriculum. Pre-Conference Workshop at Digital Humanities 2011. Stanford, CA. https://dh2011.stanford.edu/?page_id=499