Biodiversitätsforschung und Linguistik – die DFG fördert zwei neue Fachinformationsdienste

Die Universitätsbibliothek wirbt 2,2 Millionen Euro Fördermittel ein

Die Universitätsbibliothek Johann Christian Senckenberg betreut zahlreiche Spezialsammlungen, die eine lange Geschichte aufweisen und überregional Bedeutung haben. Für den Ausbau der Sammlungen und die Transformation von Sammlungsinhalten in das Zeitalter von Semantic Web und Big Data spielen die Fachinformationsdienste (FID) eine große Rolle; sie werden von der Deutschen Forschungsgemeinschaft (DFG) in einem speziellen Programm gefördert. Betreute die Universitätsbibliothek zuletzt bereits vier Fachinformationsdienste (Afrikastudien, Allgemeine und Vergleichende Literaturwissenschaft, Jüdische Studien, Darstellende Kunst), so sind mit Jahresbeginn 2017 zwei weitere Projekte hinzugetreten, die hier vorgestellt werden.

FID Biodiversitätsforschung

Wie kann historische und aktuelle Literatur zur Biodiversität in zeitgemäßen Formaten für die Forschung zur Verfügung gestellt werden? Im Hinblick auf diese Herausforderung haben drei Antragsteller unter Federführung der Universitätsbibliothek ein Konzept erstellt. Beteiligt sind auch die von Prof. Dr. Mehler geleitete AG Texttechnologie am Institut für Informatik der Goethe-Universität sowie die Senckenberg Gesellschaft für Naturforschung. Ein fruchtbarer Austausch mit der Fachcommunity hat zu einem FID-Konzept beigetragen, das auch die DFG überzeugt hat.

Die Erforschung der Biodiversität hat jüngst einen großen Aufschwung genommen. Sie wird zunehmend zu einer von enormen Datenmengen getriebenen Forschungsrichtung. Für ihre weitere Entwicklung sind jedoch nicht nur neuere Fachliteratur oder neueste Forschungsdaten relevant, sondern gerade auch die ältere Literatur und die darin gespeicherten Daten und Erkenntnisse. Erst der Vergleich früherer und heutiger Zustände ermöglicht es beispielsweise, Veränderungen der Verbreitung von Organismen zu rekonstruieren; dies ist essenziell, um den Einfluss von Umweltveränderungen auf Biodiversität und Ökosystemleistungen verstehen zu können. So kommt gerade bei ökologischen Fragestellungen die Rolle der Bibliothek als Wissensspeicher zum Tragen.

Wesentlicher Schwerpunkt des FID Biodiversitätsforschung ist ein Pilotvorhaben im Bereich Text-Mining. Automatisierte Verfahren erschließen und strukturieren dabei die Inhalte von Literatur. Damit wird eine Mobilisierung von Daten zur Biodiversität erreicht; das heißt, die aufbereiteten Daten aus der Literatur sind zukünftig sowohl leichter auffindbar als auch besser in die Strukturen der Biodiversitätsinformatik integrierbar.

Ein zweiter Schwerpunkt ist die Digitalisierung von Literatur des 20. Jahrhunderts. Im Gegensatz zu (noch) älterer Biodiversitätsliteratur ist gerade in diesem Segment bislang wegen urheberrechtlicher Probleme eine elektronische Verfügbarkeit oft nicht gegeben. Die Digitalisierung dient einerseits der Generierung des notwendigen Textkorpus für das Text-Mining; andererseits bildet sie die Basis für ein zu etablierendes Angebot für Open-Access-Zeitschriften.

FID Linguistik

Der FID Linguistik verfolgt das Ziel, eine leistungsfähige, überregional zugängliche Forschungsplattform für die Allgemeine Linguistik, die Allgemeine und Vergleichende Sprachwissenschaft und die Linguistik der einzelphilologischen Sprachwissenschaften zu etablieren.

Als erster Baustein dient dabei das Linguistik-Portal, ein internetbasiertes Fachportal, welches konventionelle gedruckte Werke gemeinsam mit digitalen Veröffentlichungen unter einer einheitlichen Oberfläche präsentiert, durch standardisierte Metadaten erfasst und im Idealfall direkt zur Verfügung stellt. Die Hauptmodule des Linguistik-Portals, die in den Jahren 2012 bis 2014 und 2015 bis 2017 mit DFG-Förderung aufgebaut wurden, werden in die Webpräsenz des Fachinformationsdienstes übernommen und weiter ausgebaut. Darüber hinaus werden spezielle Services für die Fachcommunity angeboten.

Kerngedanke der beiden Antragsteller Universitätsbibliothek und Prof. Dr. Chiarcos/AG Angewandte Computerlinguistik ist, sowohl die Recherche nach relevanter Information als auch den Zugriff darauf weiter zu verbessern. Dies betrifft nicht nur Sekundärliteratur, sondern vor allem auch Forschungsprimärdaten wie Sprachkorpora.

Um dies zu erreichen, werden die Vernetzung des Linguistik-Portals mit Linked Open Data (LOD) und die Erweiterung der entsprechenden LOD-Funktionalitäten vorangetrieben. Mithilfe von LOD-Mechanismen und Werkzeugen des Natural Language Processing wird ein automatisiertes Verfahren entwickelt, das Forschungsprimärdaten, die bislang nur formal ausgezeichnet sind, um inhaltliche Schlagwörter anreichert und dadurch besser auffindbar macht.

Die Sichtbarkeit von Sprachkorpora und ihrer wissenschaftlichen Analyse wird zudem dadurch erhöht, dass geeignete Publikationen aus der Bibliography of Linguistic Literature auf Metadatenebene mit den jeweils behandelten Sprachkorpora verlinkt werden.

Als weitere Maßnahme, um der besonderen Rolle der digitalen Korpora in der Sprachwissenschaft gerecht zu werden, ist im Rahmen des FID geplant, überregionale Lizenzen für kommerzielle, multilinguale Sprachkorpora anzubieten.

Sowohl der FID Biodiversitätsforschung als auch der FID Linguistik bieten wissenschaftlichen Open-Access-Zeitschriften als langfristigen Service eine technische und organisatorische Plattform für redaktionelle Bearbeitung und Hosting an (siehe auch UniReport Nr. 1/2017).

Beide Fachinformationsdienste werden darüber hinaus durch gezielten Erwerb und überregionale Bereitstellung dafür Sorge tragen, dass die Literatur des Spitzenbedarfs – sei es in gedruckter oder in elektronischer Form – auch weiterhin für die Forschung zur Verfügung steht.

Heike Renner-Westermann, Gerwin Kasperek

Weitere Informationen und Links

Projektinfos zum FID Linguistik
➤ www.ub.uni-frankfurt.de/projekte/fid-linguistik.html

Sammlung Linguistik
➤ www.ub.uni-frankfurt.de/linguistik/home.html

Projektinfos zum FID Biodiversitätsforschung
➤ www.ub.uni-frankfurt.de/projekte/biofid.html

Sammlung Biologie
➤ www.ub.uni-frankfurt.de/biologie

Universitätsbibliothek Johann Christian Senckenberg

www.ub.uni-frankfurt.de

Campus Bockenheim

Zentralbibliothek
Tel: (069) 798-39205/-39208
auskunft@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de/zentrale/so.html

Bibliothek Kunstgeschichte/Städelbibliothek und Islamische Studien
Tel: (069) 798-24979
kunstbibliothek@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de/kunstbibliothek/kmb_so.html

Mathematikbibliothek
Tel: (069) 798-23414
mathebib@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de/bnat/mathematik/home.html

Informatikbibliothek
Tel: (069) 798-22287
informatikbib@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de/bnat/informatik/home.html

Campus Westend

Bibliothek Recht und Wirtschaft (BRuW)
Tel: (069) 798-34965
bruw-info@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de/bruw/so_bruw.html

Bibliothek Sozialwissenschaften und Psychologie (BSP)
Tel: (069) 798-35122
bsp@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de/bsp/so.html

Bibliothekszentrum Geisteswissenschaften
Tel: (069) 798-32500 (Q1)
Tel: (069) 798-32653 (Q6)
bzg-info@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de/bzg/so_bzg.html

Campus Riedberg

Bibliothek Naturwissenschaften
Tel: (069) 798-49105
bnat@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de/bnat/bnat_so.html

Campus Niederrad

Medizinische Hauptbibliothek
Tel: (069) 6301-5058
h.krueger@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de/medhb/mallg.html

Sport-Campus

Bibliothek für Sportwissenschaften
Tel: (069) 798-24521
sportbib@ub.uni-frankfurt.de
www.ub.uni-frankfurt.de/bsp/sport

Einführungen und Schulungsveranstaltungen in der Universitätsbibliothek Johann Christian Senckenberg

Termine und Anmeldung
www.ub.uni-frankfurt.de/benutzung/literatursuche.html