

Herausforderungen bei der Dokumentation

Ein Gespräch mit den Direktoren
der Zentralen Fachbibliotheken TIB, ZB MED und ZBW

Am 18. und 19. März lud der Leibniz-Bibliotheksverbund Forschungsinformation GOPORTIS¹ zur Fachkonferenz „Non-Textual Information Strategy and Innovation Beyond Text“² nach Hannover ein. Die Direktoren der in GOPORTIS engagierten Zentralbibliotheken, **Uwe Rosemann**, Direktor der Technischen Informationsbibliothek (TIB), **Ulrich Korwitz**, Direktor der Deutschen Zentralbibliothek für Medizin (ZB MED) und Professor **Dr. Klaus Tochtermann**, Direktor der Deutschen Zentralbibliothek für Wirtschaftswissenschaften (ZBW) sprachen am Rande der Veranstaltung mit b.i.t.online über die aktuellen Herausforderungen bei der Dokumentation der Forschungserkenntnisse, die von der Wissenschaft mit Hilfe von Software und Computern erarbeitet werden.



Mit ihren Vorträgen über hoch visionäre Datenforschungsansätze und innovative Aufbereitung von Digitalvideos wirkt diese Konferenz streckenweise wie Science Fiction. Was hat das mit GOPORTIS zu tun?

ROSEMANN Der Bezug zu Goportis ist erst einmal ein allgemeiner. Wir bearbeiten in diesem Verbund ja bestimmte strategische Themenfelder und dazu gehört die Entwicklung von Verfahren für die Dokumentation von „Nicht-textuellen Materialien“. Der konkrete Bezug zu Hannover

als Ausrichtungsort dieser Konferenz ist, dass wir hier nach einer entsprechenden Evaluierung durch den Senat der Leibniz-Gemeinschaft ein Kompetenzzentrum für „Nicht-Textuelle Materialien“ aufbauen.

Gibt es dieses Kompetenzzentrum schon?

ROSEMANN Seit zwei Jahren befindet es sich im Aufbau und seine Fertigstellung hängt auch von zukünftigen Zusatzfinanzierungen ab. Wir sind aber ganz zuversichtlich, diese zu bekommen.

Warum beschäftigen Sie sich so intensiv mit Nicht-textuellen Materialien?

KORWITZ Wir stellen in unseren Bibliotheken einen gewissen Mediendruck fest: Immer unterschiedlichere

und verschiedenartigere Medien müssen bedient werden. Wir haben gestern in einem Vortrag gehört, dass in einer medizinischen Zeitschrift 95% der Artikel mit „Supplementary Material“ ausgestattet sind. Dieses Zusatzmaterial muss nicht nur verfügbar, sondern auch durchsuchbar gemacht werden. Die dafür notwendige semantische Bearbeitung beinhaltet große Arbeits- und Forschungsfelder im Rahmen von Science 2.0.

TOCHTERMANN Die Anforderungen an die Bibliotheken verändern sich, weil sich die Publikationsprozesse ändern. Während Bibliotheken in der Vergangenheit mehr das textuelle Material aufbereitet und verbreitet haben, geht es jetzt in zunehmendem Maß darum, dass man Forschungsdaten als wissenschaftliche Publikation hat, mitunter auch Software, mit der man eben mal Forschungsergebnisse generieren kann. Ergänzt wird das noch durch die vielen Multimediadaten, die heute eben durchaus auch Publikationen ausmachen können, z.B. Simulationen, Animationen, die Teil der Publikation sind. Vor diesem Hintergrund müssen die textuellen Dokumente um Nicht-textuelle Komponenten angereichert werden. Als Bibliotheken müssen wir uns darum genauso kümmern wie um die Texte, die in der Vergangenheit im Mittelpunkt standen.

Wo liegen denn die größten Probleme technisch, organisatorisch, politisch? In welchem dieser Felder?

TOCHTERMANN Die organisatorische Problematik besteht darin, dass die Forschungsgemeinschaft erst einmal überzeugt werden muss, ihre „Supplementary

¹ <http://www.goportis.de/ueber-goportis.html>

² Ein Konferenzbericht erscheint in der nächsten Ausgabe von b.i.t.online 16 (2013) | Ausgabe 3.



Sie arbeiten im GOPORTIS-Verbund gemeinsam an wichtigen Forschungsfragen, weil eine Bibliothek alleine die Aufgaben nicht mehr lösen kann. (v.l.n.r.) Professor Dr. Klaus Tochtermann, Direktor der ZBW, Uwe Rosemann, Direktor der TIB und Ulrich Korwitz, Direktor der ZB MED.

Data“ bei den Bibliotheken abzugeben. Ein Großteil dieser ergänzenden Daten liegt heutzutage noch auf den lokalen Arbeitsplatz-Rechnern der Forschenden und nur die Publikation, also das textuelle Dokument, das pdf oder die gedruckte Version, landet in den Bibliotheken.

Teilen Sie diese Meinung?

» **KORWITZ** ◀ Ja. Hier sind Anreize, „Incentives“, für die Forscher notwendig, ihre Forschungsdaten an einem sicheren Ort abzulegen. Wo dieser sichere Ort ist, muss dann noch geklärt werden. Ein solcher Anreiz wäre beispielsweise ein „Citation Index“ für Forschungsdaten, der ja langsam auch gebildet wird. Mit diesem Index bekommt der Forscher quasi durch das Ablegen seiner Daten ein Qualitätsmerkmal, das ihm in seiner wissenschaftlichen Karriere weiterhilft. Letztendlich muss es das eigene Interesse der Forscher werden. Zur Zeit fragen sich viele noch, was sie davon haben außer Arbeit. Da muss etwas geschehen, auch politisch.

Anreize für Forschende schaffen klingt interessant.

» **ROSEMANN** ◀ : Für mich ist es primär eine politische Herausforderung, da wir in Deutschland nicht die notwendigen Infrastrukturen haben, um das Problem zu

bewältigen. Es gibt Ansätze wie die KII-Konzeption³, die auch bei der GWK⁴ und im Wissenschaftsrat diskutiert wird. Aber das muss eben auch umgesetzt werden. Das setzt eine finanzielle Mittelbereitstellung in einer Größenordnung voraus, die im Kontext von Forschungsdaten für Deutschland ungewöhnlich hoch sein dürfte. Und es setzt Konsens von Bund und Ländern voraus, was bekanntermaßen nicht ganz einfach ist in Deutschland.

Unterstützen Sie die Umsetzung und Weiterentwicklung des KII-Infrastrukturkonzeptes?

» **ROSEMANN** ◀ Ja, wir alle drei haben ja im Steuerkreis und in den Arbeitsgruppen der KII mitgearbeitet und wir versuchen jetzt natürlich auch, die politische Umsetzung mit zu gestalten.

Wo sehen Sie den größten Informationsbedarf in Richtung Ihrer Bibliothekskunden, der Wissenschaft?

» **KORWITZ** ◀ Sie dafür zu gewinnen, ihre Forschungsdaten in den Portalen, die wir aufbauen, abzulegen und mit diesen Portalen zu arbeiten

3 <http://www.leibniz-gemeinschaft.de/infrastrukturen/kii/>

4 <http://www.gwk-bonn.de/>

Wer soll sich also Ihrer Meinung nach um Forschungsdaten kümmern?

» **ROSEMANN** ◀ Bislang haben wir ein Drei-Gruppen-Modell im Blick: die Wissenschaft, die wissenschaftliche Einrichtung und die Bibliothek. Die Wissenschaftler produzieren die Forschungsdaten, die wissenschaftlichen Einrichtungen fungieren als Datenzentrum und wir weisen die Metadaten dieser Datensammlung nach. Das sind Modelle, die kann man vielleicht nicht auf alle Fachgebiete anwenden. Hier muss man schauen, dass man Infrastrukturen schafft, die Dienstleistungen für Fachgebiete bereitstellt, die das nicht auf eigene Füße stellen können. Für mich ist das kein Informationsproblem, sondern ein Infrastrukturproblem.

Sollen Bibliotheken diese Dienstleistungen bereitstellen?

» **ROSEMANN** ◀ Nicht unbedingt. Aktuell haben wir mit über 50 wissenschaftlichen Instituten Verträge, in denen sie sich verpflichten, die Daten nachhaltig zu speichern: Die Bibliotheken referenzieren diese Metadaten z. B. bei DataCite. Unsere drei Bibliotheken sind ja Mitglieder von DataCite. Es gibt noch ein viertes deutsches Mitglied, die GESIS⁵. Dieses Modell funktioniert in bestimmten Fachgebieten bereits sehr gut, aber nicht in allen. Wo es nicht so gut funktioniert, kann man Dienstleistungen anbieten, wie sie z.B. im RADAR⁶-Projekt konzipiert sind, das uns gerade im Rahmen der DFG-Ausschreibung „Forschungsnahe Informationsinfrastruktur“ bewilligt wurde.

» **TOCHTERMANN** ◀ Eine große Schwierigkeit liegt unter anderem darin, dass in den ganzen Disziplinen und innerhalb der Disziplinen die Forschungsdaten unterschiedlich behandelt werden. In den Naturwissenschaften haben wir die Sensordaten, oft Gemeingut, in den Wirtschaftswissenschaften haben wir empirische Daten, die sind irgendwie geschützt, in den Sozialwissenschaften haben wir Umfragedaten aus der Bevölkerung, da kommt man in den Privacy-Bereich hinein. Wir haben also offene Daten, wir haben geschützt Daten, wir haben Privatsphären, die gewahrt werden müssen. Vor diesem Hintergrund ist wirklich jede einzelne Disziplin auch wieder einzeln zu behandeln. Das ist der große Unterschied zu den klassischen Publikationen, wo wir ein Textdokument haben, das in einer Zeitschrift, in einem Buch oder als pdf erscheint. Da hat man standardisierte Metadaten und gut ist. Bei Forschungsdaten ist die Bandbreite in den Daten und in den Metadaten viel größer.

5 <http://www.gesis.org/das-institut/>

6 http://www.dfg.de/foerderung/info_wissenschaft/info_wissenschaft_13_11/index.html

» **KORWITZ** ◀ Dazu kommt noch, dass die Publikationen früher nicht frei waren. Heute haben wir Open Access und wir versuchen bei den Forschungsdatensätzen auch nach Möglichkeit, eine Bereitstellung als Open Data zu erreichen. Ein Vertreter der STM⁷-Publisher war ja in Hannover anwesend und hat ganz klar Stellung bezogen. Seine Aussage: Wenn die Forschungsdaten Rohdaten sind, sind sie Open Data, wenn sie in irgendeiner Weise bearbeitet sind, befinden sie sich im Besitz der Verleger und müssten erworben werden. Das ist auch noch eine Sache, über die wir zu sprechen haben, denn wir streben Open Data an, natürlich mit Ausnahmen, wie z.B. Krankenakten in der Medizin. Aber das ist machbar, das kann man steuern.

» **ROSEMANN** ◀ An diesem Punkt ist aber auch die Wissenschaft nicht einheitlich aufgestellt. Wir vertreten beispielsweise das Fach Chemie, in dem es große Bedenken gegen Open Data gibt. Eine weitere Besonderheit: Die meiste Chemieforschung findet in der Industrie statt, wo natürlich an Verwertung und Patente gedacht wird. Das muss also auch alles sehr differenziert betrachtet werden.

Was glauben Sie: Wie lange wird es noch dauern, bis eine tragfähige Infrastruktur aufgebaut ist?

» **ALLE** ◀ Mindestens fünf bis acht Jahre, vielleicht auch zehn.

Nur noch so kurz?

» **ROSEMANN** ◀ Auf jeden Fall. Die Notwendigkeit ist erkannt, es gibt seit 2010 entsprechende Voten auch der EU-Kommission. „Riding the wave“⁸ ist hier der immer wieder zitierte Report und dieser sprach von zehn bis fünfzehn Jahren. Wir reden also von 2020 bis 2025. Aber dazu müssen Hunderte von Millionen, wenn nicht Milliarden in die Hand genommen werden.

» **KORWITZ** ◀ Die Politik ist aber auch aufgewacht; das Bundesforschungsministerium hat Big Data entdeckt. Das bringt am Anfang mit Sicherheit auch finanzielle Mittel mit sich.

Es bleibt also weiterhin spannend, ob und wie Forschungserkenntnisse im Computerzeitalter nachhaltig dokumentiert und nachgewiesen werden können. Vielen Dank für das Gespräch.

Das Interview führte Vera Münch für b.i.t.online.

7 Scientific, Medical and Technical Publishers; Verlage, die in den Bereichen Wissenschaft, Medizin und Technik tätig sind.

8 <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>