

Langzeitarchivierung von Forschungsdaten – Bericht von der nestor-School im Februar 2016

Irina Chaplinskaya-Sobol und Hendrik Wesner

► Bereits zum neunten Mal seit 2007 hatten die im nestor-Kompetenznetzwerk¹ organisierten Qualifizierungseinrichtungen im Rahmen der bewährten Weiterbildungsreihe der „nestor-Schools“ eingeladen. Die diesjährige Qualifizierungsveranstaltung zum Thema „Langzeitarchivierung von Forschungsdaten“ fand vom 15. bis 17. Februar in der Landesmusikakademie Hessen Schloss Hallenburg in Schlitz statt und erfolgte in Kooperation mit dem EU-Projekt PERICLES². Umgeben von der malerischen Kulisse und der künstlerisch-musischen Atmosphäre der Schlossakademie bekamen die TeilnehmerInnen einen umfassenden Überblick über die Thematik mit vielen praktischen Beispielen und Insiderwissen aus dem aktuellen Kontext.

Ganz in der Tradition der nestor-Schools erwies sich auch dieses Mal die gemischte Zusammensetzung der 38 TeilnehmerInnen und ReferentInnen als Erfolgskonzept. Der breite Teilnehmerkreis mit VertreterInnen verschiedener Informationseinrichtungen aus Deutschland und der Schweiz förderte einen regen disziplinübergreifenden Austausch.

Zum Auftakt lieferten die Organisatoren Stefan Strathmann (SUB Göttingen), Achim Obwald (TH Köln) und Jonas Recker (GESIS, Köln) einen breiten Überblick über die Aktivitäten, Kooperationspartner und Qualifizierungsangebote des nestor-Netzwerkes. Bewusst verzichtete man auch

diesmal auf eine theoretische Einführung in das Thema Langzeitarchivierung. Stattdessen fokussierten die Veranstalter einzelne wichtige Aspekte, die in Form von vier Arbeitsblöcken durch renommierte ReferentInnen präsentiert wurden. Nach einer Einführung durch die Experten wurden die Themen im Rahmen von Gruppenübungen vertieft und anschließend im Plenum diskutiert. Die Teilnehmenden konnten so auch ihren Background und die Perspektive der eigenen Einrichtung in die Diskussion einbringen.

Zunächst thematisierte Jens Ludwig (Staatsbibliothek zu Berlin, Stiftung Preußischer Kulturbesitz, Berlin) die Besonderheiten im Verständnis von Forschungsdaten und grenzte verschiedene Ebenen sowie Interessengruppen auf diesem Gebiet ein und machte auf die Unterschiede zwischen den Aspekten der Langzeitarchivierung und des Forschungsdatenmanagements aufmerksam. Auch die Wichtigkeit der Kooperationsaspekte und die Perspektiven der Neuausrichtung von generischen Infrastruktureinrichtungen wie Bibliotheken und Rechenzentren wurden im Vortrag betont.

Die weiterführende Gruppenübung hatte eine Einordnung von Forschungsdatenmanagementkonzepten in den Universitätskontext aus Sicht verschiedener Akteure und Zielgruppen zum Ziel. Die TeilnehmerInnen wurden angeregt, in verschiedenen Rollen innerhalb der Entscheidungsprozesse im Rahmen des FDM im universitären Kontext zu schlüpfen und die Thematik somit aus verschiedenen Blickwinkeln zu

betrachten. Die anschließende Gegenüberstellung von Interessengruppen in einer fiktiven Diskussionsrunde machte Interessenkonflikte und Lösungspotentiale sichtbar.

Am zweiten Tag stellte Fabian Cremer (Max Weber Stiftung, Bonn) die erweiterten Potenziale eines Embedded Data Managers (EDM) in der Forschungslandschaft vor. Sein Vortrag fokussierte die Neupositionierung der Informationsdienstleister durch seine „Einbettung“ in den Forschungsprozess. Auf der Grundlage der praktischen Beispiele aus dem Sonderforschungsbereich „Informationsinfrastrukturen“ der Deutschen Forschungsgemeinschaft thematisierte der Beitrag die konkrete Beschreibung der Aufgabengebiete, Kompetenzen und Erfolgskriterien eines EDM. Vor allem wurde betont, dass es sich dabei aktuell um kein etabliertes Berufsbild, sondern um eine spezifische Tätigkeitsform im Forschungskontext handelt.

Die anschließende Übungseinheit vertiefte das Thema, indem im Rahmen der Arbeitsgruppen die praktischen Anforderungen an einen EDM reflektiert wurden. Dabei standen viele Fragen zur Diskussion: Wie sieht das Idealbild und das optimale Modell eines Embedded Data Managers aus? Welche Anforderungen können an seine Kompetenzen gestellt werden? Auf welcher Ebene bringt sich ein Embedded Data Manager effizient ein und welche Fragen sind in seiner konkreten Tätigkeit zu berücksichtigen? Welche Rolle übernimmt er in einem konkreten Projekt und mit welcher Einbettungstiefe? Als besonders positiv wurde

1 http://www.langzeitarchivierung.de/Subsites/nestor/DE/Home/home_node.html [08.03.2016].

2 <http://pericles-project.eu> [08.03.2016].

in der Diskussion der Effekt betont, der aus der Kompetenznutzung der bereits vorhandenen Informationsdienstleister für die Forschungsvorhaben entsteht.

Wie können Forschungsdaten bewertet und auf dieser Grundlage relevante Daten für die Langzeitarchivierung ausgewählt werden? Mit dieser Frage beschäftigte sich der zweite Vortrag des zweiten School-Tages von Reiner Maurer (GESIS, Köln). Forschungsdaten werden in erster Linie nicht zum Zweck der reinen Archivierung von Daten aufbewahrt, sondern um eine Nachnutzung für die Wissenschaft zu ermöglichen. Forschungsdaten sind wertvolles Material, allerdings gibt es dennoch verschiedene beschränkende Faktoren wie zum Beispiel das Datenvolumen oder gesetzliche Bestimmungen. Aber auch hohe Kosten für die umfangreichen Prozesse der Weiterverarbeitung machen eine Auswahl von zur Archivierung vorgesehenen Forschungsdaten erforderlich. Die Trennung von relevanten und weniger relevanten Daten muss durch einen geregelten Auswahlprozess erfolgen. Dabei sind transparente und nachvollziehbare bzw. messbare Kriterien erforderlich.

Die Gruppenarbeit bot eine sehr gute Möglichkeit sich für das Thema zu sensibilisieren, indem relevante Kriterien für die Auswahl und Bewertung von persönlichen Daten in verschiedener Form aufgestellt wurden. Quintessenz der Übung war es, dass sich Nutzungsszenarien für Forschungsdaten nur schwer für die Zukunft abschätzen lassen. Fehleinschätzungen lassen sich zwar nicht unbedingt vermeiden, da es auch unter den TeilnehmerInnen zum Teil keinen Konsens über die Relevanz verschiedenster Datentypen gab, was zu einer interessanten und ausgiebigen Diskussion führte.

In den vierten und abschließenden Themenblock führte Anna Eggers von der SUB Göttingen am dritten Veranstaltungstag ein. Der Vor-

trag betrachtete Langzeitarchivierung bereits im Entstehungskontext von Forschungsdaten. Dabei werden relevante Umgebungsinformationen während der Entstehung von digitalen Objekten gesammelt und idealerweise mit den Objekten verkapselt. Für beide Anwendungszwecke wurden beim PERICLES-Projekt zwei Software-Tools entwickelt und stehen Interessenten frei zur Verfügung³. Dabei handelt es sich um PET (PERICLES Extraction Tool) zur Live-Analyse von digitalen Objekten und PeriCAT (PERICLES Content Aggregation Tool) zur Informationsverkapselung durch verschiedene technische Methoden.

Die beiden Tools wurden in der anschließenden praktischen Übung in den Gruppen näher betrachtet. Durch eine praxisorientierte Aufgabe konnte der Umgang mit den Tools geübt und ein Verständnis für die Funktionsweise aufgebaut werden. Es hat sich gezeigt, dass trotz des Status der Software als Entwicklerversion eine Einarbeitung in kurzer Zeit problemlos möglich war.

Zusammengefasst lässt sich die Veranstaltung als äußerst gelungen bezeichnen. Dies wurde auch deutlich an der durchweg positiven Resonanz der TeilnehmerInnen in der Abschlussrunde. In allen Beiträgen wurde die hohe thematische Relevanz der Langzeitarchivierung deutlich. Nicht nur die Vorträge und Übungssektionen, sondern auch die anregenden Diskussionen, an denen sich die Gäste mit wichtigen Impulsen beteiligten, trugen zum Gelingen der Veranstaltung bei.

In allen Beiträgen konnte das Thema Langzeitarchivierung aus sehr unterschiedlichen Perspektiven beleuchtet werden. Die TeilnehmerInnen hatten die Möglichkeit, sich beim gemeinsamen Abendprogramm auszu-

³ <https://github.com/pericles-project> [08.03.2016].



Schöner Tagungsort, funktionierendes Konzept: 38 Teilnehmende und Referierende aus Deutschland und der Schweiz trafen sich in der Landesmusikakademie Hessen Schloss Hallenburg zur nestor-School.

tauschen. Die gute Organisation trug ebenfalls dazu bei, dass die diesjährige nestor-School als positive und voranbringende Erfahrung bleibt. ■



Irina Chaplinskaya-Sobol
Medienzentrum
Institut für Medizinische Informatik
der Universitätsmedizin Göttingen



Hendrik Wesner
Hochschul- und Landesbibliothek
Fulda
Bereich IT-Projektmanagement und
Softwareentwicklung