

Policy der Deutschen National- bibliothek für dissertationsbezogene Forschungsdaten

Version 1.1

Stand: 11. August 2017

Im Auftrag der Deutschen Nationalbibliothek
erstellt von Dr. Dirk Weisbrod

Deutsche Nationalbibliothek (Leipzig, Frankfurt am Main)
2017

<urn:nbn:de:101-2017092701>

Policy der Deutschen Nationalbibliothek für dissertationsbezogene Forschungsdaten

Als zentrale Archivbibliothek und nationalbibliografisches Zentrum der Bundesrepublik Deutschland hat die Deutsche Nationalbibliothek (DNB) die Aufgabe, alle deutschen und deutschsprachigen Publikationen ab 1913, seit 2006 auch in digitaler Form zu sammeln. Dissertationen fallen selbstverständlich in die Sammelpflicht der DNB. Doch wie verhält es sich mit den Forschungsdaten, die zu diesen Dissertationen gehören? Gerade das digitale Publizieren bietet verschiedene Möglichkeiten, Forschungsdaten in Dissertationen zu integrieren oder ihnen beizufügen, sodass an dieser Stelle ein grundsätzlicher Klärungsbedarf besteht. Aufgabe der Policy ist es, die Zuständigkeit der DNB für dissertationsbezogene Forschungsdaten zu definieren.

1. Gesetzliche Grundlagen und Grundsätze guter wissenschaftlicher Praxis

Grundlage für den Sammelauftrag ist die zuletzt 2014 geänderte Pflichtablieferungsverordnung (PflAV). Während „selbstständig veröffentlichte Primär-, Forschungs- und Rohdaten“ nach § 9, Absatz 10 der Pflichtablieferungsverordnung (PflAV) vom Sammelauftrag ausgeschlossen sind, bestimmt § 7, Absatz 2, PflAV eindeutig, dass die „Ablieferungspflicht (...) auch alle Elemente, Software und Werkzeuge [umfasst], die in physischer oder in elektronischer Form erkennbar zu den ablieferungspflichtigen Netzpublikationen gehören, auch wenn sie für sich allein nicht der Ablieferungspflicht unterliegen.“ Die Grundsätze zur Sicherung guter wissenschaftlicher Praxis setzen zudem die Sichtbarkeit und Nachprüfbarkeit jener Forschungsdaten voraus, die einer Dissertation zugrunde liegen.

2. Welche Forschungsdaten sammelt die DNB? – Kriterium der Unverzichtbarkeit

Die DNB wird aufgrund der gesetzlichen Bestimmungen und unter Berücksichtigung der Grundsätze zur Sicherung guter wissenschaftlicher Praxis jene öffentlich zugänglichen Forschungsdaten sammeln, die für eine ablieferungspflichtige Dissertation und die Nachprüfbarkeit der darin enthaltenen Forschungsergebnisse *unverzichtbar* sind. Über die Unverzichtbarkeit von dissertationsbezogenen Forschungsdaten entscheiden entweder die Promovierenden selbst oder die für die Ablieferung der Dissertation zuständigen Hochschulen, indem sie der abzuliefernden Dissertation Forschungsdaten beifügen oder den Persistent Identifier (PID) von Daten angeben, die bereits auf anderen Repositorien veröffentlicht sind. Diese Daten werden als Teil der Dissertation selbst betrachtet.¹ Die DNB fordert die Ablieferung von Forschungsdaten nicht aktiv ein, sondern überlässt die Definition entsprechender Rahmenbedingungen der jeweiligen Hochschule. Ebenso entscheidet sie nicht über die Unverzichtbarkeit von Forschungsdaten.

¹ Die abgelieferten Forschungsdaten müssen maschinenlesbar sein.

Die DNB unterscheidet somit technisch die folgenden Arten von Forschungsdaten:

1. in die Dissertation eingebundene Daten (z.B. Grafiken oder Tabellen. Diese Daten werden bereits gesammelt),
2. der Dissertation als Supplement beigegebene Daten (einzelne Daten oder Datencontainer, die bei der Ablieferung der Dissertation als separate Dateien beigelegt werden),
3. externe Daten; das sind Daten, die in einer Dissertation ausgewertet wurden und nach Auffassung der Promovierenden oder der ablieferungspflichtigen Instanz für die Dissertation unverzichtbar im oben ausgeführten Sinne sind, aber bereits von Dritten (Forschungsdatenrepositorien, Verlage etc.) archiviert werden. Dabei muss bei Ablieferung explizit auf diese Daten hingewiesen werden, etwa indem der DNB der Persistent Identifier (PID) der Daten übermittelt wird.

Die DNB bezieht sich damit auf Typ (1) bis (3) der „ODE Data Publication Pyramid“². Aufgrund des weiter oben bereits zitierten § 9, Absatz 10, PflAV, schließt sie explizit selbstständige Datenpublikationen (4) und Rohdaten (5) aus ihrem Sammelauftrag aus.

Um Redundanzen zu vermeiden, sammelt die DNB keine Daten, die für die betreffende Forschungsdisziplin allgemeine Bedeutung haben und sehr häufig auch in anderen Publikationen genutzt oder zitiert werden (etwa Panel-Daten in der Sozialwissenschaft), auch wenn sie im oben ausgeführten Sinne unverzichtbar sind. Die Promovierenden oder die abliefernde Institution haben jedoch die Möglichkeit, die PIDs dieser Daten bei der Ablieferung in einer Form anzugeben, die es der DNB erlaubt, diese PIDs in den Metadaten zur Dissertation zu zitieren.

3. Unterstützung eines verteilten Archivierungsansatzes

Unter Berücksichtigung der Empfehlungen des Rates für Informationsinfrastrukturen (RfII), wonach „Aufgaben der Langzeitarchivierung als informationsinfrastrukturelle Daueraufgaben durch ggf. miteinander vernetzte Einrichtungen im Rahmen der NFDI“ (Nationalen Forschungsdateninfrastruktur) zu leisten und zu koordinieren sind³, unterstützt die DNB bei Forschungsdaten einen verteilten Archivierungsansatz. Vielfach liegen die Daten heute schon bei der Ablieferung auf einem Hochschul- oder Verlagsserver oder auf einem Forschungsdaten-Repositorium vor. Die DNB übernimmt deshalb extern gespeicherte Daten (Typ 3) in der Regel nicht, sondern referenziert diese Daten nur, indem sie sie in geeigneter Weise mit dem Metadatensatz der entsprechenden Dissertation verknüpft. Sie behält sich jedoch eine Übernahme für den Fall vor, dass die archivierende Instanz keine Langzeitarchivierung ermöglichen kann und ihren Sitz oder eine Betriebsstätte in Deutschland hat. Derzeit wird eine Zertifizierung nach dem Data Seal of Approval (DSA) oder dem nestor-Siegel als ausreichend angesehen.⁴ Supplemente (Typ 2) übernimmt die DNB in jedem Fall. Daten, die auf ausländischen Repositorien liegen, werden von der DNB grundsätzlich nicht von diesen übernommen. Sie können der Dissertation aber vom Ablieferer als Supplement beigelegt werden. Die Ablieferungsprozesse und Kriterien hierfür werden entsprechend definiert. Wie schon bei der Ablieferung von Dissertationen werden die

² Reilly, Susan [u. a.]: Opportunities of Data Exchange: Report on Integration of Data and Publications, hdl:10013/epic.40198.d001, S. 18ff.

³ Rat für Informationsinfrastrukturen [Hrsg.]: Leistung aus Vielfalt: Empfehlungen zu Strukturen, Prozessen und Finanzierung des Forschungsdatenmanagements in Deutschland. Göttingen: 2016, S. 47

⁴ Dies entspricht den Empfehlungen des Rates für Informationsinfrastruktur. Vgl. Rat für Informationsinfrastrukturen 2016, S. 30.

Hochschulen gebeten, die für die Übernahme oder die Referenzierung nötigen Angaben und Daten zu erheben.

4. Langzeitarchivierung von Forschungsdaten

Je nach Wissenschaftsdisziplin kann es sich bei Forschungsdaten um Messdaten, Beobachtungsdaten, Umfrageergebnisse, oder andere Arten von Daten handeln. Aufgrund der Vielfalt und Komplexität der Datenformate kann die Langzeitarchivierung im Sinne einer unbeschränkten Nutzbarkeit über lange Zeiträume nicht für alle Forschungsdaten von der DNB gewährleistet werden. In Ergänzung ihrer derzeit gültigen Dateiformat-Policy übernimmt die DNB dissertationsbezogene Forschungsdaten und setzt nach Möglichkeit alle geeigneten Langzeitarchivierungs-Strategien ein.

Auf Grundlage dieser Policy wird die DNB ihren Pflichtablieferungsworkflow für Publikationen so weiterentwickeln und anpassen, dass die Übernahme und Langzeitarchivierung von dissertationsbezogenen Forschungsdaten im Rahmen der jeweils bestehenden Möglichkeiten gewährleistet ist. Prozesse und Kriterien, die aus dieser Policy resultieren, werden zu einem späteren Zeitpunkt veröffentlicht.