

# Forschungsdaten in Theorie und Praxis. Das DARIAH-DE Repository und die DARIAH-DE Collection-Registry

*Sektionsvorschlag - DHd2015 „Von Daten zu Erkenntnissen“, eingereicht von:  
Peter Andorfer, Johanna Puhl, Stefan Schmunk*

## Abstract zur Sektion

Das Thema „Forschungsdaten“ ist auch innerhalb DARIAH-DEs von zentraler Bedeutung. Dies gilt sowohl für die theoretisch-methodische Verortung dieses Begriffes als auch hinsichtlich des praktischen Umgangs mit Forschungsdaten in den kultur- und geisteswissenschaftlich arbeitenden Disziplinen. Die konkrete Arbeit kreist dabei vor allem um folgende Fragestellungen und Aufgabengebiete:

- (1) Was sind Forschungsdaten in den Kultur- und Geisteswissenschaften? Kann angesichts der hohen Heterogenität in den einzelnen Disziplinen, deren vielfältigen Forschungsinteressen, -materialen und Methoden überhaupt eine allgemein verbindliche Definition des Begriffes Forschungsdaten gefunden werden und falls ja, was sind deren Kriterien und welche technisch-praktischen Konsequenzen lassen sich daraus wiederum für die Generierung, Sicherung und Distribution von Forschungsdaten ableiten.
- (2) In engem Zusammenhang dazu stehen die Fragen zum Lebenszyklus von Forschungsdaten: So können Forschungsdaten in unterschiedlichen Phasen eines Projektes auf unterschiedliche Art und Weise erzeugt, gesammelt, aufbereitet und/oder analysiert werden. Forschungsdaten können dabei Ergebnis und/oder Quelle eines Forschungsprojektes sein. Diese Dynamik soll in einem eigenen, speziell auf die Eigenschaften kultur- und geisteswissenschaftlicher Forschungsdaten abgestimmten Modell abgebildet werden. Gleichzeitig soll dieses Modell eines Forschungsdatenzyklus auch (technische) Anforderungen an Storage- und Publikationssysteme für digitale geistes- und kulturwissenschaftliche Forschungsdaten beschreiben können. Besonders interessant ist an dieser Stelle die Frage, inwiefern sich mithilfe von automatischen Prozessen in einer den Forschungsdatenzyklus unterstützenden Infrastruktur Erkenntnisse gewinnen lassen.

(3) Die in den Punkten eins und zwei ausgearbeiteten Anforderungen werden bei der Entwicklung des DARIAH-DE Repositoriums und der DARIAH-DE Collection Registry aufgegriffen und realisiert. Forschungsdaten, die in das Repository zur langfristigen Archivierung hochgeladen werden, werden in der Collection Registry kontextualisiert, als Sammlung von Forschungsdaten einem konkreten Forschungsprojekt zugeordnet und somit für die Nachnutzung durch andere aufbereitet.

Die Vorträge der Sektion „Forschungsdaten in Theorie und Praxis“ folgen diesen eben skizzierten Themenkomplexen. **Der erste Vortrag mit dem Titel „Forschungsdaten – Versuch einer Definition“ (Peter Andorfer, HAB Wolfenbüttel)** versucht in einem ersten Schritt eine generische Definition von „digitalen geistes- und kulturwissenschaftlichen Forschungsdaten“ und testet die Funktionalität dieser Definition anhand eines konkreten (geschichts)wissenschaftlichen Forschungsprojektes bzw. der darin gesammelten, beschrieben und/oder erzeugten (Forschungs?)Daten. **Im zweiten Vortrag „Definition des DARIAH Research Data LifeCycle“ (Johanna Puhl, HKI Köln)** werden der „DARIAH Research Data LifeCycle“ vorgestellt, die Besonderheiten und Spezifika gegenüber bereits bestehenden Modellen herausgearbeitet und die technische Anforderungen an eine Infrastruktur zur Speicherung und Publikation von Forschungsdaten formuliert. Eine solche von DARIAH-DE entwickelte Infrastruktur wird **im dritten Vortrag „Die Nutzung von Geistes- und kulturwissenschaftlichen Forschungsdaten - Das DARIAH-DE Repository“ (Stefan Schmunk, SUB Göttingen)** vorgestellt. Neben den technischen und administrativen Aspekten (wer kann, wie, unter welchen Voraussetzungen, ab wann und wie lange Repository und Collection Registry nutzen) soll anhand der bereits aus dem ersten Vortrag bekannten Forschungsdaten der Vorgang des Dateneingest in das Repository und deren Registrierung in der Collection Registry exemplarisch vorgeführt werden.

Mit Hilfe der hier vorgestellten Sektionen sollen vornehmlich zwei Ziele erreicht werden. Einerseits geht es darum, die DARIAH-DE Collection Registry sowie das DARIAH-DE Repository in der DH-Community und über diese hinaus bekannt zu machen. Andererseits sollen die innerhalb von DARIAH-DE erarbeiteten Konzepte und Definitionen zum Forschungsdatenbegriff und zum Research Data Lifecycle mit Vertretern der unterschiedlichen geistes- und kulturwissenschaftlichen Disziplinen diskutiert werden.