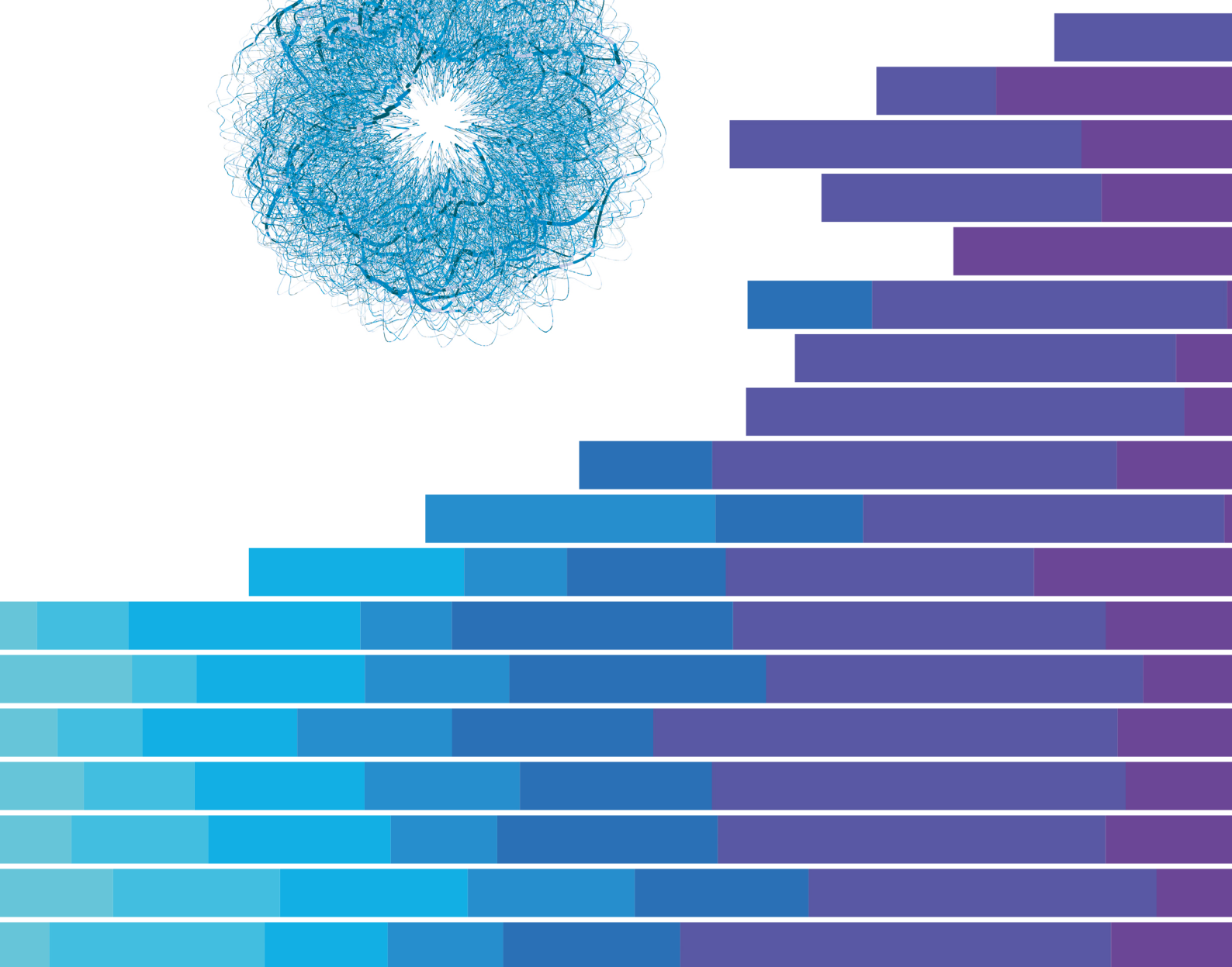


# **Forschende und ihre Daten**

## **Ergebnisse einer österreichweiten Befragung**

### **Executive Summary**



## Einleitung

Dieser Report gibt einen Überblick über die österreichweite Befragung zu Forschungsdaten, die im Rahmen des Projekts e-Infrastructures Austria<sup>1</sup> zu Jahresbeginn 2015 durchgeführt wurde. Diese richtete sich an das wissenschaftliche und künstlerisch-wissenschaftliche Personal aller 21 öffentlich-rechtlicher Universitäten sowie an drei außeruniversitäre Forschungseinrichtungen in Österreich.

Die Teilnehmenden wurden zu folgenden Themenbereichen befragt:

- Datentypen und Formate
- Datenarchivierung, -sicherung und -verlust
- Ethische und rechtliche Aspekte
- Zugänglichkeit und Nachnutzung
- Infrastruktur und Services

Die in diesem Kontext erstmals auf nationaler Ebene durchgeführte Befragung diente der Erhebung des praktischen Umgangs mit Forschungsdaten in Österreich und ist somit die Basis für eine konsekutive Optimierung der zweckdienlichen Infrastruktur, für eine Anpassung der Serviceangebote sowie für eine Neuorientierung bei der Ermittlung von Ressourcen in diesem strategischen Bereich entsprechend der geäußerten Bedürfnisse der im Forschungsprozess Tätigen.

## Hintergrund

Ein solides Forschungsdatenmanagement ist die Grundlage für eine kooperative, offene Wissenschaft und somit für ihre Nachvollzieh- und Überprüfbarkeit. Auch Sichtbarkeit und Reputation der österreichischen Forschungslandschaft spielen eine wichtige Rolle. Die Thematik ist für Forschende, Fördergeber und leitende Stellen wissenschaftlicher Einrichtungen gleichermaßen relevant und aktuell, was gegenwärtig durch den *Pilot for Research Data* der Europäischen Kommission<sup>2</sup> demonstriert wird.

## Methodik

Die Befragung basiert auf bereits durchgeführten, institutionellen oder disziplinspezifischen Umfragen an Hochschulen und Forschungseinrichtungen in anderen Ländern. Ein besonderer Fokus wurde bei der Erstellung des eigens entwickelten Fragebogens auf die Berücksichtigung der unterschiedlichen Kulturen in Wissenschaft und Kunst gelegt. Er wurde in deutscher und englischer Sprache erstellt und mittels der Open Source-Software *LimeSurvey* programmiert. Der Durchführungszeitraum war zwischen 19. Januar und 31. März 2015. Die 3026 vollständig ausgefüllten Fragebögen (entspricht einer durchschnittlichen Rücklaufquote von 9 %) wurden mittels dem Open Source-Statistikprogramm *R* sowie dem OpenSource-Tabellenkalkulationsprogramm *OpenOffice* statistisch ausgewertet.

## Wichtigste Ergebnisse

Die Ergebnisse der Studie bestätigen die gängigen Erwartungen hinsichtlich des Umgangs mit Forschungsdaten und sichern diese statistisch ab. Für jeden Themenbereich wurden sowohl fächerübergreifende Gemeinsamkeiten als auch disziplinspezifische Besonderheiten – sofern relevant – ermittelt.

---

1 Projekt e-Infrastructures Austria, Website. Online unter: <http://e-infrastructures.at/startseite> (Zugriff: 30.09.2015).

2 European Commission: HORIZON 2020, Open Science (Open Access). Online unter: <http://ec.europa.eu/programmes/horizon2020/en/h2020-section/open-science-open-access> (Zugriff: 30.09.2015).

## ***Datentypen und Formate***

- Vom Großteil der Forschenden werden Forschungsdaten in Form unstrukturierter Textdateien, Grafiken und Tabellen generiert. Ein Viertel nutzt strukturierten Text, ein Viertel Videos, Datenbanken und Quellcode, ein Fünftel Audio und Software. Während in den technischen Disziplinen – wie erwartet – häufig Quellcode- und Konfigurationsdaten generiert werden, fällt in den Geisteswissenschaften und der Medizin besonders die vergleichsweise häufige Erzeugung von Datenbanken auf.
- Der Großteil der Forschenden erzeugt mehr als drei Viertel seines Forschungsdatenvolumens in digitaler Form; analoge Daten werden nur von etwas mehr als jeder/jedem Zwanzigsten (insbesondere in den Geisteswissenschaften) häufig genutzt.

## ***Datenarchivierung, -sicherung und -verlust***

- Die Mehrzahl der Befragten nutzt mehrere Speichermöglichkeiten, wobei eine klare Präferenz für die Nutzung des eigenen dienstlichen oder privaten Rechners sowie externer Festplatten und USB-Laufwerke ersichtlich ist.
- Zwei Drittel der Forschenden benötigen Speicherplatz in einer Größenordnung von bis zu 100 GB pro Jahr. Ein höherer Speicherbedarf lässt sich sowohl für die medizinischen als auch die künstlerischen Universitäten belegen.
- Mehr als zwei Drittel der Forschenden geben an, ihre Forschungsdaten individuell und uneinheitlich zu beschreiben, und mehr als neun von zehn, dass sie für die Archivierung der Forschungsdaten selbst zuständig sind.
- Mehr als ein Drittel hat bereits Erfahrungen mit Datenverlust gemacht.

## ***Ethische und rechtliche Aspekte***

- Während je ein Drittel der Forschenden angibt, nie oder selten mit rechtlichen Unklarheiten bei Fremddatennutzung konfrontiert zu sein, stellen sich für ein Fünftel zumindest manchmal rechtliche Unklarheiten.
- Bei einem Institutionswechsel verbleiben die Forschungsdaten tendenziell an der betreffenden Einrichtung; von zirka der Hälfte der Forschenden werden diese aber auch mitgenommen.
- Sensible Daten fallen bei jeder/jedem siebenten Forschenden an. Eine große Rolle spielen diese in der Medizin, wo vier von zehn Forschenden angeben, solche Daten oft zu nutzen.

## ***Zugänglichkeit und Nachnutzung***

- Die Nutzung von Fremddaten wird von vielen Forschenden als wesentlicher Aspekt zur Durchführung ihrer Forschung angesehen, während ein Viertel überhaupt keine Fremddaten verwendet.
- Zugriff auf selbst generierte Forschungsdaten für Dritte ermöglichen Forschende in der Regel nur eingeschränkt. Während etwas mehr als die Hälfte der Befragten den Zugang nur auf Anfrage ermöglicht, stellt nur jede/jeder Zehnte seine Forschungsdaten als *Open Data* für die Öffentlichkeit zur Verfügung; ebenso viele verweigern den Zugriff ganz.
- Zugriff auf Forschungsdaten wird von der Mehrheit der Forschenden entweder über physische Datenträger oder per E-Mail ermöglicht. Mehr als zwei Drittel der Forschenden setzen hierfür Cloud- oder Website-Anwendungen ein; Datenarchive/Repositorien werden nur von jeder/jedem siebenten Forschenden genutzt.
- Ungefähr ein Drittel der Befragten ermöglicht die Nachnutzung ihrer eigenen Forschungsdaten; tendenziell häufiger geschieht dies in der Geographie, Biologie und Chemie, verhältnismäßig seltener in der Medizin, den Sozialwissenschaften und Geisteswissenschaften. Nutzungsvereinbarungen werden von mehr als einem Drittel der Forschenden abgeschlossen.
- Von mehr als der Hälfte der Forschenden wurden erhöhte Sichtbarkeit und Impact, neue Kooperationsmöglichkeiten, Anerkennung in der Fachöffentlichkeit sowie deren Berücksichtigung als wissenschaftlicher Output als attraktivste Anreize bezeichnet, um ihre Daten zu teilen.
- Vor allem der erhöhte Zeit- und Kostenaufwand, ein möglicher Datenmissbrauch, rechtliche Unsicherheiten, eine potenzielle Datenverfälschung, unerwünschte Kommerzialisierung sowie Erhöhung des Konkurrenzdrucks wurden als hauptsächliche Hinderungsgründe genannt. Rechtliche Einschränkungen

stellen dabei insbesondere in der Medizin, den Sozial- und Verhaltenswissenschaften sowie den Ingenieurwissenschaften die größten Hindernisse dar.

### **Infrastruktur und Services**

- Bezüglich des bevorzugten Datenarchivs zeigen die Forschenden keine eindeutige Präferenz. Relativ häufig wurden in diesem Zusammenhang das internationale fachspezifische Datenarchiv, das institutionelle Repositorium, das internationale multidisziplinäre Datenarchiv und das landesweite fachspezifische Repositorium genannt.
- Die Mehrzahl der Forschenden wünscht sich technische Infrastrukturen sowie projektspezifische Unterstützung für das Forschungsdatenmanagement. Darüber hinaus zeigt mehr als ein Drittel Interesse an Rechtsberatung, einem allgemeinen Helpdesk sowie an Schulungsangeboten.
- Mehr als die Hälfte der Forschenden erwartet die Bereitstellung von zusätzlichem qualifiziertem Personal sowie die Verabschiedung von Policies zum Umgang mit Forschungsdaten. Ein Fünftel wünscht eine Aufnahme von Forschungsdatenmanagement als Lehrinhalt im Curriculum sowie eine Verankerung als Dienstpflicht.

### **Empfehlungen**

Auf Basis der vorliegenden Umfrageergebnisse wird die Umsetzung folgender Maßnahmen für den Umgang mit Forschungsdaten in Österreich dringend empfohlen:

- Schaffung einer flächendeckenden technischen Infrastruktur in Österreich unter Berücksichtigung von disziplinären Bedürfnissen
- Verabschiedung von institutionellen Policies
- Bestellung von Datenfachleuten
- Einrichtung von unterstützenden Services für die Forschenden
- Implementierung von geeigneten Anreizsystemen
- Förderung internationaler und interdisziplinärer Zusammenarbeit

Diese Empfehlungen verfolgen das Ziel, leistungsstarke Infrastrukturen für einen adäquaten Umgang mit Forschungsdaten in der österreichischen Wissenschaftslandschaft zu initiieren. Bei der konkreten Umsetzung wie auch bei der Weiterentwicklung dieser Infrastrukturen sind nicht nur die rasanten Veränderungen auf diesem Gebiet im internationalen Kontext zu berücksichtigen, sondern auch internationale Kooperationen anzustreben, um Synergien zu entwickeln. Folgewirkungen der Etablierung von Infrastrukturen für Forschungsdaten sind eine Erhöhung der Sichtbarkeit und Reputation der einzelnen beteiligten Forschungseinrichtungen sowie der österreichischen Forschung als Ganzes.

Gesamtreport siehe:

PI: [phaidra.univie.ac.at/o:407513](https://phaidra.univie.ac.at/o:407513)

DOI: [10.5281/zenodo.32043](https://doi.org/10.5281/zenodo.32043)



Dieses Werk bzw. dieser Inhalt steht unter einer Creative Commons Namensnennung 4.0 International Lizenz.  
<http://creativecommons.org/licenses/by/4.0/legalcode>