

Informationsinfrastrukturen für Forschungsdaten

Frequently Asked Questions (FAQs)

Die FAQs zum Förderprogramm werden kontinuierlich aktualisiert und sind als ergänzende Erläuterungen zum Programmmerkblatt¹ zu verstehen

Versionsdatum 07.04.2022

Allgemeine Informationen

Was sind Forschungsdaten?

- Forschungsdaten² sind eine wesentliche Grundlage für das wissenschaftliche Arbeiten. Die Vielfalt solcher Daten spiegelt die Vielfalt unterschiedlicher wissenschaftlicher Disziplinen, Erkenntnisinteressen und Forschungsverfahren wider. Zu Forschungsdaten zählen u. a. Messdaten, Laborwerte, audiovisuelle Informationen, Texte, Surveydaten oder Beobachtungsdaten, methodische Testverfahren sowie Fragebögen. Korpora und Simulationen können ebenfalls zentrale Ergebnisse wissenschaftlicher Forschung darstellen und werden daher ebenfalls unter den Begriff Forschungsdaten gefasst.
- Da Forschungsdaten in einigen Fachbereichen auf der Analyse von Objekten basieren (z. B. Gewebe-, Material-, Gesteins-, Wasser- und Bodenproben, Prüfkörper, Installationen, Artefakte und Kunstgegenstände), muss der Umgang mit diesen ebenso sorgfältig sein.
- Im Rahmen des Förderprogramms „Informationsinfrastrukturen für Forschungsdaten“ umfasst der Begriff „Forschungsdaten“ nicht Forschungssoftware.

Was ist der Unterschied zwischen Infrastruktur- und Forschungsförderung bei der DFG?

- Forschungsförderung zielt wesentlich darauf, neue Einsichten und Erkenntnisse zu gewinnen. Infrastrukturförderung zielt hingegen auf die (Weiter-)Entwicklung von Technologien, Systemen und Werkzeugen, mit denen wissenschaftlich relevante Daten z.B. erhoben, analysiert, verbreitet oder langfristig gesichert werden können.
- Bei der Infrastrukturförderung stehen Aufbau und Weiterentwicklung leistungsstarker Informationsinfrastrukturen für die Wissenschaften im Zentrum der Förderung.
- In der Infrastrukturförderung – zu der alle LIS-Förderprogramme gehören – kann keine Grundlagenforschung gefördert werden. Somit sollen die in Infrastrukturvorhaben zu bearbeitenden Fragestellungen auch nicht als „Forschungsfragen“ bezeichnet werden. Dies gilt in diesem Programm insbesondere auch für förderbare Studien, die ebenfalls nicht unter die von der DFG geförderte Grundlagenforschung fallen. Im Programmkontext dienen Studien vielmehr dazu, Fragen rund um das Forschungsdatenmanagement zu klären, die für den Aufbau und die Weiterentwicklung von Forschungsdateninfrastrukturen relevant sind.
- Mit der Infrastrukturförderung kann keine wissenschaftliche Qualifikation gefördert werden.

¹ www.dfg.de/formulare/12_14/

² Weitere Informationen der DFG zum Thema Forschungsdaten: www.dfg.de/antragstellung/forschungsdaten

Was sind Informationsinfrastrukturen für Forschungsdaten?

- „Informationsinfrastrukturen für Forschungsdaten“ oder „Forschungsdateninfrastrukturen“ umfassen sowohl technische als auch organisatorische Strukturen, die den Wissenschaften den Umgang mit Forschungsdaten ermöglichen.

Was ist Informationsinfrastruktursoftware?

- Ein Aspekt der technischen Grundlagen von Informationsinfrastrukturen ist Informationsinfrastruktursoftware. Diese beruht überwiegend auf einem Softwarestack. Die einzelnen hierarchischen Komponenten dieses Stacks sind häufig nachnutzbare und etablierte Software (z.B. Betriebssystem, Web-Server-Technologie, Laufzeitumgebung, Programmiersprachen, Compiler, Datenbanken, Frameworks, Bibliotheken und Schnittstellen).
- Mit der Bezeichnung „Informationsinfrastruktursoftware“ wird verdeutlicht, dass es sich bei Informationsinfrastrukturen auf der Anbieter- und Betriebsseite um Softwaresysteme handelt, deren Entwicklung, Implementierung und funktionelle Weiterentwicklung im Rahmen des Förderprogramms gefördert werden kann.

Was bedeutet vertikale Integration einer Forschungsdateninfrastruktur?

- Die vertikale Integration meint die nutzerorientierte, technische und ggf. organisatorische Einbindung einer Informationsinfrastruktur beginnend am Arbeitsplatz der Wissenschaftlerinnen und Wissenschaftler in die lokalen Strukturen und Prozesse über die regionale (z.B. Forschungsdatenmanagement-Landesinitiativen) und nationale (z.B. Nationale Forschungsdateninfrastruktur (NFDI)) bis hin zur internationalen Ebene (z.B. European Open Science Cloud (EOSC)).
- Antragstellenden wird empfohlen, sich frühzeitig mit nationalen und internationalen Initiativen zu befassen und zu vernetzen.
- Insbesondere im Sinne der vertikalen Integration, aber auch im Sinne der Harmonisierung von Standards, Organisationsmaßnahmen etc. ist eine Interoperabilität und Nutzung von Synergieeffekten unbedingt anzustreben.

Fragen zur Antragstellung/ Begriffserläuterungen im Sinne des Programms

Können in diesem Programm mehrere Antragstellende gemeinsam einen Antrag stellen?

- Die Kriterien zur Antragsberechtigung sind in Abschnitt 2.1 des Programmmerkblasss dargelegt. Daraus wird deutlich, dass Wissenschaftlerinnen und Wissenschaftler und auch Einrichtungen antragsberechtigt sind.
- Es wird ausdrücklich eine kooperative Antragsstellung von Infrastrukturanbietern sowie Nutzerinnen und Nutzern empfohlen, sodass Vorhaben frühzeitig und konsequent auf den Nutzen für Wissenschaftlerinnen und Wissenschaftlern ausgerichtet werden können.
- Personen oder Institutionen, die z.B. in NFDI, EOSC, Fachinformationsdiensten (FID) oder einer anderen Initiative mitwirken, können selbstverständlich Anträge im Förderprogramm „Informationsinfrastrukturen für Forschungsdaten“ stellen, sofern die formalen Voraussetzungen erfüllt sind. Dabei ist eine Doppelförderung desselben Vorhabens ausgeschlossen.

Was ist eine Bedarfsanalyse?

- Alle Vorhaben müssen bedarfsorientiert sein, also spezifische Anforderungen einschlägiger Communities erfüllen; dieser Bedarf kann in den Wissenschaften oder auch bei wissenschaftlichen Infrastruktureinrichtungen bestehen. Dies aufzuzeigen ist das Ziel der Bedarfsanalyse.
- Grundlagen für die Bedarfsanalyse können Ergebnisse aus Workshops, Umfragen, Letters-of-Support, etc. sein. Bereits vorhandene Informationen können nachgenutzt werden.

Was ist eine Umfeldanalyse?

- Die Umfeldanalyse zeigt auf, dass es für den ermittelten Bedarf noch keine angemessene oder ausreichende Lösung gibt. Dies begründet das beantragte Vorhaben.
- Im Hinblick auf die Ziele des Vorhabens zeigt die Umfeldanalyse auf, welche technischen und organisatorischen Lösungen ggf. nachgenutzt werden können und wie sich neue Lösungen von vorhandenen abgrenzen.
- Gleichzeitig geht aus der Umfeldanalyse hervor, in welche horizontalen und vertikalen Strukturen ein Vorhaben eingebettet werden kann bzw. soll.

Was ist eine Risikoanalyse?

- Eine Risikoanalyse beschreibt mögliche Abweichungen von der Projektplanung, die bei der Durchführung des Projekts eintreten können. Darüber hinaus zeigt sie Strategien auf, wie Risiken minimiert werden können bzw. wie mit Abweichungen angemessen umgegangen werden kann.
- Die Risikoanalyse kann zum Beispiel personelle, technische, organisatorische, inhaltliche oder andere Aspekte betreffen.

Können auch Sachmittel für juristische Beratung beantragt werden?

- Alle juristischen Fragen, die unmittelbaren Einfluss auf die Durchführbarkeit des Projekts haben, müssen vor der Antragsstellung geklärt und im Antrag dargelegt werden.
- Juristische Fragen, die die Durchführung des Projekts nicht gefährden, aber für den Umgang mit Forschungsdaten relevant sind, können im Rahmen des Projekts geklärt und in diesem Rahmen gefördert werden.

Können Investitionsmittel zur Beschaffung von Hardware beantragt werden?

- Dies ist nur dann möglich, wenn der projektspezifische Bedarf überzeugend dargelegt wird. Wenn es z.B. um dauerhaften Speicherplatz geht, muss dieser aus der Grundausstattung bestritten werden.

Kann die Integration und Kuration von Forschungsdaten gefördert werden?

- Vorhaben oder Arbeitspakete, in denen es alleine um die Integration neuer Datensätze in eine Informationsinfrastruktur oder einzig um die Kuration von Daten geht, ohne dass diese im direkten Zusammenhang mit der (Weiter-)Entwicklung der Forschungsdateninfrastruktur stehen, sind von der Förderung ausgeschlossen.

Gestaltung des Antrags bzw. des Vorhabens**Wie kann das Leistungs- und Funktionsspektrum einer Forschungsdateninfrastruktur ausgestaltet sein?**

- Die Gestaltung des Leistungs- und Funktionsspektrums ist immer an den wissenschaftlichen Bedarfen auszurichten. Dabei verdeutlichen die nachfolgenden Beispiele verschiedene Aspekte innerhalb des Leistungs- und Funktionsspektrums und können bei der Planung einer Forschungsdateninfrastruktur als Orientierung dienen. Die Aufzählung ist nicht abschließend.
 - **Information und Planung** für das Management von Forschungsdaten vor, in und nach Forschungsprojekten (z.B. Planung des Datenmanagements, Erfassung des Datenlebenszyklus, Einhaltung von Förderrichtlinien, Informationsangebote, uvm.)
 - **Organisation und Aufbereitung** von Forschungsdaten (z.B. Definition eines Datenkurationsprofils, Rechte- und Zugangsmanagement, uvm.)
 - **Beschreibung und Dokumentation** von Forschungsdaten mit strukturierten Informationen (z.B. Nutzung von Metadaten und Metadatenstandards, kontrollierten Vokabularen, Normdaten, Ontologien, uvm.)

- **Speicherung und Anschlussfähigkeit** von Forschungsdaten für die Analyse (z.B. Wahl und Einrichtung der Speicherorte, Authentifikations- und Autorisierungsinfrastruktur, Strategien und Maßnahmen für die Datensicherheit und das Datenbackup, Planung von Schnittstellen zu Analyse- und Visualisierungssoftware sowie dem wissenschaftlichen Rechnen, uvm.)
- **Veröffentlichung und Archivierung** von Forschungsdaten (z.B. Publikation nach den FAIR-, CARE- oder FACT-Prinzipien in einem geeigneten Repositorium, Nutzung persistenter Identifikatoren, Maßnahmen für die Langzeitarchivierung, Wahl offener und maschinenlesbarer Dateiformate, uvm.)
- **Auffindbarkeit und Nachnutzung** von Forschungsdaten (z.B. Verzeichnisse von Repositorien, Metadatendienste und -indizes, Datenzitation, uvm.)
- **Rechte und Pflichten** für den Umgang mit Forschungsdaten (z.B. Klärung und Festlegung urheber- und datenschutzrechtlicher Spezifikationen für die Nachnutzung, die Zugriffsmöglichkeiten und die Publikation von Daten sowie Festlegung von Nutzungslizenzen, uvm.)
- **Ethik und gute wissenschaftliche Praxis** (z.B. Klärung und Einhaltung von Ethikrichtlinien entlang des Lebenszyklus von Daten, Erarbeitung und Umsetzung disziplinärer Richt- und Leitlinien zum Umgang mit Forschungsdaten, uvm.)

Wie misst sich der Erfolg eines Prototyps? Wie gestaltet sich der Übergang in eine weitere Entwicklungsphase?

- Mit der Entwicklung eines Prototyps wird die technische Machbarkeit einer avisierten Forschungsdateninfrastruktur aufgezeigt. Ein Ziel des Projekts ist es dabei, zu einer Beurteilung zu kommen, ob der Prototyp zu einem verlässlichen Dienst ausgebaut und danach betrieben werden kann und soll.
- Daher ist in Anträgen zu Prototypen darzulegen, nach welchen Kriterien die Eignung des Prototyps festgestellt werden kann. Es können Kriterien z.B. zur technischen Funktionalität (Skalierbarkeit, etc.), dem Nutzen für Anwenderinnen und Anwender (Stichwort Usability: Bedienfreundlichkeit, Passgenauigkeit, etc.) usw. formuliert werden. Anhand dieser Kriterien kann im Projektverlauf und zum Projektende die Evaluation des Prototyps erfolgen; dies wird im Zwischen- oder Abschlussbericht dokumentiert.
- Der Übergang in eine weitere Entwicklungsphase erfordert einen Fortsetzungsantrag, dem ein Zwischenbericht über die Evaluation des Prototyps beigelegt wird. Sind Prototyp und Folgeantrag im Begutachtungsverfahren überzeugend, so kann eine Weiterentwicklung gefördert werden.

Was ist eine Implementierung?

- Unter einer Implementierung wird im Kontext des Förderprogramms die Umsetzung eines Prototyps bzw. Softwareentwurfs in einen verlässlichen Dienst verstanden.

Was ist eine funktionelle Weiterentwicklung einer Informationsinfrastruktur?

- Unter einer funktionellen Weiterentwicklung ist die Erweiterung des Funktions- und Leistungsspektrums einer bestehenden Informationsinfrastruktur zu verstehen. Die Weiterentwicklungen sind an den sich verändernden Bedarfen aus den Wissenschaften ausgerichtet.

Was ist eine rein technische Neuerung einer Informationsinfrastruktur?

- Unter einer rein technischen Neuerung wird die Integration neuer bzw. aktualisierter Hardware- oder Softwarekomponenten (beispielsweise Softwareupdates) verstanden. Dies fällt unter die Wartung einer Informationsinfrastruktur und ist in der Regel Aufgabe des Betreibers. Die rein technische Neuerung ist nicht förderbar.

Was sind digitale Dienste im Bereich der Wissenschaft?

- Digitale Dienste sind Software und deren organisatorische Umgebung für die wissenschaftliche Nutzung. Die Nutzung eines Dienstes muss nicht mit Kosten verbunden sein.
- Je nach Fokus können digitale Dienste kategorisiert³ werden in:
 - Digitaler wissenschaftlicher Dienst: Informationstechnische Dienstleistung, die z.B. Umgebungen, Werkzeuge und Lösungskomponenten für die wissenschaftliche Arbeit von Forschenden und Forschungsgruppen bereitstellt.
 - Digitaler generischer Dienst: Infrastrukturell bereitgestellter Dienst z.B. für das Identitäts- und Berechtigungsmanagement, das Übertragen, Speichern, Verarbeiten, Teilen, Archivieren sowie Auffinden von Daten und Informationen.

Was sind Organisations- und Vernetzungsformen?

- Die Organisation einer Informationsinfrastruktur umfasst die Ausgestaltung sämtlicher Arbeitsabläufe, welche im Kontext des Auf- und Ausbaus, sowie des Betriebs einer Informationsinfrastruktur für Forschungsdaten anfallen. Bei der Konzeption einer Organisationsform werden Zuständigkeiten, Rechte und Pflichten rund um die Gestaltung und Nutzung der Informationsinfrastruktur definiert.
- Bei Vernetzungsformen geht es um die Einbindung der (zu entwickelnden) Informationsinfrastruktur für Forschungsdaten in die wissenschaftlichen Fachcommunities, und/oder in die Entwickler- bzw. Betreibercommunities. Hierbei können unterschiedliche Formate (Konferenzbeiträge, Workshops, Rundgespräche, Zeitschriftenartikel, Öffentlichkeitsarbeit, uvm.) genutzt werden, um die Etablierung der Informationsinfrastruktur zu befördern.

Was ist ein Betriebsmodell?

- Ein Betriebsmodell beschreibt Strukturen und Abläufe, die den nachhaltigen Betrieb der Informationsinfrastruktur ermöglichen. Entsprechend können hier Regelungen zur Nutzung der Informationsinfrastruktur getroffen werden (Umfang der Nutzung, für ggf. unterschiedliche Nutzergruppen, Rechte, Pflichten, ggf. Kosten, Datenclearingstelle, etc.). Auch organisatorische Abläufe und Zuständigkeiten können definiert werden (z.B. für die Beantragung eines Speicherkontingents, Wartungszyklen, Aufgabenverteilung zwischen verschiedenen am Projekt beteiligten Institutionen, etc.).
- Ein Betriebsmodell kann auch ein Geschäftsmodell beinhalten, bei dem finanzielle Aspekte wie beispielsweise Nutzungsgebühren geregelt werden.
- Die Entwicklung eines Betriebsmodells ist im Programm „Informationsinfrastrukturen für Forschungsdaten“ förderbar; dabei ist wichtig ist, dass auch im Falle von zu entwickelnden Betriebsmodellen alle Projektergebnisse unter einer freien Lizenz zur Nachnutzung zur Verfügung gestellt werden müssen.

Wie können Nachhaltigkeitskonzepte gestaltet sein?

- Ein Nachhaltigkeitskonzept soll passend auf die Entwicklungsphase der Informationsinfrastruktur abgestimmt sein. Je reifer ein Projekt ist, desto verbindlicher ist der langfristige Betrieb der Informationsinfrastruktur technisch, finanziell und organisatorisch abzusichern.
 - Die Ergebnisse aus jeder Entwicklungsphase müssen FAIR⁴ (findable, accessible, interoperable, reusable) sein. Bei der Entwicklung von Prototypen kann ggf. die Sicherung der Nachnutzbarkeit bereits als eine Form der Nachhaltigkeit anerkannt werden. Zu beachten ist hierbei, dass auch bei negativer Evaluierung eines Prototyps eine Verpflichtung besteht, die Projektergebnisse in wohl dokumentierter Form zu sichern und verfügbar zu machen.
 - Führt ein Projekt zum Betrieb einer Informationsinfrastruktur für Forschungsdaten, steht die Organisation des Dauerbetriebs im Fokus des Nachhaltigkeitskonzepts. Zentral ist

³ vgl. Digitale Dienste für die Wissenschaft – wohin geht die Reise?, DOI: [10.5281/zenodo.4301924](https://doi.org/10.5281/zenodo.4301924), Seite 7

⁴ The FAIR Guiding Principles for scientific data management and stewardship, DOI: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)

hierbei ein langfristig tragfähiges Konzept, das den Betrieb der Informationsinfrastruktur sicherstellt. Dabei sind unterschiedliche Organisationsformen denkbar. Neben der Verantwortung der antragstellenden Organisationen ist auch die Übergabe der Verantwortung an andere Institutionen oder Organisationen (wie beispielsweise NFDI-Konsortien) denkbar, sofern diese die Organisation des Dauerbetriebs verbindlich zusagen.

- Für die kontinuierliche Weiterentwicklung einer Informationsinfrastruktur kann der Aufbau einer Entwicklergemeinschaft angestrebt werden.
- Sollte sich im Projektverlauf ergeben, dass ein Dienst nur für eine überschaubare Zeit in Betrieb genommen wird, ist ein End-of-life Konzept zu erarbeiten.
- Zur Sicherung der Nachhaltigkeit sollen sogenannte Lock-In-Effekte, also Abhängigkeiten von kommerziellen Anbietern, vermieden werden.

Was ist ein Datenkurationsprofil?

- Ziel der Datenkuration ist es, kohärente, reproduzierbare und nachnutzbare Datensätze zusammenzustellen.
- Das Datenkurationsprofil beschreibt die Kriterien zur Auswahl von Datensätzen (z.B. Qualität, Inhalt, Format, Metadaten, Auszeichnungstiefe, etc.) und definiert einen Beschreibungsstandard für diese Daten (Metadatenschema, Normdaten, Persistent Identifier, Auszeichnungsformat, etc.).

Einbindung externer Dienstleistungen in das Projekt

- Werden externe Dienstleistungen im Projekt benötigt, so müssen bei der Antragstellung zwei Vergleichsangebote vorgelegt werden.
- Grundsätzlich wird empfohlen, die technische Expertise zum Aufbau bzw. der Weiterentwicklung von Informationsinfrastrukturen in den Einrichtungen der Antragstellenden zu verankern.
- Sollen Arbeiten an Dritte vergeben werden, so ist zu beachten, dass Werkverträge seitens der DFG zustimmungspflichtig sind. Näheres regeln die Verwendungsrichtlinien⁵.

Lizensierung von Daten und Software

- Die Lizenzierung der Informationsinfrastruktursoftware wird durch das Programmmerkblatt geregelt. Hierbei sind open-source Lizenzen zu wählen, die eine kostenlose Nachnutzung durch Dritte ermöglichen.
- Die Lizenzierung von Forschungsdaten wird als so frei und offen wie möglich empfohlen.

Nach dem Einreichen eines Antrags

Nach welchen Kriterien werden Anträge in diesem Programm begutachtet?

- Bei der Begutachtung werden die allgemeinen Begutachungskriterien für LIS Programme angewandt. Diese sind auf der DFG-Website veröffentlicht⁶.

Wie ist der Abschlussbericht zu gestalten?

- Der Abschlussbericht orientiert sich am Leitfaden für Projektberichte⁷ im Bereich „Wissenschaftliche Literaturversorgungs- und Informationssysteme“. Bitte gehen Sie insbesondere auf das Thema „Nachhaltigkeit“ ein.

⁵ Verwendungsrichtlinien der Deutschen Forschungsgemeinschaft: https://www.dfg.de/formulare/2_00/

⁶ http://www.dfg.de/formulare/12_107/

⁷ https://www.dfg.de/formulare/12_02/