

WEBINAR

Forschungsdatenmanagement in der
Ost-, Ostmittel- und Südosteuropaforschung

OstData: Forschungsdatendienst für die Ost-, Ostmittel- und Südosteuropaforschung

- gefördert durch DFG (1. Projektphase: 2019-2022; 2. Förderphase geplant)
- Aufbau und Konzeption einer modularen und netzwerkartigen technischen Infrastruktur zur
 - **Speicherung**
 - **Veröffentlichung**
 - **Langzeitarchivierung**
 - **Suche**von Forschungsdaten.
- Vielfältige Beratungsleistungen für die wissenschaftliche Community
 - Workshops, [Webinare](#) und Vorträge
 - Handreichungen und Informationsmaterial (über [Osmikon](#) & [Zenodo](#) zugänglich)
 - Einwerbungsstrategien für die Publikation von Forschungsdaten

BSB Bayerische
Staatsbibliothek
Information in erster Linie

 COLLEGIUM CAROLINUM

 **IOS** Leibniz-Institut für
**OST- UND SÜDOST-
EUROPAFORSCHUNG**

OSTDATA

 **Ówzo** Leibniz-Institut für
Geschichte und Kultur
des östlichen Europa

 **HERDER-INSTITUT**
für historische Ostmitteleuropaforschung
INSTITUT DER LEIBNIZ-GEMEINSCHAFT

Module

① Open Access

- Open Access – Open Science
- Open Access – Forschungsdaten

② Forschungsdaten

- Nachvollziehbarkeit und
Nachnutzbarkeit von Forschung
- Publikation

③ Forschungsdaten- management

- Notwendigkeit und Fachspezifik
- Forschungsdatenlebenszyklus

④ Datendokumentation

- Datenmanagementplan
- OstData-Metadatenchema

⑤ Beratungsangebote

MODUL ①

Open Access

OSTDATA

Open Access und Open Science

- Open Science als Bewegung hin zu mehr Transparenz in Forschung und Wissenschaftskommunikation zwecks bestmöglicher Verbreitung wissenschaftlicher Informationen
- Forderung nach Open Access, also **freiem, uneingeschränktem und kostenlosen Zugang** zu wissenschaftlicher Literatur und anderen Materialien im Internet
- Open Data als weiteres Kernprinzip, das für Forschungsdaten fordert, dass diese ebenfalls **frei zugänglich** veröffentlicht werden sollen
- Aber: rechtliche oder ethische Gründe können auch unterschiedliche Arten von Zugangsbeschränkungen zur Folge haben, z. B. nur Zugang zu Metadaten oder ein eingeschränkter Zugriff*

* vgl. z. B. das Controlled Vocabulary for Access Rights: http://vocabularies.coar-repositories.org/documentation/access_rights/ mit den Kategorien open access, embargoed access, restricted access, metadata only access, letzter Zugriff am 06.01.2021.

Open Access und Forschungsdaten*

Die **Allianz der Deutschen Wissenschaftsorganisationen** spricht sich für einen offenen Zugang zu Forschungsdaten aus. Für einen offenen Zugang sprechen:

- herstellbare Transparenz
- Qualitätssicherung durch prinzipielle Reproduzierbarkeit der Forschung (**Nachvollziehbarkeit**)
- Effizienzsteigerung und Kostenersparnis durch mögliche Sekundäranalysen (**Nachnutzbarkeit**)

Forschungsdaten in aktueller wissenschaftspolitischer Perspektive:

- „Datenstrategie der Bundesregierung“ (01/2021)
- Etablierung einer „Nationalen Forschungsdaten Infrastruktur“ NFDI
- DFG:
 - Fachspezifische Leitlinien zum Umgang mit Forschungsdaten
 - „Leitlinien zur Sicherung guter wissenschaftlicher Praxis“, Leitlinien 15-17
 - langfristige Sicherung und öffentliche Zugänglichkeit von Forschungsdaten und –ergebnissen
 - Institutionen müssen infrastrukturelle Grundlagen zur Verfügung stellen, z. B. Anbindung an fachspezifische Repositorien
- BMBF: „Aktionsplan Forschungsdaten“ (10/2020)
- Horizon 2020: Erstellung eines Datenmanagementplans verpflichtend

* <https://open-access.net/informationen-zu-open-access/open-access-bei-forschungsdaten>, letzter Zugriff am 06.01.2021.

FAIRe Forschungsdaten

Auffindbar
(**F**indable)



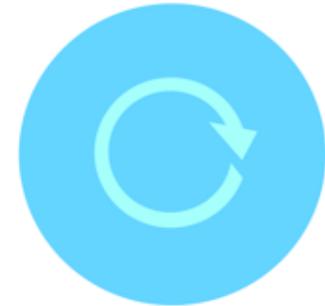
Zugänglich
(**A**ccessible)



Interoperabel
(**I**nteroperable)



Wiederverwendbar
(**R**eusable)



MODUL ②

Forschungsdaten

OSTDATA

Input

Welche Daten sehen Sie
als (Ihre) Forschungsdaten an?

Was sind Forschungsdaten?

Forschungsdaten stellen all jene Daten dar, die auf Grundlage wissenschaftlicher Methoden gesammelt, erzeugt, entwickelt, beschrieben und/oder ausgewertet sowie angemessen dokumentiert werden.*

- Ton- und Videoaufnahmen von Interviews
- unmittelbare digitale Repräsentationsformen kultureller Objekte (z. B. Digitalisate)
- Transkripte von Texten
- Sprachaufnahmen
- Annotationen und Exzerpte
- Bibliografien, Taxonomien, Ontologien, kontrollierte Vokabularien, Normdaten
- Mess-, Umfrage- und Erhebungsdaten (z. B. aus empirischer Forschung)
- Text- und Sprachkorpora
- Auswertungen quantitativer Sozialforschung
- Computerskripte und Quellcode
- Abbildungen von physischen Objekten, Grafiken oder Diagrammen
- Netzwerkanalysen und -visualisierungen
- Geodaten
- kritische Editionen von Quellen und Texten
- digitale Forschungsumgebungen
- traditionelle und neue Formen der Ergebnispräsentation (Blogserien, komplexe Narrative mit Multimedia- und Hypertextelementen)
- Algorithmen und fachspezifische Softwarewerkzeuge

* entwickelt anhand von Puhl, Johanna; Andorfer, Peter; Höckendorff, Mareike; Schmunk, Stefan; Stiller, Juliane, Thoden, Klaus: „Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften“, DARIAH-DE Working Papers 11, Göttingen 2015, S. 12-14 (urn:nbn:de:gbv:7-dariah-2015-4-4), letzter Zugriff am 11.01.2021.

Publikation von Forschungsdaten

Anforderungen an Forschungsdaten für Publikation:

- Sicherstellung langfristiger Auffindbarkeit durch dauerhafte Adressen (z. B. Persistent Identifiers (PID) wie Digital Object Identifier (DOI), persistente URLs)
- Beschreibung von Daten und Datensammlungen durch **Metadaten**:
 - Erfassung eines Datensatzes über **standardisiertes Metadatenschema**
 - Eingabe **dateiinterner Metadaten** zu Autoren, beteiligten Personen/Institutionen, Datenprovenienz
 - Empfehlung: Verwendung von Dateiformaten, die in der jeweiligen Disziplin **gängig** sind, **Langzeitarchivierung** unterstützen und **möglichst frei und anbieterunabhängig** genutzt werden können (z. B. .tiff, .pdf)

Die **Open Access-Publikation von Forschungsdaten** steht unter spezifischen **Anforderungen**, die vorab ausgelotet werden müssen. Daten können...

- daten-, urheber- oder personenschutzrechtlich geschützt sein
- durch Speicherung in einer anderen Datenbank Rechte Dritter betreffen
- durch heterogene Speicher- und Dateiformate nicht die erforderlichen Qualitätsstandards erfüllen

Forschungsdaten

Forschungsdaten sind, grob gefasst...

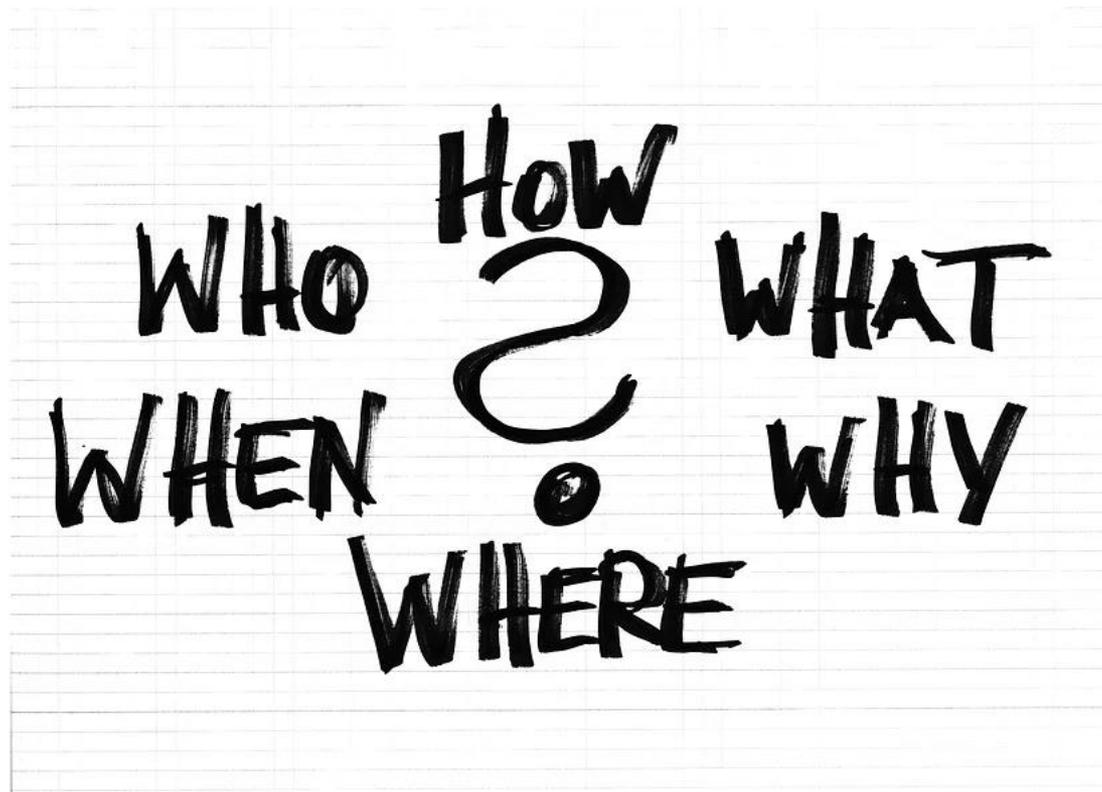
Daten von und für die Forschung

Forschungsdatenmanagement zielt auf die

Nachvollziehbarkeit und **Nachnutzbarkeit**

von Forschung, Forschungsprozessen und -ergebnissen

Fragen

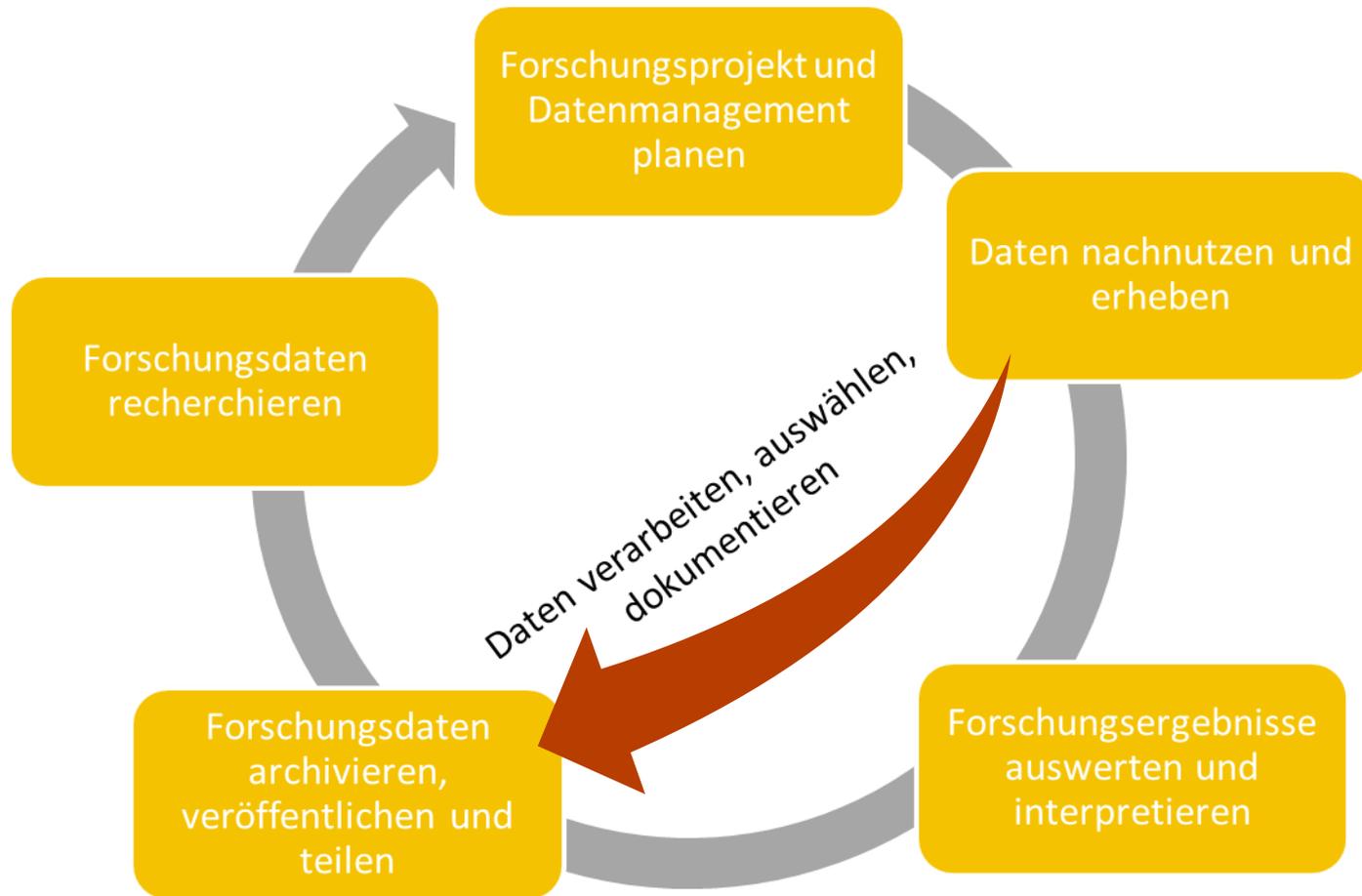


MODUL ③

Forschungsdatenmanagement

OSTDATA

Forschungsdatenlebenszyklus



Was ist Forschungsdatenmanagement (FDM)?

Rat für Informationsinfrastrukturen (2016):
Das **Forschungsdatenmanagement** umfasst alle [...] Maßnahmen, die getroffen werden müssen, um **qualitätsvolle Daten** zu gewinnen, um die **gute wissenschaftliche Praxis im Datenlebenszyklus** einzuhalten, um Ergebnisse **reproduzierbar** und Daten zur **Nachnutzung verfügbar** zu machen und um ggf. bestehenden **Dokumentationsverpflichtungen** (z. B. im Gesundheitswesen) Rechnung zu tragen.*

* Rat für Informationsinfrastrukturen: „Bericht des Redaktionsausschusses Begriffe an den RfII, RfII Berichte 1, Göttingen 2016, S. 11 (urn:nbn:de:101:1-201607146410), letzter Zugriff am 11.01.2021.

FDM – Notwendigkeit und Fachspezifik

Warum ist FDM (plötzlich) so wichtig?

- Forschung ging schon immer mit Datenmanagement einher (z. B. Verwaltung von Archivalien, Scans) → Datenmengen steigen
- ‚Krise der **Reproduzierbarkeit**‘ wissenschaftlicher Studien
- **Nachvollziehbarkeit** von Forschung → Auswirkungen gefälschter Daten sind z. T. gravierend
- Vermeidung doppelter Arbeit → stattdessen Aufbau auf und **Nachnutzung** von Daten
- FDM als transdisziplinäre Aufgabe

FDM in den Geisteswissenschaften

Große Datenbestände entstehen seit Jahrzehnten. Die Digitalisierung verstärkt das Wachstum. Es entsteht die Notwendigkeit, Daten aufzubereiten, zu archivieren und für eine Nachnutzung bereitzustellen.

- **Reproduzierbarkeit**: es wird erwartet, dass gleiche Methoden, angewandt auf weitere Untersuchungsgegenstände, unterschiedliche Resultate ergeben
 - FDM: Prozess der Untersuchung nachvollziehbar machen
- **Nachvollziehbarkeit** von Forschung: Zugang zu Quellen kann z. T. erschwert sein (Archivgut)
 - FDM: Untersuchungsgegenstände zugänglich machen
- **Nachnutzung** generieren: Entwicklung innovativer Fragestellungen auf der Basis erschlossener Daten, z. B. umfänglicher Sprachkorpora oder Quellensammlungen

FDM – Gute Argumente und neue Hürden

An sich eine gute Sache...*	...aber
<ul style="list-style-type: none"> – eventl. Auflagen zur Publikation von Ergebnissen werden durch FDM von Beginn an berücksichtigt 	<ul style="list-style-type: none"> – FDM erfordert durch Planung zu Projektbeginn mehr Zeit
<ul style="list-style-type: none"> – Einarbeitung in Projekte wird durch organisiertes FDM erleichtert und spart Zeit 	<ul style="list-style-type: none"> – kontinuierliche Dokumentation sowie Beschreibung größerer Datenmengen mit (dateiinternen) Metadaten kann sehr arbeitsintensiv sein (u. a. durch Einarbeitung in Software)
<ul style="list-style-type: none"> – fördernde Institutionen können bei Reviewprozessen Daten anfragen, die durch FDM organisiert vorliegen 	<ul style="list-style-type: none"> – werden am Projektende keine Daten publiziert, kann der Eindruck entstehen, der Mehraufwand für FDM stehe in keinem Verhältnis zur Forschungsarbeit
<ul style="list-style-type: none"> – standardisierte FDM-Prozesse ermöglichen eine leichtere Nachnutzung hausinterner und extern publizierter Daten 	<ul style="list-style-type: none"> – gutes FDM kann die Kosten für ein Forschungsprojekt steigern, z. B. durch zusätzliche Personalmittel für die Dokumentation, Rechtsberatung oder Kosten für die technische Infrastruktur
<ul style="list-style-type: none"> – Risiko eines Datenverlusts wird durch FDM-Maßnahmen minimiert, z. B. durch Datendokumentation, Datensicherung und eine geeignete Langzeitarchivierung (LZA) 	<ul style="list-style-type: none"> – Forschende müssen bereit sein, sich mit FDM auseinanderzusetzen und sich ggf. in Metadatenstandards und Lizenzfragen einzuarbeiten

* <https://www.forschungsdaten.info/themen/informieren-und-planen/wie-beginnt-forschungsdatenmanagement/>, letzter Zugriff am 12.01.2021.

MODUL ④

Datendokumentation

OSTDATA

OstData – Metadatenschema

- basiert auf dem **DataCite**-Metadatenschema:
<https://schema.datacite.org/>
- Entscheidung, DataCite-Metadatenschema zu erweitern
 - OstData als **inter- und transdisziplinäres Repositorium** mit diversen und vielfältigen Forschungsdatentypen
 - Bedarf, Forschungsdaten detaillierter beschreiben zu können
 - besseres und schnelleres Verständnis bei Recherche
 - Nachnutzung fördern
 - Forschende wünschen sich, die Provenienz von Forschungsdaten genauer beschreiben bzw. einsehen zu können
- **Erweiterung des DataCite-Metadatenschemas in OstData:**
 - Fachspezifische Zonen
 - Differenziertere Sacherschließung
 - Erweiterte Angaben zur Datenprovenienz

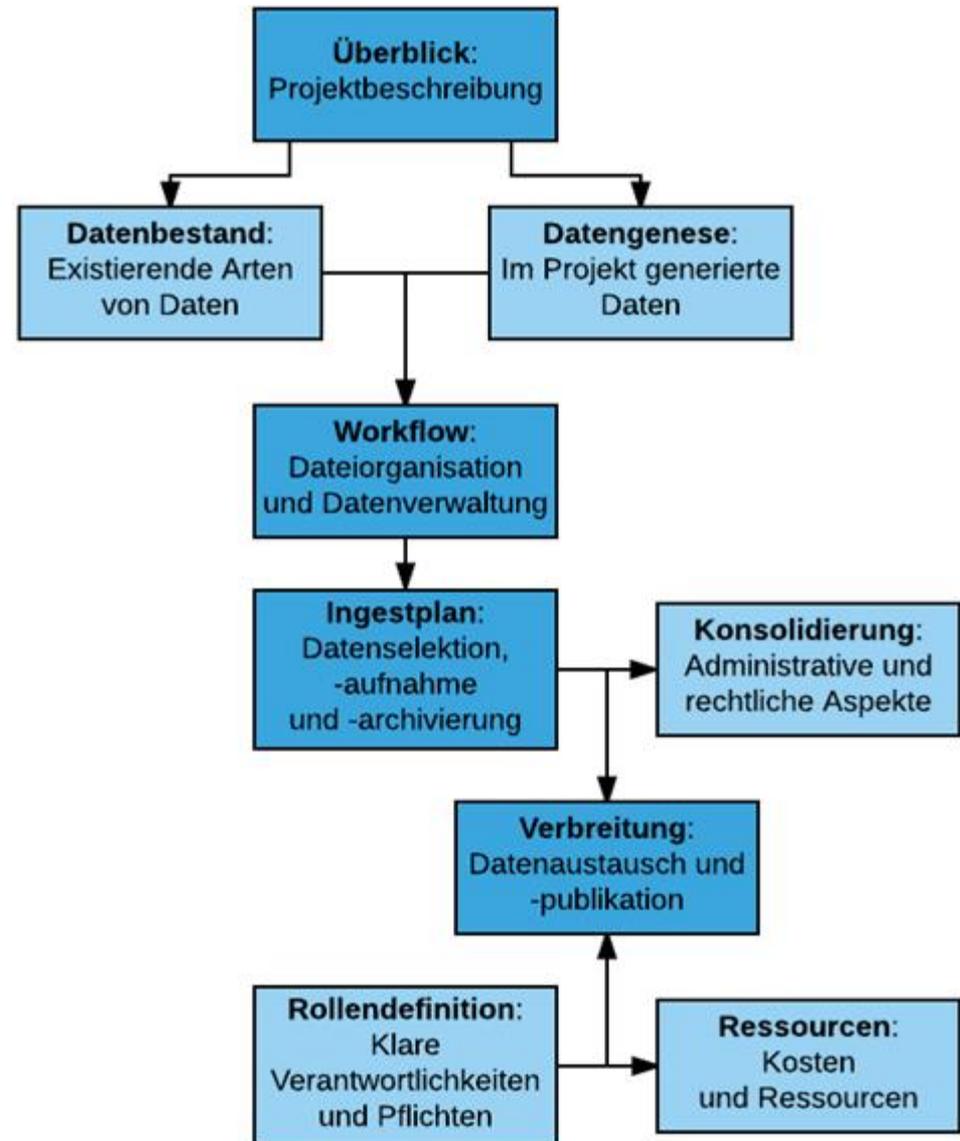
Datenmanagementplan (DMP)

Datenmanagementpläne unterstützen die systematische Ablage, Benennung und Speicherung von Daten in einem Forschungsprojekt.* Ein Datenmanagementplan sollte bereits **vor Antragstellung/Planung** eines Projekts erstellt und als **„lebendes“ Dokument** mindestens jährlich aktualisiert werden.

- Erstellung eines DMP ist für FDM-Neulinge mit hohem zeitlichem Aufwand verbunden
 - allerdings: zeitlicher Aufwand lohnt sich durch effektivere Strukturierung, Dokumentation und Organisation von Daten
 - sehr viele Beratungsangebote zur DMP-Erstellung und Pflege online
 - **OstData** will Angebote schaffen, die **bedarfsgerecht auf Zielgruppe** zugeschnitten sind

* <https://www.forschungsdaten.info/themen/informieren-und-planen/datenmanagementplan/>, letzter Zugriff am 12.01.2021.

Aufbau und Struktur eines DMP



Datenmanagementplan von OstData

- Aufbau und Struktur des OstData-DMP:
 - **Metadaten** (u. a. Bearbeitende, Versionierung)
 - Allgemeine **Angaben zum Forschungsprojekt**
 - **Inhaltliche Beschreibung** von Forschungsdaten (u. a. Forschungsfrage, Sacherschließung)
 - **Umgang mit Daten** während des Forschungsprozesses (u. a. Speicherorte, -kapazitäten, kollaboratives Arbeiten)
 - **Rechtliche und ethische Aspekte** bei der Erhebung von Forschungsdaten (u. a. Personenschutzrechte, Pseudo- und Anonymisierung)
 - **Veröffentlichung** von Forschungsdaten (u. a. Lizenzen, Urheberrechte)

Datenmanagementplan – Argumente*

Guter Plan...	...aber
<ul style="list-style-type: none"> – strukturiertes Datenmanagement als Gewinn für alle am Projekt Beteiligten (z. B. bei Einarbeitungen, Übergaben) 	<ul style="list-style-type: none"> – Erstellung eines DMP erfordert v. a. bei Erstkontakt erhöhte zeitliche Kapazitäten
<ul style="list-style-type: none"> – eröffnet Zugang zu bestimmten Förderlinien 	<ul style="list-style-type: none"> – Pflege eines DMP während der Projektlaufzeit erfordert zeitliche, technische und personelle Kapazitäten, die von Beginn an eingeplant werden müssen (z. B. durch Stellenanteile oder entsprechende technische Infrastruktur)
<ul style="list-style-type: none"> – erleichtert die Dokumentation bei Berichtspflicht 	<ul style="list-style-type: none"> – Dateiformate, die für eine Nachnutzung in Frage kommen, müssen ggf. konvertiert werden, um Anforderungen des Datenmanagements zu genügen
<ul style="list-style-type: none"> – vereinfacht die Nachnutzung von Forschungsdaten für die eigene Forschung durch klare Strukturierung 	
<ul style="list-style-type: none"> – schafft organisierten Zugang zu Nachnutzungsmöglichkeiten für Dritte 	
<ul style="list-style-type: none"> – Publikation von Forschungsdaten gilt als eigenständige Publikation (vgl. OstData-Zitierhinweis) 	
<ul style="list-style-type: none"> – verringert die Gefahr von Datenverlust 	
<ul style="list-style-type: none"> – erhöht die Chance, dass Datenträger und Dateiformate nach zehn Jahren noch lesbar sind 	

* <https://www.forschungsdaten.info/themen/informieren-und-planen/datenmanagementplan/>, letzter Zugriff am 08.03.2021.

MODUL ⑤

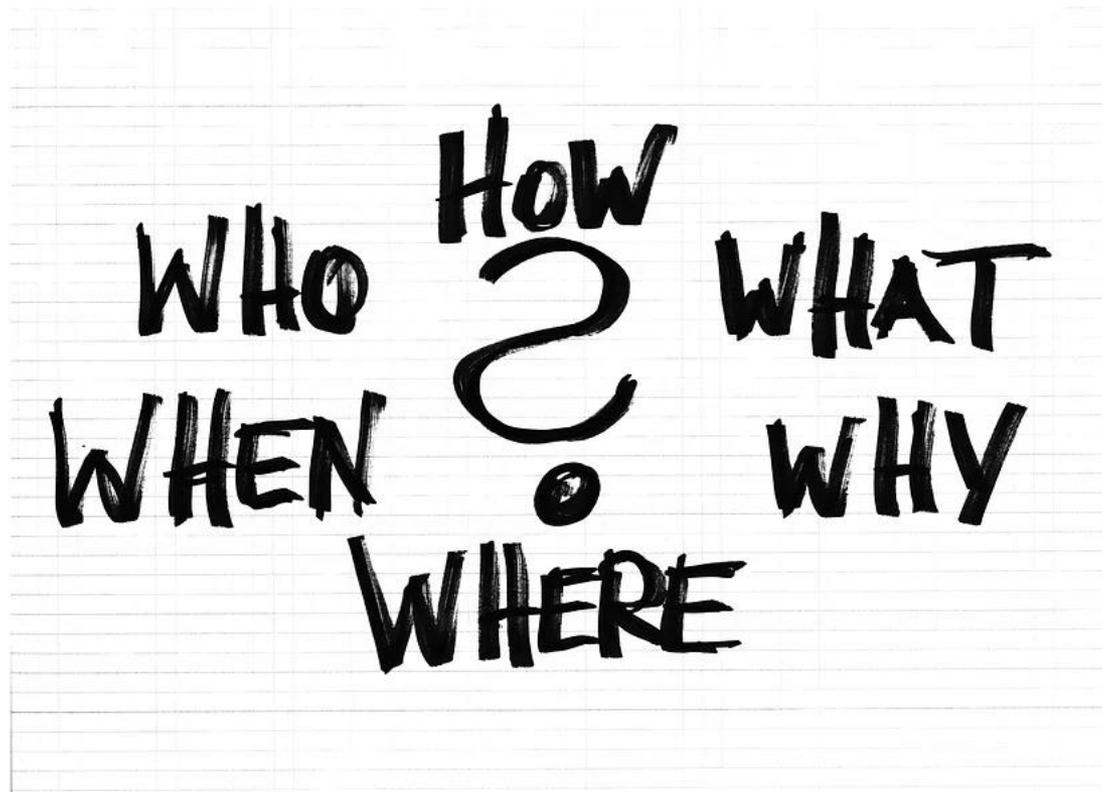
Beratungsangebote

OSTDATA

Beratung durch OstData

- **Workshops, Webinare und Vorträge** zum Forschungsdatenmanagement:
<https://www.youtube.com/watch?v=2IDsHNIFg7E>
- Einstiegsinformationen zu Forschungsdaten- und Forschungsdatenmanagement in der Ost-, Ostmittel- und Südosteuropaforschung über **Osmikon**:
<https://www.osmikon.de/forschungsdaten>
- Handreichungen und Informationsmaterial über die **OstData-community** auf **Zenodo**: <https://zenodo.org/communities/ostdata/>
- Einzelberatungen für Forschende und Forschungsprojekte
- Kooperationen mit **Fachinformationsdiensten** und **NFDI-Konsortien**

Fragen



Diskussion und Fragen



Lizenzhinweis

Die Präsentation ist lizenziert unter

Creativ Commons Namensnennung 4.0 International (CC BY 4.0)



Teile der Präsentation gehen auf unpublizierte Arbeiten aus dem Projekt „Strategische Weiterentwicklung des Forschungsdatenmanagements am Herder-Institut für historische Ostmitteleuropaforschung (FDMHerder)“ zurück (2017-2019, BMBF; [zur Projektwebsite](#)).