



Forschungsdatenmanagement: Schikane oder (persönlicher) Vorteil?

12.01.2023

Was sind digitale Forschungsdaten?

- Keine feste Definition!
- Allgemein:
„alle digital vorliegenden Daten, die während des Forschungsprozesses entstehen oder ihr Ergebnis sind“¹

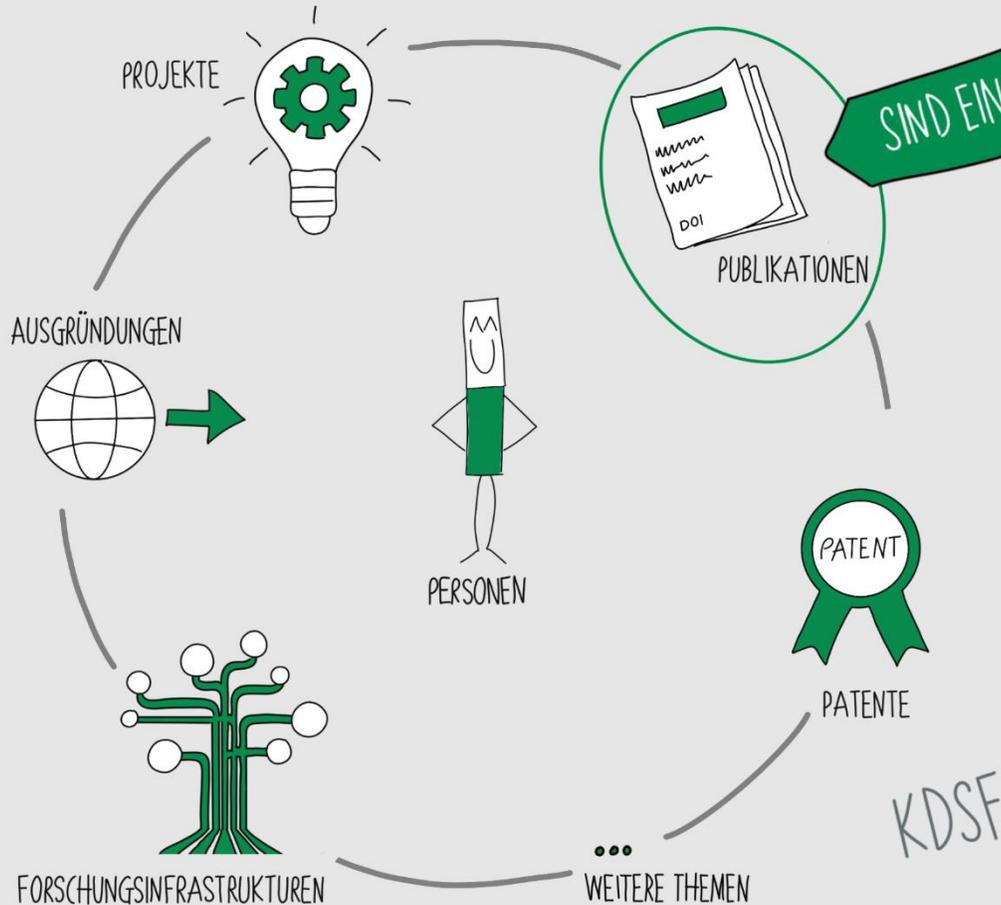
FORSCHUNGSINFORMATIONEN

BESCHREIBEN FORSCHUNG

INFORMATION

ADMINISTRATION

VERNETZUNG



& IHRE SICHT AUF

FORSCHUNGSDATEN

ENTSTEHEN WÄHREND/SIND GRUNDLAGE VON FORSCHUNG

FORSCHUNG

ENTWICKLUNG



MESSDATEN

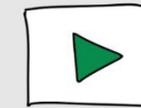


UMFRAGEN

CODE



VERSUCHE



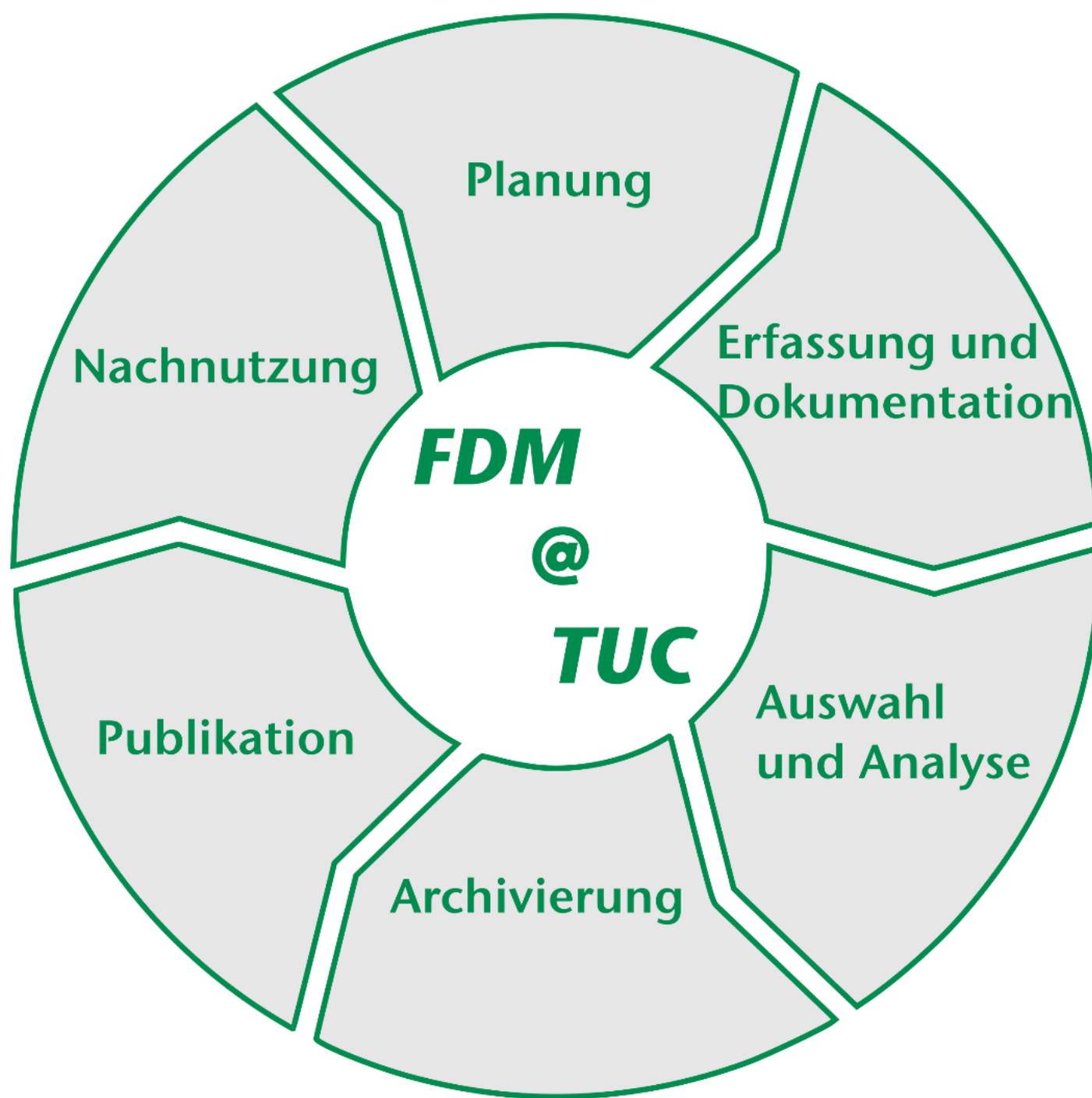
MULTIMEDIA

WEITERE TYPEN

Forschungsdatenmanagement (FDM) umfasst alle Aktivitäten, die mit → **Aufbereitung** → **Speicherung** → **Archivierung** und → **Veröffentlichung** von Forschungsdaten verbunden sind.

FDM begleitet den Forschungsprozess von den ersten Planungen bis zur Archivierung, Nachnutzung oder Löschung der Daten.

Quelle: Maxi Kindling, Peter Schirnbacher, Elena Simukovic: *Forschungsdatenmanagement an Hochschulen: das Beispiel der Humboldt-Universität zu Berlin*. LIBREAS. *Library Ideas*, 23 (2013). Online verfügbar unter: <http://libreas.eu/ausgabe23/07kindling/> DOI: 10.18452/9041

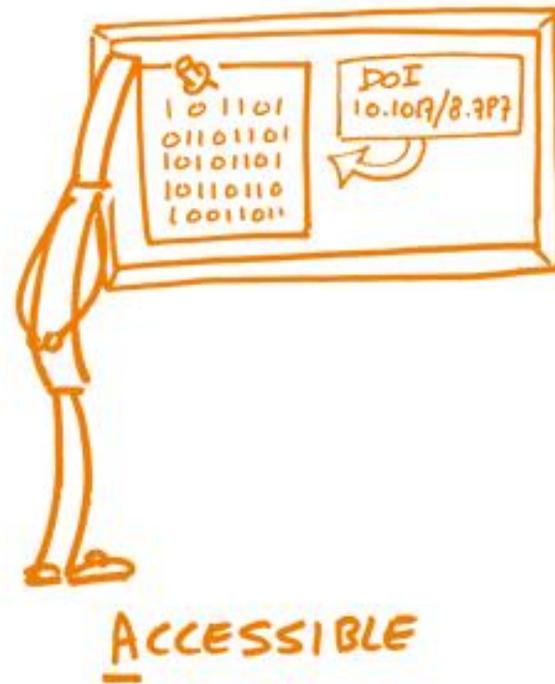




“As a scientist, you should
treat your data like a love letter
to your future self.”

Lambert Heller, Leiter Open Science Lab der TIB

FAIR DATA PRINCIPLES



[Link zum „Volltext“](#)

Assignment The Contractor shall not
Agreement, or delegate the performance
the prior written consent of the Committee



SIGNATURE

Forschungsdaten-Policies

Arten von Policies:

- *Zeitschriften- oder Verlags-Policies*
- *Förderrichtlinien*
- *Fachspezifische Policies*
- *Institutionelle Policies*





Forschungsdaten-Leitlinie der Technischen Universität Clausthal

Präambel

Forschungsdaten sind Grundlage und Ergebnis der wissenschaftlichen Arbeit. Das Management von Forschungsdaten muss daher nach anerkannten Standards erfolgen und hohen Anforderungen genügen. Nachvollziehbarkeit und Überprüfbarkeit der Forschung sowie die möglichst breite und langfristige Nachnutzung der Daten sind dabei gleichberechtigte Ziele. Dies sind zentrale Aspekte guter wissenschaftlicher Praxis. Die Universität fördert und unterstützt den freien Zugang zu Forschungsdaten gemäß ihrer Open Access Policy. Rechtliche und ethische Verpflichtungen sind zu beachten.

Definitionen

1. Forschungsdaten sind Daten, die im Zuge der Forschung gesammelt, beobachtet, simuliert, abgeleitet oder generiert werden.
2. Das Management von Forschungsdaten umfasst deren Planung, Erfassung, Verarbeitung, Dokumentation, Aufbewahrung und gegebenenfalls planmäßige Löschung. Es sichert den Zugang durch die Festlegung und Umsetzung von Zugangs- und Nutzungsregeln. Es ermöglicht die Nachnutzbarkeit, Reproduzierbarkeit und Qualitätssicherung aller Forschungsdaten, die wissenschaftlichen Ergebnissen zugrunde liegen.

Grundsätze

3. Die Projektleiterinnen und Projektleiter sind für das Forschungsdatenmanagement ihrer Forschungsvorhaben verantwortlich. Sie sind insbesondere verpflichtet, die Einhaltung der guten wissenschaftlichen Praxis und der Fachstandards sicherzustellen.
4. Forschungsprojekte mit Forschungsdaten erfordern einen Datenmanagementplan. Dieser benennt u.a. Verantwortlichkeiten für die Datenpflege, beschreibt Art und Umfang der zu erwartenden Daten sowie Zugangsrechte und -vorbehalte. Die Universität rät, den Plan vor Beginn des wissenschaftlichen Projektes zu erstellen und dem jeweils aktuellen Verlauf des Vorhabens anzupassen. Die Universität empfiehlt instituts- und/oder projekinterne Richtlinien auszuarbeiten, die detaillierte Regeln zum Umgang mit Forschungsdaten enthalten.
5. Die Speicherung und Archivierung digitaler Forschungsdaten erfolgt in der IT- und Informationsinfrastruktur der Universität oder in geeigneten und vertrauenswürdigen externen Repositorien.
6. Die Universität und ihre Wissenschaftlerinnen und Wissenschaftler beachten beim Forschungsdatenmanagement ethische, datenschutz- und urheberrechtliche oder geheimhaltungswürdige Belange.
7. Bei einer Übertragung von Nachnutzungs- oder Veröffentlichungsrechten soll darauf geachtet werden, dass die Daten für wissenschaftliche Zwecke frei verfügbar bleiben.
8. Die Universität verankert die Bedeutung eines strukturierten Forschungsdatenmanagements in der Lehre und bei der Betreuung des wissenschaftlichen Nachwuchses.

Unterstützung

9. Die Universität berät beim Forschungsdatenmanagement in Forschungsvorhaben von der Planung, über die Durchführung bis über das Vorhabensende hinaus und bietet geeignete Aus- und Fortbildung an.
10. Die Universität sorgt für eine Grundausstattung an Forschungsdateninfrastruktur und stellt damit eine angemessene Aufbewahrung und die technische Verfügbarkeit von digitalen Forschungsdaten sicher. Dies geschieht vor Ort und durch Beteiligung an standortübergreifenden Verbundlösungen. Spezifische Anforderungen sind abzustimmen und ggf. zusätzlich zu finanzieren.

Verabschiedet durch den Senat der Technischen Universität Clausthal am 17.12.2019

TO DO

1. WAKE UP

2. COFFEE

3. THE REST...

Datenmanagementpläne

Enthält alle Informationen, die die Sammlung, Aufbereitung, Speicherung, Archivierung und Veröffentlichung von Forschungsdaten im Rahmen eines Forschungsprojekts hinreichend beschreiben und dokumentieren

→ „[...] Analyse des Workflows von der Erzeugung der Daten bis zu deren Nutzung“²

Anforderungen der Förderorganisationen

Förderer	Forderung	Abgabe bei Antrag	Inhalt	Updates
EC Horizon Europe	DMP	Nein. Erster Plan innerhalb der ersten 6 Projektmonate	Inhalte des Horizon Europe Template	Update, falls signifikante Änderungen auftreten sowie zum Projektende
DFG	Angaben zum Umgang mit Forschungs-daten	Ja	Inhalte der Leitlinie zum Umgang mit Forschungsdaten	Nein
BMBF	Plan manchmal erforderlich, abhängig vom Programm	Falls notwendig, ja	Inhalt hängt vom jeweiligen Programm ab; Bildungsforschung: Checkliste	Kommt auf das Programm an
VWStiftung	DMP	Ja	Inhalte des Science Europe Template	Nein

RDMO

Das Werkzeug zur Unterstützung von Planung und Umsetzung des Forschungsdatenmanagements an der TU Clausthal

[Bilderquelle: Christopher Bala](#) / CC BY 4.0

Willkommen bei RDMO

Als Mitglied der TU Clausthal können Sie den Webservice RDMO nutzen, um Ihr Forschungsdatenmanagement strukturiert zu planen und durchzuführen. Erfassen und verwalten Sie alle relevanten Planungsinformationen und alle Datenmanagementaufgaben über den gesamten Datenlebenszyklus in einem (anpassbaren) Dokument – dem Datenmanagementplan (DMP). Im RDMO können Sie mithilfe verschiedener Vorlagen einen DMP für Ihr Projekt erstellen.

Diese Instanz von RDMO ist ein System, das sich in der Entwicklung befindet. Wir bitten, etwaige noch vorhandene Einschränkungen in der Nutzbarkeit zu entschuldigen. Bei Fragen oder Problemen wenden Sie sich gerne an die Servicestelle Forschungsdatenmanagement.

Diese Videos helfen Ihnen für einen Einstieg in RDMO

[Erste Schritte](#)

[Was kann man mit RDMO machen?](#)

[RDMO Funktionen im Überblick](#)

Anmelden

[Mit RZ-Account anmelden](#)

Über RDMO

Dieser Webservice unterstützt alle Mitglieder und Angehörigen der TU Clausthal beim Erstellen eines Datenmanagementplans. Die zugrundeliegende Software stammt vom DFG-geförderten Projekt RDMO.

Kontakt

Servicestelle
Forschungsdatenmanagement
Florian Strauß
fdm@tu-clausthal.de

Impressum

[Nutzungsbedingungen](#)
[Privacy Policy](#)



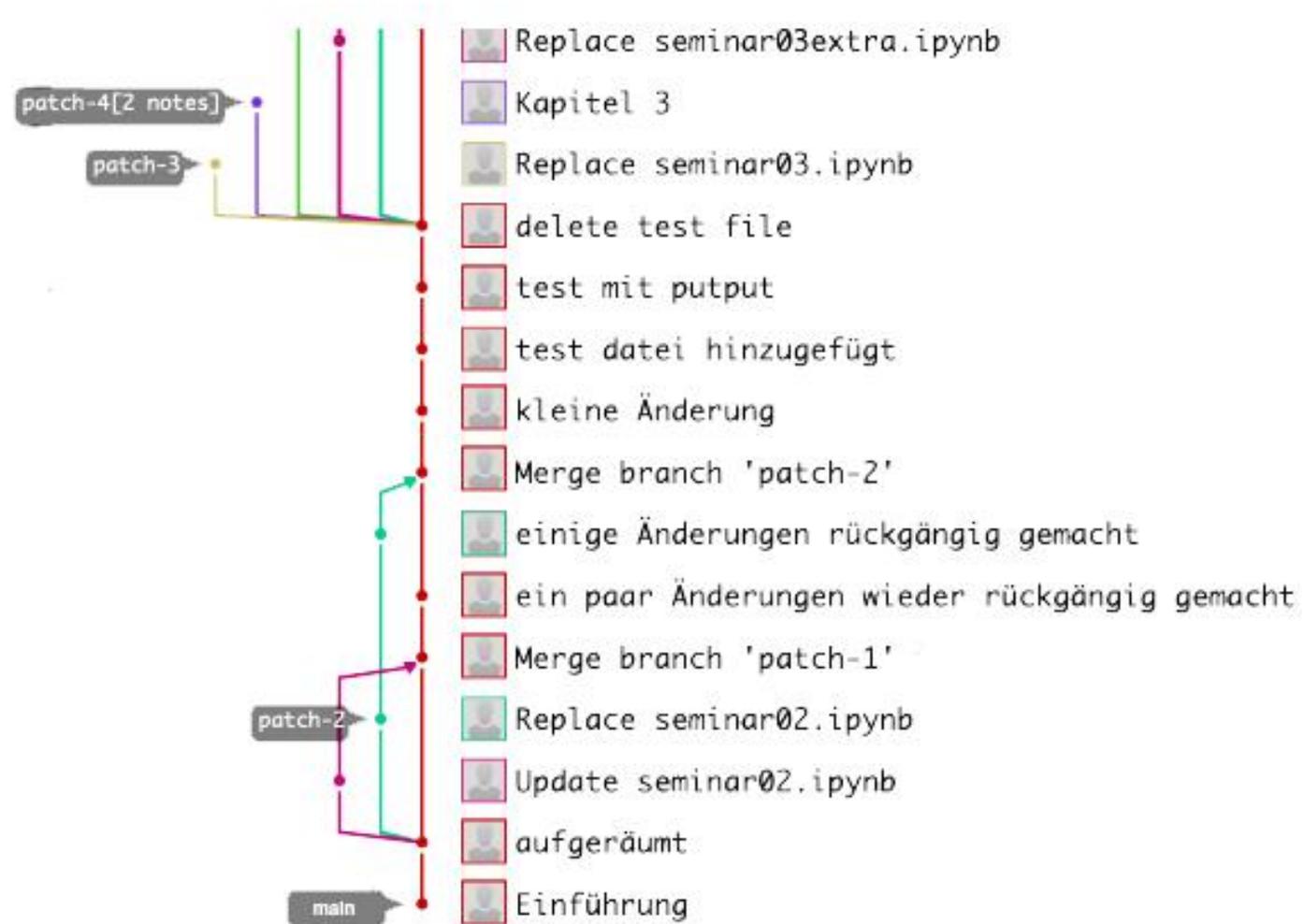
Gefördert durch
DFG Deutsche
Forschungsgemeinschaft
Projektnummer 442146713

Ordnung und Struktur

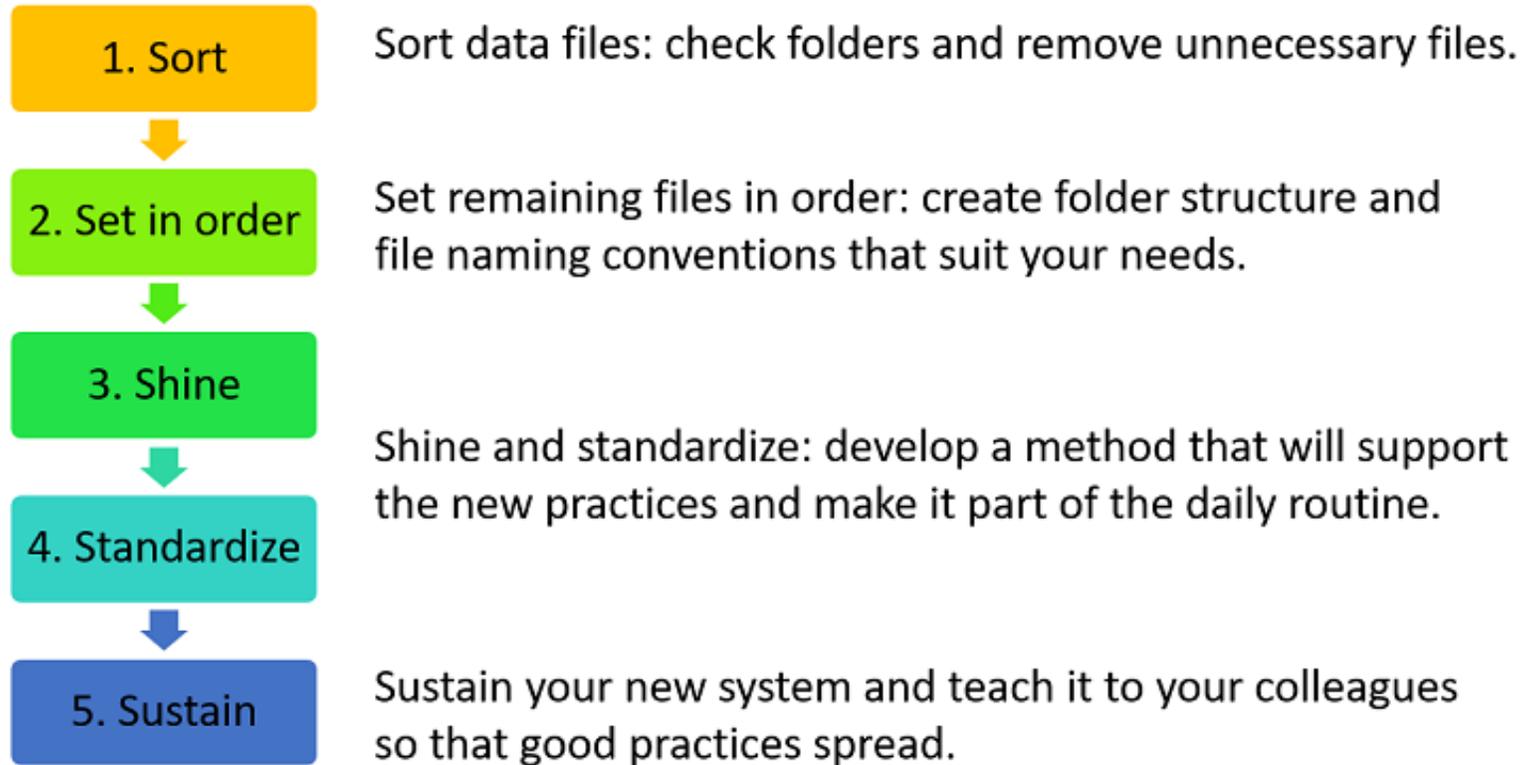
Benennung von Dateien und Ordnern

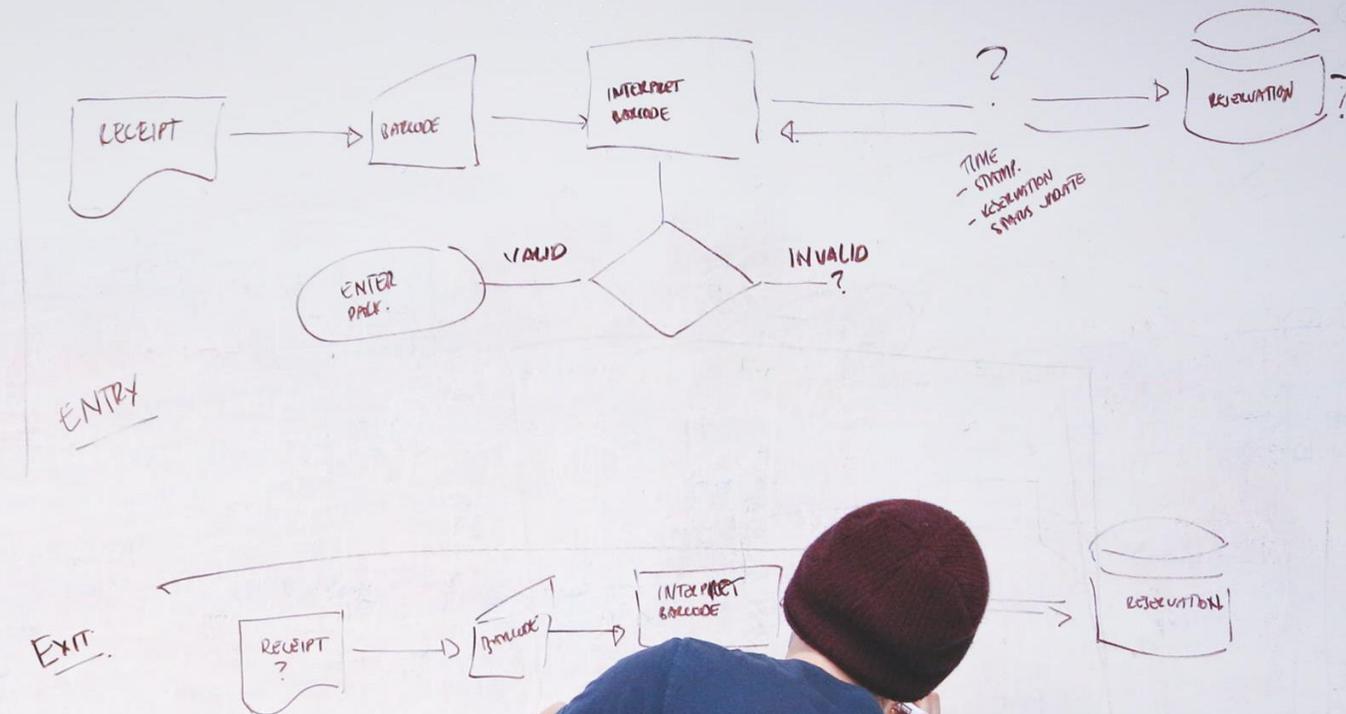
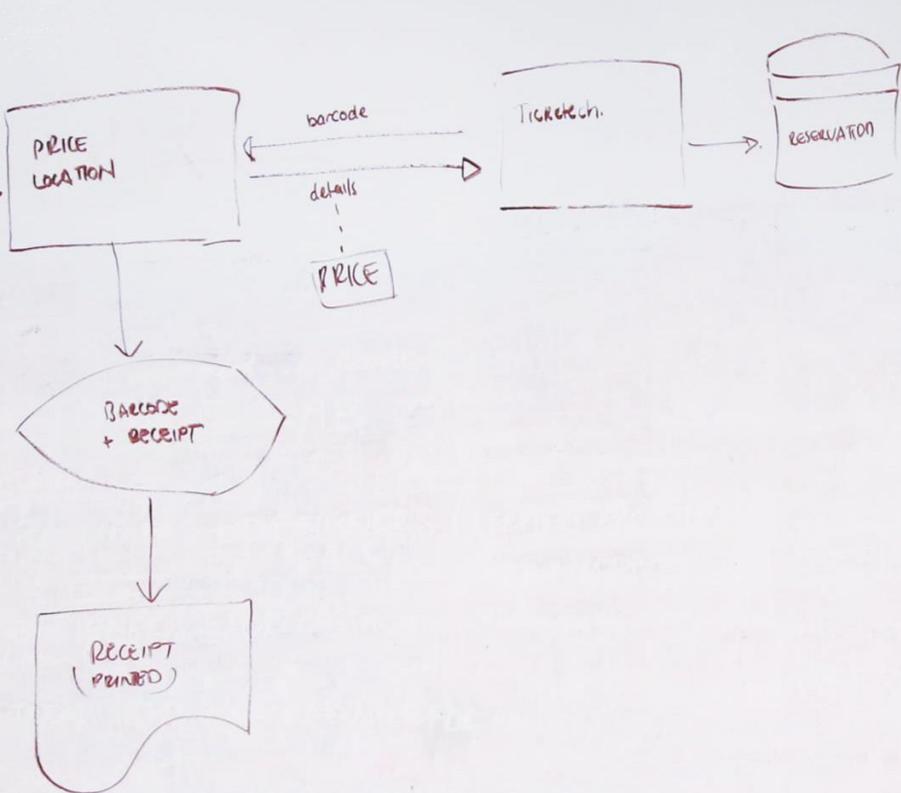
- Aussagekräftige Namen → keine „Phantasienamen“
- Einheitliches Schema
- Logische Struktur
- Datumsangabe zur chronologischen Sortierung in folgender Form:
JJJJMMTT
- Vermeidung von Leer- und Sonderzeichen
- Dokumentierte Namenskonventionen oder genutzte Abkürzungen, z. B.
[Sediment]_[Probe]_[Instrument]_[JJJJMMTT].csv
[Projekt]_[Interview]_[Ort]_[Personen-ID]_[JJJJMMTT].mp4

Beispiel für Versionskontroll-Software (gitLab)

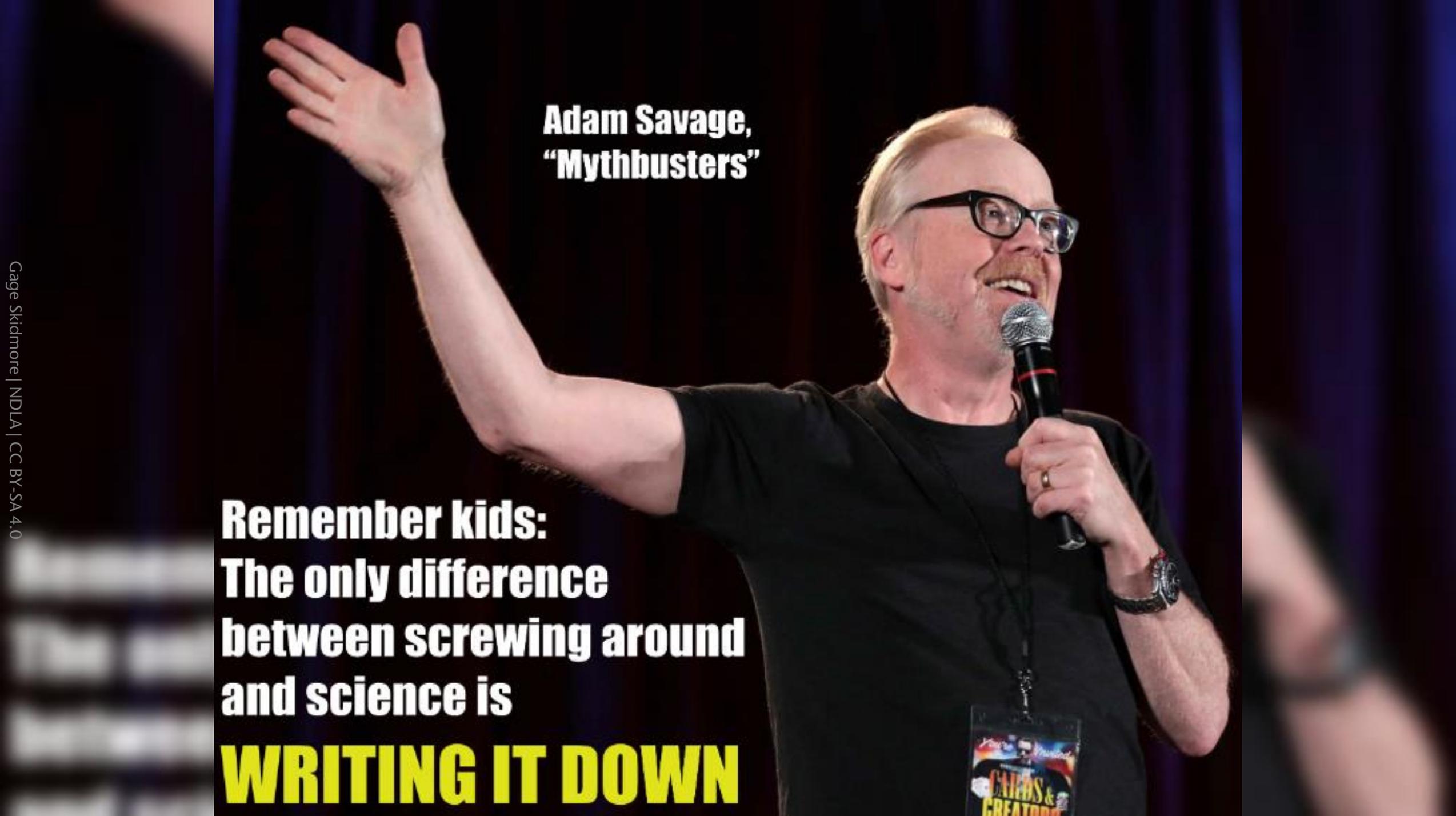


5S-DATA A lean tool for organization and improving efficiency.





Datendokumentation

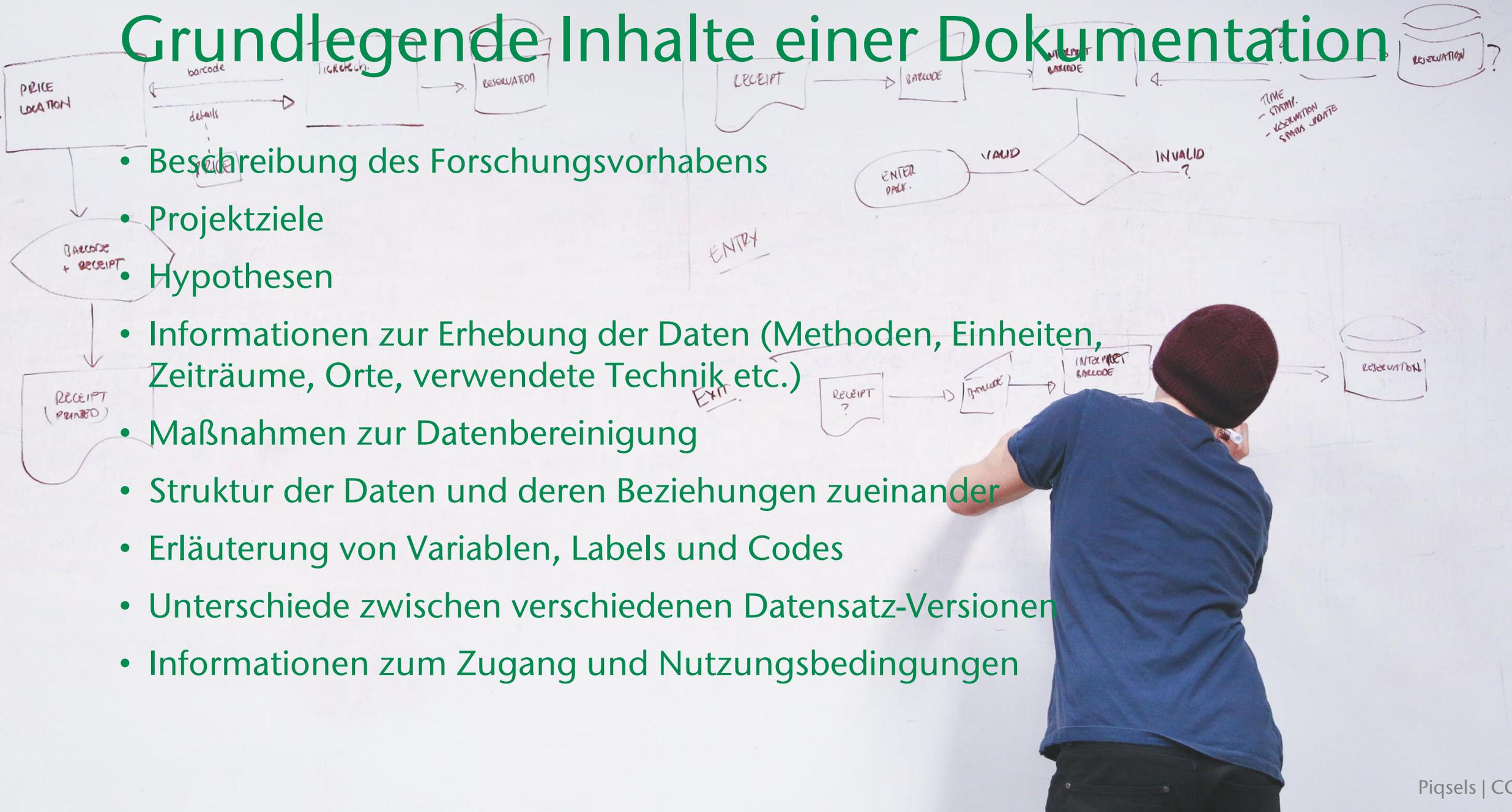
Adam Savage is shown from the chest up, wearing a black t-shirt, glasses, and a watch. He is holding a microphone in his left hand and gesturing with his right hand raised. A lanyard with a badge is around his neck. The background is dark with blue vertical stripes.

**Adam Savage,
"Mythbusters"**

**Remember kids:
The only difference
between screwing around
and science is
WRITING IT DOWN**

Grundlegende Inhalte einer Dokumentation

- Beschreibung des Forschungsvorhabens
- Projektziele
- Hypothesen
- Informationen zur Erhebung der Daten (Methoden, Einheiten, Zeiträume, Orte, verwendete Technik etc.)
- Maßnahmen zur Datenbereinigung
- Struktur der Daten und deren Beziehungen zueinander
- Erläuterung von Variablen, Labels und Codes
- Unterschiede zwischen verschiedenen Datensatz-Versionen
- Informationen zum Zugang und Nutzungsbedingungen



Dokumentation kann vorliegen als...?

- README Files

- Data Dictionaries

- Codebooks

- Electronic Lab Notebooks

- Artikel in einem Data Journal



27 genes in your body now have a new name because Microsoft Excel confused them for dates

Twenty-seven genes have got a new name over the past year to tackle a longstanding problem researchers faced while keying data into Excel, a powerful research tool.

SANDHYA RAMESH 8 August, 2020 09:05 am IST

<https://theprint.in/science/27-genes-in-your-body-now-have-a-new-name-because-microsoft-excel-confused-them-for-dates/477398/>



NEWS | 13 August 2021 | Correction 25 August 2021

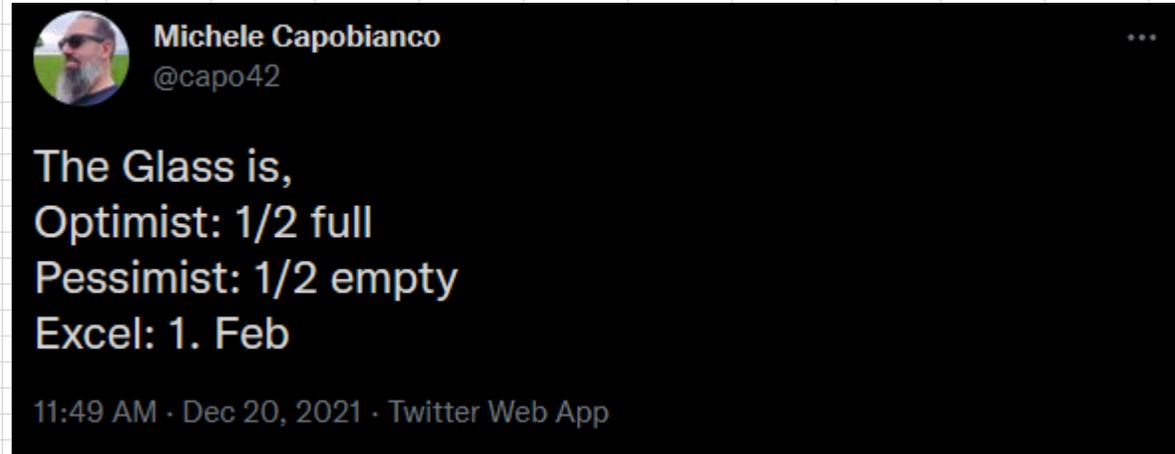
Autocorrect errors in Excel still creating genomics headache

Despite geneticists being warned about spreadsheet problems, 30% of published papers contain mangled gene names in supplementary data.

Dyani Lewis



<https://www.nature.com/articles/d41586-021-02211-4>



<https://twitter.com/capo42/status/1472881830034563077>

Comment | [Open Access](#) | [Published: 23 August 2016](#)

Gene name errors are widespread in the scientific literature

[Mark Ziemann](#), [Yotam Eren](#) & [Assam El-Osta](#) ✉

[Genome Biology](#) 17, Article number: 177 (2016) | [Cite this article](#)

138k Accesses | 58 Citations | 3249 Altmetric | [Metrics](#)

Abstract

The spreadsheet software Microsoft Excel, when used with default settings, is known to convert gene names to dates and floating-point numbers. A programmatic scan of leading genomics journals reveals that approximately one-fifth of papers with supplementary Excel gene lists contain erroneous gene name conversions.

<https://genomebiology.biomedcentral.com/articles/10.1186/s13059-016-1044-7>



eLabFTW



la durée
force

$$F = \frac{mg}{2}$$



Login through your institution

Select an identity provider

Technische Universität Clausthal ▼

Remember me

LOGIN



a free and open source electronic lab
notebook

Powered by eLabFTW

[Privacy policy](#)

Page generated in 0.0076 seconds

Metadaten – Daten über Daten

Inhaltliche Metadaten

- Titel
- Beschreibung
- Autor*in
- Urheberrechts-Inhaber*in
- Kontaktdaten
- Lizenzangaben
- Schlüsselwörter

Technische Metadaten

- Aufnahmezeitpunkt
- Brennweite
- Blende
- Belichtungsdauer
- Geographische Koordinaten
- und viele weitere

Beispiele für Metadaten-Standards

Übersicht unter: <https://rdamsc.bath.ac.uk/>

- Fachübergreifende Metadatenstandards:
 - Dublin Core
 - MARC21
- Geowissenschaften:
 - ISO 19115
- Naturwissenschaften:
 - ICAT Schema
 - Crystallographic Information Framework



Speicherung und Backups

→ Katastrophen

→ Technische Defekte

→ Diebstahl

→ Vergesslichkeit



University loses 77TB of research data due to backup error

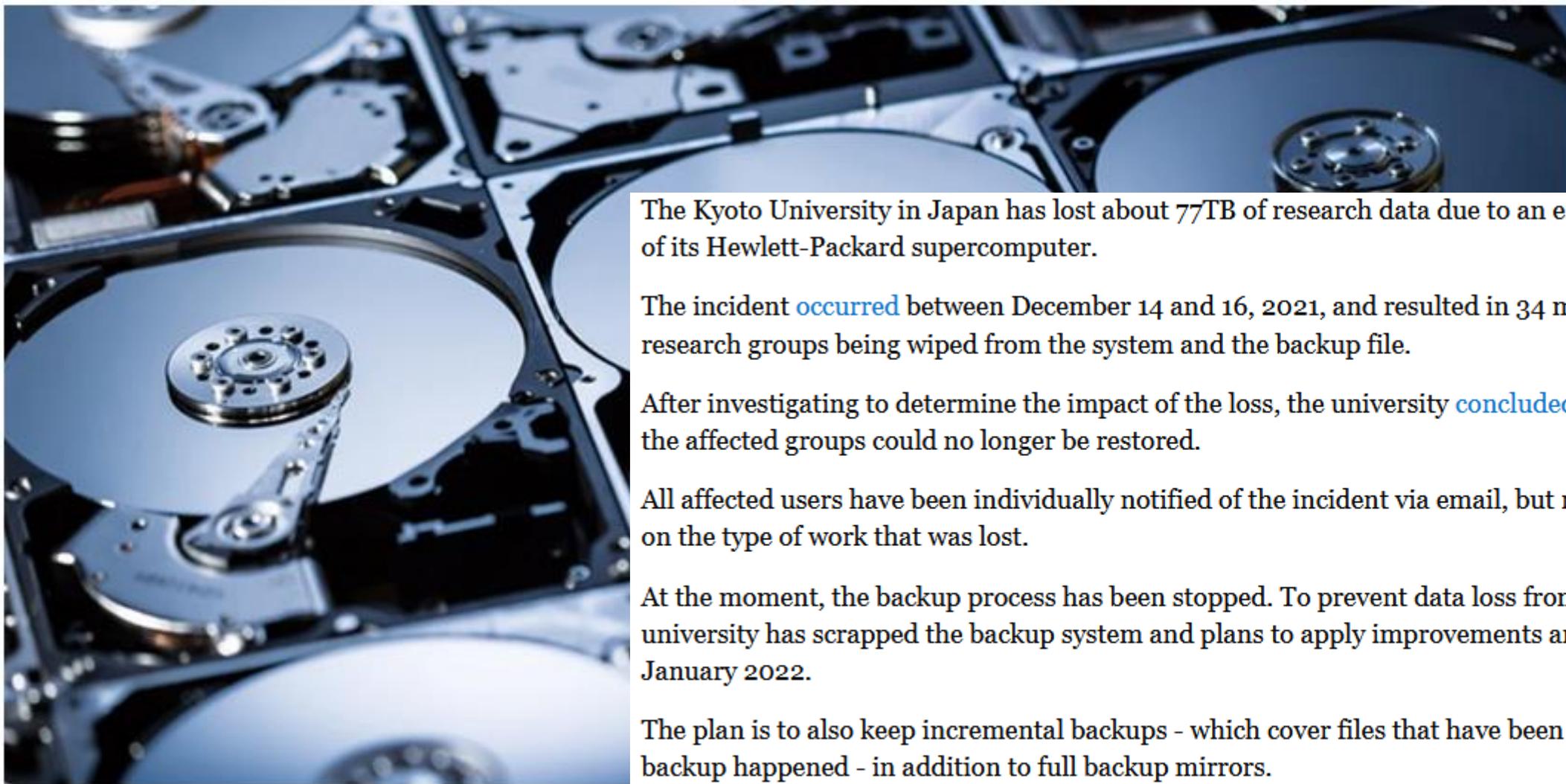
By **Bill Toulas**

December 30, 2021

11:02 AM

5

<https://www.bleepingcomputer.com/news/security/university-loses-77tb-of-research-data-due-to-backup-error/>



The Kyoto University in Japan has lost about 77TB of research data due to an error in the backup system of its Hewlett-Packard supercomputer.

The incident **occurred** between December 14 and 16, 2021, and resulted in 34 million files from 14 research groups being wiped from the system and the backup file.

After investigating to determine the impact of the loss, the university **concluded** that the work of four of the affected groups could no longer be restored.

All affected users have been individually notified of the incident via email, but no details were published on the type of work that was lost.

At the moment, the backup process has been stopped. To prevent data loss from happening again, the university has scrapped the backup system and plans to apply improvements and re-introduce it in January 2022.

The plan is to also keep incremental backups - which cover files that have been changed since the last backup happened - in addition to full backup mirrors.

University loses 77TB of research data due to backup error

By **Bill Toulas**

December 30, 2021

11:02 AM

5

<https://www.bleepingcomputer.com/news/security/university-loses-77tb-of-research-data-due-to-backup-error/>



3-2-1-Regel

Sicherung wichtiger Dateien in mindestens drei Kopien auf räumlich getrennten Datenträgern, davon einer off-site oder portabel



After investigating to determine the impact of the loss, the university **concluded** that the work of four of the affected groups could no longer be restored.

All affected users have been individually notified of the incident via email, but no details were published on the type of work that was lost.

At the moment, the backup process has been stopped. To prevent data loss from happening again, the university has scrapped the backup system and plans to apply improvements and re-introduce it in January 2022.

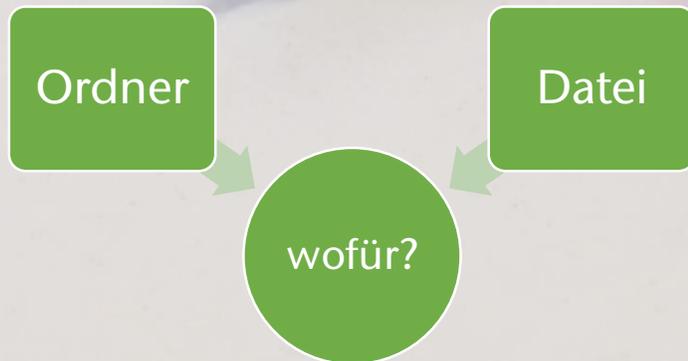
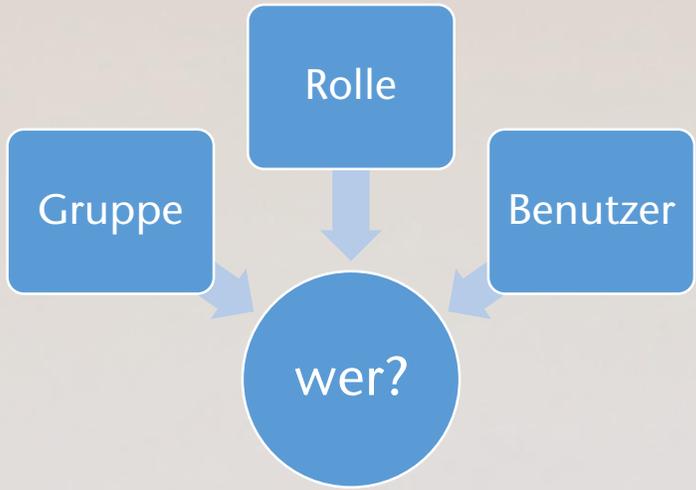
The plan is to also keep incremental backups - which cover files that have been changed since the last backup happened - in addition to full backup mirrors.



*Datenschutz Good Practice:
Schutz durch „Hardware“ (abschließbarer Raum)
Dateiverschlüsselung
Passwortsicherheit
Zugang für mindestens zwei Personen sicherstellen*

Backup	Archivierung
(Automatische) Sicherung aller Daten, um Datenverlust vorzubeugen (technisch, z. B. defekt, oder menschlich, z. B. versehentlich gelöscht)	Sicherung ausgewählter Daten, um diese langfristig aufzubewahren
Alle Versionen	Nur endgültige Versionen
	Integritätssicherung
	Langzeitspeicherung
	Durchsuchbarkeit

- Vertrauenswürdigkeit
 - Kosten
 - Zugänglichkeit
 - Langlebigkeit





Publikation von Forschungsdaten

*Wege der Datenveröffentlichung:
Supplemental Material
Enhanced Publication
Data Journals
Repositorien*





Das letzte Gerät, das die Rohdaten
für dein Paper lesen konnte.

**Wir helfen auch beim
Forschungsdatenmanagement.**

Sprecht uns an!



re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES

Persistente Identifier 1 – der DOI

DOI sind zitierfähig

Kostenlos für akademische Einrichtungen in Deutschland

Einheitliche Standards garantiert

ISO Standard 26324

Präfix einheitlich für Organisation

Suffix Kennung für Objekt

10.5281/zenodo.3723939



Persistente Identifier 2 – die ORCID
Für Personen, nicht Objekte
Eindeutige Zuordnung wissenschaftlicher Arbeiten
Kostenlos, in 30 s erstellt
Normdatum für Förderer, Journals, etc.
Wird selbst gepflegt

Wer von euch hat bereits eine?



ACADEMIC CLOUD Services Help Terms of Use Florian Strauß Logout EN DE

Account	Rocket.Chat	Mastodon	Meet
Cloud	GitLab	ePIC	ShareLaTeX
Dataverse	URL shortener	LimeSurvey	Whiteboard

<https://academiccloud.de>



reuse reduce

Nachnutzung von Forschungsdaten

Wie finde ich Forschungsdaten?

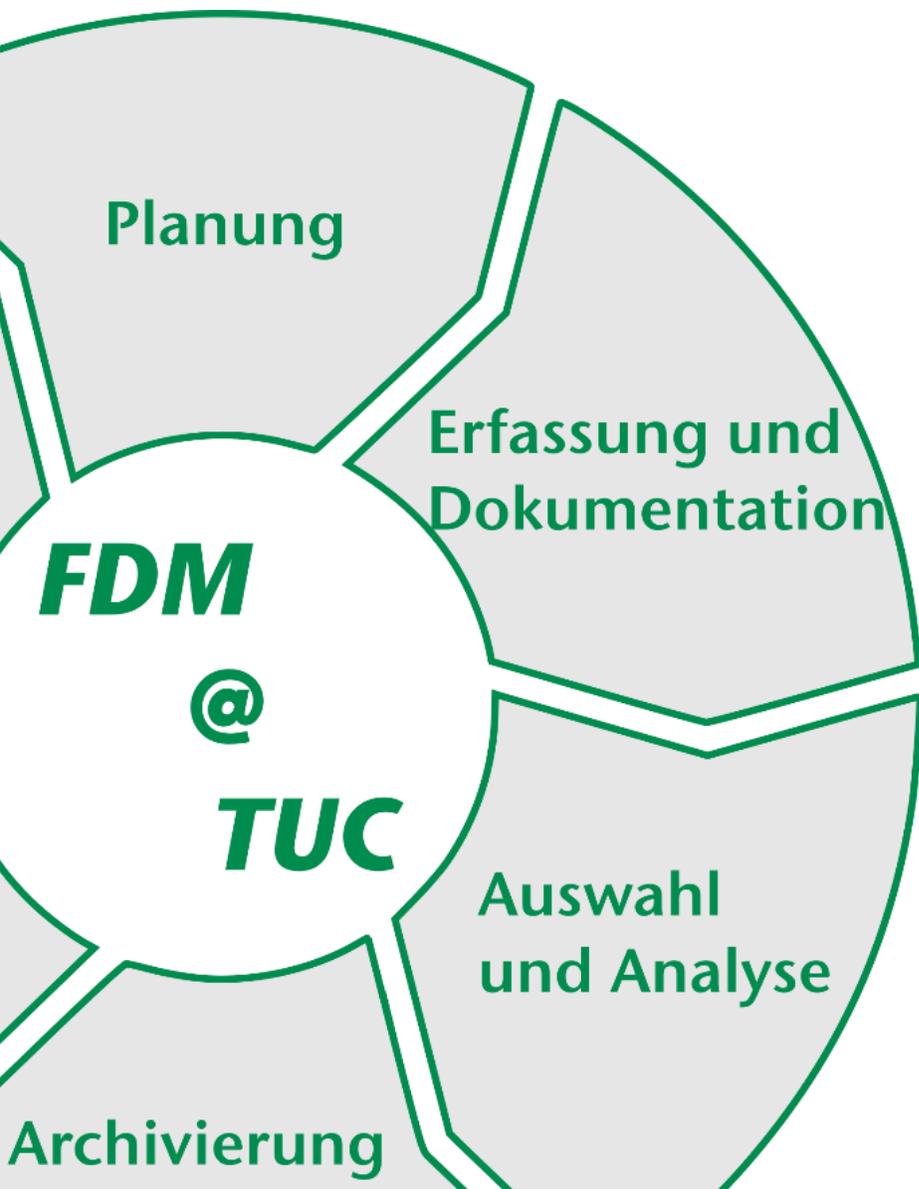
Direkt in Repositorien (Unterschied fachlich/generisch)

Metasuchmaschinen (b2find.eudat.eu, datasearch.gesis.org)

Bibliothekarische Suchmaschinen (base-search.net)

Google ;) (hat dafür eine gesonderte Oberfläche)





Fragen

Jetzt und in Zukunft



rocket.chat



Discord

Klassisch per E-Mail:

florian.strauss@tu-clausthal.de

Telefon:

05323 72 2276

Oder vorbeikommen:

UB, Raum 109