

# Datenmanagement im Fokus Organisation, Speicherstrategien und Datenschutz

Love Data Week | 19.04.2024

Carolin Hundt (carolin.hundt@zv.uni-leipzig.de) Pia Voigt (pia.voigt@zv.uni-leipzig.de)

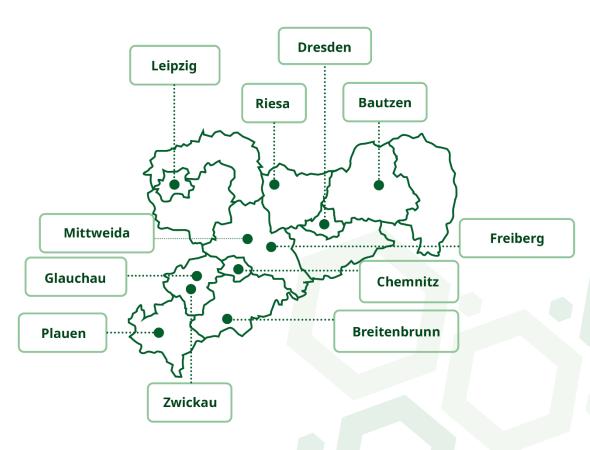
## SaxFDM

## Sächsische Landesinitiative für Forschungsdatenmanagement

- 2019 als Bottom-Up-Initiative entstanden
- derzeit 24 Einrichtungen an 11 Standorten

#### <u>Ziele</u>

- Bündelung und Koordination existierender, Dienste zum Forschungsdatenmanagement (FDM)
- Aufbau eigener FDM-Dienste
- Etablierung als zentraler Ansprechpartner in FDM-Fragen
- Zielgruppe: Forschende, Forschungseinrichtungen und FDM Personal an öffentlichen Forschungseinrichtungen



## SaxFDM: Services und Aktivitäten



## Beratungen

- Beratung für Einzelne
- Beratung für Institutionen
- Terminbuchungstool <u>https://saxfdm.de/beratung/</u>
- Fragen per Mail beratung@saxfdm.de



## Schulungen

- FDM-Einführung
- Projektanträgen & Datenmanagementplanung
- Datenmanagement in Verbundprojekten
- Neuer DFG-Kodex (GWP) was bedeutet das für den Umgang mit Forschungsdaten?



## **Vernetzung/Events**

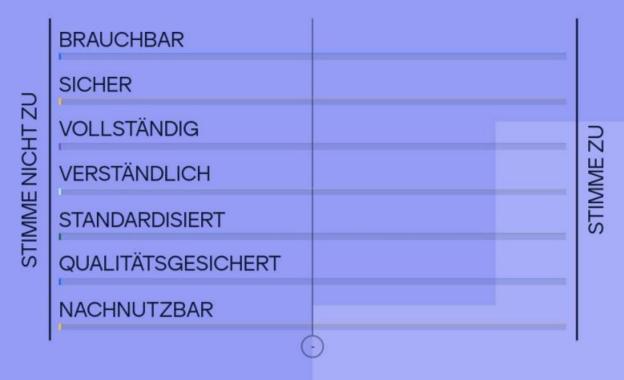
- SaxFDM-Tagung
- Digital Kitchen



# **Ein System finden – Wie ist der Zustand Ihrer Daten?**

# Wie beurteilen Sie aktuell den Zustand Ihrer Daten?





https://shorturl.at/hkJX8

Zugangscode: 43 90 64 5

## **Warum** Daten organisieren?



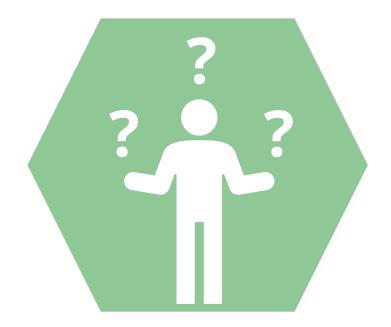
Daten auch nach längerer Zeit noch verstehen



Anderen ermöglichen, Ihre Dateien zu verstehen (Stellenwechsel, Übergabe, längere Abwesenheit)



Überblick über Ihre Daten behalten





Effizienz im Arbeitsalltag steigern



**Dateien schnell auffinden** 

## Ein System finden

#### Aller Anfang ist...

Wie ist mein Projekt strukturiert?

Welche Arbeitsschritte umfasst mein Projekt?

Wie werden die digitalen Ergebnisse meiner Arbeit strukturiert sein?

Welche Ordnungskategorien spiegeln mein Projekt am besten?

## Dokumentation in Readme-File (einfache Textdatei)

- Struktur und Ordnungskategorien
- Ordnerstruktur
- Dateibenennung



## Ein System finden | Ordnerstrukturen erstellen

#### Projekt

#### **Projektmanagement**

- 0\_Antrag
- 1\_Finanzierung
- 2\_Berichte
- **3**\_Genehmigungen

#### **Empirische Daten**

- 0 Rohdaten
- 1 Code
- 2\_Datenauswertung

#### Paper

- 0\_Literatur
- 1\_Texte

#### **Veröffentlichungen**

- 0 Präsentationen
- 1\_Journal Artikel
- 2\_Bücher

### **Obere Ebene**

 Welche Begriffe sind allen bekannt und werden von allen verstanden?

### **Untere Ebene**

- Welche Begriffe sind allen bekannt und werden von allen verstanden? und/oder
- Welche Ansätze bietet meine Arbeitspraxis, um Dateien und Ordner zu organisieren?
  - Zweck der Datenerhebung/-erfassung?
  - Verarbeitungsgrad der Daten?
  - Zeit?
  - ...?

## Wie erstelle ich eine (gemeinsame) Ordnerstruktur?

- Obere Ebene(n):
  - Welche Begriffe sind allen bekannt und werden von <u>allen</u> verstanden?

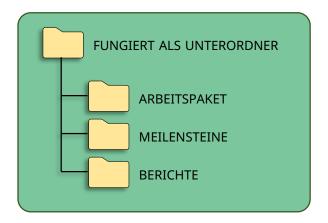




## Wie erstelle ich eine (gemeinsame) Ordnerstruktur?

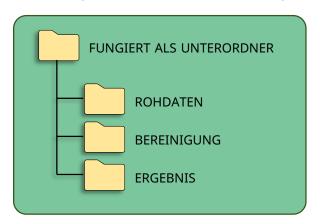
NACH "ZWECK"

WARUM WERDEN DATEN AUFBEWAHRT?



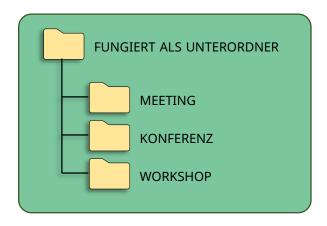
NACH "VERARBEITUNGSGRAD"

WIE WURDEN DATEN VERARBEITET?



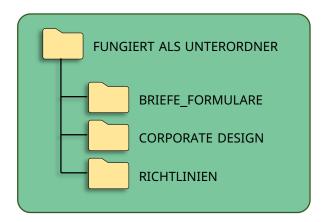
#### NACH "KOMMUNIKATION"

WO FALLEN DATEN AN?



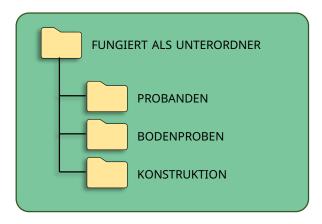
NACH "ART"

WELCHE HILFSMITTEL EXISTIEREN ZUM THEMA?



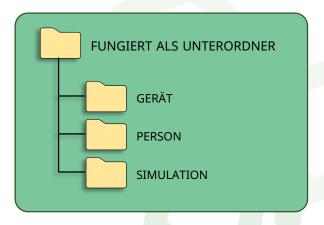
#### NACH "FORSCHUNGSOBJEKT"

WAS HABEN DATEN ZUM THEMA?

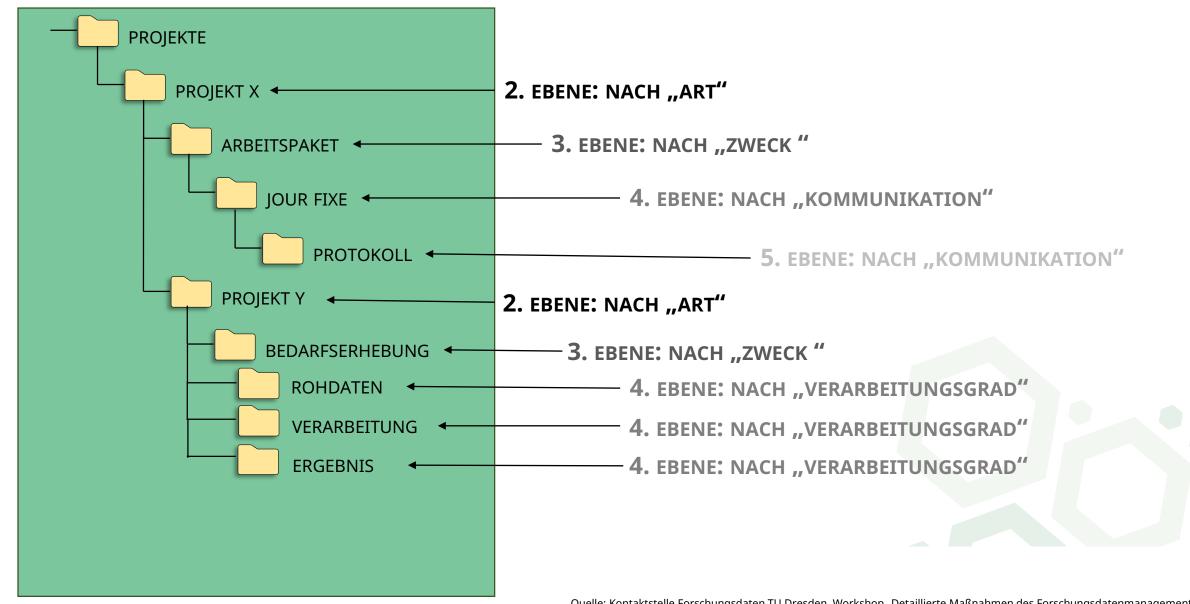


NACH "URSPRUNG"

WELCHEN URSPRUNG HABEN DATEN?



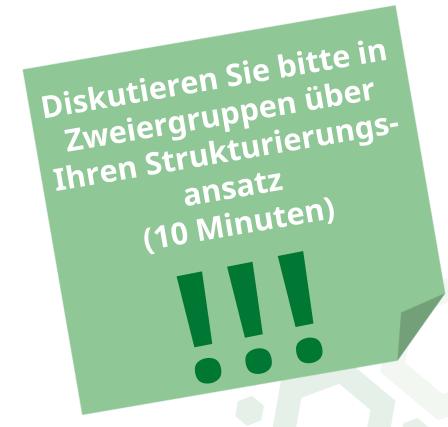
## Wie **erstelle** ich eine (gemeinsame) Ordnerstruktur? Beispiel



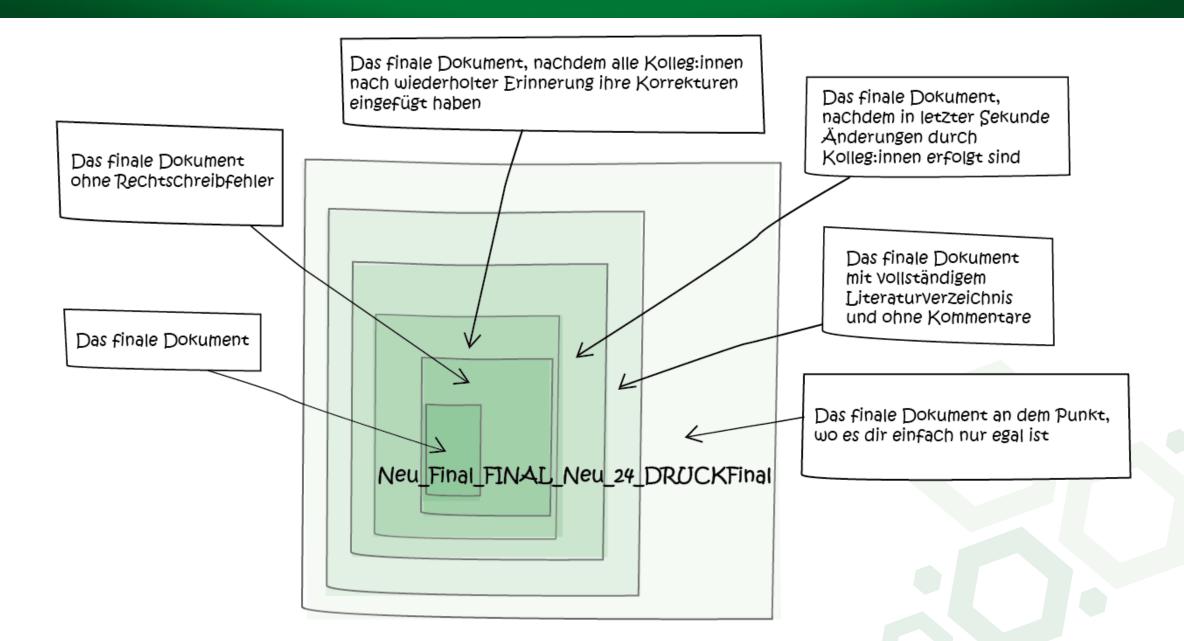
## Übung

Stellen Sie sich bitte gegenseitig kurz die Ordnerstruktur in einem aktuellen Projekt vor.

- Können Sie einen Strukturierungsansatz für Ordner oder Unterordner finden, nach welchem Sie vorgegangen sind?
- Welche Unterordner könnten gut nach einem konsistenten Ansatz strukturiert werden? Wo ist es schwierig?



## Dateibenennung | Motivation



## Übung | Dateibenennung

#### **Gute Dateinamen?**

Olga\_170413\_sample17k 170413\_sample17k\_olga

Crst\_765\_spkt\_1203 Crystal\_765\_spektr\_20161203

Naturepaper karl britta james done! Naturepaper+karl+britta+james done! edited Nature\_karlbrittajames\_endendversion



## Maßnahmen zur Dateiorganisation

## **Nutzung von Namenskonventionen**

#### Dateireihenfolge

Nummern linksbündig mit anführenden Nullen

01\_Entwurf\_\*, 02\_Entwurf\_\*

**Trennzeichen** "\_" oder "-,, verwenden, um **Schlüsselwörter** und **Metadaten** einzurahmen

2024-04-19\_ProjektX\_Meilenstein1.csv

#### Großbuchstaben verwenden

\*\_Meeting\_Protokoll.txt

#### **Chronologische Reihenfolge**

Beginn mit Datum in der Form YYYY-MM-DD

2024-04-19\_Meeting\_Protokoll\_\*

#### **Dateiversionsverwaltung**

\*\_Rohdatensatz\_v01.txt

Keine Sonderzeichen verwenden.

%**&**\$§!?#</>

maschinen lesbar

## Tools zur gleichzeitigen Umbenennung

## **Windows**

Ant Renamer
(www.antp.be/software/renamer)

**RenameIT** (sourceforge.net/prpjects/renameit)

Bulk Rename Utility (www.bulkrenameutility.co.uk/)

**Total Commander** (https://www.ghisler.com/deutsch.htm)

## Unix

rename command

#### Mac

Renamer 6 (renamer.com/)

Name Changer (mrrsoftware.com/namechanger/)

ExifRenamer

(<a href="https://www.qdev.de/?location=mac/exifrenamer">https://www.qdev.de/?location=mac/exifrenamer</a>)

## Linux

**GNOME Commander** 

(www.nongnu.org/gcmd/)

**GPRename** 

(<a href="http://gprename.sourceforge.net/">http://gprename.sourceforge.net/</a>)





## Anforderungen an Speicher

## Worauf man bei der Wahl eines Speichermediums achten sollte:

- Metadaten, Skalierbarkeit
- Dateisicherheit, Möglichkeit zur Verschlüsselung
- Größe, Kosten
- Zugriffsgeschwindigkeit, Lese- und Schreibgeschwindigkeit
- Zugangskontrolle, Rechteverwaltung
- Wartbarkeit und Zuverlässigkeit
- Gemeinsamer Zugriff, Schnittstellen
- Backup

## Charakteristika von Speichermedien

Übung – Gruppenaufgabe

Finden Sie sich in vier Gruppen zusammen und bewerten Sie die verschiedenen Speichermedien! (Pro & Contra)

Gruppe 1: Eigener PC

Gruppe 2: Mobile Speichermedien

Gruppe 3: Institutionelle Speicherorte

Gruppe 4: Externe Speicherorte



## Charakteristika von Speichermedien | Eigener PC

Vorteile	Nachteile



## Charakteristika von Speichermedien | Eigener PC

Vorteile	Nachteile
<ul> <li>Selbst verantwortlich für Sicherheit und Backup</li> </ul>	<ul> <li>Was mit dem PC geschieht, geschieht mit dem Backup</li> </ul>
Maximale Kontrolle	<ul> <li>Evtl. fehlende Ressourcen und Know-how</li> </ul>
	Einzellösungen aufwendig



# Charakteristika von Speichermedien | Mobiles Speichermedium (Stick, externe Festplatte)

Vorteile	Nachteile



# Charakteristika von Speichermedien | Mobiles Speichermedium (Stick, externe Festplatte)

Vorteile	Nachteile
<ul> <li>Einfach zu transportieren</li> <li>Kann im verschließbaren Schrank oder Safe aufbewahrt werden</li> </ul>	<ul> <li>Verlust, Diebstahl</li> <li>Bei Verlust: Inhalte ungeschützt, falls nicht verschlüsselt</li> <li>Externe Festplatte: stoß- und verschleißanfällig</li> </ul>



# Charakteristika von Speichermedien | Institutionelle Speichermedien (Cloud, virtuelle Laufwerke)

Vorteile	Nachteile



## Charakteristika von Speichermedien | Institutionelle Speichermedien (Cloud, virtuelle Laufwerke)

Vorteile	Nachteile
Backup der Daten ist sichergestellt	<ul> <li>Geschwindigkeit eventuell vom Netzwerk abhängig</li> </ul>
<ul> <li>Professionelle Durchführung und Wartung</li> </ul>	<ul> <li>Zugriff auf Backups evtl. verzögert durch Dienstweg</li> </ul>
Speicherung gem. Datenschutz- richtlinien der Institution	<ul> <li>Evtl. unklar welche Sicherheits- kriterien und -strategien eingesetzt werden</li> </ul>



## Charakteristika von Speichermedien | <u>Externe</u> Speichermedien (Cloud eines ext. Anbieters)

Vorteile	Nachteile



## Charakteristika von Speichermedien | Externe Speichermedien (Cloud eines ext. Anbieters)

<ul> <li>Einfache Nutzung/Verwaltung</li> <li>Backup der Daten ist sichergestellt</li> <li>Für mobiles Arbeiten nutzbar</li> <li>Professionelle Durchführung und Wartung</li> <li>Je nach Anbieter unsichere Verbindung</li> <li>Abhängig vom lokalen Internet (Up- &amp; Download evtl. langsam)</li> <li>Zugriff auf Backups verzögert</li> <li>Datenschutz?</li> </ul>



## LEITFRAGEN BEI DER WAHL DES SPEICHERMEDIUMS

Welche Datentypen habe ich und wie häufig werde ich diese ersetzen?

Wie viel Speicherplatz benötige ich?

Wer benötigt Zugang?

Ist es notwendig einen Remote-Zugang zu den Daten zu haben?

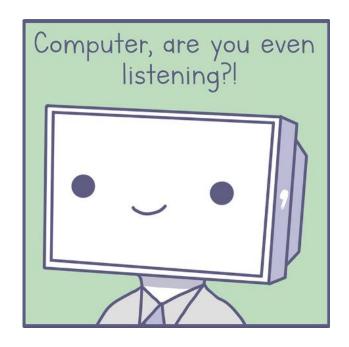
Wie wichtig ist schneller Zugriff?

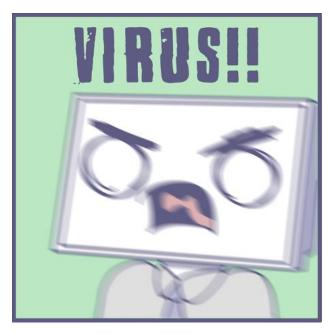


## Datensicherung und Backup | Motivation

## Risiko von Datenverlust bspw. durch

- Technische Defekte
- Vergesslichkeit
- Katastrophen
- Diebstahl







## Datensicherung und Backup | Maßnahmen

#### Erstellen von Kopien der Dateien:

Wiederherstellen nach z.B. Defekt des Speichermediums/unbeabsichtigte Löschung

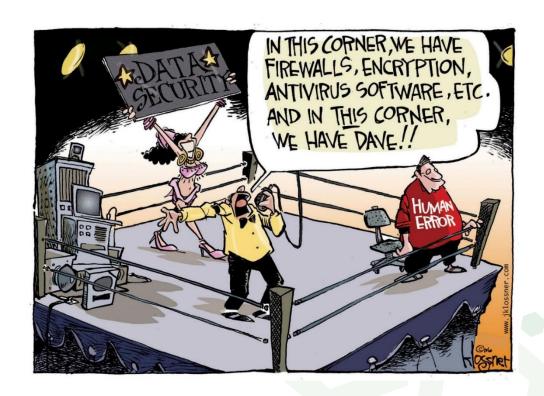
#### **Backupstrategie aufsetzen (3-2-1 Regel)**

#### **Nutzung sicherer Speichermedien**

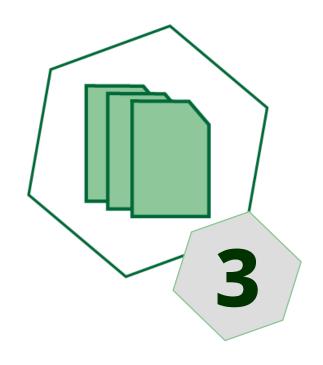
Server mit regelmäßigem automatischem Backup

#### (Sensible) Daten schützen

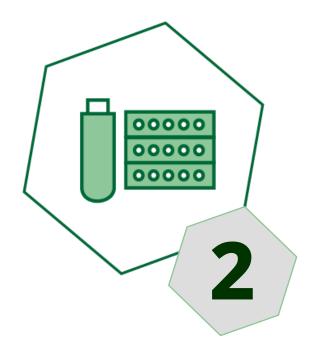
- Hardware (bspw. separat abschließbarer Raum)
- Dateiverschlüsselung
- Sicherheit der Passwörter
- Mindestens zwei Personen sollten Zugang zu Ihren Daten haben



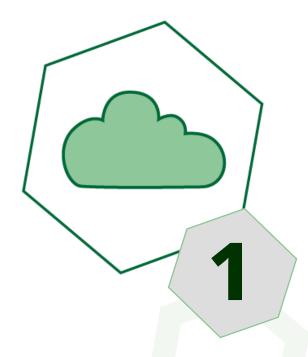
## **Backupstrategie 3-2-1-Regel**



Erstellen Sie **mindestens drei** Kopien einer Datei
1 lokal, 1 automatisches
Backup, 1 externe
Festplatte



Speichern Sie ihre Daten auf **mindestens zwei** unterschiedlichen Medien.



Bewahren Sie **ein**Backup dezentral auf.

1 lokal, 1 externe
Festplatte



## Was ist das? Datenschutz



#### Grundgesetz

Recht auf informationelle Selbstbestimmung

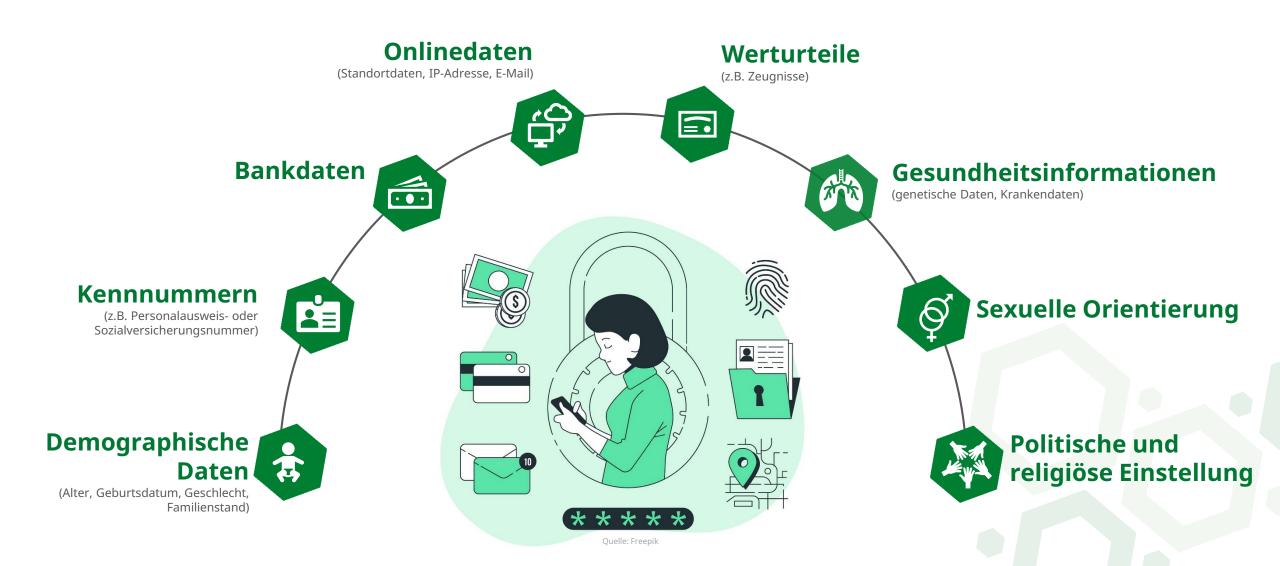
#### Datenschutzgrundverordnung

(DSGVO, EU-Ebene), Bundesdatenschutzgesetz, Landesdatenschutzgesetz

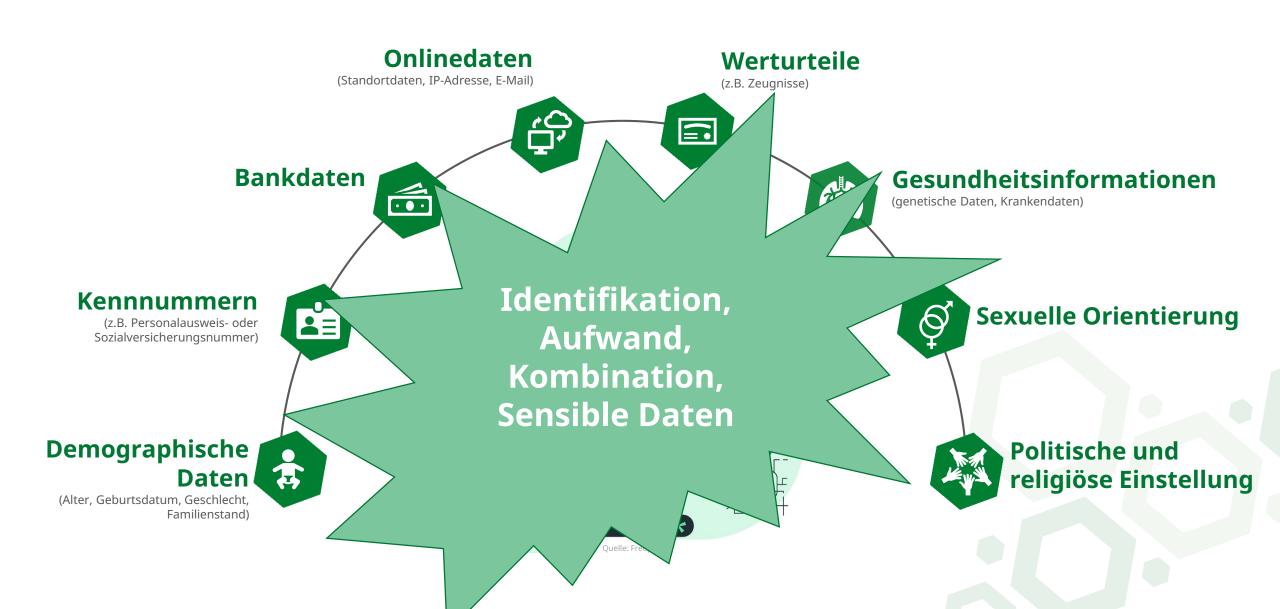
#### **Datenlebenszyklus**

Erhebung, Speicherung, Verarbeitung, Archivierung (Löschung), Nachnutzung

## **Was ist das? - Personenbezug**



## Was ist das? - Personenbezug

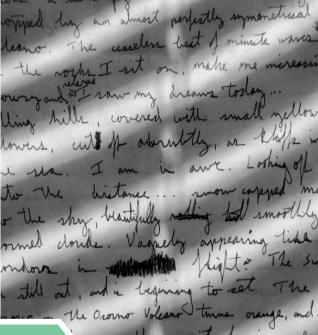


## Datenarten mit Personenbezug









walk back through the Plaza, I see the met he the bring out were for the young people at right. At the same time it where the aller generation takes the walks.



### Verarbeitung personenbezogener Daten

#### Art. 6, DSGVO: Rechtmäßigkeit der Verarbeitung

- **Einwilligung** der betroffenen Person(en) (zweckgebunden)
- Verarbeitung zur Erfüllung einer Aufgabe im öffentlichen Interesse oder eines Vertrages erforderlich
- •

## <u>Art. 89, DSGVO</u>: Verarbeitung von personenbezogenen Daten zu wissenschaftlichen Zwecken (u.a.)

- Zu definierten wissenschaftlichen Zwecken Datenauswertung möglich
- Rechte betroffener Personen müssen gewahrt bleiben,
   Grundsatz der Datenminimierung
   technische oder organisatorische Maßnahmen (z. Bsp. Pseudonymisierung, informierte Einwilligung)

#### Verarbeitung personenbezogener Daten im öffentlichen Interesse

- Öffentliches Interesse = Vorrang vor Individualinteresse
- Durchführung des (Forschungs-)Vorhabens betrifft/ist gegenwärtiges Anliegen der Allgemeinheit
- Prinzip der Datensparsamkeit im Interesse der Rechte der betroffenen Personen
- Grundrechtseingriff im Verhältnis zu Erhebungszweck (Kinder!)
- Beispiel: Forschung mit/an beeinträchtigten Personen (psychisch, physisch)

#### Nachweispflichten:

➤ Forschungsinteresse, wissenschaftlich und/oder historisch, Nutzen für die Allgemeinheit

## Pflichten | Erhebung und Verarbeitung



Abwägung Erhebungsinteresse und Schutzbelangen Betroffener von Anfang an

Welche Angaben werden <u>zwingend</u> benötigt?



Getrennte, sichere Speicherung von personenbezogenen und Erhebungsdaten

- Pseudonymisierung
- Verschlüsselung
- Zugriffskontrolle



Verfahrensverzeichnis (mit Datenschutzbeauftragtem) anlegen (Art. 30, DSGVO), Muster BfDI)



Löschen (Löschkonzept) oder Anonymisieren nach Abschluss der Forschungs-/Verarbeitungstätigkeit



Betroffene über (Forschungs-)Zweck aufklären

• Informierte Einwilligung

Prinzip de nkert

### **Informierte Einwilligung**

#### I.d.R. für jede Erhebung und Nutzung personenbezogener Daten

(Erhebung, Verarbeitung, Archivierung, Veröffentlichung, Anschlussnutzung)

**Einwilligung und Informationen trennen** 



Nachweisbare Form (schriftlich, mündlich)





Einfache, zielgruppengerechte Sprache



Widerrufsrecht hervorheben



Umstände, Zweck, Art der Datenerhebung und -verarbeitung











## Informierte Einwilligung: Inhalte

- Name, Kontaktdaten der Verantwortlichen
- Name, Kontaktdaten der Datenschutzbeauftragten
- Begleitinformationen zum Vorhaben
- Zweck(e) und Art der Datenverarbeitung
- Rechtliche Grundlage der Verarbeitung (Informationspflicht)
- Rechte des/der Betroffenen ggü. der/dem Verantwortlichen
- Hinweis auf Beschwerderecht bei übergeordneten Stellen
- Ggf. Empfänger bei Weitergabe der Daten an externe Stellen (bspw. Transkription)
- Kategorien, (öffentliche) Quelle bei Verarbeitung bereits existierender Daten (Dritterhebung)



## Informierte Einwilligung: Betroffenenrechte (DSGVO)

## Art. 7 Abs. 3 DSGVO: Recht auf Widerruf der Einwilligung

Sie haben das Recht, Ihre Einwilligung jederzeit mit Wirkung für die Zukunft zu widerrufen

## Art. 15 DSGVO: Auskunftsrecht

Sie haben das Recht, Auskunft darüber zu erhalten, welche Daten wir zu Ihrer Person verarbeiten.

## Art. 16 DSGVO: Recht auf Berichtigung

Sollten die Sie betreffenden Daten nicht richtig oder unvollständig sein, so können Sie die Berichtigung oder die Vervollständigung Ihrer Angaben verlangen

## Art. 17 DSGVO: Recht auf Löschung

Sie können jederzeit die Löschung ihrer Daten verlangen.

#### Art. 18 DSGVO: Recht auf Einschränkung der Verarbeitung

Sie können die Einschränkung der Verarbeitung der Sie betreffenden personenbezogenen Daten verlangen.

## Art. 21 DSGVO: Widerspruchsrecht

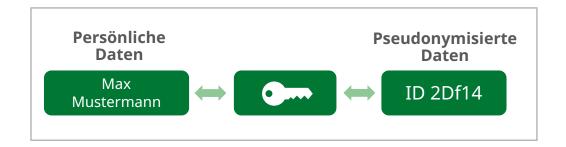
Sie können jederzeit gegen die Verarbeitung der Sie betreffenden Daten Widerspruch einlegen.

## Pseudonymisierung und Anonymisierung

#### Quantitative vs. qualitative Daten

#### **Pseudonymisierung**

- Personenbezogene Merkmale ersetzen → Schlüsselliste
- Sichere, getrennte Aufbewahrung (Zugriff mind. 2 Personen)
- Zusammenführung zu angegebenem Zweck
- Löschen der Daten und der Schlüsselliste nach Projektende



#### **Anonymisierung**

- Bezug zu natürlicher Person auch mit großem Aufwand nicht (mehr) herstellbar
- Keine absolute Anonymität (elektronische Datenanalyseverfahren)
- Möglichst hoher Aufwand zur De-Anonymisierung → faktische Anonymität



## Anonymisierung: Möglichkeiten



## Fallbeispiele Anonymisierung

Bildung von Klassen oder Kategorien

Löschen von Merkmalen/Platzhalter, starke Abstraktion

Ersetzen durch Merkmale mit vergleichbarer Bedeutung

Originalangabe, Identifikator		(3) Aggregation
Name	Aishe	Frau
	Christian	Mann
Alter	27	20 bis 30 Jahre alt
Ort	Gelsenkirchen	Ruhrgebiet
	Berlin-Neukölln	Migrantenbezirk einer Großstadt
Bundesland	Sachsen	Ostdeutschland
Arbeitgeber	Bäckerei Schmidt	Handwerk (Branchenangabe)
Beruf	Friseur	Handwerker (z. B. Angabe des ISCO-Codes)
Datumsangaben	12. 01. 1998	Zeitraum zwischen 1990 und 2000; 01/1998

Originalangabe, Ident	ifikator	(2) Pseudonym	
Name	Aishe	Nesli	
	Christian	Thomas	
Arbeitgeber	Bäckerei Schmidt	Metzgerei	
Beruf	Friseur	Kosmetiker	
		•	

Originalangabe, Identif	fikator	(4) Ersetzung
Name	Aishe	Freundin
	Christian	Bruder
Alter	27	Person der gleichen Alterskohorte/Generation
Ort	Gelsenkirchen	Wohnort, Arbeitsort, Ausbildungsort
	Berlin-Neukölln	
Bundesland	Sachsen	Bundesland des Geburtsorts

## **CHECKLISTE**DATENSCHUTZ

- Datensparsamkeit, Datenminimierung
- Informierte Einwilligungserklärung, laufende Information bei Zweckänderung
- Datensicherheit: getrennte Speicherung, Pseudonymisierung (Schlüsselliste), Passwort, Backup
- Zugriffsregelungen (Nutzungsvereinbarungen)
- Anonymisierung, sobald Forschungszweck dies erlaubt
- Fristgerechtes Löschen oder Widerspruch oder keine Einwilligung/Anonymisierung





# Vielen Dank für Ihre Aufmerksamkeit

Wir unterstützen Sie jederzeit bei allen Anliegen rund um das Thema Forschungsdatenmanagement!

#### **BERATUNG**

- für Einzelpersonen und Institutionen
- Terminvereinbarung per Mail (<u>beratung@saxfdm.de</u>) oder das Buchungstool auf unserer Website <u>www.saxfdm.de/beratung</u>

#### **SCHULUNG**

- FDM-Einführung
- Projektanträge, Datenmanagementplanung uvm.

## Acknowledgements

In diesem Workshop/Foliensatz nutzen wir Folien (mit Änderungen/Anpassungen), Inhalte und Abbildungen aus folgenden Quellen:

- Biernacka et al. 2021: Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement. Version 4. (https://doi.org/10.5281/zenodo.5773203)
- Denise Dörfel, Robert Müller, Torsten Gille (Kontaktstelle Forschungsdaten Dresden): Umgang mit Daten in der alltäglichen Forschungspraxis.
- Pia Voigt, Stephan Wünsche, Sebastian Frericks (FDM Serviceteam Leipzig): Crashkurs Forschungsdatenmanagement.
- Pia Voigt, Stephan Wünsche (FDM Serviceteam Leipzig): Rechtssicherheit im Umgang mit Forschungsdaten. (https://doi.org/10.5281/zenodo.7773801)
- Carolin Hundt, Johannes Sperling: Einführung ins Forschungsdatenmanagement.
- Meyermann, Alexia, Porzelt, Maike (2014): Hinweise zur Anonymisierung von qualitativen Daten. Forschungsdaten Bildung informiert, Nr.1 (2014). https://www.forschungsdaten-bildung.de/files/fdb-informiert-nr-1.pdf [Zuletzt abgerufen am 15.04.2024]
- System32Comis (2024). Episode 12,13, 14. Online unter: <a href="https://tapas.io/series/System32Comics/info">https://tapas.io/series/System32Comics/info</a>. [Zuletzt abgerufen am 04.03.2024]
- Klossner, John (2006): Human Security. Online unter: https://www.jklossner.com/humannature/hu5aw8dw1f2ss2xfuy89v84cpgcqk1 [Zuletzt abgerufen am 04.03.2024]

