



# Keine Chance dem Datenverlust

---

Forschungsdaten nachhaltig sichern  
Jessica Rex & Nadine Neute

# Agenda

---

1. Kurze Vorstellung des Thüringer Kompetenznetzwerks  
Forschungsdatenmanagement
2. Aktive Phase des FDM
3. FDM kurz nach dem Projektende
4. FDM lange nach dem Projektende
5. Q&A



# Strategie zur Digitalisierung (12/2017)

---

## Handlungsfelder:

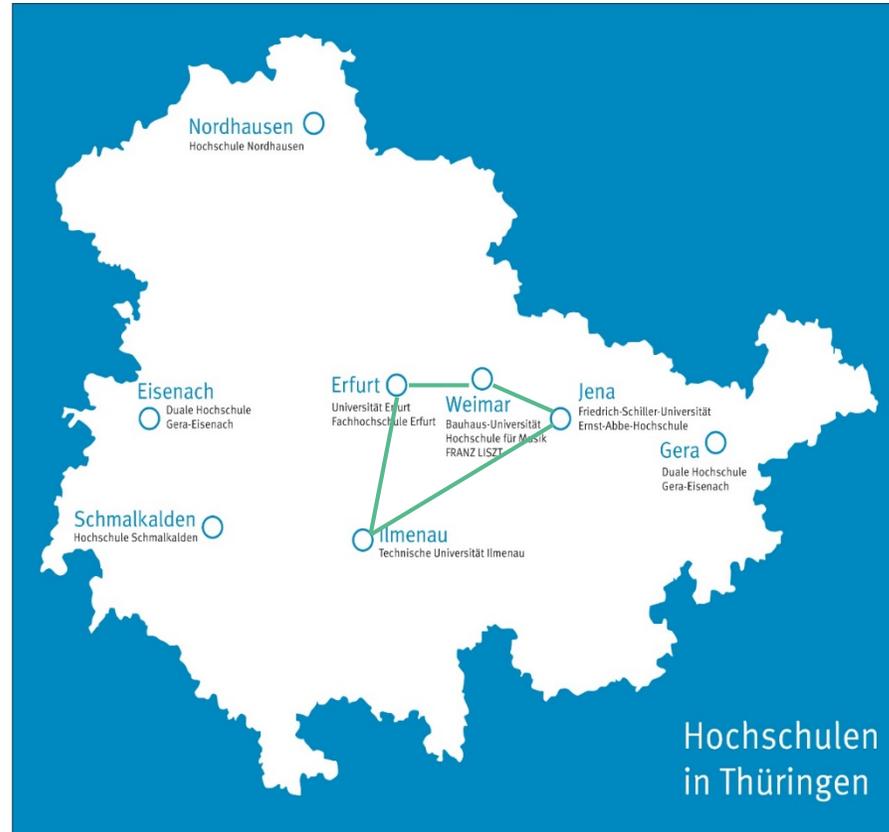
- Digitale Hochschullehre
- Digitalisierte Forschungsprozesse
- Open Access und Lizenzierung elektr. Informationsressourcen
- Forschungsinformationssysteme und Wissenstransfer
- Digitale Infrastruktur an Hochschulen



[https://www.tlpk.de/fileadmin/Downloads/Allgemeines/20171215\\_Digitalisierungsstrategie\\_Thuer\\_Hochschulen\\_signed.pdf](https://www.tlpk.de/fileadmin/Downloads/Allgemeines/20171215_Digitalisierungsstrategie_Thuer_Hochschulen_signed.pdf)

# Das Kompetenznetzwerk Forschungsdatenmanagement der Thüringer Hochschulen

- Laufzeit: 2 Jahre (seit 10/18)
- Koordination: Kontaktstelle Forschungsdatenmanagement der FSU Jena
- lokale Ansprechpartner in Ilmenau, Weimar, Erfurt und Jena
- Aufbau von Beratungs- und Schulungsangeboten sowie einer gemeinsamen Online-Präsenz
- Organisation der Thüringer FDM-Tage
- Entwicklung + Bereitstellung von Informationsmaterialien



# Ansprechpartner im Netzwerk



## Kontakt



**Roman Gerlach**  
Friedrich-Schiller-Universität Jena  
Netzwerkkordinator  
JenTower, Raum 08N01  
Leutragraben 1  
07743 Jena  
Tel.: +49 (0)3641 9-46340  
E-Mail: [roman.gerlach@uni-jena.de](mailto:roman.gerlach@uni-jena.de)  
FDM-Seite der FSU Jena



**Dr. Volker Schwartze**  
Friedrich-Schiller-Universität Jena  
Ansprechpartner:  
» Friedrich-Schiller-Universität Jena  
» Lebenswissenschaften  
» Naturwissenschaften  
» Sozialwissenschaften  
JenTower, Raum 08N01  
Leutragraben 1  
07743 Jena  
Tel.: +49 (0)3641 9-46344  
E-Mail: [volker.schwartze@uni-jena.de](mailto:volker.schwartze@uni-jena.de)  
FDM-Seite der FSU Jena



**Dr. Nadine Neute**  
Universität Erfurt  
Ansprechpartnerin:  
» Universität Erfurt  
» Geisteswissenschaften  
» Pädagogik  
» Wirtschaftswissenschaften  
Universitätsbibliothek, Raum 140  
Nordhäuser Straße 63  
99089 Erfurt  
Tel.: +49 (0)361 737-5707  
E-Mail: [nadine.neute@uni-erfurt.de](mailto:nadine.neute@uni-erfurt.de)  
FDM-Seite der Universität Erfurt



**Jessica Rex**  
Technische Universität Ilmenau  
Ansprechpartnerin:  
» Technische Universität Ilmenau  
» Ingenieurwissenschaften  
» Mathematik  
Ernst-Abbe-Zentrum (EAZ), Raum 3315  
Ehrenbergstraße 29  
98693 Ilmenau  
Tel.: +49 (0)3677 69-2502  
E-Mail: [jessica.rex@tu-ilmenau.de](mailto:jessica.rex@tu-ilmenau.de)  
FDM-Seite der TU Ilmenau



**Kevin Lang**  
Bauhaus-Universität Weimar  
Ansprechpartner:  
» Bauhaus-Universität Weimar  
» Medienwissenschaften  
» Kunst & Musikwissenschaften  
Mediathek, Raum 131  
Bauhausstr. 11  
99423 Weimar  
Tel.: +49 (0)3643 - 58 3703  
E-Mail: [kevin.lang@uni-weimar.de](mailto:kevin.lang@uni-weimar.de)  
FDM-Seite der Bauhaus-Universität Weimar



# Die FAIR-Prinzipien

---

**F**

**indable** (Auffindbar)

Persistente Adressierung

**A**

**ccessible** (Zugänglich)

Vorhaltung in öffentlichen Datenrepositorien

**I**

**nteroperable** (Interoperabel)

Verwendung einheitlicher Standards, Formate und Metadaten

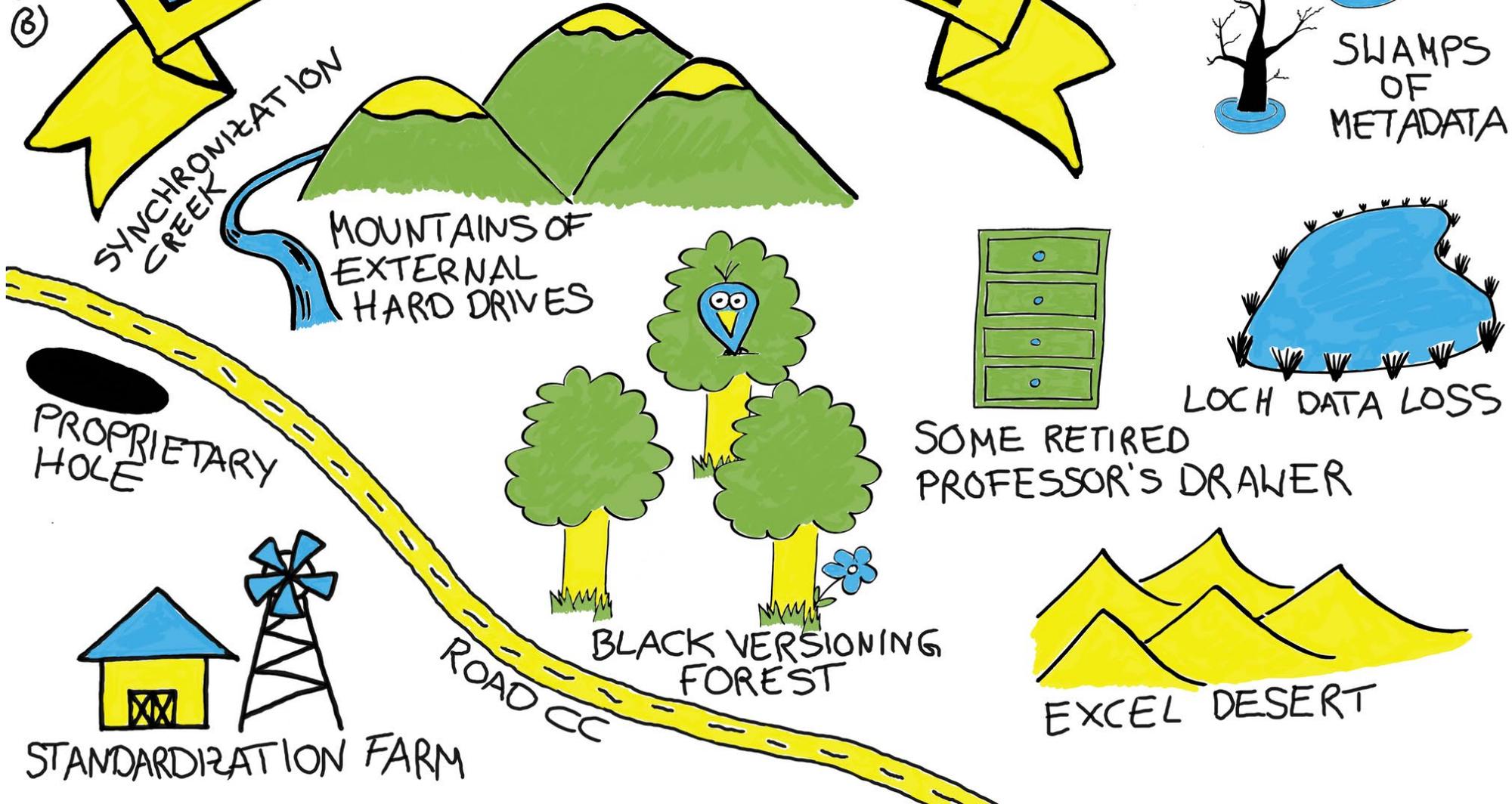
**R**

**eousable** (Wiederverwertbar)

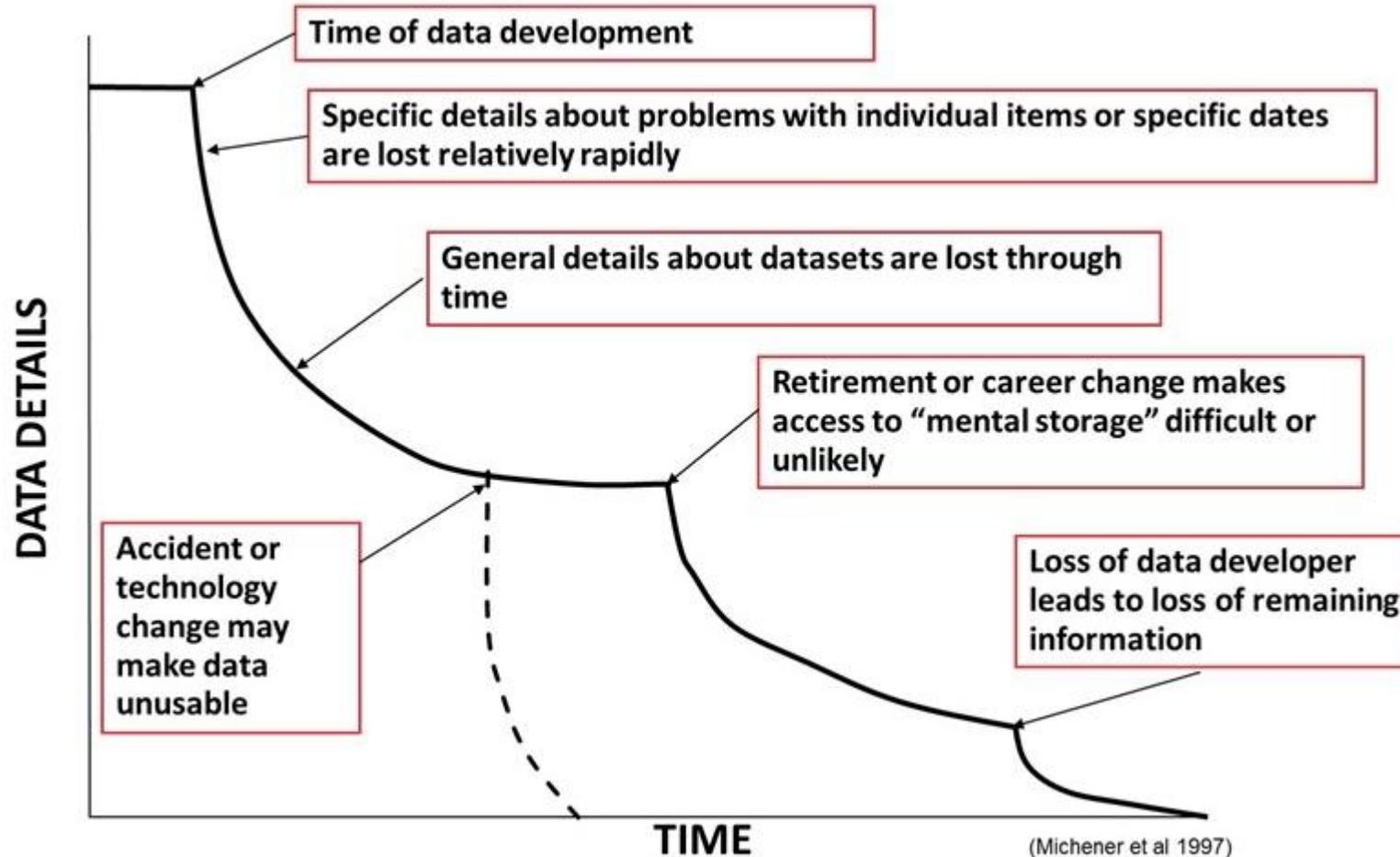
Klare, offene Lizenzregelungen; Communitystandard

Mehr Informationen auf [force11.org](https://force11.org)

# LOST DATA MAP



# Was wissen Sie noch über ihre Daten?



# 1. Aktive Phase

---

- I. Aktive Arbeit mit den Daten/Datenerfassung
- II. Speichern der Daten nach Beendung der Datenerfassung  
ggf. auch nach Beendung der Analyse für den  
spezifischen Datensatz im laufenden Projekt
- III. Speichern nach Projektende

# I – Datengenerese

---

- Konventionen erarbeiten:
  - Variablennamen (personenunabhängig, sachbezogen)
  - Kodierungen
  - Zusammenhänge
  - zeitlicher Ablauf
  - Vertrauenswürdigkeit
- Dem Bearbeiter präsent
- Der Bearbeiter ist greifbar

# I - Datengenerese

---

- Trotzdem: Konventionen vereinbaren und schriftlich erfassen
- Ordnerstruktur (hierarchisch, aber möglichst flach; inhaltsbezogen)
- Dateibenennung (systematisch; Datumsangaben im Format JJMMTT)
- Versionskennzeichnung
- Zugangssicherung über Passworte (Memnotechniken oder Tools wie Keepass)& bestenfalls automatisches Hinterlegen der Passworte für den Notfall

# I Datengenerese

---

- Umgang mit physischen Objekten
  - Forschungstagebücher
  - Interviewmitschnitte
  - Proben?
  - Etc.
- ✓ Backup über Digitalisate soweit als möglich einplanen
- ✓ ggfs. Probenmanagement über Etikettierung mit Barcode/QR-Code

# II - Datenspeicherung

---

- Klare Versionsdokumentation (SVN, GitLab, Text)
- Zusatzinformation für die **Nutzbarkeit** von Dateien
  - Format
  - Zeichenkodierung
- Zusatzinformation für die **Verständlichkeit** der Daten?
  - Maßeinheiten
  - Strukturzusammenhänge

Innerhalb einer Projektstruktur häufig nicht mitgeliefert, da alle in der Umgebung mit den Konventionen vertraut sind/dieselben Programme nutzen etc.

# II Backup (redundantes Sichern)

---

Nach der 3-2-1-Regel vorgehen:

- mind. 3 Kopien einer Datei (LOCKSS-Prinzip: Lots of Copies KeepStuff Safe)
- auf mind. 2 unterschiedlichen Medien
- wovon mind. 1 davon dezentral ist (sprich: nicht im selben Gebäude)

In regelmäßigen Abständen die Effektivität der Datenwiederherstellung/des Backups testen

# II Datensicherheit

---

Sensible und/oder personenbezogene Daten schützen durch:

- Aufbewahrung in sicheren, abschließbaren Räumlichkeiten oder Möbeln
- Passwortschutz/Verschlüsselung verwenden (unterschiedliche PW für jede zu sichernde Datei)
- sichere (!) Passwörter verwenden (ggfs. Software wie KeePass oder ein Merksystem einsetzen)
- Passwörter alle 12 Monate ändern und sicher dokumentieren
- mind. eine weitere Person sollte Zugang haben

# III – Archivierung & Veröffentlichung

---

- Rückbezug zum Projekt ermöglichen
  - Eindeutige Identifikatoren der Beteiligten (ORCID ID) und des Projektes
  - Verlinkung zu Publikationen über Digital Object Identifier o.ä.
- Veröffentlichungsbedingungen ebenfalls dokumentieren
  - Lizenzen
  - Zugangsbedingungen





## 2. Kurz nach dem Projekt (2-5 Jahre)

---

- I. Einpflegen in Repositorien (wenn nicht bei Veröffentlichung)
- II. Nutzung für weitere Projekte
- III. Replizierbarkeit

# Datenpublikationsmöglichkeiten

---

1. Fachspezifische Datenarchive, Datenzentren und Repositorien
2. Übergreifende Angebote (bspw. zenodo, Digitale Bibliothek Thüringen)
3. Datensupplemente in Fachzeitschriften
4. Eigenständige [Datenzeitschriften](#) (bspw. Biomedical Data Journal, Data Observer, Journal of Open Psychology Data, Open Health Data, Research Data Journal für the Humanities and Social Sciences)
5. Institutionelle Repositorien

# Repositorien

---

- Datenbanken, in denen digitale Objekte dokumentiert, archiviert und publiziert werden können
- disziplinspezifische, fachübergreifende und institutionelle Repositorien
- Wo finden:

**re3data.org**  
REGISTRY OF RESEARCH DATA REPOSITORIES

über 2500  
registrierte  
Repos

# Wahl eines Repositoriums

---

Generell gilt:

Disziplinspezifisches vor generischem Repository

Weitere Kriterien:

- Core Trust Seal, Data Seal of Approval (Vorgänger des Core Trust Seal) oder DINI-Zertifikat vorhanden
- Vergabe von persistenten Identifikatoren möglich (DOI, URN, o.ä.)
- Transparenz (Nutzungs- und Lizenzbedingungen werden klar genannt)
- Bei Open Data Verpflichtung: Open Access ist gewährleistet



# Digitale Bibliothek Thüringen

---

- gemeinsame Plattform Thüringer Hochschulen für das Elektronische Publizieren und Langzeitarchivieren multimedialer Dokumente und Datensätze
- Hosting und Support durch die ThULB Jena
- Publizieren dürfen alle Angehörigen der Thüringer Universitäten und Hochschulen (nach Anmeldung)



# Nutzung für weitere Projekte

---

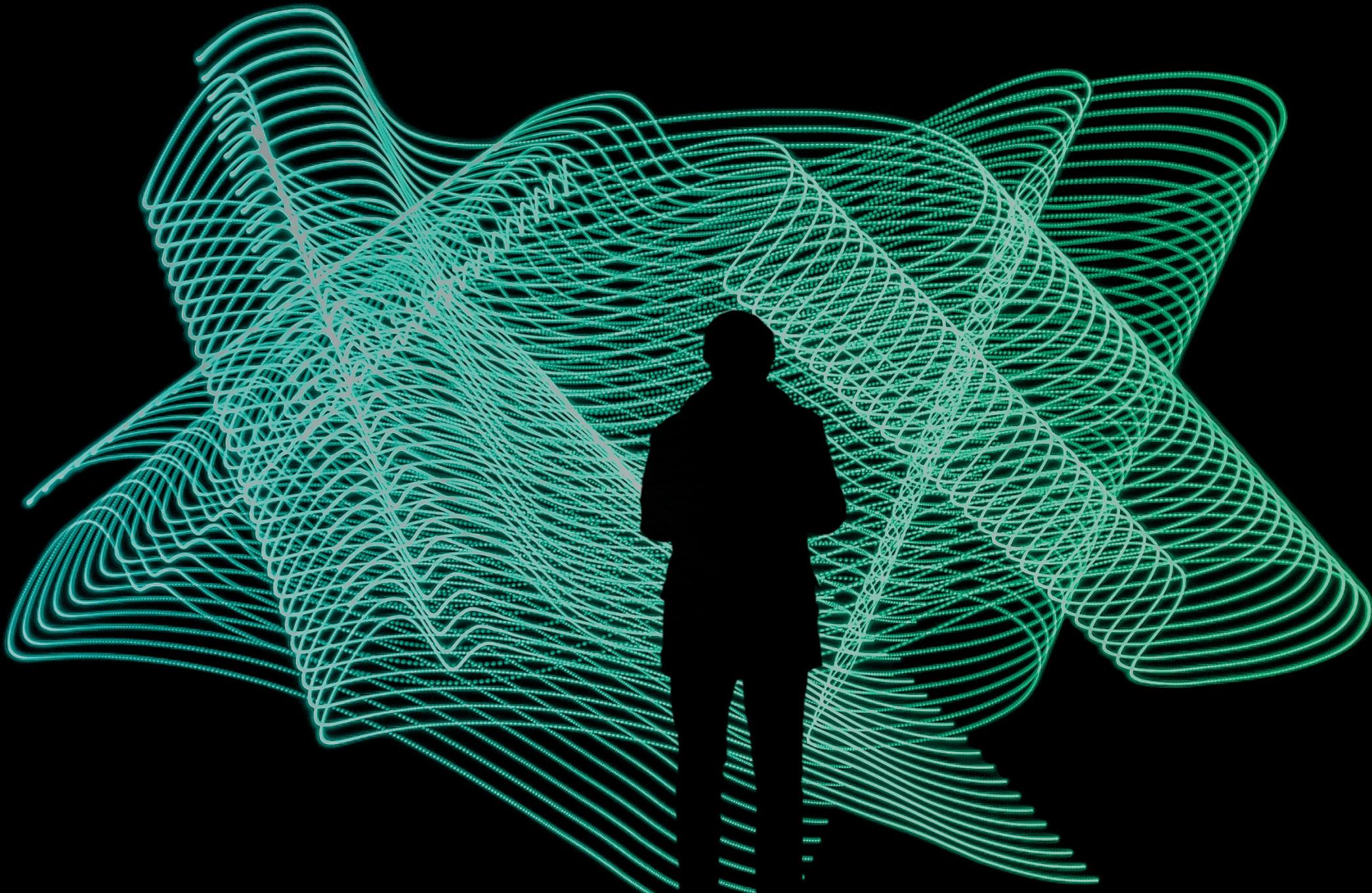
- Veröffentlichte Daten gewähren Zugang auch nach:
  - Projektende
  - Institutionswechsel der Beteiligten
  - Organisatorischen Veränderungen (bspw. Schließung von Instituten)

Rechte bezüglich der Nachnutzung wurden bereits bei Veröffentlichung geklärt (über Lizenzen)

# Replizierbarkeit

---

- Nachvollziehbarkeit des Forschungsansatzes:
  - Datenmanagementplan
  - veröffentlichte Metadaten
  - Dokumentation der Datengenerese und der Forschung (nicht zwangsläufig veröffentlicht)
- Für GWP über mindesten 10jährige Archivierung dieser Dokumente
  - gepresste DVDs, BluRays oder zukünftig Millenium Discs
  - Achtung: zentrale Archivierung an der Hochschule institutionalisieren/anfragen



# 3. Lange nach dem Projekt

---

## Ausscheiden von Personen und Technik

1. Wie gehen die Strukturen, in denen gearbeitet wurde, damit um?
2. Was muss individuell erledigt werden?

# Institutionelles Problembewußtsein

---

- Seit einiger Zeit Versuche über Policies den Umgang mit Forschungsdaten zu regeln
- Immer mehr Universitäten geben sich eigene Policies (in Thüringen u.a. eine der Pflichtaufgaben aus der Digitalisierungsstrategie)
- Parallel: Aufbau einer Forschungsdateninfrastruktur (fachspezifische und fachübergreifende Repositorien)
- Leider keine Rückwirkung...

# Policies und Leitlinien

---

- Dokumente, die den zentralen Handlungsrahmen für den Umgang mit Forschungsdaten abstecken
- Viele Förderer, Wissenschaftsorganisationen, Fachgebiete, Journals und zunehmend mehr Hochschulen haben eigene Forschungsdaten-Policies
- unterschiedliche Verbindlichkeitsgrade und Charaktere

# Policies und Leitlinien I

---

DFG:

- Leitlinien zur Sicherung guter wissenschaftlicher Praxis Kodex (2019)
- Sicherung guter wissenschaftlicher Praxis (1998/2013)
- Deutsche Forschungsgemeinschaft (DFG): Leitlinien zum Umgang mit Forschungsdaten (2015)

„[...] Wenn aus Projektmitteln systematisch Forschungsdaten oder Informationen gewonnen werden, die für die Nachnutzung durch andere Wissenschaftlerinnen und Wissenschaftler geeignet sind, legen Sie bitte dar, ob und auf welche Weise diese für andere zur Verfügung gestellt werden. Bitte berücksichtigen Sie dabei auch - sofern vorhanden - die in Ihrer Fachdisziplin existierenden Standards und die Angebote existierender Datenrepositorien oder Archive. [...]“

„Die für die Nachnutzung der Forschungsdaten anfallenden projektspezifischen Kosten können Sie im Rahmen des Projekts beantragen.“

# Policies und Leitlinien II

---

- Allianz der Wissenschaftsorganisationen: Grundsätze zum Umgang mit Forschungsdaten (2010)
- Europäische Kommission: [Guidelines on Data Management in Horizon 2020](#) (2016)  
Open Data Mandat: 100% Open Data in allen Programmbereichen ab 01.01.2017,  
Datenmanagement-Plan obligatorisch, Opt-Out möglich

**Datenverlust fügt Ihnen  
und den Menschen in Ihrer  
Umgebung Schaden zu**

Brauchen Sie Unterstützung? Wir helfen:  
[www.forschungsdaten-thueringen.de](http://www.forschungsdaten-thueringen.de)



Thüringer Kompetenznetzwerk  
Forschungsdatenmanagement

**Schlechtes  
Datenmanagement führt  
zum Verlust kostbarer  
Lebenszeit**

Brauchen Sie Unterstützung? Wir helfen:  
[www.forschungsdaten-thueringen.de](http://www.forschungsdaten-thueringen.de)



Thüringer Kom  
Forschungsdatem

**Datenverlust kann zu  
Stress und schweren  
Wutanfällen führen**

Brauchen Sie Unterstützung? Wir helfen:  
[www.forschungsdaten-thueringen.de](http://www.forschungsdaten-thueringen.de)



Thüringer Kompetenznetzwerk  
Forschungsdatenmanagement

**Datenverlust kann zu  
Kopfschmerzen und  
Panikattacken führen**

Brauchen Sie Unterstützung? Wir helfen:  
[www.forschungsdaten-thueringen.de](http://www.forschungsdaten-thueringen.de)



enznetzwerk  
management

**Schlechtes  
Datenmanagement verkürzt  
die freie Lebenszeit**

Brauchen Sie Unterstützung? Wir helfen:  
[www.forschungsdaten-thueringen.de](http://www.forschungsdaten-thueringen.de)



Thüringer Kompetenznetzwerk  
Forschungsdatenmanagement

Vielen Dank für Ihre Aufmerksamkeit.