

# Datenmanagementplan

LOVE  
DATA  
WEEK



13. – 17. 02. 2023

**zhb** Zentral- und  
Hochschulbibliothek  
Luzern

**HSLU** Hochschule  
Luzern

**BH** Berner Fachhochschule  
Haute école spécialisée bernoise  
Bern University of Applied Sciences

*u<sup>b</sup>*

UNIVERSITÄT  
BERN

 University  
of Basel

 UNIVERSITÉ  
DE GENÈVE

 DaSCH  
EPFL

# Datenmanagementplan

*Den Umgang mit Forschungsdaten planen:  
dokumentieren, organisieren, schützen,  
publizieren - mach es FAIR*



Clemens Trautwein  
Fachstelle FDM HSLU Bibliotheken  
17. Februar 2023

LOVE



DATA  
WEEK

13. - 17. 02. 2023



**zhb** Zentral- und  
Hochschulbibliothek  
Luzern



**HSLU** Hochschule  
Luzern



**BH** Berner Fachhochschule  
Haute école spécialisée bernoise  
Bern University of Applied Sciences



**EPFL**

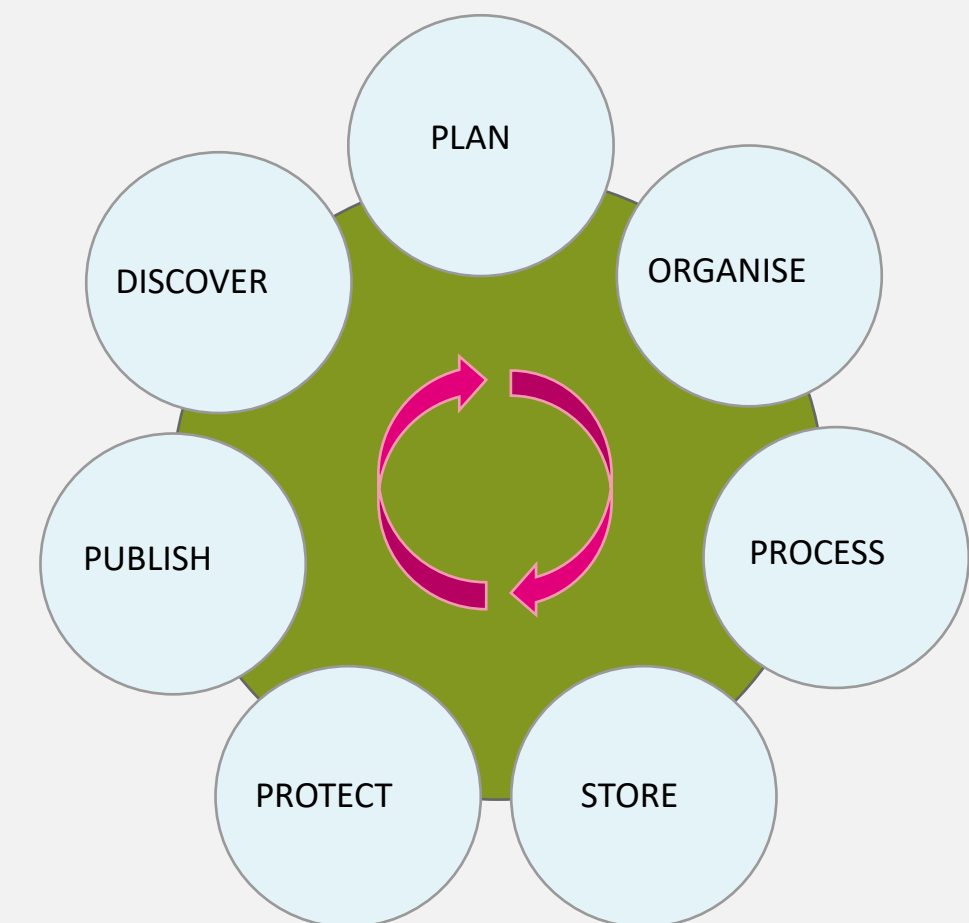


# Inhalte

Datenmanagementplan

Warum ein DMP?

Lebenszyklus von Forschungsdaten



[www.cessda.eu/DMGuide](http://www.cessda.eu/DMGuide)



# Was ist ein Datenmanagementplan?

**«Schriftliches Dokument,  
das den Umgang mit Daten in einem Forschungsprojekt  
über alle Stationen des Lebenszyklus beschreibt.»**

## **SNF Datenmanagementplan:**

<https://www.snf.ch/de/FAiWVH4WvpKvohw9/thema/forschungspolitische-positionen>

The DLCM project: <https://www.dlcm.ch/resources/dlcm-dmp>

Anleitung der ETH zum Datenmanagementplan des SNF: <https://documentation.library.ethz.ch/guidance-snsf-dmp>

## **Checklisten:**

[EPFL RDM guides and templates](#)

[Digital Curation Centre \(DCC\): Checklist Data Management Plan](#)

## **Beispiele**

Collection of DMPs of students at TU Wien <https://zenodo.org/communities/dast-2021/?page=1&size=20>

# Wie erstellt man einen DMP? Bestandteile

## 4 Hauptbereiche

1. Datenerhebung und Dokumentation
2. Ethische, rechtliche und Sicherheitsfragen
3. Datenspeicherung und Erhalt
4. Datenaustausch und Zugänglichkeit

## Leitfragen

1. Welche Daten werden wie erhoben, beobachtet, generiert, wiederverwendet?  
Welche Metadaten sehen sie vor?
2. Wie gestalten sich der Umgang mit und die Behandlung von ethischen Fragen und Urheberrechtsfragen?  
Wie werden der Zugriff auf Daten und die Datensicherheit verwaltet?
3. Wie werden Ihre Daten während der Forschungsarbeiten gespeichert und gesichert?  
Wie gestaltet sich Ihre Planung für den Datenerhalt?
4. Wie und wo werden die Daten zugänglich gemacht?  
Sind bestimmte Einschränkungen erforderlich, um sensible Daten zu schützen?



## Weshalb Forschungsdatenmanagement?

### **Verdacht auf Datenmanipulation**

Letztes Jahr wurde ein ETH-Biologe wegen gravierender Fehler beim Veröffentlichen von Forschungsergebnissen verurteilt. Neue Verdachtsmomente betreffen Fachartikel, an denen dieser Biologe beteiligt war.

### **Datenfälschung**

## **Datenfälschung an der ETH aufgedeckt**

ZÜRICH - An der ETH Zürich sind drei Forschungsarbeiten gefälscht worden.

### **Datenfälschung: Ex-Professor der Uni Zürich unter Verdacht**

Der Schweizer Chemieprofessor Reto Dorta soll die Co-Autorin einer Arbeit aufgefordert haben, Daten zu fälschen. Die Universität Zürich will die Vorkommnisse prüfen.

<https://www.tagesanzeiger.ch/zuerich/stadt/Datenfaelschung-ExProfessor-der-Uni-Zuerich-unter-Verdacht-/story/23027273>

<https://www.blick.ch/news/schweiz/datenfaelschung-datenfaelschung-an-der-eth-aufgedeckt-id1502292.html>

<https://www.nzz.ch/zuerich/aktuell/faelschungsvorwuerfe-in-mikrobiologie-eth-untersucht-fachpublikationen-ld.115622>

# Weshalb Forschungsdatenmanagement?

Schutz vor Verlust

Durch technische Defekte, Unfall, Brand, Diebstahl

Unberechtigten Zugriff und Löschung...

## Was sind Forschungsdaten?

***Forschungsdaten sind (digitale) Daten, die während wissenschaftlicher Tätigkeit (z. B. durch Messungen, Befragungen, Quellenarbeit) entstehen. Sie bilden eine Grundlage wissenschaftlicher Arbeit und dokumentieren deren Ergebnisse.***

Das können Messergebnisse, Erschließungsdaten wissenschaftlicher Sammlungen, Studienterhebungen, Digitalisate historischer Quellen, Zellkulturen, Materialproben, die Dokumentation archäologische Funde oder Fragebögen sein.



# Was sind Forschungsdaten?



9) **Wie viele Bücher gibt es ungefähr bei dir zu Hause?**  
 Zähle Schulbücher nicht mit!

keine

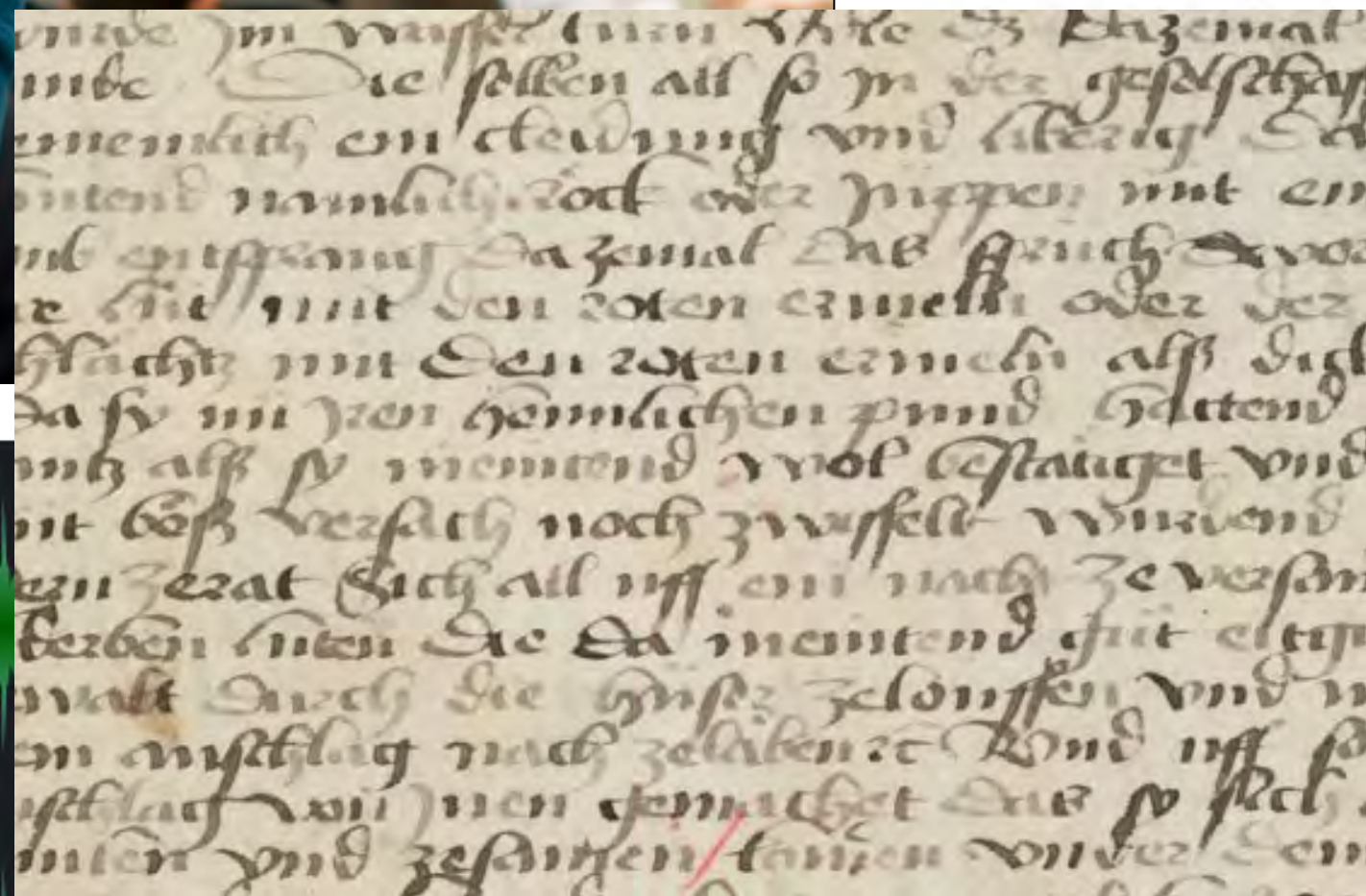
1 - 10 Bücher       101 - 250 Bücher

11 - 50 Bücher       251 - 500 Bücher

51 - 100 Bücher       mehr als 500 Bücher

10) **Wurden die Schülerinnen und Schüler in deiner Klasse auf verschiedene Englischkurse aufgeteilt, je nach dem, wie gut sie in Englisch sind?**  
 (z.B. in einen A-, B- und C-Kurs, oder Grundkurs und Erweiterungskurs, oder Grundkurs und Leistungskurs)

ja       nein (dann geh weiter zu Frage 1)

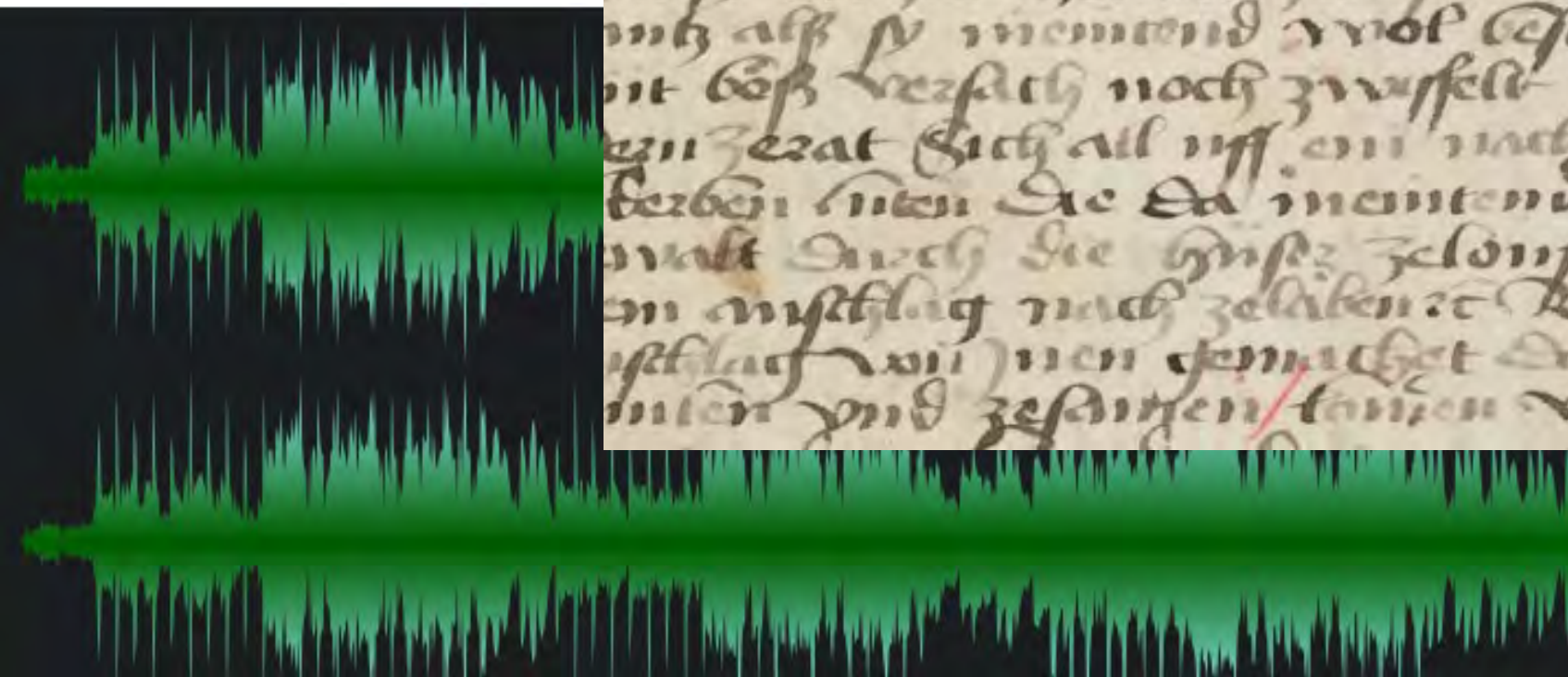


en Kurs besuchst du zurzeit?  
 rs  
 Kurs

Winddaten						
Windgeschw. in m/s		Windrichtung				
Windgeschw.	Imm.-Ort	Nabenhöhe				
6	0,9	236				
2	0,4	249				
1	0,1	166 - 178				
0	1,3	58	308 - 115			
1	1,1	50	321 - 104			
0	0,9	52	316 - 97			
03:00 - 04:00	7,4	1,3	40 - 60	311 - 69	Gegenwind	43,3
05:00 - 06:00	9,7	1,3	59	296 - 101	Gegenwind	44,8
05:00 - 06:00	8,8	k. M.	194	k. M.	Mitwind	40,9
04:00 - 05:00	10,5	k. M.	263	k. M.	Mitwind	41,8

**Dynamic MTs in PDUS** 03.04.2023

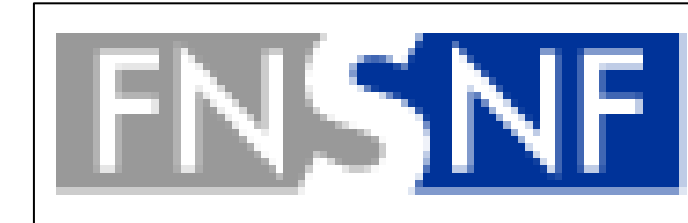
1) **MT seeds** - 5µl Biotinylated Biotinylated tubulin (lab. ratio 2%)  
 2) **Casine** - 10µl aliquot aus Dize lab  
 3) **Imaging buffer** - GD- and C-suspensions according to protocol  
 4) **Tubulin mix** - according to protocol (18µl tubulin)  
 5) **Final solution** - 50µl MT seeds + 50µl imaging buffer  
 6) **PDUS** - 1<sup>st</sup> chamber -> no holes (forgetten)  
 2<sup>nd</sup> chamber -> all cuts right on 1<sup>st</sup> attempt  
 -> no holes much right  
 -> plasma according to protocol



M. keine Messergebnisse



# Forschungsförderer & Forschungsdaten



## Schweizerischer Nationalfonds

«Forschungsdaten sollen offen und für alle zugänglich sein - für die Wissenschaft wie für die Gesellschaft.»




Deshalb erwartet der SNF von allen von ihm geförderten Forschenden, dass sie

- › die Forschungsdaten, an denen sie gearbeitet und die sie während ihrer Forschungsarbeit produziert haben, archivieren,
- › diese Daten anderen Forschenden zugänglich machen, sofern dem keine rechtlichen, ethischen oder urheberrechtlichen Gründe, Vertraulichkeitsklauseln o. ä. entgegenstehen,
- › ihre Daten und Metadaten in bestehenden öffentlichen Datenbanken in Formaten veröffentlichen, die es jedem Interessierten ohne Einschränkungen ermöglicht, sie zu finden, darauf zuzugreifen und sie weiterzuverwenden.

## ➤ FAIRer Umgang mit Forschungsdaten

# FAIR Data Principles

Graphik: [Paulina Halina Sieminska](#) / Bearbeitung: Clemens Trautwein / [CC BY-SA 4.0](#)

<p>Auffindbar (<b>F</b>indable)</p> 	<p>Zugänglich (<b>A</b>ccessible)</p> 	<p>Interoperabel (<b>I</b>nteroperable)</p> 	<p>Wiederverwendbar (<b>R</b>eusable)</p> 
<p>umfangreiche Metadaten, eindeutige Kennung</p>	<p>Zugänglich (Meta- ) Daten über ein offenes Protokoll</p>	<p>allgemeine Standards für Daten/Metadaten</p>	<p>Zitierbarkeit und eindeutige Lizenz</p>

Übersichtsartikel und Hilfestellung zur Umsetzung der FAIR Prinzipien:

<https://forschungsdaten.info/themen/veroeffentlichen-und-archivieren/faire-daten/>

FAIR Data Principles bei Force11 (2016)

<https://www.force11.org/group/fairgroup/fairprinciples>

# Forschungsdaten !

Daten bewerten  
und für die Fragestellung  
aufbereiten

Datennutzung (Analyse)

Daten erzeugen  
oder suchen

Teilen und Publizieren

Datenablage

Archivieren

Externe Nachnutzung





# Warum einen DMP?

Dokumentation und Diskussionsgrundlage

- Mögliche Fragen im Vorhinein beantworten

Aufwand macht sich später bezahlt

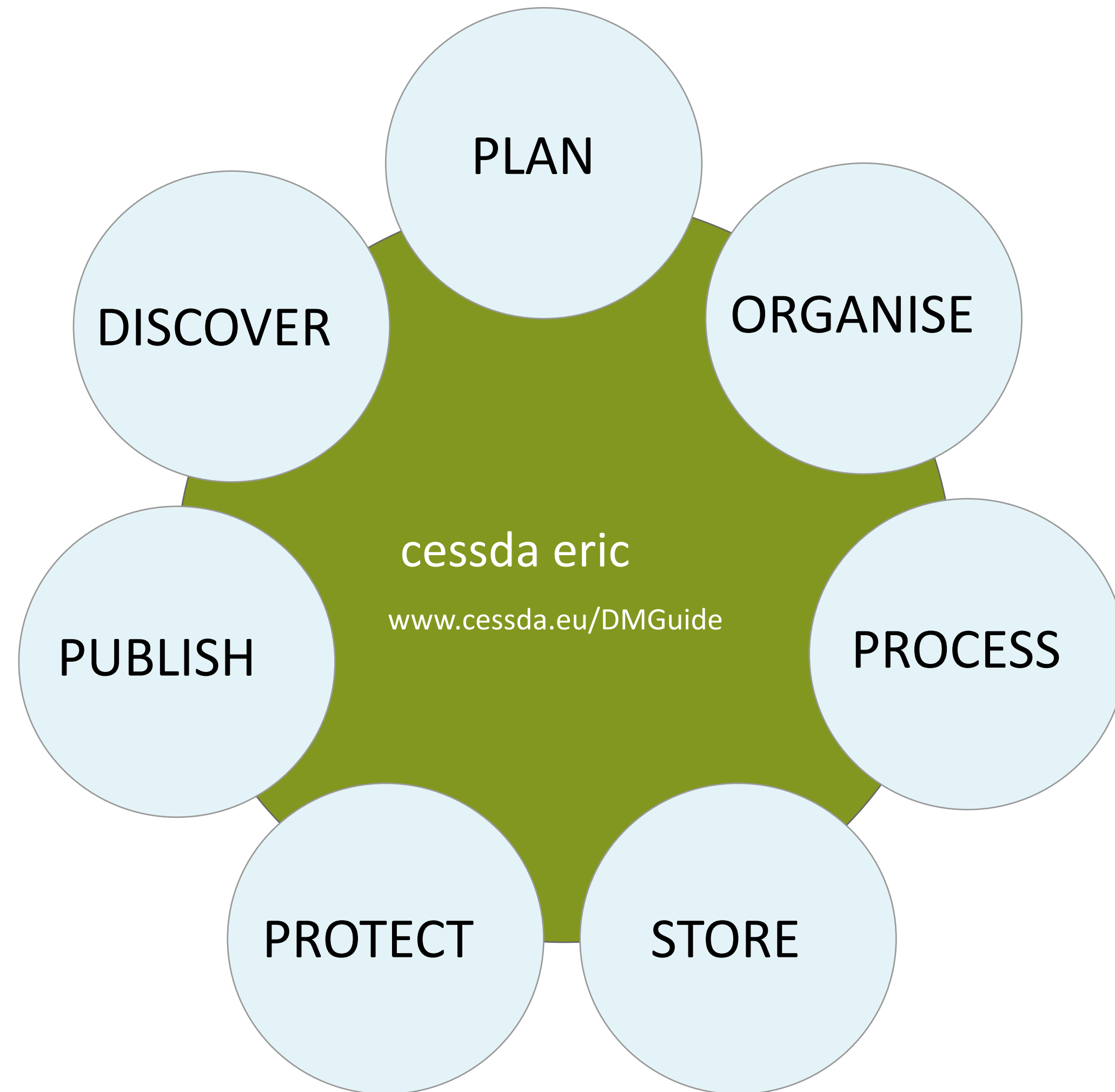
Lebendiges Dokument. Änderungen einpflegen

Kann jederzeit erstellt werden (am besten vorgängig, aber es ist nie zu spät)

## **Stakeholder/ Anspruchsgruppen**

- Forschungsteam
- Institution
- Forschungsförderer
- Partner
- Wirtschaftspartner
- IT

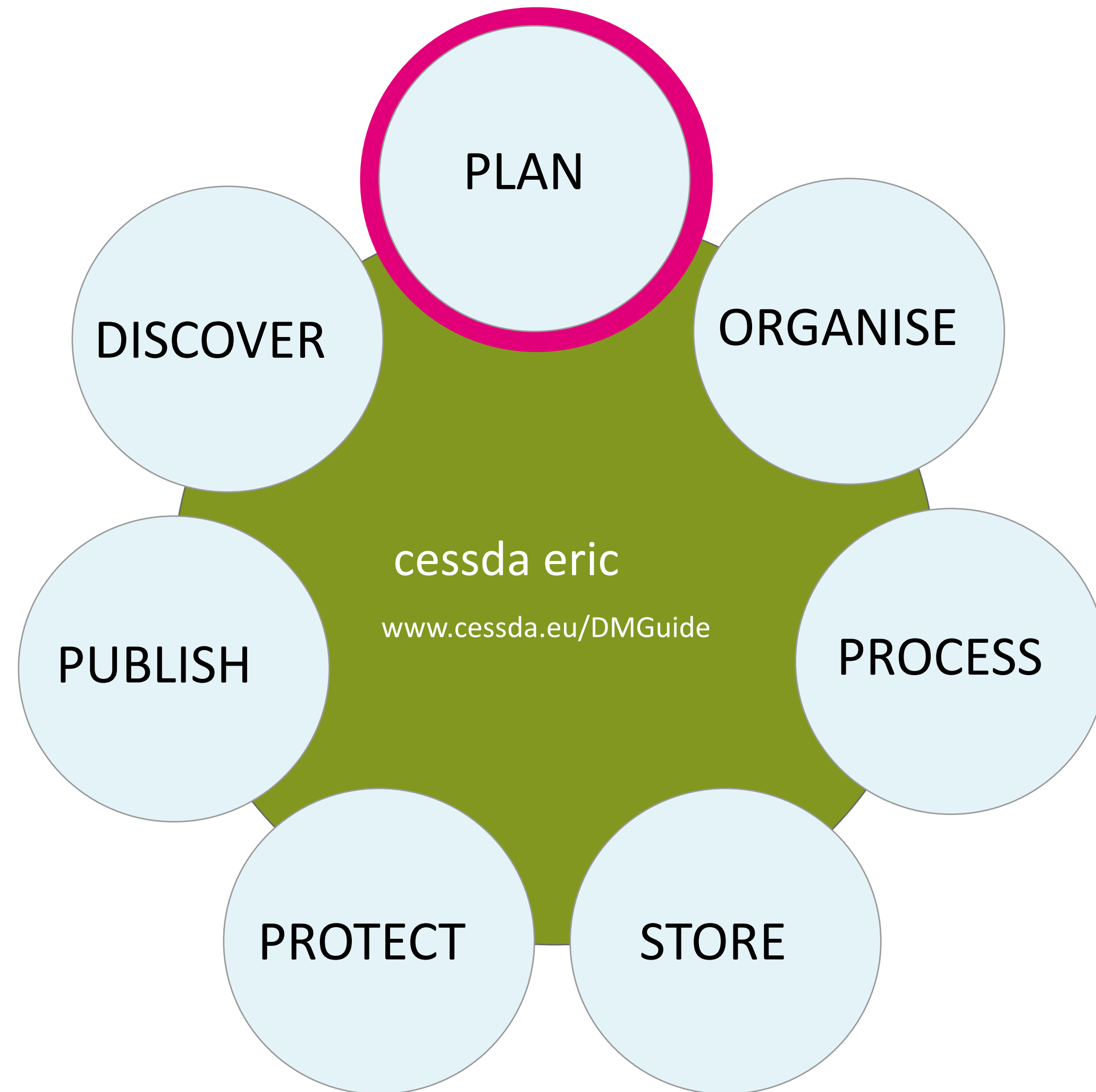
# Lebenszyklus der Forschungsdaten



## Forschungsdaten

- Management?
- Komplex: Zusammenspiel Forschung, IT und Recht
- Bezeichnet den Prozess die Forschungsdaten während ihres Lebenszyklus zu verwalten:
  - Planung
  - Organisation & Dokumentation
  - Datenverarbeitung & Datensicherung
  - Ethik & Recht
  - Publikation, Nachnutzung & Archivierung
- Ziel: Verantwortungsbewusster, nachhaltiger und FAIRer Umgang mit Forschungsdaten

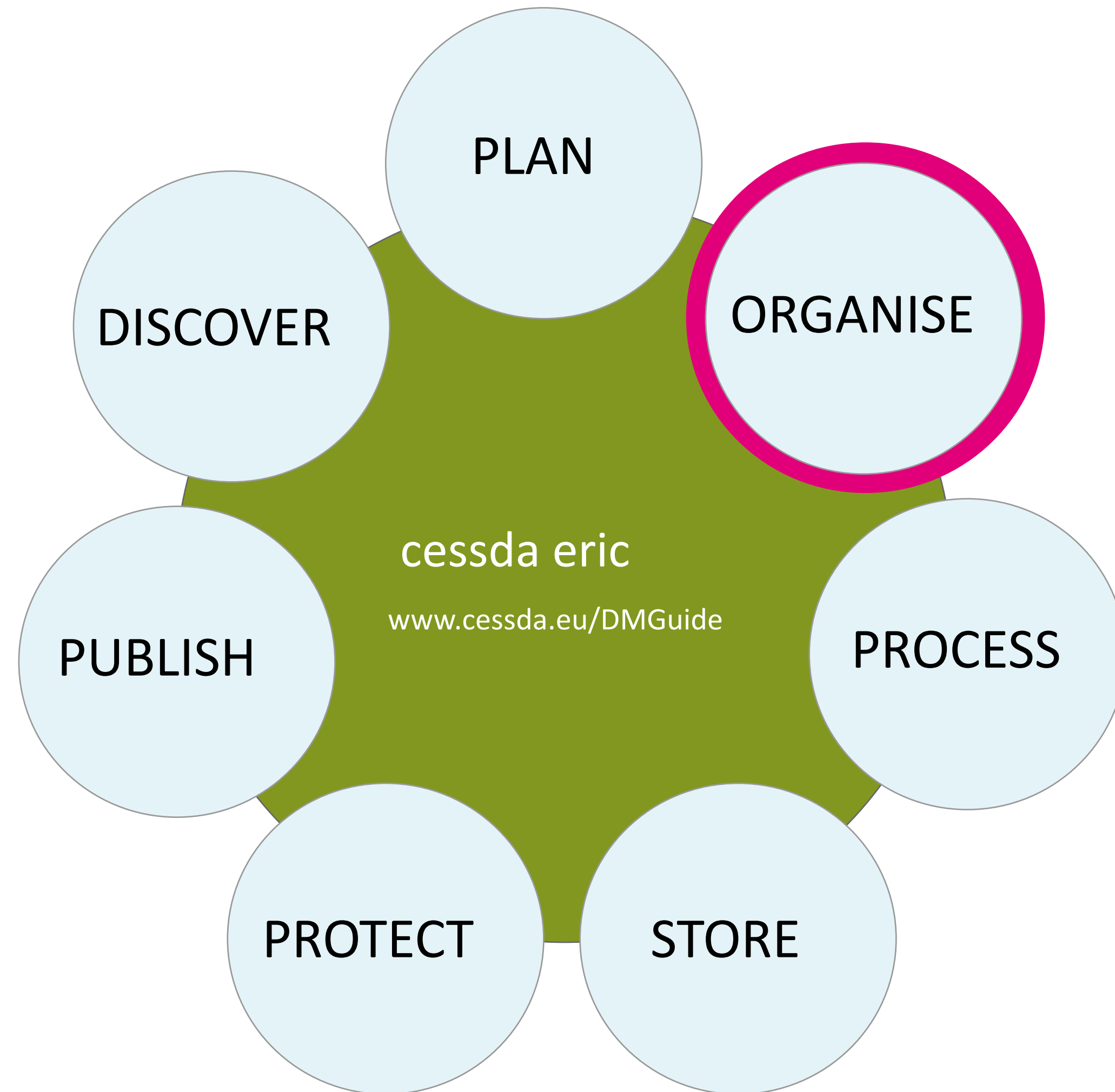
# Lebenszyklus der Forschungsdaten



## Planung

- Effizientes Forschungsdatenmanagement bedingt gute Planung
- Datenmanagementplan (DMP):
  - Klare Struktur
  - Identifikation von Problemen
  - verantwortungsbewusster Umgang mit Forschungsdaten
- Erstellung DMP zu Projektbeginn und kontinuierliche Überarbeitung

# Lebenszyklus der Forschungsdaten



## Organisation & Dokumentation

- Ablage- und Ordnersystem
- Einmal gewählte Organisation, Struktur, Benennung beibehalten.
- Dokumentation der Forschungsdaten und Metadaten wichtig für Reproduzierbarkeit & Nachnutzung



# Organisation von Forschungsdaten: Ablagestruktur

Daten müssen verständlich und wiederauffindbar sein für:

- einen selber
- Kollegen/innen
- weitere/spätere Interessierte

Datenablage und Ablagestruktur planen

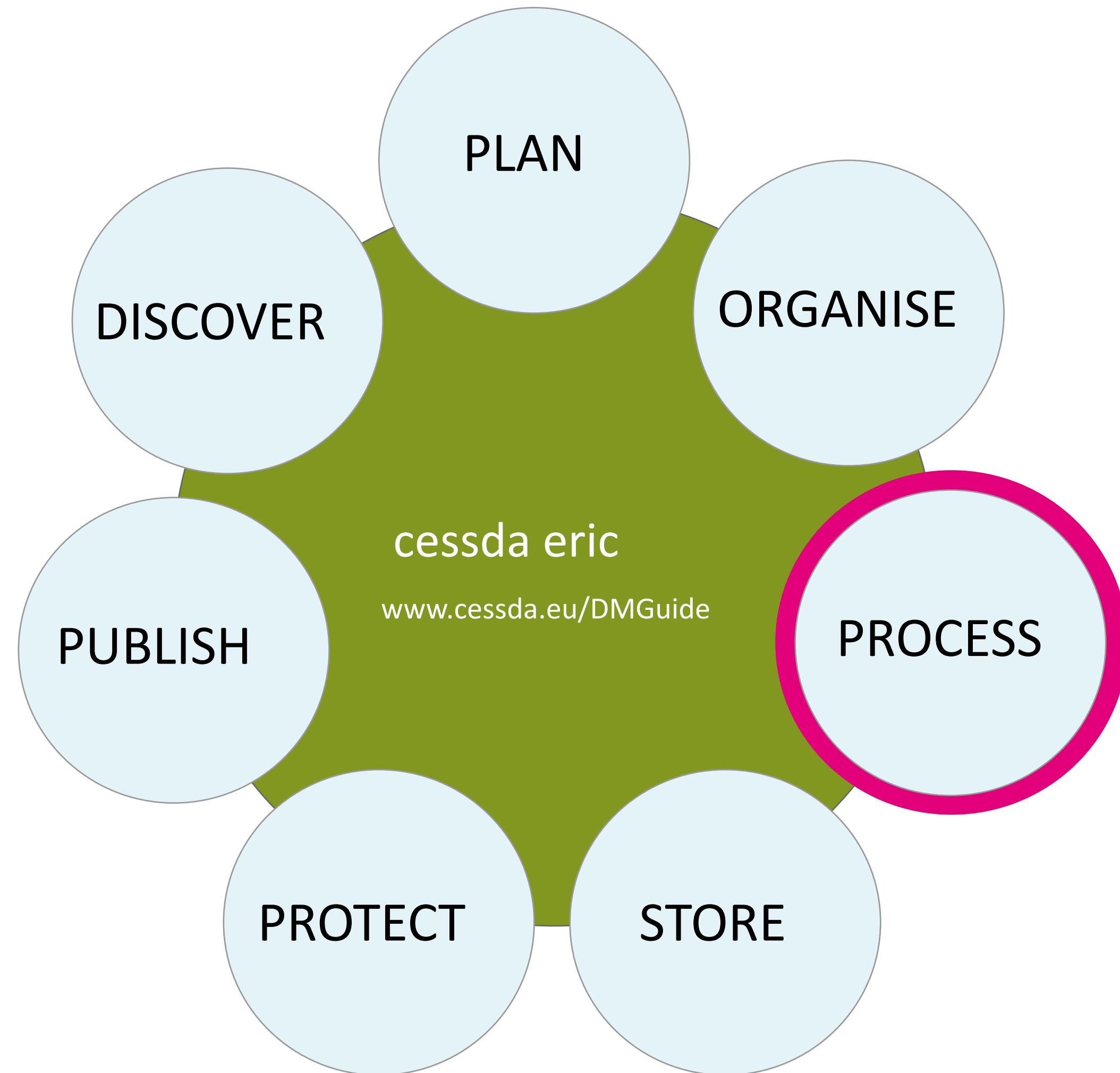
Datenstruktur von Anfang an planen:

- ✓ Nach welchen Kriterien soll strukturiert werden....
- ✓ Konsistente sprechende und informative Namenskonventionen
- ✓ Versionenkontrolle (Versionsnummer und Speicherdatum in Format YYYYMMDD-Format)
- ✓ Bei Teamarbeiten: «Meilensteinversionen», die nicht mehr geändert werden, an speziellem Ort speichern.

Beispiele:

<https://www.jisc.ac.uk/guides/managing-information/good-file-name>

# Lebenszyklus der Forschungsdaten



## Datenverarbeitung

- Während Forschungsprojekt: einfache & rasche Bearbeitung
- Nach Forschungsprojekt: Nachhaltigkeit
- short term data processing vs. long term data preservation
- Dateiformat sollte den verschiedenen Bedürfnissen Rechnung tragen

## Dateiformate?

**Datenverarbeitung**      bearbeitbare, z.T. proprietäre Formate

**Nachnutzung**          bearbeitbare, möglichst nicht proprietäre Formate

**Archivierung**          für Langzeitarchivierung geeignete, möglichst nicht proprietäre Formate

evtl. Speicherung/Publizierung in mehreren Formaten

# Datenformate für Langzeitarchivierung



<b>Text</b>	MS Word Open/Libre Office MS Power Point	PDF A/1-b
<b>Tabelle</b>	Excel	CSV
<b>Bild</b>	Jpeg Png GIF	TIFF PDF A/1-b
<b>Audio/Video-Material</b>	Windows Player	MPEG-4_AVC WAV

Empfehlungen der ZHB:

<http://www.zhbluzern.ch/dienstleistungen/forschen-publizieren/>

Archivtaugliche Dateiformate (ETH Fachstelle digitaler Datenerhalt)

<https://documentation.library.ethz.ch/display/DD/Archivtaugliche+Dateiformate>



# Lebenszyklus der Forschungsdaten



## Datensicherung

- Ohne Forschungsdaten keine Forschung
- Schutz vor Datenverlust und Datenmanipulation
- Datenspeicherung ist abhängig von den produzierten Forschungsdaten (Datenmenge, Datengrösse, gemeinsames Projekt? → Notebook, Externe Festplatte, Cloud etc.)
- Regelmässige Datensicherung schützt vor Datenverlust und Datenbeschädigung

# Lebenszyklus der Forschungsdaten



## Datenschutz & Urheberrecht

- Bei Erhebung, Verwendung, Aufbewahrung und Publikation von Daten sind rechtliche und ethische Richtlinien zu beachten. Insbesondere:
  - Datenschutz
  - Urheberrechte

## Ethik & Datenschutz



Bei Forschung am/mit Menschen:

- ✓ Entsprechende Erlaubnis eingeholt (z.B. Ethikkommissionen?)
- ✓ Einwilligungserklärung (Hinweis auf den Zweck der Datenerhebung und allfällige Publikation der Daten)
- ✓ Anonymisierung der Daten

Ethikkommission HSLU: beurteilt ethische Unbedenklichkeit eines Projekts, wenn von Forschenden oder Dritten verlangt

<https://inside.hslu.ch/rk/Seiten/InfoWiki/Ethikkommission.aspx>

Weiterführende Informationen:

- Schweizerische Ethikkommissionen:  
<http://www.swissethics.ch/index.html>
- Textbausteine informierte Einverständniserklärungen:  
[https://www.forschungsdaten-bildung.de/get\\_files.php?action=get\\_file&file=fdbinfo\\_4.pdf](https://www.forschungsdaten-bildung.de/get_files.php?action=get_file&file=fdbinfo_4.pdf)

## Urheberrecht



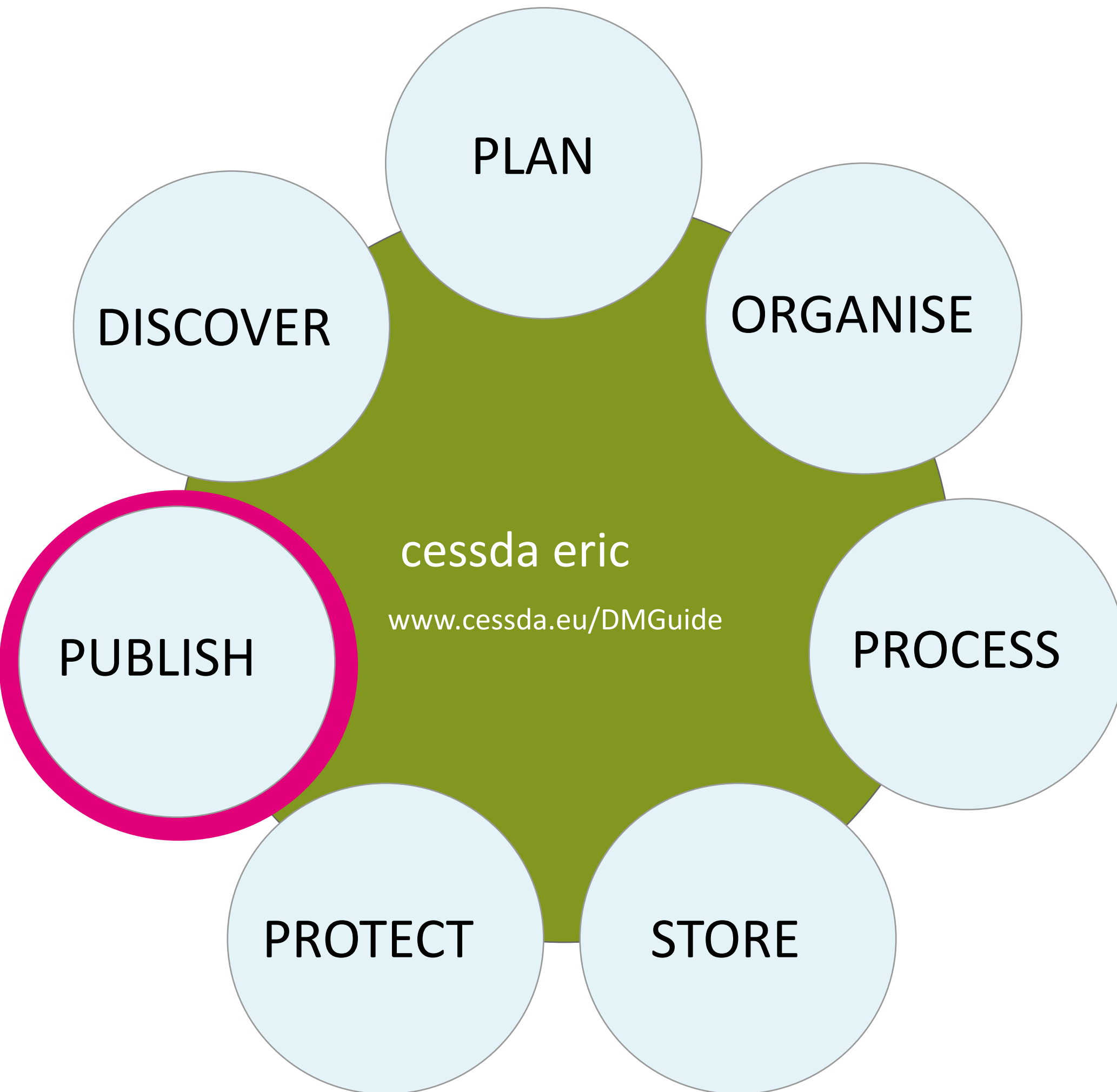
- ✓ Urheberrechtliche Vorgaben der HSLU:  
*«Für urheberrechtlich geschützte Werke, welche die Mitarbeitenden während der Ausübung der dienstlichen Tätigkeit schaffen, steht ihnen das Urheberrecht zu. Der Hochschule Luzern steht daran ein zeitlich unbeschränktes, unentgeltliches und umfassendes Verwendungsrecht zu.»*
- ✓ Vorgaben von anderen Institutionen?
- ✓ Bei mehreren Beteiligten: Ist abgeklärt, wem die Daten gehören und was mit den Daten geschehen soll/darf?
- ✓ Lizenzierung der eigenen Daten: [Creative Commons](#)

### Tipps:

- Rechtliche Lage am Anfang klären: Wem gehören die Daten - besonders wenn Partner dabei sind!
- Geplanter späterer Umgang mit Daten (z.B. Veröffentlichung) bei Einverständniserklärungen, Verträgen etc. einbeziehen.



# Lebenszyklus der Forschungsdaten



## Publikation

- Reproduzierbarkeit der Forschungsergebnisse → Transparenz
- Nachnutzung für weitere Forschungsprojekte
- Förderung der Reputation
- Publikation auf einem fachspezifischen oder institutionseigenen Repository (z.B. LORY)
- Langzeitarchivierung

# Warum Daten teilen/publizieren?



  
CC-BY Danny Kingsley & Sarah Brown

[Danny Kingsley & Sarah Brown](#), CC BY 4.0 via Wikimedia Commons

## Daten publizieren: Welche Daten?

Daten, die einer Publikation zugrunde liegen

Einmalige Daten

Daten, die mit hohem Aufwand/hohen Kosten generiert wurden

Zur Weiterverwendung geeignete Daten

➤ Nicht alle Daten können zugänglich gemacht werden (Datenschutz, Urheberrecht, Geheimhaltungsverpflichtungen, Patentrechte etc.)

➤ Nicht alle Daten sollen und müssen erhalten bleiben (Qualitätsmanagement)

# Wo kann ich meine Daten publizieren?

## Beispiele **Fachrepositorien**:

Beispiele für Repositorien können der institutionelle Publikationsserver einer Universität, ein fachliches (Open-Access) Repository, ein fachliches Datenrepositorium oder ein Langzeitarchiv für Daten und Publikationen sein.

–Liste institutioneller und disziplinspezifischer Repositorien in der Schweiz

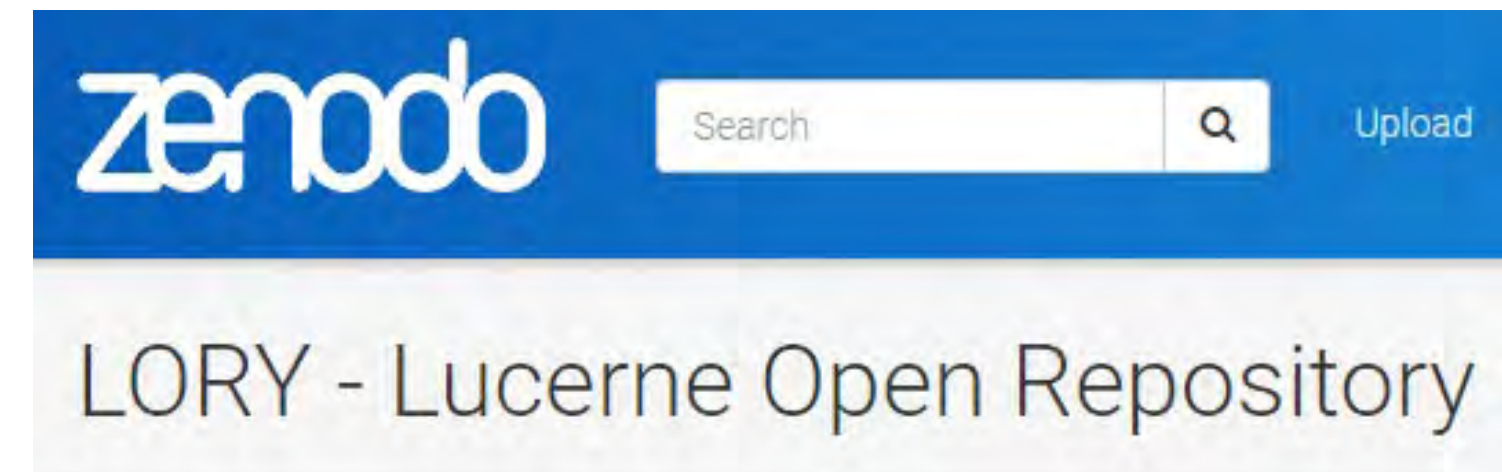
– [forschungsdaten.info/schweiz](https://forschungsdaten.info/schweiz)

–Campus Luzern: [Lucerne Open Repository](https://lory.ch/)

## Übersicht:

Registry of Research Data Repositories:

<https://www.re3data.org/>





July 2, 2021

Dataset **Open Access**

# Knowledge-intensive firms, connectivity and spatial restructuring: dynamics and differences in Germany and Switzerland

Bentlage, Michael; Heidinger, Mathias; Lüthi, Stefan; Müller, Christiane; Thierstein, Alain; Wenner, Fabian; Zöllner, Silke

This dataset is based on the interlocking network model developed by the Globalization and World Cities Study Group (GaWC) at Loughborough University (Taylor 2004). It provides one specific way to address the question how inter-city relations can be empirically measured despite the chronic lack of data on inter-city information flows. The model uses a proxy – intra-firm networks of multi-branch, multi-location enterprises – to estimate potential flows of knowledge-creating information between cities and towns. The method was originally developed to measure the connectivity between global cities based on multi-branch advanced producer services (APS) firms as they organize business activities across their offices worldwide. In this dataset, the model is adapted to measure relations between functional urban areas (FUAs) within and beyond the German and Swiss space economy, for both APS and High-Tech firms. In contrast to the original methodology, in this dataset the company locations were localized by the exact, georeferenced address, not only on the level of FUAs.

The present dataset is created as part of the SNF research project "Knowledge-intensive firms, connectivity and spatial restructuring: dynamics and differences in Germany and Switzerland".

Files (257.6 MB)	
Name	Size
R-Matrices_FINAL.zip	257.4 MB
md5:e09e03d7bcd85fc2c4f1a0f5849ee68e	
Read_me_FINAL.pdf	206.2 kB
md5:32a00f5b9c5a4975945ca5ae79e3e614	

Daten in zip-Paket

Metadaten in 'ReadMe' file

Preview

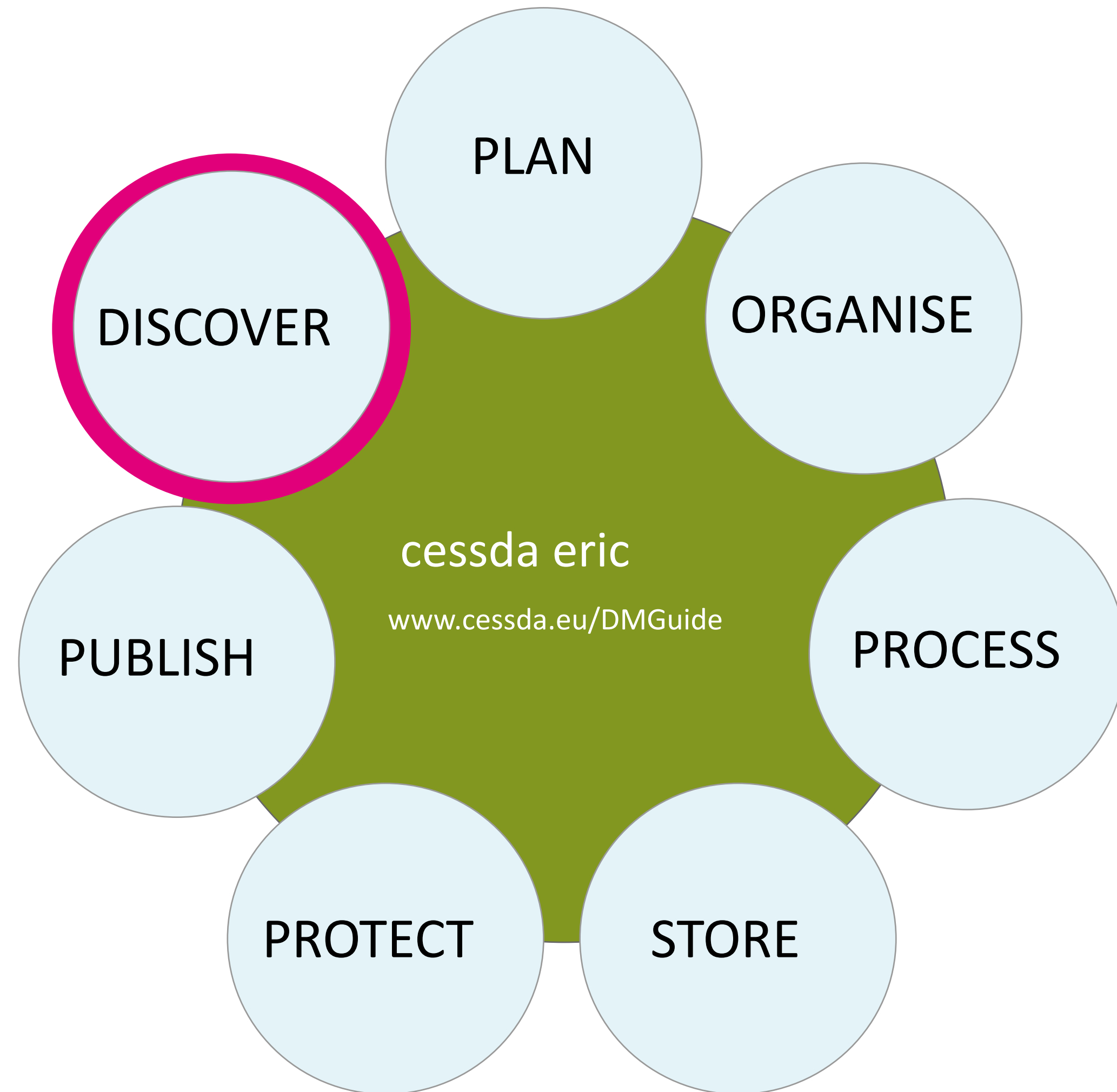
R-Matrices\_FINAL.zip

R-Matrices\_FINAL

- Germany
  - Advanced\_Producer\_Services
    - R-Matrix\_Germany\_Accounting\_2019\_FINAL.csv 15.8 MB
    - R-Matrix\_Germany\_Accounting\_2019\_FINAL.xlsx 23.7 MB
    - R-Matrix\_Germany\_Advertising-Media\_2019\_FINAL.csv 15.1 MB
    - R-Matrix\_Germany\_Advertising-Media\_2019\_FINAL.xlsx 21.1 MB
    - R-Matrix\_Germany\_Banking-Finance\_2019\_FINAL.csv 15.2 MB
    - R-Matrix\_Germany\_Banking-Finance\_2019\_FINAL.xlsx 21.3 MB
    - R-Matrix\_Germany\_Design-Architecture\_2019\_FINAL.csv 15.2 MB
    - R-Matrix\_Germany\_Design-Architecture\_2019\_FINAL.xlsx 22.7 MB
    - R-Matrix\_Germany\_Information-Communication\_2019\_FINAL.csv 15.2 MB
    - R-Matrix\_Germany\_Information-Communication\_2019\_FINAL.xlsx 21.4 MB
    - R-Matrix\_Germany\_Insurance\_2019\_FINAL.csv 15.2 MB
    - R-Matrix\_Germany\_Insurance\_2019\_FINAL.xlsx 21.4 MB
    - R-Matrix\_Germany\_Law\_2019\_FINAL.csv 15.2 MB
    - R-Matrix\_Germany\_Law\_2019\_FINAL.xlsx 21.2 MB
    - R-Matrix\_Germany\_Logistics\_2019\_FINAL.csv 15.4 MB



# Lebenszyklus der Forschungsdaten



## Daten Nachnutzen

- Prüfen bisheriger Ergebnisse
- Weitere Untersuchungen mit bestehenden Daten
- Bestehende Forschung rezensieren
- Lehre

## Daten nachnutzen – Wo finden?

Verzeichnis von Datenrepositories:

<https://www.re3data.org/>

### **Portale**

Datenportale und Index-Services ermöglichen die Suche über mehrere Repositorien hinweg

[DataCite](#) - Suche von registrierten Datensätzen

[European Union Open Data Portal](#) - Offenes Datenportal der Europäischen Union

Google Dataset Search

<https://datasetsearch.research.google.com/>

## Daten zitieren

- **Grundsätzlich:**  
Daten werden ähnlich wie ein Buch zitiert.
- **Minimale Angaben:**  
Creator/Autor/in (Publikationsjahr): Titel. Publisher. Identifier.
- **Optimale Angaben:**  
Creator/Autor/in (Publikationsjahr): Titel. Version. Publisher. Typ der Ressource. Identifier. Zugriffsdatum.

### **Beispiel:**

Zwally, H.J., R. Schutz, C. Bentley, J. Bufton, T. Herring, J. Minster, J. Spinhirne, and R. Thomas. 2003. *GLAS/ICESat L1A Global Altimetry Data V018, 15 October to 18 November 2003*. National Snow and Ice Data Center. Data set accessed 2011-07-21 at doi:10.3334/NSIDC/gla01.

## Bezug zu Open Science

**Open Science:** offener Zugang zu wissenschaftlichen Informationen für jeden

**Open Access:** unbeschränkter online Zugang zu wissenschaftlicher Literatur

**Open (Research) Data:** Daten, die für alle zugänglich sind und weiter ohne Einschränkungen genutzt werden können

**Research Data:** Daten, die im Forschungsprozess erhoben und bearbeitet wurden; mit oder ohne Zugriffsmöglichkeiten

<https://www.zhbluzern.ch/love-data-week>

Hochschule Luzern  
Technik & Architektur  
Bibliothek  
Clemens Trautwein  
Fachreferent Bibliothek

T direkt +41 41 349 32 71  
clemens.trautwein@hslu.ch

LOVE



DATA  
WEEK

13. - 17. 02. 2023

