



Projekt Forschungsdatenmanagement - FoDaMa



Elfi Hesse

elfi.hesse@htw-dresden.de

0351 462-2067



Juliane Baier

juliane.baier@htw-dresden.de

0351 462-3839

fdm@htw-dresden.de

www.htw-dresden.de/fdm

Data Sharing and Management „Horror-Story“

Video

<https://www.youtube.com/watch?v=N2zK3sAtr-4>

Und wer sind Sie?

Was erwarten Sie von diesem Workshop?

Haben Sie Fragen/Themen mitgebracht?



Forschungsdaten

alle (digitalen) Daten, die während eines Forschungsprozesses generiert, erhoben oder genutzt werden

Forschungsdatenmanagement

Forschungsdatenmanagement umfasst alle Aktivitäten von der

- Planung eines Forschungsprojekts über die
- Generierung,
- Erhebung,
- Verarbeitung,
- Speicherung,
- Langzeitarchivierung bis zur
- Veröffentlichung und
- Nachnutzbarmachung von Forschungsdaten.

→ verantwortungsvoller Umgang mit Forschungsdaten



Audio



Video



Beobachtungs-
daten



Fotos



Text



Simulations-
daten



Fallstudien



Befragungs-
daten



Messdaten



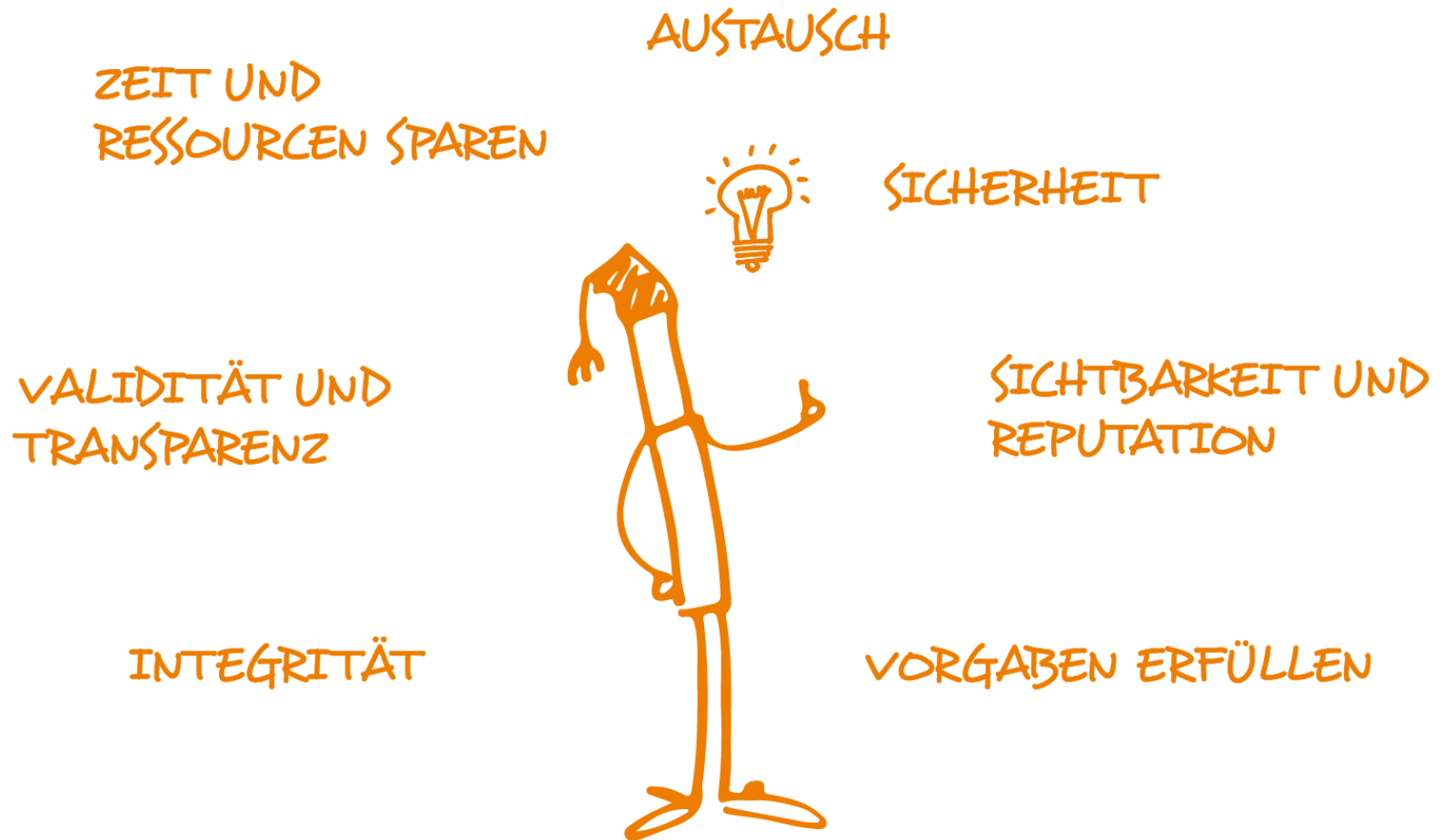
Laborwerte



Software



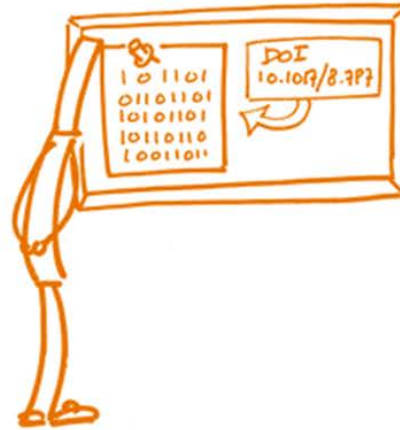
Karten



FAIR DATA PRINCIPLES



FINDABLE



ACCESSIBLE



INTEROPERABLE



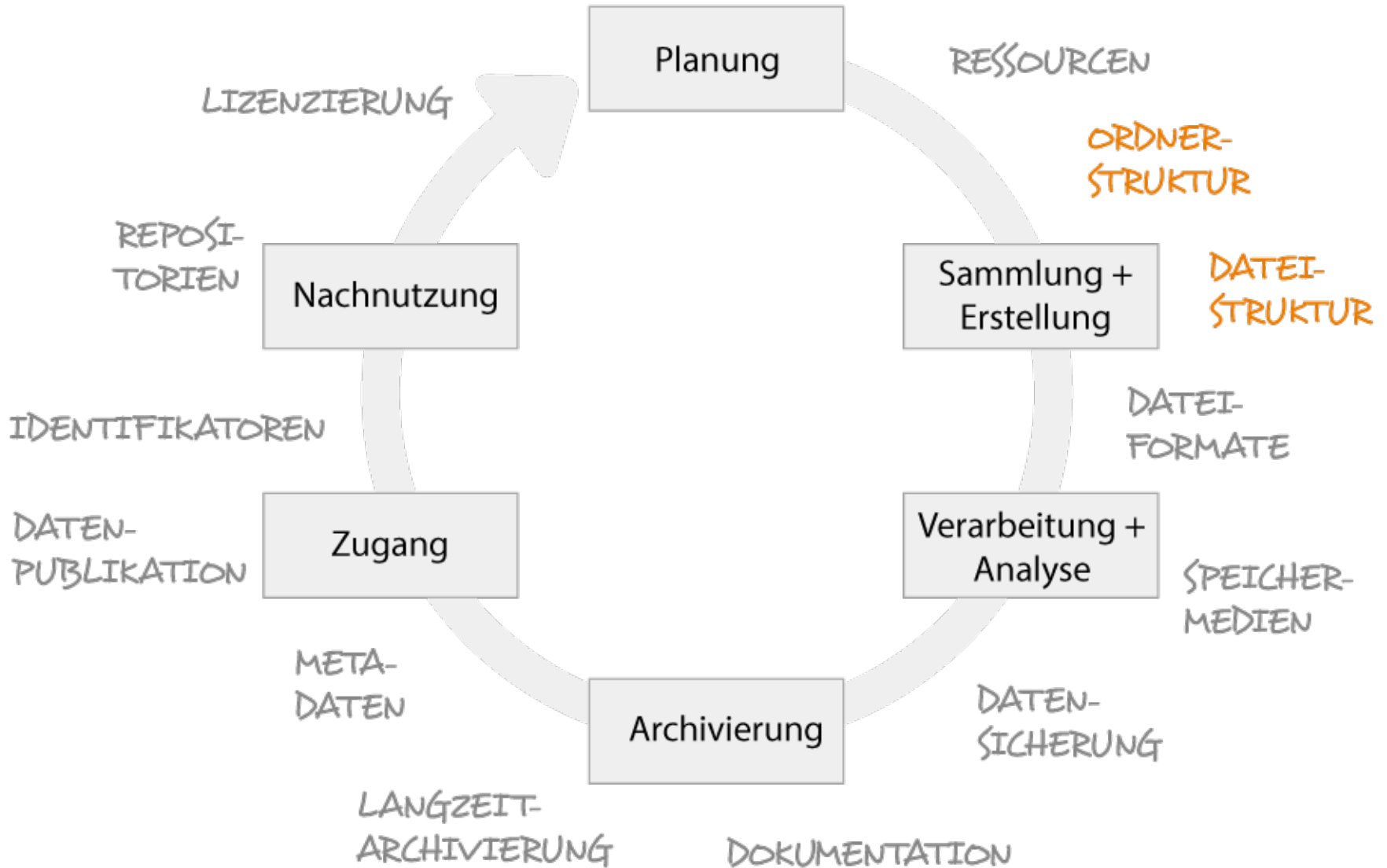
REUSABLE

- Findable (Auffindbarkeit)
- Accessible (Zugänglichkeit)
- Interoperable (Interoperabilität)
- Reusable (Nachnutzbarkeit)

DATENMANAGEMENTPLAN



DATENMANAGEMENTPLAN



Einfache Grundregel

Benennen und verwalten Sie Ihre Dateien so, dass Sie (auch langfristig) und andere Personen diese Dateien wieder auffinden können.

- spart Zeit und Aufwand bei der täglichen Arbeit mit (Forschungs-) Daten
- vereinfacht die Nachnutzung von Daten durch klare Strukturen
- verringert die Gefahr von Datenverlust

Wo haben Sie Ordnung und Struktur Ihrer Daten schmerzlich vermisst? Notieren Sie (anonym) ein Beispiel.

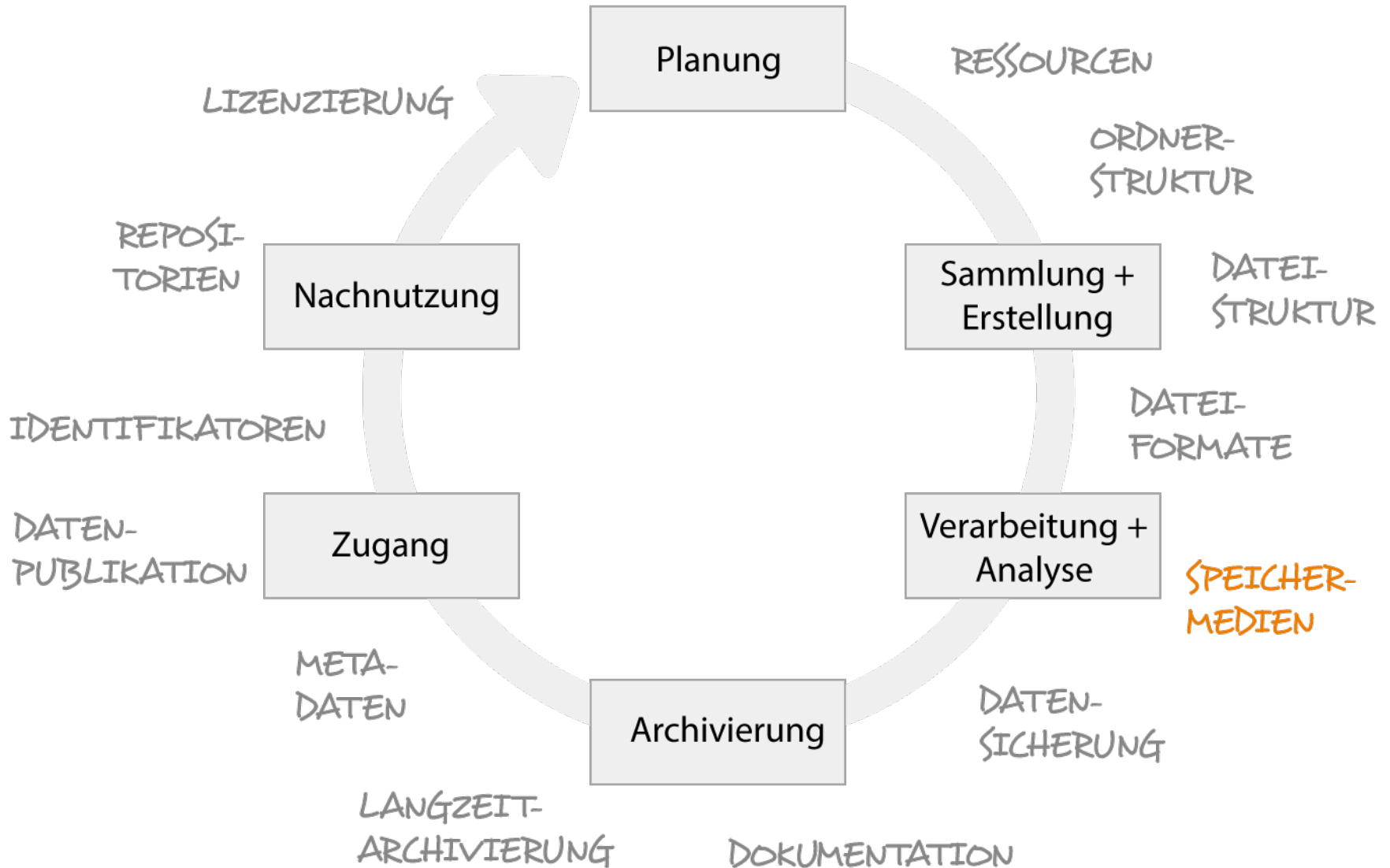


Tipps zur Ordnung und Struktur Ihrer Dateien:

- intuitiv verständliche Ordnerstruktur
- eindeutige Dateibenennung
- Verwendung von Namenskonventionen
- Kennzeichnung verschiedener Dateiversionen
- Dokumentation von Abkürzungen
- eine weitere Person sollte Zugriff haben



DATENMANAGEMENTPLAN

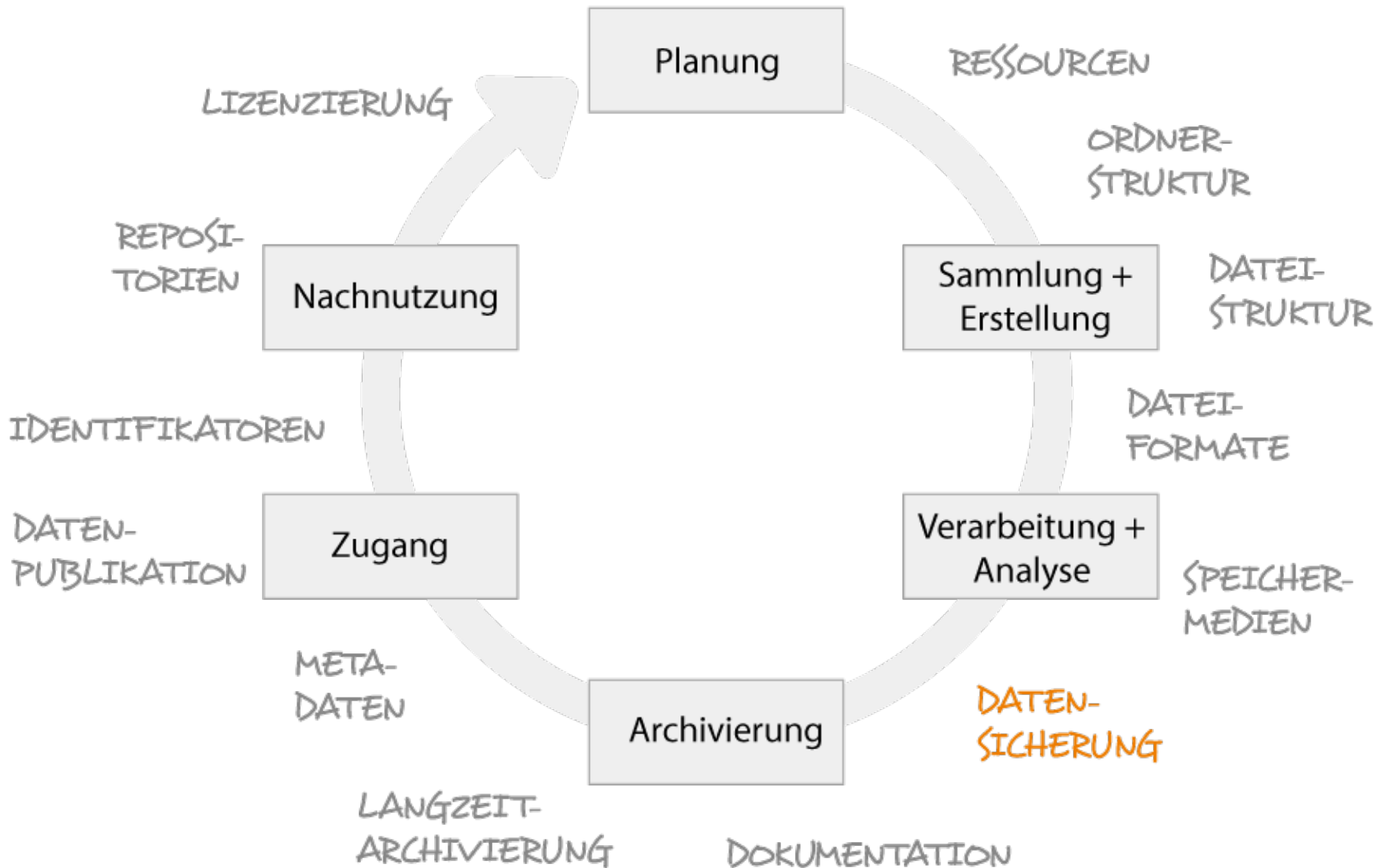


Wohin mit den Daten?

Sammeln Sie Vor- und Nachteile der üblichen Speichermedien: eigener PC, mobiles Speichermedium (z.B. USB-Stick), Netzlaufwerk der HTW Dresden und externe Cloud.



DATENMANAGEMENTPLAN

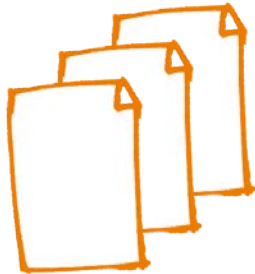


Lesen Sie sich in Gruppen die Beispiele durch. Präsentieren Sie den Fall anschließend im Plenum. Was hätte getan werden können, um Datenverlust zu verhindern?



3, 2, 1... BackUp!

3



3 Kopien

2



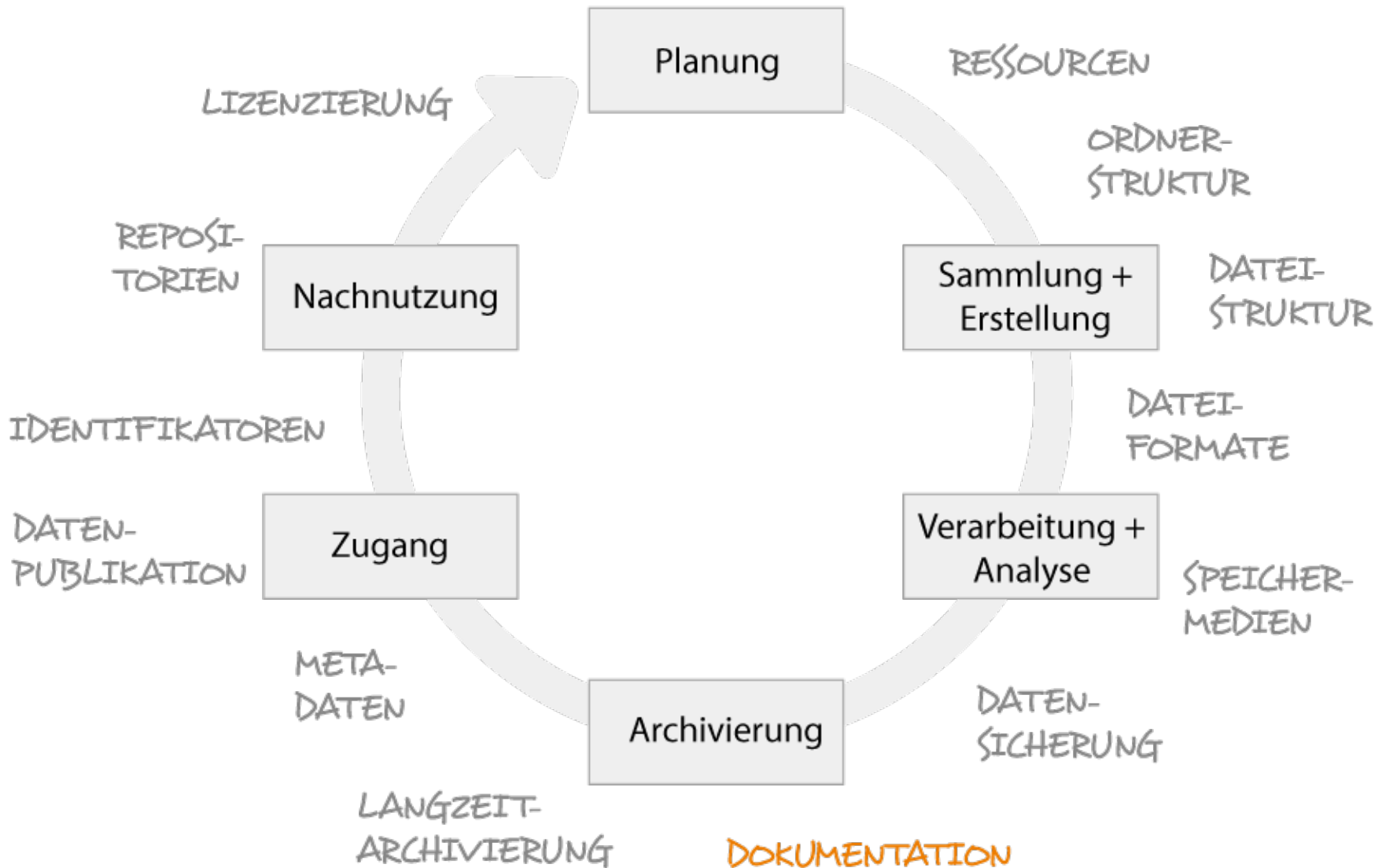
2 verschiedene
Speichermedien

1



1 davon an einem
anderen Ort

DATENMANAGEMENTPLAN



= Informationen über Entstehungsprozess, Inhalt und
Nachnutzungsmöglichkeiten von Daten

Dokumentieren Sie immer, was Sie tun!



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

→ bessere Nachvollziehbarkeit

→ bessere Auffindbarkeit

DATENMANAGEMENTPLAN



Daten sollten mindestens 10 Jahre verfügbar sein

Relevante Frage: „Lässt sich die Datei in 10 Jahren noch öffnen?“

Welche Dateien sollen archiviert werden?

- Schwerpunkt auf wichtige Daten setzen
- nicht alle Dateien müssen archiviert werden



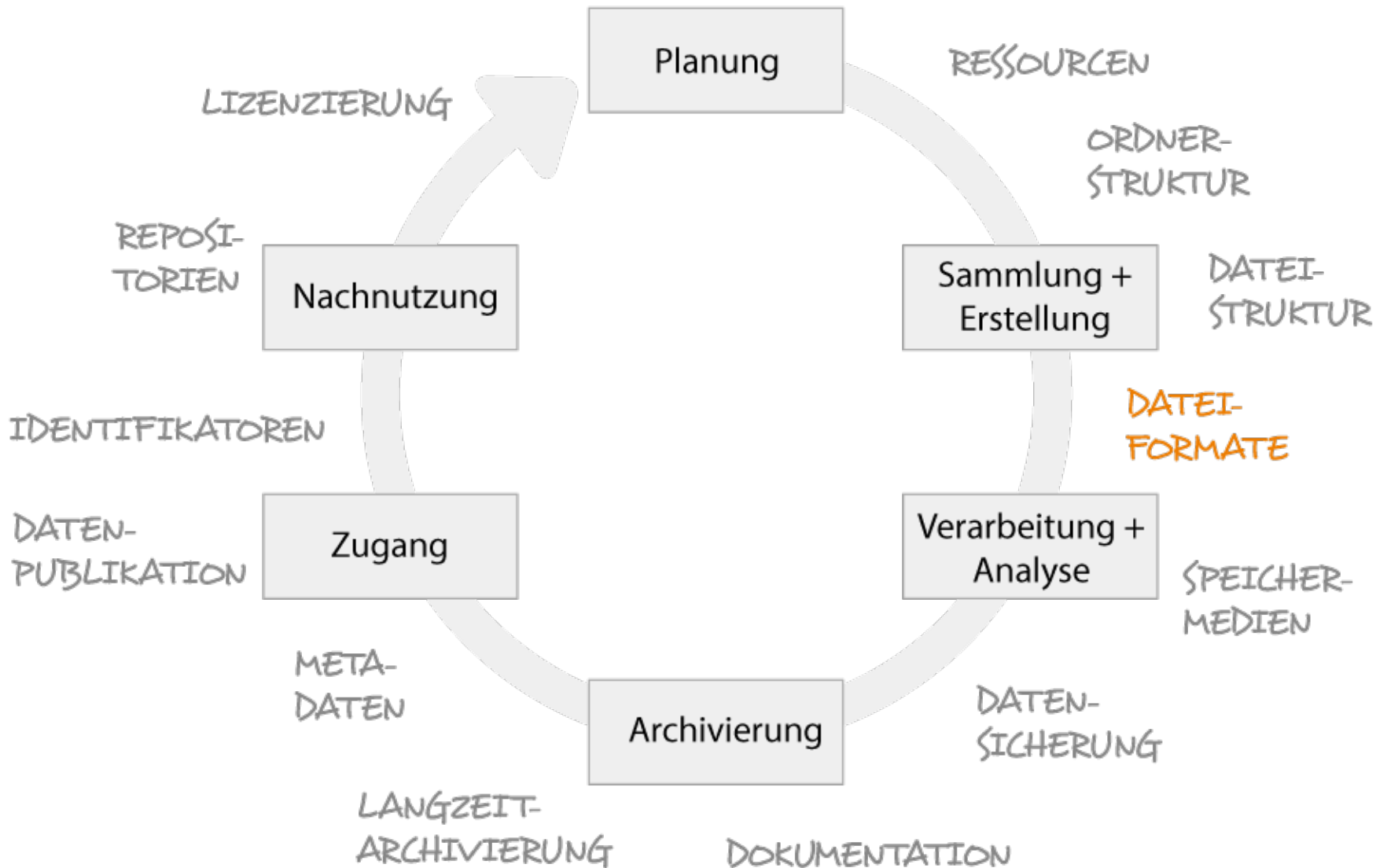
Manchmal sinnvoll: Archivierung von Software

Welche Speichermedien sollen verwendet werden?

- Lebenszeiten verschiedener Datenträger meist 5-10 Jahre
- Verwendung eines Repositoriums (online)

Welche **Dateiformate** eignen sich für eine Langzeitarchivierung?

DATENMANAGEMENTPLAN



Eine Frage des Formats...

Digitale Dateiformate entwickeln sich mit der Zeit weiter:
Forschungsdaten lassen sich dann im schlimmsten Fall nicht mehr
öffnen oder bearbeiten.

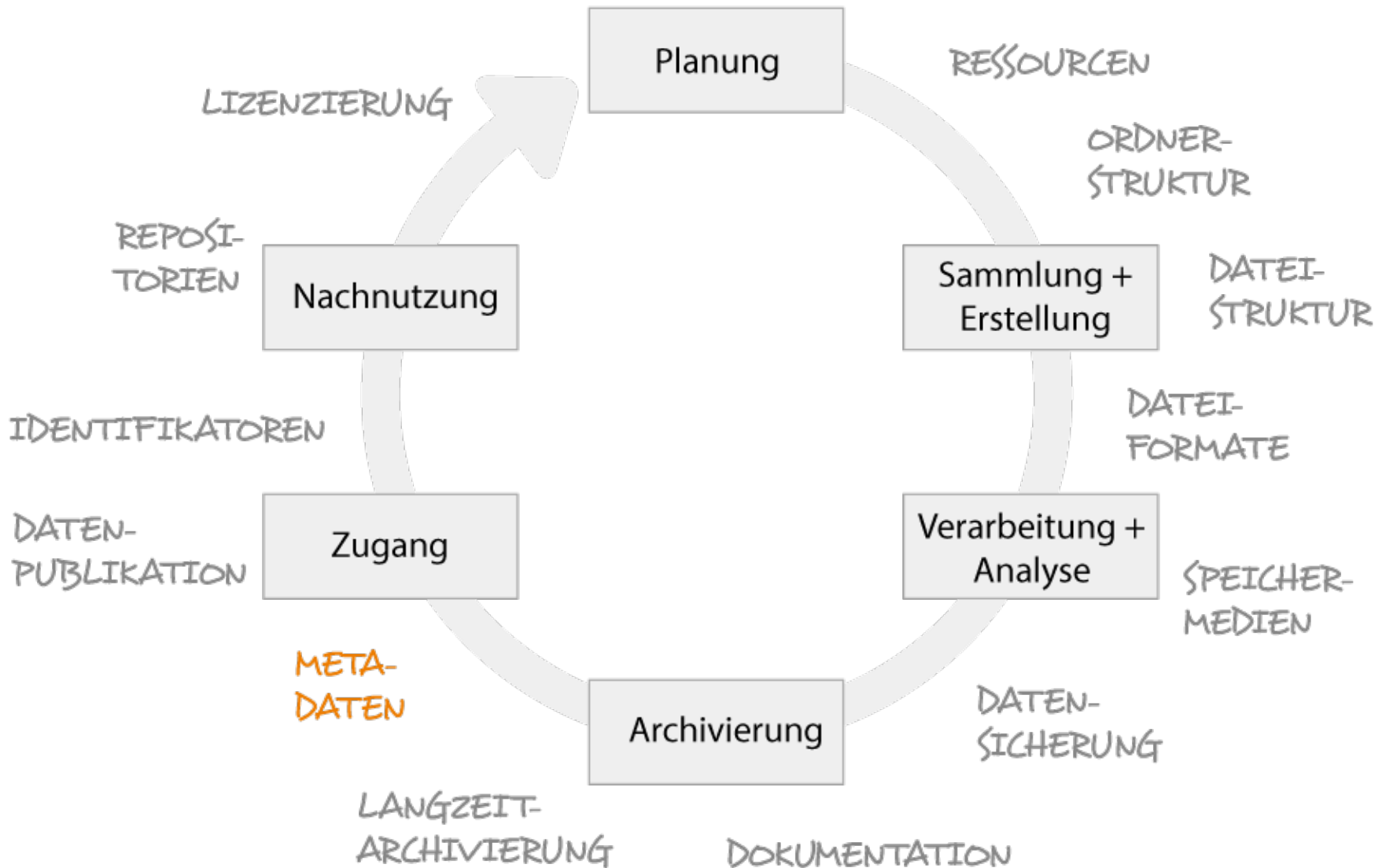
- sinnvoll: offene, freie, gut dokumentierte Formate
- besonders empfehlenswert: reine Textdateien (txt, xml, csv)

Welche Dateiformate nutzen Sie?

Wie gut sind diese Formate für eine Langzeitarchivierung geeignet?
In welche Formate könnten diese eventuell konvertiert werden?



DATENMANAGEMENTPLAN



Wissen, wonach man sucht

Metadaten = Daten über Daten

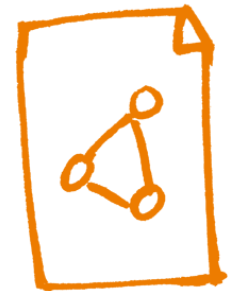
Das Etikett auf einer Dose sagt, was drin ist. Genauso sorgen Metadaten dafür, dass digitale Daten und Objekte gefunden und genutzt werden können. Deshalb ist es wichtig, auch die eigenen Forschungsdaten mit „Etiketten“ zu versehen.

Stil des Etiketts: **Metadatenstandard**

inhaltlich und strukturell gleichförmige Beschreibung von ähnlichen Daten

bibliographische bzw. administrative Daten

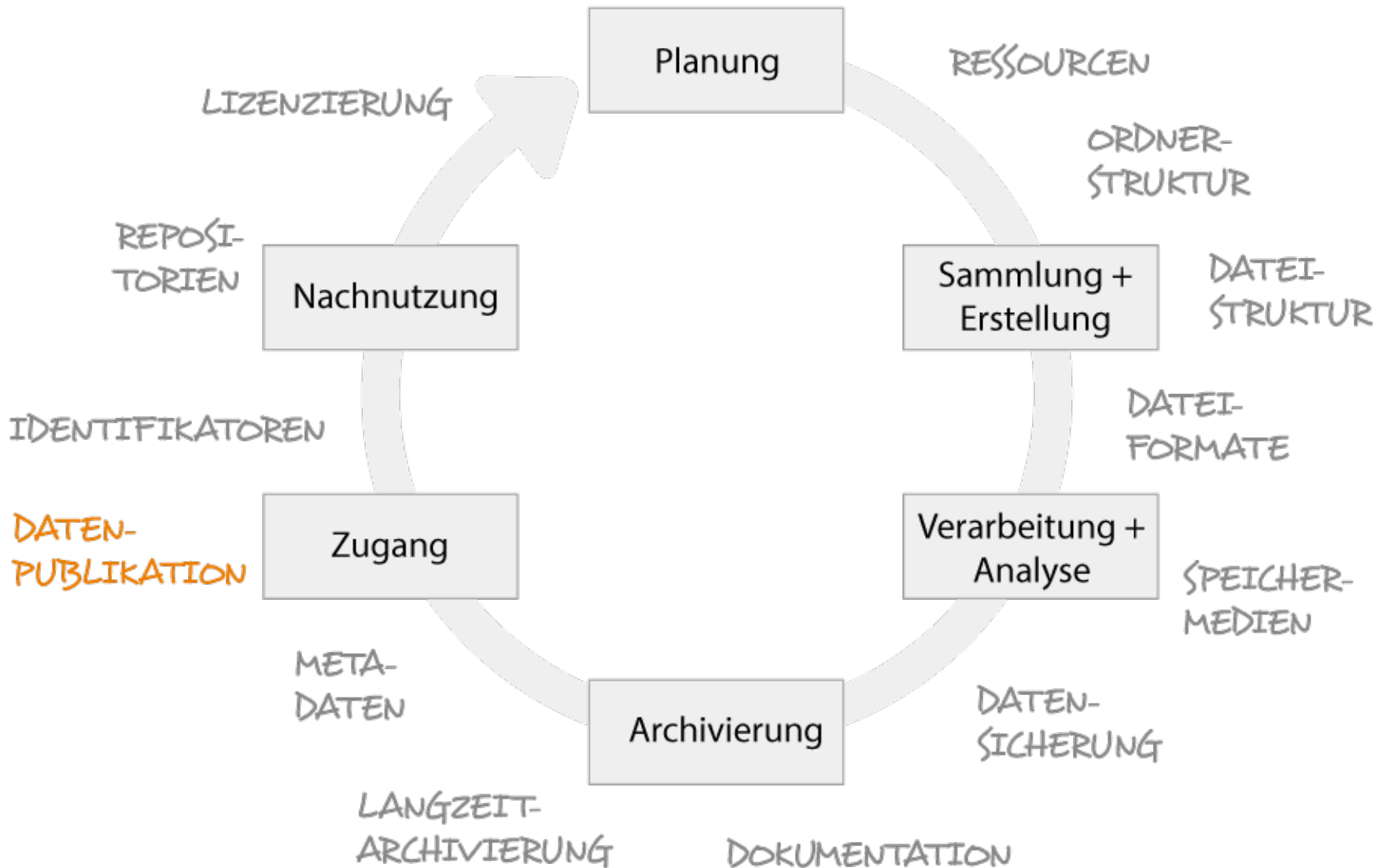
- Informationen zur Verwaltung und Entstehung der Daten
- allgemein, disziplinspezifisch
- **Beispiele:** Autor, Jahr, Titel, Format



inhaltsbeschreibende bzw. fachliche Daten

- beschreiben einzelne Aspekte oder Datensätze genauer
- bieten zusätzliche Informationen
- disziplinspezifisch
- **Beispiele:** Messmethode, Probe, Gerätetyp, Interviewperson, Koordinaten, Standort

DATENMANAGEMENTPLAN



„Sharing is Caring“

Teilen lohnt sich: denn Daten sind nicht nur für die eigene Forschung wertvoll. Wer seine Forschungsdaten mit anderen teilt, unterstützt die Forschungscommunity.

Von einem offenen Umgang mit Daten profitieren nicht nur „die Anderen“:

Vorteile

- die eigene Forschung bekommt mehr Aufmerksamkeit
- Forschende zitieren die eigenen Datensätze
- untermauert die eigenen Forschungsergebnisse
- Ergebnisse lassen sich einfacher vergleichen
- unterstützt interdisziplinäre Forschung
- ermöglicht Metaanalysen

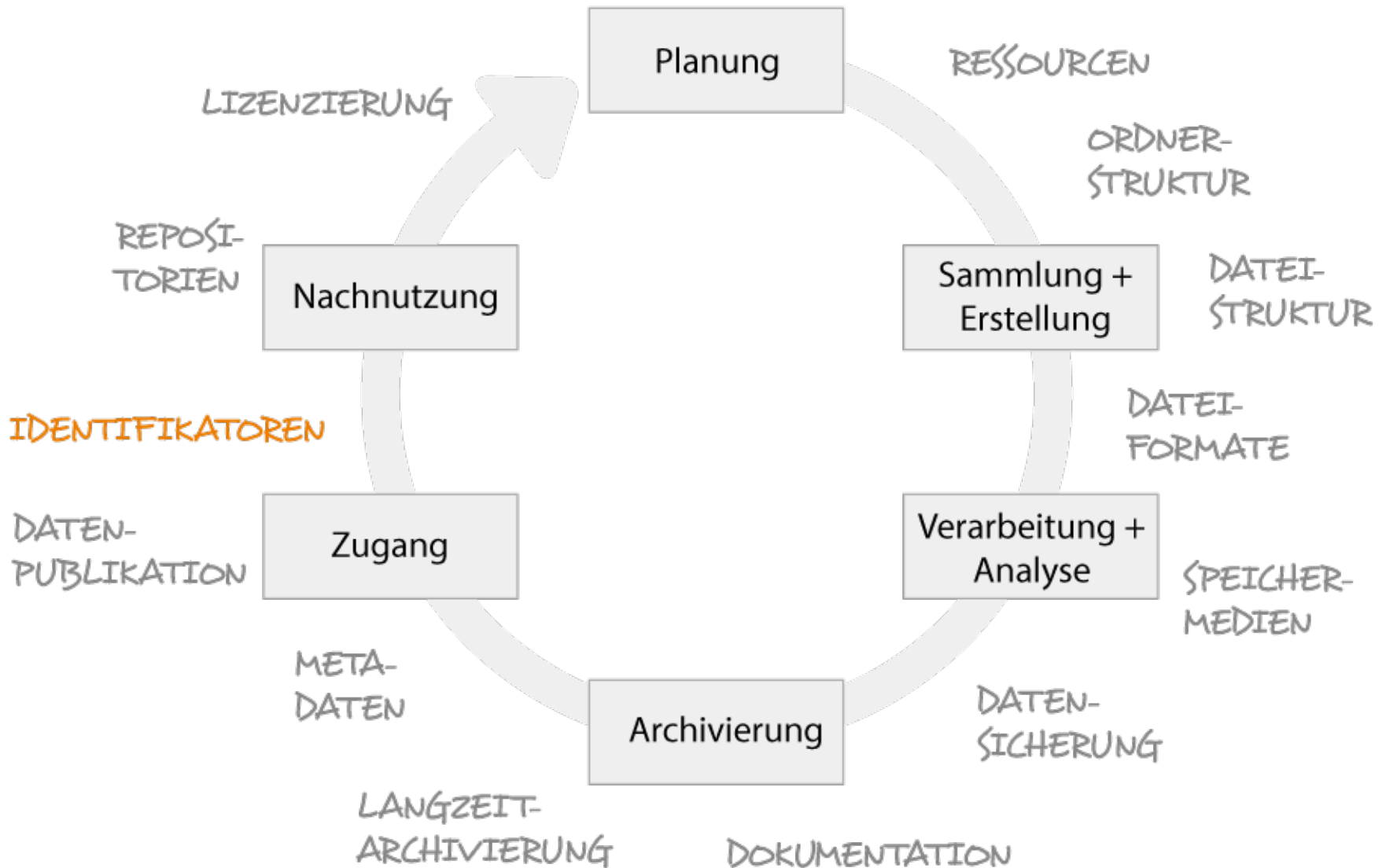


Möglichkeiten zur Publikation

- disziplinspezifische Repositorien
- disziplinübergreifende/generische Repositorien
z. B. Zenodo, Radar
- institutionelle Repositorien
→ OpARA
- Data Journals
(Fokus auf Beschreibung und Methodik der Datengewinnung)
- Fachzeitschriften
(ergänzend zu Textpublikationen)



DATENMANAGEMENTPLAN



Daten und Autoren wiederfinden – und gefunden werden!

Publikationen und Daten – Persistent Identifier (PID)

- eindeutige Bezeichnung
- garantiert dauerhafte Erreichbarkeit
- Beispiel: DOI (digital object identifier)

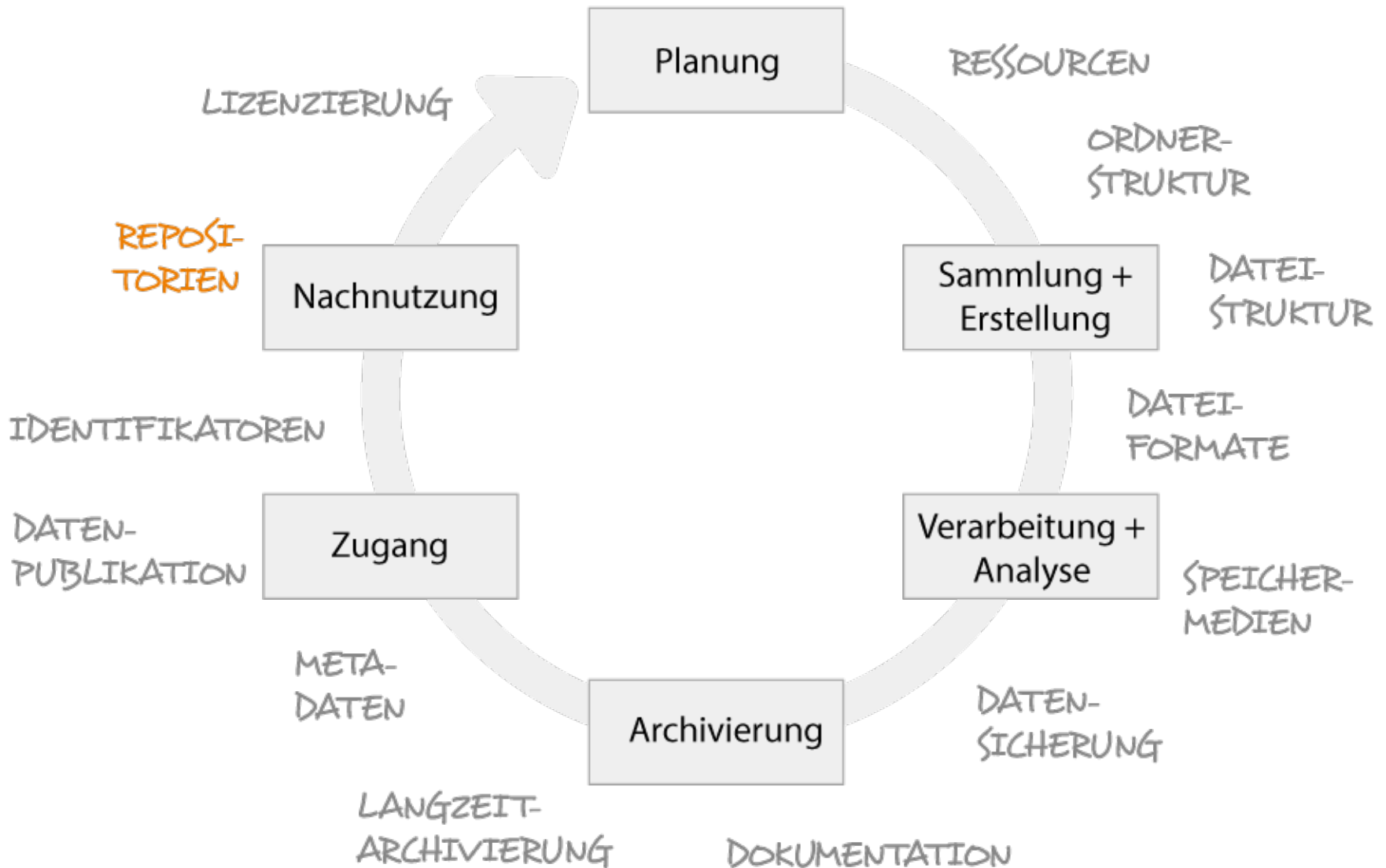
Personen – Open Researcher and Contributor ID (ORCID ID)

Mehrfach vorkommende Namen, unterschiedliche Schreibweisen oder eine Namensänderung können dazu führen, dass Autoren nicht eindeutig zuzuordnen sind.

- dauerhafte digitale Kennung
- eindeutig zuordbar
- lässt sich in wichtige Forschungsabläufe wie die Einreichung von Manuskripten und Anträgen integrieren
- <https://orcid.org/>



DATENMANAGEMENTPLAN



In online abrufbaren Daten-Repositorien können Forschungsdaten sicher für einen längeren Zeitraum abgelegt und gefunden werden.

- bessere Auffindbarkeit und Nachnutzung
- Absicherung der langfristigen Verfügbarkeit

Auswahl eines Repositoriums

1. disziplinspezifisches Repository
2. disziplinübergreifendes/generisches Repository
3. institutionelles Repository (OpARA)

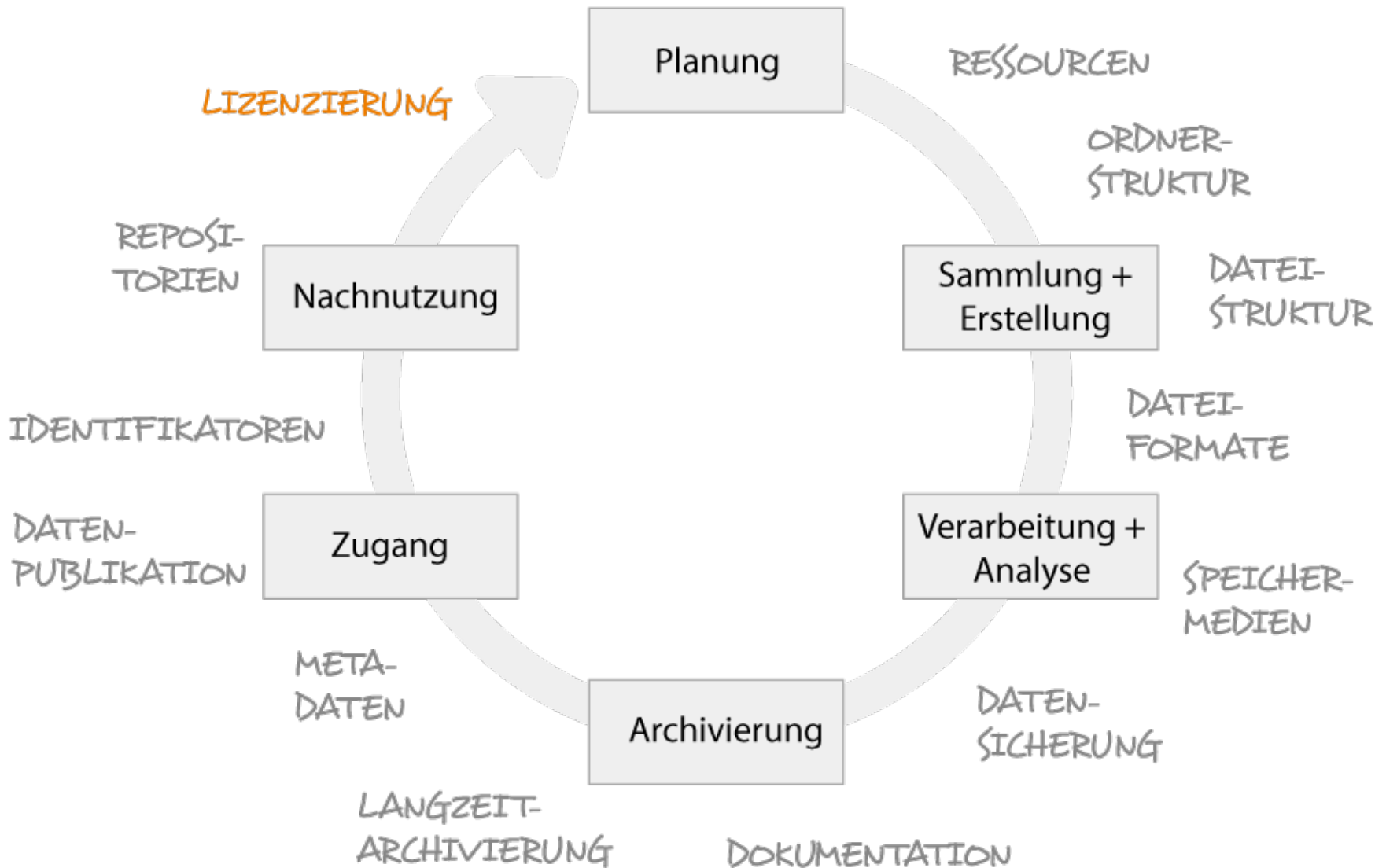
Repositorien Recommender der HTW: www.htw-dresden.de/~fdm

Registry of Research Data Repositories: <https://www.re3data.org/>

Ungeeignet zur Veröffentlichung von Forschungsdaten:

- eigene Website
 - soziale Forscher-Netzwerke (wie z.B. ResearchGate)
- Links setzen

DATENMANAGEMENTPLAN



Es ist gut, wenn Texte einfach gelesen werden können (und wissenschaftliche Daten eingesehen werden können). Doch was kann man damit machen – jenseits des Lesens und Zitierens? Schon das Weitergeben an andere Kolleginnen oder Kollegen bedarf der Erlaubnis der Urheberinnen oder Urheber.

Creative-Commons-Lizenzen

- modulare Struktur
- sehr bekannt
- für Fotos, Videos, Texte, ...
- <https://creativecommons.org/>



Open Data Commons

- Lizenzen speziell für Datenbanken
- Inhalte einer Datenbank können gesondert lizenziert werden
- <https://opendatacommons.org/>

Untersuchen Sie gemeinsam die verschiedenen Lizenzmodule.
Welches Icon gehört zu welchem Text?

Welche Kombinationen sind aus diesen Lizenmodulen möglich?





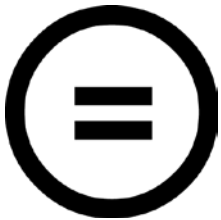
BY – Namensnennung

Es müssen angemessene Urheber- und Rechteangaben gemacht, ein Link zur Lizenz beigefügt und angegeben werden, ob Änderungen vorgenommen wurden.



SA – Weitergabe unter gleichen Bedingungen

Wenn das Material geremixt, verändert oder anderweitig direkt darauf aufgebaut wird, dürfen die Beiträge nur unter derselben Lizenz wie das Original verbreitet werden.



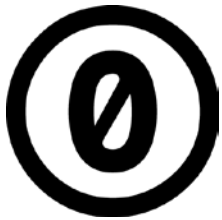
ND – keine Bearbeitungen

Wenn das Material geremixt, verändert oder darauf anderweitig direkt aufgebaut wird, darf die bearbeitete Fassung des Materials nicht verbreitet werden.



NC – nicht kommerziell

Das Material darf nicht für kommerzielle Zwecke verwendet werden.



ZERO – kein Urheberrechtsschutz

Weltweiter Verzicht auf alle urheberrechtlichen und verwandten Schutzrechte, soweit das gesetzlich möglich ist.

Das Werk darf kopiert, verändert, verbreitet und aufgeführt werden, sogar zu kommerziellen Zwecken, ohne um weitere Erlaubnis bitten zu müssen.

Sortieren Sie die sechs verschiedenen CC-Lizenzen nach dem Grad ihrer Offenheit bzw. Einschränkung.





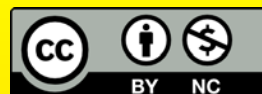
Kein Urheberrechtsschutz



Namensnennung



Namensnennung - Weitergabe unter gleichen Bedingungen



Namensnennung - Nicht kommerziell



Namensnennung - Nicht kommerziell - Weitergabe unter gleichen Bedingungen



Namensnennung - Keine Bearbeitung



Namensnennung - Nicht kommerziell - Keine Bearbeitung

DATENMANAGEMENTPLAN



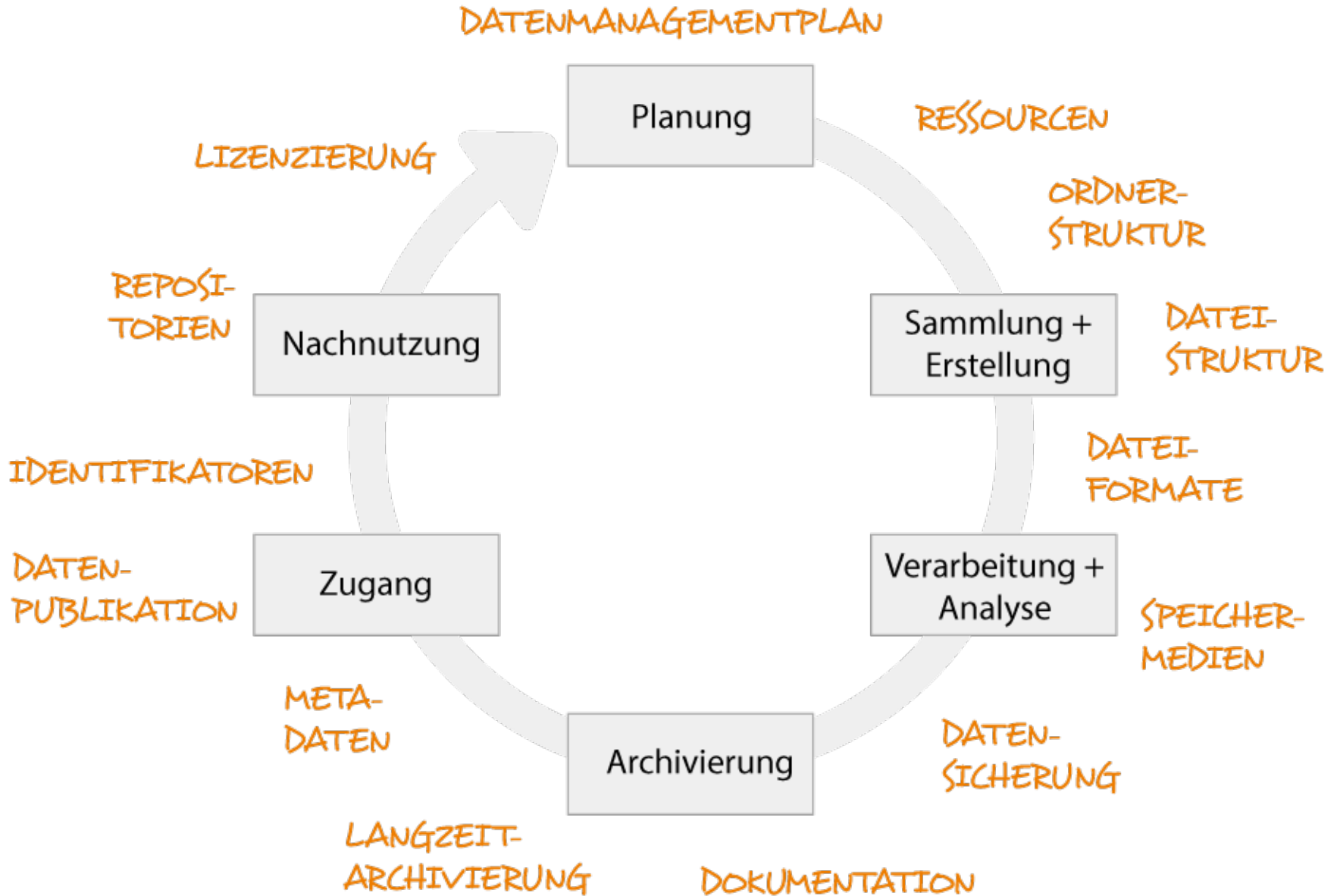
Welche Daten müssen neu erzeugt werden?
Können bestehende Daten nachgenutzt werden?

Möglichkeiten:

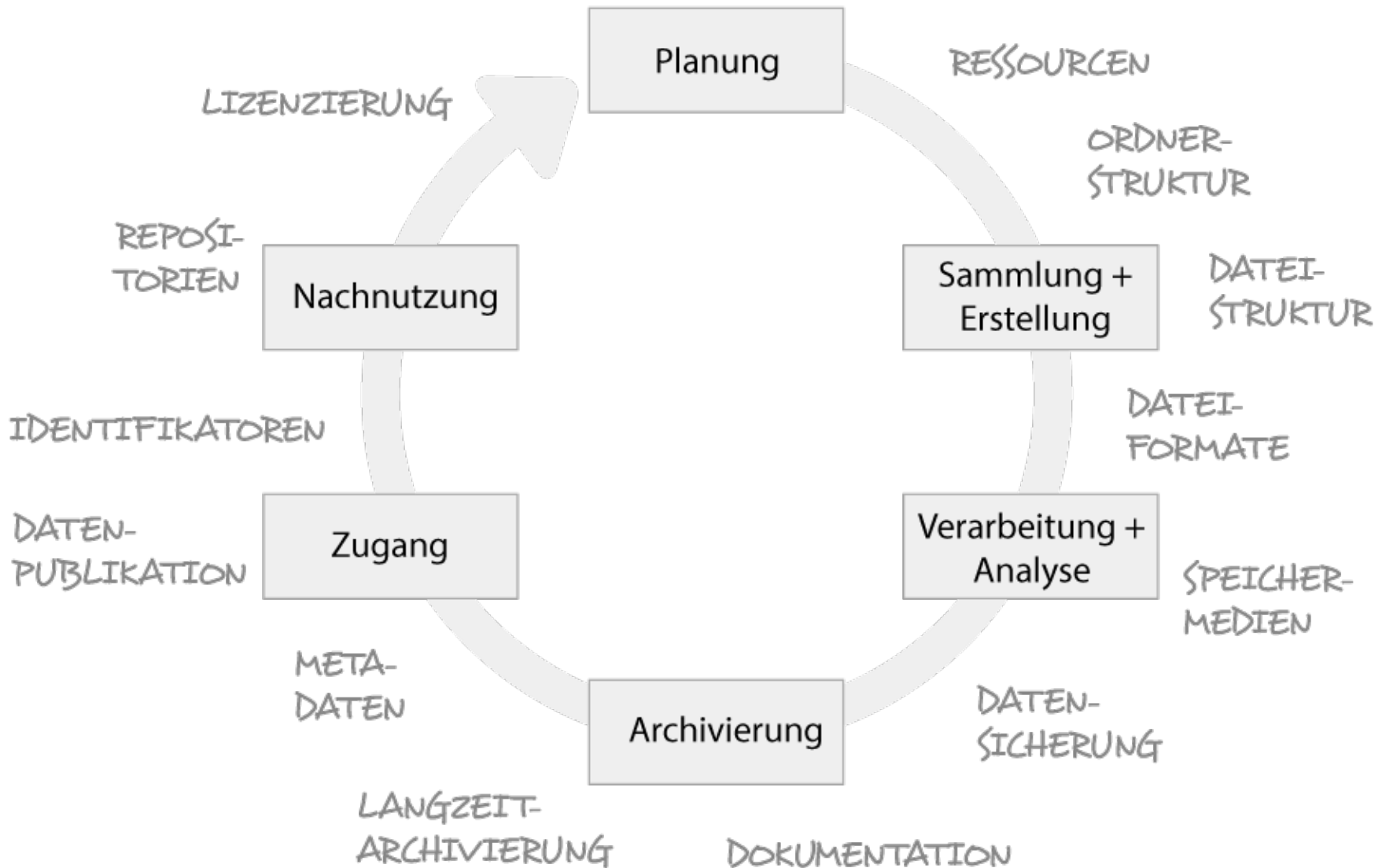
- Projekt- und fachspezifische Datenerhebung
- Nutzung bereits vorhandener Daten

Nutzung bereits vorhandener Forschungsdaten

- Recherche (Repositorien → RR, Projektpartner, ...)
- Kosten
- Einschätzung der Qualität (Dokumentation, Format, Methoden ...)
- Integration
- Zitation



DATENMANAGEMENTPLAN



Was ist ein Datenmanagementplan (DMP)?

Ein DMP strukturiert den Umgang mit den eigenen Forschungsdaten.

DMP dienen als Checklisten und zur laufenden Dokumentation: von der Erhebung bis zur langfristigen Speicherung bzw. Veröffentlichung der Daten.

Immer mehr Forschungsförderer, wie die EU oder das BMBF verlangen einen DMP.

Vorteile

- erleichtert Absprachen zum Datenaustausch bei Kooperationen
- erleichtert die Dokumentation bei Berichtspflichten
- erleichtert die Nachnutzung der eigenen Daten
- verringert die Gefahr von Datenverlust

Ein DMP enthält:

- Informationen zur Vorgehensweise im Umgang mit Forschungsdaten
- Dokumentation von Verlauf, Ergebnissen, verwendeten Werkzeugen, angewendeten Verfahren
- Zuständigkeiten

Umfang ca. 1-2 Seiten



Beispiele:

- [Muster-Datenmanagementpläne der HU Berlin](#)
- [internationale Beispiele für Datenmanagementpläne](#)

Bitte entfernen Sie Ihre beantworteten Fragen/Themen vom Themenparkplatz.

Notieren Sie weitere Fragen/Themen auf den bereitliegenden Kärtchen oder kleben Sie Klebepunkte auf bereits genannte Themen.



OpARA-Schulung

07.10.2019 14:00 bis 15:30 Uhr

Kontakt:

Fragen, Anmerkungen, Themenwünsche

fdm@htw-dresden.de



Allgemein

- <http://www.forschungsdaten.org>
- <https://www.forschungsdaten.info>

Online-Kurse:

- <http://www.researchdatamanagement.ch>
- <http://datalib.edina.ac.uk/mantra/>

FDM an der HTW Dresden:

- <http://www.htw-dresden.de/fdm>
- [Leitlinie zum Umgang mit Forschungsdaten an der HTW Dresden](#)
- [Open-Access-Erklärung der HTW Dresden](#)

Die Präsentation zum Workshop “Forschungsdatenmanagement” vom Projekt FoDaMa der HTW Dresden ist lizenziert unter einer [Creative Commons Namensnennung 4.0 International Lizenz](http://creativecommons.org/licenses/by/4.0/). (<http://creativecommons.org/licenses/by/4.0/>)

Ausgenommen davon sind die Porträtfotos und das Logo der HTW Dresden.