

# Datenformate und Datenflüsse im Digitalisierungsprozess, in der Präsentation und in der Langzeitarchivierung am Beispiel der ZBW

---

*Berlin, 19. März 2026*

*Dr. Monika Zarnitz*

---

# Um was geht es?

---

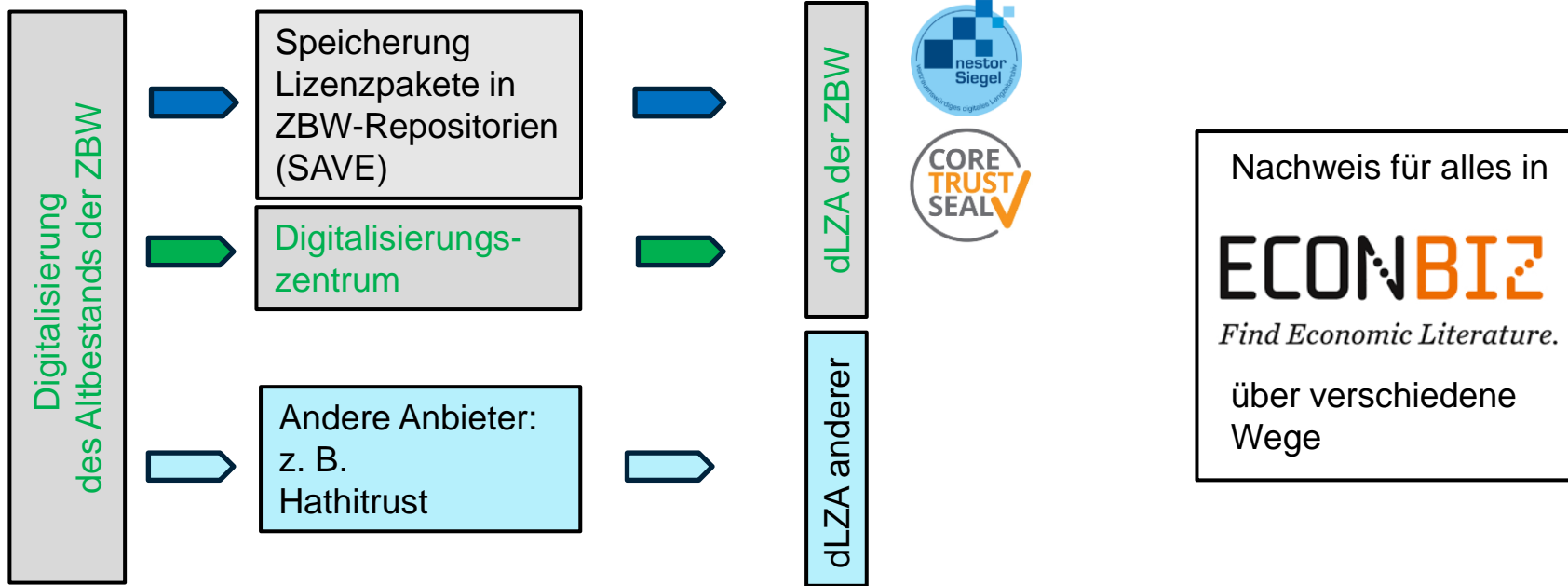
- Anlass
  - Speicher insbesondere für Bilddateien, die bei der Retrodigitalisierung entstehen, ist teuer
  - Plan der ZBW: Die umfangreichen Tiff-Dateien und weitere Daten werden im Produktionssystem gelöscht
  - Wie verfahren mit nachträglichen Korrekturen an abgeschlossenen Vorgängen?
- Sinn der Lösung
  - Sicherstellen, dass notwendige Änderungen möglich sind
  - Aufwand in Grenzen halten
- Ziel des Vortrags
  - Vorstellung von Datenzyklus und Datenhaltung / Lösungsideen der ZBW / umgesetzt ist das noch nicht

# Überblick

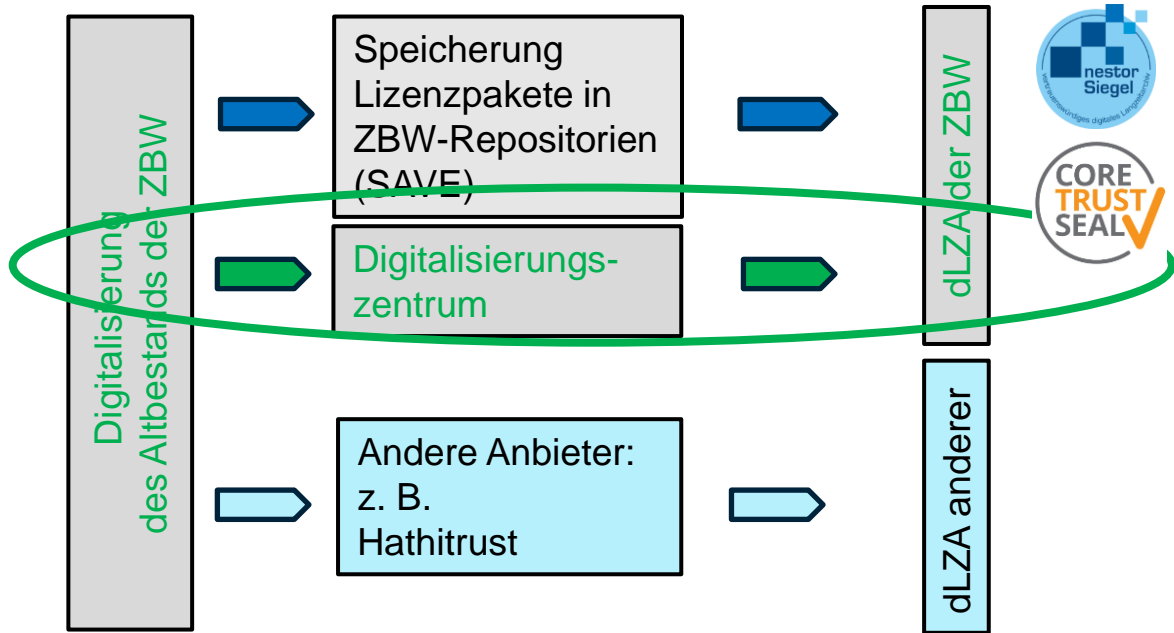
---

- Digitalisierungsstrategie der ZBW zum Erhalt der Inhalte des Altbestands
- Massendigitalisierung im eigenen Digitalisierungszentrum
- Datenfluss bei Retrodigitalisierung – Referenz für Fall „ohne Korrekturen“
- Speicher sparen – aber wie?
- Korrekturen I – V / Was überhaupt und wie? / Datenfluss und -haltung bei Korrekturen
- Fazit aus den Überlegungen der ZBW

# Digitalisierungsstrategie der ZBW I

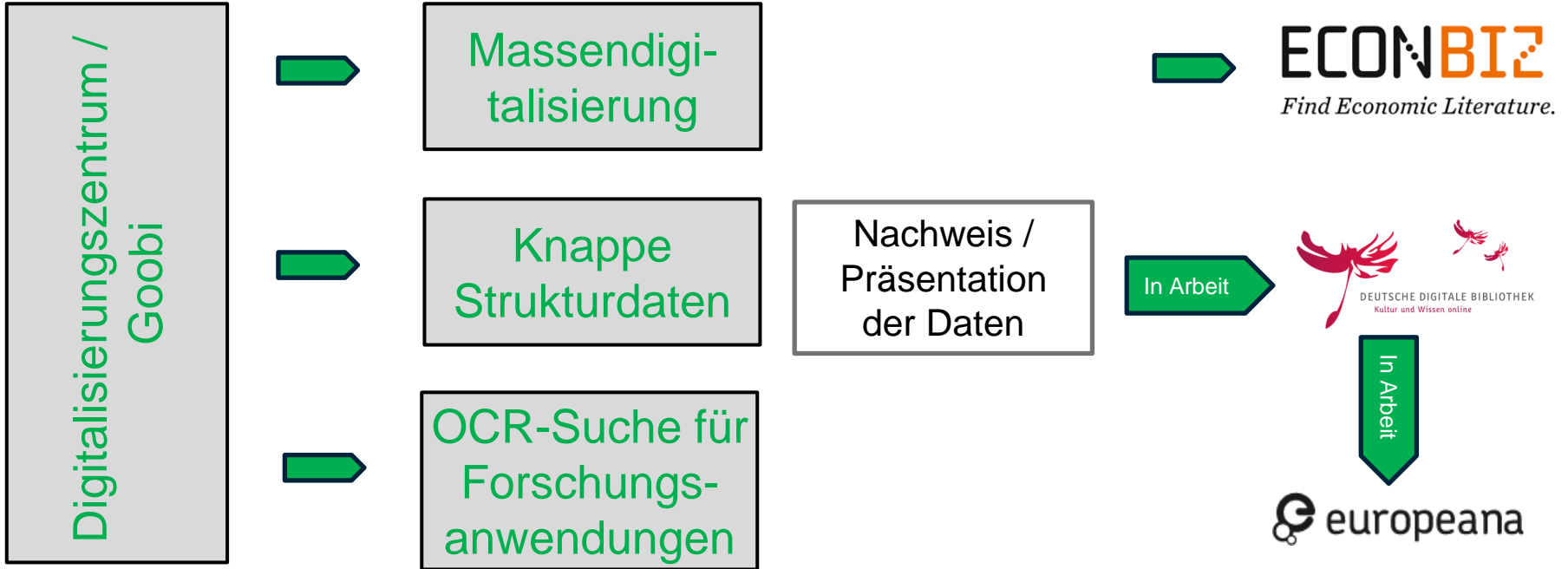


# Digitalisierungsstrategie der ZBW II

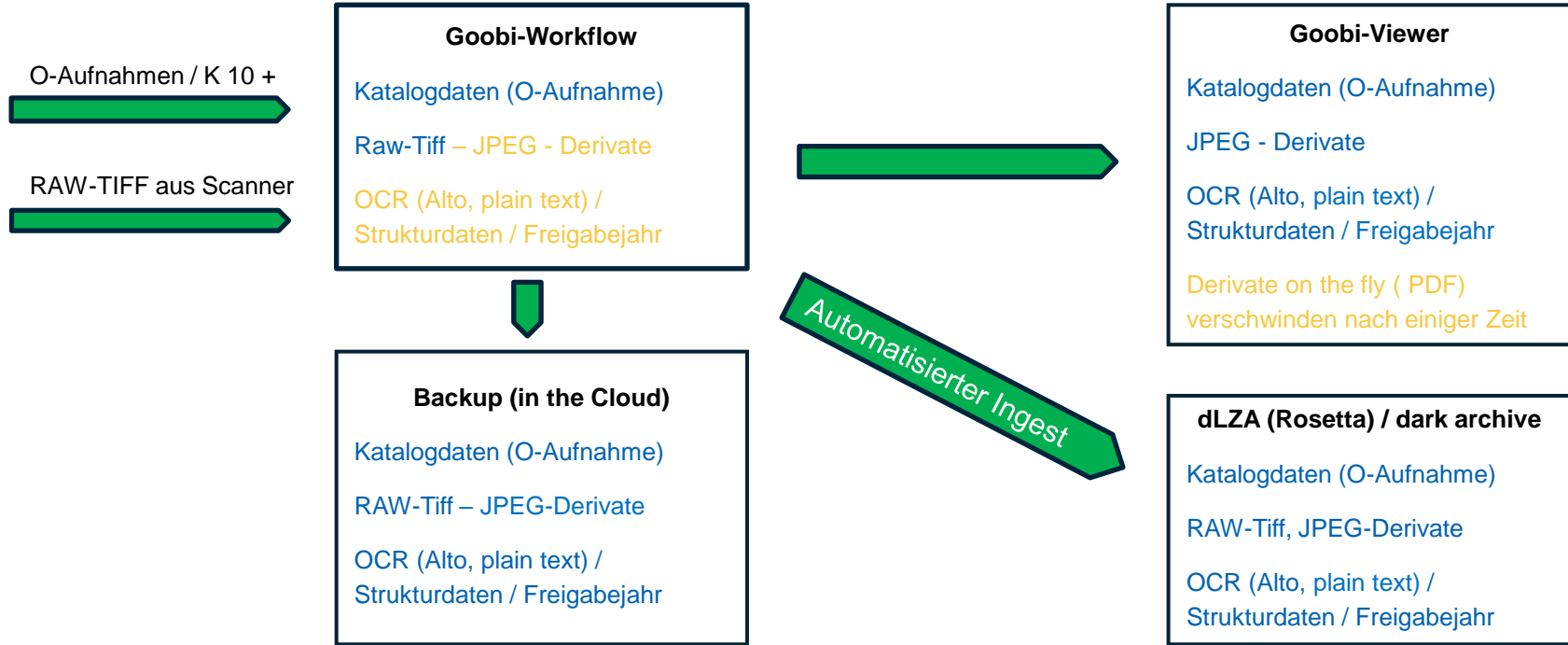


Nachweis für alles in  
**ECONBIZ**  
*Find Economic Literature.*  
über verschiedene  
Wege

# Digitalisierungszentrum



# Datenfluss bei Retrodigitalisierung – Referenz



# Exkurs I: Begriff „Vorgang“

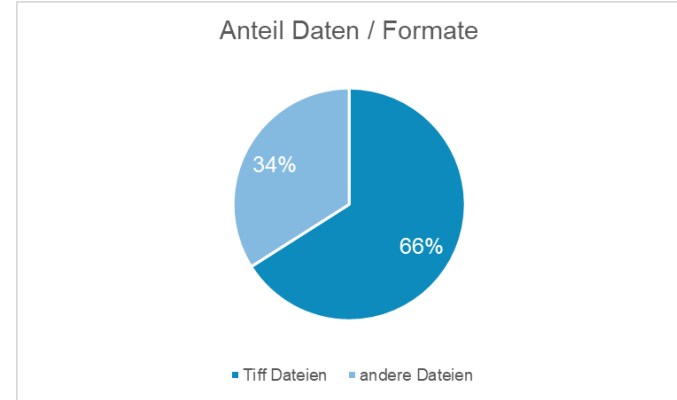
---

- Der Vorgang wird zu Beginn des Prozesses in Goobi Workflow angelegt
- Katalogdaten aus K10+ über die ppn importieren, Angaben z. B. zur Schriftart, zum Freigabejahr usw. ergänzen
- Am Vorgang „hängen“ die Bilddateien, die OCR-Ergebnisse und die Metadaten
- Den Vorgang kann man als Ganzes löschen oder einzelne Daten daraus

# Speicher sparen – aber wie korrigieren?

---

- Vorgang erhalten (zur Identifikation und wegen der daran „hängenden“ Metadaten)
- Idee: Tiffs, Alto, plain text sowie Tiffs mit JPEG-Derivaten aus Goobi Workflow löschen, wenn bestimmte Bearbeitungsstadien erreicht sind (Ingest in dLZA und Überführung in das Back-up)
- Was aber, wenn Korrekturen nötig werden? – „Modell“ ist in Arbeit



# Datenhaltung bei Retrodigitalisierung – nach Löschungen in Goobi-Workflow

## Goobi-Workflow

Katalogdaten (O-Aufnahme)

Strukturdaten / Freigabedatum

## Backup (in the Cloud)

Katalogdaten (O-Aufnahme)

RAW-Tiff – JPEG Derivate

OCR (Alto, plain text) / Strukturdaten /  
Freigabedaten

## Goobi-Viewer

Katalogdaten (O-Aufnahme)

JPEG Derivate

OCR (ALTO, plain text) / Strukturdaten  
/ Freigabedaten

(Derivate on the fly ( PDF))

## dLZA (Rosetta) – dark archive

Katalogdaten (O-Aufnahme)

RAW-Tiff, JPEG Derivate

OCR (Alto, plain Text) / Strukturdaten

# Korrekturen | Art und Ursachen von Fehlern

---

- An Metadaten
  - Eigene Fehler oder neue Erkenntnisse, z. B. bei Kennzeichnung Freigabejahr
  - „Leben“ im K10+: Dublettenabgleich des Verbundes / andere Bibliotheken ändern usw.
- An den Bilddateien
  - Fehler beim Scannen, die zu spät entdeckt werden
  - Vollständige Version eines Werks taucht z. B. beim Dublettenabgleich auf usw.
  - Wichtig, da ggf. Ersatz für gedrucktes Werk
- Aufwand für Korrekturen in Goobi-Workflow bei abgeschlossenen Vorgängen ist hoch, da sie in Goobi Viewer, Back-up und dLZA nachzuziehen sind

# Korrekturen II Abmachungen dLZA / DZ

---

- Änderungen an Metadaten
  - Nur ppn – pica production number = Identifier im K10+
  - Nur 2. Stelle im Feld 0500 = „Bibliografische Gattung und Status“ davon  
2. Stelle = „Bibliografische Erscheinungsweise“
  - Freigabejahr – „Schalter für die öffentliche Zugänglichkeit“
- Begründung Änderungen Metadaten
  - Da Identifikation Katalogdaten im K10+ nach Herstellung aus dLZA nicht möglich wäre
  - Um Objekttyp und Relationen analog zum Katalog korrekt abzubilden, z. B. Teile MTM
  - Neue Informationen zu Schutzfristen (nachträgliche Freigabe oder Sperrung)
- Fehler in den Bilddateien
  - Nur, wenn z. B. vollständigere Version eines Buchs auftaucht und relevant ist
  - Korrekturbedarf nach Begutachtung der Relevanz des Fehlers festlegen

# Korrekturen III Häufigkeiten

---

	Seit 2019
Datenproduzent zieht Rechte zurück	1204
PPN ändern	366
PDF ausgetauscht	576
Datei hinzugefügt	37

- Änderungen an Daten aus EconStor – fachliches Repository
- Trifft das auf Retrodigitalisate auch zu? – Die Erfahrung wird es zeigen...

# Exkurs II Rolle der ppn als Identifier

---

- Die PPN (Pica Production Number) der Katalogdaten ist für die diversen Systeme der Identifier oder ein Teil davon
  - Viewer = ppn Identifier / kann nicht geändert werden: löschen + neu importieren
  - ppn ist Teil des Titels für den Scanjob und den Goobi-Vorgang (tsl\_ppn)
  - ppn ist Teil der Metadaten des tiff – Headers (kann nur der Admin händisch ändern)
  - Technisch nicht notwendig aber Frage der Datenorganisation

# Korrekturen III ppn

O-Aufnahmen / K 10 +

## Goobi-Workflow

Katalogdaten (O-Aufnahme) /  
Strukturdaten / Freigabedaten

Ggf. Löschen alter  
Vorgang / Import  
neuer Vorgang

## Backup (in the Cloud)

Katalogdaten (O-Aufnahme)  
RAW-TIFF – JPEG-Derivate  
OCR (Alto, plain text) /  
Strukturdaten / Freigabedaten

Löschen alter Vorgang  
/ Import neuer  
Vorgang

Löschen alter Vorgang /  
Import neuer Vorgang

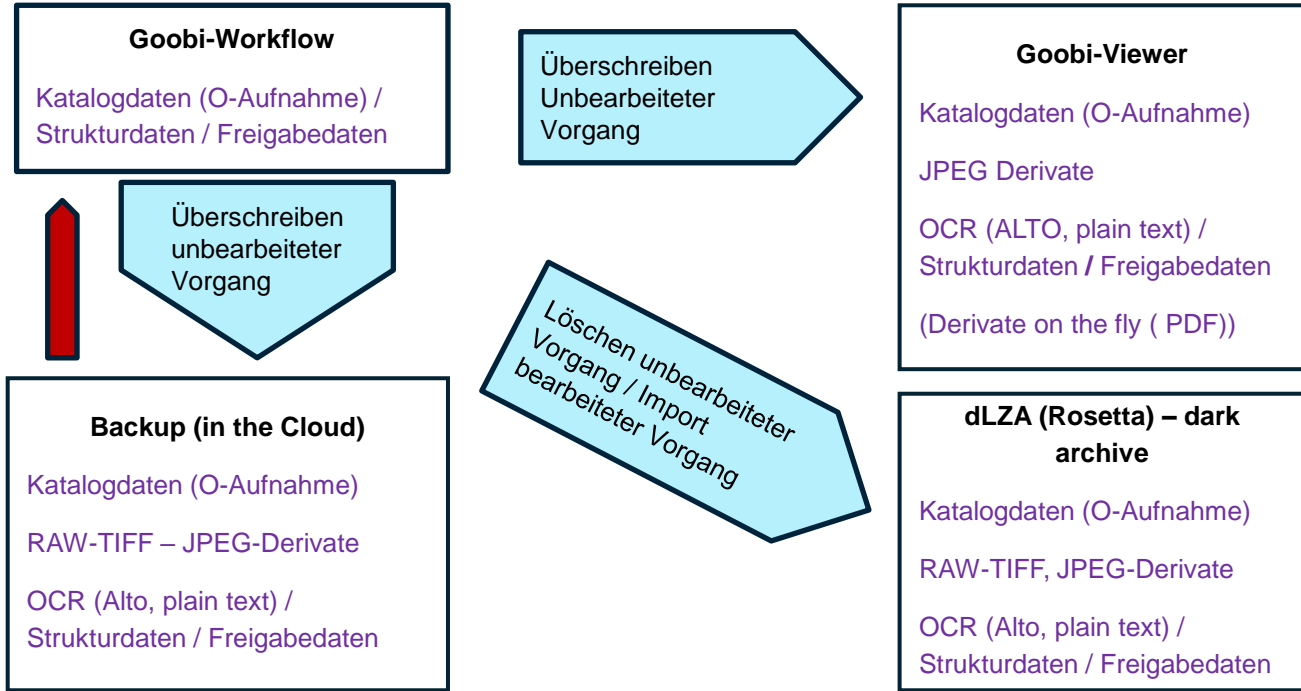
## Goobi-Viewer

Katalogdaten (O-Aufnahme)  
JPEG Derivate  
OCR (ALTO, plain text) /  
Strukturdaten / Freigabedaten  
(Derivate on the fly ( PDF))

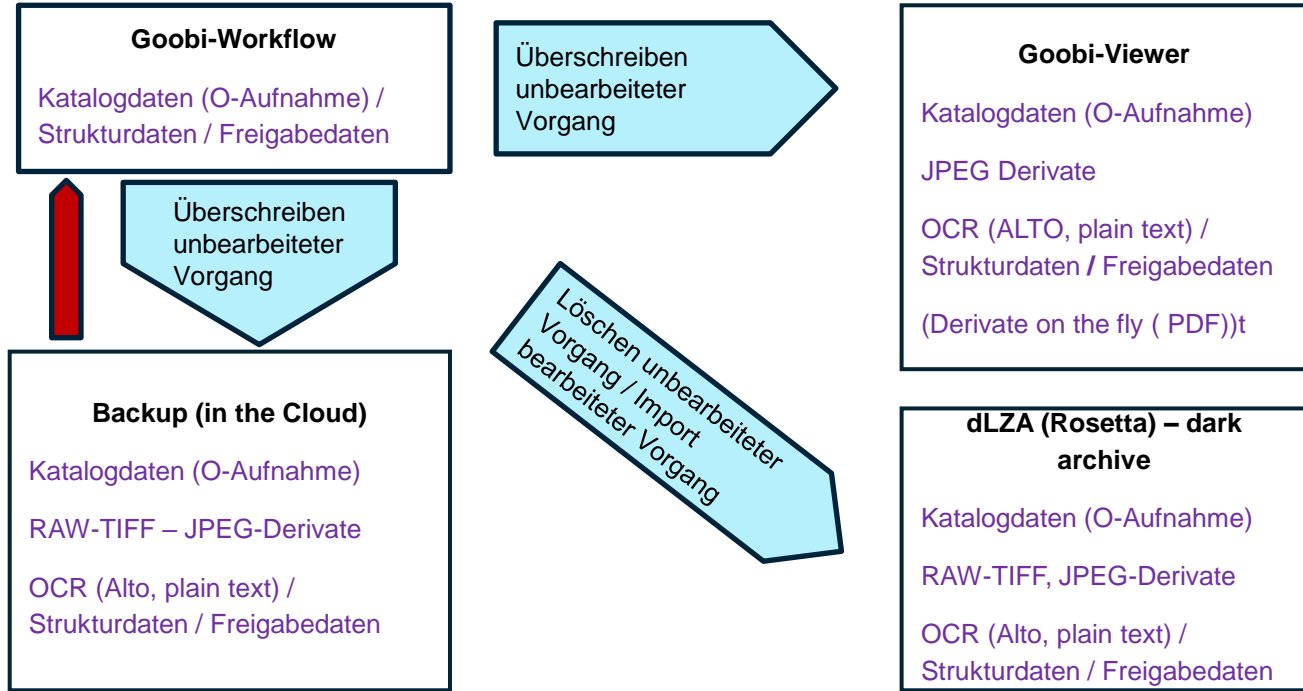
## dLZA (Rosetta) – dark archive

Katalogdaten (O-Aufnahme)  
RAW-TIFF, JPEG-Derivate  
OCR (Alto, plain text) /  
Strukturdaten / Freigabedaten

# Korrekturen IV 0500 / Freigabejahr



# Korrekturen V Bilddateien



# Fazit für den Datenfluss in der ZBW

---

- Änderungen an Daten müssen in diversen Systemen vorgenommen werden
- Daten (Metadaten und Bilddateien sowie Ergebnisse der OCR) sind nicht statisch – sie „leben“ – auch wenn der Vorgang abgeschlossen ist
- Es gibt einen Trade-Off zwischen Aufwand und Nutzen von Änderungen, den man pragmatisch gestalten sollte
- Es müssen Kompromisse für den Umgang mit Änderungen gefunden werden
- Klare Absprachen und Regeln helfen, den Kompromiss durchzusetzen

---

**Vielen Dank für Ihr Interesse!**

**Gern nehme ich Fragen entgegen...**

Dr. Monika Zarnitz  
Tel.: 0431 8814 431  
m.zarnitz@zbw.eu