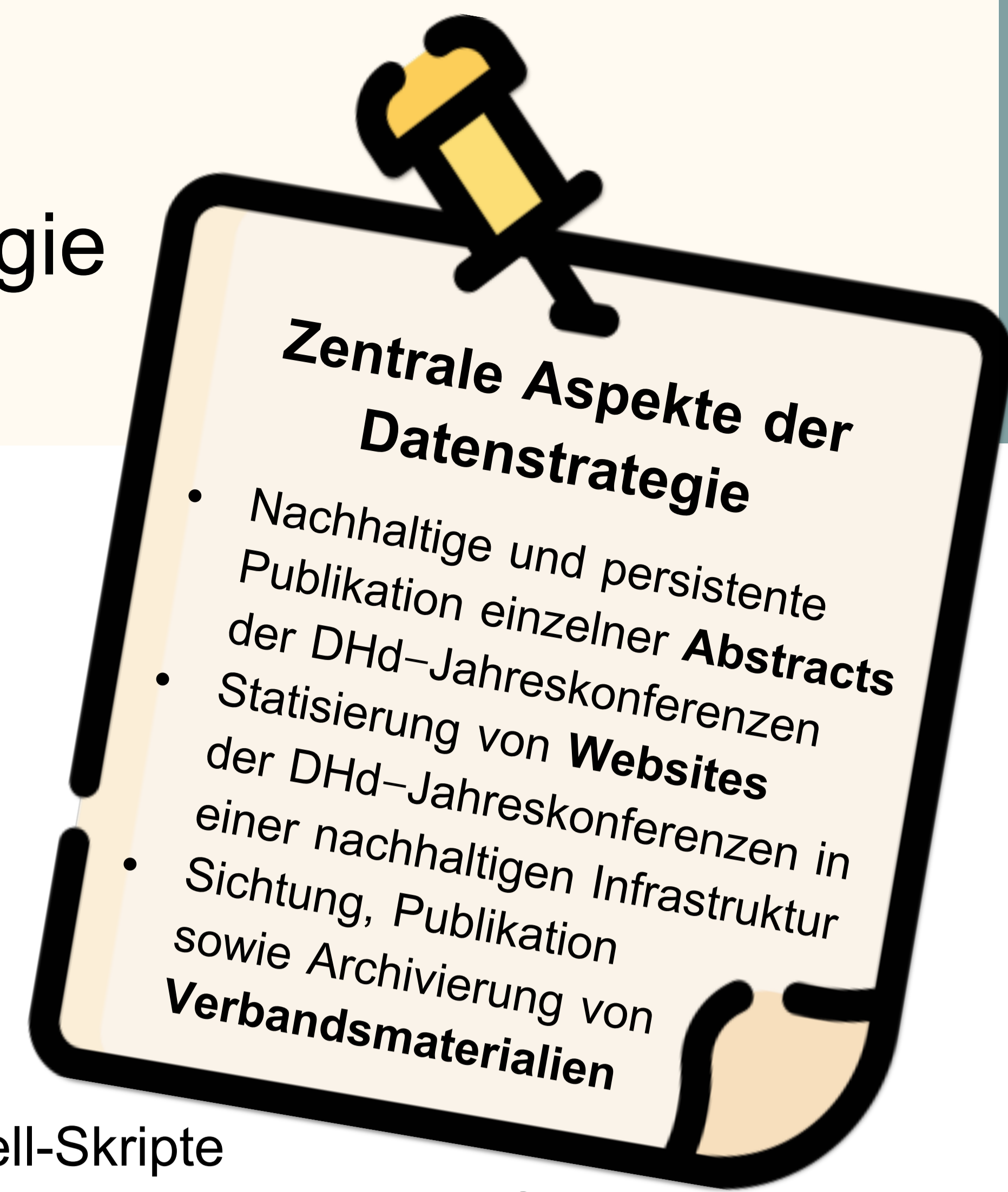


# Der DHd Data Steward

Maßnahmen zur Entwicklung einer nachhaltigen Datenstrategie für die Digital Humanities im deutschsprachigen Raum



## Ausgangspunkt und Zielsetzung des Data Stewards

- **Ernennung:** durch den DHd-Vorstand als kooptiertes Vorstandsmitglied
- **Aktiv seit:** der DHd-Jahreskonferenz 2020 in Paderborn
- **Ziel:** Entwicklung einer umfassenden Datenstrategie für den DHd-Verband
- **Zentrale Aufgaben:** Entwicklung von Workflows und Prozessen zum FAIRen Umgang mit verschiedenen Materialien aus dem Kontext des DHd-Verbands

## 1. Nachhaltige Publikation aller einzelnen DHd-Abstracts

- Vorverarbeitungs- und Publikationsworkflows → XSLT, Python- und Shell-Skripte
- Persistente Zitierbarkeit aller individuellen DHd-Abstracts → Zenodo, DataCite-Metadatenchema und DOIs
- Erfassung von DHd-Abstracts in Onlinekatalogen → Indexierung durch das Digital Bibliography & Library Project
- Verbesserung der Nachnutzbarkeit von DHd-Abstracts → Publikation aller Daten über GitHub-Repositorien

### 1.1 Vorverarbeitungs- und Publikationsworkflows

#### a) Vorverarbeitung und Metadaten

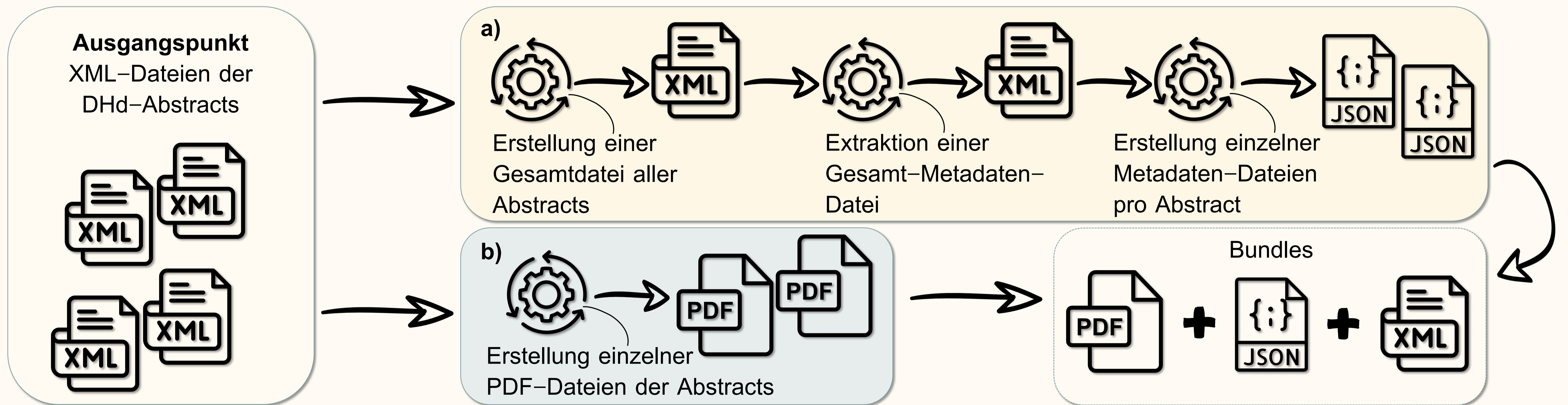
Alle XML-Dateien einer Jahreskonferenz werden zu einer Gesamtdatei zusammengeführt (XSLT), aus der eine Gesamt-Metadaten-Datei aller Abstracts eines Jahrgangs generiert wird (XSLT). Hierbei wird sich an dem DataCite-Metadatenchema orientiert. Für jedes einzelne Abstract werden die Metadateneinträge schließlich in einzelne JSON-Dateien transferiert (Python).

#### b) Erstellung von PDF-Dateien

Aus den XML-Dateien einer Jahreskonferenz werden einzelne PDF-Dateien generiert. Hierbei wird auf schon bestehende Transformationskripte zur Erstellung von Book of Abstracts aus XML-Dateien zurückgegriffen (Shell). Diese Skripte wurden entsprechend strukturell angepasst, für das Layout wurde auf Voreinstellungen zum Book of Abstracts 2018 zurückgegriffen.

#### c) Publikationsprozess

Es werden automatisiert Bundles erstellt, die jeweils aus der PDF-Datei, der XML-Datei und der Metadaten-Datei eines Abstracts bestehen. Die Bundles werden über die API-Schnittstelle von Zenodo skriptbasiert als Drafts hochgeladen. Nach einer manuellen Überprüfung und ggf. Korrektur der Drafts werden diese automatisiert in der DHd-Community auf Zenodo mit DOIs publiziert (Python).



## 2. Statisierung der DHd-Konferenzwebsites

- **Zentrale Technologie-Stacks:** Wordpress, JavaSkript, Typo3, Joomla etc.
- **Strategie:** Erstellung statischer HTML-Versionen der Websites und Überführung in die technische Infrastruktur der Alliance of Digital Humanities Organizations (ADHO)

## 3. Management von Verbandsmaterialien

- **Sichtung und Kategorisierung** von öffentlichen und privaten Materialbeständen
- Einrichtung einer entsprechenden **Zenodo-Community** → DHd-Verbandsaktivitäten
- **Archivierung** und ggf. Publikation des Materials



### Bisherige Achievements

- ✓ Publikation von 917 DHd-Abstracts (ink. Datenveröffentlichung auf GitHub)
- ✓ Publikation von 61 Abstracts der vDHd Konferenz „Experimente“
- ✓ Einrichtung der Zenodo-Community „Verbandsaktivitäten“
- ✓ Statisierung der Website zur DHd-Jahreskonferenz 2020
- ✓ Bildung einer community-getragenen TaskForce

