

# **AREDO – Das Langzeitarchivierungsangebot der Deutschen Nationalbibliothek**

Erfahrungen zwischen Konzeption und Praxis

Karlheinz Schmitt

05.06.2014

# Übersicht

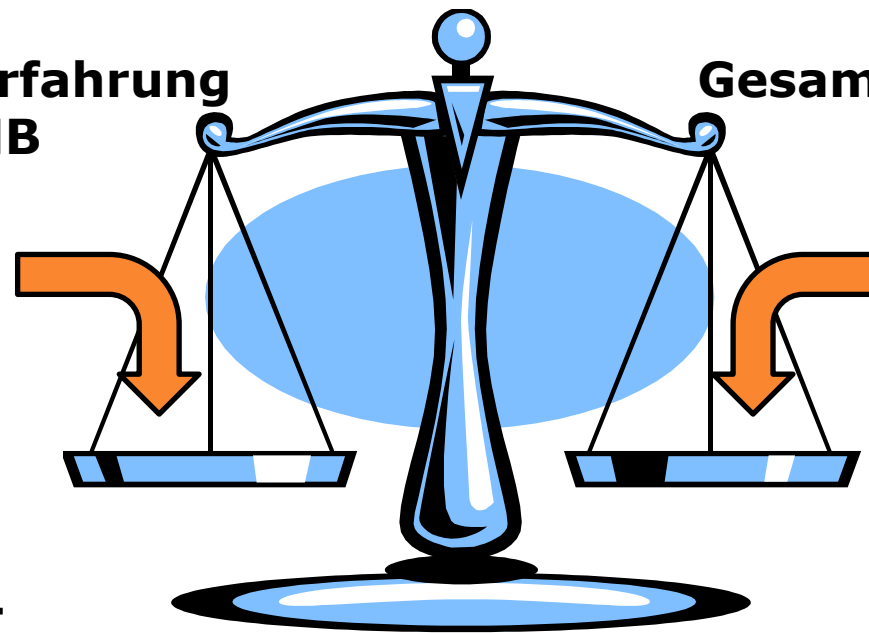
- **AREDO** –  
**AR**chivierung und dauerhafte **E**rhaltung **D**igitaler **O**bjekte
  
- Erfahrungen zwischen Konzeption und Praxis
  - Schnittstellen
  
  - Risiken

## **Notwendigkeiten**

- **Erfahrungsaustausch**
- **Zusammenarbeit**
- **Grenzen überwinden**

## Kooperative Langzeitarchivierung

**Gesammelte Erfahrung  
der DNB**



**Gesammelte Erfahrungen  
der Partner**

**in Form einer  
„LZA-Angebotsleistung“**

**als „Gegenleistung“**

## Das Leistungs-Katalog

### Basis-Angebot:

Integritätsprüfung aller digitalen Objekte

Automatische Generierung technischer Metadaten

Qualitätsprüfung (technisch) aller digitaler Objekte

Bereitstellung von Kontrollmöglichkeiten

Preservation Planning: Risikomanagement

Bereitstellung von Massenimport- und -exportschnittstellen

**Rechenzentrum-Standard**

## Automatisierte Massenschnittstellen

*Ziel: Flexibilität in den Schnittstellen*

### ● Anlieferung:

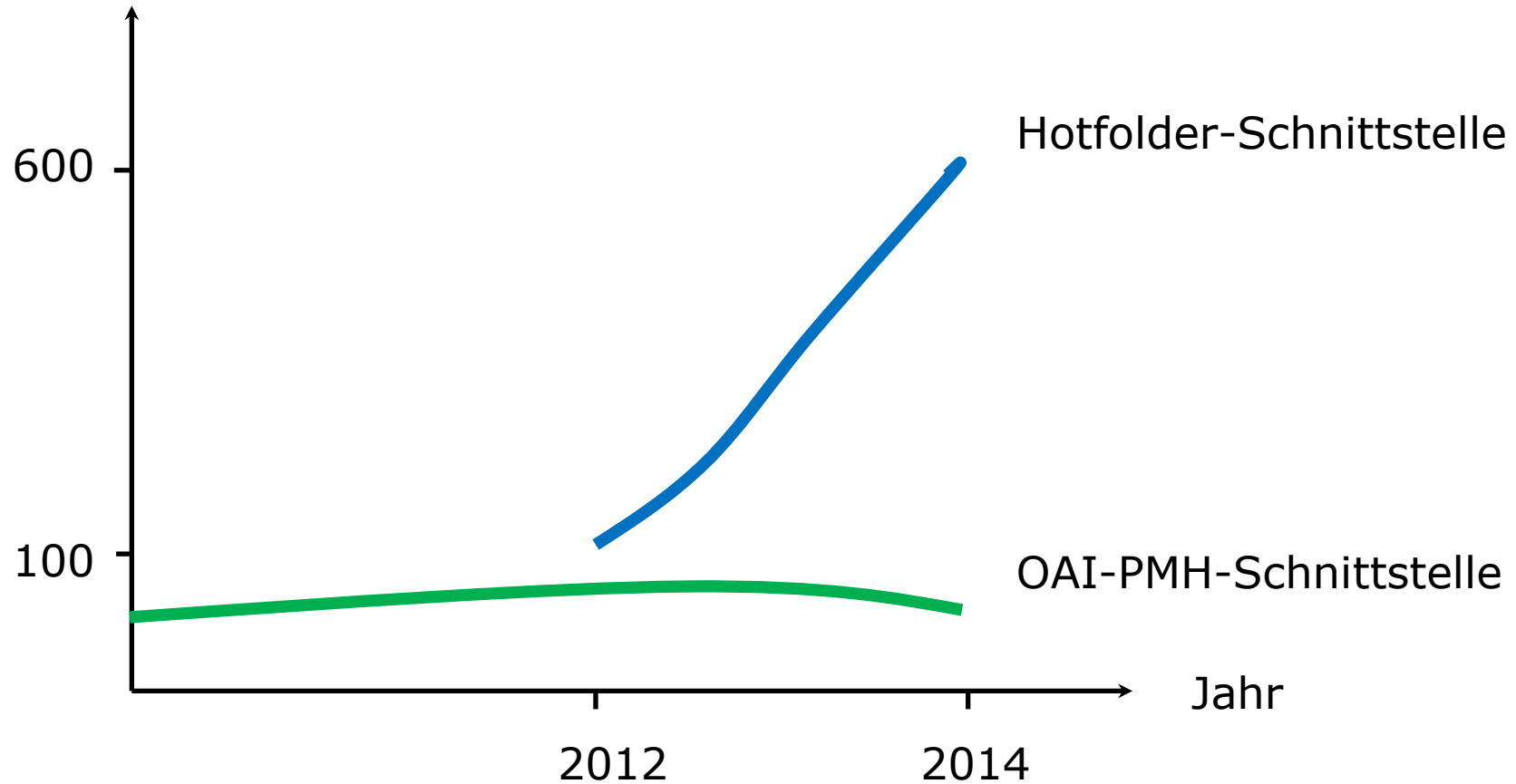
- Hotfolder
- OAI-PMH-Protokoll

### ● Auslieferung:

- SOAP-Schnittstelle
- REST-Schnittstelle
- WEB-GUI (Benutzerschnittstelle)

Anzahl Ablieferer

(Skizze)



## Austauschobjekte

### Anlieferungspakete:

- Beliebige Zusammenstellung zu „semantischen Einheiten“
- Beliebige Dateiformate
- Mitlieferung von deskriptiven Metadaten möglich
- Checksummen pro digitales Objekte und/oder pro Anlieferungspaket

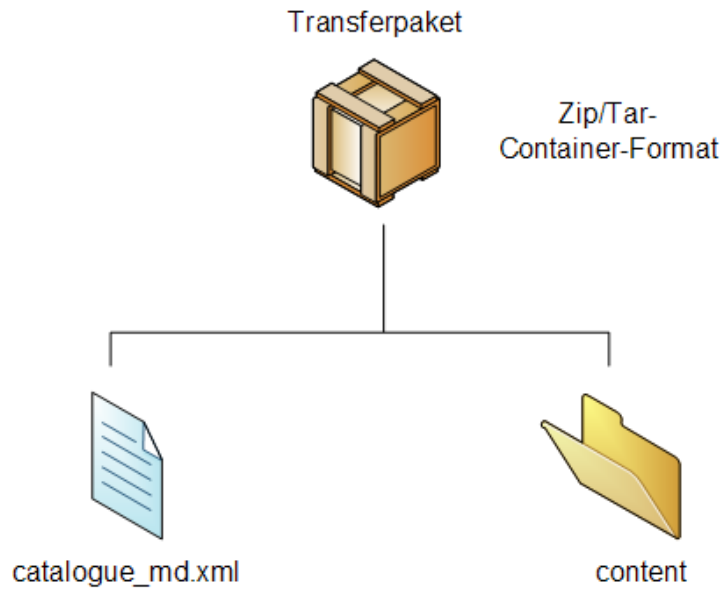
### Rücklieferungsobjekte

Je nach Bedarf können einzelne Archivpakete, oder ganze Migrationshistorien samt generierte Metadaten zurückgeliefert werden.



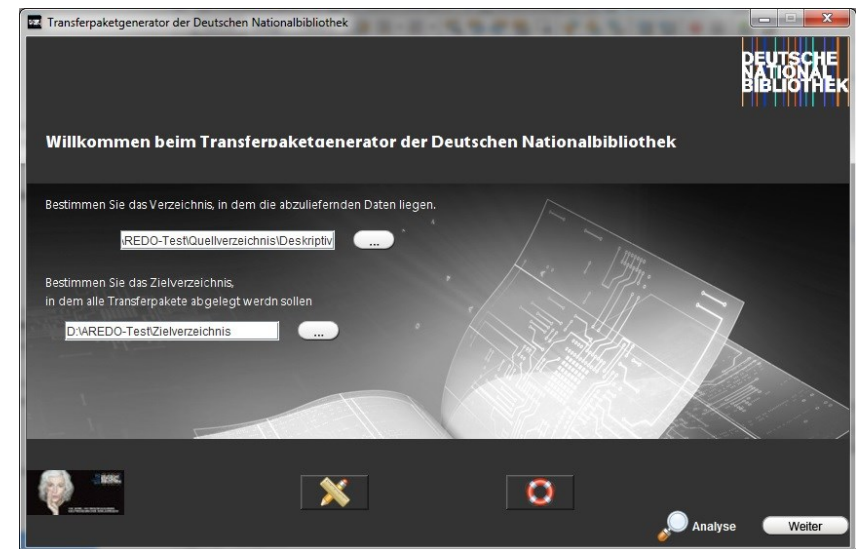
## Aufbau Anlieferungspaket

### Konzept



### Praxis

#### Transferpaket-Generator (Prototyp)



# Risikomanagement

## 1.) Technische Qualitätsprüfung

- Anhand aller technischer Metadaten und
- Vereinbarter gemeinsamer Workflows

## 2.) Klassifizierung in Ingest-Level

Ein Ingest-Level repräsentiert das Ergebnis eines mehrstufigen Prüfverfahrens, welches zurzeit durch mehrere aufeinander aufbauende Dienste innerhalb der DNB realisiert wird.

## Risikomanagement

# Ingest-Level - Erklärung

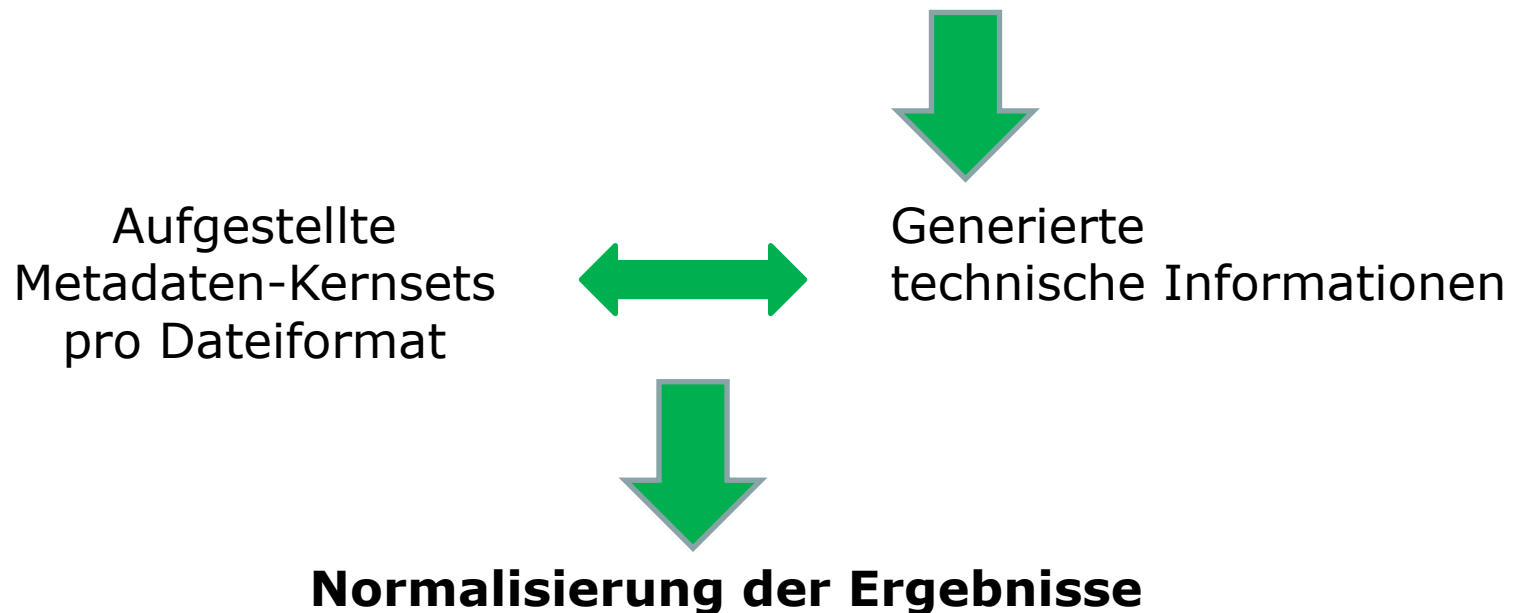
Der Ingest-Level ist eine Abschätzung bzgl. der wahrscheinlichen Langzeitarchivierbarkeit digitaler Objekte, die auf Grundlage momentan zur Verfügung stehender technischer Mittel und den etablierten Geschäftsprozessen der DNB getroffen wird. Der Ingest-Level stellt jedoch keine Aussage über die allgemeine Qualität der geprüften digitalen Objekte dar.



## Risikoabschätzung

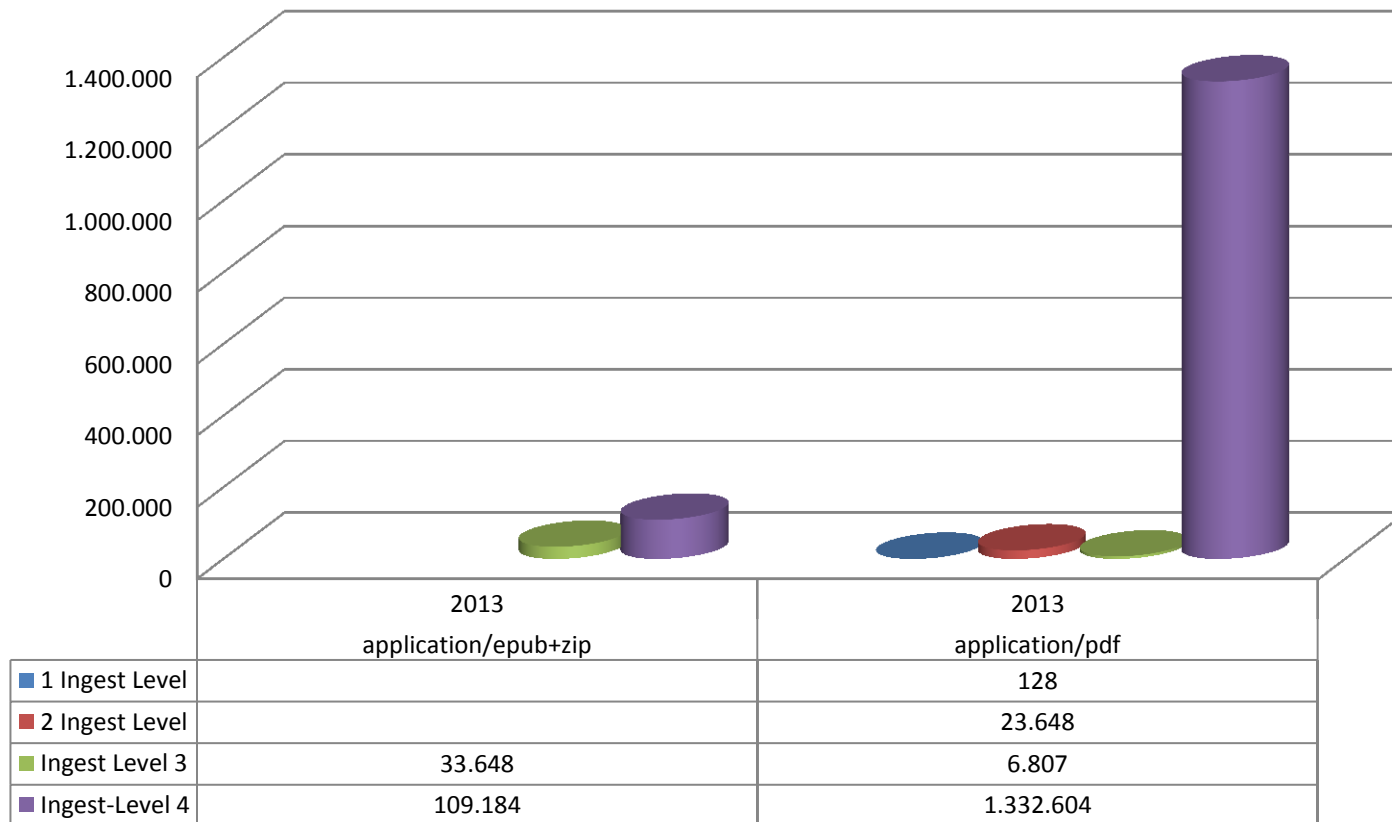
## Generierung technischer Metadaten

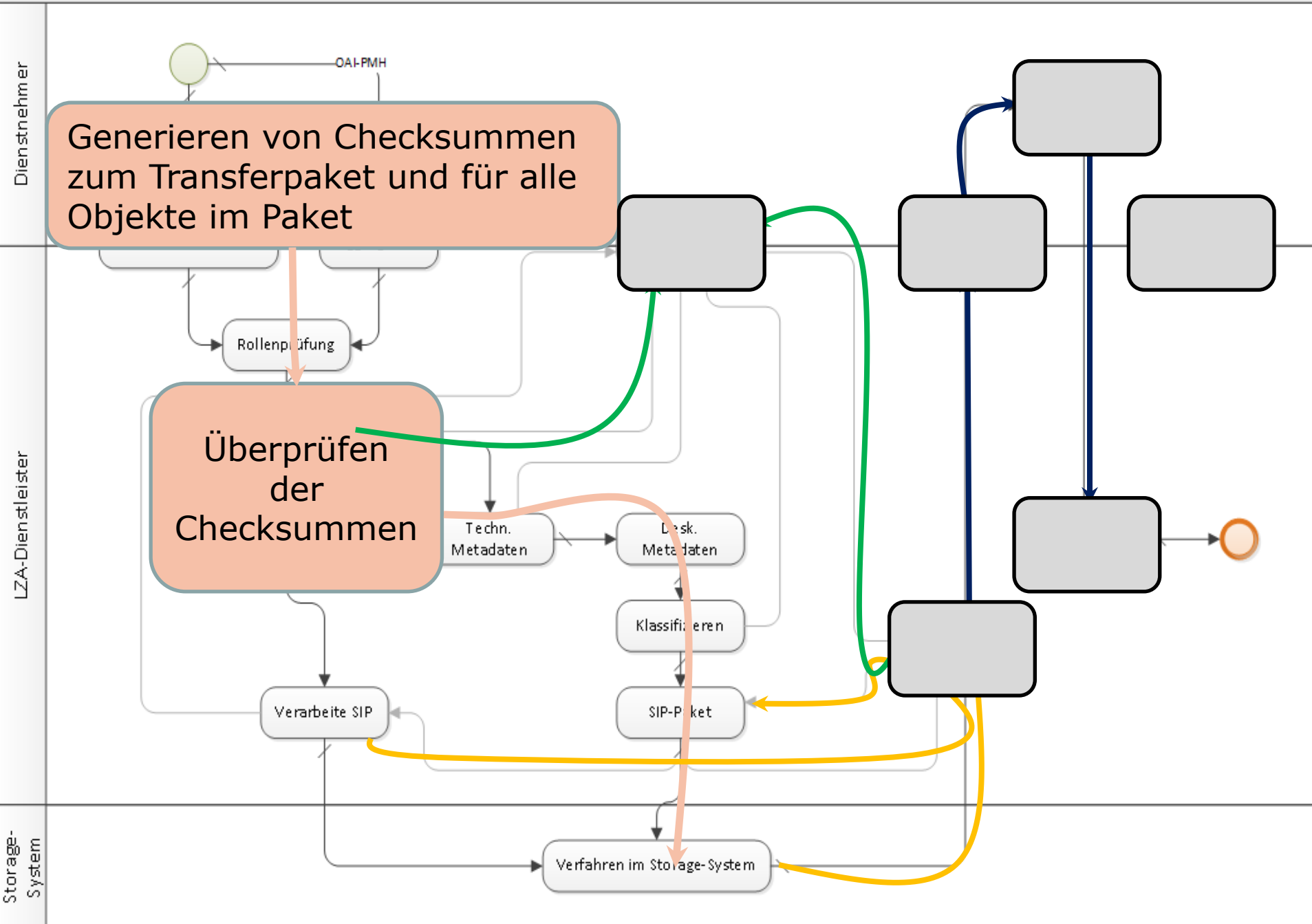
- FITS-Framework
- didigo-Tool (Eigenentwicklung)
- aus Informationssystemen



Ingest-Level	Daten-Integrität	Identifikation	Beschränkungsfreiheit	Format-Metadaten	Validität
Level 0	<b>sehr eingeschränkt LZA-tauglich</b>				
Level 1					
Level 2	<b>bedingt LZA-tauglich</b>				
Level 3					
Level 4	<b>LZA-tauglich</b>				

## Ingest-Level EPUBs & PDFs





## ACCESS

# AREDO - LTP Access Service



### Liste aller abgelieferten Objekte

Sie sehen hier alle von Ihnen abgelieferte Objekte und haben die Möglichkeit sich weitere Informationen anzeigen zu lassen.

angemeldet als DP4lib-Verlag  
(ID: 151453696)

### Ausgabe

urn	Ingest-Transaktions-ID	Eingangsdatum	Transferpaketname	Status	Retrieval
<a href="#">urn:nbn:de:101:1-201202141614</a>	<a href="#">1329242058461479</a>	14.02.2012 18:54:21	20120214_17.zip		
<a href="#">urn:nbn:de:101:1-201112254501</a>	<a href="#">13292286364151323</a>	14.02.2012 15:10:41	20120214_11.zip		

Bei Fragen wenden Sie sich bitte an Stefan Hein (s.hein@dnb.de) oder Karlheinz Schmitt (k.schmitt@dnb.de)  
Version 0.9.4 (13.02.2012)





### Ingest-Verlauf

Sie sehen hier den aktuellen Verlauf einer Archivierung durch den Ingest-Workflow. Sollten Fehler oder Probleme bei gewissen Schritten innerhalb der Prozesskette aufgetreten sein, werden diese hervorgehoben.

ID: 20120214\_11\_err.zip

### Ausgabe

Didigo

Transferpaket: 20120214\_11\_err.zip  
 abgeschlossen am: 14.02.2012 15:00:14  
 Paketbestandteil:  
 Status: DONE  
 Modulenachricht: alle Paketbestandteile erfolgreich gepueft

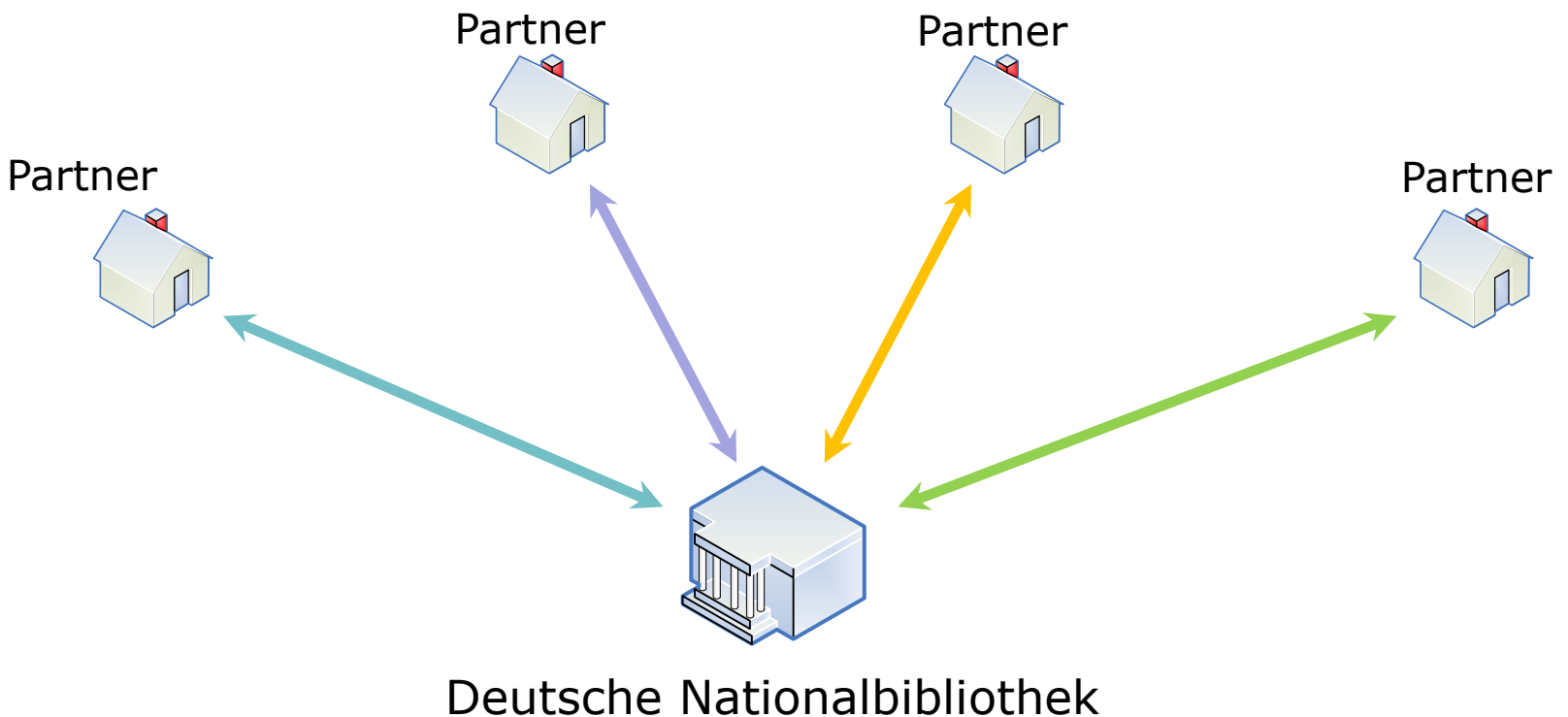
IngestLevelCheck

Transferpaket: 20120214\_11\_err.zip  
 abgeschlossen am: 14.02.2012 15:00:15  
 Paketbestandteil: LVZ\_LPZ\_19\_11\_11.pdf  
 Status: ERROR  
 Modulenachricht: Ingest-Policy: IngestPolicy für DP4lib-Verlag (151552000) - Wohlgeformtheit bzw. Formatvalidität verletzt -> ermitteltes Ingest-Level 3 fuer Format fmt/18 liegt unter der Policy-Vorgabe (4) - Abbruch der Verarbeitung - error-Messages: No file trailer offset=62942880

SubmitErrorProtocol

Transferpaket: 20120214\_11\_err.zip  
 abgeschlossen am: 14.02.2012 15:00:15  
 Paketbestandteil: LVZ\_LPZ\_19\_11\_11.pdf  
 Status: DONE  
 Modulenachricht: Error-Protokoll an <https://dp4lib.langzeitarchivierung.de/report.php> versendet. message: protocol\_type=error; delivery\_role=DP4lib; tp\_name=20120214\_11\_err.zip; error\_module=IngestLevelCheck; error\_message=Ingest-Policy: IngestPolicy für DP4lib-Verlag (151552000) - Wohlgeformtheit bzw. Formatvalidität verletzt -> ermitteltes Ingest-Level 3 fuer Format fmt/18 liegt unter der Policy-Vorgabe (4) - Abbruch der Verarbeitung - error-Messages: No file trailer offset=62942880

**Mit jedem Partner kann im Rahmen des Leistungskatalogs  
ein individueller LZA-Workflow  
vereinbart und umgesetzt werden**



# Vielen Dank!

## Kontakt:

Karlheinz Schmitt  
AREDO

[k.schmitt@dnb.de](mailto:k.schmitt@dnb.de)  
[aredo@dnb.de](mailto:aredo@dnb.de)

Informationen

<http://www.dnb.de/aredo>