

DP4*lib*

Digital Preservation for libraries

Als Beispiel einer spartenübergreifenden
technisch-organisatorischen Organisation

Reinhard Altenhöner & Karlheinz Schmitt

Inhaltsverzeichnis

- 1. Ziel & Grundlagen**
- 2. Kooperationsmodell**
- 3. Aufgaben**
- 4. Vorgehen**
- 5. Ausblick**

Ziel

DP4*lib*

Organisatorische und technische
Weiterentwicklung

der kopal-Lösung zu einem integrierten Dienst

mit dem Ziel der
kommissarischen Langzeitarchivierung digitaler Objekte

Grundlagen

Das kopal-Projekt

KOoPerativer Aufbau eines *Langzeitarchivs* digitaler Informationen

Förderprojekt des Bundesministeriums für
Bildung und Forschung (BMBF)

Laufzeit: 1.7.2004 – 30.9.2007

Projektziel: Aufbau einer von Gedächtnisorganisationen
nachnutzbaren technischen und organisatorischen
Infrastruktur zur Sicherung der Langzeitverfügbarkeit
elektronischer Publikationen

Leitbilder: Kooperation, Nachnutzbarkeit, Nachhaltigkeit

Grundlagen

Digital Information Archiving System (DIAS):

- Archivsystem nach dem OAIS-Referenzmodell
- Backend-System mit definierten Schnittstellen (Internettechnologie)
- Zukunftsfähige Trennung von Speicherkonzept und Datenverwaltung

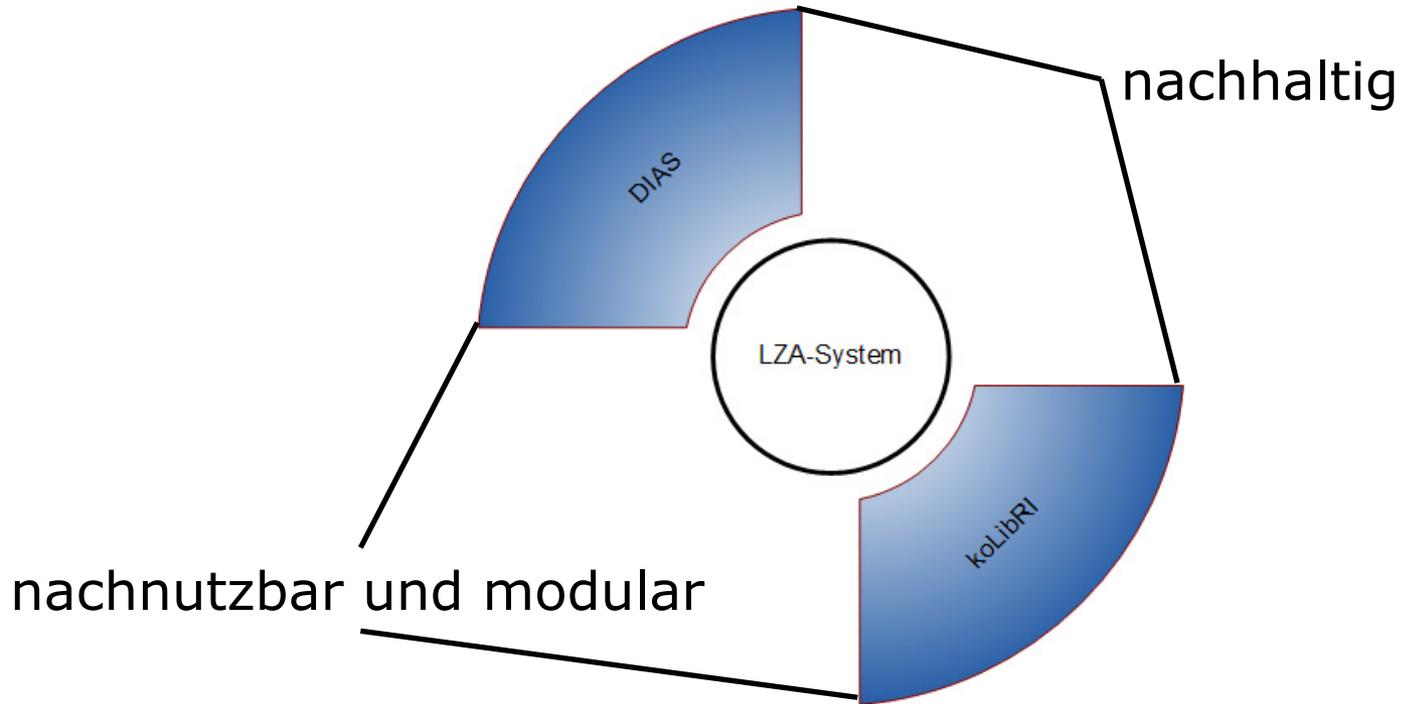
koLibRI-Software:



- Generische freie Softwarebibliothek zur Ein- und Anbindung der unterschiedlichen IT-Infrastrukturen
- Automatische Erstellung von technischen Metadaten
- Konfigurierbare Workflows für das Verwalten, das Einspielen, den Zugriff und die Migration

Grundlagen

kopal-Langzeitarchivierungssystem



Bisher als Inhouse-Lösung eingesetzt

Ziel

... aber wie



Teilaufgaben:

1. Aufbau des Geschäftsprozesses
 2. Überführung in den Produktivbetrieb
 3. Betrieb des Geschäftsprozesses
- Ausreichende finanzielle Mittel zur Verfügung stellen?
Sicher, aber: Was ist ausreichend?)
Kann dies eine Institution alleine?
 - Weitere Ressourcen zwingend notwendig:
 - technisch
 - personell
 - organisatorisch
 - politisch (Awareness!?)
 - ...



Kooperations- Modell

Durchführung:



Niedersächsische Staats- und
Universitätsbibliothek Göttingen

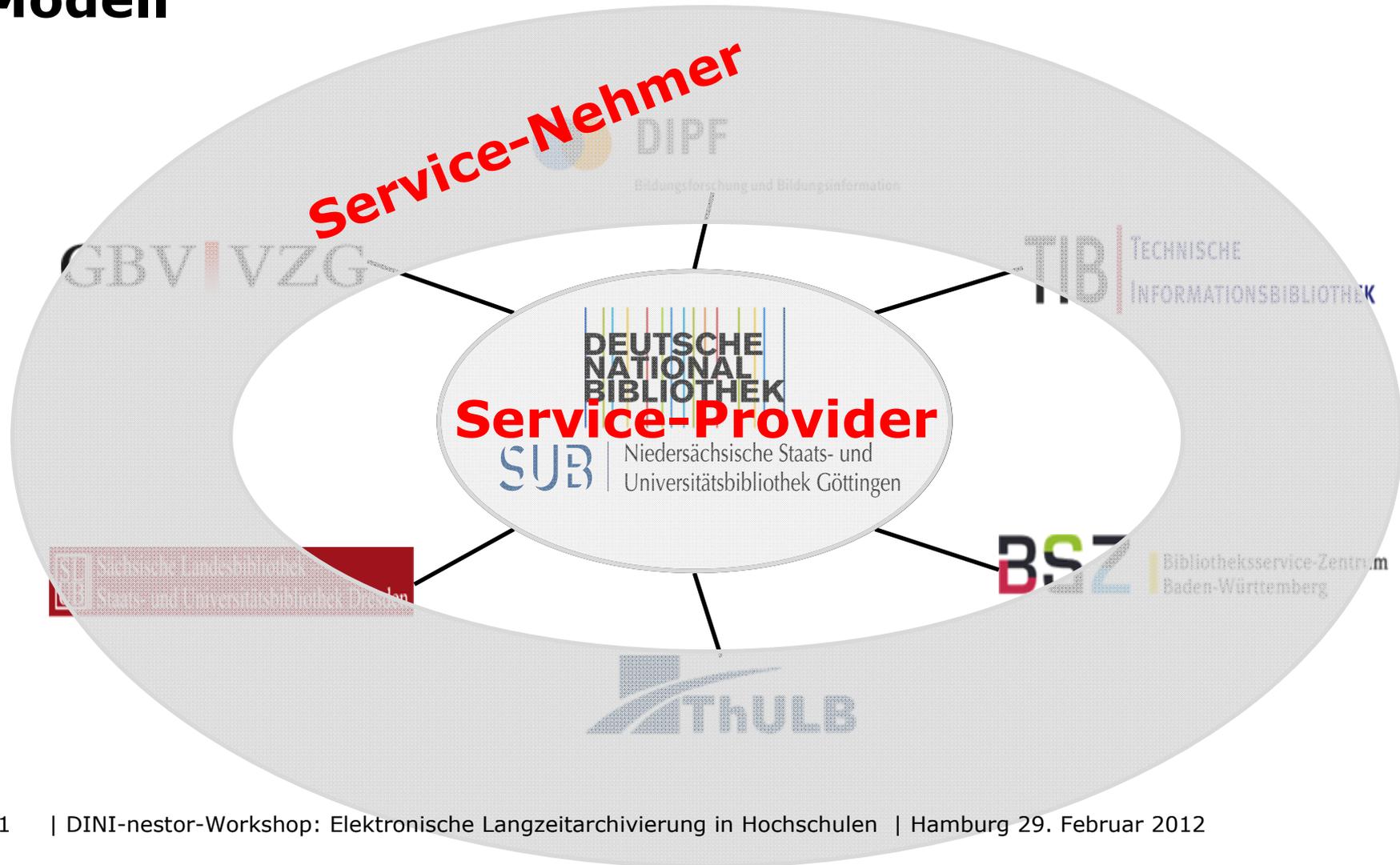
Laufzeit:

2 Jahre (Abschluss Februar 2012)

Gefördert von:



Kooperations- Modell



Aufgaben

Aufbau der Dienstleistung

Technischer Aufbau

- Weiterentwicklung des koLibRI-Toolsets
- Integration in die jeweiligen IT-Workflows
- Auf- und Ausbau der IT-Infrastruktur
- ...

Organisatorischer Aufbau

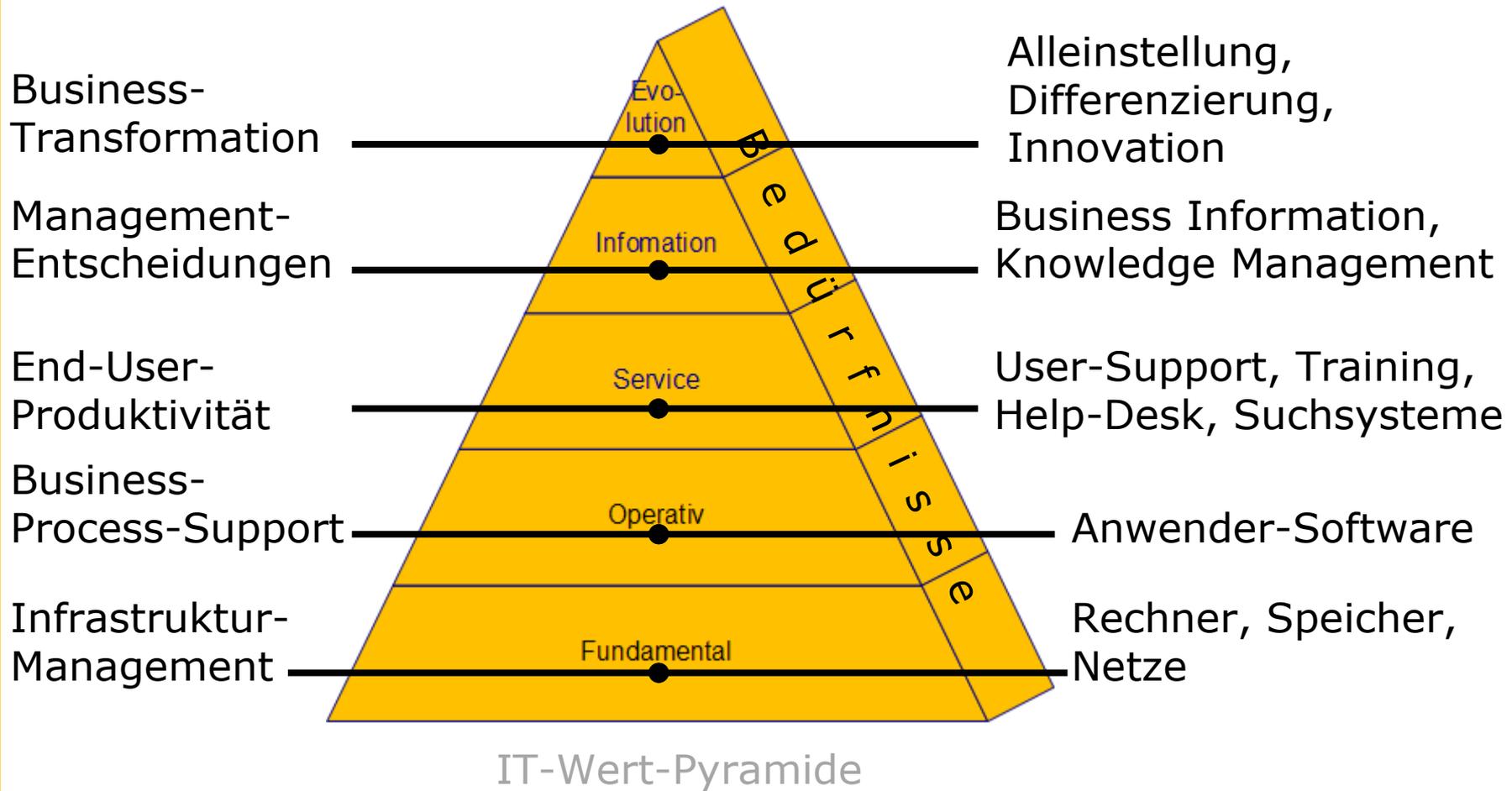
- Aufbau eines gemeinschaftlichen Geschäftsprozesses
- Aufbau eines Betriebsmodells
- Aufbau eines Kostenmodells
- ...

Rechtlich-/vertraglicher Aufbau

- Evaluierung rechtlicher Fragestellungen
- Erstellung von Service Level Agreements (SLA)
- Erstellung von Underpinning Contracts (UC)
- ...

Aufgabe

Aufbau der Dienstleistung



Aufgabe

Anforderungen, die an die Dienste der Langzeitarchivierung gestellt werden, erheben und genau spezifizieren.

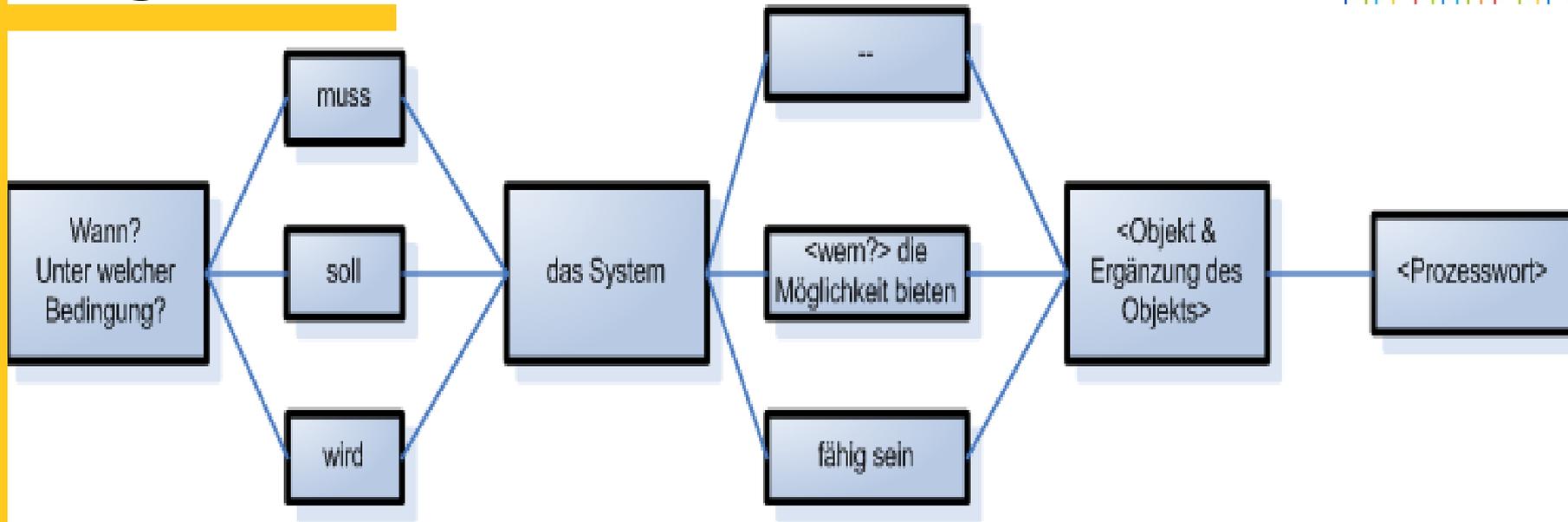


Bilaterale Gespräche,
Erhebungsbögen,
Dokumentation,
Qualitätsmanagement



Anforderungskatalog

Aufgabe: I



Sofort, wenn der Ablieferer ein Transferpaket (TP) an den Hotfolder übertragen hat, muss das System fähig sein das TP an den Importprozess zu übergeben.

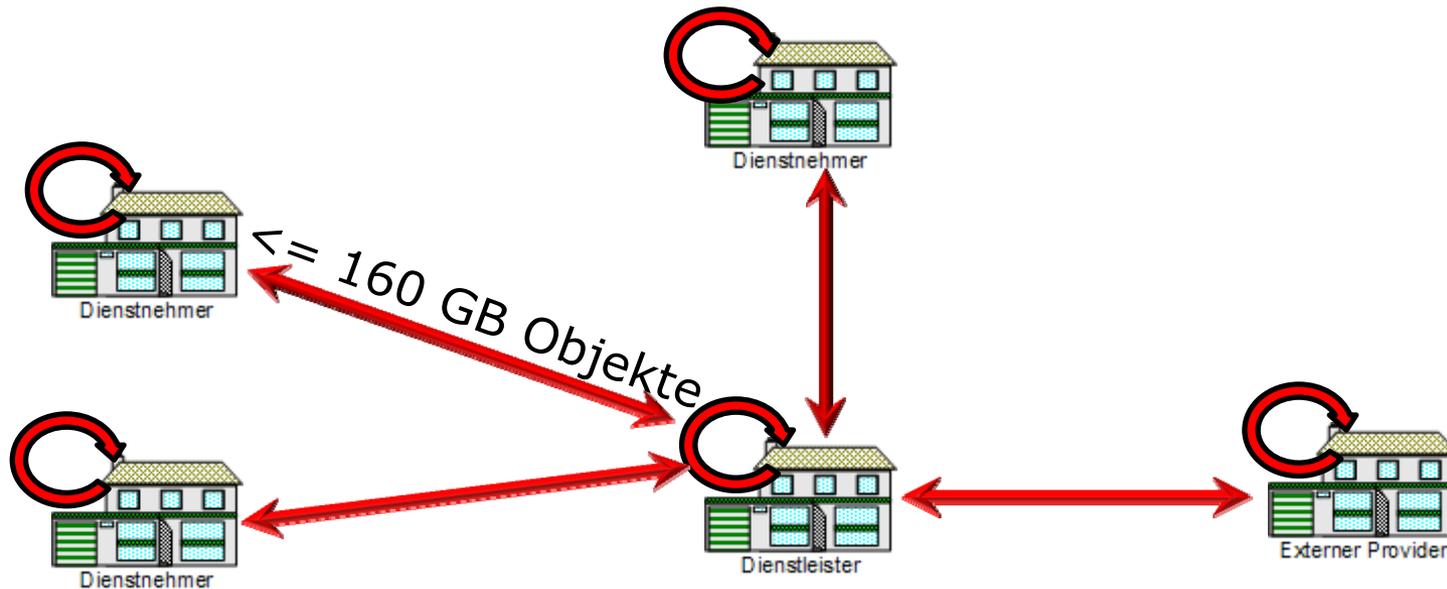
Akzeptanzkriterium: Sofort:

Alle 20 Sekunden wird Hotfolder überprüft.

Technisch

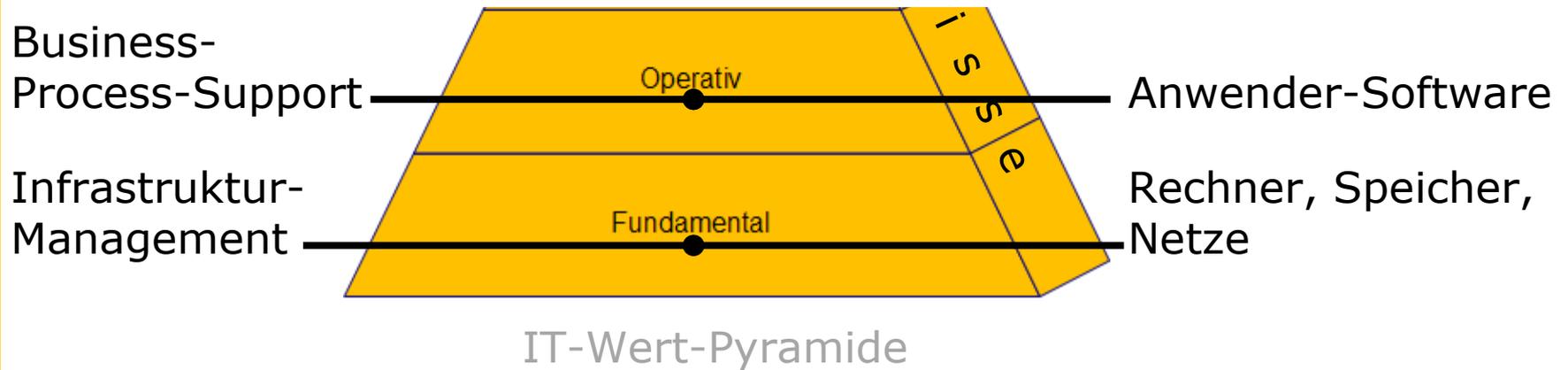
Infrastrukturmaßnahmen an alle Orten notwendig:

- Dienstnehmer
 - Dienstleister
- Outsourcing
Fokussierung auf Kernkompetenzen
- Externer Provider (Rechenzentrum)



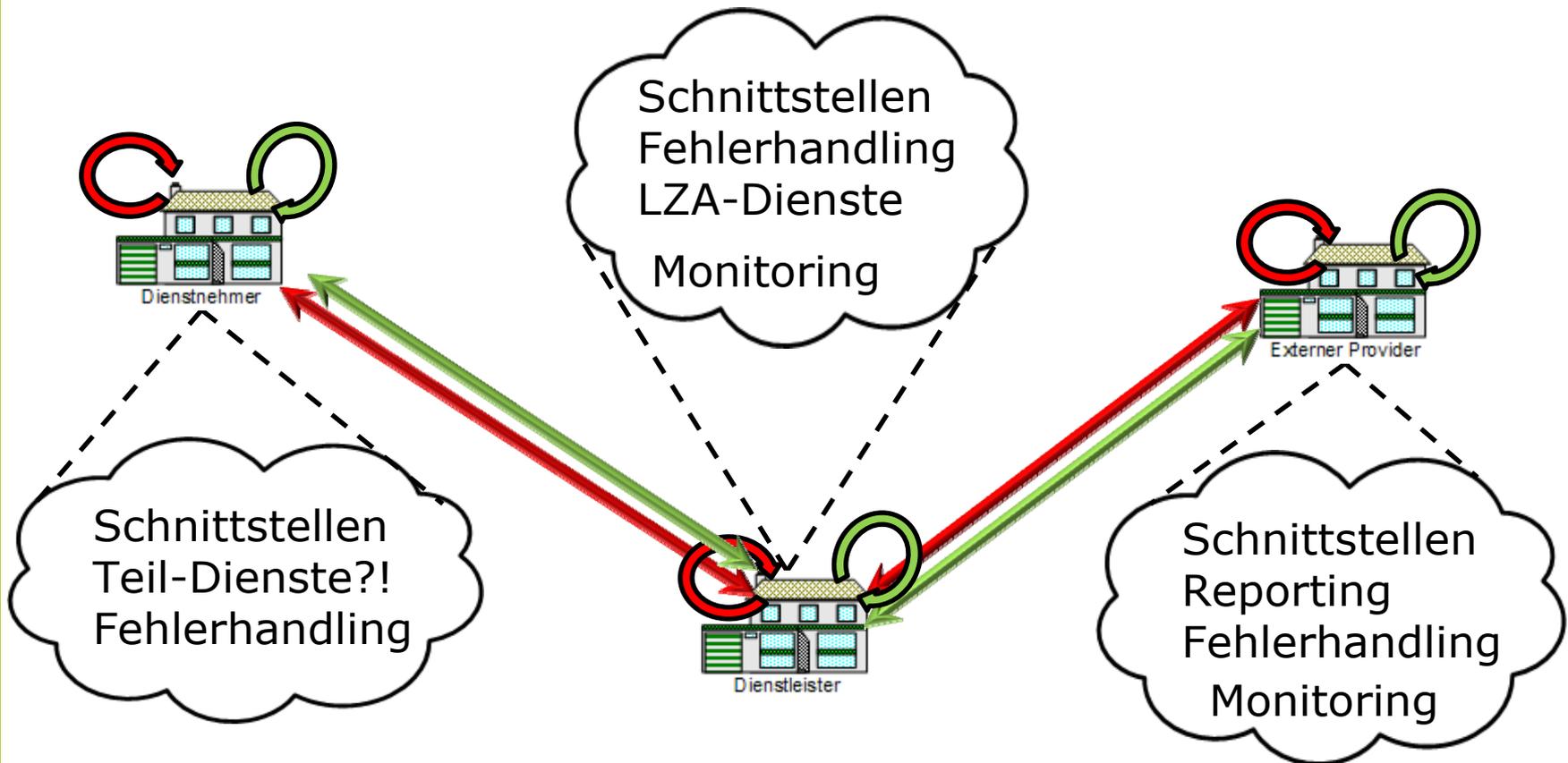
Aufgabe

Aufbau der Dienstleistung



Operativ

↔ Operative Maßnahmen
Implementation der geforderten Funktionalitäten

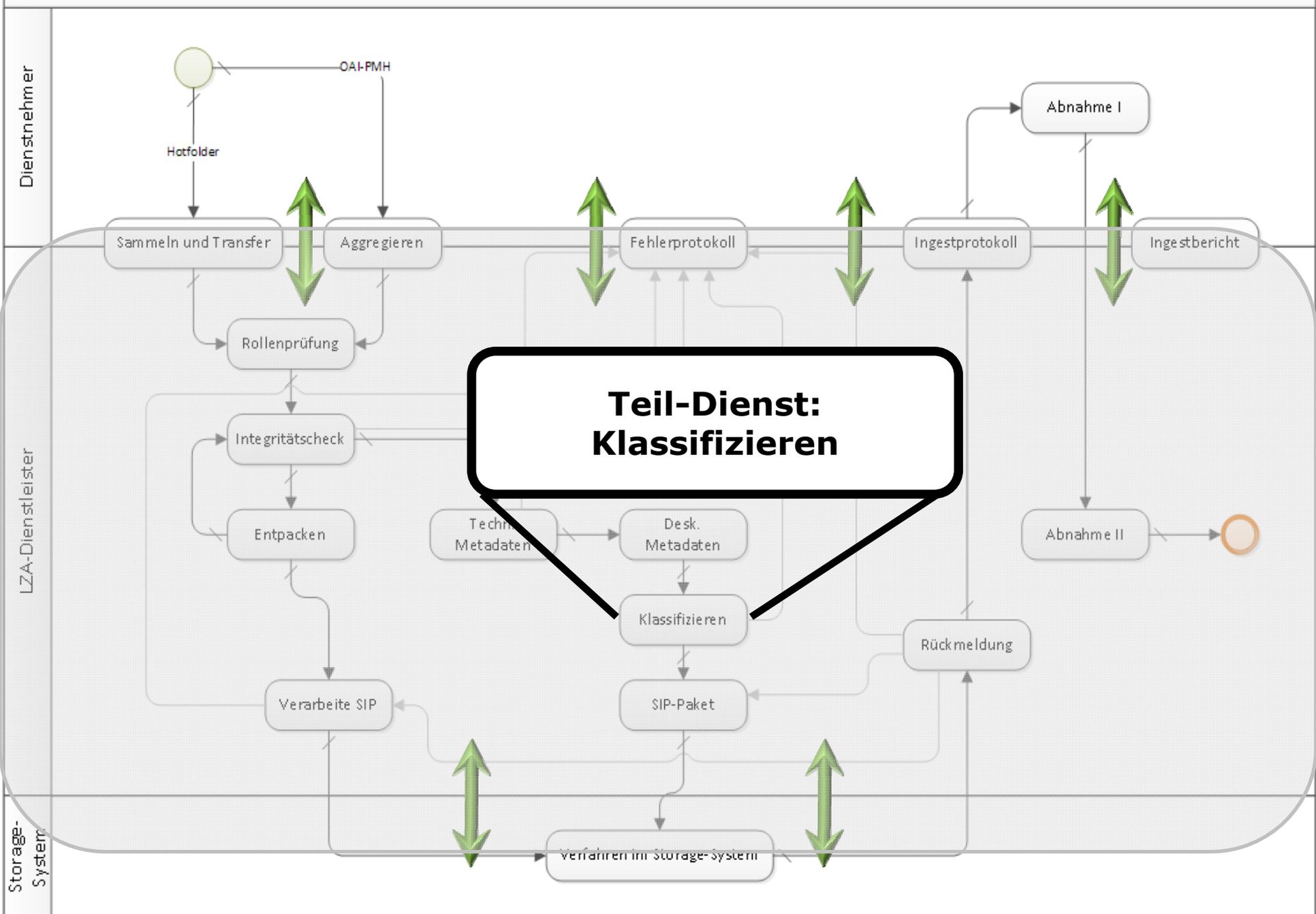


Aufbau der Workflows/Dienstleistungen für:

- den Ingest
- den Access
- das „Curation“

Bitstream Preservation
Preservation Planning

auf Basis der erhobenen Anforderungen.



Klassifizieren



„Risikomanagement“

Problem: Zustand der digitalen Objekte ist unbekannt, dennoch soll die Dienstleistung der Migration möglich oder gar garantiert sein.

DEFINITION: INGEST-LEVEL

Ein Ingest-Level repräsentiert das Ergebnis eines mehrstufigen Prüfverfahrens, welches zurzeit durch mehrere aufeinander aufbauende Dienste innerhalb von DP4lib realisiert wird. Ein Ingest-Level wird jedem digitalen Objekt zugewiesen.

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 des Dienstleisters getroffen wird. Der Ingest-Level stellt jedoch keine Aussage über die allgemeine Qualität der geprüften digitalen Objekte dar.



Risikoabschätzung

Kriterien

- **DATENINTEGRITÄT (DI)**

Der vom Dienstnehmer übermittelte Bitstream des digitalen Objekts hat sich im Zuge der Datenübertragung nicht verändert.

- **IDENTIFIKATION (ID)**

Das digitale Objekt wurde hinsichtlich seines Dateiformats identifiziert.

Service

- BESCHRÄNKUNGSFREIHEIT (BF)

Das digitale Objekt ist beschränkungsfrei, d. h. es existieren keine - seitens des Dienstleisters - technisch erkennbare Beschränkungen, welche die Nutzung des digitalen Objektes beeinträchtigen oder unmöglich machen.

- EXTRAKTION FORMATSPEZIFISCHER TECHNISCHER METADATEN (MD)

Formatspezifische technische Metadaten konnten gewonnen werden.

- FORMAT-VALIDITÄT (V)

Das Dateiformat des digitalen Objekts ist gegenüber einer Formatspezifikation valide.

Ingest-Level

	DI	ID	BF	MD	V
Level 0	X	O	O	O	O
Level 1	X	X	O	O	O
Level 2	X	X	X	O	O
Level 3	X	X	X	X	O
Level 4	X	X	X	X	X

Teildienstleistungen des Ingest-Dienstes

Paket-Integrität	Datei-Integrität	Deskriptive Metadaten	Pflicht-abgabe	Technische Metadaten	Ingest Level Check	Fehler Protokoll	Ingest Protokoll
							

Format-Policy:

Format	Ingest Level
PDF	4
JPEG	4
TXT	4

Absprache von Policies am Beispiel des Ingests

Ingest-Policy:

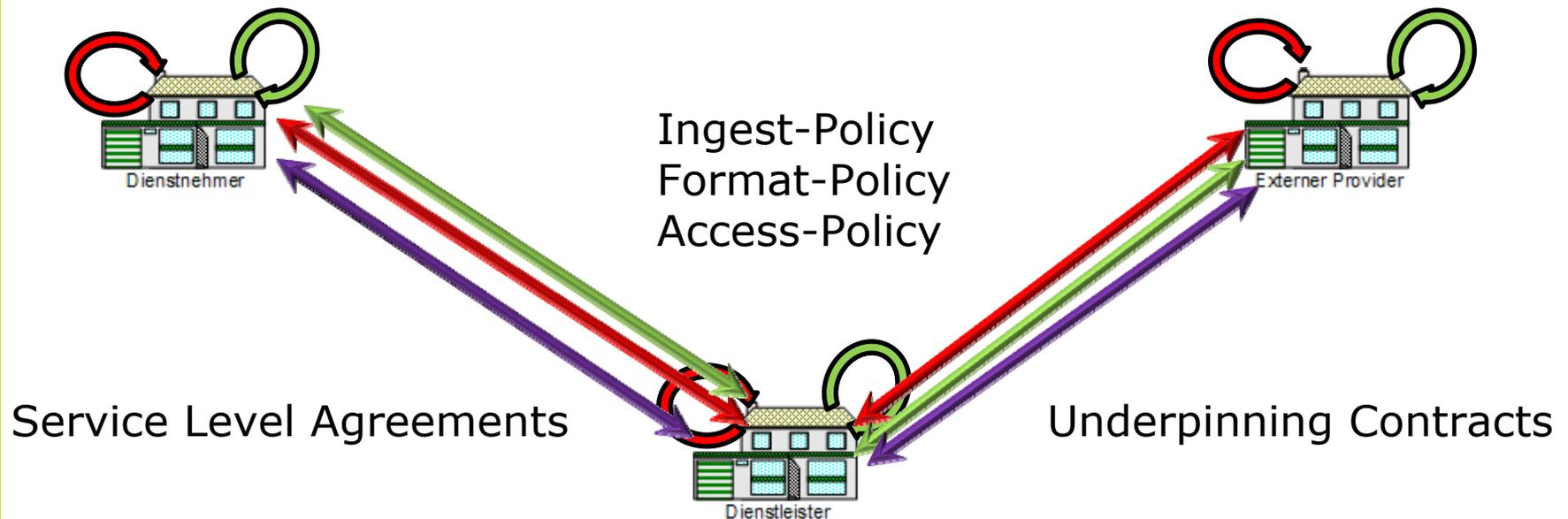
- Welche Ingest-Dienstleistungen werden gewünscht?
- Ansprechpartner (Operativ, Strategisch)
- Schnittstellen
- User-Support
- Eskalationswege
- ...

Service

Welche Dienstleistungen sollen erbracht werden?

Von wem sollen die Dienstleistungen erbracht werden?

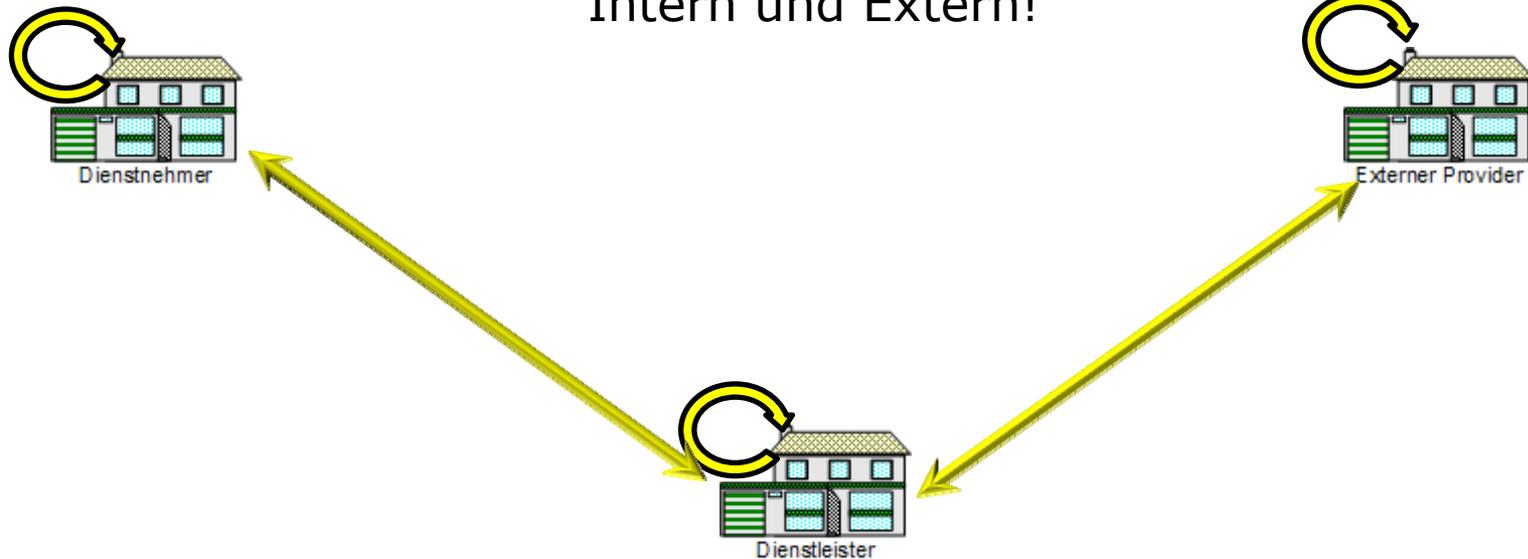
In welcher Qualität sollen die Dienstleistungen erbracht werden? (Service Level)



Incident-Management

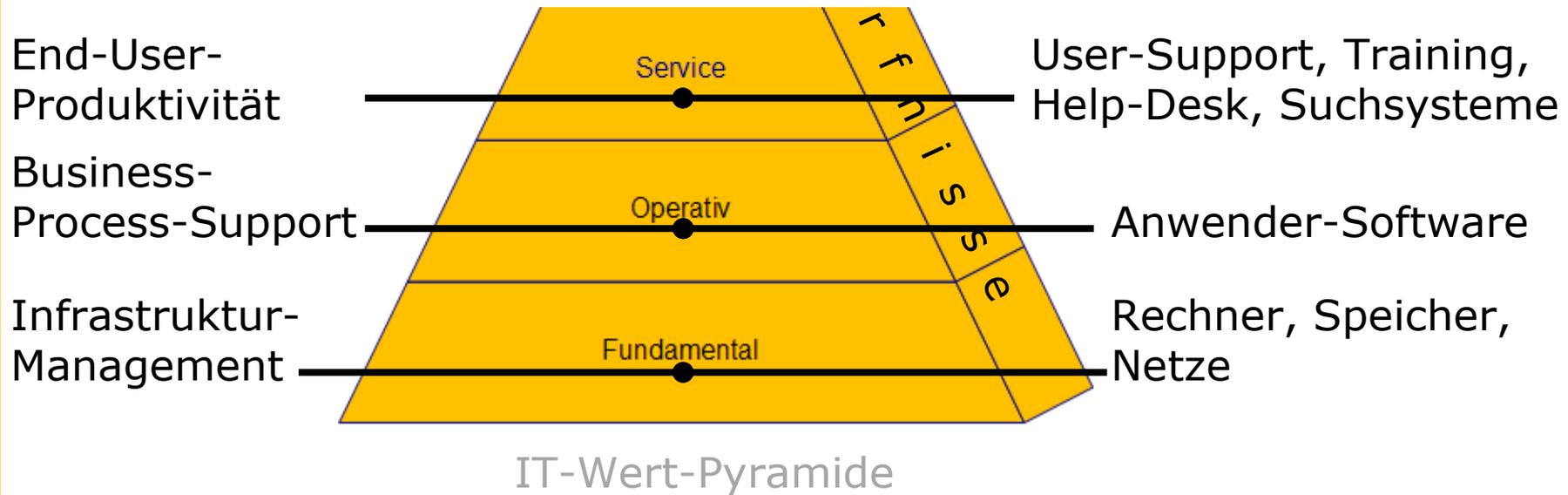
Berichtswesen
Fehler-Handling

Intern und Extern!

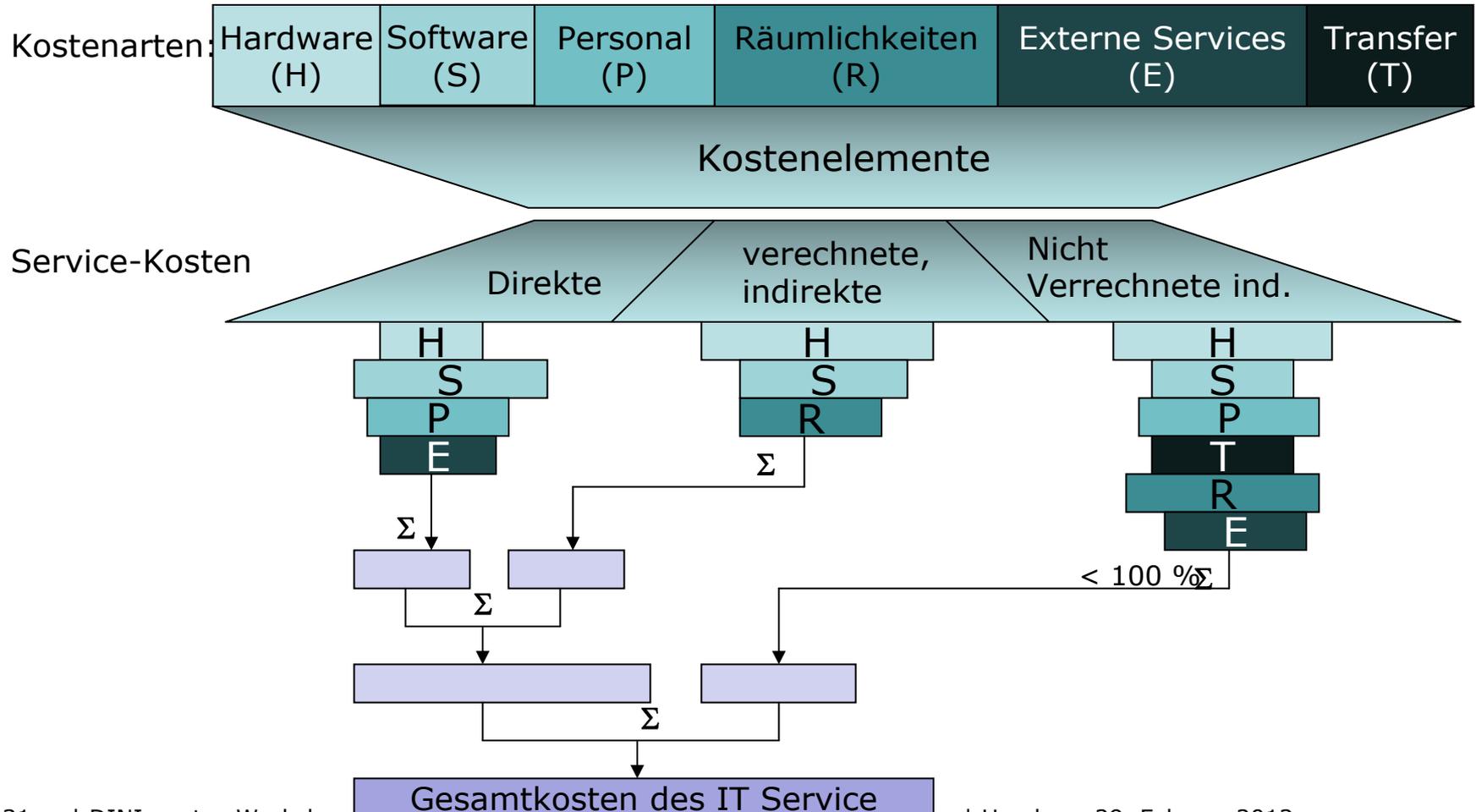


Aufgabe

Aufbau der Dienstleistung

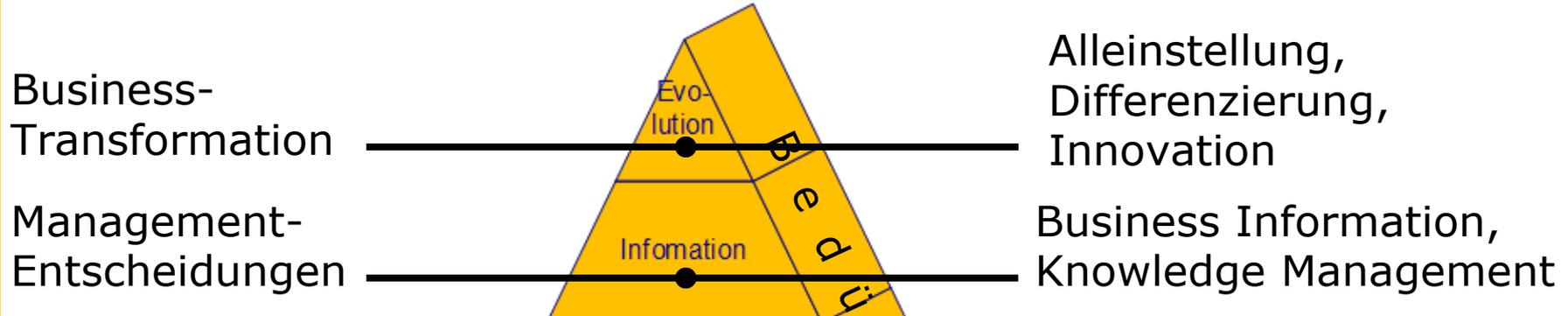


Was Kostet der Aufbau und der Betrieb?



Ausblick

Aufbau der Dienstleistung



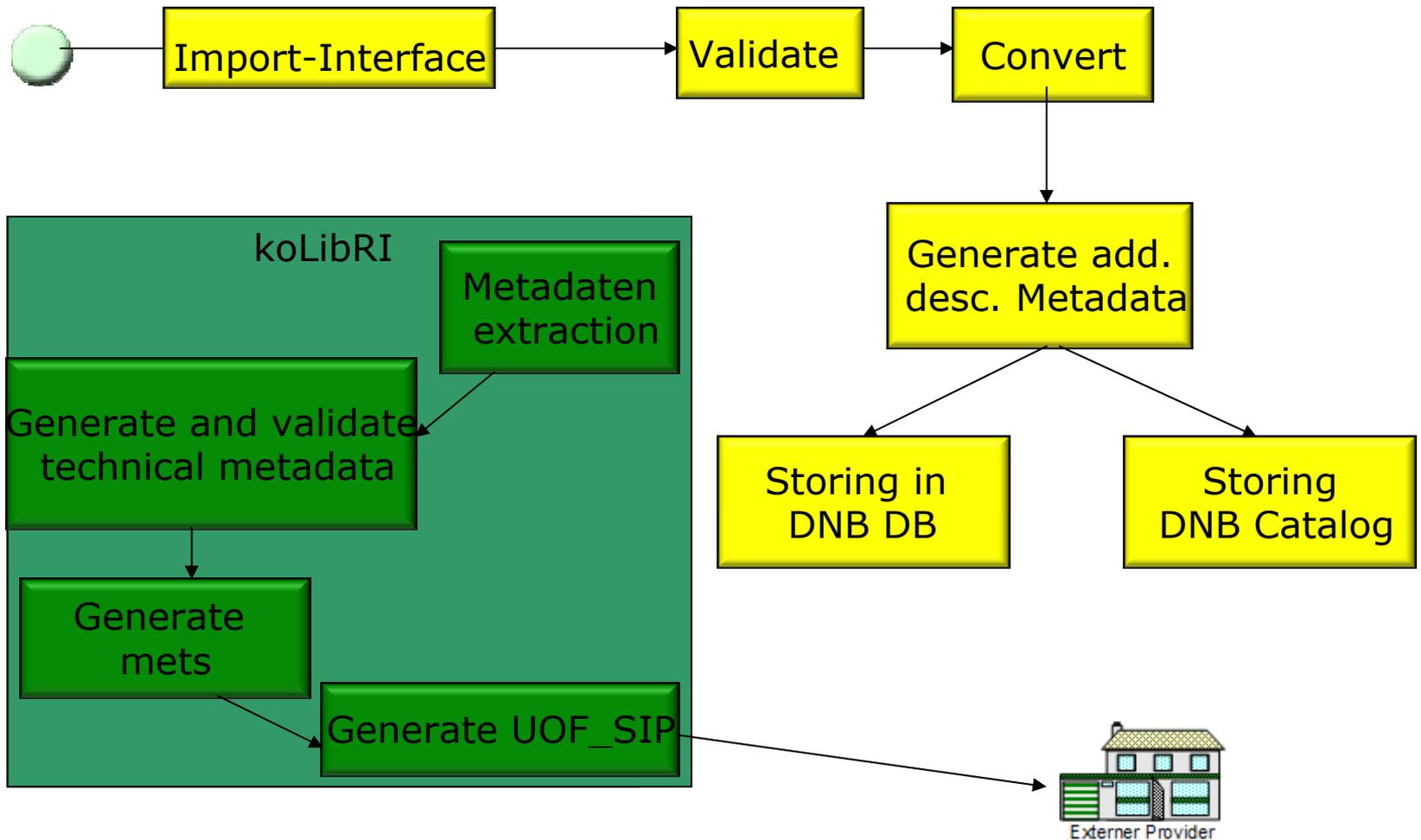
Wie soll der LZA-Dienst weiter entwickelt werden?

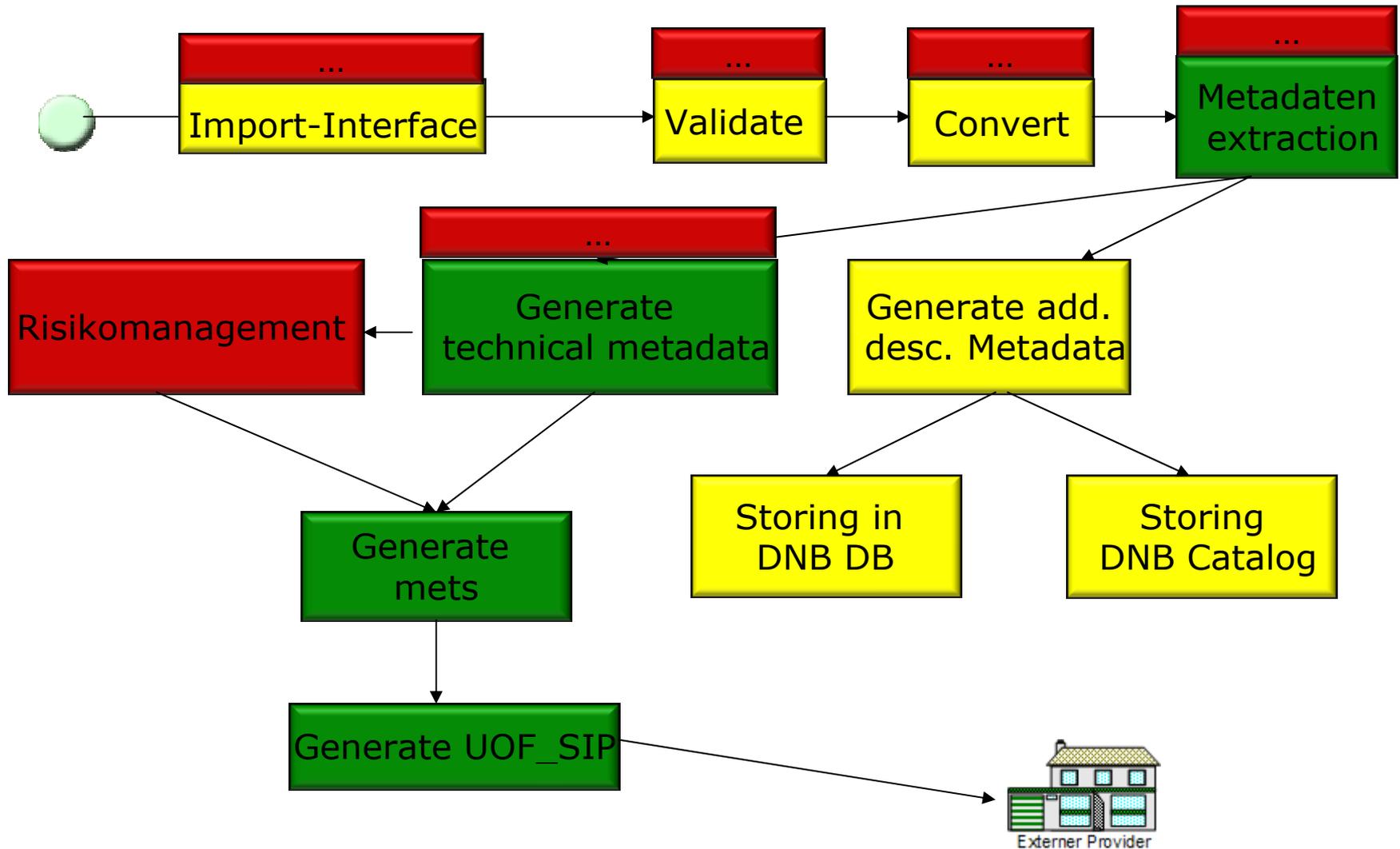
IT-Wert-Pyramide

Vielen Dank für die Aufmerksamkeit!

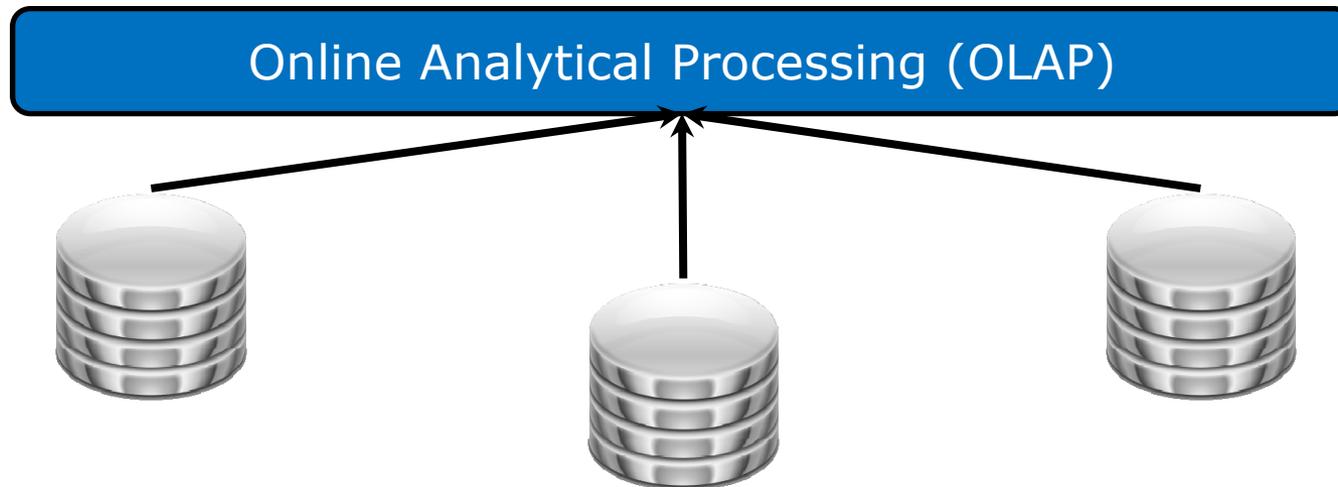


Operativ





- Welche Datenmengen sind verarbeitet worden?
- Welche Datenformate?
- Welche Tendenzen sind zu erkennen?
- Welche Fehler traten vermehrt auf?
- ...

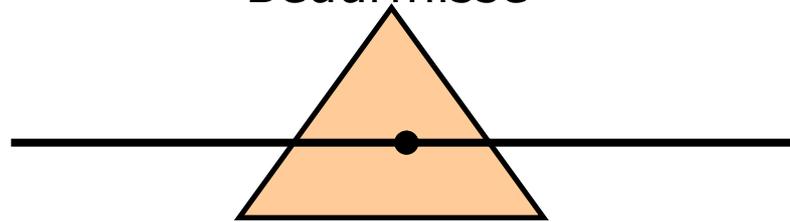


Evolution



Business-
Transformation

Evolution's
Bedürfnisse



Alleinstellung,
Differenzierung,
Innovation

Wie soll der LZA-Dienst weiter entwickelt werden?

Vergleich Anforderungen vs. DIAS-System

- Zu archivierende Objekte bis zu 160 GB groß
- Es können mehrere Millionen einzelne Dateien in einem zu archivierenden Paket enthalten sein
- Abgabe ganzer Kollektionen auf einmal



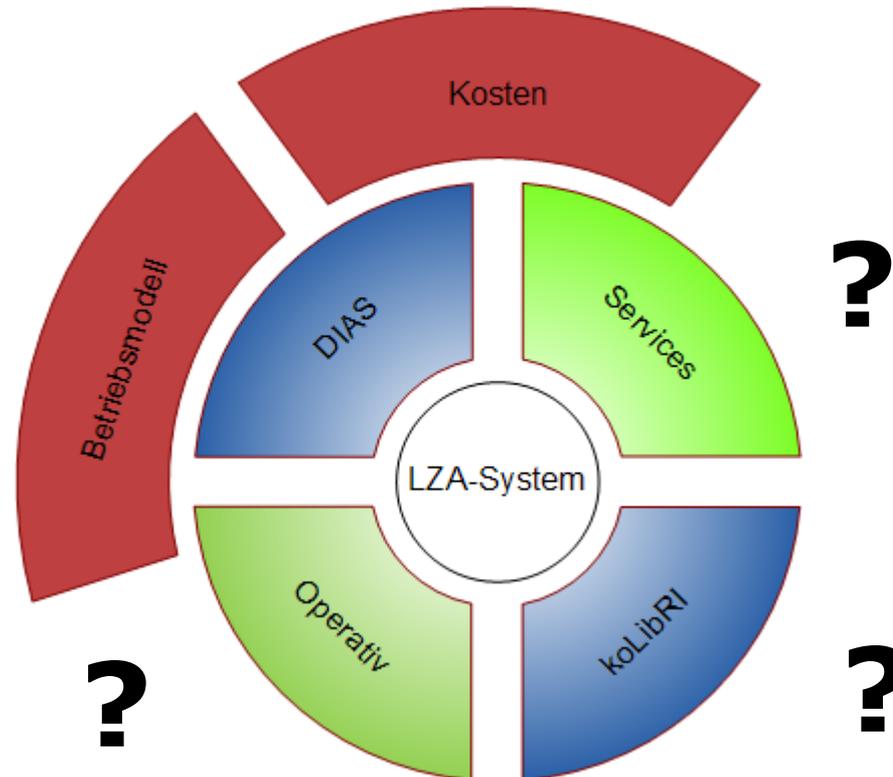
Neukonfiguration des DIAS-Systems wurde durchgeführt um Performance-Anforderungen zu erfüllen.

Vergleich Anforderungen vs. Interne IT-Infrastruktur



Hotfolder zur Massenan- und auslieferung
Parallelverarbeitung ermöglichen

LZA-Dienstleistung



Nun dienstleistungsorientierte Lösung!