

**One-stop publishing and archiving:
Forschungsdaten für Promotionsvorhaben über Repositorien publizieren und archivieren: Eine landesweite Initiative im Rahmen des Projekts bwDataDiss am Beispiel des Karlsruher Instituts für Technologie (KIT)**

Tobias Kurze & Regine Tobias, E-Science Tage in Heidelberg am 17.03.2017

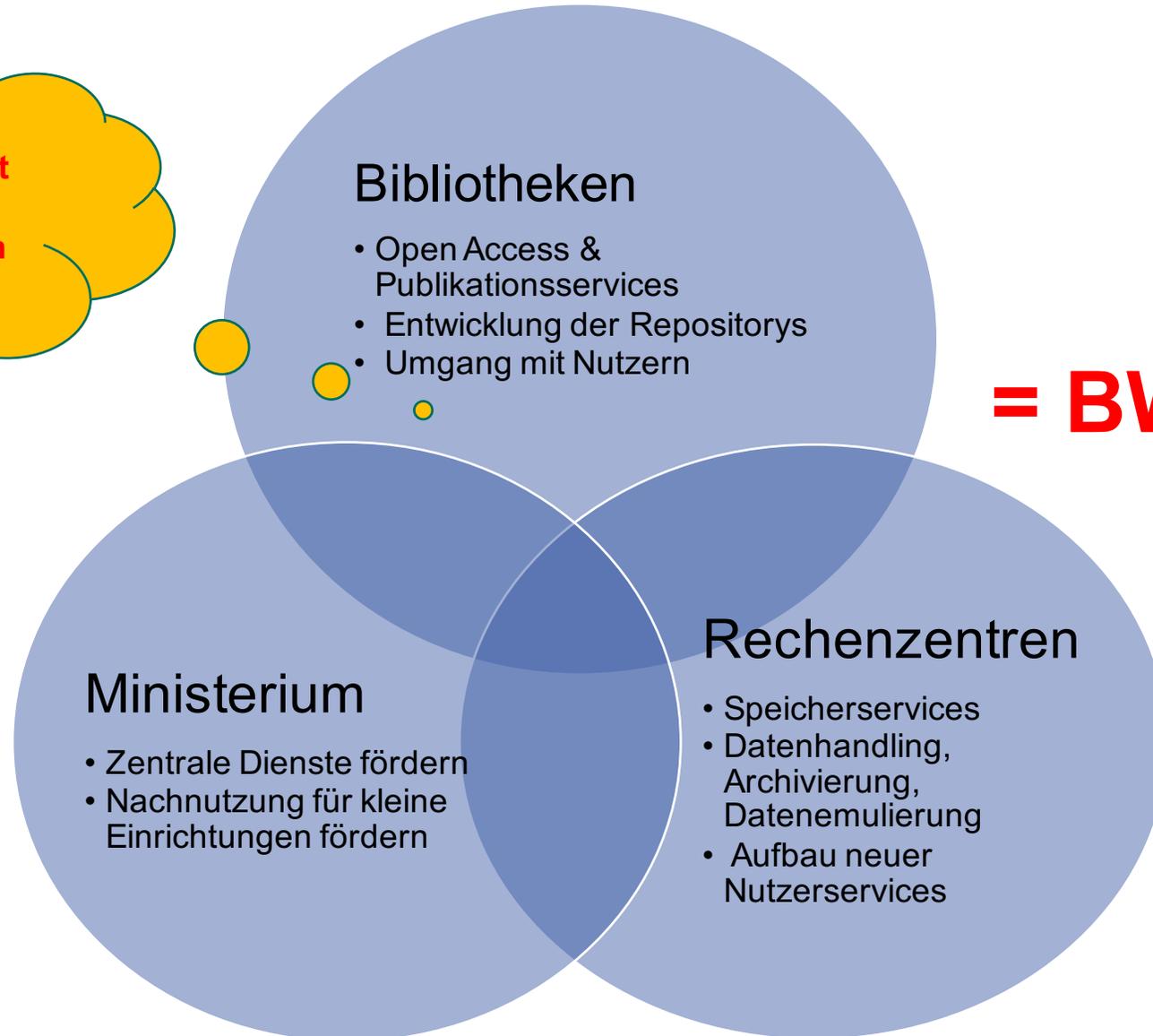
KIT-BIBLIOTHEK





Projektidee

**Forschungsdaten mit
den Workflows der
Veröffentlichung von
Dissertationen
verknüpfen**



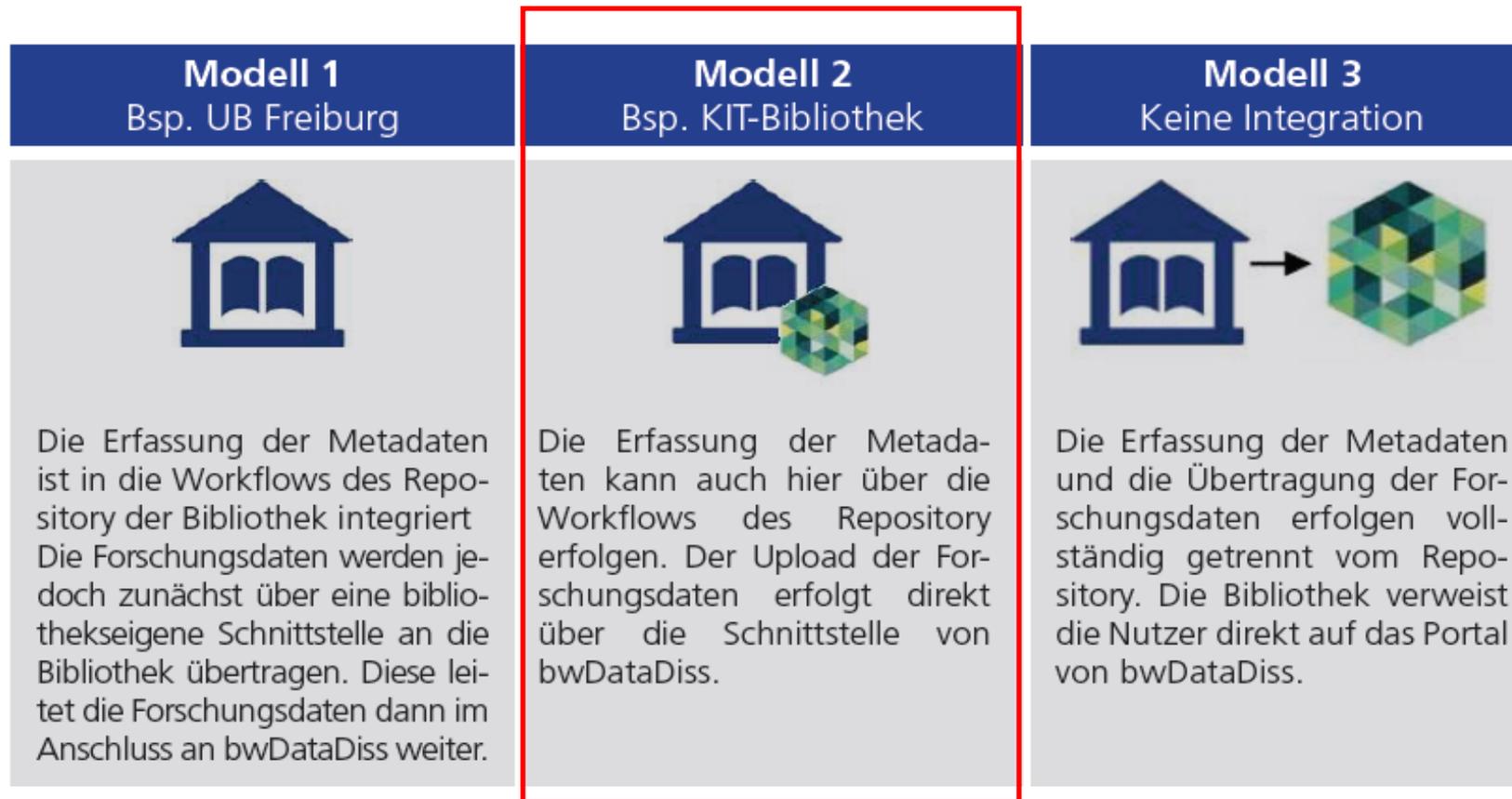
= BW DATA DISS

bwDataDiss – Was und wer ist das?

- Außerdem:
 - Vor der Archivierung führt bwDataDiss eine Charakterisierung der erhaltenen Forschungsdaten durch
 - Die Ergebnisse der Charakterisierung werden den Bibliotheken als Grundlage zur Bewertung der Daten zur Verfügung gestellt
 - Unterstützt die Qualitätssicherung der Bibliotheken
 - Basis für neue Services
- Projekt & Partner: Juli 2014 – August 2017
 - Universität Freiburg
 - Bibliothek
 - Rechenzentrum → Charakterisierung
 - KIT
 - Bibliothek
 - SCC → Archivierung

Ergebnisse (1) – Nachnutzung von bwDataDiss

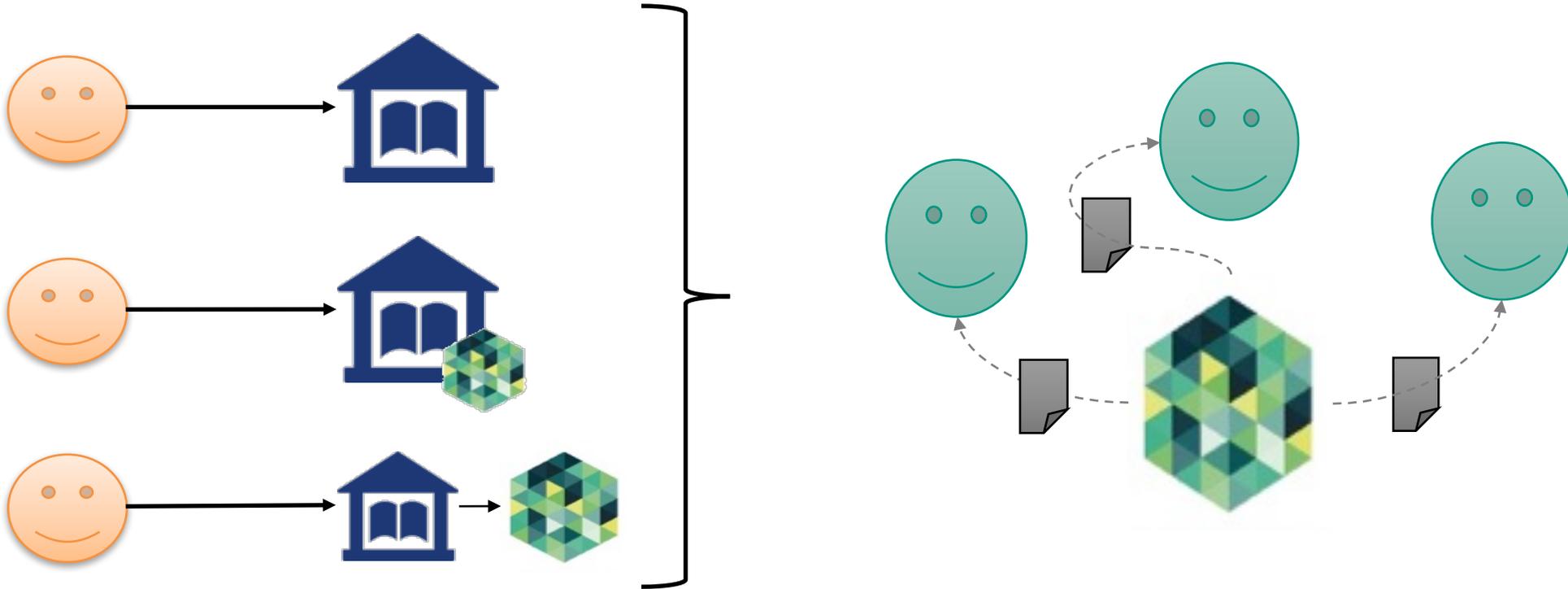
- Erarbeitung von verschiedenen Modellen für die Verknüpfung mit den lokalen Repositorys:





Integrationsmodelle

- bwDataDiss ist sehr flexibel bzgl. Nutzung und Integration
- Alle Varianten und Modelle nutzen das bwDataDiss-Portal für den Zugriff auf die Daten.



Ergebnisse (2) - Aufbau eines neuen Forschungsservices am KIT



Willkommen bei KITopen-Erfassung

Login Bibliotheksausweis Login bwDataDiss

	Kontonummer des Bibliotheksausweises / @Alias
	Passwort
Anmelden	

Forschende und Promovierende des KIT können über KITopen ihre Publikationen (bibliographische Angaben, Volltexte, Dissertationen) und weitere Daten (Forschungsdaten, Indikatoren des Großforschungsbereichs) erfassen. Vor der Veröffentlichung erfolgt eine Validierung durch die KIT-Bibliothek. Weitere, allgemeine Informationen zu KITopen

KITopen für Forschende ▼
In 3 Schritten zum Ziel

- 1. Registrierung (einmalig) ▼
- 2. Anmeldung ▼
- 3. Publikationen und weitere Daten erfassen ▼

KITopen für Promovierende ▼
In 3 Schritten zum Ziel

- 1. Registrierung (einmalig)/ Anmeldung ▼
- 2. Dissertation publizieren ▼
- 3. Forschungsdaten publizieren (fakultativ) ▼



Integration des Workflows der Pflichtabgabe der Dissertationen

i Über die „Suche“ im linken Menübereich können Sie prüfen, ob Ihre Publikationen bereits erfasst sind.

Was möchten Sie tun? **i**



**Publikationen in KITopen
erfassen und publizieren**



**Meine Dissertation in KITopen
veröffentlichen (Pflichtabgabe)**



Login Bibliotheksausweis



Login bwDataDiss



Kontonummer des Bibliotheksausweises / @Alias



Passwort

Anmelden

Interaktion der Dienste KITopen und bwDataDiss



Forschungsdaten in bwDataDiss hochladen

Daten Upload

Prüfsummenberechnung: 0% Datentypenübersicht (inges. / gewählt: 0) Uploadfortschritt: 0%

Upload-Fortschritt und Datentypenübersicht

Lokale Dateien Dateien auf dem Server

Upload starten / fortsetzen Upload abschließen Upload später fortsetzen

Zuletzt eingegebener Datensatz

Publikationstyp	Hochschulschrift
KITopen-ID	1000065973
Titel	test
Verfasser	test
Jahr	2017
Zugehörige Institution(en) des KIT	Institut für Angewandte Informatik/Automatisierungstechnik
Verlag, Ort	Karlsruhe
Abschlusstyp	Dissertation
Fakultät	Fakultät für Bauingenieur-, Geo- und Umweltsenschaften
Institut	Institut für Angewandte Informatik/Automatisierungstechnik
Referent / Betreuer	test
Eingabedatum	07.03.2017
Status	Erfasst
Erstgende Person	Danielsch, Tobias (regne.tobias@kit.edu)
Erstgende Institution	KIT Bibliothek (IBI)
Promotionsordnung	Fakultät für Bauingenieur-, Geo- und Umweltsenschaften
Prüfungstermin	23.03.2017
DOI der Dissertation	10.5072/R/1000065973

- Forschungsdatenaupload
- Charakterisierungsergebnisse



Bewahren

Archivieren Sie Ihre Forschungsdaten

Finden

Forschungsdaten in unserem Archiv suchen

Neueste Daten

Jüngste Forschungsdaten in bwDataDiss

Willkommen bei bwDataDiss!

Der Dienst bwDataDiss stellt Bibliotheken in Baden-Württemberg Infrastruktur zur Verfügung, um einerseits den Erhalt von Forschungsdaten - die im Rahmen von Dissertationen entstanden sind - sicherzustellen, und andererseits diese Forschungsdaten anderen Forschenden zur Nachnutzung bereitzustellen.

Langzeitarchivierung
bwDataDiss speichert Forschungsdaten auf Speicherbänder wo diese für einen Zeitraum von mindestens zehn Jahren sicher aufbewahrt werden.

Open Access
bwDataDiss bekennt sich zu Open Access und stellt Forschungsdaten in der Regel frei zur Verfügung. Die Forschungsdaten sind mit Open-Access-Lizenzen versehen, die die Bedingungen der Nachnutzung regeln.

Embargo
Nicht immer kann ein promovend sofort den Zugang zu den Forschungsdaten ermöglichen. Daher kann der Zugriff auf Forschungsdaten während eines Embargozeitraumes

Nachhaltige Bereitstellung
Forschende können die in bwDataDiss aufbewahrten Forschungsdaten für ihre Forschungsvorhaben und

- Erfassung der Metadaten der Forschungsdaten
- Kommunikation mit Nutzer



Details

Titel	Validation Framework for RDF-based Constraint Languages - PhD Thesis Research Data
Verfasser	Hartmann, Thomas
Zugehörige Institution(en) am KIT	Institut für Angewandte Informatik und Formale Beschreibungsverfahren (AIFB)
Publikationstyp	Forschungsdaten
Jahr	2017
Erstellungszeitraum	31.12.2016
Identifikator	DOI(KIT): 10.5445/IR/1000067357 KITopen ID: 1000067357
Lizenz	CC BY 3.0 DE: Creative Commons Namensnennung 3.0 Deutschland
Abstract (englisch)	Research data, research results, and publications of the PhD thesis entitled 'Validation Framework for RDF-based Constraint Languages', submitted to the Department of Economics and Management at the Karlsruhe Institute of Technology (KIT).
Liesmich	# PhD Thesis --- * Title: *[Validation Framework for RDF-based Constraint Languages][http://dx.doi.org/10.5445/IR/1000067357] * Author: *[Thomas Hartmann][http://www.dr-thomashartmann.de]** * Examination Date: 08.07.2016 * University: [Karlsruhe Institute of Technology (KIT)][https://www.kit.edu/english/] * Chair: [Institute of Applied Informatics and Formal Description Methods][http://www.aifb.kit.edu/de/haendeler/] * Department: [Department of Economics and Management]

KITopen
> XML-Format
> PDF-Format
[Download \(bwDataDiss\)](#)



KIT-Katalog: Verweis auf Forschungsdaten in bwDataDiss

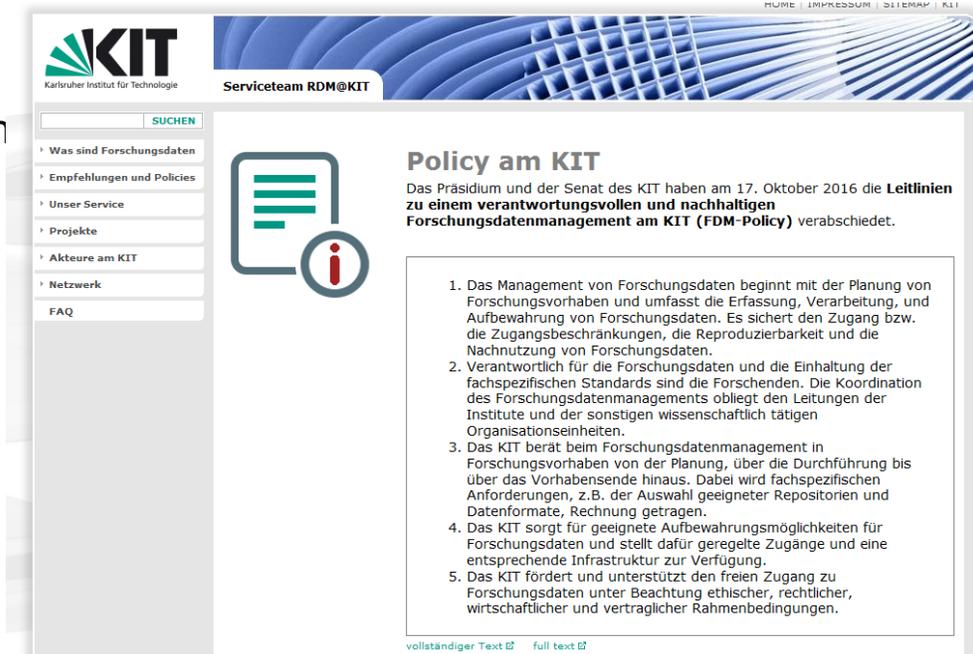
Dank bwDataDiss → nun Forschungsdaten in KITopen

■ Eckdaten:

- 2 Ziele der KIT-Bibliothek: Forschungsdaten publizieren & archivieren
- Generisches Repository: Alle Dateitypen und alle Fächer
- Forschungsdaten sind eigener Publikationstyp mit Verknüpfung zur Publikation
- Step by Step: Zunächst ist der Forschungsdatenworkflow mit der Pflichtabgabe für Doktoranden verbunden

■ Neuer Service ist erst am Entstehen...

- Umgang und Annahme von Dateien (Datei mit „finalelem Charakter“)
- Beratung im Bereich der Langzeitarchivierung von Forschungsdaten
- Marketing für neuen Forschungsdienst im KIT



KIT
Karlsruher Institut für Technologie

Serviceteam RDM@KIT

SUCHEN

- Was sind Forschungsdaten
- Empfehlungen und Policies
- Unser Service
- Projekte
- Akteure am KIT
- Netzwerk
- FAQ

Policy am KIT

Das Präsidium und der Senat des KIT haben am 17. Oktober 2016 die **Leitlinien zu einem verantwortungsvollen und nachhaltigen Forschungsdatenmanagement am KIT (FDM-Policy)** verabschiedet.

1. Das Management von Forschungsdaten beginnt mit der Planung von Forschungsvorhaben und umfasst die Erfassung, Verarbeitung, und Aufbewahrung von Forschungsdaten. Es sichert den Zugang bzw. die Zugangsbeschränkungen, die Reproduzierbarkeit und die Nachnutzung von Forschungsdaten.
2. Verantwortlich für die Forschungsdaten und die Einhaltung der fachspezifischen Standards sind die Forschenden. Die Koordination des Forschungsdatenmanagements obliegt den Leitungen der Institute und der sonstigen wissenschaftlich tätigen Organisationseinheiten.
3. Das KIT berät beim Forschungsdatenmanagement in Forschungsvorhaben von der Planung, über die Durchführung bis über das Vorhabensende hinaus. Dabei wird fachspezifischen Anforderungen, z.B. der Auswahl geeigneter Repositorien und Datenformate, Rechnung getragen.
4. Das KIT sorgt für geeignete Aufbewahrungsmöglichkeiten für Forschungsdaten und stellt dafür geregelte Zugänge und eine entsprechende Infrastruktur zur Verfügung.
5. Das KIT fördert und unterstützt den freien Zugang zu Forschungsdaten unter Beachtung ethischer, rechtlicher, wirtschaftlicher und vertraglicher Rahmenbedingungen.

[vollständiger Text](#) [full text](#)



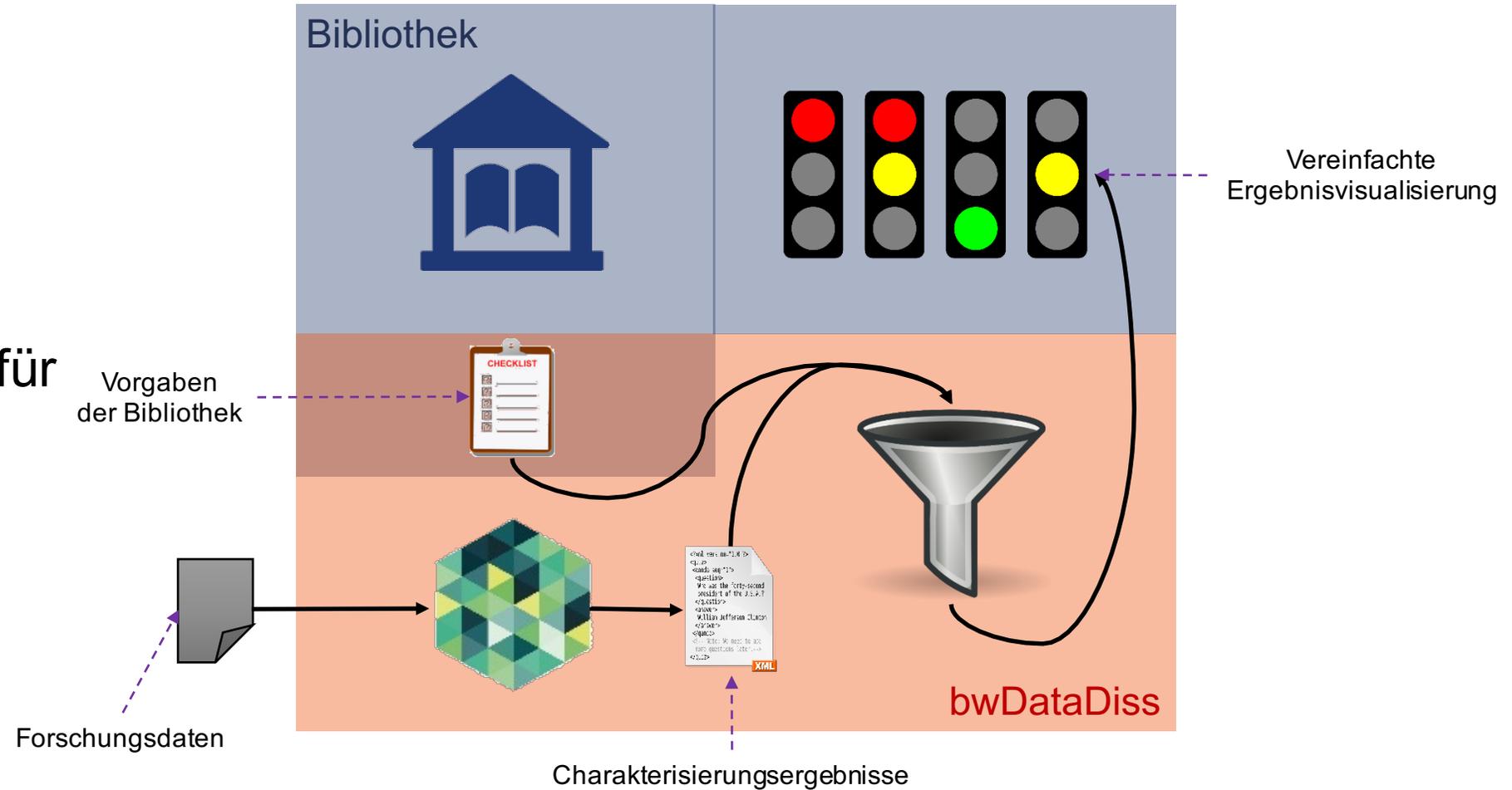
Detailaspekte

Charakterisierung, Authentifizierung und Metadaten



Charakterisierung - Prinzip

- Unterstützung für Bibliotheken bei der Qualitätssicherung
- Grundlage bzw. Entscheidungshilfe für die Erstellung von Erhaltungsstrategien





Charakterisierung – Policies (1)

- Policies gehören immer zu einer Einrichtung und können nur von dieser verändert werden
 - Alle Policies stehen aber immer allen Einrichtungen zur Nutzung offen und können geklont und dann bearbeitet werden
- ← Policy Editor bietet eine grafische Oberfläche die eine einfache Bearbeitung und Erstellung von Policies ermöglicht

The screenshot shows the 'Charakterisierungspolicy bearbeiten' interface. It includes a description field, a dropdown for 'Wert bzw. Farbe für nicht konfigurierte PRONOM Typen' (set to 'rot'), and a list of PRONOM types. A legend allows setting colors: 'Set as good (green)', 'Set as mediocre (yellow)', 'Set as bad (red)', and 'Reset value(s)'. The list shows various file formats like Acrobat PDF/A and Adobe Flash, with some highlighted in green, yellow, or red. A 'Text der Charakterisierungspolicy' section at the bottom lists entries like 'fmt/969 GREEN'.



Charakterisierung – Policies (2)

- "Basis-Policy" zur Bewertung von Dateiformaten
 - 3 (bzw. 4) Klassen (Ampel-Modell)
 - "grün", alle Formate, die mit sehr großer Wahrscheinlichkeit in Zukunft interpretiert werden können
 - ohne spezifische Software, wie "plain-text" Formate (ASCII, UTF-8)
 - Formate, die explizit zur Archivierung entwickelt wurden (zBsp. PDF/A)
 - "hellgrün" (→ "grün"), alle anderen binären Formate, die mit sehr großer Wahrscheinlichkeit in Zukunft interpretiert werden können
 - Z.Bsp. Formate mit weiter Verbreitung, (offener) Dokumentation und quelloffener (Referenz-)software (PNG, GIF, etc.)
 - "gelb", alle Formate, die wahrscheinlich noch in Zukunft interpretiert werden können
 - Formate, für die es eine Migrationsstrategie gibt
 - Formate, deren Software ebenfalls archiviert wird und langfristig z.B. mittels Emulation nutzbar bleibt
 - Restrisiko
 - "rot", alle anderen Formate

Charakterisierung – Ergebnisse

fmt/18 (Acrobat PDF 1.4 - Portable Document Format)	Ø Beschreibung des Problems und der Algorithmen.pdf	143.1 KiB	10.01.17 14:48:25	Details
fmt/17 (Acrobat PDF 1.3 - Portable Document Format)	Ø FüßeSchreibtisch1.pdf	418.89 KiB	10.01.17 14:48:25	Details
fmt/531 (AutoCAD Drawing)	Ø FüßeSchreibtisch1.bak	401.58 KiB	10.01.17 14:48:25	Details
fmt/18 (Acrobat PDF 1.4 - Portable Document Format)	Ø html5boilerplate-pdf.pdf	228.23 KiB	10.01.17 14:48:27	Details
x-fmt/263 (ZIP Format)	Ø classicthesis-pkrefined.zip	143.91 KiB	10.01.17 14:48:27	Details
fmt/20 (Acrobat PDF 1.6 - Portable Document Format)	Ø haushaltsplan_bw_2015_16.pdf	45.99 MiB	16.01.17 09:24:28	Details

Charakterisierungsergebnisse

Der Datensatz wurde charakterisiert. Nachfolgend die Ergebnisse.



```

[
  {
    "value": "GREEN",
    "type": "fmt/279",
    "count": 1,
    "files": [
      ".\BIS1447-002-flac_24.flac"
    ]
  },
  {
    "value": "GREEN",
    "type": "fmt/20",
    "count": 1,
    "files": [
      ".\haushaltsplan_bw_2015_16.pdf"
    ]
  },
  {
    "value": "RED",
    "type": "x-fmt/263",
    "count": 2,
    "files": [

```

- Ergebnisse der Charakterisierung
 - werden für jede Datei farblich hervorgehoben
 - und für den Gesamtdatensatz als Ampel dargestellt
 - können auch über die bwDataDiss API abgerufen werden

Authentifizierung

■ bwIDM

- Bei bwDataDiss können sich Nutzer per bwIDM anmelden
- Bei Integration z.B. des Uploaders auf der Bibliothekswebseite, kümmert sich bwDataDiss um die bwIDM Anmeldung

■ bwDataDiss Kontennutzung

- Die Bibliothek kann Nutzernamen + Passwort Kombinationen ihrer zugeordneten Nutzer überprüfen und dies als Authentifizierung nutzen

■ Einladungskomponente von bwDataDiss

- Nutzer können von der Bibliothek zu bwDataDiss eingeladen werden
- Eingeladene Nutzer können sich dann – wie oben beschrieben – authentifizieren



Authentifizierung - Benutzereinladung

Bibliothek: Benutzer zu bwDataDiss einladen

Mit diesem Formular können Sie Benutzer zu bwDataDiss einladen. Die Funktion ist für Benutzer vorgesehen, die sich anderfalls (über bwIDM) nicht anmelden können, weil z.B. der benötigte Universitätsaccount deaktiviert wurde. Sollte der Benutzer die Einladung annehmen, wird er ihrer Bibliothek zugeordnet.

* Vorname: Max
Vorname des

* Nachname: Mustermann
Nachname des

* E-Mail Adresse des Empfängers: Empfängerad...
E-Mail Adress

Gültigkeit der Einladung: Bis wann soll die Einladung gültig sein?

* Betreff: Einladung zu bwDataDiss
Der Betreff der E-Mail Nachricht an den einzuladenden Benutzer

* Nachricht an den Einzeladenden: <DOCTYPE HTML> <html> <head> <meta http-equiv="Content-Type" content="text/html; charset=UTF-8" /> </head> <body> <h3>Accounteinladung für bwDataDiss</h3>
 Hallo {{ firstName }} {{ surName }}

 Sie wurden von Admin User (SYSTEM) zu bwDataDiss eingeladen.

 Um ihren Account zu aktivieren, klicken Sie auf nachfolgenden Link
 Einladung annehmen

 </body> </html>

Absenderadresse: admin@bwdd.kit.edu
Ihre Absenderadresse

Einladung abschicken

Zur Liste der Benutzereinladungen

Hinweise zur Gültigkeit der Einladung: Sie können eine Gültigkeitsbegrenzung für die Einladung hinterlegen. Ist diese überschritten wird die Einladung ungültig und kann nicht mehr akzeptiert werden.

- Können Nutzer bwIDM nicht mehr nutzen, können neue Accounts durch Einladung zu bwDataDiss durch die Bibliothek angelegt werden
 - Benutzereinladung ist direkt über das Portal möglich (als Bibliotheksmitarbeiter)

Metadaten

- **Klassifizierung:**
 - Neben DFG (nach Fachsystematik) und DDC ist es nun auch möglich, beliebige andere Klassifizierungen anzugeben
 - Für DDC und DFG bietet das bwDD Portal eine Auswahlhilfe
- **MARC21 Format**
 - In Zusammenarbeit mit UBFR wird bwDataDiss Metadatenschema in MARC21 abgebildet
 - bwDataDiss unterstützt MARC21 per OAI-PMH
- **Metadaten können auch über die API „gepusht“ werden**
- **DublinCore kann auch (per OAI-PMH) angeboten werden**
 - Entsprechende Erweiterung nötig
 - Mapping kann in Abstimmung mit zukünftigem Erstnutzer festgelegt werden

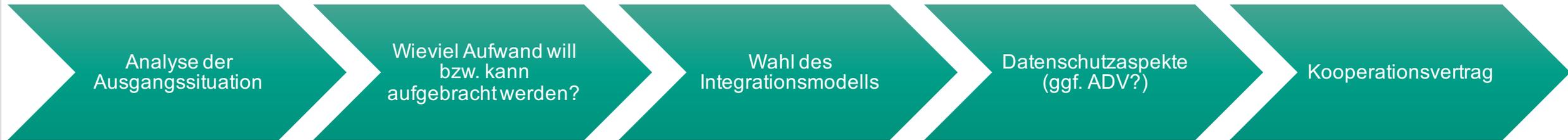


Überlegungen zur Nutzung von bwDataDiss

Schritte zum neuen Forschungsdienst

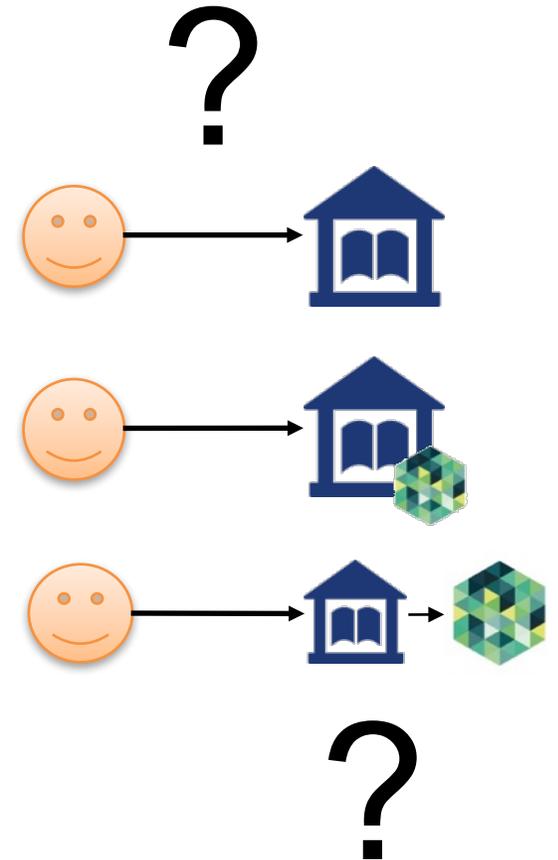


BWDATADISS



Entscheidungshilfe - Integrationsmodelle

- Unterstützt ihr Repository die Erfassung von Metadaten zu Forschungsdaten o. ist dies demnächst geplant?
 - Ja → Modell 1 o. 2 geeignet
 - Nein → Modell 3
- Können Sie als Bibliothek ggf. auch Speicherplatz in größerem Umfang zur (Zwischen-)Speicherung von Daten bereitstellen?
 - Ja → Modell 1 möglich
 - Nein → Modell 2 o. 3 eher geeignet





Zukünftige Erweiterungsmöglichkeiten

Erweiterungsmöglichkeiten: wie geht's weiter?

- Struktur der Dateiabgabe:
 - Umgang mit ZIP-Archiven
 - Umgang mit hierarchisch organisierten Dateien (in Verzeichnissen)
- Öffnung gegenüber Spezialumgebungen von Forschungsdaten
 - Statt Upload kann Datenübertragungen zwischen den Systemen erfolgen



BWDATADISS

Besten Dank

Bei Interesse und Fragen wenden Sie sich gerne an uns!

regine.tobias@kit.edu & tobias.kurze@kit.edu