



Daten- und AI-Governance mit dem Datenmanagementportal BKA

Manfred Gruber, Bundeskanzleramt

Alessandro Volcich, Bundesrechenzentrum

Omar Ismail, Bundesrechenzentrum

Einleitung

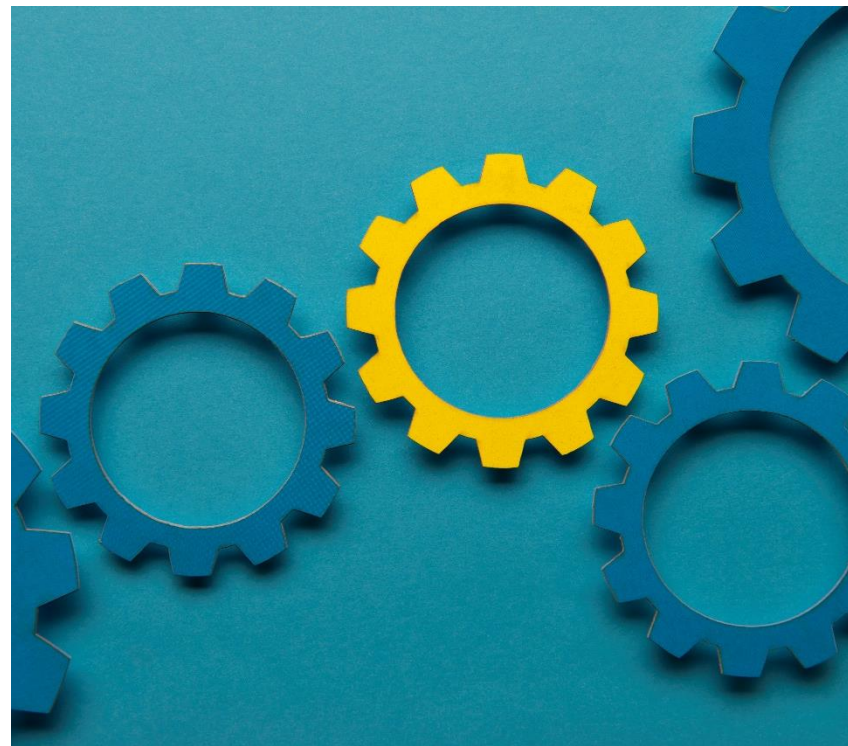
Manfred Gruber

Pflicht

- Public Sector Information Richtlinie (PSI)
- Data Governance Act
- AI Act
- Data Act
- DSGVO
- ...

Möglichkeiten

- Schaffen von Überblick
- Vernetzen der Daten
- Erhöhen oder ermöglichen von Wertschöpfung
- Grundlage für neue Projekte (AI, Data Science, ...)
- ...





Beispiel Erzeugung des Digitalen Gedächtnisses

Österreichisches Staatsarchiv

350 km Regalbestände
davon

1,5 Millionen Karten (mehr
als 60 Petabyte an Daten)

Bundesdenkmalamt

1,5 Millionen Bilder (Dias,
Fotos, Glasplatten, born-
digital)

60.000 Pläne

Bis 2025 **2,1 Petabyte**
Daten

Datenmanagementportal BKA

Service Stack

Alessandro Volcich

Orchestrierung

Prozesslandkarte

Technische Workflows
Camunda

Entscheidungs-Workflows
ELAK

Service- und Tool
Management

Prozesse

Datenerzeugung

- Ingest
- Automatisieren
- Metadaten erzeugen
- Qualität sichern

Daten(er)haltung

- Containerverwaltung
- Aktualität sicherstellen
- Integrität sicherstellen
- Erhalten und Archivieren

Datennutzung

- Daten freigeben
- Daten bereitstellen
- Daten präsentieren
- Daten nutzen

open Enterprise Data Hub
(openEDH)

Core
Kernfunktionalitäten, Data Layer Management, Daten Container Management

Security
Portal Austria, Standardportal

Services

Data Integrity & Documentation
Daten-Beweisketten, Integrität, Dokumentation

Persistent Identifier
urn:gvat:#####, Handle System, Weltweit eindeutige ID, delegierbar

Datenerzeugung

Goobi Workflow
Automatisierung Digitalisierungsprozessen, Analog/Digital, Digital/Digital

Fachdaten Ingest
Daten aus Fachanwendungen (z.B. Log-Daten)

BKA AI Hub

BRZ AI Hub
Modellerstellung, Inferenz-Service

AI Services
Gesichtsdetektion, Gebäudeklassifizierung, Bildoptimierung

Übersicht

Datenkatalog
Bestandsübersicht, Abruf

Data Governance Monitor
Technische Übersicht über die Daten und Prozesse

Schnittstellen

GAIA-X
Eclipse Data Connector, Europeana Schnittstellen

data.gv.at
Schnittstelle zu data.gv.at

Erhaltung

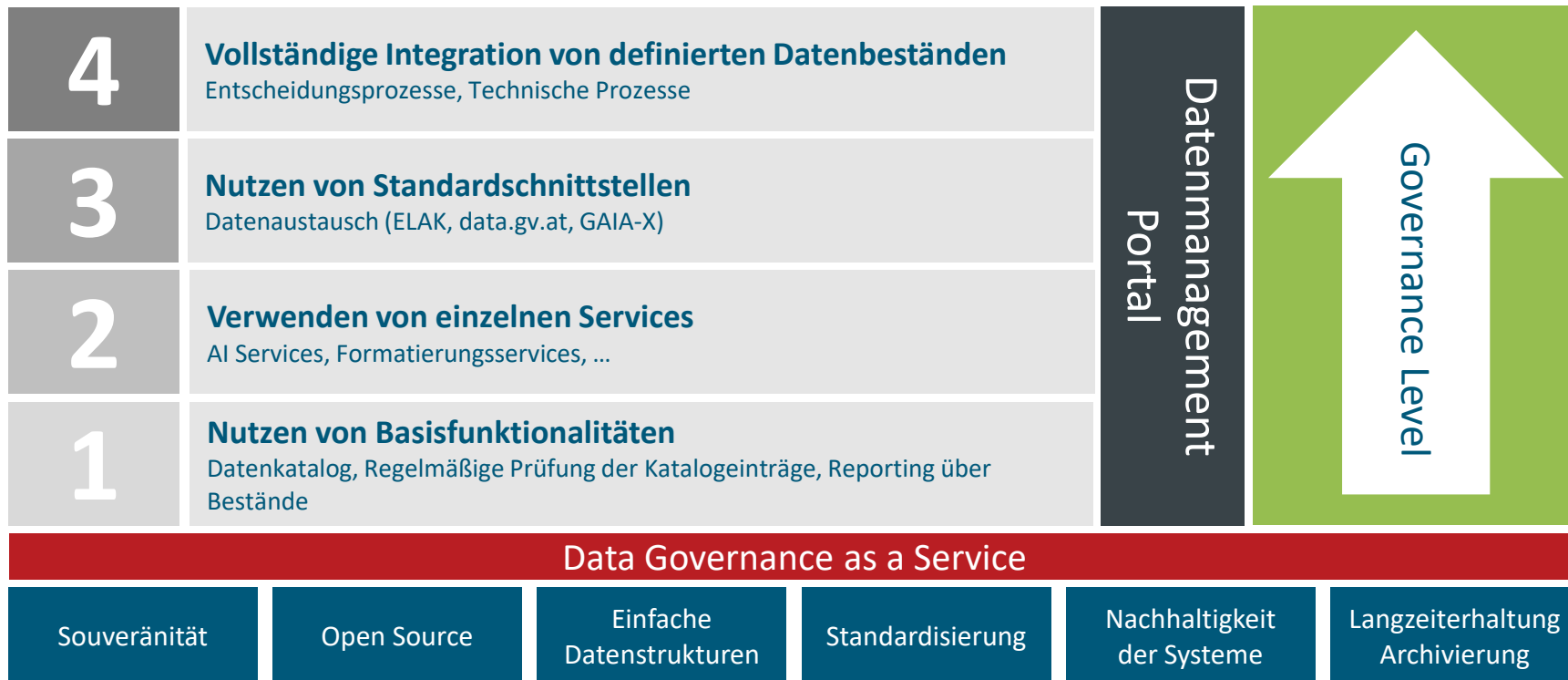
BALES
Bundesarchiv- und Langzeiterhaltungssystem, Digitales Archiv

Digitale Arche
Daten-Resilienz im Bunker

Visualisierung

Goobi Viewer
Visualisierung digitalisierte Bestände im optimierten Design

Beyond Arts App
Vermittlung der Bundesdenkmalamt-Bestände



Der Weg eines Datenpakets im Datenmanagementportal

Onboarding und Katalogisierung



Datenkatalog

Datensätze Organisationen Gruppen Über uns Suche

Datensätze

Organisationen
Bundeskanzleramt 45

Gruppen
Für diese Suche wurden keine Gruppen gefunden.

Tags
OSIA 27
Bundeskanzleramt 1
Bundesdenkmalamt 1
Qualitätssicherung 1
Bundeskanzler 1
Bundesminister 1
Bücher 1
Denkmale 1
Fotografien 1
Mehr Tags anzeigen

45 Datensätze gefunden Sortieren nach: Relevanz

Organisationen: Bundeskanzleramt

BDA_Archivologiedaten
Dieser Container enthält Aufzeichnungen zu den archaischen Ausgrabungsstätten in Celerech seit 2000(?)

BKA_Fotoarchiv_Backup
Dieser Container enthält Sicherungskopien der Fotodateien von Veranstaltungen mit Bundesministern und Bundeskanzleramt aus dem Foto- und Videoarchiv des Bundeskanzleramts

BKA_Videoarchiv_Backup
Dieser Container enthält Sicherungskopien der Videodateien von Veranstaltungen mit Bundesministern und Bundeskanzleramt aus dem Foto- und Videoarchiv des Bundeskanzleramts

BKA_Fotoarchiv
Dieser Container enthält Fotodateien von Veranstaltungen mit Bundesministern und Bundeskanzleramt aus dem Foto- und Videoarchiv des Bundeskanzleramts

Organisationen / Bundeskanzleramt / BKA_Fotoarchiv

BKA_Fotoarchiv

Follower: 0

Folgen

Organisation

Bundeskanzleramt

Bundeskanzleramt
Es gibt keine Beschreibung für diese Organisation

Sozial
Twitter

Daten und Ressourcen

BKA_Fotoarchiv **Entdecken**

Bundeskanzler Bundeskanzleramt Bundesminister Fotografien

Zusätzliche Informationen

Feld	Wert
Letzte Änderung der Metadaten	23. Februar 2023
Datenverantwortliche*	arthur.werz@bka.gv.at
Datenverantwortliche* Email	arthur.werz@bka.gv.at
Datenverantwortliche Organisation	Bundeskanzleramt
Technischer Kontakt (Ansprechperson)	
Technischer Kontakt (Organisation)	
Technischer Kontakt (Email)	oedh@brz.gv.at
Digitale Arche?	Ja
Feldbeschreibungen bei strukturierten Daten	
Herkunft der Daten/Datenlieferant	Bundeskanzleramt - Foto- und Videoarchiv
Zweck und rechtliche Grundlage	Erhaltungsmaßnahmen
IT-Verfahren	openEDH
Beschreibung zum IT-Verfahren	
Ist Open-Data? Zugriffsmöglichkeiten auf Nutzdaten	Ja
Nutzungs- und Abrufmöglichkeiten	
Kategorien	Verwaltung und Politik

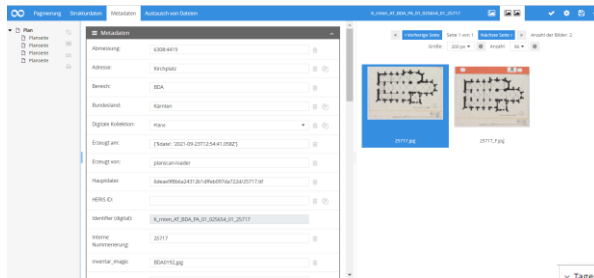
Onboarding und Katalogisierung

- Katalogisierung im Datenkatalog
 - Angaben zu Dateninhaber:innen und fachliche Metadaten
 - Portalverbund-kompatibles Berechtigungskonzept
 - Unterstützt DCAT-AP Metadatenschema zur Beschreibung der Datensätze
 - Schnittstelle zu data.gv.at

Der Weg eines Datenpakets im Datenmanagementportal

Onboarding und
Katalogisierung

Datenerzeugung und
Anreicherung



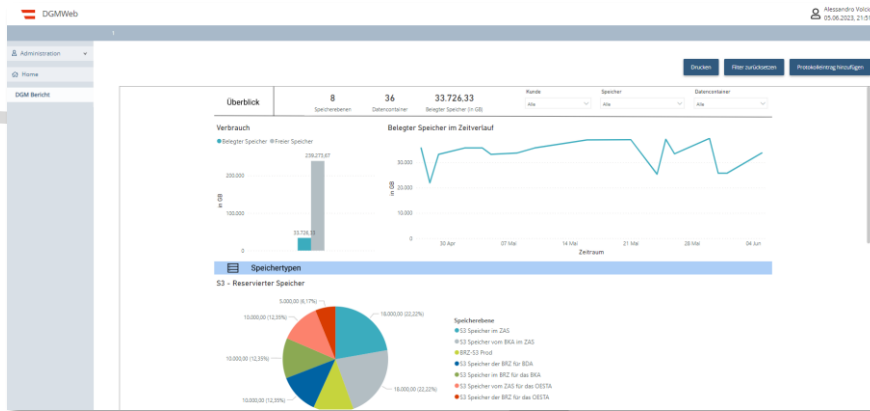
Tagesproduktion

Nr.	Titel	Status
1	Job anlegen	■
2	Zuweisung eines Fotografen	■
3	Fotografieren und Bilder einspielen	■
4	Qualitätskontrolle	■
5	Erzeugung von Derivaten	■
6	Überprüfung der Metadaten	■
7	Generierung von Signaturen	■
8	Veröffentlichung im Goobi Viewer	■

Datenerzeugung und Anreicherung

- Einfache Steuerung von Digitalisierungsprojekten
 - Born digital
 - Analoges Material
- Wiederverwendbare Aktivitätsschritte in modularen Workflows
- Integration von Third Party Services in den Workflows
- Offene Standards: METS, MODS

Der Weg eines Datenpakets im Datenmanagementportal



Einspeisen des Datenbestands



Einspeisen des Datenbestands

- Storage Management mit openEDH
 - S3 Cloud Storage der BRZ
 - Filebasierter Storage
 - Datenbestand bleibt physisch in der dateninhabenden Organisation
- Eliminierung historisch gewachsener Fileablagen und von Shadow Data
- Monitoring und Tracking des Speicherverbrauchs in der Reporting Platform

Der Weg eines Datenpakets im Datenmanagementportal

Date	Type	Detail	Outcome
2023-05-16 10:10:00	format identification	Identified the object's file formats and versions using Sigfig.	Success
2023-05-12 13:15:09	ingest end	The ingest process has ended.	Success
2023-05-12 13:15:09	accession	Added package to the inventory. After this point, the responsibility for the digital content's preservation is passed on to the repository.	Success
2023-05-12 13:15:09	authorization check	User permissions have been checked to ensure that he has sufficient authorization to store the SIP under the assigned node of the classification scheme.	Success
2023-05-12 13:15:09	format identification	Identified the object's file formats and versions using Sigfig.	Success
2023-05-12 13:15:08	message digest calculation	Created base PREMIS objects with file original name and file fluxy information (SHA-256).	Success
2023-05-12 13:15:08	wellformedness check	Checked whether the descriptive metadata is included in the SIP and if this metadata is valid according to the established policy.	Success
2023-05-12 13:15:08	virus check	Scanned package for malicious programs using ClamAV.	Success
2023-05-12 13:15:08	wellformedness check	Checked that the received SIP is well formed, complete and that no unexpected files were included.	Success
2023-05-12 13:15:08	unpacking	Extracted objects from package in E-ARK SIP 2 format.	Success
2023-05-12 13:15:07	ingest start	The ingest process has started.	Success

Einspeisen des Datenbestands

Digitale Archivierung

Information package processing report

Entity: `BRZ/70764-6c7c4318-9550-e95005wellfdoc_obj_3`

Source: `transfer/receive`

Outcome: `individual entity`

Run tasks:

- E-ARK SIP 2
- Hydrex detector

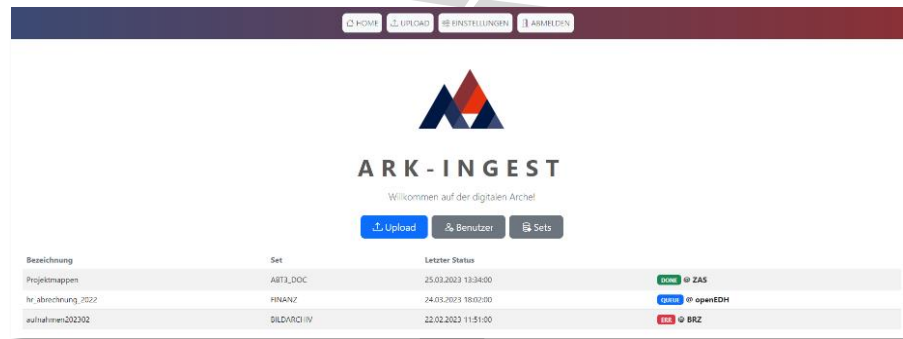
Digitale Archivierung

- Hoher Grad an Automatisierung durch Verwendung von Data Pipelines
- Verwendung offener Standards:
 - E-ARK2, EAD
 - BALES für ÖStA
 - Archivierung > 1 Mio. ELAK-Akten pro Jahr
 - Edidoc, Heterogene Datenformate

Der Weg eines Datenpakets im Datenmanagementportal

Onboarding und
Katalogisierung

Datenerzeugung und
Anreicherung



The screenshot shows the ARK-INGEST web interface. At the top, there is a navigation bar with links for HOME, UPLOAD, EINSTELLUNGEN, and ABMELDEN. Below the navigation bar is the ARK-INGEST logo and the text 'Willkommen auf der digitalen Arche'. There are three buttons: Upload, Benutzer, and Sets. Below this is a table with the following data:

Bezeichnung	Set	Letzter Status	
Projektmappen	ARIT_DOC	25.03.2023 13:34:00	Done @ ZAS
hr_abrechnung_2022	FINANZ	24.03.2023 18:02:00	Done @ openEDH
aufnahmen202202	DILDARO IV	22.02.2023 11:51:00	Done @ BRZ

Digitale Archivierung

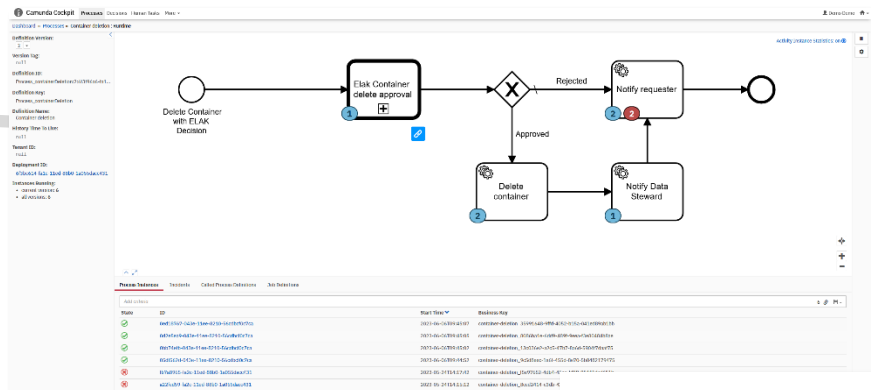
Sicherung in Digitaler Arche



Digitale Arche

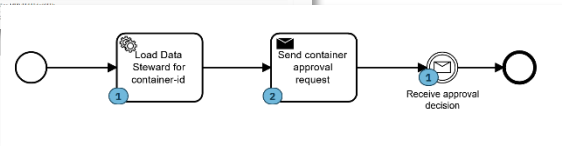
- Krisensicherer Datenspeicher im Zentralen Ausweichsystem
- Kontrolliertes Einspeisen von Datenpaketen über Schnittstellen oder mit Ark-Ingest
- Sicherer Datenzugriff im Katastrophenfall und bei Ausfall der operativen Systeme

Der Weg eines Datenpakets im Datenmanagementportal



Einspeisen des Datenbestands

Digitale Archivierung



Sicherung in Digitaler Arche

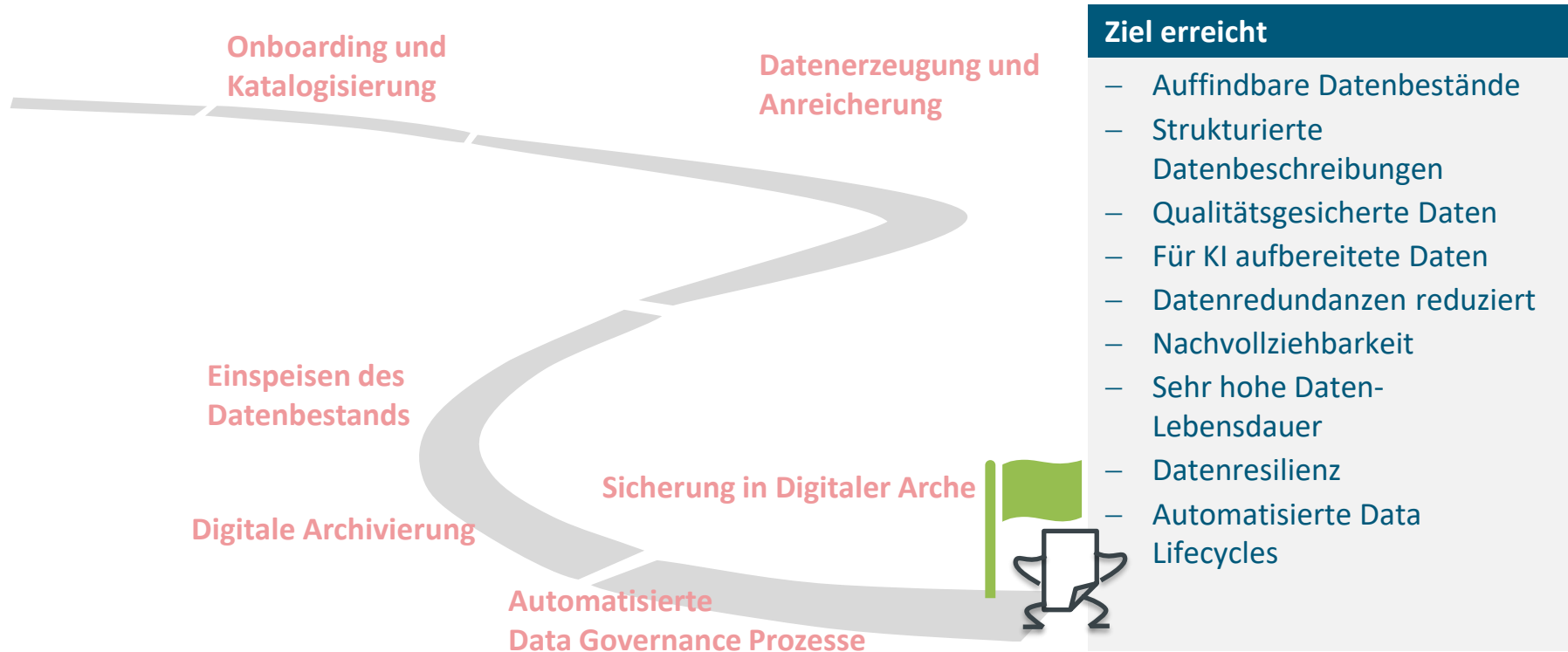
Automatisierte Data Governance Prozesse



Automatisierte Data Governance

- jederzeit nachvollziehbare Datenoperationen
- Einwilligungsmanagement mit ELAK
 - Datenanfragen
 - Löschroutinen
 - Aktualisierung der Dateneigner
- Automatisierte Bereitstellungsmechanismen

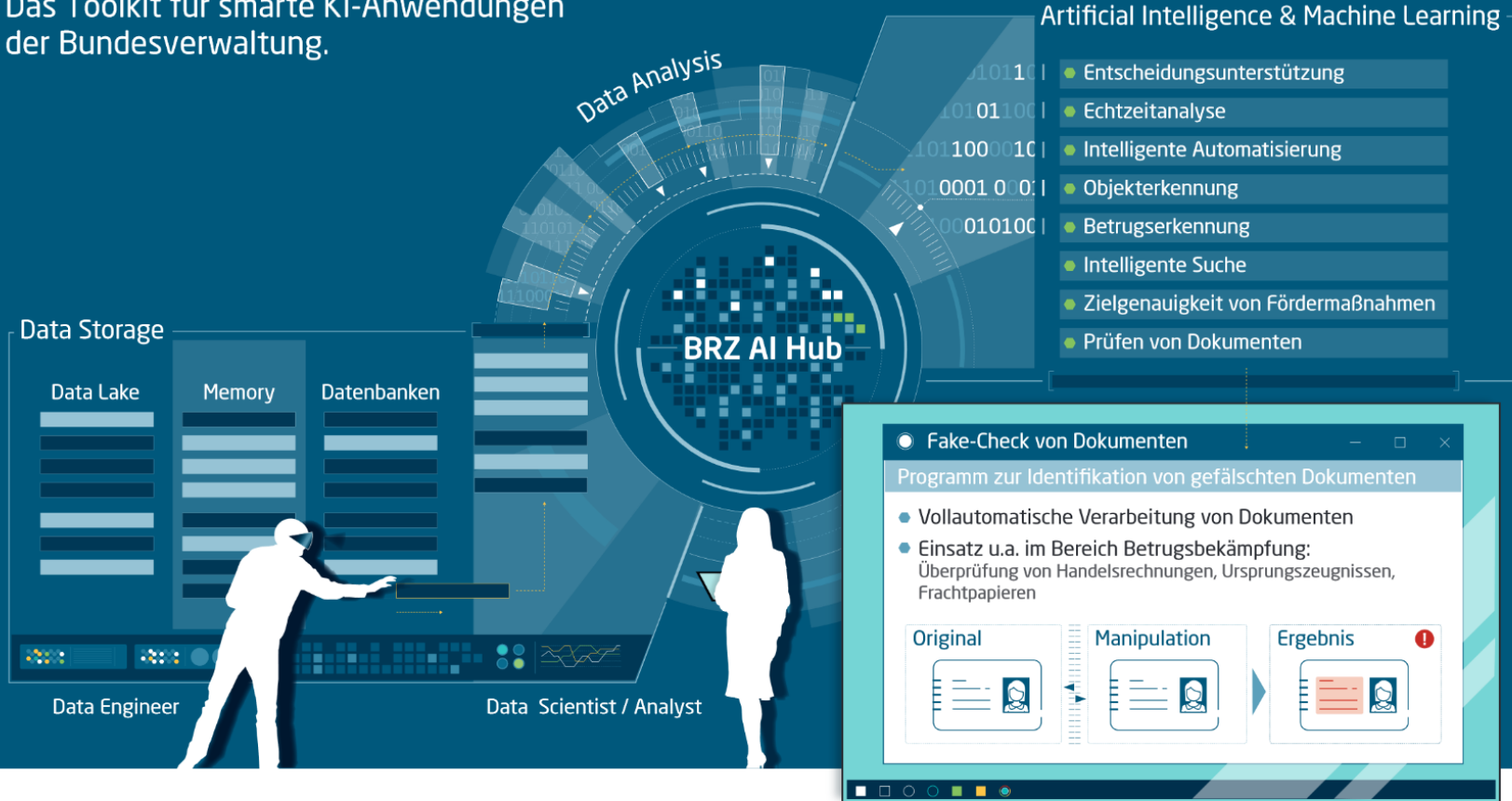
Der Weg eines Datenpakets im Datenmanagementportal



Datennutzung mit der Forschungsplattform

Omar Ismail

Das Toolkit für smarte KI-Anwendungen
der Bundesverwaltung.



Trustworthy AI – Wo fängt es an?

Treffer pro Seite: 25 Sortierung: Zufällig

Dokumenttyp
Foto (31147)
Kleinbildla (143)

Kath. Wallfahrtskirche hl. Maria, Maria im Mösl in Niederarnsdorf

Ruine Thürndl

Volksgarten

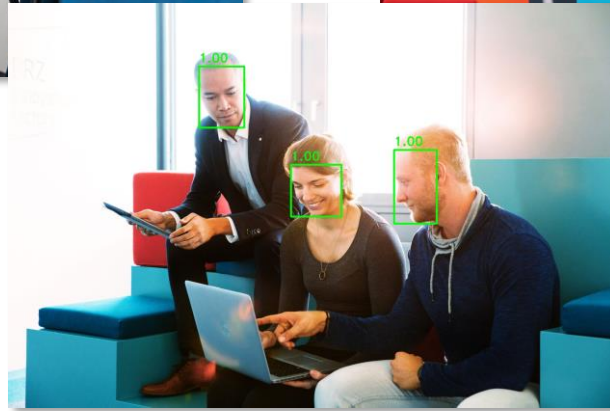
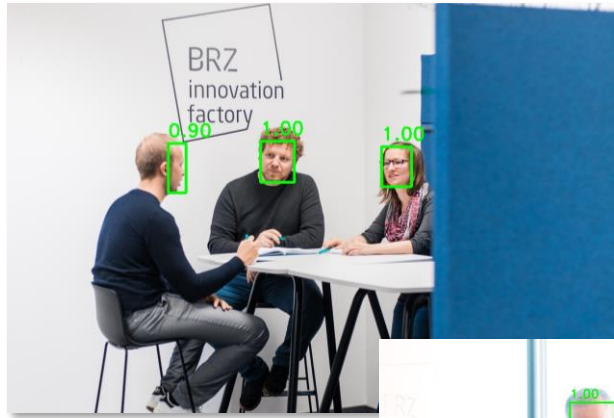
Palais Schönburg, ehemals Gartenpalais Starhemberg

Schloss Urstein

Stephansdom, Dom- und Metropolitankirche St. Stephan

Secession

Ehem. Stadtpalais des Prinzen Eugen, Bundesministerium für Finanzen





- **Ziel:** KI-Service zur Gebäudeerkennung und -klassifizierung von Bildern in Gebäudekategorien.
- **Datensatz:** Vom Bundesdenkmalamt (BDA) zur Verfügung gestellte Fotodatenbank IMAGO mit einer hohen Anzahl an Gebäudefotographien und dazugehörigen Metadaten.
In den Metadaten vorhandene Kategorien sind in drei Stufen eingeteilt: Grob, Mittel und Fein.
- **Ergebnisse:**
 - Genauigkeit von 69,41%/90,83% (mittel) bzw. 62,62%/83,27% (fein)
 - Verarbeitung von bis zu 11.520 Bildern möglich (bei Bildgrößen von 1-3 MB).



03.07.2023

10156987181Neubauer_Pregl_Bettina2013

Metadaten Austausch von Dateien

Metadaten

BDA-Bereich: Fotoarchiv

Bundesland: Wien

Digitale Kollektion: Fotoaufträge des BDA zu Denkmaldokumentationen seit de

DMD-ID: 101569

Elternelement: 3599

Entstehungsjahr:

Erhaltungszustand des Objekts:

Fotografin:

Gemeinde (Hauptadresse): Wien 19 - Döbling #MIG-#H

Grundstücksnummer: 371, 366/1, 401/1

Hauptkategorie fein:

Hauptkategorie grob:

Hauptkategorie mittel:

HERIS ID: 87181

Identifizier (digital): 10156987181Neubauer_Pregl_Bettina2013

Katastralgemeinde (KG): 01511 Salmannsdorf

Objekttitle: Höhenstraße

< Vorherige Seite Seite 1 von 1 Nächste Seite > Anzahl der Bilder: 4

Größe: 200 px Anzahl: 96

bda_Hoehenstr_033.jpg

bda_Hoehenstr_034.jpg

bda_Hoehenstr_035.jpg

bda_Hoehenstr_036.jpg

AI Hub – Gebäudererkennung

Stift Melk, Melk



Hauptkategorie Mittel

- Klöster/Stifte: 0.85
- Kirchen: 0.04
- Verwaltungsbauten: 0.03

Hauptkategorie Fein

- BenediktinerInnenkloster: 0.63
- Pest-/Dreifaltigkeitssäule: 0.15
- Postgebäude: 0.03

AI Hub – Gebäudererkennung

Hanne-Sobek-Sportanlage, Berlin



Hauptkategorie Mittel

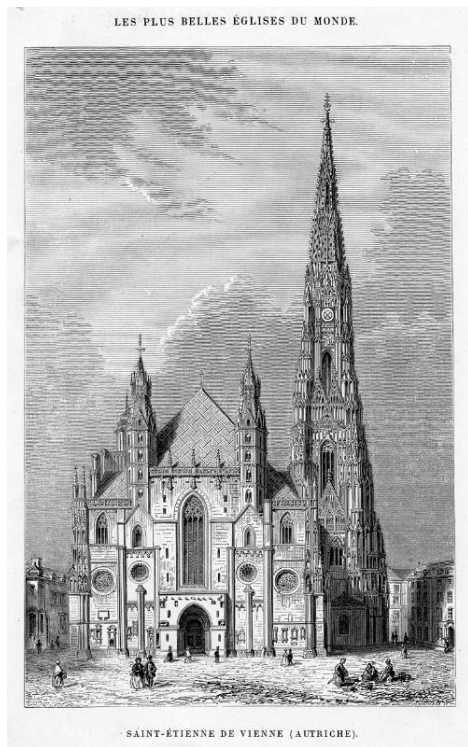
- Friedhöfe/Begräbnisplätze: 0.77
- Gartenbaudenkmäler: 0.11
- Erinnerungs/Kleindenkmäler: 0.03

Hauptkategorie Fein

- Gartenbaudenkmale: 0.41
- Sonstige Kleindenkmale: 0.12
- Badeanlage: 0.03

AI Hub – Gebäudererkennung

Stephansdom, Wien (Holzschnitt)



© Scan von Dmitry Makeev; CC-BY-SA 4.0

Hauptkategorie Mittel

- Kirchen: 0.75
- Karner: 0.02
- Klöster/Stifte: 0.02

Hauptkategorie Fein

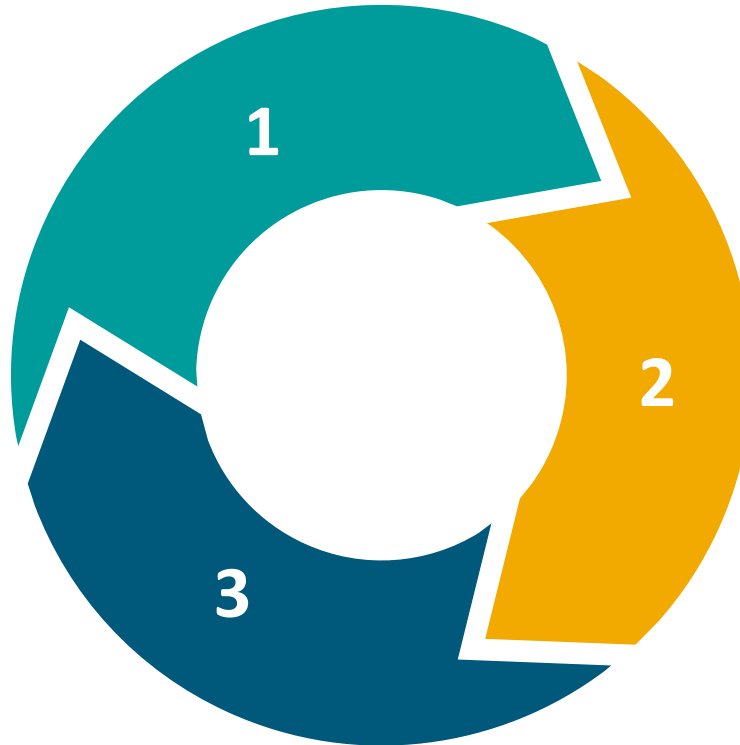
- Kloster-/Stiftskirche: 0.58
- Wallfahrtskirche: 0.20
- Katholische Kirche: 0.05

Datenerstellung

- Daten sammeln
- Daten definieren

Datennutzung

- Daten wählen
- Daten verwenden
- Neue Daten erstellen



Datenmanagement

- Daten verwalten
- Daten anreichern
- Daten bereinigen

Danke für Ihre Aufmerksamkeit.



- **Manfred Gruber**
Bundeskanzleramt
manfred.gruber@bka.gv.at
- **Alessandro Volcich**
Bundesrechenzentrum
alessandro.volcich@brz.gv.at
- **Omar Ismail**
Bundesrechenzentrum
omar.ismail@brz.gv.at

