

# IANUS: Archivierung von digitalen Forschungs- daten in den Altertumswissenschaften

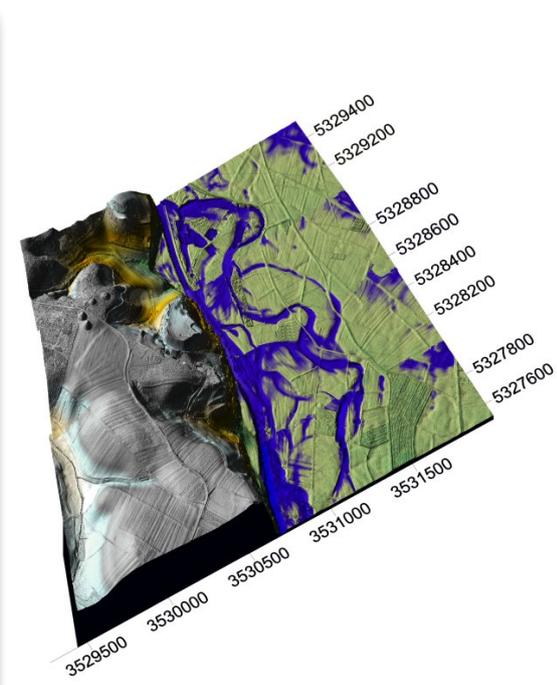
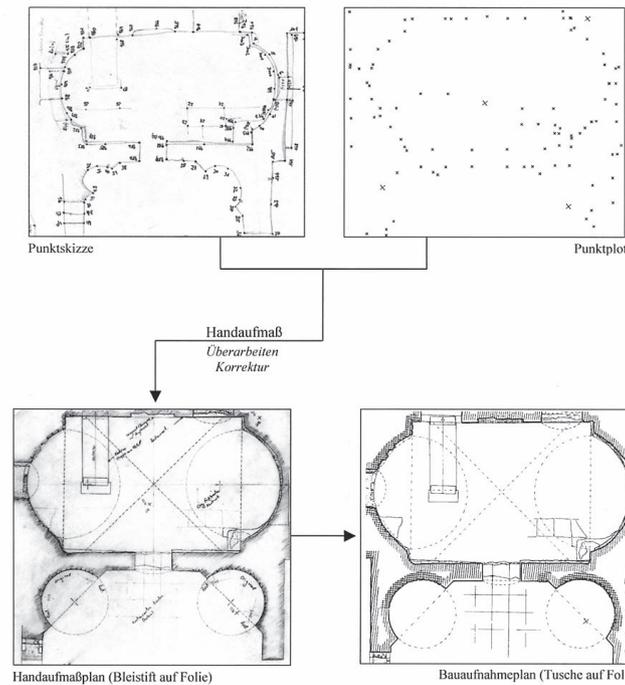
---

Reithaus, Weimar

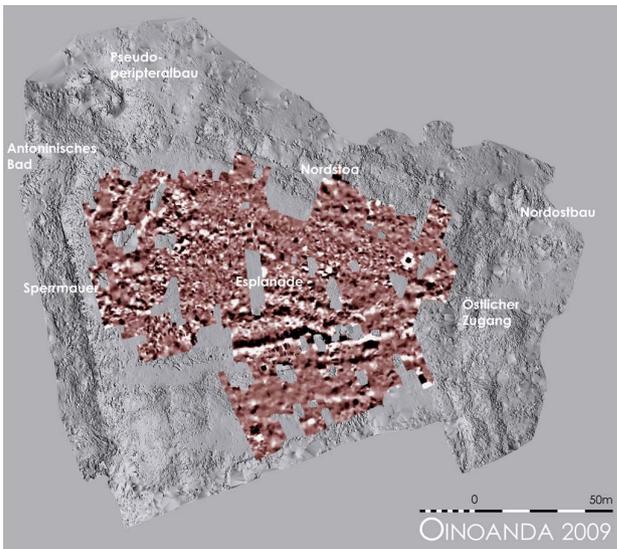
11. März 2014

Die Archäologische und altertumswissenschaftliche Forschung ist geprägt von einer wachsenden Menge komplexer digitaler Daten, die

- von **(Einzel-)Disziplinen** wie Archäologie, Philologie, Alte Geschichte, Anthropologie, Archäometrie, etc. erzeugt werden
- auf **unterschiedlichen Methoden** wie Ausgrabungen, Surveys, Bau- forschung, Paläogenetik, Knochenanalyse, Materialuntersuchungen, Fernerkundung, Geodäsie, Textanalyse etc. basieren



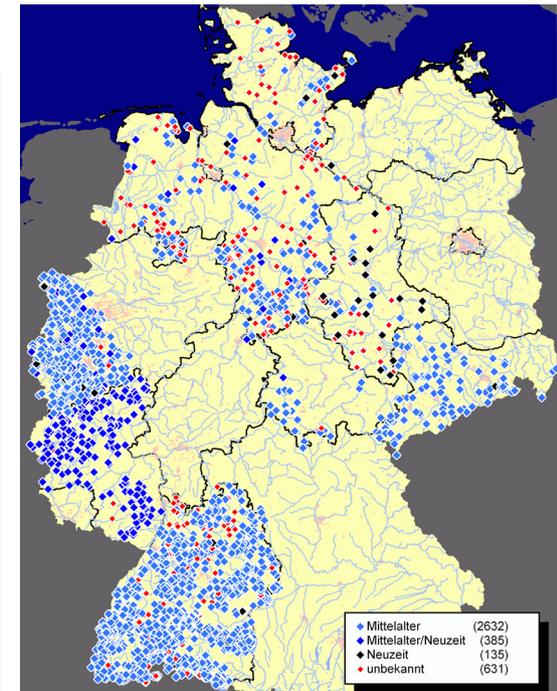
- **komplexe Fragestellungen** zu anspruchsvollen Konzepten (z. B. Raum, Erinnerung, Gedächtnis, Landschaft, Macht) beantworten
- sich auf **geographische & regionale Schwerpunkte** (Afrika, Germanen, Kelten, Monumentalität, Limes etc.) beziehen
- das **kulturelle Erbe** der Vorgeschichte und der Antike dokumentieren
- **Schnittmengen** zur Ethnologie, Anthropologie, den Sozial-, Geschichts-, Geo- und Naturwissenschaften, Baugeschichte, Denkmalpflege, ...



Ἐρε πνακαριος εφρεμ ο νιδρεφνιμ ἀνιμυ εφρασοϋ · η  
 20 ἀνιμυ εφρασοϋ · η ου  
 25 τηρ εφραμ εφραμ  
 30 ποσ · αυω πε  
 Τοϋναοϋνιμ ριτ  
 βασηλοομ εφρατ

6 Path: apophthegmata patrum 11 > AP 216 besarion\_merge  
 ἀλχοος παπα βησαριον σε παειωτ ροβε τονου . ρελλο δε λαφωνα

cb	cb													
dipl	α	ι	σο	ο	η	απα	βησαριον	σε	πα	ειωτ	†	οβε	τονου	
dipl_word	αλχοος		παπα		βησαριον		σε	παειωτ		†οβε	τονου			
lb_n	11		12		13		14							
norm	α	ι	σοο	ο	η	απα	βησαριον	σε	πα	ειωτ	†	οβε	τονου	
pb	pb													
pos	APST	PPERS	V	PPERO	PREP	N	NPROP	CONJ	PPOS	N	PPERS	VSTAT	ADV	
source_lang	Hebrew													
translation	I said to Abba Besarion:						"My father, I am very thirsty."							
word	αλχοος	παπα	βησαριον	σε	παειωτ	†οβε	τονου							
tok	α	ι	σο	ο	η	απα	βησαριον	σε	πα	ειωτ	†	οβε	το	νου



Projekte / Aktivitäten  
(Individuen, Gruppen, Institutionen)  
mit dynamischen Daten

Infrastrukturen  
mit statischen Daten

Web-Portale  
mit Daten

Erstellung

Verarbeitung

Analyse

Auswahl

Archivierung

Zugang

Nachnutzung

Lehre/  
Schulung

Werkzeuge

Dokumentation

Qualitäts-  
kontrolle

technische  
Infrastruktur

online/  
offline

Forschung

Ratgeber

Beratung

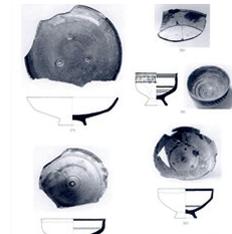
Anwendungen

Metadaten-  
standards

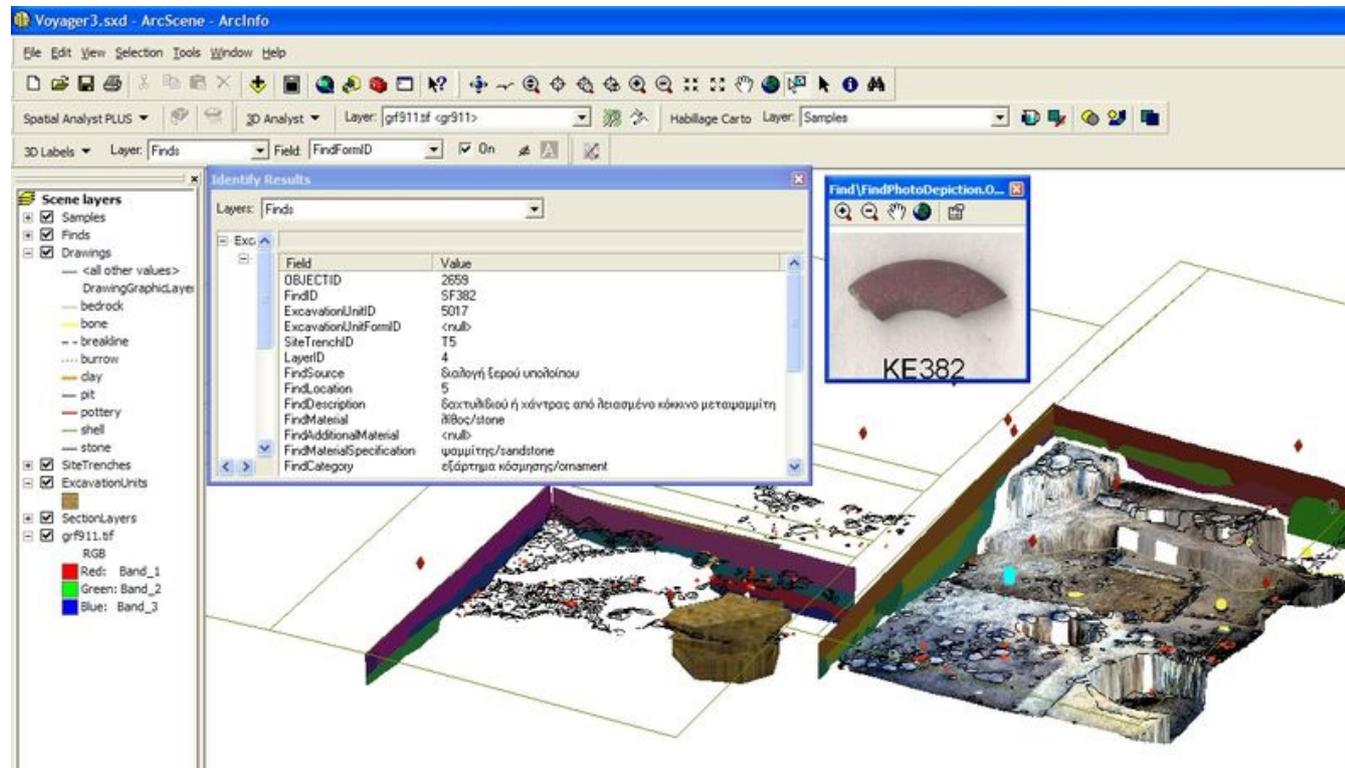
Datenpflege

persistente  
Identifikatoren

Interoperabilität  
& Web-Services



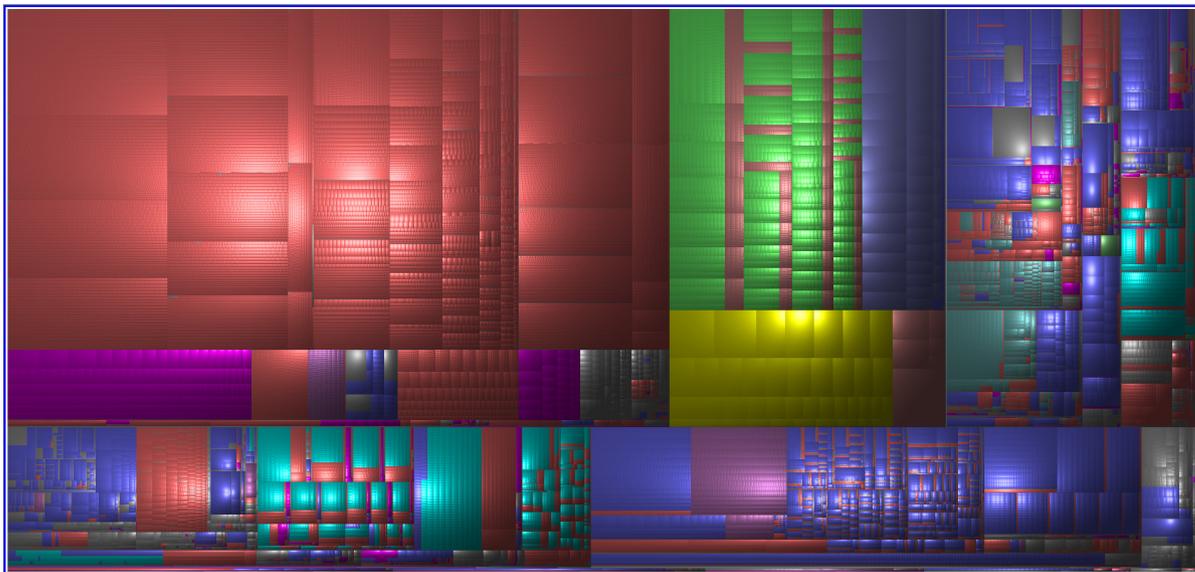
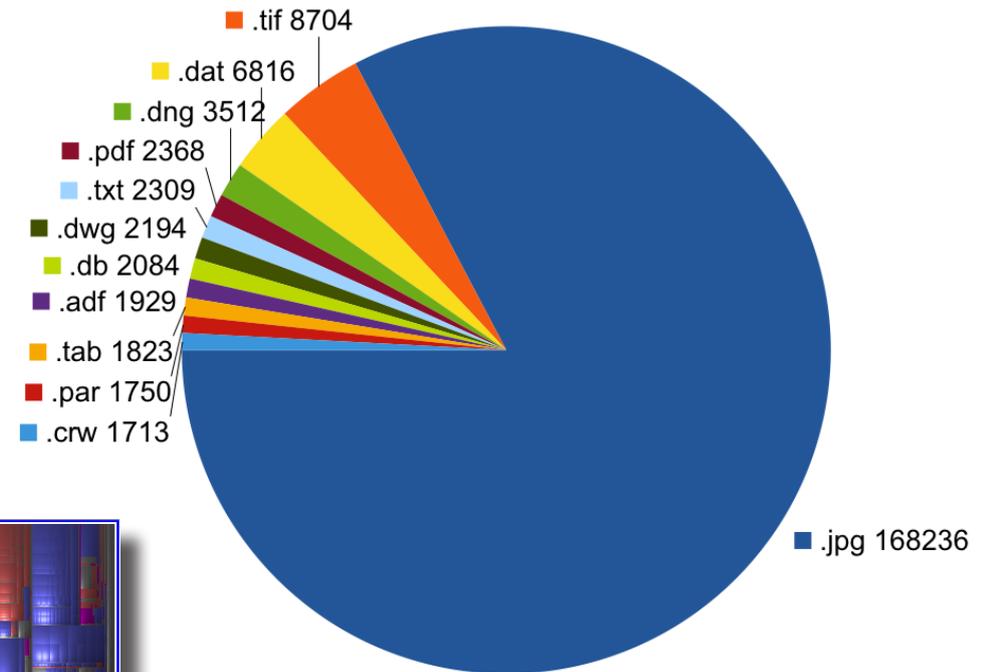
- Audio
- Vektordaten / CAD
- Datenbanken
- Fernerkundung / Satelliten
- Filme
- Geophysik
- GIS / Vermessung
- Laserscanning
- Mark-Up Text
- Photogrammetrie
- Rasterbilder
- Tabellen
- Statistik
- (Frei-)Texte
- 3D / Virtual Reality
- Webseiten



- 19 Sammlungen von Testdateien unterschiedlicher Einrichtungen und Disziplinen
- Nicht für Archivierung vorbereitete Datenbestände
- ca. 240.000 Dateien, ca. 7.500 Verzeichnissen, ca. 700 GB Speicher
- 462 unterschiedliche Dateiformate

Dateitypen nach Anzahl

Alle Testdaten



Forschungsdaten in der Archäologie und den Altertumswissenschaften umfassen

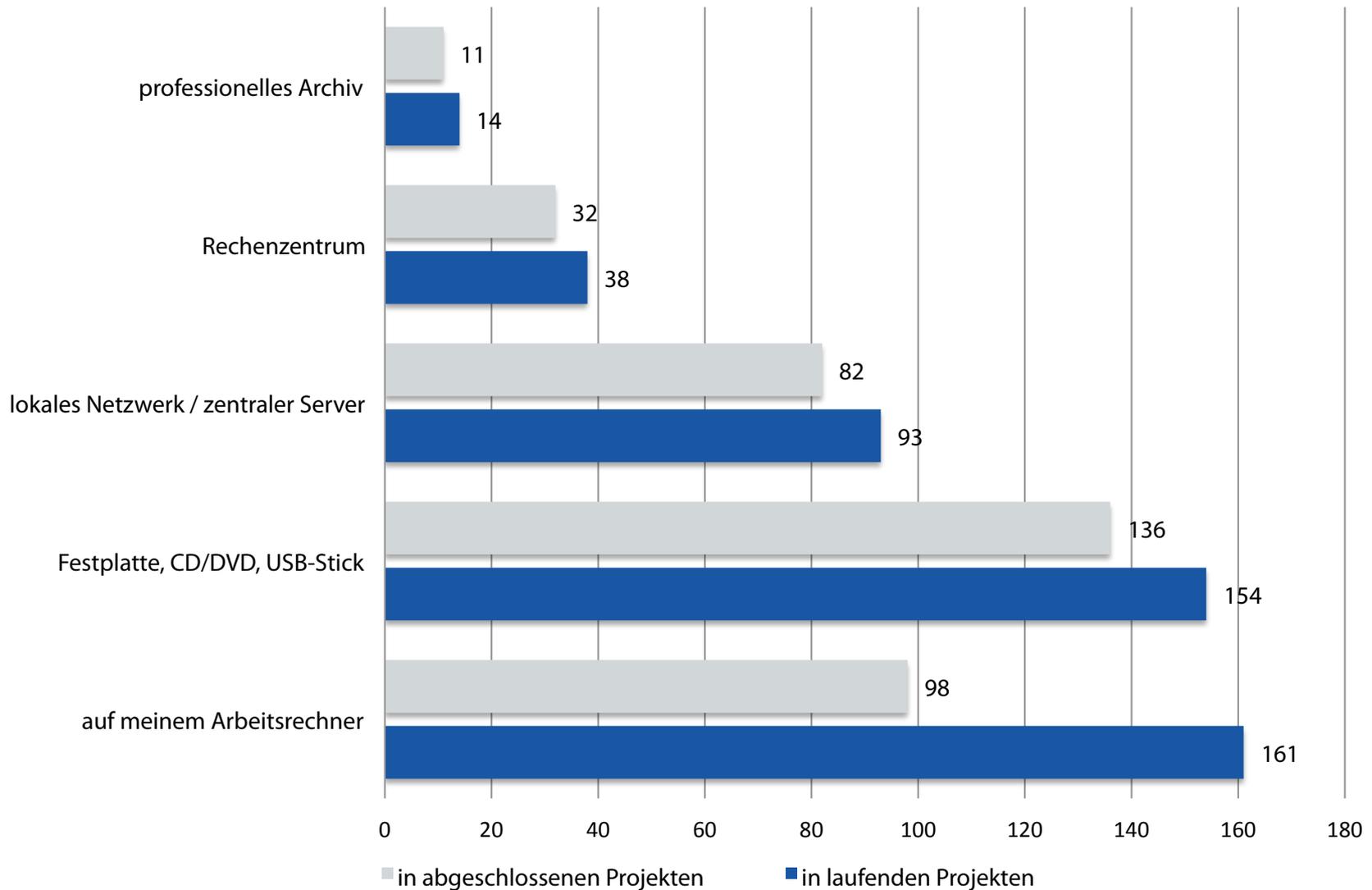
- unbearbeitete / primäre Daten
- prozessierte / sekundäre Daten
- publizierte / unpublizierte Daten
- Daten aus Projekten, Verwaltungsaufgaben und anderen Arbeitsprozessen

Sie sind

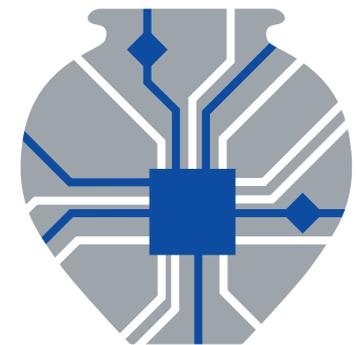
- **einzigartig**, weil nicht reproduzierbar
- institutionell **verteilt** und disparat
- **heterogen** in Inhalt und Form
- für lange Zeiträume **forschungsrelevant**
- für die Zukunft **gefährdet**



## Wie sichern Sie Ihre digitalen Forschungsdaten?



- DFG-Forschungsprojekt koordiniert durch das Deutsche Archäologische Institut in Berlin
- Reaktion auf die Anforderungen im Umgang mit digitalen Daten und Methoden
- Auftrag & Ziel:  
Aufbau und langfristiger Betrieb eines nationalen Forschungsdatenzentrums für die Archäologie und Altertumswissenschaften
- Kernaufgaben
  - › Langfristige Kuratierung, Archivierung und Bereitstellung von Forschungsdaten
  - › Beratung zu Fragen des Forschungsdatenmanagements
  - › Etablierung von Standards und Best-Practices
- 1. Förderung: 09.2011 - 08.2014 (laufend)  
2. Förderung: 09.2014 - 08.2017 (beantragt)

**DFG****IANUS**

## Fachrichtungen

- Materialorientierte Archäologien und Bauforschung
- auf Schrift/Text ausgerichtete Philologien und historische Fächer
- kulturhistorisch forschende Naturwissenschaften

## Institutionen & Personen

- Uni-Institute, Drittmittelprojekte, Forschungsverbünde
- Akademien, Museen, Forschungseinrichtungen
- Bodendenkmalämter
- Privatwirtschaft (z.B. Grabungsfirmen, Prospektionsbüros)
- Qualifizierungsarbeiten, interessierte Öffentlichkeit



## IANUS Forschungsdatenzentrum Archäologie & Altertumswissenschaften

Koordinierung: Deutsches Archäologisches Institut (DAI)



Steuerung: DFG-Arbeitsgruppe



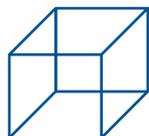
**rem**

Curt-Engelhorn-Zentrum  
Archäometrie gGmbH

berlin-brandenburgische  
AKADEMIE DER WISSENSCHAFTEN



EXCELLENCE  
CLUSTER



**TOPOI**

**SUB**

NIEDERSÄCHSISCHE STAATS- UND  
UNIVERSITÄTSBIBLIOTHEK GÖTTINGEN



human development in landscapes

archäologische  
staatssammlung  
münchen



Verband der  
Landesarchäologen  
in der Bundesrepublik  
Deutschland

Basis: Mitglieder (ca. 100) der verschiedenen Arbeitsgruppen aus unterschiedlichen Institutionen (ca. 40)

## Disziplinspezifische Vorbilder im Ausland



## Datenzentren/Archive anderer Disziplinen im Inland



## Vorbereitung durch Datengeber

- Erstellung Metadaten gemäß Vorgaben
- Vorbereitung Dateiformate gemäß Vorgaben
- Einhaltung Benennungsregeln
- Klärung Urheber- und Nutzungsrechte
- Erstellen von Metadatensatz über Online-Portal  
= Auslösung eines Tickets im Archiv-Management-System
  
- Übertragung der Daten:
  - › Kleine Datenmengen - Upload mittels Online-Formular
  - › Mittlere Datenmengen - FTP-Server / Cloud-Service
  - › Große Datenmengen - postalisch DVD / Festplatte

## Präferrierte und akzeptierte Dateiformate

Vektor-Grafiken / CAD			
✓	AutoCAD Drawing interchange Format Scalable Vector Graphics	dwg dxf svg	<b>ai, svg:</b> Verknüpfungen zu anderen Dokumenten, Titel <b>dwg, dxf:</b> Programmversion, Vorgaben/Elemente (Ebenen, Farben, Konturen, Muster, Symbole, etc.), Verknüpfungen zu anderen Dateien (Datenbanken, Objektbibliotheken, etc.) <b>dxf:</b> die ASCII-Version des Formates sollte bevorzugt werden <b>pdf:</b> CAD-Dateien können nicht mehr importiert und in CAD-Programmen wiederverwendet werden <i>Template für Metadaten verfügbar!</i>
~	CorelDraw Adobe Illustrator	cdr ai	
Rastergrafiken			
✓	Uncompressed Baseline TIFF v. 6 Adobe Digital Negative	tif & tiff dng	Titel, Photograph, Aufnahmedatum, Ort (Land- Bezirk-Gemeinde), Urheberrecht, Schlagwörter <i>Template für Metadaten verfügbar!</i>
~	Portable Network Graphics Joint Photographic Expert Group Graphics Interchange Format Bit-Mapped Graphics Format PhotoCD PICT Image Photoshop (Adobe) CorelPaint JPEG2000 RAW image format	png jpg & jpeg gif bmp pcd pict psd cpt jp2 & jpx verschiedene	

## Manuelle inhaltliche Überprüfung

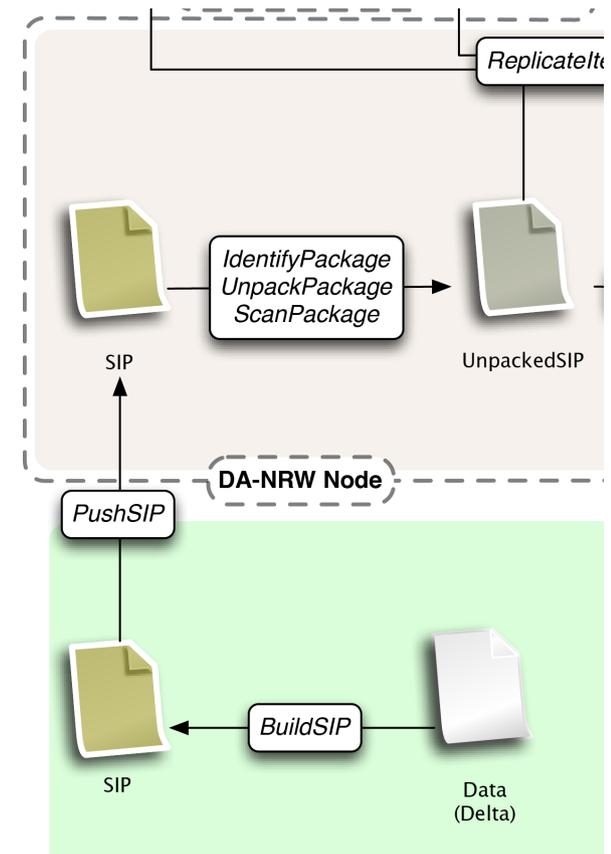
- Virus-Check
- Überprüfung Lesbarkeit des Datenträger und der Dateien
- Vergleich der Dateien mit zugehöriger Dokumentation, insbesondere Dateilisten
- Öffnen der Dateien (je nach Größe der Datensammlung exemplarisch)
  - › Feststellung korrupter Dateien
  - › Feststellung passwortgeschützter Dateien
- Überprüfung akzeptierte Dateiformate
- Überprüfung Vollständigkeit der Dokumentation
  - › Überprüfung geforderte Metadaten für gesamte Datensammlung
  - › Überprüfung geforderte Metadaten für einzelne Dateitypen

## Manuelle inhaltliche Überprüfung

- Festlegungen zur Datenbereitstellung
  - › Kurztexpte und Abbildungen für Web-Bereitstellung der Daten
  - › Klärung individueller Wünsche zur Darstellung
- Festlegungen Nutzungsrechte / Lizenzbedingungen für Weiternutzung
- Abschluss Datenübergabevertrag
- Einscannen analoger Dokumentation und sonstiger Texte (z.B. Korrespondenz, Lizenz, Vertrag)
- SIP-Builder starten
- Originale Datenträger beschriften und aufbewahren
- Arbeitsschritte in Archiv-Management-System dokumentieren

## Automatisierte technische Aufbereitung (DA-NRW: SIP-Builder)

- Ersetzung von Leerzeichen, Sonderzeichen durch Unterstriche
- Strukturierung Dateiablage (für Versionierung)
- Dateiliste erstellen
- Checksummen für jede Datei erzeugen (md5, BagIt)
- PREMIS Datei anlegen  
(v.a. mit Rechteinstellungen, basalen Metadaten)
- Komprimierung zur schnelleren Datenübertragung (tar)
- Kopieren der Daten auf Archiv-Server
- Überprüfung Kopiervorgang

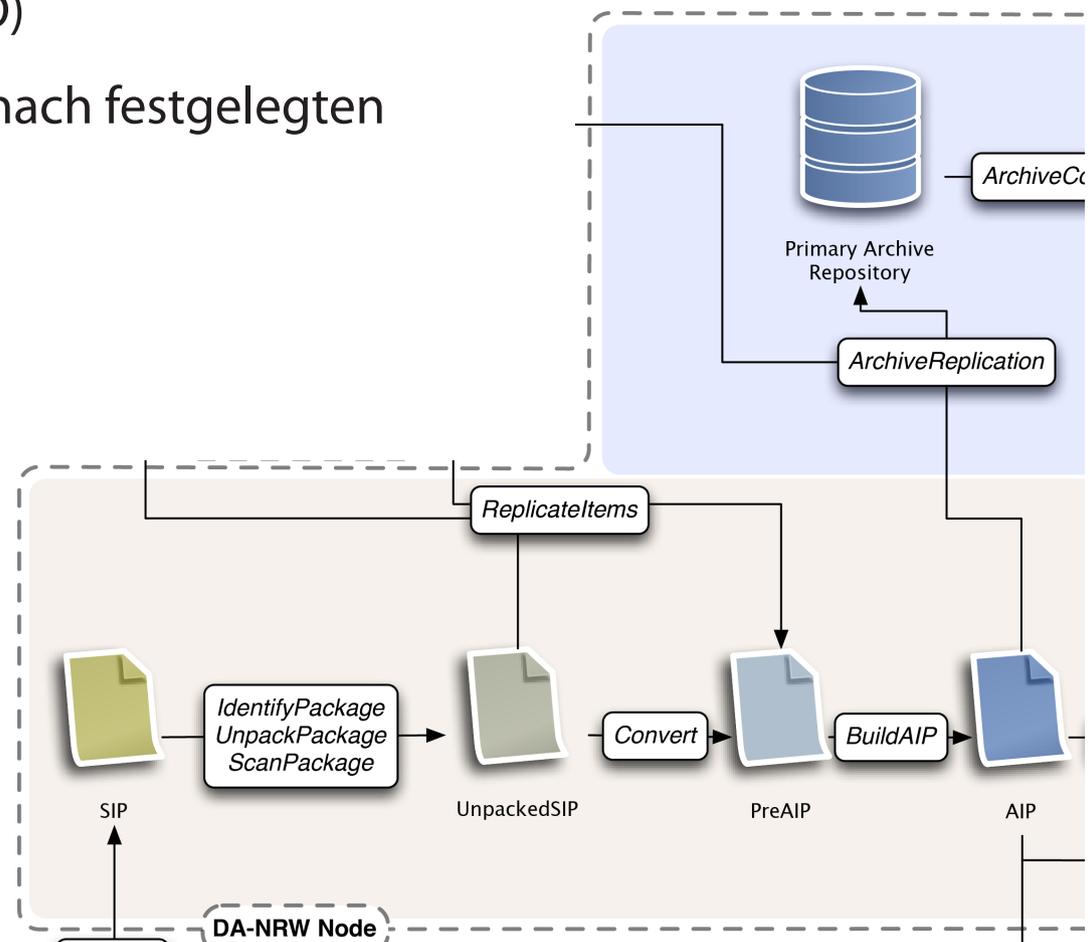


## Manuelle inhaltliche Kuratierung

- Vervollständigung und Standardisierung Metadaten
- Datenstruktur aufräumen und Dateinamen vereinheitlichen
- Verweise zu anderen AIPs
- Provisorische DOI reservieren, Granularität mit Datengeber festlegen
- Festlegung Migrationsplan / Archivierungsstrategie
  - › ggf. Vorbereitung zur technischen Umsetzung
- interne Kontrolle (2. Person, Checklisten)
- Metadaten aus Archiv-Management-System exportieren
- Content-Broker starten (SIP zu AIP)
- Arbeitsschritte in Archiv-Management-System dokumentieren

## Automatisierte technische Aufbereitung (DA-NRW: Content-Broker)

- SIP entpacken
- Formatidentifikation (FIDO / DROID)
- Formatkonvertierung der Dateien nach festgelegten Migrationspolicies
- Extraktion Metadaten
- Protokollierung Migration
- Erzeugung AIP
- AIP wird DOI zugewiesen
- Kopieren des AIP auf Archiv-Server
- Überprüfung Kopiervorgang

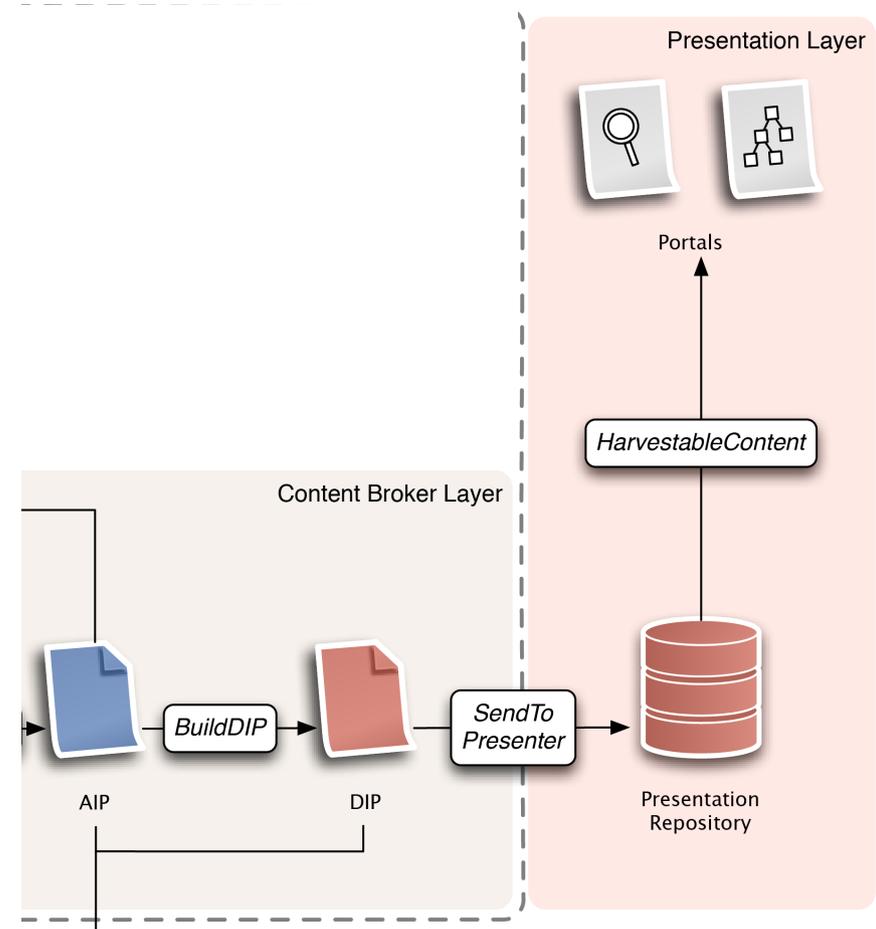


## Manuelle Vorbereitung für Bereitstellung

- Vorbereitung Landingpage (Beschreibung Datensammlung) für Web-Portal
- Umsetzung der Zugriffsregelungen mittels Rechte-Rollen-Konzept, nach Festlegungen durch Datengeber
- Rücksprache mit Datengeber zur Bereitstellung
- Zusendung GUI-Entwurf zur Diskussion
- Formulierung Nutzungsbedingungen mit Zitations-Hinweis
- Content-Broker starten (AIP zu DIP)
- interne Kontrolle (2. Person, Checklisten)
- Darstellung und Funktionalität über Web-Portal prüfen
- Datengeber informieren

## Automatisierte technische Aufbereitung (DA-NRW: Content-Broker)

- Migration der Dateien gemäß FIDO-Ergebnis, individueller Festlegungen und umgesetzter Policies
- Erzeugung Prüfsumme für Dateien, die Online bereitgestellt werden
- Registrierung DOIs
- Überprüfung DIP
- DIP auf Web-Server kopieren
- Überprüfung Kopiervorgang

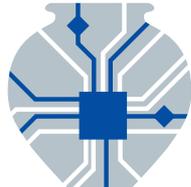


<http://www.ianus-fdz.de>

Anschrift  
c/o Deutsches Archäologisches  
Institut  
Podbielskiallee 69-71  
D-14195 Berlin  
Tel.: +49-(0)30-187711-359

Projektleitung  
Prof. Dr. Ortwin Dally  
Prof. Dr. Friederike Fless

Projektkoordinatoren  
Dr. Felix F. Schäfer  
Maurice Heinrich



**IANUS**  
Forschungsdatenzentrum  
Archäologie &  
Altertumswissenschaften

Datenerhaltung  
Planung  
Archivierung  
Austausch  
Forschung  
Digitale Daten  
Dokumentation  
IT-Empfehlungen  
Nachnutzung  
Metadaten



DFG  
e  
us  
en



Forschungsdatenzentrum  
Archäologie & Altertumswissenschaften

Suche:

[Ergebnisse](#) [Projektübersicht](#) [IT-Empfehlungen](#) [Partner](#) [Veranstaltungen](#)

IANUS - Forschungsdatenzentrum Archäologie &  
Altertumswissenschaften

[Zur Projektübersicht](#)

Herzlich willkommen auf den Seiten von IANUS !

der DFG gefördertes Projekt zum Aufbau eines  
ngsdaten zentrums für die Archäologien und  
haften in Deutschland, beantragt von einer DFG-  
koordiniert durch Frau Prof. Dr. Friederike Fless und  
twin Dally am Deutschen Archäologischen Institut in

ltung "IANUS - Alte Kulturen. Digitale Welten.  
ties"

iermit herzlich zum **IANUS-Empfang am Dienstag  
2014** im Auditorium Friedrichstraße einladen. Es werden  
NUS und den digitalen Geisteswissenschaften gehalten  
Benden Empfang besteht die Möglichkeit für Fragen und  
ltungen.

Anfahrt, Einladungskarte und Programm finden Sie hier

ranstaltung wird eine [Pressekonferenz](#) stattfinden, zu der  
erzlich einladen möchten.

ent/Leitbild

es, auf die Möglichkeiten und Herausforderungen zu  
durch die Arbeit mit digitalen Daten und Methoden  
nehmend den Alltag in den Altertumswissenschaften  
tschland bislang eine zentrale, disziplinspezifische  
Fragen und Antworten zum Forschungsdatenmanagement  
zyklus von forschungsrelevanten Informationen bündelt  
asste technische Lösungen umsetzt, wird eine Daten-  
struktur aufgebaut. Diese soll – nach einer Konzeptphase –  
von Bundesländern, Institutionen und Projektaufzeiten  
egelbetrieb IT-Dienstleistungen anbieten.

Koordination



Förderung



Kooperationspartner



Hochschule für Technik  
und Wirtschaft Berlin

University of Applied Sciences

Ankündigungen

IANUS-  
Informationsveranstaltung  
am 18.02.2014

Datenerhaltung  
Forschung  
Nachnutzung  
Digitale Daten