

Generische Komponenten und fachspezifische Anforderungen im KA³-Projekt

27. November 2017

Christoph Stollwerk
stollwerk@uni-koeln.de
Jonathan Blumtritt
blumtritt@uni-koeln.de



KA³



KA³ im Überblick



GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Förderkennzeichen: 01UG1511A

Kölner Zentrum **A**nalyse und **A**rchivierung von **AV**-Daten

Ein Forschungsdatenrepositorium für die Erhebung, Analyse, Archivierung und Bereitstellung von multimodalen Sprachdaten aus der Linguistik.

Beteiligung des RRZK im Verbundprojekt:

- Auf- /Ausbau eines fach- und standortübergreifenden institutionellen Repositoriums
- Expertise in den Bereichen Softwareentwicklung, Nachhaltigkeit und Betrieb einbringen



Kooperationspartner des RRZK



MAX PLANCK INSTITUTE
FOR PSYCHOLINGUISTICS

Repository

- Austausch von Basiskonzepten
- Erarbeitung von Metadatenformaten
- Entwicklung von Software



Fraunhofer
IAIS

Analyse

- Algorithmen
 - auf Datenmengen
 - Transkripte von Audio & Video
 - Segmentierung
 - von Dokumenten



hbz

FDM

- Pilotprojekt zum Thema Langzeitverfügbarkeit mit
 - HBZ, RWTH, USB
 - Evaluation der Software Rosetta
 - als Dienst des RRZK



Diversität im KA³-Projektverbund

Komplexe **Anforderungen** ...

- User Stories der Fachwissenschaften funktional abdecken
- Solide Software-Komponenten bereitstellen, die einen nachhaltigen Betrieb und eine effiziente Wartung ermöglichen

... und zahlreiche **Features** ...

- Ingest, Store, Preserve, Access, Search, Analyze, Transcript, Manage, Version, Secure, Publish, etc.

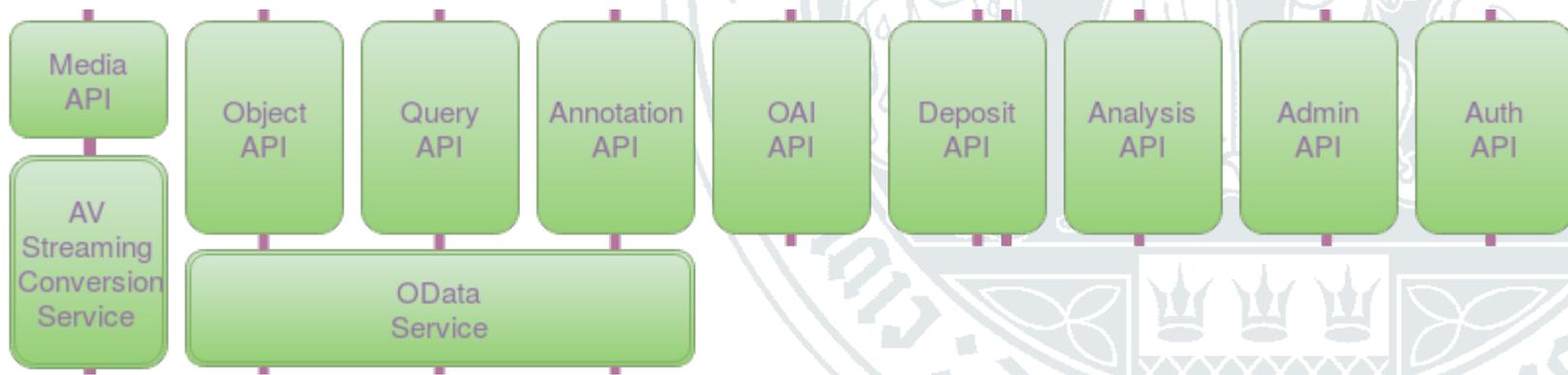
... erfordern sehr sorgfältig ausgewählte **Technologien**.

- Groovy/Grails 3, MariaDB, Elasticsearch 5, NetApp-Storage, Jenkins, Docker, Rosetta, Angular 2, , etc.



Generische Komponenten

Workflows wissenschaftlicher Arbeitsprozesse
werden durch zahlreiche API's ermöglicht:



Generische und institutionell etablierte Dienste dienen dabei als Schnittstellen, die fächerübergreifend genutzt werden.



Institutionell etablierte Dienste



Shibboleth®



vmware



Ausgewählte Technologien



docker



Apache Tomcat



Unterstützte Standards



KA³



Symbiose von Feature & Technik

Komplexe Anforderungen ...

(Fachwissenschaften funktional bedienen & Nachhaltige Software)

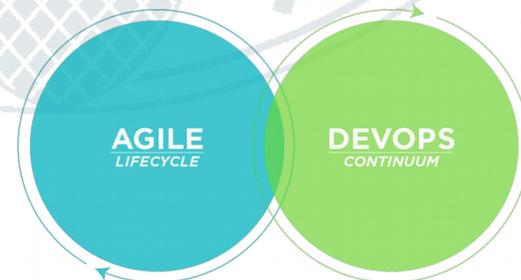
... und zahlreiche Features ...

(Ingest, Store, Access, Search, Analyze, Transcript, Manage, Publish, etc.)

... erfordern **nicht nur** sehr sorgfältig ausgewählte **Technologien**.

Sondern auch eine **Systematik**, die fähig dazu ist, die **Diversität** auf

technologischer und
nicht-technologischer
Ebene zulässt.





Fragen ??



