

Sichtweise einer Humanities-Einrichtung

Kai Wörner, Universität Hamburg



Universität Hamburg

DER FORSCHUNG | DER LEHRE | DER BILDUNG

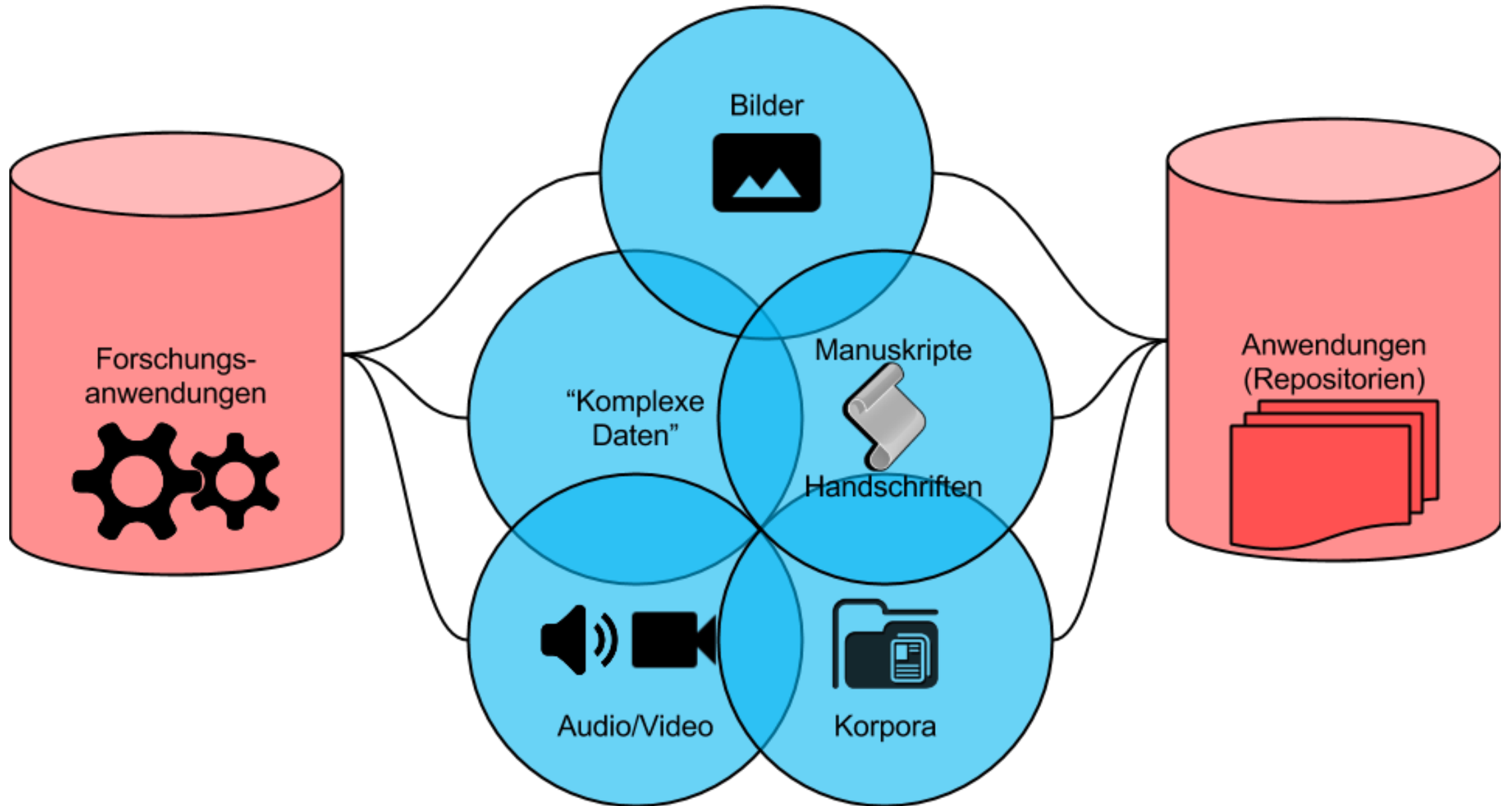
gwin

Geisteswissenschaftliche Infrastruktur für Nachhaltigkeit

<https://www.gwiss.uni-hamburg.de/gwin/>

- Gefördert durch das Präsidium der UHH, zunächst 3 Jahre
- 3 volle Stellen (1 Koordination, 2 Kuration)
- „Das Projekt ‚Geisteswissenschaftliche Infrastruktur für Nachhaltigkeit‘ (gwin) baut eine technische und organisatorische Infrastruktur auf, um die langfristige Verfügbarkeit von Forschungsdaten und Forschungsanwendungen in der Fakultät für Geisteswissenschaften sicherstellen zu können.“

UHH - Fakultät für Geisteswissenschaften



UHH - Fakultät für Geisteswissenschaften



Common Trends – „widely agreed“

- Layers of enabling technologies – agreed!
- Data Management Commons – agreed!
- Central Role of PIDs – agreed!
- Registered Data & Trusted Repositories – agreed!
- Physical and Logical Store – agreed!

Common Trends – „widely agreed“

- Layers of enabling technologies – agreed!
Aber: Es gibt mehr als „Datasets“
- Data Management Commons – agreed!
Aber: siehe oben!
- Central Role of PIDs – agreed!
- Registered Data & Trusted Repositories – agreed!
- Physical and Logical Store – agreed!
„as of now, there is no common agreement“

Common Trends – „widely agreed“

- Automatic Workflows – don't care (yet)
Bislang keine konkreten Anwendungsfälle
- Federations – very much agreed!
Aber: Schwierige Aufgabe, zusätzlicher Overhead.

Consequences of Principles – Change Data Culture

- “Make research Data open”
 - Beratung in diese Richtung wo immer möglich – ist aber häufig einfach nicht möglich (Datenschutz, Copyright auf Primär(?) -Daten, ...)
- Convince researchers to adhere to a simple high-level data model...
 - Ja, wo immer möglich (Koppelung an Dienstleistungszusagen)
- Educate researchers to do proper data citation to acknowledge data-related work
 - Ja, im Rahmen unserer Möglichkeiten
- Help change the existing culture to make data work be a recognized part of CVs...
 - “...for granting tenure.” – offensichtlich nicht in unserer Hand.
- Help to define proper mechanisms to use data citations in impact metrics.
 - Definition erarbeiten steht nicht auf dem Arbeitsplan.
- Help to train a new generation of data professionals.
 - Aktivitäten in der Lehre

Consequences of Principles: Discoverability

- Describe each digital object with adequate metadata to support data discovery.
 - Neue Projekte: Immer. Kuratiere Daten: Soweit möglich.
- Register the digital objects and make the discovery metadata available via machine-readable interfaces, e.g. OAI-PMH.
 - Werden wir machen – soweit leistbar.
- Register metadata schemas and their semantic categories in open registries to facilitate the process of metadata interpretation.
- Register metadata vocabularies that are being used in open registries.
 - Gerne, aber: (Siehe CLARIN CMDI/VLO-Diskussionen)
- Associate suitable information with PIDs to make it possible to trace digital objects back.
 - Ja
- Create provenance records that make it possible to trace back digital objects history.
 - Gute Idee, aber noch nicht

Consequences of Principles: Interpretation, Re-Use

- Associate information with the assigned PIDs that makes it possible to prove identity and integrity.
 - Wünschenswert
- Describe each digital object with metadata – including contextual information such as **prose text describing the creation process** and its manifold relationships) – that supports interpretation and re-use of the data.
 - Unbedingt – Unser eigenes FIS?
- Register schemas and their semantic categories in open registries to make it possible to interpret content.
- Register vocabularies being used in open registries.
- Register data types in type registries and associate executable data processing functions with them.
 - Alle drei: Siehe vorherige Seite

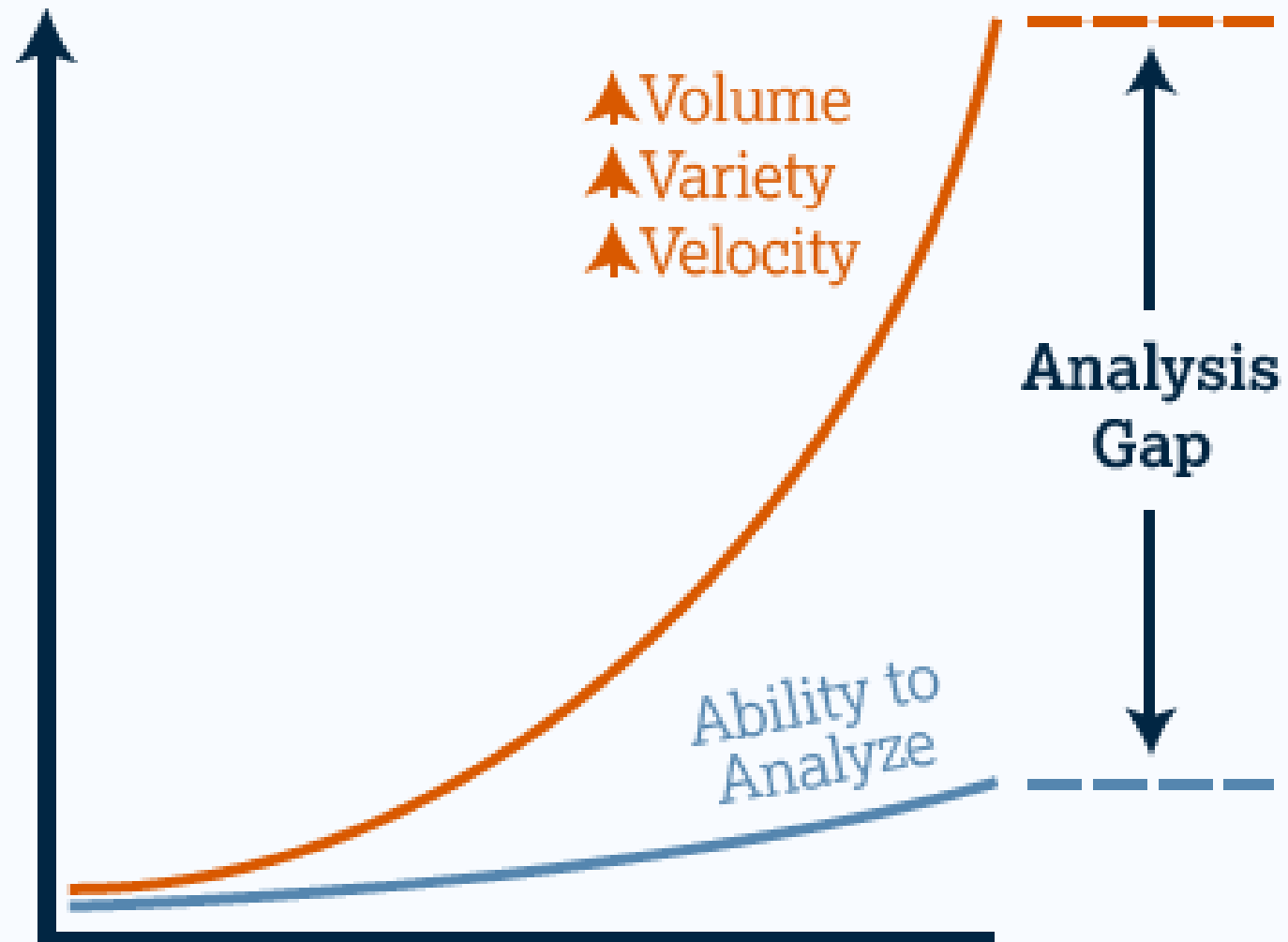
Consequences...: Management and Stewardship

- Define policies that guarantee accessibility, preservation and re-usability of data over time.
- Develop practical policies that turn these policies into executable and verifiable procedures and register them in open registries.
 - Beides Aufgaben mit hoher Priorität. Die weniger technischen Policies sind das größere Problem.
- Create policy frameworks that make it possible to easily integrate scientific algorithms into pre-fabricated workflows.
- Create registry systems for different purposes (such as storing metadata and information about schema, semantics, vocabularies, and practical policies) where necessary **and maintain them**.
- Agree on rules for data creation and handling that allow us to assess the quality of the data.
 - Alles wünschenswert, aber ohne Priorität.

Technical Components

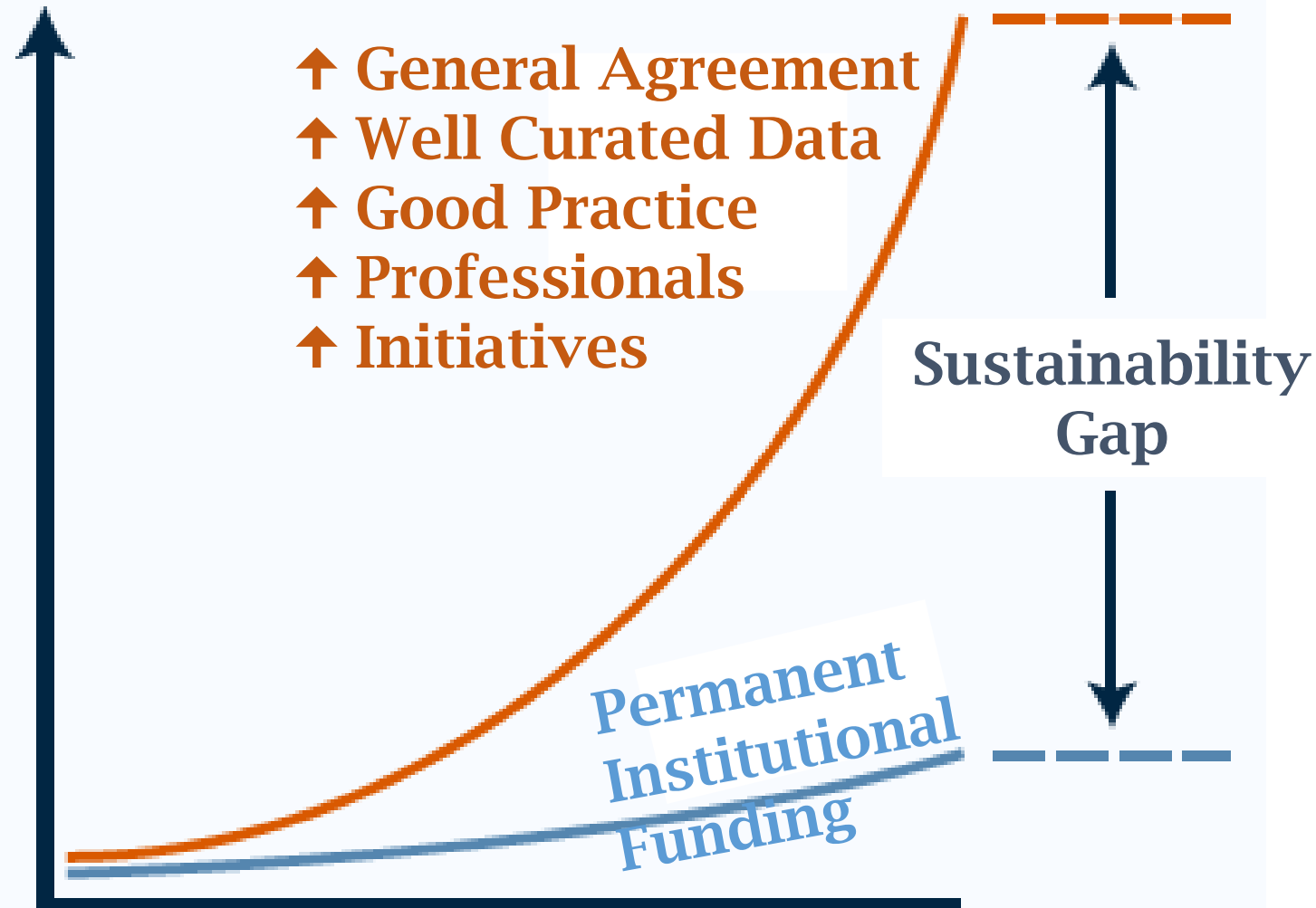
- **PID-System:** Ja! PIT-API offenbar eine gute Idee, aber auch hier viel wichtiger: Niedrige Hürde, um an PIDs zu kommen, langfristige (Finanzierungs-) Perspektive.
- **ID-System for Actors:** Unbedingt! Betörende Zukunftsmusik!
- **Trusted Repository Registries:** Bestehende Initiativen ausbauen/unterfüttern (re3data)
- **Registries, registries, registries...** Haben alle ihren Zweck, heiligt dieser Zweck alle Mittel?
- **Distributed Authentication System:** Ebenfalls wundervolle Zukunftsmusik!
- (...)

Information Explosion



Data Management Surrender

„Gefördert durch das
Präsidium der UHH,
zunächst 3 Jahre“



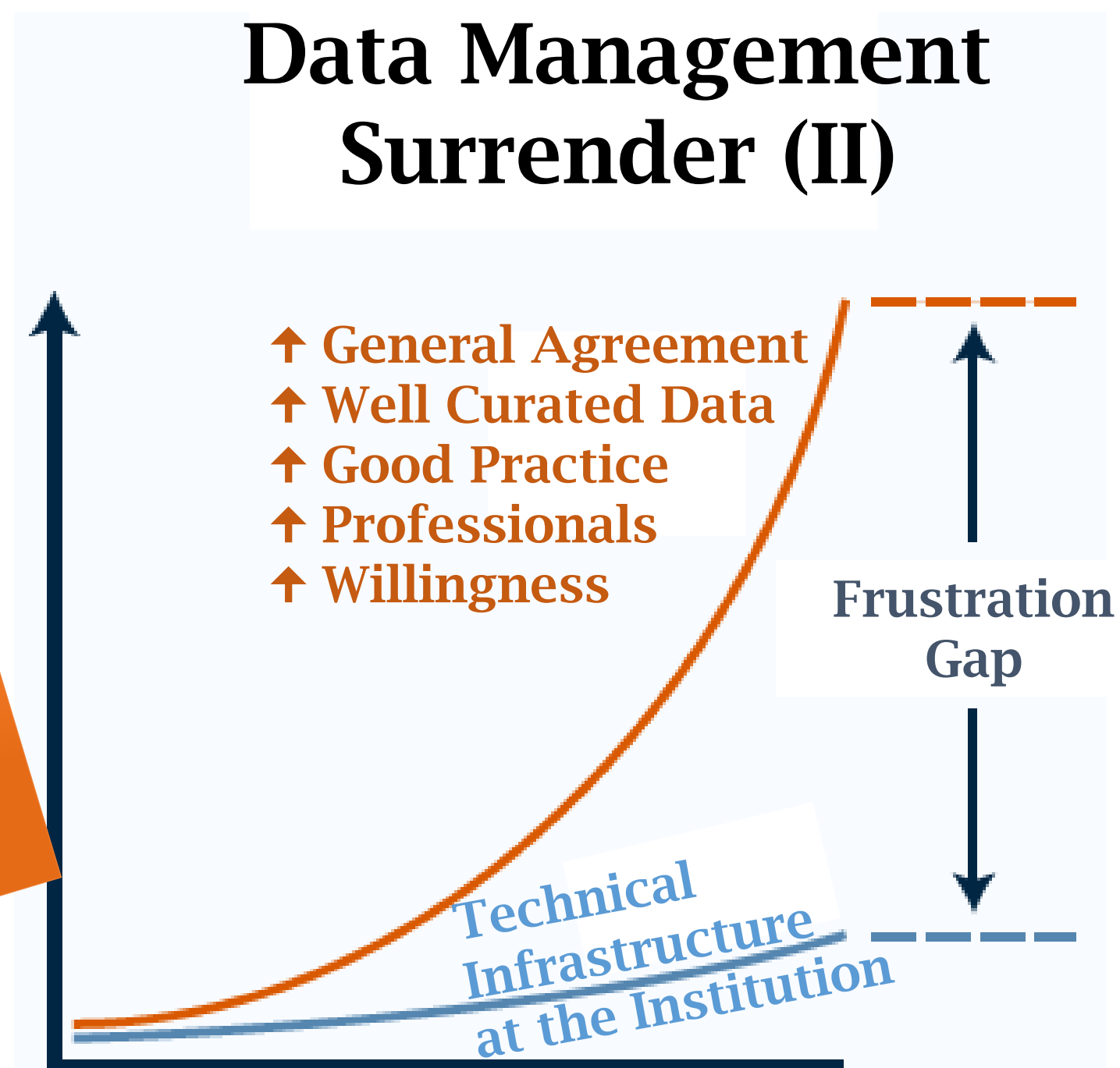
Data Management Surrender (II)

- ↑ General Agreement
- ↑ Well Curated Data
- ↑ Good Practice
- ↑ Professionals
- ↑ Willingness

Frustration
Gap

Technical
Infrastructure
at the Institution

Kein IDM
Kein Langzeitarchiv
Keine DFN-AAI
Kein FIS



RDA-Angebote

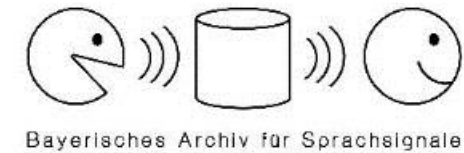
(Data Citation, Metadata Standards, Management Plans, PID Information Types, Practical Policy, Ethics and Social Aspects, **Long Tail of RD, Archiving multimedia interactive/dynamic data and projects, GW-IG in RDA?, ...**

Nützlich, unterstützenswert, in die richtige Richtung gedacht, aber...

Vor allem argumentatives Futter für die Lösung der Nachhaltigkeits- und Frustrationsprobleme.

Wie werden RDA-Resultate zu „meinen“ Resultaten?

Mitmachen! Aber wo?



UK Data Service



FORGE 2015

Forschungsdaten in den Geisteswissenschaften

- 16-18. September 2015, Universität Hamburg
- Vorläufiges Programm Anfang Juni

<http://uhh.de/forge2015>



Forschungsdaten in den

FORGE 2015

Geisteswissenschaften

16.-18. September 2015, Universität Hamburg

Fachspezifische Repositorien
Nachhaltigkeit für Forschungsanwendungen und Software
Vernetzung, Verbände und Initiativen
Forschungsdateninfrastrukturen vs. Datenzentren
Betriebs- und Geschäftsmodelle
Sichtbarkeit von Forschungsdaten
Rechtliche Fragen



Universität Hamburg, Von-Melle-Park 6 • Workshops: 15.–16. September • Hauptkonferenz: 16.–18. September • Fachmesse am 17. September

gwin Geisteswissenschaftliche
Infrastruktur
für Nachhaltigkeit

<http://uhh.de/forge2015>

UH
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG