

Consulting Workflow for Humanities Research Data

FORGE 2016, 15. September 2016
Jonathan Blumtritt

1

Consulting Workflow for Humanities Research Data, FORGE 2016,
Jonathan Blumtritt

15. 09 2016



Data Center for the Humanities

Kölner Datenzentrum für die Geisteswissenschaften

- zentrale Serviceeinrichtung der Philosophischen Fakultät der Universität zu Köln
- Gegründet 2013
- unterstützt Forscher an der Fakultät dabei, die dauerhafte Sicherung, Verfügbarkeit und Präsentation von Forschungsdaten und Ergebnissen der geisteswissenschaftlichen Forschung sicherzustellen

-> dch.uni-koeln.de

Forschungsdaten in den Geisteswissenschaften



- Philosophische Fakultät, Uni Köln
 - 27 Institute
 - wissenschaftliches Personal?
- auffällige Diversität:
 - Standards und Formate
 - Methodik und Fachtraditionen
- große Anzahl von projektspezifischen Lösungen/Formaten
- besondere Bedeutung von zugeschnittenen Präsentationssystemen für Forschungsprojekte
 - Projekte produzieren „lebende Systeme“

Angebot des DCH



Beratungsangebot Projektanträge

Das DCH hilft Ihnen für Ihren Projektantrag ein Konzept für Datenhaltung, Archivierung und Bereitstellung Ihrer Forschungsdaten zu entwickeln und umzusetzen.



Projektentwicklung

Das DCH führt in Kooperation mit Fachwissenschaftlern und unseren lokalen Partnern Entwicklungs- und Kurationsprojekte durch.



Digitale Ressourcen erhalten

Laufend entstehen digitale Ressourcen, Datenbanken, Webseiten und Forschungsdatenbestände im Rahmen von zeitlich befristeten Projekten. Was geschieht mit diesen, nach dem Ende des Projekts?

- Beratung und Vermittlung
- Kooperationsprojekte/Entwicklungsprojekte
- Betrieb und Betreuung von Softwarelösungen/Repositorien

Angebot des DCH



Beratungsangebot Projektanträge

Das DCH hilft Ihnen für Ihren Projektantrag ein Konzept für Datenhaltung, Archivierung und Bereitstellung Ihrer Forschungsdaten zu entwickeln und umzusetzen.



Projektentwicklung

Das DCH führt in Kooperation mit Fachwissenschaftlern und unseren lokalen Partnern Entwicklungs- und Kurationsprojekte durch.

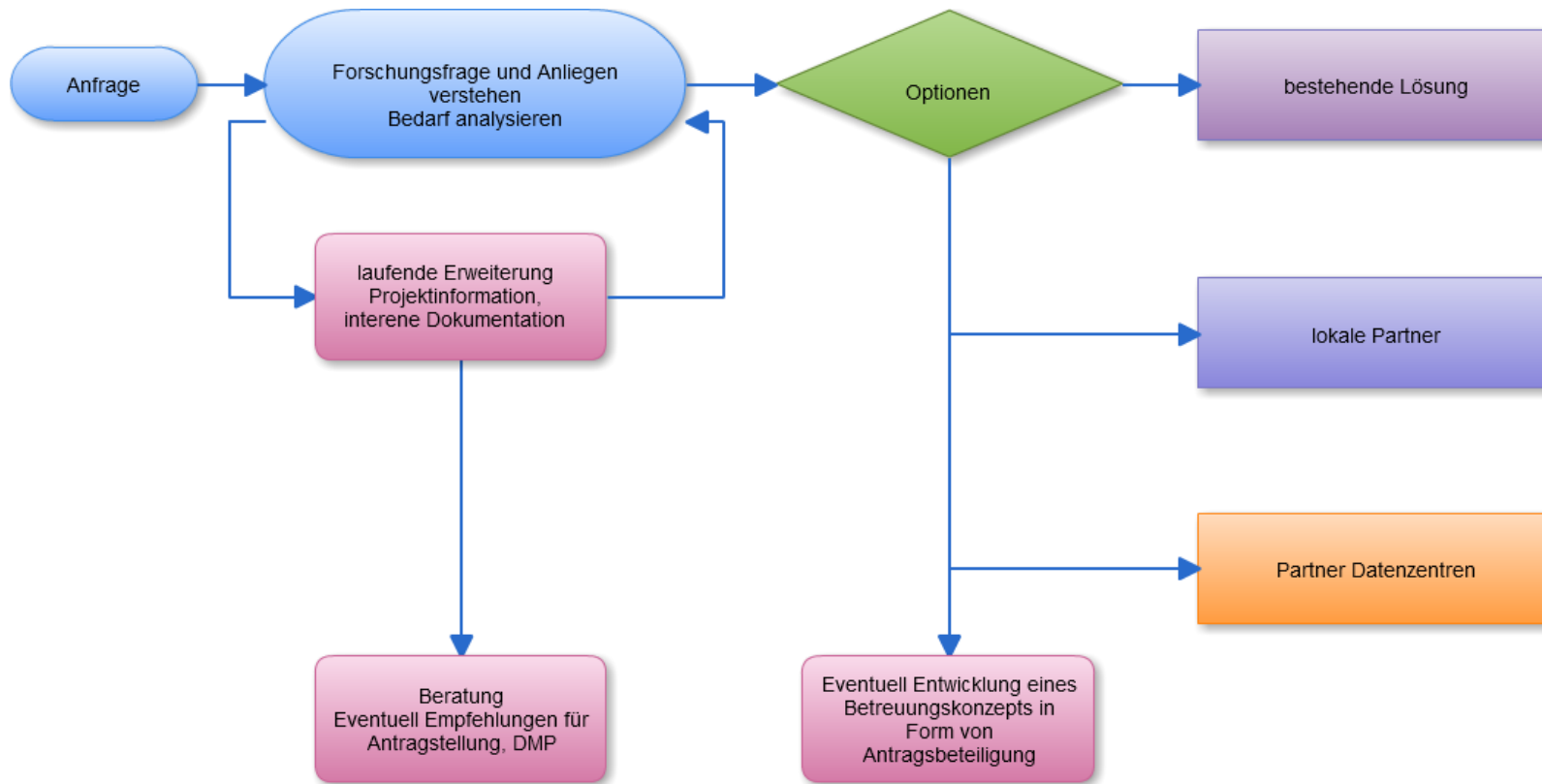


Digitale Ressourcen erhalten

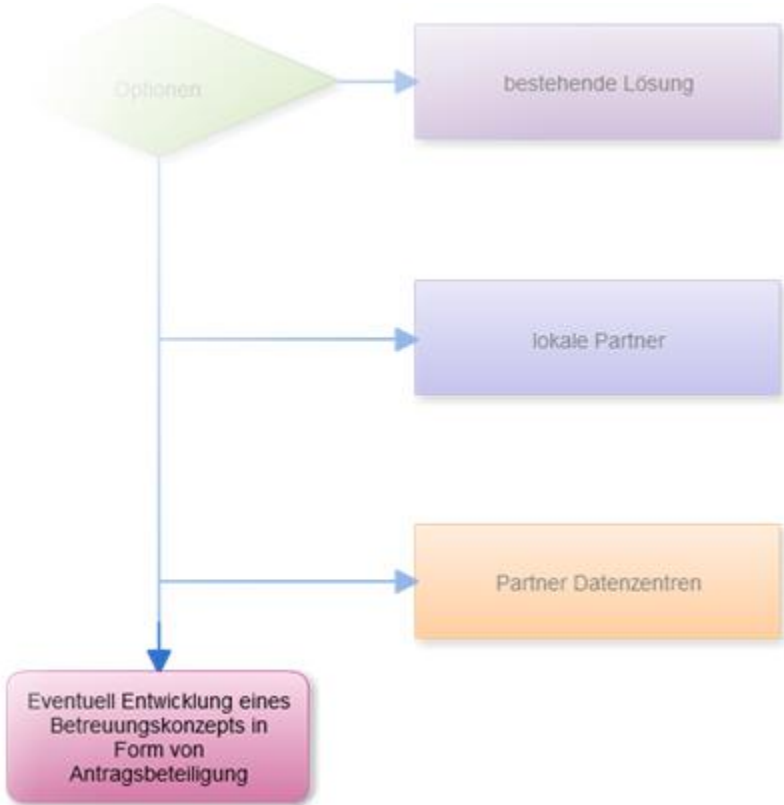
Laufend entstehen digitale Ressourcen, Datenbanken, Webseiten und Forschungsdatenbestände im Rahmen von zeitlich befristeten Projekten. Was geschieht mit diesen, nach dem Ende des Projekts?

- Beratung und Vermittlung
- Kooperationsprojekte/Entwicklungsprojekte
- Betrieb und Betreuung von Softwarelösungen/Repositorien

Beratung und Vermittlung: Beratungsworkflow



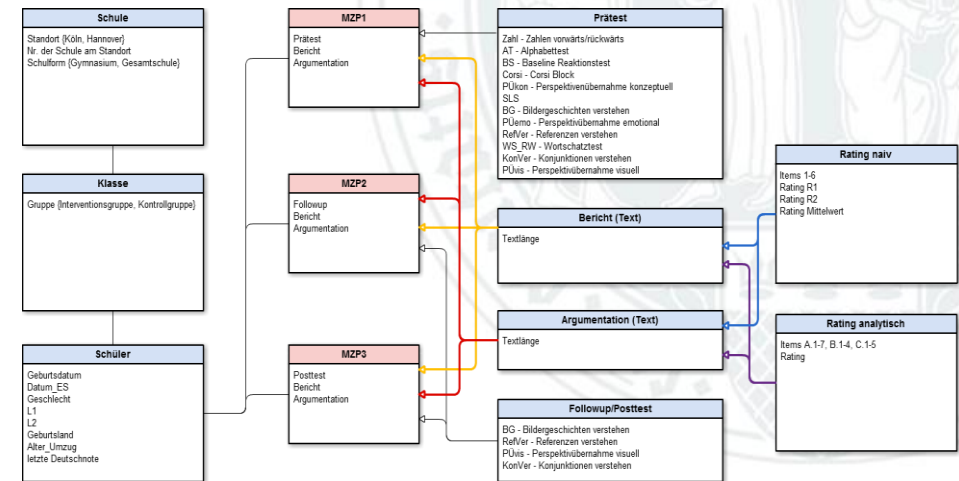
Kooperationsprojekte/Entwicklungsprojekte



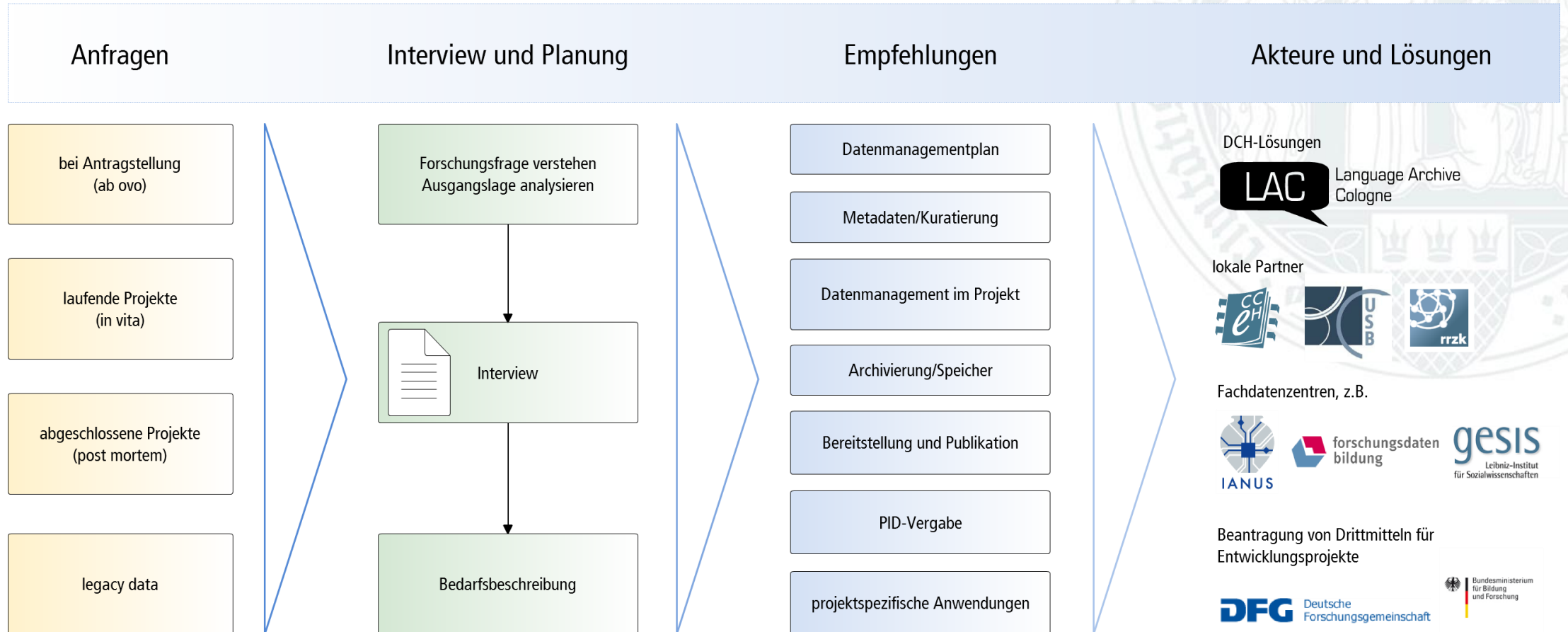
- Entwicklung nur
 - wenn keine bestehende Lösung greift
 - um Lücken zu schließen und bestimmte Ziele zu erreichen
- Bisher enge Zusammenarbeit mit dem CCeH
 - Expertise: Datenbasis XML-encodierte/mixed-content Texte, Datenmodellierung, X-Technologien
 - Häufig eXist-db als Entwicklungs-Plattform

Ablauf

- Interview, Gespräche
- Analysieren und Verstehen
- Kurzbericht
 - 1-2 Seiten Kurzbericht mit den wesentlichen Informationen
- Handlungsempfehlung
- Entscheidung in Teamsitzung



Beratungsworkflow



Interview, Ressourcenbeschreibung, Dokumentation (die Langform)

- Projektbeschreibung
 - Entstehungskontext
 - ...
- Daten
 - Methoden der Datenerzeugung
 - Datengeneese
 - Datenformate
 - ...
- Digitaler Zugang
 - Datenpublikation
 - Präsentationssystem
 - Technische Realisierung
 - ...
- Nachhaltigkeitsplanung
 - Übergabvereinbarung
 - Erhaltungsmaßnahmen
 - Technische Abhängigkeiten, Gefahren
 - Disaster Recovery Plan
 - ...
- Administrative und rechtliche Aspekte
 - Nachfolgeregulung
 - ...
- Weitere Dokumentation: Erfahrungswerte, lessons learned

Universität zu Köln, Datum

Ressourcenbeschreibung

Durch das Data Center for the Humanities (DCH) an der Universität zu Köln

Ressource: **Projektname**

Geschichte			
ID	Autor	Datum	Beschreibung

Zustimmung			
ID	Rolle / Autor	Partner	(Beschreibung / Grund)
	DCH		DCH Intern, bis Version 1.0 erreicht ist. [z.B. O habe intern, go 1.0 darf veröffentlicht, weitergegeben werden.]

Zustimmung			
ID	Rolle / Autor	Partner	(Beschreibung / Grund)
	DCH		DCH Intern, bis Version 1.0 erreicht ist. Erste Prüfung durch B. Abnahme durch Patrick Sahle erforderlich.
			Bei externen Projekten: Zustimmung zur Publikation von Projektbetreibern einholen.

Literatur / Referenzen	
Böttner, Stephan, Hans-Christoph Hobohm; Lars Müller (Hrsg.) Handbuch Forschungsdatenmanagement. Bad Honnef: rock + merches, 2013.	Datenmanagementplan Universität Bielefeld. https://data.uni-bielefeld.de/de/data-management-plan . (abgerufen April 2015).
Lühning, Jens, Harry Enke (Hrsg.) Leitfaden zum Forschungsdaten-Management, Handreichungen aus dem WISGRID-Projekt. http://www.wisgrid.de/publikationen/leitfaden_data-management-wisgrid.pdf . (abgerufen Mai 2015).	Deutsche Forschungsgemeinschaft, Leitlinien zum Umgang mit Forschungsdaten. http://www.dfg.de/foerderung/interregional_beguetzung_entscheidung/entragstellende/entragstellung_nachnutzung_forschungsdaten/ . (abgerufen April 2015).
Data Seal of Approval Guidelines. http://datasealapproval.org/media/ffler_public/2013/09/27/dataseal_2014-2015.pdf . (abgerufen Mai 2015).	Interne Umfrage des CCH (Cologne Center for the Humanities), 2014, Universität zu Köln.

Inhaltsverzeichnis

1. Projektbeschreibung	1
a) Projektinformation	1
b) Entstehungskontext der Ressource	1
c) Nutzungsschema, Nutzungsintention	2
d) Dokumentation	2
e) Datenmanagement	2
2. Daten	2
a) Datenbereiche, Formate, Modelle	2
b) Daten und Umfänge	2
c) Datengeneese, Datenerzeugung	2
d) Metadaten	3
e) Identifikatoren	3
f) Interoperabilität	3
3. Digitaler Zugang	3
a) Schnittstellen	3
b) Speicherung	4
c) Datenpublikation, Präsentationssystem, technische Realisierung	4
d) Monitoring	4
4. Nachhaltigkeitsplanung	4
a) Langzeitarchivierung	4
b) Übergabe der Ressource an das DCH	5
c) Erhaltungsmaßnahmen	5
d) Protokoll	5
e) Kosten und Ressourcen	5
f) Technische Abhängigkeiten, Gefahren	6
g) Disaster Recovery Plan	6
5. Administrative und rechtliche Aspekte	6
a) Vorgaben, Anforderungen	6
b) Nachfolgeregulung, Kassation	7
c) Urheberrecht, Eigentümer der Daten, Datenschutz	7
d) Zugriff und Nutzung, Nutzungsbeschränkungen	7
6. Weitere Dokumentation: Erfahrungswerte, Lessons learned	8

10

Consulting Workflow for Humanities Research Data, FORGE 2016,
Jonathan Blumtritt

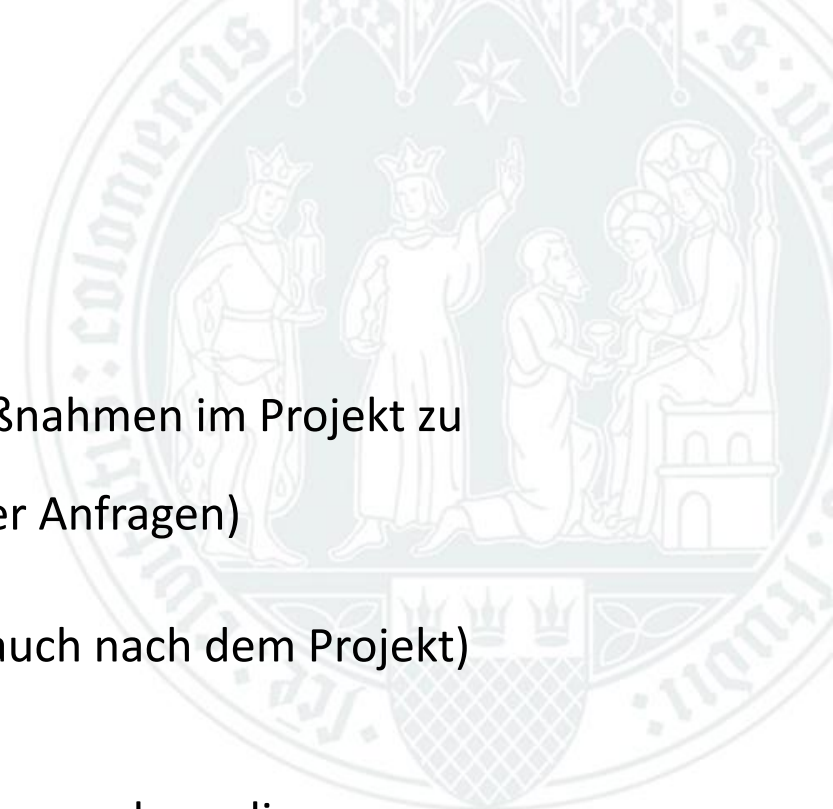
15.09.2016

Ressourcenbeschreibung: Quellen

- Büttner, Stephan; Hans-Christoph Hobohm; Lars Müller (Hrsg.). Handbuch Forschungsdatenmanagement. Bad Honnef: Bock + Herchen, 2011.
- Datenmanagementplan Universität Bielefeld: <https://data.uni-bielefeld.de/de/data-management-plan>, (abgerufen April 2015).
- Ludwig, Jens; Harry Enke (Hrsg.). Leitfaden zum Forschungsdaten-Management, Handreichungen aus dem WissGrid-Projekt: http://www.wissgrid.de/publikationen/Leitfaden_Data-Management-WissGrid.pdf, (Kurzfassung, S.98-100), (abgerufen Mai 2015).
- Deutsche Forschungsgesellschaft, Leitlinien zum Umgang mit Forschungsdaten: http://www.dfg.de/foerderung/antragstellung_begutachtung_entscheidung/antragstellende/antragstellung/nachnutzung_forschungsdaten/, (abgerufen April 2015).
- Data Seal of Approval Guidelines: http://datasealofapproval.org/media/filer_public/2013/09/27/guidelines_2014-2015.pdf (abgerufen Mai 2015).
- Interne Umfrage des CCeH (Cologne Center for the Humanities), 2014, Universität zu Köln.

Typologie Beratungsanfragen

- vor Antragstellung
 - mit genügend Vorlauf, um Strategie und Finanzierung von FDM-Maßnahmen im Projekt zu verankern
 - kurzfristig (1-2 Wochen) vor Antragseinreichung (große Mehrheit der Anfragen)
- laufende Projekte
 - Fragen rund um Datenmanagement, Sicherung und Verfügbarkeit (auch nach dem Projekt) ergeben sich erst in der Erhebungsphase
- nach Projektende
 - unbetreute Ressourcen durch personelle/institutionelle Veränderungen; ehemalige Projektmitarbeiter oder Mitglieder der Fachcommunity, die an der weiteren Nutzung interessiert sind
- legacy data
 - Projekt liegt zu weit zurück; niemand, der aktiv Verantwortung übernimmt; Keine funktionsfähige Oberfläche; aufwändige Maßnahmen zur Revitalisierung notwendig.



Typologie Beratungsanfragen: Kategorisierung der Bedürfnisse



Workspace



Archive



Publication



- Datenmanagement und Austausch von Daten im Team (aka “Cloud”)
- kollaboratives Bearbeiten den Daten im Team
- Tools, VRE, ...



- Archivierung, Bitstream preservation
- Langzeitverfügbarkeit, Langzeitarchivierung
- self-archiving



- zitierfähige digitale Publikation
- Projektspezifische Präsentationsplattform?
Zugangskontrolle, Access Rights Management
- Discovery:
 - Metadatenkatalog, Browsing, facettierte Suche
 - Volltextsuche, Content-Search
- Ressourcen herunterladen, anschauen, abspielen
- Auffindbarkeit/weitere Verbreitung Metakataloge

Typologie Beratungsanfragen: Datenbestand vs. Software

- Datenbestand und/oder Forschungsanwendung
- „lebendes System“



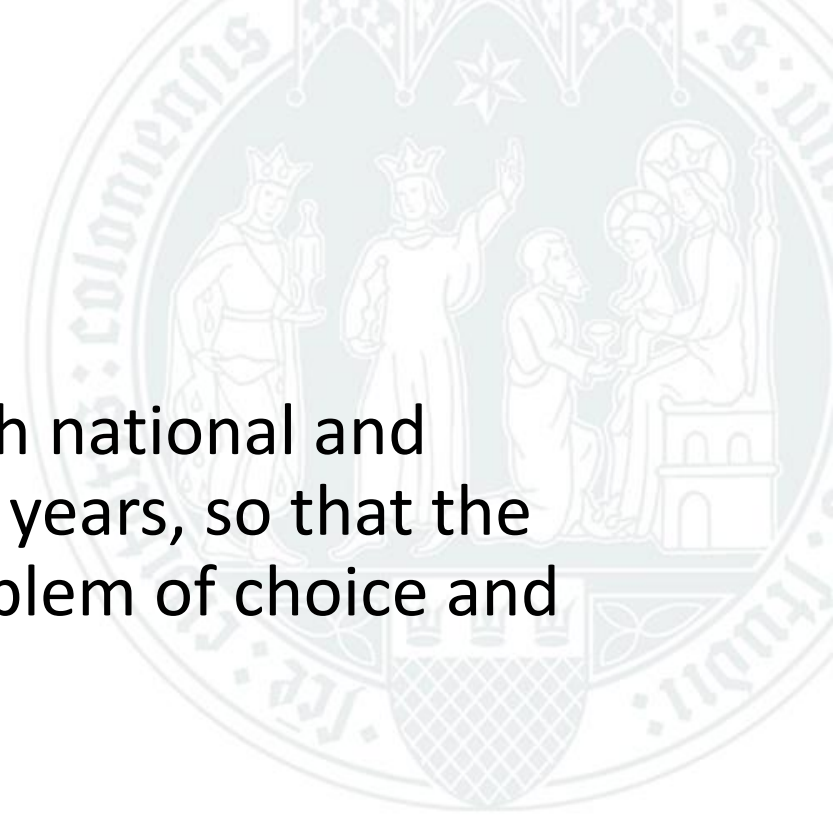
Beratungsanfragen: Übersicht

#	Institut, Projekt	Phase	LS/SW	Daten	W	A	P
1		4		x		x
2		1		x		x
3		2		x	x	x
4		1	x	x		x
5		1		x		x
6		1	x			x
7		3	x	x		x
8		2		x	x	x
9		3	x			x
10		3	x			x
11		1	x			x
12		3		x		x
13		1		x		x
14		1	x			x
15		2		x		x
16		1		x		x
17		1		x		x
18		1		x		x
19		1		x		x
20		1		x		x



Beratung und Vermittlung

Abstract: „A plethora of options for storage on both national and international level have been established in recent years, so that the problem of basic preservation is increasingly a problem of choice and awareness.“



Beratung und Vermittlung

Partner Fachdatenzentrum:

"Momentan besteht keine Möglichkeit, die Daten des Projektes [...] in der gewünschten Form und vor allem mit den gewünschten Filterfunktionen [...] abzubilden. [...] **Auch eine gröbere Erschließung und Bereitstellung der Daten auf Kollektionsebene ist in diesem Fall schwierig, da die Daten gerade durch ihre Verknüpfung zueinander interessant sind.**"



Fazit

- Das Angebot wird angenommen
- „Workspace“-Problematik scheint gelöst?
- Die Anzahl von Erstkontakten VOR Antragstellung könnte höher sein
- Vermittlung ist in vielen Fällen erfolgreich
- projektspezifische Anwendungen, „lebende Systeme“ bleiben in vielen Fällen unumgänglich

