
Abschluss-Workshop KResCo

Wolfgang Beinhauer, Dr. Daniela Beyer, Siegrid Hecht-Veenhuis, Kathrin Tuppi, Johannes Sautter, Andrea Wuchner



AP2 Forschungsdatenmanagement, 31.01. 2022

Verschiedene externe Stakeholder stellen zunehmend Anforderungen an das Forschungsdatenmanagement

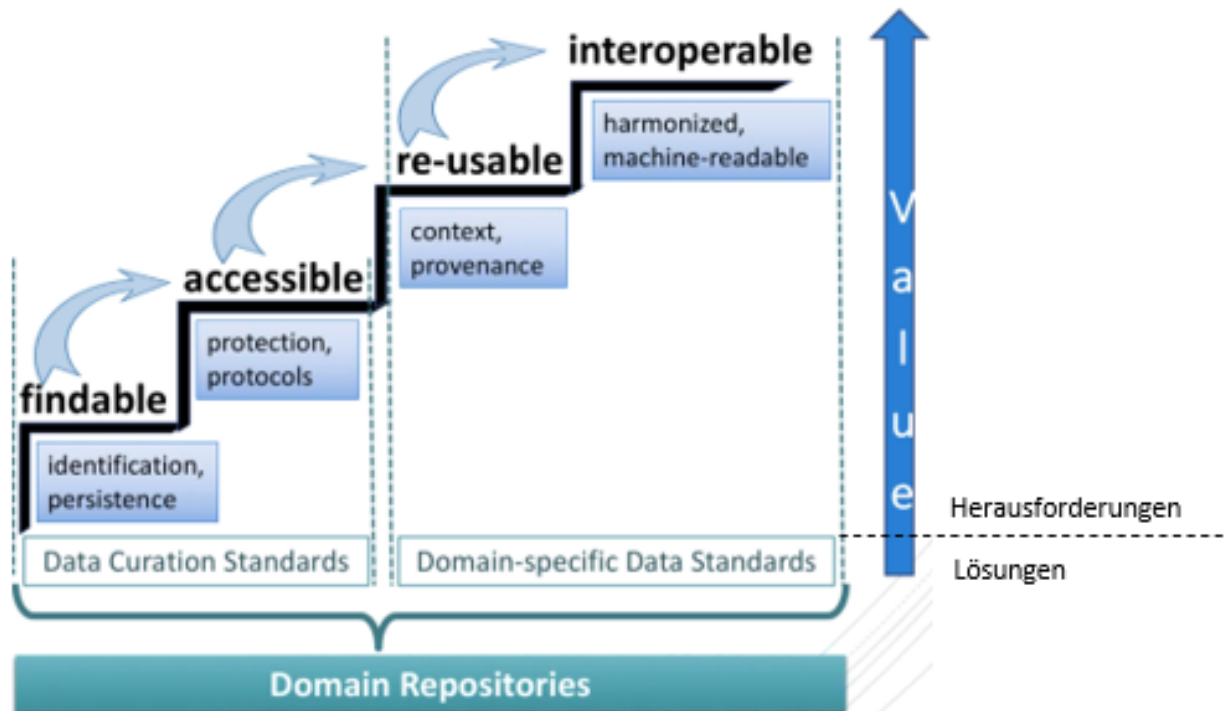
- DFG:
 - [Leitlinien für den Umgang mit Forschungsdaten](#) (2015)
 - [Leitlinien zur Sicherung der guten wissenschaftlichen Praxis](#) (2019)

- Europäische Kommission:
 - [Horizon Europe Model Grant Agreement](#) (2021)
 - [FAIR-Prinzipien](#) (2016)

- [Datenschutzgrundverordnung](#) (2018)

Wie können wir sukzessive den Wert der Daten erhöhen?

Datenperspektive



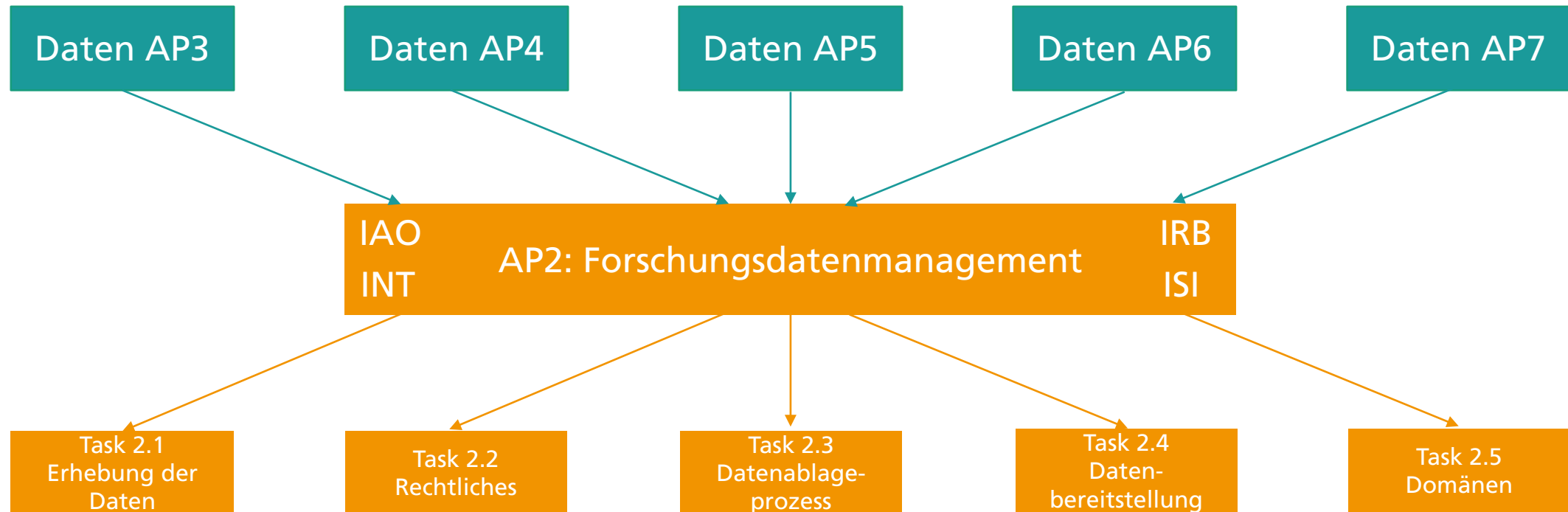
Nach Tim Ahern, Ted Haberman, Kerstin Lehnert: The "I&R" in FAIR: Who develops, approves, and governs domain-specific standards in the Earth, Space, & Environmental Sciences?, 2019 ESIP Winter Meeting

Organisationsperspektive Herausforderungen

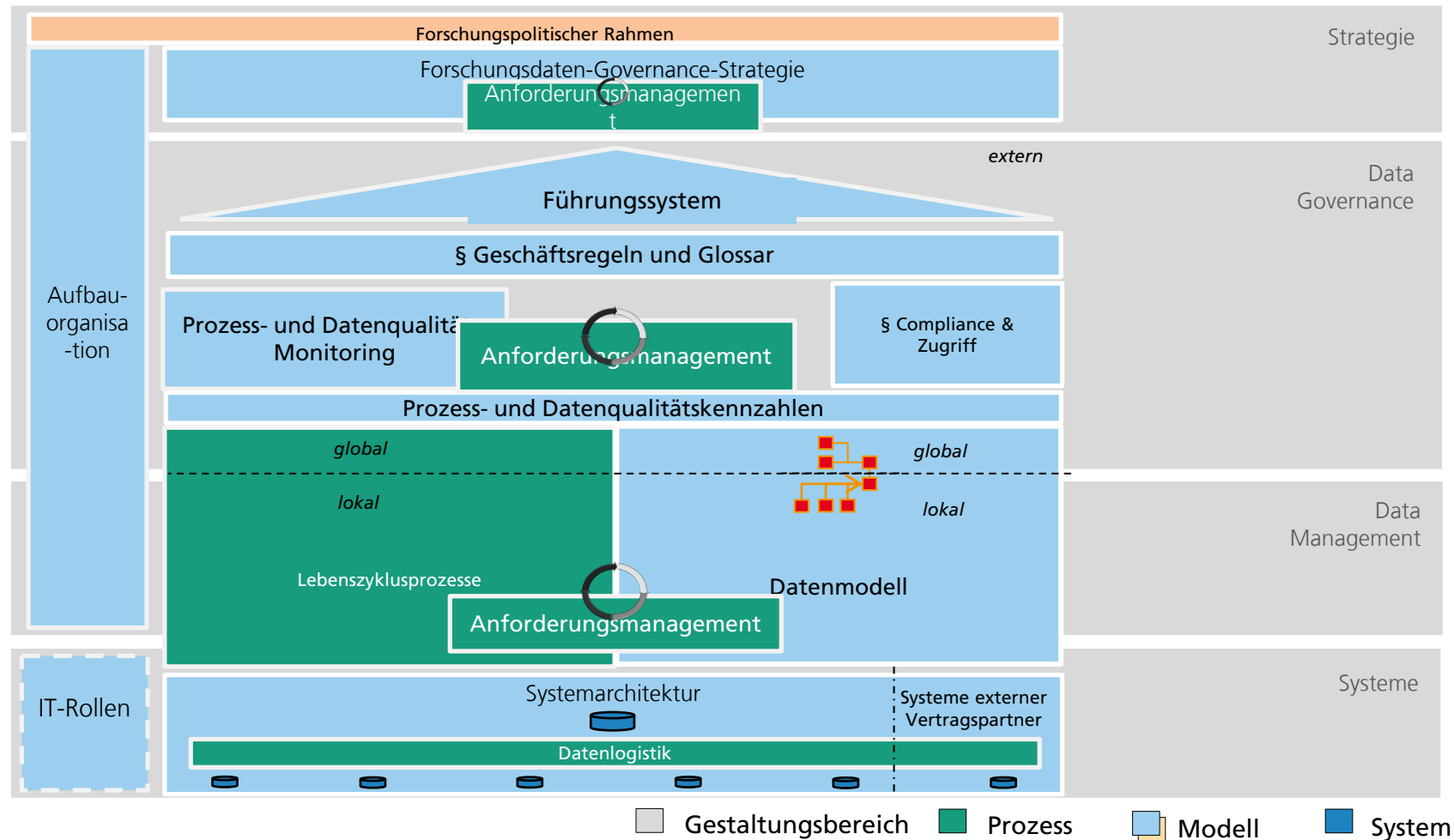
- **Operative Exzellenz**
 - Operative Effizienz
 - Austausch und Kollaboration
 - Datenverfügbarkeit
 - Operative Integration und Interoperabilität
 - **Rechtliche Exzellenz**
 - Operative rechtliche Anforderungen
 - Kenntnis der Vorschriften
 - **Exzellenz in der Prozessqualität**
 - Klare und zuverlässige Verantwortlichkeiten, Prozesse und Richtlinien
 - Datentransparenz und Überprüfbarkeit
 - Zentraler Datenzugriff
- **Datenqualität**
 - Intrinsisch
 - Zugänglichkeits-betr.
 - Kontextuell
 - Repräsentativ

Quelle: Sautter, J., Litauer, R., Fischer, R., Klages, T., Wachner, A., Müller, E., Schaj, G., Dobrovkatova, E., Drems, P. and Rieß, S.: Beyond Data Quality: Data Excellence Challenges from an Enterprise, Research and City Perspective. In Proceedings of the 7th International Conference on Data Science, Technology and Applications (DATA 2018), pages 245-252, ISBN: 978-989-758-318-6

Projektstruktur für das Forschungsdatenmanagement

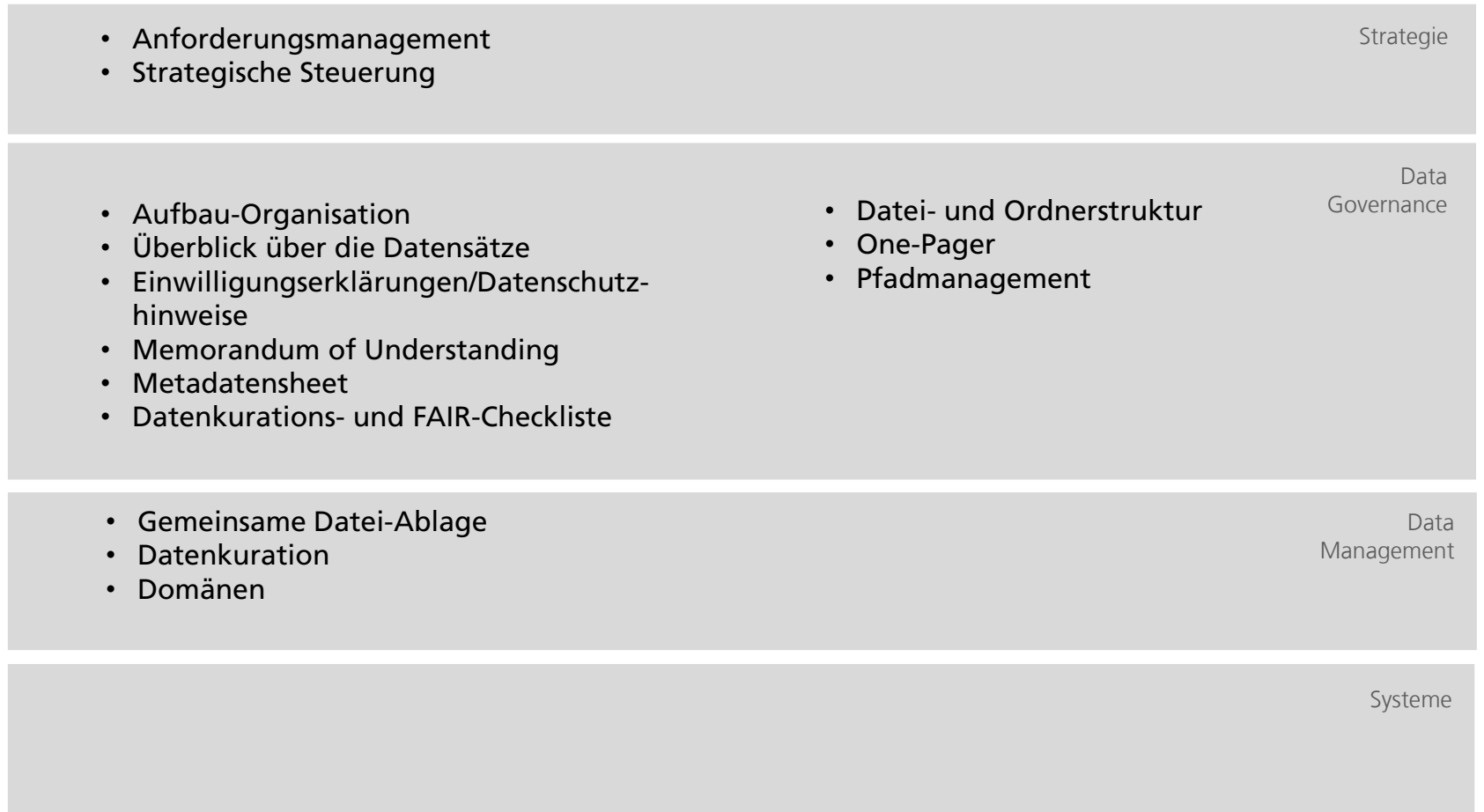


Das Data-Governance-Referenzmodell als Grundlage



Nach J. Sautter, S. Riess, L. Loch, D. Kopperger, R. Litauer, M. Stanisic-Petrovic, C. Schelling: Data Governance (2019). Unveröffentlicht.

Umsetzung des Data-Governance-Ansatzes



□ Gestaltungsbereich ■ Prozess □ Modell ■ System

Nach J. Sautter, S. Riess, L. Loch, D. Kopperger, R. Litauer, M. Stanic-Petrovic, C. Schelling: Data Governance (2019). Unveröffentlicht.

»KResCo-Space « in der Owncloud

Name	Änderungsdatum
!Akquisition	25.11.2020 12:26
!Ergebnisse	25.11.2020 12:26
!Forschungsdaten	25.11.2020 12:26
!Kommunikation	25.11.2020 12:26
!Verträge	25.11.2020 12:26
Abrechnungen	25.11.2020 12:26
Arbeitspakete	25.11.2020 12:26
Meetings_und_Veranstaltungen	25.11.2020 12:26
Projektmanagement	25.11.2020 12:26
Publikationen	25.11.2020 13:06
Sonstiges	25.11.2020 12:26
Unteraufträge	25.11.2020 12:26
Readme.txt	27.08.2019 10:31

Name	Date modified	Type
AP1_Projektmanagement	18.11.2020 11:08	File folder
AP2_Forschungsdatenmanagement	18.11.2020 11:08	File folder
AP3_PolitischeEntscheidungen	18.11.2020 11:17	File folder
AP4_Wirtschaft&Gesellschaft	18.11.2020 11:09	File folder
AP5_Innovation	18.11.2020 11:09	File folder
AP6_Bevölkerungsschutz	18.11.2020 11:09	File folder
AP7_Forschung	18.11.2020 11:09	File folder
AP8_Handlungsempfehlungen	18.11.2020 11:09	File folder

■ 2 Mal die AP-Struktur

■ Aktivitäten von AP2 nur unterhalb !Forschungsdaten

Name	Änderungsdatum
AP3_Politische Entscheidungen	25.11.2020 13:05
AP4_Wirtschaft_und_Gesellschaft	25.11.2020 13:05
AP5_Innovation	25.11.2020 13:06
AP6_Bevoelkerungsschutz	25.11.2020 13:06
AP7_Forschung	25.11.2020 13:06
AP8_Handlungsempfehlungen	25.11.2020 13:06
Erhebung_Datensätze_Kresco_V2.xlsx	04.11.2020 10:24
Readme.txt	27.08.2019 10:31

Memorandum of Understanding: Vereinbarung zur gemeinsamen Datennutzung

- Das MoU regelt die Nutzung der Forschungsdaten zwischen den Projektpartnern. Kernpunkte sind folgende:
 - “Alle Projektbeteiligten erhalten Zugriff und Einsicht in offene, unsensible und sensible Daten. Hochsensible Daten bleiben am jeweiligen Institut. In der Regel sind das Daten, die vertraulichen Personenbezug enthalten.”
 - “Geteilte Datensätze dürfen nicht zweckentfremdet werden”
 - “Das Teilen der Datensätze befähigt andere Projektinstitute allein zur Nutzung und zweckmäßiger Verarbeitung unter den gemeinsamen Forschungszielen des Projekts.”
 - “Die Projektpartner verpflichten sich hiermit, die geteilten Daten nur unter Einbezug des Datenurhebers außerhalb des Projektes weiter zu nutzen. Eine geplante Nutzung oder Veröffentlichung ist im Konsortium anzuzeigen. Im Falle von personenbezogenen Daten ist die Einwilligungserklärung des Datengebers zu beachten.”

Dokumentation der Forschungsdaten mit dem Metadatenheet

1								https://www.dnb.de/DE/Professionell/DDC-Deutsch/DDCuebersichten/ddcUebersichten_no_de.html#doc259496bodyText3	https://science-matrix.com/sites/default/files/matrix/sm_journal_classification	
2	Datensatz_ID*	Titel des Forschungsdatensatzes*	Zusammenfassung/Beschreibung des Forschungsdatensatzes*	Forschungsfrage*	Forschungsgegenstand*	Untersuchungseinheit*	Schlagwörter zum Forschungsdatensatz	Sprache des Forschungsdatensatzes	DDC-Klassifikation*	ScienceMatrix*
3	3.1	Basisdaten politische Entscheidungen	Der Datensatz beinhaltet die bereinigten Datensätze anderer Institutionen als Grundlage	Welche politischen Entscheidungen wurden in der Pandemie getroffen?	Maßnahmen mit Nachkodierung in Ordinalskalierung und Indizes	Maßnahme / Zeiteinheit	Politische Maßnahmen; Political decisions; Corona Regulation;	englisch		300 Political Science & Public Adr
4	3.2	kodierte Reden	Sammlung aller Reden an die Nation der Regierungsoberhäupter, die in 2020-2021 gehalten wurden. Die Reden wurden in quasi-Sätze aufgeteilt und entsprechend eines abgestimmten Codebooks numerisch kodiert	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Reden	Quasi-Sätze	Reden an die Nation; Reden; Corona; Speeches; MP; Head of State; Staatsoberhaupt; Kanzler; Ministerpräsident	verschiedene, je nach Landessprache		300 Political Science & Public Adr
5	3.3	Politische Entscheidungen erweitert	Erweiterte, nachkodierter Datensatz politischer Entscheidungen. Auf Basis einer Abwägung der Vor- und Nachteile einzelner existierender Datensätze wird ein eigener Datensatz aufgebaut.	Welche politischen Entscheidungen wurden in der Pandemie getroffen?	Maßnahmen mit Nachkodierung in Ordinalskalierung und Indizes	Paneldatensatz - Tag	Politische Maßnahmen; Political decisions; Corona Regulation;	englisch		300 Political Science & Public Adr
6	3.4	Gesamtdatensatz kodierte Reden	Überblick über alle kodierten Reden, indem die Anzahl der vergebenen Codes jeweils aufsummiert und so eine Vergleichbarkeit erreicht wurde	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Anzahl Codes	Reden an die Nation; Reden; Corona; Speeches; MP; Head of State; Staatsoberhaupt; Kanzler; Ministerpräsident	numerisch		300 Political Science & Public Adr
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										

Dokumentation der Forschungsdaten mit dem Metadaten sheet

1								https://www.dnb.de/DE/Professionell/DDC-Deutsch/DDCuebersichten/ddcUebersichten_no_de.html#doc259496bodyText3	https://science-matrix.com/sites/default/files/matrix/sm_journal_classification	
2	Datensatz_ID*	Titel des Forschungsdatensatzes*	Zusammenfassung/Beschreibung des Forschungsdatensatzes*	Forschungsfrage*	Forschungsgegenstand*	Untersuchungseinheit*	Schlagwörter zum Forschungsdatensatz	Sprache des Forschungsdatensatzes	DDC-Klassifikation*	ScienceMetric*
3	3.1	Basisdaten politische Entscheidungen	Der Datensatz beinhaltet die bereinigten Datensätze anderer Institutionen als Grundlage	Welche politischen Entscheidungen wurden in der Pandemie getroffen?	Maßnahmen mit Nachkodierung in Ordinalskalierung und Indizes	Maßnahme / Zeiteinheit	Politische Maßnahmen; Political decisions; Corona Regulation;	englisch		300 Political Science & Public Adr
4	3.2	kodierte Reden	Sammlung aller Reden an die Nation der Regierungsoberhäupter, die in 2020-2021 gehalten wurden. Die Reden wurden in quasi-Sätze aufgeteilt und entsprechend eines abgestimmten Codebooks numerisch kodiert	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Reden	Quasi-Sätze	Reden an die Nation; Reden; Corona; Speeches; MP; Head of State; Staatsoberhaupt; Kanzler; Ministerpräsident	verschiedene, je nach Landessprache		300 Political Science & Public Adr
5	3.3	Politische Entscheidungen erweitert	Erweiterte, nachkodierter Datensatz politischer Entscheidungen. Auf Basis einer Abwägung der Vor- und Nachteile einzelner existierender Datensätze wird ein eigener Datensatz aufgebaut.	Welche politischen Entscheidungen wurden in der Pandemie getroffen?	Maßnahmen mit Nachkodierung in Ordinalskalierung und Indizes	Paneldatensatz - Tag	Politische Maßnahmen; Political decisions; Corona Regulation;	englisch		300 Political Science & Public Adr
6	3.4	Gesamtdatensatz kodierte Reden	Überblick über alle kodierten Reden, indem die Anzahl der vergebenen Codes jeweils aufsummiert und so eine Vergleichbarkeit erreicht wurde	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Anzahl Codes	Reden an die Nation; Reden; Corona; Speeches; MP; Head of State; Staatsoberhaupt; Kanzler; Ministerpräsident	numerisch		300 Political Science & Public Adr
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										

Dokumentation der Forschungsdaten mit dem Metadaten sheet

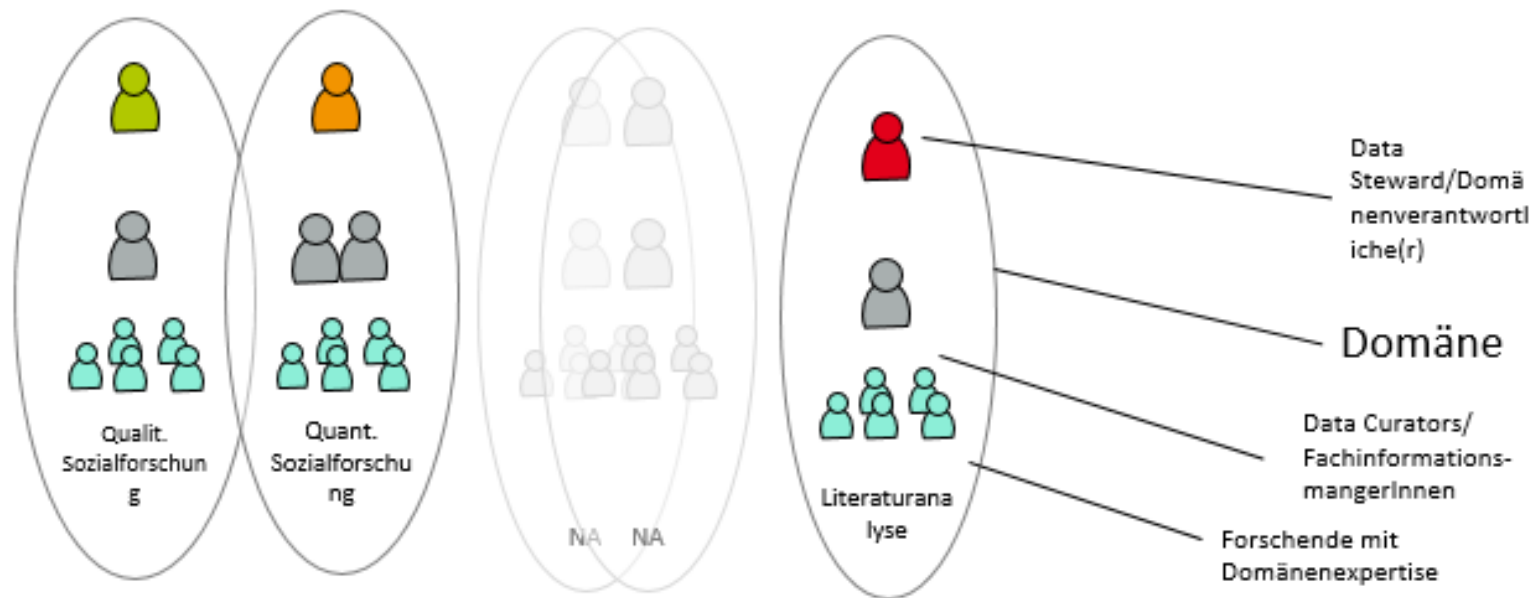
	Datensatz_ID*	Titel des Forschungsdatensatzes*	Zusammenfassung/Beschreibung des Forschungsdatensatzes*	Forschungsfrage*	Forschungsgegenstand*	Untersuchungseinheit*	Schlagwörter zum Forschungsdatensatz	Sprache des Forschungsdatensatzes	DDC-Klassifikation*	ScienceMetric*
1									https://www.dnb.de/DE/Professionell/DDC-Deutsch/DDCuebersichten/ddcUebersichten_no.html#dcr259496bodyText3	https://science-metric.com/sites/default/files/metric/sm_journal_classification
2										
3	3.1	Basisdaten politische Entscheidungen	Der Datensatz beinhaltet die bereinigten Datensätze anderer Institutionen als Grundlage	Welche politischen Entscheidungen wurden in der Pandemie getroffen?	Maßnahmen mit Nachkodierung in Ordinalskalierung und Indizes	Maßnahme / Zeiteinheit	Politische Maßnahmen; Political decisions; Corona Regulation;	englisch		300 Political Science & Public Adr
4	3.2	kodierte Reden	Sammlung aller Reden an die Nation der Regierungsoberhäupter, die in 2020-2021 gehalten wurden. Die Reden wurden in quasi-Sätze aufgeteilt und entsprechend eines abgestimmten Codebooks numerisch kodiert	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Reden	Quasi-Sätze	Reden an die Nation; Reden; Corona; Speeches; MP; Head of State; Staatsoberhaupt; Kanzler; Ministerpräsident	verschiedene, je nach Landessprache		300 Political Science & Public Adr
5	3.3	Politische Entscheidungen erweitert	Erweiterte, nachkodierter Datensatz politischer Entscheidungen. Auf Basis einer Abwägung der Vor- und Nachteile einzelner existierender Datensätze wird ein eigener Datensatz aufgebaut.	Welche politischen Entscheidungen wurden in der Pandemie getroffen?	Maßnahmen mit Nachkodierung in Ordinalskalierung und Indizes	Paneldatensatz - Tag	Politische Maßnahmen; Political decisions; Corona Regulation;	englisch		300 Political Science & Public Adr
6	3.4	Gesamtdatensatz kodierte Reden	Überblick über alle kodierten Reden, indem die Anzahl der vergebenen Codes jeweils aufsummiert und so eine Vergleichbarkeit erreicht wurde	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Anzahl Codes	Reden an die Nation; Reden; Corona; Speeches; MP; Head of State; Staatsoberhaupt; Kanzler; Ministerpräsident	numerisch		300 Political Science & Public Adr
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										

Dokumentation der Forschungsdaten mit dem Metadaten sheet

1								https://www.dnb.de/DE/Professionell/DDC-Deutsch/DDCuebersichten/ddcUebersichten_no_de.html#doc259496bodyText3	https://science-matrix.com/sites/default/files/matrix/sm_journal_classification	
2	Datensatz_ID*	Titel des Forschungsdatensatzes*	Zusammenfassung/Beschreibung des Forschungsdatensatzes*	Forschungsfrage*	Forschungsgegenstand*	Untersuchungseinheit*	Schlagwörter zum Forschungsdatensatz	Sprache des Forschungsdatensatzes	DDC-Klassifikation*	ScienceMatrix*
3	3.1	Basisdaten politische Entscheidungen	Der Datensatz beinhaltet die bereinigten Datensätze anderer Institutionen als Grundlage	Welche politischen Entscheidungen wurden in der Pandemie getroffen?	Maßnahmen mit Nachkodierung in Ordinalskalierung und Indizes	Maßnahme / Zeiteinheit	Politische Maßnahmen; Political decisions; Corona Regulation;	englisch		300 Political Science & Public Ad
4	3.2	kodierte Reden	Sammlung aller Reden an die Nation der Regierungsoberhäupter, die in 2020-2021 gehalten wurden. Die Reden wurden in quasi-Sätze aufgeteilt und entsprechend eines abgestimmten Codebooks numerisch kodiert	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Reden	Quasi-Sätze	Reden an die Nation; Reden; Corona; Speeches; MP; Head of State; Staatsoberhaupt; Kanzler; Ministerpräsident	verschiedene, je nach Landessprache		300 Political Science & Public Adr
5	3.3	Politische Entscheidungen erweitert	Erweiterte, nachkodierter Datensatz politischer Entscheidungen. Auf Basis einer Abwägung der Vor- und Nachteile einzelner existierender Datensätze wird ein eigener Datensatz aufgebaut.	Welche politischen Entscheidungen wurden in der Pandemie getroffen?	Maßnahmen mit Nachkodierung in Ordinalskalierung und Indizes	Paneldatensatz - Tag	Politische Maßnahmen; Political decisions; Corona Regulation;	englisch		300 Political Science & Public Adr
6	3.4	Gesamtdatensatz kodierte Reden	Überblick über alle kodierten Reden, indem die Anzahl der vergebenen Codes jeweils aufsummiert und so eine Vergleichbarkeit erreicht wurde	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Welchen Fokus legen Staatsoberhäupter in Reden an die Nation?	Anzahl Codes	Reden an die Nation; Reden; Corona; Speeches; MP; Head of State; Staatsoberhaupt; Kanzler; Ministerpräsident	numerisch		300 Political Science & Public Adr
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										

Das Konzept der Domänen

Domäne = Logische Gruppierung bestimmter Forschungsdaten nach deren Erhebungsmethode



Querschnittsaufgaben als projektinterne Dienstleistung:

- Kuratierte Methodenliste
 - Literaturverweise
 - Verweise zu KResCo Datensätzen
- Datenkurationsprofil
 - Metadaten
 - Dateiformate
- Beratung zu domänenspezifischen Datenthemen

Datenkurationsprofil der Domäne „Qualitative Sozialforschung“

Data Curation Profile – Datentabelle

Data Stage	Output	# of Files / Typical Size	Format	Other / Notes	Method (siehe Tabelle oben)	Sensibilitätsstufe
Primary Data						
Raw	Video	400-550 MB	MP4		Interview	
Processed	Text	200-300 KB	Word	Transkribiert	Transkription	
Analyzed			MAXQDA	Transkribiert, codiert, proprietäres Format		
Analyzed	Text		CSV (XLSX?)	Codiert	Codierung (Nr x)	
Analyzed	Text	-	RTF (TBD aus Word)	Analysiert und zusammengefasst	Text-Zusammenfassungen entlang des Codebaums	
Finalized	-	-	- TBD Finale Date-Transfer in Unsens-Dateien für interne und externe Veröff	Transferprozesse FDM hin zu unsens-Dateien		

Fazit

- Data Governance für ein Forschungsprojekt war neuartig.
- Auch für das Projekt und alle Beteiligten „neu“.
- Die Hauptaktivitäten im Projekt „KResCo“ fanden auf der Data-Governance-Ebene statt
- Viele Materialien/Vorlagen wurden erarbeitet
- Diese wurden noch nicht konsequent genutzt
- Das Datenmanagement muss „vorauslaufen“

Ausblick

- Der Data-Governance-Ansatz lässt sich erweitern: auf verschiedene Projekte, Organisationsteile, Organisationen und auf ganze Forschungs-Communities
- Die Strategie-Ebene ist noch ausbaufähig, sowie das Monitoring ist noch ausbaufähig
- Das Konzept lässt sich auf beliebige Forschungsprojekte anwenden
- Aus dem Projekt gehen zwei Masterarbeiten hervor:
 - Eine Masterarbeit stellt eine Anbindung zwischen dem Metadaten-sheet und dem Forschungsdaten-Repository Fordatis her
 - Eine Masterarbeit untersucht den Zusammenhang zwischen Forschungsdaten und Patenten

Kathrin Tuppi

Fraunhofer INT
Kathrin.Tuppi@int.fraunhofer.de
Tel: +49 (0)2251 18-316



Siegrid Hecht-Veenhuis

Fraunhofer INT
Siegrid.Hecht-Veenhuis@int.fraunhofer.de
Tel: +49 (0)2251 / 18-233



Wolfgang Beinhauer

Fraunhofer IAO
Wolfgang.Beinhauer@iao.fraunhofer.de
Tel: +49 (0)711 / 970-2376



Andrea Wuchner

Fraunhofer IRB
Andrea.Wuchner@irb.fraunhofer.de
Tel: +49 (0)711 / 970-2714



Johannes Sautter

Fraunhofer IAO
Johannes.Sautter@iao.fraunhofer.de
Tel: +49 (0)711 / 970-2387



Daniela Beyer

Fraunhofer ISI
Daniela.Beyer@isi.fraunhofer.de
Tel: +49 (0)721 / 6809-187

