

Analyse des Forschungsoutputs als Grundlage für Forschungsdatenmanagement



Thomas Seyffertitz & Michael Katzmayr
Wirtschaftsuniversität Wien
Universitätsbibliothek

107. DEUTSCHER BIBLIOTHEKARTAG, BERLIN, 12.-15.6.2018



- **Ausgangslage**
- **Motivation** und **Ziel**
- **Methodische Vorgehensweise**
- **Ergebnisse**
- **Schlussfolgerungen**

- **Akademische Einheiten:** 11 Departments (60 Institute), 15 Forschungsinstitute, 6 Kompetenzzentren
- **Wissenschaftliches Personal** ca. 770 VZÄ
- produziert jährlich **ca. 600 Beiträge** in wiss. Fachzeitschriften
- Noch **kein** durchgehend **institutionalisiertes Forschungsdatenmanagement**
- **FDM** als **neues „Geschäftsfeld“** in der UB (Teil der Strategie 2016-2018)
- Auftrag Vizerektor für Forschung → **FDM Konzeptentwicklung**

Motiv

Kenntnis der **Forschungsdatenlandschaft** eine wichtige Voraussetzung für Forschungsdatenmanagement

Zielsetzung

Forschungsdatenlandschaft: „*Welche Daten haben wir an der WU?*“

Forschungsdaten: Umgang, Erfahrungen, Trends, Bedürfnisse forschenseitig

Idee zur Umsetzung

- Analyse des **Forschungoutputs**: Dokumentenanalyse der wiss. Zeitschriftenartikel (quantitative Komponente)
- Umgang und Erfahrungen mit FD entlang Forschungszyklus
semi-strukturierte Interviews mit ForscherInnen (qualitative Komponente)

Methodische Vorgehensweise Teil 1

Dokumentenanalyse – Zeitschriftenartikel

Zitat	Beze
32 Hofmann, Eva, Penz, Elfriede. 2016. The moral decision-making process of unauthorized downloading. <i>International Journal of Business Environment</i> 8 (4), 385-407.	NO
36 Collien, Isabel, Sieben, Barbara, Müller-Camen, Michael. 2016. Age work in organizations: Maintaining and disrupting institutionalized understandings of higher age. <i>British Journal of Management (BJM)</i> 27, 778-795.	NO



The image shows a research article page with several sections highlighted in blue boxes. These sections are labeled 'Methodische Vorgehensweise', 'Erhebung', 'Auswertung', and 'Diskussion'. Two blue arrows point from the citation table on the left to the 'Methodische Vorgehensweise' and 'Erhebung' sections, indicating the flow of information from the citation to the full text.



						empirisch	quantitativ	qualitativ	Mix	explorativ	deskriptiv	explorativ	explorativ	explorativ	explorativ	explorativ	explorativ	explorativ	explorativ
Quantitative Methoden	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Qualitative Methoden	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0
Mix	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
empirisch	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
quantitativ	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
qualitativ	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Mix	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
explorativ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
deskriptiv	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
explorativ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
explorativ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
explorativ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Artikelzitat aus Forschungsinformationssystem „FIDES“:

- ID, Artikelzitat
- akademische Einheit
- DOI → Volltext

Extraktion der Informationen zu Daten aus Volltext

- Liegen Daten vor?
- Autorenschaft, Fördergeber, Verlag
- Kontextinformationen zu Daten
- Forschungsmethoden
- Datenformen
- Worüber sind Daten (Domänen)
- Formattypen (Zahlen, Bild, ...)

Erfassung/Aufbereitung in Datenmatrix

- Kodierungsschema
- Variable und abgeleitete Variablen zu FD
- deskriptive qualitative Merkmale
- Datenbereinigung

Dokumentenanalyse – Zeitschriftenartikel

Ergebnisse allgemein

Merkmale	WU		SoWi	
	abs.	rel.	abs.	rel.
Forschungsdaten	250	41,95%	247	57,18%
Keine FD	346	58,05%	185	42,82%
Summe alle Artikel	596	100,00%	432	100,00%
Fördergeber	126	21,14%	110	25,46%
Ohne Fördergeber	470	78,86%	322	74,54%
Summe alle Artikel	596	100,00%	432	100,00%
Datenform	WU		SoWi	
	abs.	rel.	abs.	rel.
nur quantitative FD	141	56,40%	139	56,28%
nur qualitative FD	35	14,00%	34	13,77%
beide Formen	74	29,60%	74	29,96%
Anzahl Artikel mit FD	250	100,00%	247	100,00%

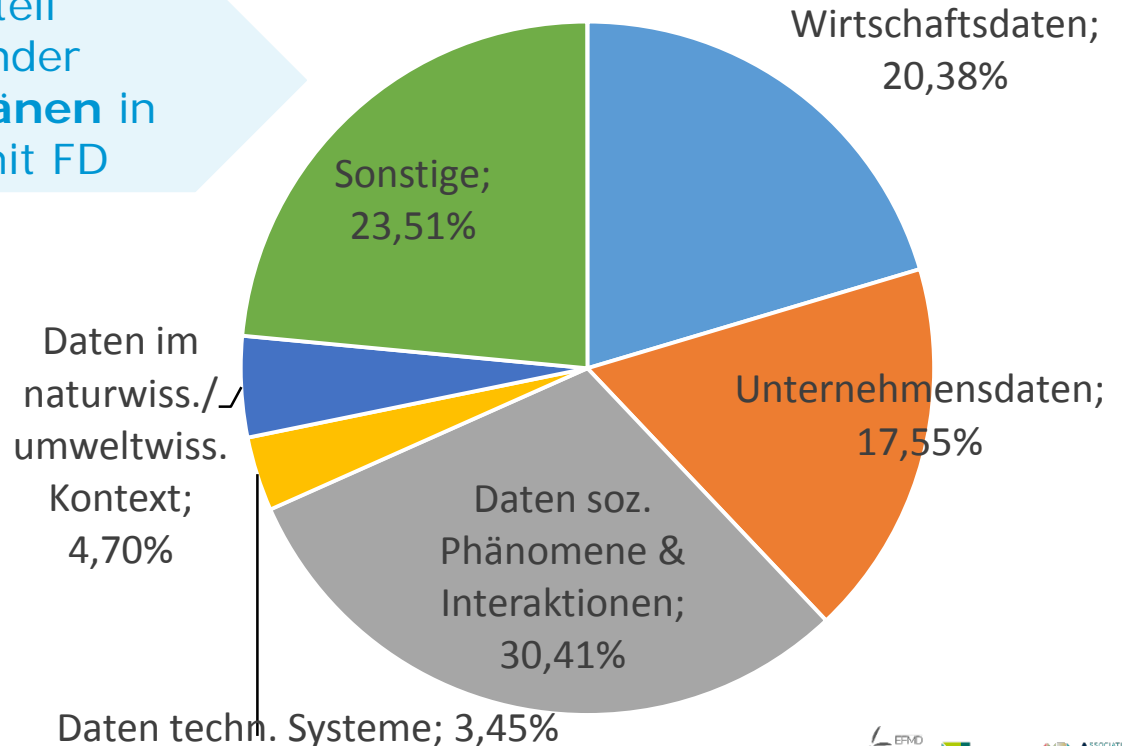
- ~80% ohne Förderung (SoWi: ~75%)
- aber Artikel ohne FD nur 11% gefördert
- Artikel mit FD ca. 1/3 gefördert
- SoWi: weit über 50% mit Forschungsdaten
- über 85% quantitative Forschungsdaten
- knapp 30% beide Formen

Dokumentenanalyse

Ergebnisse Datendomänen auf WU-Ebene

Rel. Anteil
vorliegender
Datendomänen in
Artikeln mit FD

Datendomänen	WU
Wirtschaftsdaten	65
Unternehmensdaten	56
Soziale Phän. & Interaktionen	97
Daten technischer Syst.	11
Daten im umwelt- naturwiss. Kontext	15
Sonstige	75
Häufigkeit über alle 250 Artikel (WU)	319

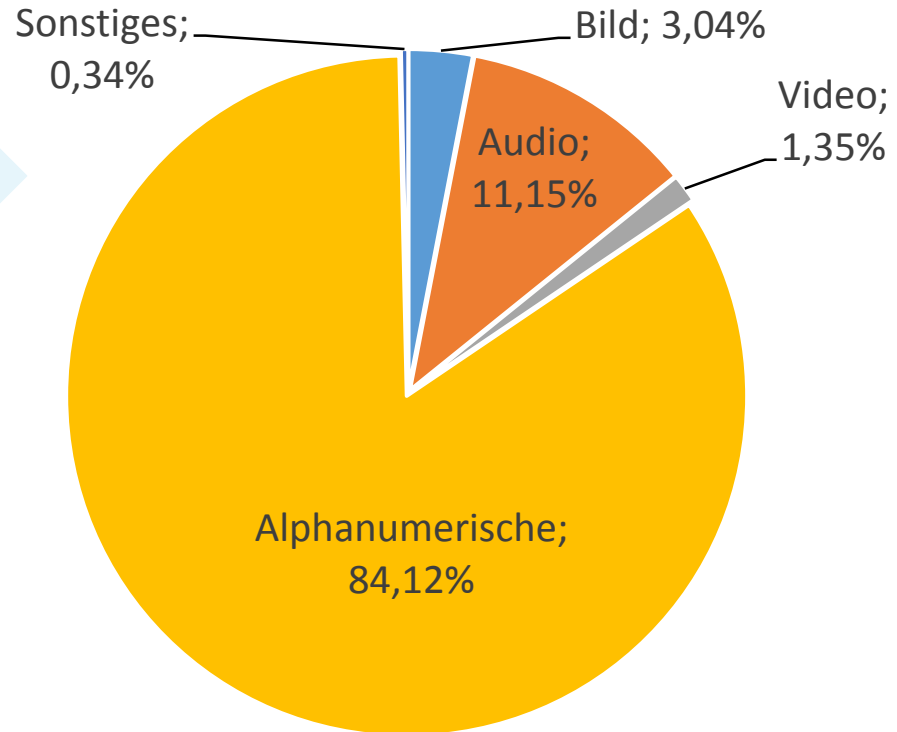


Dokumentenanalyse

Ergebnisse Datenformattypen auf WU Ebene

Rel. Anteil
vorliegender
Datenformattypen
in Artikeln mit FD

Datenformattypen	WU (abs.)
Bild	9
Audio	33
Video	4
Alphanumerische	249
Sonstige	1
Gesamthäufigkeit über alle 250 Artikel (WU)	296



Methodische Vorgehensweise Teil 2

Semi-strukturierte Leitfadeninterviews

Attraktivität von Interviews

Beliebtes und verbreitetes Instrument zur Exploration; Forschungsökonomische Argumente; **Inhalte** zu den Befragten **transportieren**

Typologie

Explorativ (qualitatives Paradigma), Systematisierend, Theoriegenerierend

Expertenwissen

Technisches Wissen; **Prozesswissen**; **Deutungswissen**

Auswahl der Befragten

unterschiedliche Fächer/Departments, befasst mit **FD**, **Hierarchiestufe** kein Kriterium; **Pre-Test**; 25 Interviews

Rollenzuschreibung im Interview

Experte anderer Wissenskultur; Teil der Vorbereitung: Rolle bewusst anlegen, um Diskurs zu steuern

Methodische Vorgehensweise Teil 2

Leitfaden – Forschungszyklus

- **Allgemeines**

Welche Daten im Fachbereich, Charakteristika der Daten, Methoden

- **Forschungsprozess**

DMP, Fördergeber, Datenverwaltung, Datenschutz, Datenverlust; Trends & Entwicklungen im Fachbereich

- **Publikationsprozess**

Verlage, Veröffentlichung von Daten

- **Nach Abschluss des Forschungsprozesses**

Archivierung, Datenmanagement bei internationalen Projekten; Teilen von Daten (Hemmnisse/Anreize)

- **Services rund um Forschungsdatenmanagement**

Bedürfnisse, Erwartungen an Uni, FD-Policy, Erfahrungen woanders

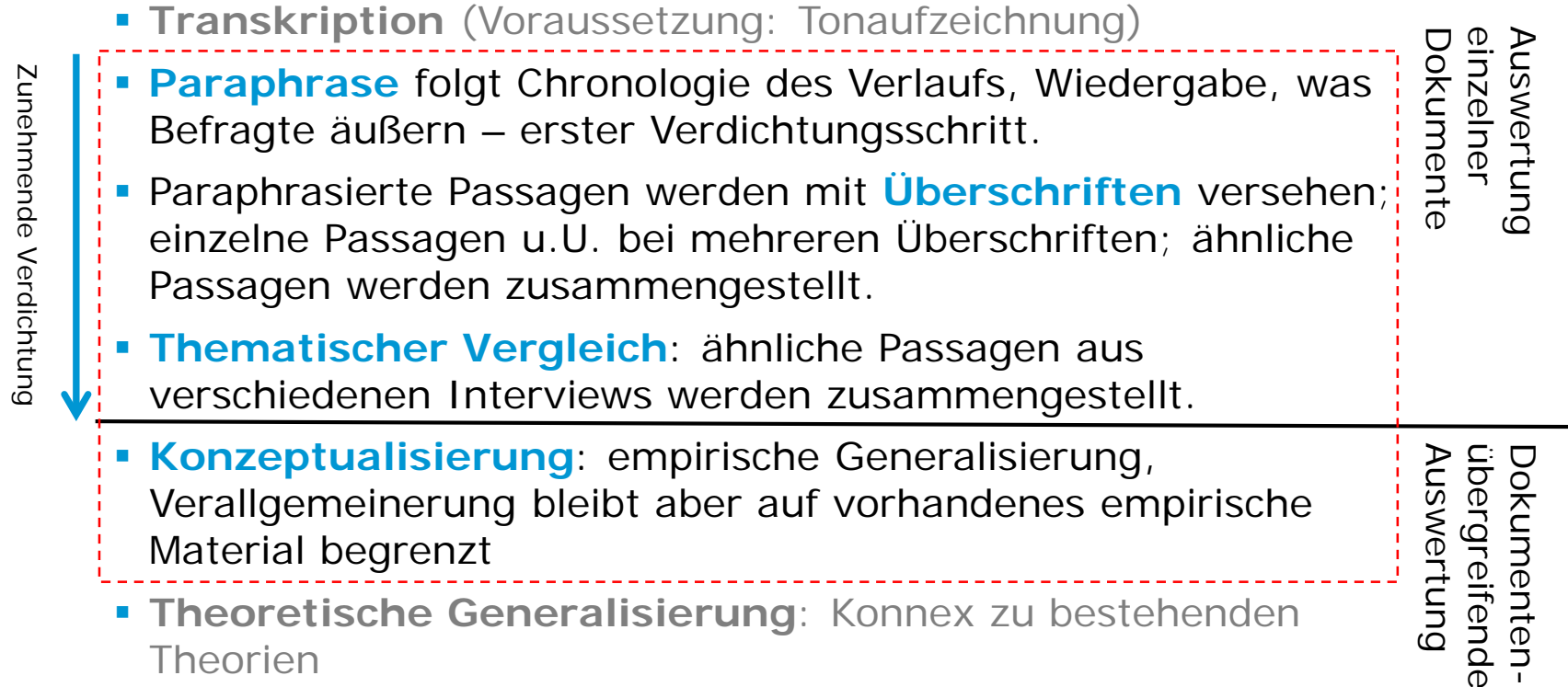
Auswertung Leitfadeninterviews

Generelle Überlegungen

- **25 Interviews** durchgeführt (26 ForscherInnen, ca. 60-70min)
- Auswertung orientiert sich an **thematischen Einheiten**, an inhaltlich zusammengehörige, **über einzelne Interviewtexte verstreute Passagen**
- Ziel ist es, das **überindividuell Gemeinsame** herauszuarbeiten
- Es stehen verschiedene Methoden der **qualitativen Inhaltsanalyse** zur Auswertung zur Verfügung (z.B. Beispiel: Meuser und Nagel (1991))

Auswertung Leitfadeninterviews

Beispiel: Methode nach Meuser/Nagel 1991



Ergebnisse Leitfadeninterviews (I)

DMP, Auflagen der Fördergeber

- Erfahrungen mit Anforderungen von Fördergebern
- DMP wurden für die *DFG*, *ERC*, *ESRC* und *Horizon 2020* erstellt

Trends/Entwicklungen hinsichtlich FD

- Quantitative Methoden: Big Data
- Quellen: zunehmend WWW, soziale Medien
- Speicherplatz und Rechenleistung

Verwalten von FD im Forschungsprozess

- Daten verstreut, Datenmanagementstrategie eher Ausnahme
- Externe Cloud-Systeme für Datenaustausch

FD im Publikationsprozess

- Quantitative FD: Data-Policies von ZS-Verlagen relevant, FD werden auch veröffentlicht
- Qualitative FD: eher keine Datenpublikation, keine Policies

Ergebnisse Leitfadeninterviews (II)

Langfristige Archivierung von FD

- findet manchmal statt, individuelle Herangehensweisen
- Beschreibung von FD ist sehr aufwändig – Nutzenfrage

Erfahrungen mit Datenverlust

- Kommt durchaus vor bei schon bearbeiteten Daten
- Unwiederbringlicher Verlust wurde selten berichtet

Teilen & Nachnutzung von FD

- Nachnutzung eine relevante Option (Reproduzierbarkeit)
- Daten nach einmaliger Verwendung „verbraucht“
- Weitergabe von FD wird eher positiv gesehen

FD-Policy

- Unterschiedliche (disziplinäre) Kulturen, Sorge vor Überreglementierung, Aufwand

FD-Management als Service

- Frage nach Erfahrungen an anderen Unis wenig ergiebig
- Information/Beratung, Archivierung, technische Services

Handlungsempfehlungen

- **Entwicklung** einer **FD-Policy**: Bewusstseins-schaffung, Forschungskulturen; WU Ebene/DP Ebene
- **Webseite** zum Thema FD: Informationen entlang FD-Lebenszyklus; Fördergeber; Ansprechpersonen an WU
- **Anlauf- und Servicestelle**: Fragen beantworten/Verweis auf Ansprechpartner; Angebot von Services

Methodisch

- Teilautomatisierung Dokumentenanalyse
- Angaben zu Forschungsdatendaten im Forschungsinformationssystem

Danke für Ihre Aufmerksamkeit!