A two-stage model to reveal a university's research data landscape and faculty's research data practices WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

Thomas Seyffertitz & Michael Katzmayr

INTERNATIONAL CONFERENCE ON ECONOMICS AND BUSINESS



Vienna University of Economics and Business Characteristics & Timeline





2



Knowledge of the RD landscape is an important prerequisite for appropriate research data services

Objectives

Motivation

- Research data landscape: "What research data do we have at WU?"
- Research culture: how do researchers deal with their data; experience, data trends, needs

Research design & method

- Case study design mixed method approach
- Stage 1 Document analysis of research output → journal articles (quantitative aspects)
- Stage 2 Semi-structured interviews with researchers, designed along the lifecycle of research data (*qualitative aspects*)

Stage 1 – Document Analysis Analysis of Journal Articles



Zitat	Bezeichnung DP, FI (einzeln)	DOI	Volitext (0/1)	4 Verla	
Gasser, Stephan, Rammerstorfer, Margarethe, Weinmayer, Karl. Forthcoming, Markowitz Revisited: Social Portfolio Engineering, European Journal of Operational Research (EJOR) ,	DFAS	<u>http://dx.doi.org/10.1016/i.eior.2016.10.043</u>	1	Elsevier	Empir Entwi ESG s Thom a ma: all st the fi in ga SP(s), -> nu
Kastner, Gregor. 2016. Dealing with Stochastic Volatility in Time Series Using the R Package stochvol. Journal of Statistical Software 69 (5): 5. 1-30.	DFAS	http://dx.doi.org/10.16637/iss.v669.i05		Condati on for Open Access Statistics (FOAS)	1. We des R Dater 2. R-C Finar
Malsiner-Walli, Gertraud, Frühwirth- Schnatter, Sylvia, Grün, Bettina. 2016. Model-based clustering based on sparse finite Gaussian mixtures. Statistics and Computing 26 (1): S. 303- 324.	DFAS	http://dx.doi.org/10.1007/s11222-014-9500-2	1	Springer	1. Kür Simul 2. Par 3. Kra Krabt 4. Iris

Reather J. Lawrence Ohio University, Alberty, Ohio, USA, and Christopher R. Moberg rhoting Datastment, Ohio University, Athene, Ohio, USA you plane which Kanal's placements, and weakers toorthing red. During the full concepts, then prime inducted they compared this paper into possible for exact, in terms line tensors, from that new whether poster members the solar well 300 tensors and improve the performance using and reasoning previous version performance. in regardings and our Sold-based based documents The second secon signing such a sufficient descent and a sufficient suff htt 2009, 2007, 2008, 2009 and 2009 presentation one of the History Senior, page projects, Lynnis Darig considerit er in Catorila Research Lane The second state data designed a replica Property many of the first housing they competing the house the promarkle of hand dama is the comparison profiled interior was der preimtage of leden place inking 1. A. D. Martin and Statements and Address of the second statements of the second statement of the ing Sports, Marketing, Them working: a last over a design of the residual of the last HE III and the second s for he lotted, when your plan, littlend and Marketing Department, Oh Correct Parigie and Publical Revelanote the south to be \$2 entering to the paint model have madel a the part of some

Abstract About the spectral sp

Profess - The party conversed, for the present differences in particular and the difference in provide a number of the The second second second second difference is a second sec Departmenter, Altraditionatori esplates devel à reservicembientes par este integrité de la définitione departer de la définition de la définit

at the 'P' white reading a minimal the particular is in the other or many



			Foerder		B. IW. DO	1			empirisch	duantation	qualitativ	selb st	frend	makročk.	untem.	201	tech. Syst	UmcoNact
¥	-	-	-	*	-	-	-	~	-	-	-	-	-	Ŧ	-	-	-	Ŧ
G 375 od.	Rantitative Methoda: Mathematisches Model; Modelmodifikation wird empirisch analysiert (Portfolio-Optimierung mittels Simulation);	0	0	1. Thomson Reuters Equity Global Index 2. Thomson Reuters ASSET4 ESG database	1	3	3	Datenset produziert	1	1	0	0	1	0	1	0	0	0
	Gunoucine mathematisches Modell, Buyes, che Simulation; MCL C	0	0	Datenset: European Central Baak's Statistical Data Warchouse - Dally bilateral prices of one recein 23 currencies from January 3, 2000 until April 4, 2012.	1	1	1	Sample selbst erstellt	1	1	0	0	1	1	0	0	0	0
	Quantitative Methode; MCMC Schötzung; Modell- (Performance)eralustion mittels Simulation mit künstlich erzeugten Daten, Bayessches modellbasiertes Clustering	FWF V170-N18	1	1. Selbsterzeugte künstliche Datensötze für Simulation 2. Krabben-Datensatz (R-Package MASS)	0	1	3	0A-Artikel	1	1	1	1	1	0	0	0	0	1
	Quantitative methode; Formal-analytisch; Graphische Analuse v.	National Institutes of Health—Nationa	1		1	1	3											





Stage 2 – Semi-structured Interviews Rationale & Methodological Aspects



- Expert interviews can be very effective
- Awareness of certain topics can be increased among the interview partners
- Objective: exploring expert knowledge (in terms of technical and process knowledge)
- Sample building criteria
 - Experience with data driven research
 - Covering all departments
 - Junior- and senior-researchers included
- Interview/Topic-guide: designed along the researcher's day-to-day research work and lifecycle of research data
- Pre-test & 25 interviews

5



Results – Analysis of the Journal Articles General



Attributes WU-level SoWI-level abs. rel. abs. rel. **Research data** 250 41,95% 247 57,18% No research data 346 58,05% 185 42,82% Total 596 100,00% 432 100,00% Funder 21,14% 25,46% 126 110 No funding 470 78,86% 322 74,54% 100,00% 100,00% Total 596 432 Data form WU-level SoWi-level abs. rel. abs. rel. Quantitative RD only 56,40% 139 56,28% 141 Qualitative RD only 35 14,00% 34 13,77% Both forms 74 29,60% 74 29,96%

250

100,00%

247

100,00%

Total

- Analysed 596 articles published in 2016
- ~80% without external/third-party funding
- Only 12% funding of papers without RD
- Whereas >33% of articles containing RD received funding
- 86% quantitative RD
- Almost 30% contained both quantitative and qualitative RD

Results – Analysis of the Journal Articles What the data are about





Results - Analysis of Journal Articles Data Format Types





UNIVERSITY O ECONOMICS AND BUSINES

Analysing the Interviews Following Meuser and Nagel (1991, 2009)



- Transcription
- Paraphrase sequencing of the text according to Analysis of a thematic units single interview **Coding** – ordering paraphrased passages thematically along our topic-guide **Thematic comparison** – grouping comparable Analysis passages from interviews across multiple **Conceptualization** - generalizing restricted to the interviews empirical data
 - Theoretical generalization



Semi-structured Interviews - Findings



DMP, research funders	 Some experience existing with funder mandates DMPs have been used for projects funded by DFG, ERC, ESRC and Horizon 2020
Data trends and developments	 Quantitative research methods: Big Data Source: increasingly WWW, Social media Storage; computer performance
Managing RD within the research process	 Data are scattered, data management strategy within departments rather an exception Use of external cloud-systems
RD in the publication process	 Quantitative RD: Publishers' data-policies relevant Qualitative RD: data usually not published; no policies (e.g. sociology)



Semi-structured Interviews – Findings ctd.

Archiving RD	In a pragmatic and case-based wayeffort required for the accurate description of RD
Data loss	Respondents have already experienced data lossRelated to processed RD
Sharing & Reuse	Most regard reuse as relevant (reproducibility)Sometimes data "used up"; dependent on discipline
Research data- Policy	Different (discipline-specific) culturesSome fear over-regulation
RDM as a service	 Awareness-raising measures, one-stop shop Need for advice; information & technical services

UNIVERSITĂT WIEN VIENNA UNIVERSITY OF ECONOMICS AND BUSINESS

- WIRTSCHAFTS UNIVERSITÄT WIEN VIENNA UNIVERSITY OF ECONOMICS
- Developing a RDM-Policy: awarness, research cultures; optional implementation at department level
- 2. Online material at the university's website summarizing relevent information
- 3. Single point of contact: providing information on RDM-services, referring to other sources etc...



Developing a RDM-Policy for WU Several Important Aspects



Organisational and contentrelated issues

- *Single policy* for the entire university <u>or</u> *framework* with optional *customization* at the academic department
- Recommending or more regulating character



Basic orientation when developing a first draft

- Based on state-of-the-art examples
- Taking into account empirical findings and the research cultures at WU
- Middle ground between directive specifications and non-binding recommendations
- The draft defines an ethical standard in dealing with research data and is therefore deliberately designed as a policy and not as guideline



Selected Literature



Methodology

- Bogner A., Menz W (2009): The Theory-Generating Expert Interview. Epistemological Interest, Forms of Knowledge, Interaction. In: A. Bogner, B. Littig and W. Menz (eds.): *Interviewing Experts*. London [UK], Palgrave Macmillan, pp. 43–80.
- Meuser M., Nagel U. (2009): The Expert Interview and Changes in Knowledge Production. In: A. Bogner, B. Littig and W. Menz (eds.): *Interviewing Experts*. London [UK], Palgrave Macmillan, pp. 17–42.
- Pickard J., Childs S (2013): *Research methods in information*. London, Facet.
- Yin R. K. (2017): Case study research and applications design and methods. London, SAGE.

Conceptualising data

 Kitchin, R. (2014): The data revolution: big data, open data, data infrastructures & their consequences. Thousand Oaks [CA], SAGE.





Thank you for your attention! Questions?



15