

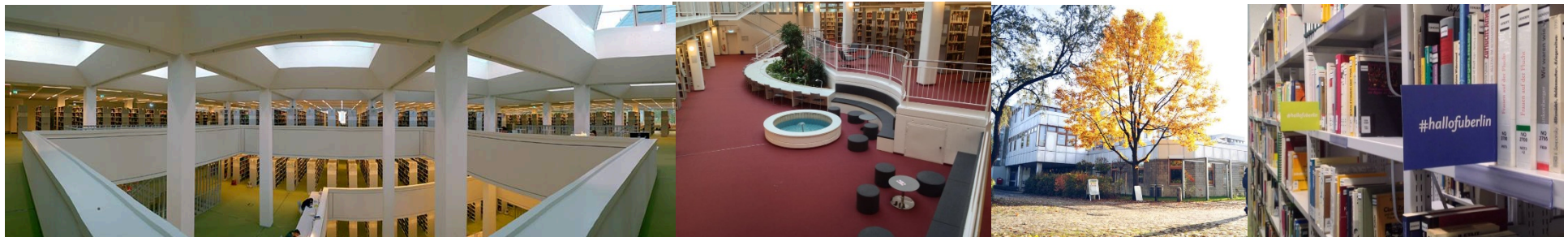
GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

„Kleine Fächer – große Forschungsdaten“- lessons learned im BMBF-Projekt „FDM_OAS-Orient“

Dr. Cosima Wagner
Universitätsbibliothek / Campusbibliothek



Gliederung

- 1. Das BMBF-Projekt FDM_OAS-Orient: lokaler Hintergrund und Skizze**

- 2. Methode Experteninterviews für FDM-Bedarfserhebungen: *lessons learned***

- 3. Diskussionsbeitrag zu den Workshop-Fragen:**
 - **Braucht jede Hochschule Ihre eigene Umfrage?**
 - **Wie lassen sich konkrete Maßnahmen aus den Bedarfserhebungen ableiten?**
 - **Desiderata**

1. Das BMBF-Projekt FDM_OAS-Orient: lokaler Hintergrund und Skizze

Konzeptphase für den Aufbau neuer Forschungsinfrastrukturen an der Freien Universität Berlin

1. Das Bibliotheksintegrationsprojekt (“24in1”, 2012-2015) → Campusbibliothek

→ Neues Bibliotheksmanagementsystem ALMA (Frontend PRIMO, Anbieter ExLibris) ist nicht für NLS* optimiert

2. Fusion von zwei zuvor getrennten Organisationseinheiten: Universitätsbibliothek und Center für Digitale Systeme (CeDiS) (seit 1/2018)

- Chance zur Entwicklung einer neuen digital scholarship / e-Research Infrastruktur an der FU
- Abteilungsübergreifende AG zum FDM (ExpertInnen zu Digital Humanities, Open Access, IT and Forschungs-Administration / Drittmittel)
- Steigendes Interesse an DH und FD, aber noch keine institutionalisierten Austauschplattformen am Campus



*Abk. NLS = nicht-lateinische Schriften



1. Das BMBF-Projekt FDM_OAS-Orient: lokaler Hintergrund und Skizze

FDM OAS-Orient: Projektziele



Konzeption von FDM-Services der Campusbibliothek, welche Forschung und Lehre der CB-Fächer mit nicht-lateinischen Schriften / internationalem Fokus über gesamten FD-Lebenszyklus begleiten

(Schwerpunkt: Ostasienwissenschaften, Orientwissenschaften, Altertumswissenschaften)



Erprobung der Konzeption von exemplarischen Metadaten-Leitfäden für Fächer mit Forschungsdaten in nicht-lateinischen Schriften



Unterstützung der gesamtuniversitären Strategie zur Etablierung von Forschungsdatenmanagementstrukturen mit der exemplarischen Erfassung von fachspezifischen Bedürfnissen in den Geistes- und Regionalwissenschaften und Vernetzung mit externen Akteuren in dem Bereich

1. Das BMBF-Projekt FDM_OAS-Orient: lokaler Hintergrund und Skizze

Daten FDM-Pilotprojekte

Sinologie	Digitalisierte Archivmaterialien ca. 160 GB JPEG, TIFF, PDF, Word
Japanologie	Digitalisierte Archivmaterialien
	Videomaterial; Blogbeiträge
Koreastudien	51 gedruckte kommentierte Quellensammlungen mit ca. 3800 Dokumenten/ digitalisierten Archivmaterialien
Altorientalistik	5 TB Bilder, Geländemodelle, MySQL-Datenbanken



FDM OAS-Orient: Ziele & Projektabschnitte

Bedarfserhebung
Experteninterviews

Erstellen von
Datenkurations-
leitfäden

Workshop
Publikation
Vernetzung

- Fachspezifische Bedarfe identifizieren
- Sensibilisierung von WissenschaftlerInnen und Informationsinfrastrukturen für die Herausforderung „nicht-lateinische Schriften“ und Forschungsinfrastrukturen
- Konzeption eines FDM-Services für Forschung und Lehre in der Bibliothek

2. Methode Experteninterviews für FDM- Bedarfserhebungen: *lessons learned*

2. Methode Experteninterviews für FDM-Bedarfserhebungen: lessons learned

Methode: ca. 2-stündige semi-strukturierte Experteninterviews in den Büros der Forschenden vor Ort

6 Befragte aus 5 verschiedenen Forschungsprojekten am Fachbereich Geschichts- und Kulturwissenschaften an der Freien Universität Berlin

Darunter Professor/innen, Projektleiter/innen, Wissenschaftliche Mitarbeiter/innen, Informatiker

Aufzeichnung der Interviews und Transkription

Disziplinen: Japanologie, Sinologie, Koreastudien, Altorientalistik

Projektstadien: Von Antragsstellung bis Abschlussphase

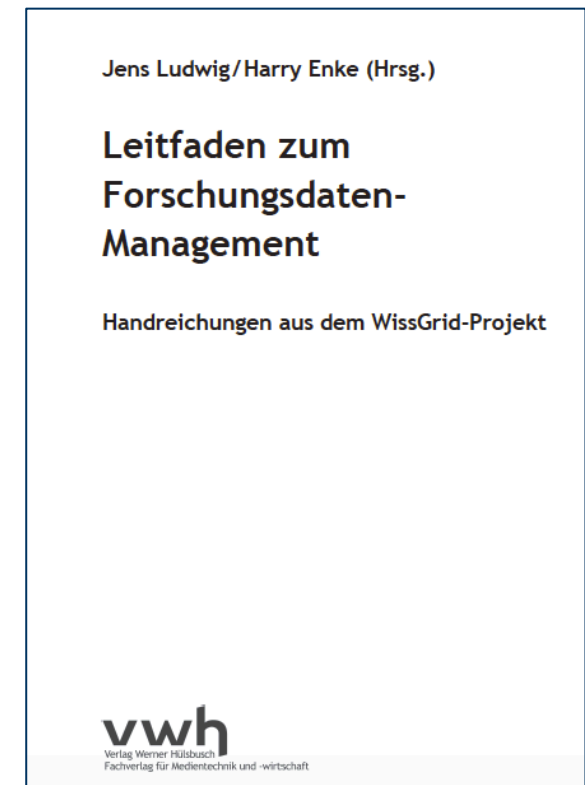
Zeitraum: Juli-September 2017

2. Methode Experteninterviews für FDM-Bedarfserhebungen: lessons learned

Interviewleitfaden:

Erstellung auf der Basis der „Checkliste zum Forschungsdatenmanagement“ von Enke, Harry; Ludwig, Jens (Hrsg.) (2013): Leitfaden zum Forschungsdaten-Management. Dortmund: Werner Hülsbusch. Glückstadt, S. 83-100

Themenfelder:



2. Methode Experteninterviews für FDM-Bedarfserhebungen: lessons learned

Vorhandene Forschungsdaten und deren Bearbeitung

Datentypen: Text-, Bild-, Audio-, Videodateien, Tabellen, 3D-Modelle, Datenbanken, Printmaterialien (Tagebücher, Archivalien, graue Literatur)

- 4 der 5 Projekte arbeiten mit Archivalien oder grauer Literatur. Für eine qualitativ-hochwertige, systematische Digitalisierung der Quellen sind Handreichungen zum fachgemäßen Vorgehen gewünscht.
- Aufgrund der Lage zB in Archiven in den Regionen vor Ort ist ein konsistentes, standardisiertes Vorgehen zum Teil nicht möglich → mangelnde (einheitliche) Qualität erschwert die Forschungsarbeit mit den Daten im Anschluss.
- Zum Teil wurden Fotokopien angefertigt und diese später eingescannt. Die Fotokopien liegen meist nur in Schwarz-Weiß vor und Details gehen verloren.
- OCR-Software für den Umgang mit Archivmaterialien in NLS ist (noch) nicht vorhanden/nicht bekannt.
- Nur in einem Projekt ist ein Mitarbeiter für die Datenverwaltung + Forschungssoftwareentwicklung angestellt, in allen anderen Projekten muss diese Aufgabe von DoktorandInnen / Postdocs neben der Forschung für die eigene wiss. Qualifizierung übernommen werden.

2. Methode Experteninterviews für FDM-Bedarfserhebungen: lessons learned

Fachspezifische Anforderungen

- Ethnologische/ sozialwissenschaftliche Forschung in den Ostasienwissenschaften: viele personenbezogene und somit sensible Daten (Archivalien, Interviews und Fragebögen, Ausgrabungstagebücher)
- In der Altorientalistik / Archäologie gefährdet Veröffentlichung von Geodaten die Unversehrtheit der Fundorte (Grabräuber etc.)
- Der „Zwischenschritt“ der fachgemäßen Digitalisierung von "analogen" Forschungsdaten bringt insbesondere bei der Feldforschung in den Regionen (Ostasien, Vorderer Orient) vor Ort eigene Herausforderungen mit sich und ist grundsätzlich zeitaufwändig.
- Daten müssen in verschiedenen Sprachen und Originalschrift beschrieben werden, damit sie für alle Interessierten/Fachcommunity auffindbar und nachnutzbar werden. Forschende sind auf ein verlässliches Daten-Retrieval angewiesen und müssen erkennen können, wie die jeweilige Suchmaschine eines Discovery-Systems/Repositoryums funktioniert (basiert sie auf cross-language information retrieval oder welche Schriftarten/Sprachen akzeptiert sie)

2. Methode Experteninterviews für FDM-Bedarfserhebungen: lessons learned

Fazit

- Die Auseinandersetzung mit FDM in den untersuchten Disziplinen steht noch am Anfang, nur eine der 6 befragten Personen war mit der Thematik vertraut und bereitet die Daten aus dem Projekt gemäß den DFG-Empfehlungen zur Nachnutzung auf.
- Alle Befragten sind am FDM interessiert und "sehr positiv" dem Thema gegenüber eingestellt, benötigen jedoch systematische Informationen und Unterstützung, um ein gutes FDM bei ihren eigenen Daten zu ermöglichen.
- Viele rechtliche Aspekte sind ungeklärt.
- Bei originalschriftlichen FD in nicht-lateinischen Schriften sind die Beschreibung (Metadaten) und das zuverlässige Retrieval besondere Herausforderungen

2. Methode Experteninterviews für FDM-Bedarfserhebungen: lessons learned

Fazit

- Beratung zu FDM, DH und Hilfe bei der Aufbereitung von Archivmaterialien bzw. grauer Literatur allgemein und für Forschungsdaten in nicht-lateinischen Schriften sind die wichtigsten Service-Wünsche, die von den Befragten geäußert wurden.
- Ebenfalls wichtig ist den Befragten eine klare, einheitliche Strategie der Universität hinsichtlich einer FDM-Policy.
- Vorhandene Ansprechpartner/innen und Service sind nicht ausreichend bekannt, es werden mehr Angebote/Schulungen von der Bibliothek gewünscht.

2. Methode Experteninterviews für FDM-Bedarfserhebungen: lessons learned

Beispiele für Fragestellungen der PilotprojektpartnerInnen

Datenaufbereitung

- Wie digitalisiert man Archivdaten am besten?
- Welche Schritte sind bei der Datenmodellierung zu beachten, z.B. OCR-Erfassung von vorliegenden pdfs, tiffs etc.? Aspekt Originalschrift?
- Welche Metadaten sind nötig? Wie sollen sie dokumentiert werden?
- Welche Transkriptionsstandards sollten eingehalten werden?
- Wie können Daten in multilingualen Umgebungen auffindbar gemacht werden?
- Wie können Daten untereinander verknüpft werden?
- Wie werden Daten interoperabel nutzbar (Projektpartner in Ostasien, Türkei)?

Wie sichert man sich in rechtlichen Fragen ab?

- Urheberrecht
- Persönlichkeitsrecht
- International gesammelte Daten
- Politische Situation kann Einfluss auf Informationszugang haben
- Dürfen Digitalisate veröffentlicht werden?
- **Wie können komplexe Präsentations- und Arbeitsumgebungen konserviert werden?**
- Gestaffelte Zugriffs-/Bearbeitungsrechte

2. Methode Experteninterviews für FDM-Bedarfserhebungen: lessons learned

Beispiele für Fragestellungen der PilotprojektpartnerInnen

Daten speichern/archivieren

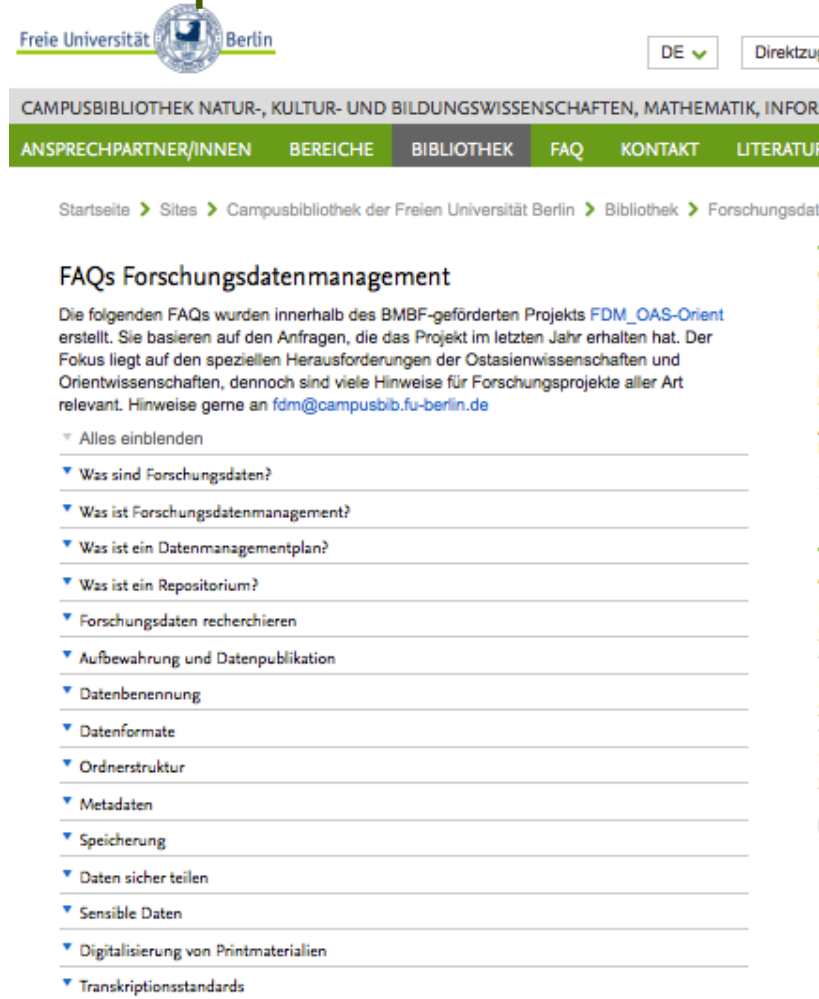
- Was ist bei Codierung zu beachten?
- Originalschriftliche Daten in der Volltextindizierung
- Originalschriftliche Metadaten
 - Schrift
 - Umschrift

Retrieval (Daten finden)

- Suchmaschinen für nicht-lateinische Schriften optimieren
 - Tokenisierung
 - Mapping verschiedener Schriften/Sprachen/Transkriptionen
 - Suchfelder für fachspezifische Beschreibungen

2. Methode Experteninterviews für FDM-Bedarfserhebungen: lessons learned

Von Rückfragen aus den Interviews zu FDM- und DH-Informationssseiten der Campusbibliothek



The screenshot shows the 'FAQs Forschungsdatenmanagement' page. It includes a navigation menu with 'ANSPRECHPARTNER/INNEN', 'BEREICHE', 'BIBLIOTHEK', 'FAQ', 'KONTAKT', and 'LITERATUR'. The main content area has a breadcrumb trail: 'Startseite > Sites > Campusbibliothek der Freien Universität Berlin > Bibliothek > Forschungsdaten'. The title is 'FAQs Forschungsdatenmanagement'. The text explains that the FAQs were created for the BMBF-funded project 'FDM_OAS-Orient' and focuses on challenges in Asian studies and orientation studies. It provides contact information: 'Hinweise gerne an fdm@campusbib.fu-berlin.de'. Below the text is a list of 16 expandable FAQ items, such as 'Was sind Forschungsdaten?' and 'Was ist ein Repositorium?'.



The screenshot shows the 'Digital Humanities Werkzeuge für Daten in nicht-lateinischen Schriften' page. It includes the same navigation menu as the first screenshot. The breadcrumb trail is: 'Startseite > Sites > Campusbibliothek der Freien Universität Berlin > Bibliothek > Forschung Daten'. The title is 'Digital Humanities Werkzeuge für Daten in nicht-lateinischen Schriften'. The text describes a list of digital tools for non-Latin scripts, noting it is under development and last updated in August 2018. Contact information is provided: 'Für Ergänzungen oder Hinweise kontaktieren Sie uns gerne unter fdm@campusbib.fu-berlin.de'. Below the text is a list of tool links: 'Alpheios', 'CATMA Computer Assisted Textual Markup and Analysis', 'Chunagon (中納言)', 'Cuneiform Analyser', 'JSesh', 'Juxta Collation Software', 'Juxta Commons (online Version)', 'KH Coder', and 'Mallet'.

<https://www.fu-berlin.de/sites/campusbib/bibliothek/Forschungsdatenmanagement/index.html>

2. Methode Experteninterviews für FDM-Bedarfserhebungen: lessons learned

Wie wesentlich sind Bedarfserhebungen für den Aufbau institutioneller Repositorien?

Im Falle des FU Repositoriums „ReFUBium“ bezüglich NLS nur im Stadium der „Awareness“ (Test von FD Retrieval in Originalschrift, ABER noch keine Umsetzung / Adaption)

- FD in Originalschrift sind nicht / nur teilweise recherchierbar/auffindbar
- Keine Anwendungsspeicherung möglich
- Gestaffeltes Rechtemanagement für die Sichtbarkeit von FD soll eingerichtet werden



The screenshot shows the homepage of the Refubium repository. At the top left is the logo of Freie Universität Berlin. To the right are links for 'Einloggen', 'Hilfe', 'Kontakt', and 'Impressum', along with a language dropdown menu set to 'Deutsch'. Below this is a navigation bar with 'REFUBIUM - REPOSITORY DER FREIEN UNIVERSITÄT BERLIN' and buttons for 'SUCHE', 'BROWSEN', and 'MEIN BENUTZERKONTO'. The main content area features a 'Willkommen im Refubium' section with a brief description of the repository's purpose and a list of 'Bereiche in Refubium' including 'Dissertationen', 'Forschungsdaten', 'Publikationen', and 'Serien und mehrbändige Werke'. To the right of the text is the 'Re|fub|ium' logo, which consists of the text 'Re|fub|ium' and a stylized geometric graphic of overlapping blue and green shapes.

<https://refubium.fu-berlin.de>

3. Diskussionsbeitrag zu den Workshop-Fragen

Braucht jede Hochschule Ihre eigene Umfrage? JA

Ja, weil

- es eine hervorragende Möglichkeit zum **differenzierten Kennenlernen der fachspezifischen Forschungsrealitäten** und –alltage ist (Interviews in Büros der Forschenden)
- es sehr gut ist, **um Vertrauen und einen kontinuierlichen Dialog** zwischen wissenschaftlicher Community und Infrastruktureinrichtungen aufzubauen → Grundverständnis aufbauen!
- es sehr gut ist, um das **Interesse für das Thema FDM** zu wecken / zu vertiefen
- auch Interviews mit wenigen ExpertInnen neben fachspezifischen Anforderungen viele **Hinweise auf weitere allgemeine Bedarfe zum FDM** ergeben
- jedes Projekt anders ist und es **nicht eine Lösung für alles** gibt
- aus den Interviews **schnelle Support-Services** entstehen können (Beratung INF-Antrag Drittmittelprojekt, Schulung DoktorandInnen, FAQs...)

3. Diskussionsbeitrag zu den Workshop-Fragen

Braucht jede Hochschule Ihre eigene Umfrage? JAIN

Nein, weil

- es neben fachspezifischen, lokalen Anforderungen auch eine Reihe an generischen Anforderungen gibt, zB Verwendungsmöglichkeit von Originalschrift in Discovery-Systemen allgemein, modulare Forschungsplattformen, Anwendungsspeicherung...
- Es zu übergreifenden Themen bereits sehr gute Positionspapiere und Handreichungen gibt, zB AG Datenzentren des DHd-Verands u.a.
- wiederholte Befragungen zu „survey fatigue“ führen

3. Diskussionsbeitrag zu den Workshop-Fragen

Desiderata: Wie lassen sich konkrete Maßnahmen aus den Bedarfserhebungen ableiten?

- **Niedrigschwellig/sofort:** Durch die Einrichtung Abteilungsübergreifender AGs, FAQ-Seiten, Schulungen in Graduiertenschulen, fachspezifischen und fachübergreifenden Workshops
- **Langfristig:** durch eine partizipative Forschungsinfrastrukturentwicklung
→weniger Umfragen, mehr „embedded“ sein in Forschungsprojekten (Institutionalisierung von fachspezifischen data librarian Services – oder Erweiterung Fachreferat?)
→Offenheit bei IT-EntwicklerInnen für Multilingualitätsproblematik
→Einbeziehung von IT-EntwicklerInnen bei den Experteninterviews vor Ort
- Etablierung einer Hochschulübergreifenden Austauschplattform zu FDM-Bedarfserhebungen, speziell Daten in NLS →Wo?
- „Mind-scripting“ Workshops mit IT-EntwicklerInnen? → Vorbild Science & Technology Studies (STS) (Doris Alhutter 2011)

3. Diskussionsbeitrag zu den Workshop-Fragen

Aktuelle Lektüre-Anregung (ERC Projekt „K-Plex“, <https://kplex-project.eu>):

Jennifer Edmond, Georgina Nugent Folan, Michelle T. Doran (2019) *'Reconciling the Cultural Complexity of Research Data: Can we Make Data Interdisciplinary without Hiding Disciplinary Knowledge'*. Data Science Journal www.tara.tcd.ie/handle/2262/83156 (preprint)

„For example, one of the major factors now inhibiting interdisciplinary data driven research is that of **how to capture provenance and facilitate the discovery, use, and reuse of discipline specific research 'data'**. All data are not not created equal, and **research disciplines are not all digital in the same way**.

Whilst Moedas (2017) sees the [EOSC] Cloud as a means of supporting those disciplines that may be 'lagging behind', the **reality is that such disciplines may be digital in ways that do not comply with the current conceptions of what makes digital data comprehensible, valuable, and reusable**. Indeed, they may resist digitisation and/or datafication.

Furthermore different researchers, research disciplines and institutions, **differ drastically over what they consider to be data in the first place.**“

Literaturhinweise

Allhutter, Doris (2012, online: 2011) Mind Scripting: A Method for Deconstructive Design. *Science, Technology & Human Values*, Bd. 37 (6), S. 684-707.

Asef, Esther; Gräff, Andreas; Lee, Martin; Wagner, Cosima (2018): Workshop-Bericht „Nicht-lateinische Schriften in multilingualen Umgebungen: Forschungsdaten und Digital Humanities in den Regionalstudien“. *DHd-Blog* vom 24.10.2018. <https://dhd-blog.org/?p=10669>

Dushay, Naomi: Discovery Grindstone (Blog) discovery-grindstone.blogspot.de/2014/01/searching-in-solr-analyzing-results-and.html

Edmond, Jennifer; Georgina Nugent Folan, Michelle T. Doran (2019): 'Reconciling the Cultural Complexity of Research Data: Can we Make Data Interdisciplinary without Hiding Disciplinary Knowledge'. *Data Science Journal*, URI preprint: <http://hdl.handle.net/2262/83156>

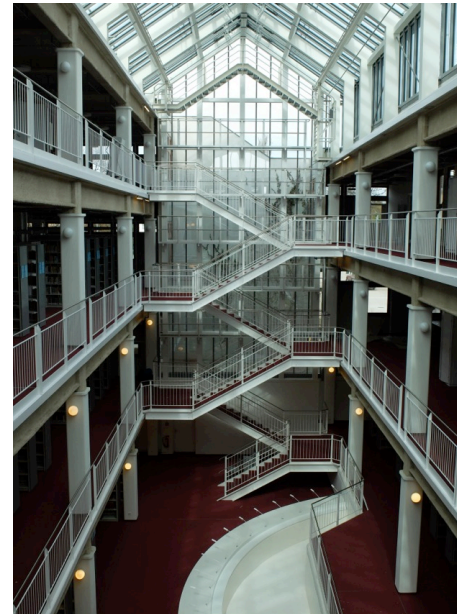
Enke, Harry; Ludwig, Jens (Hrsg.) (2013): *Leitfaden zum Forschungsdaten-Management*. Dortmund. Glückstadt: Werner Hülsbusch.

Halpern, Jack (2008?): "The Challenges of Intelligent Japanese Searching"; <http://www.cjk.org/cjk/joa/joapaper.htm>

Vogel, Iris (2014): „Phraseographie Japanisch–Deutsch: Lexikographische Perspektiven bei der Entwicklung digitaler Nachschlagewerke“ (<https://publikationen.uni-tuebingen.de/xmlui/handle/10900/52902>) Dissertation.

Ende

спасибо 谢谢
GRACIAS 谢谢
THANK YOU
ありがとうございました MERCI
DANKE धन्यवाद
شُكراً **OBRIGADO**



Aufbau und Erprobung von Strategien zum Forschungsdatenmanagement mit dem Schwerpunkt von Forschungsdaten in ‚nicht-westlichen‘ Sprachen, insbesondere aus dem Ostasiatischen Raum und dem Vorderen Orient („FDM_OAS-Orient“)

Martin Lee
m.lee@fu-berlin.de

Dr. Cosima Wagner
cosima.wagner@fu-berlin.de

Dr. Andreas Gräff
andreas.graeff@fu-berlin.de

Esther Asef
esther.asef@fu-berlin.de

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung

<http://www.fu-berlin.de/sites/campusbib/bibliothek/Forschungsdatenmanagement/index.html>