

HERBSTSCHULE 2016, Aufbaumodule

FORSCHUNGSDATEN

MODUL 8: Tools für das Forschungsdatenmanagement in Digital Humanities

Niklaus Stettler

Schweizerisches Institut für Informationswissenschaft (SII)

Zürich, 25.10.2016

Lizenz CC BY 2.5

All cartoons courtesy of Jørgen Stamp,
digitalbevaring.dk
CC BY 2.5



Ablauf

Einige Erfahrungen

Mögliche Dienste der Bibliothek

Infrastruktur für Forschungsdokumentation

Infrastruktur für Datenpublikation





Erfahrungen von TextGrid und DARIAH.de

- Die Akzeptanz der Angebote von Dariah.de ist durchzogen
- Zentrale Servicezentren haben sich nicht durchgesetzt
- Forschende der Digital Humanities arbeiten oft alleine oder in Kleingruppen
- Forschende wünschen Unterstützung / Tipps für Softwarenutzung
- Wenn sie sich auf neue Tools einlassen, stören sie sich nicht an Usability-Problemen
- Die Kommunikationsprobleme mit der Informatik sind nicht leicht zu überbrücken
- Die Bibliotheken sollten eine aktivere Rolle im FDM übernehmen



Erfahrungen mit Forschungsdatenmanagement

- Forschende nutzen besondere Dienste zur Verwaltung der Daten kaum und
 - Sie äussern kaum Bedarf für ein strukturiertes Datenmanagement
- Trotzdem sind die Forschenden mit aktueller Situation nicht zufrieden
 - Sie hätten gerne Datenmanagement für die eigene Forschung
 - Es gibt Bedarf zur Verknüpfung von Daten und Publikation (Throughput)
 - Forschende der Geistes- und Sozialwissenschaften legen grosses Gewicht auf Langzeitarchivierung

Was wir daraus lernen

- Forschungsdatenmanagement muss sich an Forschenden orientieren
- Diese haben Bedarf an Unterstützung mit Tools
- Möglichkeit zur Verwaltung der Daten muss verbessert werden

=> Bibliothek sollte Zugang zu Forschungsdaten gewährleisten.

Ablauf

Einige Erfahrungen

Mögliche Dienste der Bibliothek

Infrastruktur für Forschungsdokumentation

Infrastruktur für Datenpublikation





Forschende wünschen Unterstützung mit Tools

- Sehr viele Tools stehen zur Verfügung – doch es fehlt ein Überblick
 - Bibliothek sollte Forschende auf Tools aufmerksam machen
 - Sie kann Zugang zu Software gewähren
 - Sie sollte auch Zugang zu Schulungen (z.B. MOOCS, Videos) verschaffen
 - Wichtig: eine reine Tool-Liste reicht nicht. Forschende brauchen auch Beratung.
-
- Beispiele:
 - Jan Krause: A Selection of Research data Management Tools Throughout the Data Lifecycle EPFL, 2015: <https://infoscience.epfl.ch/record/211157>
 - TextGrid Hinweise auf Tools: <https://wiki.de.dariah.eu/display/TextGrid/Subject-specific+Tools+and+Services>
 - <https://www.dataone.org/all-software-tools>



Generische Tools für oft auftretende Aufgaben

- Verschiedene Methoden können in unterschiedlichen Projekten zur Anwendung kommen
- Bibliothek kann generische Tools zur Verfügung stellen
- Z.B.: Generische Tools für folgende Aufgaben :
 - Erstellen eines DataManagementPlans <https://dmptool.org/plans/19207.pdf>
 - Annotationen
 - Visualisierung
 - Georeferenzierung
 - Prosopographie
 - Diskursanalyse
 - Gemeinsames Schreiben

Beispiel Georeferenzierung

- In unterschiedlichsten Projekten kann der Raumbezug interessant sein
- Wenn Forschende einfache Möglichkeit haben, ihre Daten zu georeferenzieren, ergeben sich ev. neue Fragestellungen
- DARIAH stellt einen Geobrowser zur Verfügung – ein möglicher Dienst der Bibliothek
- <https://geobrowser.de.dariah.eu/edit/index.html#id=404651>



Load Data

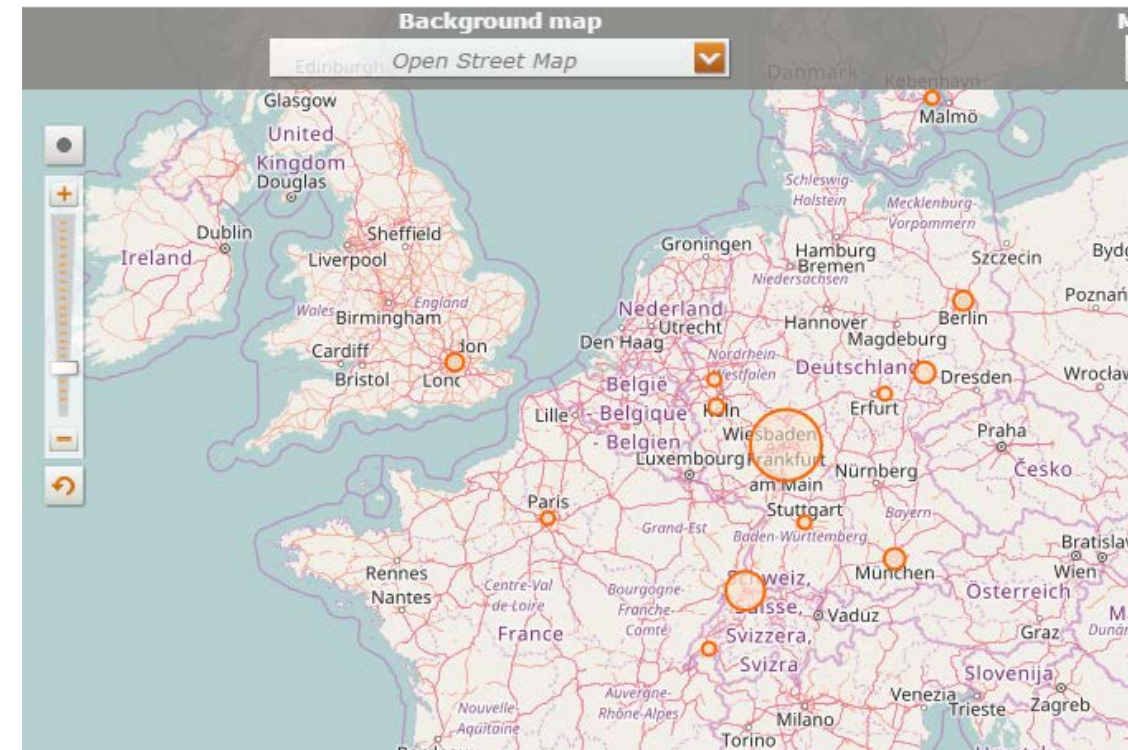
Static Data

Franz Kafka

load

Load Overlay

KML File URL



Beispiel Prosopographie

- Akteure sind in vielen Projekten wichtig
- Interessant, wenn Beziehungen zwischen den Personen dargestellt werden kann
- DARIAH stellt eine Personendatenbank zur Verfügung, die es erlaubt, unterschiedliche Aspekte zu den Akteuren zu erfassen
- Eine Personendatenbank (Archiv-Editor) für ein Institut kann viel neues Wissen generieren
- Forschende können sich gegenseitig Wissen zur Verfügung stellen
- <https://de.dariah.eu/personendatenrepositorium>

Archiv-Editor 2.3.14 lite

Datei Person Aspekt Quelle Ansicht Datenbank Einstellungen Hilfe

Aktueller Benutzer
admin admin (admin)

Aspekte-Ansicht Hannover, Ernst August von Leibniz, Gottfr...

Aktuell ausgewählte Quelle: Leibniz Universität Hannover
[Note: 3.9.13]

Kriterien für Baumdarstellung
Suche nach: Personen
Kategorie/Quellentyp: Themen

Suche

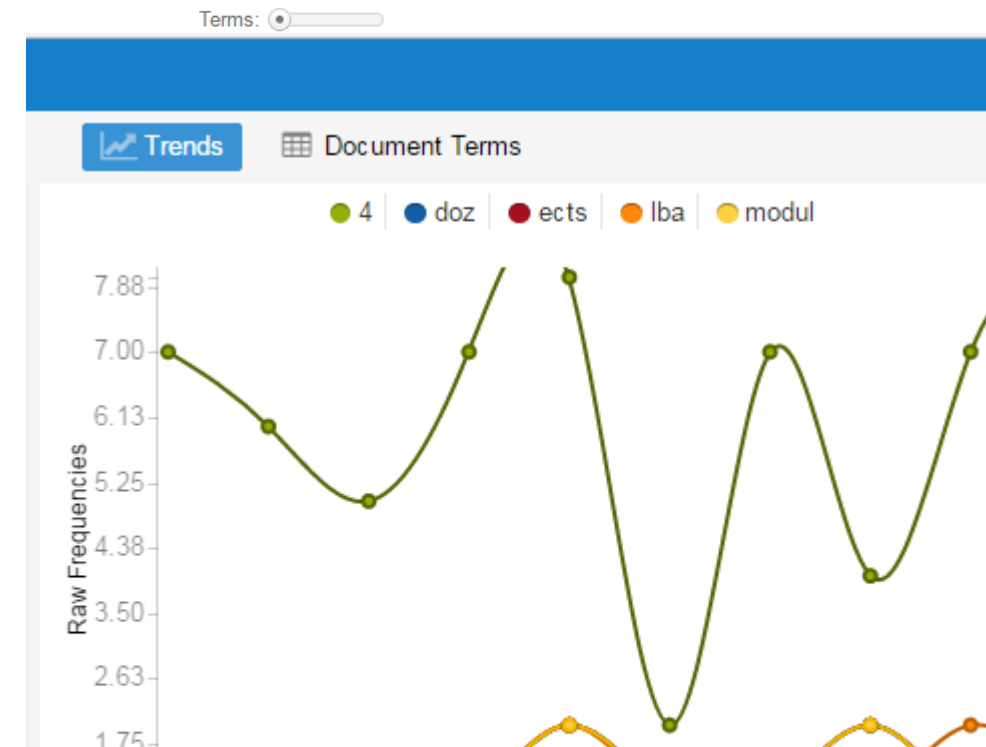
Alle Personen Alle Quellen Personen 3

- Altmitgliederverzeichnis der...
- Geburtsurkunde
- Holz, Hans Heinz: Leibniz: Das Leben eines Uni...
- Kuno, Fischer: Gottfried Wilhelm Leibniz: L...
- Leibniz Universität Hannover
- Mitgliederverzeichnis der Ro...
- Regionalgeschichte
- Sterbeurkunde
- Test, Test: dfgvfdg
- Titel
- Wikipedia

Subjekt \ Objekt	Braunschweig-...	Hannover, Ern...
Braunschweig-Lü...		
Hannover, Ernst...		
Hannover, Georg...		
Leibniz, Gottfr...	Angestellt von: 1676	Gefördert von: V...
Schönborn, Joha...		

Beispiel Textanalyse

- DARIAH stellt ein einfaches Werkzeug zur Analyse von Texten zur Verfügung
- <http://voyant-tools.org/>
- Daten aus verschiedenen Texten lassen sich leicht auswerten, z.B. in Wordclouds, Bubblelines, etc.
- Damit lassen sich sehr unterschiedliche Fragestellungen verfolgen: Begriffsverwendung in literarischen Texten genauso wie die Entwicklung von Diskursen im Laufe der Zeit.

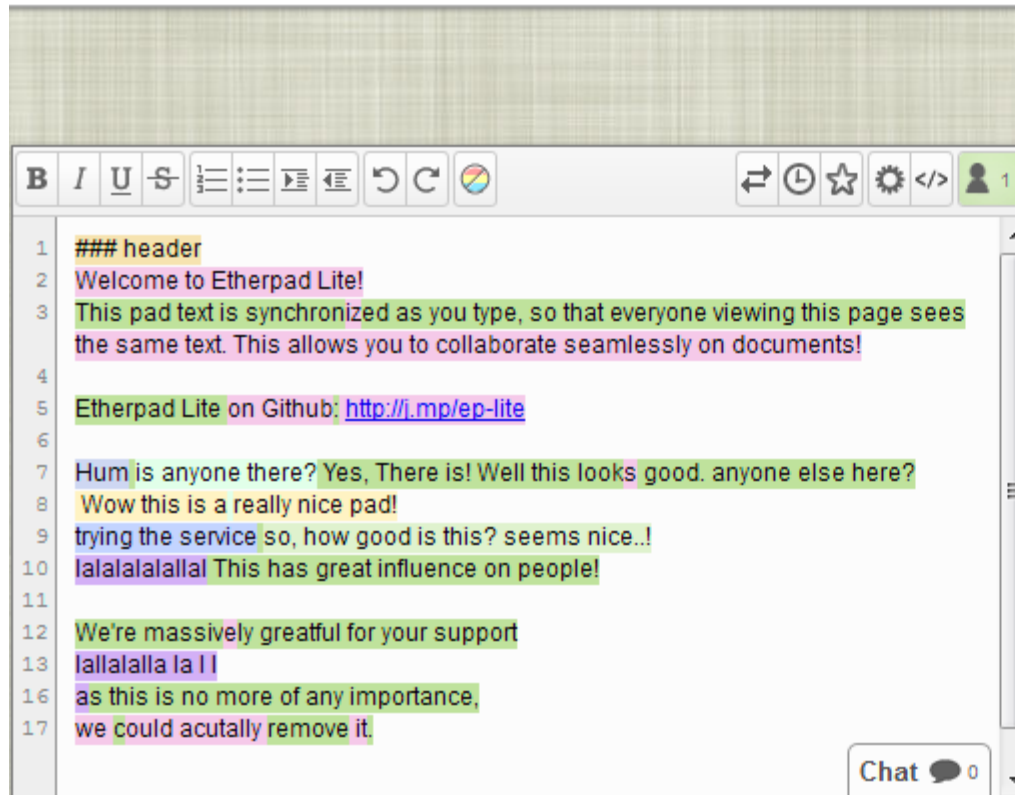


Beispiel gemeinsames Schreiben

- DARIAH verweist auf etherpad – ein einfaches Werkzeug zum gemeinsamen Schreiben
- <http://etherpad.org/>



About Download Contribute Links



Etherpad is a highly customizable **Open Source** online editor providing collaborative editing in real **real-time.**

DOWNLOAD

Version 1.6.0



Eine neue Rolle für die Bibliothek

- Unterstützung der Forschenden bei der Suche nach Tools
- Zugang zu Schulungen zu der Software
- Bereitstellen von einzelnen, generischen Tools für die Forschung

Dies kann eine neue Rolle für die Bibliothek im Forschungsdatenmanagement bringen

- Sie ist nahe an der Produktion der Forschungsdaten
- So kann die Verwaltung der Forschungsdaten vereinfacht werden
- Wer, wenn nicht die Bibliothek könnte diese Aufgaben übernehmen?

Kennen Sie Ihre Forschenden?

- Wenn die Bibliothek eine aktive Rolle im Forschungsdatenmanagement übernehmen will, muss sie die Beziehung zu den Forschenden intensivieren.

Diskutieren Sie:

- Kennen Sie die Forschenden?
- Kennen Sie den Bedarf der Forschenden an Tools?
- Soll / kann die Bibliothek für einzelne Institute / Disziplinen Dienste anbieten?
- Wie könnten diese aussehen?

Ablauf

Einige Erfahrungen

Mögliche Dienste der Bibliothek

Infrastruktur für Forschungsdokumentation

Infrastruktur für Datenpublikation





Forschungsdatenmanagement bedeutet Dokumentation der Forschungstätigkeit

- Forschungsdatenmanagement dient auch der Qualitätssicherung
 - Dazu müssen nicht nur die Daten selbst kuratiert werden, sondern auch die Notizen und Dokumente zum Entstehungsprozess der Daten
 - Dazu dient traditionell das Laborjournal
 - In der digitalen Forschungsumgebung braucht es ein digitales ‚Laborjournal‘ (ELN)
 - In den Geisteswissenschaften ist das noch kaum etabliert – trotzdem aber notwendig
 - Open Source ELN sind meist wenig intuitiv, alternativ:
-
- Nutzung von Dokumentenmanagementsystemen als Laborjournal, z.B. wenn vorhanden:
 - SharePoint
 - Confluence



Einführung von Confluence oder SharePoint ist eine Hochschulaufgabe

- Einführung von SharePoint oder Confluence ist jedoch aufwändig
- Nur gut implementiert, bringen diese Tools auch Arbeitserleichterung
- Andernfalls weichen die Forschenden auf ihre eigenen Tools aus

Niederschwellige Möglichkeiten

- Viele Forschende nutzen weit verbreitete Notiz-Lösungen. Z.B.:
 - Evernote
 - OneNote
 - etc.
- Ermöglicht: Integration von Dateien, Kalenderfunktion, Notizfunktion, Arbeit auf vielen Geräten, Teilen von Dateien in Gruppen, Export einer Seite für das Archiv
- Allerdings: Cloudlösungen – sollen die Daten z.B. bei Microsoft liegen?

Seminar 2016 - OneNote

DATEI START EINFÜGEN ZEICHNEN VERLAUF ÜBERPRÜFEN ANSICHT

Notizbücher digital Neuer Abschnitt 1

Eigene Notizbücher

- Notizbuch
- Schnelle Notizen
- Neuer Abschnitt 1
- Neuer Abschnitt 2

dasErste

- Lesenotizen
- Familie
- lod_led
- Neuer Abschnitt 1

Seminar 2016

Dienstag, 2. August 2016 14:02

Erfahrung: Bisher kritische Auseinandersetzung mit Daten

Besser: Wenn kleinere Einheiten - Auseinandersetzung mit kleinen 'Häppchen'

Durchsuchen (Strg+E)

Seite hinzufügen

- Die digitale seit
- Train2dacar 25
- Harvey dilza-Le
- Seminar 2016

Niederschwellige Möglichkeiten

- Projektmanagement-Tool als elektronisches Laborjournal nutzen?
- Boards lassen sich als CSV, JSON oder PDF exportieren
- Z.B. Trello

The screenshot displays a Trello board interface with the following components:

- Header:** 'Boards' tab, search bar, Trello logo, and user profile 'Niklaus Stettler'.
- Board Title:** '2016_Projekt_BAR-Dateiformate' (Private).
- Lists:**
 - Korrespondenz:** A list of email correspondence items, including 'Mail vom 15.9. von Niklaus', 'Mail von Niklaus an BAR vom 14. September 2016', 'Mail von Niklaus vom 9. August 2016', 'Mail von Krystyna vom 28. Juli 2016', 'Mail von Niklaus vom 21. Juni 2016', 'Mail vom 16. Juni 2016 von Krystyna', 'Mail vom 6. Juni 2016 von Krystyna Ohnesorge', and 'Mail vom 3. Mai 2016 von Krystyna'.
 - Aspekte:** Cards discussing 'Container und Codecs', 'Warum Audio und Video als Inhalte von archivtauglichen Dateien ausschliessen?', 'JPEG2000: OCR Software?', 'OOXML: wie steht es mit den Open Document Standards?', and 'JPEG2000: welches Dateiformat? Begründung?'.
 - Quellen:** Cards providing links to 'JPE2000 and Digitisation: Expert round table', 'JPEG 2000 Part 1 (Core) jp2 File Format', 'Using Lossy JPEG 2000 Compression For Archival Master Files', and 'International Image Interoperability Framework'.
 - Arbeitspakete:** Cards detailing tasks like 'Zuteilung der Arbeitspakete', 'JPEG2000 bzw. JPEG', 'OOXML bzw. Open Document', 'XML / XSD', 'Videoformate (MPEG-2, MPEG-4, andere)', 'PDF/A-2', and 'alle Überlegungen in Bezug zu Webseiten bringen'.
- Right Panel (Menü):** Options for 'Hintergrund ändern', 'Karten filtern', 'Extras', 'Sticker', 'Mehr', and 'Aktivität'.

Ablauf

Einige Erfahrungen

Mögliche Dienste der Bibliothek

Infrastruktur für Forschungsdokumentation

Infrastruktur für Datenpublikation



Wo die Daten ablegen?

- Untersuchung zum Forschungsdatenmanagement an baden-württembergischen Universitäten hat ergeben, dass Geistes- und Sozialwissenschaftler fachspezifische Repositorien klar bevorzugen
- Also: Wo die Daten ablegen?

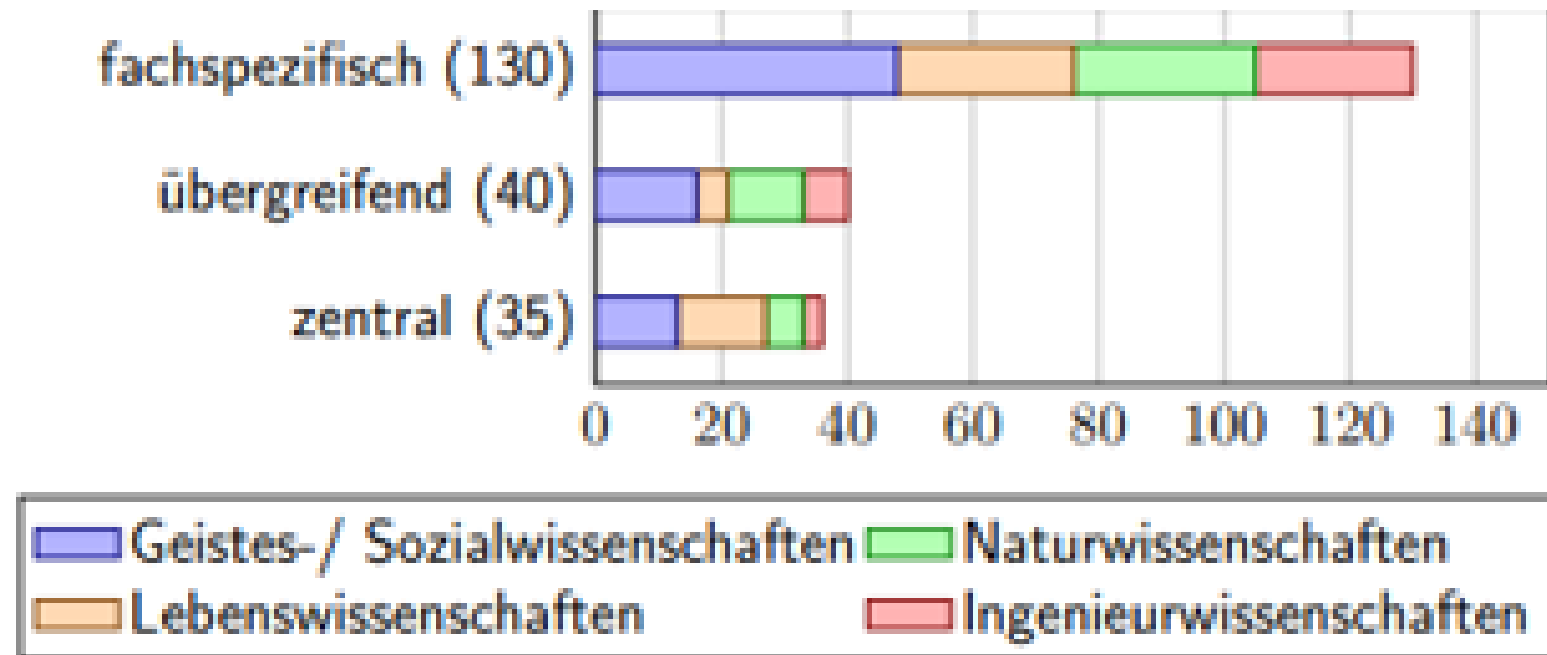


Abb. 34: Disjunkte Aufteilung nach dem Betriebskonzept des gewünschten Repositoriums in den Wissenschaftsbereichen.



Präferenz für fachspezifische Repositorien

- Um die Sichtbarkeit zu erhöhen
- Um Nachnutzung zu erhöhen

In der Schweiz sind fachspezifische Repositorien für die Digital Humanities erst im Aufbau.

Im Rahmen der Projekts ‚wissenschaftliche Information‘ von Swissuniversity könnten solche entstehen. Noch ist aber der Zeithorizont relativ offen.

Eine niederschwellige Alternative bis der nationale Dienst steht?

- Mendeley Data – verbindet Literaturverwaltung und Datenverwaltung
- Publikation und Daten können verbunden werden
- Gruppenbildung möglich
- Daten werden langzeitarchiviert
- Suchmaschinen finden die Daten

For datasets big and small
Store your research data online

Quickly and easily upload files of any type and we will host your research data for you. Your experimental research data will have a permanent home on the web that you can refer to.

Make your research data citable
DOIs and versioning following FORCE11 guidelines

Your published research data will include a FORCE11 compliant citation so that other researchers can effortlessly cite your research. We will also provide a unique DOI for each version of your dataset, so that your dataset's citation will always be valid.

Share your progress
with your colleagues and funding bodies

Privately share your unpublished data with the right people to move your research forward. You have full control over who can see and download your research data.

View and create
your dataset on any device

Mendeley Data has been created to work on a wide range of devices - from your lab workstation to your mobile phone. Browse datasets on the go, or take a picture in the lab and upload it instantly to your research data.

Long term preservation of your research data
Dark archive storage with DANS

We have teamed up with DANS (Data archiving and Networking Service) to ensure that your research data will

Eine niederschwellige Alternative?

Results: More than bin packing: Dynamic resource allocation strategies in cloud data centers

Published: 19 Sep 2015 | **Version 3** | DOI: 10.17632/xrpf54r78h.3

Viewed 524 Downloaded 175

















Contributor(s): [Andreas Wolke](#)

Description of this data

Data produced by simulations and experiments that were used in our paper "More than bin packing: Dynamic resource allocation strategies in cloud data centers" (Information Systems)

Experiment data files

[Download all files \(8\)](#)

 table_kmt_factors.csv	0.24 KB	
 sonar_cpu_mem_utilization.csv	0.15 KB	
 table_schedule_factors.csv	0.11 KB	
 table_schedule_factors_simulation.csv	0.07 KB	
 cbc_all.csv	108 KB	
 data---experiments.csv	26 KB	
 data---simulations-8.csv	108 KB	
 schedule_configurations4.csv	0.71 KB	

This data is associated with the following peer reviewed publication:

[More than bin packing: Dynamic resource allocation strategies in cloud data centers](#)

[Cite this article](#)



Published in:
Information Systems

Latest version

Version 3 2015-09-19

Published: 2015-09-19

DOI: 10.17632/xrpf54r78h.3

Vielen Dank