

Übungen zum Crashkurs Forschungsdatenmanagement

AG Forschungsdatenmanagement Onlineversion, April 2020

1. Mit welchen Forschungsdaten arbeiten Sie?

Notieren Sie einige Beispiele für Ihre Forschung.

Übungen zum Crashkurs Forschungsdatenmanagement

2. Skizzieren Sie den Forschungsdatenlebenszyklus.

Gehen Sie exemplarisch für eines Ihrer Forschungsprojekte den Forschungsdatenlebenszyklus durch.

Übungen zum Crashkurs Forschungsdatenmanagement

TIPPS

- Welche Daten erheben Sie?
- Wie verarbeiten und analysieren Sie anschließend die Daten?
- Welche Daten werden archiviert und was machen Sie, um die Daten zur Archivierung aufzubereiten?
- Sollen und wenn ja, wo sollen Ihre Daten veröffentlicht werden?
- Gibt es rechtliche Einschränkungen?
- Was benötigen andere zur Nachnutzung? Oder welche Daten von anderen würden Sie selber nachnutzen?



3. Welchen Nutzen bringt gutes Forschungsdatenmanagement (FDM) aus Ihrer Sicht?

Notieren Sie möglichst viele Gründe und ordnen Sie Ihre Antworten nach Ihren Prioritäten von der wichtigsten abwärts.

Übungen zum Crashkurs Forschungsdatenmanagement

MÖGLICHE ANTWORTEN

- Reproduzierbarkeit von Ergebnissen
- Auffindbarkeit der Daten
- Nachvollziehbarkeit der Forschung
- Transparenz der Forschung
- Qualitätssicherung und Qualitätsmerkmal

- Effizienz im Forschungsprozess
- Nachhaltigkeit
- Wissensgenerierung
- Keine Redundanz
- Strukturierung, Organisation
- Kollaborationen ermöglichen

MÖGLICHE ANTWORTEN

- Öffentliche Geldgeber, Forschungsdaten als öffentliches Gut
- Anforderungen der Fördermittelgeber
- Anforderungen von Verlagen, Repositorien
- Reputationsgewinn
- Publikationsoutput steigern durch Datenpublikationen
- Zitierbarkeit von Forschungsdaten

- Datenschutz und Datensicherheit
- Rechtssicherheit
- Datensicherung und Speicherung
- Datenzugang und Nachnutzung für Andere
- Neue Forschung ermöglichen
- Kostenersparnis

4. Nennen Sie die FAIR-Prinzipien und beschreiben Sie diese kurz.

• Findable:

Auffinden von Forschungsdaten und dazugehörigen Metadaten, menschen- und maschinenlesbar

Accessible:

Forschungsdaten und Metadaten mittels Standard-Kommunikationsprotokollen herunterladbar, abrufbar, nutzbar (lokal) machen

• Interoperable:

(teil)automatisierter Datenaustausch, Interpretation und (Neu)Kombination der Daten ermöglichen

• Reusable:

Nachnutzung von Forschungsdaten und verständlichen Bedingungen (Lizenzen, Zitierbarkeit)

5. Beurteilen Sie das FDM für Ihr eigenes Projekt anhand der FAIR-Prinzipien.

Handeln Sie nach den FAIR-Prinzipien? Welche der FAIR-Prinzipien halten Sie bereits ein? Und an welcher Stelle besteht Nachbesserungsbedarf?

5. Welche der Beispiele für Dateinamen auf der nächsten Folie folgen einer guten Benennungskonvention und warum?

BEISPIELE

- Olga_170413_probe17k
- Naturepaper karl britta james fertig!
- Vm4520132Schmidt.pdf
- 647749157.pdf
- 170413_probe17k_olga
- Naturepaper+karl+britta+james &nal
- Olga170413probe17k

- Krst_765_spkt_1203
- Naturepaper+karl+britta+james fertig! überarbeitet
- Kristall_765_spektr_20161203
- Nature_karlbrittajames_endendversion
- 28q8QGIHKwrRw.pdf
- Tagung_Digitale_Wissenschaft.pdf

- Kristall_765_spektr_20161203
- Begründung:
 - Keine Sonderzeichen
 - Sortierung nach: Bezeichnung, ID-Nummer, Datum
 - Spektr ist vermutlich das Experiment
 - Datum ist eindeutig zuzuordnen

7. Bewerten Sie die typische Benennungskonvention Ihrer eigenen Dateien oder denen Ihrer Arbeitsgruppe.

Entspricht dies dem Gelernten über gute Dateinamen?

Ja, weil

Nein, weil

Verbesserungsvorschlag:

8. Welche Standards für Dokumentationen, Metadaten oder Vokabulare kennen Sie aus Ihrem Fach?

ANTWORT: EINIGE BEISPIELE

- DIN Norm
- CAS Nummern (Chemikalien)
- ICD 10
- SI Einheiten
- SOPs (standard operating procedures)
- ISO
- Anamnesebögen
- Dublin Core

9. Nennen Sie die Vor- und Nachteile folgender Speichermedien:

eigener PC, USB-Stick, Speicherwolke der Universität Leipzig, Dropbox

ANTWORT – EIGENER PC

Vorteile

Nachteile

ANTWORT – EIGENER PC (LOKALER SPEICHERORT)

Vorteile	Nachteile
 selbst verantwortlich für Sicherheit und Backup 	 was mit dem PC geschieht, geschieht mit dem Backup
maximale Kontrolle	 eventuell fehlende Ressourcen und Know-How
	Einzellösungen aufwendig

ANTWORT – USB-STICK

Vorteile

Nachteile

ANTWORT – USB-STICK (MOBILE SPEICHERMEDIEN)

Vorteile	Nachteile
 einfach zu transportieren kann im verschließbaren Schrank oder Safe aufbewahrt werden 	 Verlust, Diebstahl, Schaden → besonders unsicher bei Verlust: Inhalte ungeschützt, falls nicht verschlüsselt externe Festplatte: stoß- und verschleißanfällig

ANTWORT – SPEICHERWOLKE DER UNIVERSITÄT LEIPZIG

<u>Vorteile</u> •

- •
- •

ANTWORT – SPEICHERWOLKE DER UNIVERSITÄT LEIPZIG (INSTITUTIONELLER SPEICHER)

Vorteile	Nachteile
 Backup der Daten ist sicherge-	 Geschwindigkeit eventuell vom
stellt	Netzwerk abhängig
 professionelle Durchführung un	 Zugriff auf Backups eventuell
Wartung	verzögert durch Dienstweg
Speicherung gemäß Datenschurichtlinien der Institution	• eventuell unklar, welche Sicherheitskriterien und - strategien eingesetzt werden

ANTWORT – DROPBOX

Vorteile

Nachteile

ANTWORT – DROPBOX (EXTERNE SPEICHERORTE)

Vorteile	Nachteile
 einfach zu nutzen und zu verwalten 	 je nach Anbieter kann die Verbindung auch unsicher sein
Backup der Daten ist sichergestellt	 abhängig von Internetzugang (Up- & Download eventuell langsam)
 für mobiles Arbeiten nutzbar professionelle Durchführung und Wartung 	Zugriff auf Backups eventuell verzögertDatenschutz?

9. Erläutern Sie den Unterschied zwischen Backup und Archivierung!

ANTWORT

• Backup:

automatische Sicherung aller Daten, um Datenverlust (technische Störung, menschliches Versagen, Unfall etc.) vorzubeugen, alle Versionen bleiben erhalten

Archivierung:

Sicherung ausgewählter Daten, um diese langfristig aufzubewahren, nur endgültige Versionen bleiben erhalten, dient der Integritätssicherung, Langzeitspeicherung, Durchsuchbarkeit von Daten

10. Worauf legen Sie Wert bei der Auswahl eines Langzeitarchivs für Ihre Daten?

Nennen Sie einige Beispiele.

ANTWORT

UNIVERSITÄT

LEIPZIG

MÖGLICHE ANTWORT

- Gesicherte Langzeitverfügbarkeit
- Zertifizierung, Siegel
- "Sicherer" Anbieter, Reputation des Anbieters, Langlebigkeit des Anbieters
- Technische Anforderungen
- Strategie zur Migration und Datenkonvertierung, regelmäßige Prüfung der Lesbarkeit
- Kosten
- Zugänglichkeit der Daten, zum Beispiel: Gibt es auch eine Repositoriumsfunktion?

11. Recherchieren Sie auf http://re3data.org/ nach einem Repositorium, das Sie für Ihre Forschungsergebnisse nutzen können!

Haben Sie ein passendes gefunden?

ANTWORT

Ja, und zwar

Nein, weil

12. Welche Eigenschaften eines Repositoriums sind für Sie wichtig?

Begründen Sie Ihre Entscheidung.

ANTWORT

- •

MÖGLICHE ANTWORT

- Open-Access-Veröffentlichung
- Nutzungsbedingungen
- Policy des Repositoriums
- Möglichkeit, Lizenzen zu vergeben
- Vergabe eines Persistent Identifiers
- Zertifizierung

13. Welche Vorteile sehen Sie im Publizieren von Forschungsdaten?

Nennen Sie einige Thesen.

ANTWORT

MÖGLICHE ANTWORT

- Nur durch öffentliches Teilen von Forschungsdaten können globale Ereignisse, wie Pandemien, erforscht werden.
- Ich werde meine Daten publizieren, damit mein Artikel häufiger zitiert wird.
- Forschung wird zu großen Teilen öffentlich finanziert, daher sind die dabei entstandenen Daten auch ein öffentliches Gut und sollten daher frei verfügbar sein.
- Durch die Nachnutzung meiner Daten k\u00f6nnen sich spannende neue Kollaborationen ergeben.
- Forschungsdaten sind ein Gut, dessen langfristiger Erhalt und Sicherung für die Zukunft einen Wert darstellt.
- Durch das Teilen meiner Forschungsdaten wird meine Reputation und Glaubwürdigkeit in der Community gesteigert.

14. Was spricht thesenhaft gegen die Publikation von FD?

Notieren Sie mögliche Gegenargumente für das Veröffentlichen von Forschungsdaten.

ANTWORT

MÖGLICHE ANTWORT

- Meine Forschungsdaten gehören mir.
- Die Publikation von Forschungsdaten trägt nicht zum Reputationsgewinn bei.
- Wenn ich meine Forschungsdaten publiziere, könnte mir Jemand zuvorkommen und Erkenntnisse veröffentlichen, die auf meinen Daten basieren.
- Natürlich werde ich immer eigene Daten erheben, statt bereits vorhandene zu nutzen: Ich werde meine Fragestellungen nicht an vorhandene Daten anpassen.

MÖGLICHE ANTWORT

- Wenn ich meine Daten publiziere, wird meine Forschung völlig transparent und selbst kleinste Fehler werden offenbar.
- Veröffentlichte Daten bringen keinen weiteren Nutzen.
- Management und Publikation von Forschungsdaten verursachen Kosten, die ich nicht tragen kann/möchte und in keinem Verhältnis zum Aufwand stehen.
- Das Anlegen von Persistenten Identifiern, wie ORCID, kostet Zeit und bringt keinen gesteigerten Nutzen.

15. Wie könnten Sie Gegenargumente kontern und Zweifel beseitigen?

Finden Sie im Folgenden Gegenargumente für jede These.

ANTWORT "Meine Forschungsdaten gehören mir."

- •
- •
- •
- •

MÖGLICHE ANTWORT "Meine Forschungsdaten gehören mir."

- Daten sollten als Allgemeingut angesehen werden, wenn diese aus öffentlich finanzierten und geförderten Projekten stammen.
- Ein alleiniger Besitzanspruch auf Daten lässt sich außerdem nur schwer erheben, da hierbei auch das jeweilige Anstellungsverhältnis, vertragliche Vereinbarungen mit Projektpartner_innen und/oder Verlagen berücksichtigt werden müssen.

ANTWORT "Die Publikation von Forschungsdaten trägt nicht zum Reputationsgewinn bei."

- •
- •
- •

MÖGLICHE ANTWORT "Die Publikation von Forschungsdaten trägt nicht zum Reputationsgewinn bei."

- Datenpublikationen z\u00e4hlen als Forschungsoutput, k\u00f6nnen zitiert und somit als eigenst\u00e4ndige Publikation angesehen werden.
- Außerdem steigen Glaubwürdigkeit, Transparenz und Ansehen der eigenen Forschung in der Fachcommunity, wenn qualitativ hochwertige Forschungsdaten veröffentlicht werden.

ANTWORT "Wenn ich meine Forschungsdaten publiziere, könnte mir Jemand zuvorkommen und Erkenntnisse veröffentlichen, die auf meinen Daten basieren."

- •

MÖGLICHE ANTWORT "Wenn ich meine Forschungsdaten publiziere, könnte mir Jemand zuvorkommen und Erkenntnisse veröffentlichen, die auf meinen Daten basieren."

- Über den Zeitpunkt der Veröffentlichung von Forschungsdaten können Sie i. d. R. allein bestimmen.
- Bei vielen Repositorien gibt es zudem eine sogenannte Embargo-Funktion, mit deren Hilfe Forschungsdaten erst nach einem bestimmten Zeitraum veröffentlicht werden können.

ANTWORT "Natürlich werde ich immer eigene Daten erheben, statt bereits vorhandene zu nutzen: Ich werde meine Fragestellungen nicht an vorhandene Daten anpassen."

- •
- •

MÖGLICHE ANTWORT "Natürlich werde ich immer eigene Daten erheben, statt bereits vorhandene zu nutzen: Ich werde meine Fragestellungen nicht an vorhandene Daten anpassen."

 Nachnutzen von Daten bedeutet natürlich nicht, keine eigenen Daten und Fragestellungen zu entwickeln. Nachgenutzte Daten können die eigenen Daten ergänzen, zur Überprüfung dienen oder auch zu neuen Hypothesen führen.

ANTWORT "Wenn ich meine Daten publiziere, wird meine Forschung völlig transparent und selbst kleinste Fehler werden offenbar."

- •
- •
- •
- •

MÖGLICHE ANTWORT "Wenn ich meine Daten publiziere, wird meine Forschung völlig transparent und selbst kleinste Fehler werden offenbar."

 Transparenz und Reproduzierbarkeit sind wichtige Bestandteile der guten wissenschaftlichen Praxis. Die Datenqualität und Überprüfbarkeit der Ergebnisse sollten demnach immer gewährleistet werden.

ANTWORT "Veröffentlichte Daten bringen keinen weiteren Nutzen."

- •
- •
- •
- •

MÖGLICHE ANTWORT "Veröffentlichte Daten bringen keinen weiteren Nutzen."

- Veröffentlichte Daten bringen sehr großen Nutzen. Nur so kann z. Bsp. die globale, interdisziplinäre Forschung über Pandemien ermöglicht und daraus neue Erkenntnisse gewonnen werden.
- Fragestellungen können über die Jahre hinweg neu entwickelt und die Daten für Forschungsfragen genutzt werden, die bei der ursprünglichen Erhebung noch gar nicht angedacht waren.

ANTWORT "Management und Publikation von Forschungsdaten verursachen Kosten, die ich nicht tragen kann/möchte und in keinem Verhältnis zum Aufwand stehen."

- •
- •
- •

MÖGLICHE ANTWORT "Management und Publikation von Forschungsdaten verursachen Kosten, die ich nicht tragen kann/möchte und in keinem Verhältnis zum Aufwand stehen."

- Ja, Kosten entstehen. Aber bei fast allen Fördermittelgebern (DFG, EU, BMBF etc.) können Mittel für das Datenmanagement beantragt werden.
- Außerdem werden viele Services wie Speicher, Backup, Tools, Schulungen und Beratung durch zentrale Services und Infrastrukturen der Universität zur Verfügung gestellt.
- FDM und Datennachnutzung sparen außerdem Kosten und Ressourcen, da viele Daten gar nicht noch einmal erhoben werden müssen und können (Stichwort Beobachtungsdaten).
- Die zukünftige Bereitstellung von Forschungsdaten wird erleichtert.

ANTWORT "Das Anlegen von Persistenten Identifiern, wie ORCID, kostet Zeit und bringt keinen gesteigerten Nutzen."

- •
- •
- •
- •

MÖGLICHE ANTWORT "Das Anlegen von Persistenten Identifier, wie ORCID, kostet Zeit und bringt keinen gesteigerten Nutzen."

 Im Vergleich zum Nutzen ist der Aufwand gering. Mit ORCID ist man als Forschender eindeutig identifizierbar. ORCID ist bspw. für Veröffentlichungen nutzbar. 16. Markieren Sie im folgenden Text jedes Wort, jede Phrase oder jeden Textbestandteil, die eine Anonymisierung notwendig machen.

TEXT: Fallbeschreibung und Transkript des Interviews

Herr Tom Jeavons, 63 Jahre alt, litt an metastasierendem Krebs, der sich aus der Primärstelle in der Harnblase entwickelt hat. Er hatte monatelang mit starken Schmerzen, Angstzuständen und anderen Symptomen zu kämpfen. In dieser Zeit kümmerte sich hauptsächlich seine Ehefrau Sue (58) um ihren Mann. Schließlich erhielt sie Hilfe vom "Hospiz zu Hause"-Team vom nahe gelegenen Hospiz St. Barbara. 11 Tage vor seinem Tod wurde er stationär aufgenommen, er verstarb dort. Die Ärzte stuften den Fall als äußerst kritisch ein, insbesondere unter Berücksichtigung der palliativen Sedierung und der bisherigen Erfahrung des Personals, diese komplexen Symptome unter Kontrolle zu bringen. Aus dem Hospiz befanden sich unter den Interviewteilnehmern die Fachärztin Dr. Jane O'Connor sowie drei Krankenpfleger/innen: Elaine McDonald, Claire Smith und Mark Ferguson. Des Weiteren wurde Dr. Paul Hyde, der Hausarzt von Frau und Herrn Jeavons befragt, was der Untersuchung noch eine andere medizinische Perspektive hinzufügte, die die Ungewöhnlichkeit des Falles stärker verdeutlicht.

Die zentralen Themen in allen Interviews bildeten die hartnäckigen und gualvollen Symptome und die wiederholten Anfragen von Herrn Jeavons nach Sterbehilfe. Frau Jeavons erwähnte frühere Diskussionen mit ihrem Mann über die Möglichkeit, in eine Dignitas-Klinik zu gehen. Zu dieser Zeit war er aber bereits zu krank, um zu reisen. Sie machte in der Befragung ebenfalls deutlich, wie besorgt sie darüber war, was ihre erwachsenen Kinder vielleicht beim Sterben ihres Mannes im Hospiz mit ansehen müssen.

ANTWORT: ANONYMISIERUNG

Herr Tom Jeavons, 63 Jahre alt, litt an metastasierendem Krebs, der sich aus der Primärstelle in der Harnblase entwickelt hat. Er hatte monatelang mit starken Schmerzen, Angstzuständen und anderen Symptomen zu kämpfen. In dieser Zeit kümmerte sich hauptsächlich seine Ehefrau Sue (58) um ihren Mann. Schließlich erhielt sie Hilfe vom "Hospiz zu Hause"-Team vom nahe gelegenen Hospiz St. Barbara. 11 Tage vor seinem Tod wurde er stationär aufgenommen, er verstarb dort. Die Ärzte stuften den Fall als äußerst kritisch ein, insbesondere unter Berücksichtigung der palliativen Sedierung und der bisherigen Erfahrung des Personals, diese komplexen Symptome unter Kontrolle zu bringen. Aus dem Hospiz befanden sich unter den Interviewteilnehmern die Fachärztin Dr. Jane O'Connor sowie drei Krankenpfleger/innen: Elaine McDonald, Claire Smith und Mark Ferguson. Des Weiteren wurde Dr. Paul Hyde, der Hausarzt von Frau und Herrn Jeavons befragt, was der Untersuchung noch eine andere medizinische Perspektive hinzufügte, die die Ungewöhnlichkeit des Falles stärker verdeutlicht.

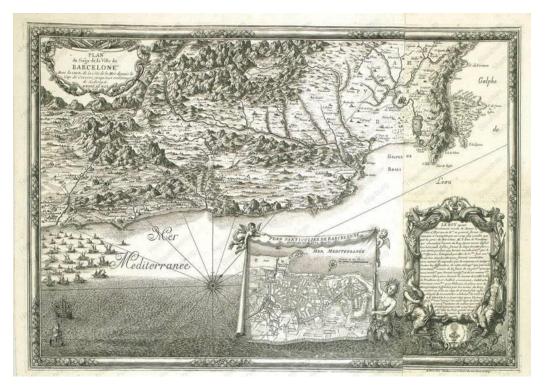
Die zentralen Themen in allen Interviews bildeten die hartnäckigen und qualvollen Symptome und die wiederholten Anfragen von Herrn Jeavons nach Sterbehilfe. Frau Jeavons erwähnte frühere Diskussionen mit ihrem Mann über die Möglichkeit, in eine Dignitas-Klinik zu gehen. Zu dieser Zeit war er aber bereits zu krank, um zu reisen. Sie machte in der Befragung ebenfalls deutlich, wie besorgt sie darüber war, was ihre erwachsenen Kinder vielleicht beim Sterben ihres Mannes im Hospiz mit ansehen müssen.

17. Wenn Sie das Bild auf der folgenden Folie nachnutzen möchten, welche Angaben zur Lizenz müssen Sie machen?

BILD ZU AUFGABE 17

Nutzen Sie folgenden Link zur Information:

https://bit.ly/3cpjB6h



- •
- •
- •
- •

- Da die Karte 1698 erstellt wurde, ist sie gemeinfrei und kann ohne Lizenzangabe genutzt werden. Das Urheberrecht ist abgelaufen.
- Man kann annehmen, dass die Karte gescannt wurde, um sie zu digitalisieren. Beim Scannen entsteht jedoch in der Regel kein neuer Urheberrechtsschutz.
- Demnach ist die auf der Webseite angegebene Lizenz (CC-BY 3.0) weder bindend noch verbindlich. Bei der Nachnutzung ist keine Lizenzangabe notwendig.
- Auf Grund der guten wissenschaftlichen Praxis empfiehlt es sich jedoch immer die Quelle anzugeben.

18. Zitieren Sie den Datensatz auf der nachfolgenden Folie.

DATENSATZ

- Identifier: https://doi.org/10.1594/PANGAEA.905293
- Creator: Simon van Bellen
- Title: Holocene temperature reconstructions for north-eastern North America and the north-western Atlantic
- Publisher: PANGAEA Data Archiving & Publication
- Publication year: 2019
- Resource type: Dataset

- •
- •
- •
- •

MÖGLICHE ANTWORT NACH da|ra

 Van Bellen, Simon (2019): Holocene temperature reconstructions for north-eastern North America and the north-western Atlantic. PANGAEA Data Archiving & Publication. Dataset. https://doi.org/10.1594/PANGAEA.905293

Weitere Informationen: <u>Zitieren von Daten – forschungsdaten.info</u>

19. Was sind die wesentlichen Bestandteile eines Datenmanagementplans?

- •
- •
- •
- •

- Autor, Projekt und Verantwortlichkeiten
- Datentypen, -formaten, -umfang
- Datenbenennung, Dokumentation, Metadaten und Standards
- Datenaustausch
- Zugriffsregelungen
- Ethische und rechtliche Aspekte
- Sicherung der Daten
- Archivierung
- Veröffentlichung und Nachnutzung
- Kosten



HABEN SIE FRAGEN? WIR BERATEN SIE GERN!

forschungsdaten@uni-leipzig.de

Diese Präsentation wurde erstellt unter Nutzung von:

Dolzycka, Dominika; Biernacka, Katarzyna; Helbig, Kerstin; Buchholz, Petra (2019): Train-the-Trainer Konzept zum Thema Forschungsdatenmanagement, Version 2.0, Berlin, https://doi.org/10.5281/zenodo.2581292. – CC BY 4.0 (https://creativecommons.org/licenses/by/4.0/). Es wurden Änderungen vorgenommen.



Der Text dieser Präsentation steht unter der Lizenz Creative Commons Attribution 4.0 International (CC BY 4.0).

KONTAKT UND ZUSTÄNDIGKEITEN

Zentraler Kontakt:

forschungsdaten@uni-leipzig.de

Zuständigkeiten:

- Dezernat 1: Forschung und Transfer:
 - Erstberatung, Fördermittel beantragen, DMP-Erstellung, Schulungen
- Universitätsbibliothek:
 - Veröffentlichung von Forschungsdaten, PIDs, Lizenzen, Nachnutzung
- Universitätsrechenzentrum:
 - Datensicherheit, Storage, Backup, Langzeitarchivierung, HPC, Big Data



Dezernat 1: Forschung und Transfer

- Pia Voigt (pia.voigt@zv.uni-leipzig.de, Telefon: 97-35075)
- Dr. Barbara Weiner (<u>barbara.weiner@zv.uni-leipzig.de</u>, Telefon: 97-35071)

Universitätsbibliothek

Dr. Stephan Wünsche (<u>wuensche@ub.uni-leipzig.de</u>, Telefon: 97-30564 oder 97-35074)

Universitätsrechenzentrum

- Sebastian Frericks (<u>sebastian.frericks@uni-leipzig.de</u>)
- Dr. Stefan Kühne (stefan.kuehne@uni-leipzig.de, Telefon: 97-33303)
- Dr. Romy Elze (<u>romy.elze@uni-leipzig.de</u>, Telefon: 97-32304)
- Clemens Hoffmann (hoffmann@informatik.uni-leipzig.de, Telefon: 97-32347)