

Forschungsdatenmanagement im GrK MuSAMA an der Universität Rostock

Sebastian Schick¹ | **Karsten Labahn**² | **Robert Stephan**²

¹GrK MuSAMA, Institut für Informatik, Universität Rostock

²Universitätsbibliothek Rostock



Agenda

Das Graduierten Kolleg MuSAMA

Datenmanagementplan in MuSAMA

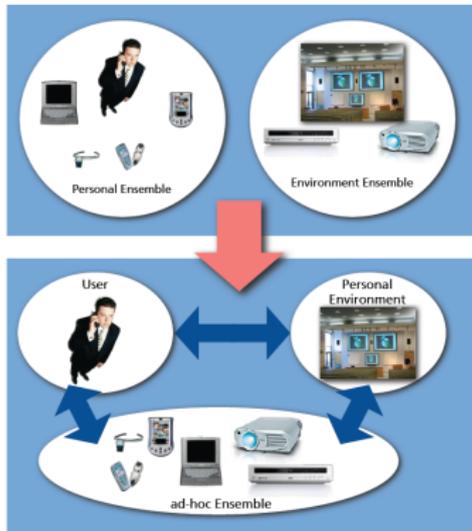
MuSAMA Forschungsprimärdaten im institutionellen Repository der Uni Rostock

Das Graduierten Kolleg MuSAMA

- MuSAMA: Multimodal Smart Appliance Ensembles for Mobile Applications
- MuSAMA untersucht Techniken, die eine spontane, kooperative Assistenz durch Geräte ermöglicht
- kooperative Assistenz von Geräte-Ensembles ohne zusätzlichen Programmier- und Entwicklungsaufwand



Herausforderung für ein solches Ensemble



- Wie kann das Ensemble herausfinden, was der Nutzer machen würde, wenn er wüsste was er machen kann?
- Und wie kann das Ensemble dann bestimmen, wie dieses Ziel erreicht wird?
- Ein Ziel von MuSAMA: wie kann das ohne Einsatz von zentralem Wissen über den Nutzer und das Ensemble gelöst werden?



Experimentalprogramm in MuSAMA

- „Ziel ist, durch einen systematischen, koordinierten Ansatz den organisatorischen und logistischen **Aufwand für experimentelle Untersuchungen** möglichst **zu minimieren**“
- „strukturiertes Experimentalprogramm [...], um die **gemeinsame Experimentalplanung** zu erleichtern“
- „ist ein weiteres Ziel [...] einen projektübergreifenden Experimentalcorpus zu erstellen, um dadurch die **wechselseitige Nutzung** von Experimentaldaten und die **Wiederholbarkeit** von Experimenten zu unterstützen“

Forschungsprimärdaten in MuSAMA

Daten:

- **Unveränderte** Forschungsprimärdaten (Messdaten, Fragebögen, etc.)
- (Daten die auf Konsistenz und Vollständigkeit überprüft wurden)
- (Zwischen) Ergebnisse aus der Datenanalyse
- Endgültige Analyseergebnisse

Software und Dokumentation:

- Laborbuchblätter
- Skripte und Programme zur Auswertung etc.
- Dokumentation von Experimenten und Analysen



Das Graduierten Kolleg MuSAMA

Datenmanagementplan in MuSAMA

MuSAMA Forschungsprimärdaten im institutionellen Repository der Uni Rostock

Datenmanagementplan in MuSAMA

- ein zweistufiges Verfahren
- **Intern:** Verwalten von Forschungsprimärdaten
 - Wiki-basierter Ansatz für die gemeinsame Experimentalplanung
 - Wiederholbarkeit von Experimenten mit Virtualisierungsplattform (Virtuelle Maschine, VMware Cluster, Storage Server)
- **Extern:** wechselseitige Nutzung von Forschungsprimärdaten
 - Repository um DFG Empfehlungen umzusetzen
 - kein geeignetes disziplinäres Forschungsdatenrepository
 - kein Repository für Forschungsprimärdaten an der Uni Rostock



Intern: Verwalten von Forschungsprimärdaten

- **Experimentalplanung:** Verwaltung von Experimenten
 - Liste von Experimenten
 - Organisation von Experimenten (Anmeldung, Planung, etc.)
- Reproduzierbarkeit gewährleisten
 - Mehrstufiges Verfahren mit Kontrollinstanz
 - Datensammlung:
 - Daten (anonymisiert/pseudonymisiert, verschlüsselt)
 - Einverständniserklärung
 - Dokumentation
 - Datenprüfung: Änderung der Zugriffsrechte auf "nur lesend"



Extern: wechselseitige Nutzung von Forschungsprimärdaten

- gesicherte Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten
- Beratung durch Universitätsbibliothek Rostock
- ein Anforderungskatalog für eine Lösung
 - Bereitstellung einer nachhaltigen und projektunabhängigen Infrastruktur
 - Zugänglichkeit im Internet (OpenAccess)
 - dauerhaften Zugriff über eine langzeitstabile und damit zitierbare URL
- derzeit diskutierte Lösung:
 - Kooperation mit Universitätsbibliothek
 - projektunabhängiges institutionelles Repository



Das Graduierten Kolleg MuSAMA

Datenmanagementplan in MuSAMA

MuSAMA Forschungsprimärdaten im institutionellen Repository der Uni Rostock



MuSAMA Forschungsprimärdaten im institutionellen Repository der Uni Rostock

- Pilotprojekt zum Thema Forschungsprimärdaten an der UB Rostock
- technische Umsetzung:
 - Publikation im Repository für elektronische Publikationen der UB Rostock
 - neuer Dokumenttyp "DataPublication"
 - ausreichende Hardwareinfrastruktur
 - Speicher, Server und Anbindung an Backup-/Archivierungsinfrastruktur
 - Kooperation mit IT- und Medienzentrum der Universität Rostock
- inhaltliche Umsetzung
 - Beschreibung mit einem Mindestmaß an Metadaten zur eindeutigen Identifikation und Recherchierbarkeit
 - Verknüpfung mit anderen Publikationsformen



Ausprägung der Metadatenbeschreibung

- Ziel der Metadatenbeschreibung:
 - ausreichende Dokumentation der Daten und Software
 - Nachvollziehbarkeit und Replizierbarkeit gewährleisten
- 3 Ebenen der Beschreibung
 1. Metadatensatz mit Kernset an Informationen (strukturiert)
 - XML/MODS (inhaltlich DataCite/RADAR angelehnt)
 - ggf. mods:extension mit Forschungsdaten-spezifischen Erweiterungen
 2. Dokumentation: pdf-Dokument mit verbaler Beschreibung des gesamten Forschungsdatensatzes (verpflichtend)
 3. detaillierte Dokumentation zu einzelnen Elementen des Datensatzes innerhalb der Datencontainer

Metadatenbeschreibung des Datensatzes

- Dokumentation im Daten-Container
 - einer/mehrere Container (zip) für Daten/Software in beliebiger Ordnerstruktur
 - `documentation.pdf` (Beschreibung des Forschungsdatensatzes)
 - Pflichtfelder: 15 Felder (Autor, Typ, Stichwörter, Aufbau, Durchführung, etc.)
- Detaillierte Dokumentation der Forschungsprimärdaten
 - je Ordner Inhalt und Struktur einzelner Dateien beschreiben
 - Pflichtfelder: Art der Daten, Variablenname und -typ, Ausgabe von Skripten ect.
 - hier nur allgemeine Vorgaben, weil nicht vollständig beschreibbar



Offen Fragen

- inhaltliche Mindestanforderungen an Datenqualität
- formale Mindestanforderungen: Lizenz, Metadatenset, Datenformate
- Metadatenset:
 - wie generisch, wie umfangreich, wie spezifisch?
 - Normdaten, Vokabulare?
- Rechte:
 - muss eine freie Lizenz (CC, ODbL) verpflichtend gefordert werden?
 - Konflikte mit enthaltenen, anders lizenzierten Inhalten
 - personenbezogene Daten
- Langzeitarchivierung