

Forschungsdatenmanagement in der Klinischen Forschung

Dr. Harald Kusch & Karoline Buckow

Institut für Medizinische Informatik
UNIVERSITÄTSMEDIZIN GÖTTINGEN



RDA/DINI-Workshop – 28. Mai 2015



<http://orcid.org/0000-0002-9895-2469>

Institut für Medizinische Informatik
UNIVERSITÄTSMEDIZIN GÖTTINGEN



Forschungsdatenmanagement in der klinischen Forschung:





Kernfragen und Herausforderungen

1. Was sind spezifische Anforderungen an die Forschungsdaten-Infrastruktur in der medizinischen Forschung?
Sind medizinische Forschungsdaten (Grundlagen- und klinische Forschung) anders als Daten anderer Fachbereiche?
2. Welche Fachverbände der medizinischen Forschung sind im Bereich des Forschungsdaten-Managements aktiv (Auswahl)?
3. Welche Bearbeitungsschwerpunkte werden im Institut für Medizinische Informatik in Göttingen bearbeitet, um die Forschungsdaten-Infrastruktur in der medizinischen Forschung voranzubringen?
4. Wie gestaltet sich die konkrete Erarbeitung von Maßnahmen zur Optimierung der Forschungsdaten-Infrastruktur bei der Förderungsform "SFB-INF-Projekt"?
 - a) Was sind fachübergreifende Herausforderungen (aus Sicht aller INF-Projekte)?
 - b) Was sind Herausforderungen im medizinischen Kontext (am Beispiel des klinischen SFB 1002)?
5. Welche Bearbeitungsschwerpunkte der RDA-Initiativen sind für die medizinische Forschung wichtig?
Welche dieser Schwerpunkte werden bereits (in der Göttinger Medizininformatik) bearbeitet?



Forschungsdatenmanagement in der klinischen Forschung:

Spezifische Anforderungen in der medizinischen Forschung

- Daten- und Persönlichkeitsschutz (datenrechtlich sensitive / schicksalhafte Daten):
 -  Management von persönlichen (identifizierenden) Daten
 -  Management von sekundär identifizierenden Daten (z.B. Genomsequenz)
 -  Löschen von Daten bzw. dem persönlichen Bezug zu Daten muss möglich sein (unter bestimmten Umständen)
 -  Datenweiterverarbeitung/-weitergabe:
 - Berücksichtigung in Einwilligungserklärung (vor Studienbeginn)



Forschungsdatenmanagement in der klinischen Forschung :



Spezifische Anforderungen in der medizinischen Forschung

- Ungenügende Standardisierung von klinischen Forschungsdaten
- Kein zentrales (überregionales) Forschungsdaten-Repository verfügbar
- Mehrfache Nutzung von Daten nicht immer möglich (z.B. Biomaterial)
- Zeitressourcen: Vereinbarkeit von Versorgung und Forschung
 - Versorgungs-Infrastruktur hat Vorrang vor Forschungs-Infrastruktur
 - Arzt und Forscher in einer Person



Forschungsdatenmanagement in der klinischen Forschung :



Medizinische Fachverbände (Auswahl):



Technologie- und Methodenplattform für die vernetzte medizinische Forschung) e.V.



Deutsche Gesellschaft für Medizinische Informatik, Biometrie und Epidemiologie e.V.



Deutscher Verband für Dokumentation und Informationsmanagement in der Medizin



Forschungsdatenmanagement in der klinischen Forschung :



Beitrag der Medizinische Informatik Göttingen zur Standardisierung des Forschungsdaten-Managements:

- Konzeption von virtuellen Forschungsumgebungen
- Entwicklung von datenschutzkonformen Identitätsmanagement-Tools
- Erarbeitung von Qualitätsmanagementverfahren für die Forschungsdaten-Dokumentation und -Archivierung
- Zusammenführung und Bereitstellung von Daten (Metadaten-Mapping)
- Erstellung von Datenschutzkonzepten nach Modellvorgaben
- Konzept-Entwicklung zur Langzeitarchivierung von Daten in der medizinischen Forschung



SFB INF-Projekte



INF = Teilprojekt Informationsstruktur

Stand Januar 2013: 27 INF-Projekte in 232 SFBs

- dienen "vor allem dem systematischen und auf langfristige Nutzung angelegten Management der im Kontext des SFBs erhobenen relevanten Daten"
- z.T. genutzt zur Etablierung von standortweiten Forschungsdaten-Infrastrukturlösungen
- untereinander kaum vernetzt



SFB INF-Projekte



Gemeinsamer Workshop der SFB-Inf-Projekte 2013

(Engelhardt & Strathmann, http://dx.doi.org/10.2312/RADIESCHEN_010)

Fachübergreifende Workshop-Ergebnisse / Ziele der INF-Projekte:

- Erhöhung der Akzeptanz: Generierung von Mehrwert für die Forschenden (z.B. besser strukturierte Datenablage, bessere Zitierbarkeit durch PIDs)
- Abbau von Vorbehalten bezüglich Daten-Teilen durch abgestufte Daten-Policy
- Institutionalisierung des Forschungsdatenmanagements in einer standortweiten Forschungsdaten-Infrastruktur (Profitieren von Erfahrung)



SFB INF-Projekte

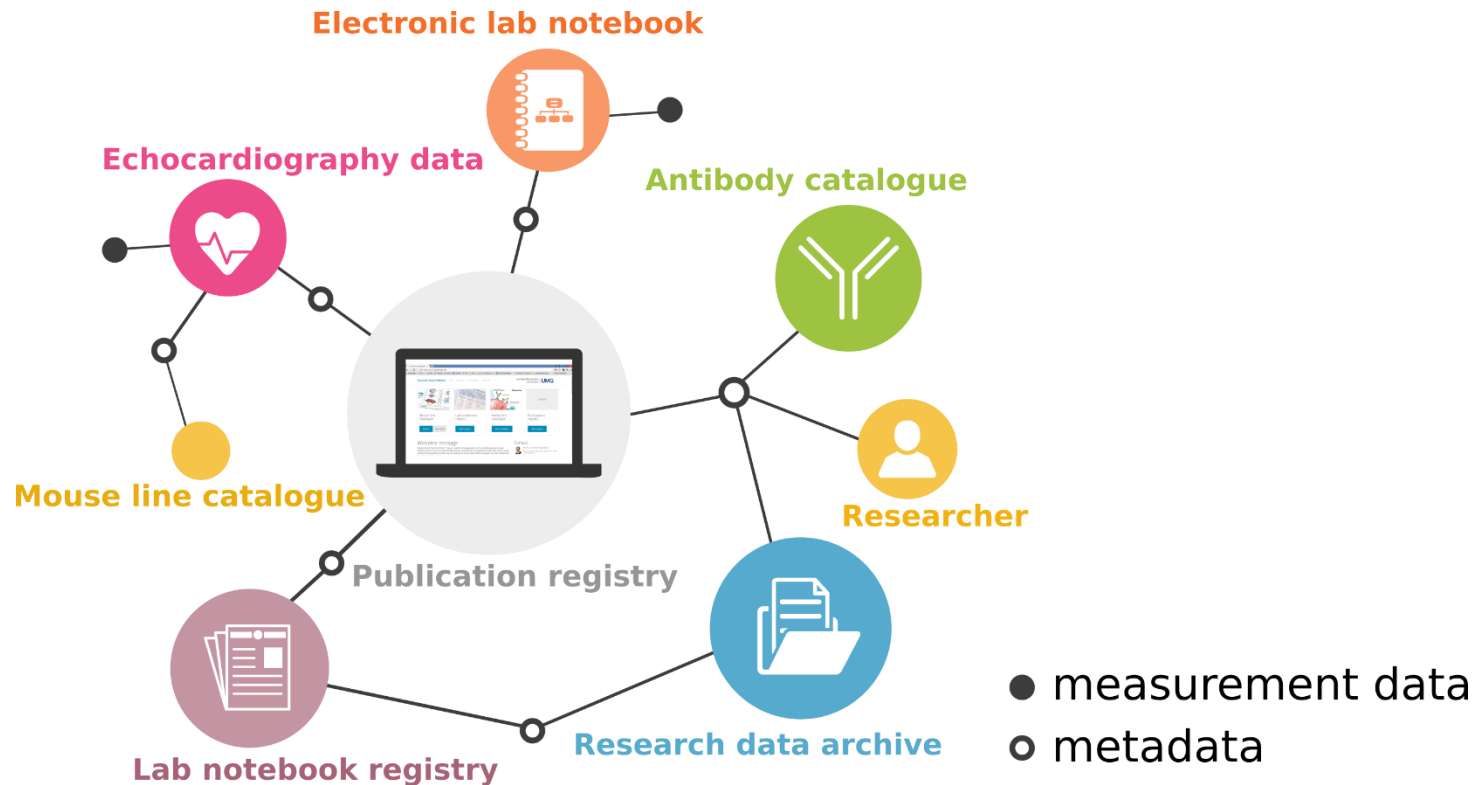
Aufgaben des INF-Projekts des SFB 1002:

- Ermittlung und Priorisierung von IT-Anforderungen der Forscher
 - Bestimmung von Kern-Komponenten der Forschungsdaten-Infrastruktur
 - Bestimmung der Komplexität der Anforderungen
- Evaluation bestehender Standards für die Verwendung im INF-Projekt
 - Standardisierung von erhobenen Daten im SFB 1002
 - Erarbeitung von Interaktionen mit bestehenden öffentlichen Datenbanken
- Evaluation bestehender Werkzeuge für die Verwendung im INF-Projekt
 - Adaptation bestehender Werkzeuge an SFB 1002 Anforderungen
 - bei Bedarf: Entwicklung von neuen Werkzeugen
- Beratung bei der:
 - Optimierung der IT-Grundversorgung
 - Automatisierung der Datenerfassung
- Eigene Forschung
 - Verbesserung der Akzeptanz von elektronischen Laborbuchsystemen



SFB 1002 Forschungsdatenplattform

Datenorganisation eines klinischen Sonderforschungsbereiches
in einer integrierten, langfristig verfügbaren Forschungsdatenplattform



Discussion

RDA Requests for comments - Technical components



RDA Components Data Management	Medizinische Informatik Göttingen
5.1 PID System	available
5.2 ID System for Actors	Useful
5.3 Registry System for Trusted Repositories	Useful
5.4 Metadata System	Under development
5.5 Schema Registry System	Under development
5.6 Registry System for Semantic Categories, Vocabularies, etc.	Under development
5.7 Registry System for Data Types	Useful
5.8 Registry System for Practical Policies	Data management policies, Use and Access Policies, SOPs
5.9 Prefabricated PP Modules	Under development
5.10 Distributed Authentication System	Useful
5.11 Authorization Record Registry System	Useful
5.12 OAI-PMH, ResourceSync, SRU/CQL	Under development (Labimi-F)
5.13 Workflow Engine & Environment	Under development (SFB: Echocardiography of Mice)
5.14 Conversion Tool Registry	Useful
5.15 Analytics Component Registry	Useful
5.16 Repository API	Useful
5.17 Repository System	Local Consortium Repositories
5.18 Certification & Trusted Repositories	Useful



Discussion



RDA Requests for comments - Data Management Plans

Desirable:

- Unification of general principles/ guidelines
- Clarify difference between Backup and re-usable archive
- Rules for responsibilities: Who makes the plan?
- Advice for putative amount of data (often not known by scientists at project start)
- Change from static to Several-Step-Adaption of Data management plans (e.g. Before project start, after acquiring initial requirements,...)
- Not only technical plans are necessary (storage, machines, software licenses)
BUT ALSO need/money for coordinating staff to
 - determine specific requirements
 - explore strategies/concepts of implementation



Discussion

RDA Requests for comments - Data sharing and Re-use



Desirable:

- Overcome low acceptance of data sharing in clinical science
 - Is clinical data special?: Sharing of Cross-Discipline Knowledge
 - Citability: Repositories for medical/clinical data
 - Versioning of datasets, Transparency of error corrections



Forschungsdatenmanagement in der klinischen Forschung

Dr. Harald Kusch & Karoline Buckow

Institut für Medizinische Informatik
UNIVERSITÄTSMEDIZIN GÖTTINGEN



RDA/DINI-Workshop – 28. Mai 2015

<http://orcid.org/0000-0002-9895-2469>

Institut für Medizinische Informatik
UNIVERSITÄTSMEDIZIN GÖTTINGEN





Diese Folie wird nicht im Vortrag verwendet sondern dient zur Dokumentation der Gliederung und Ziele des Vortrags!

Gliederung



Zahlreiche Fachverbände und Projekte im Rahmen der Medizin befassen sich seit Jahrzehnten mit der Erarbeitung von Konzepten und Entwicklungen, um Forschungsdaten, die in der klinischen Forschung erhoben werden, strukturiert zu erfassen und mit zunehmendem Maße auch dauerhaft - mindestens über die Projektlaufzeit hinaus - verfügbar zu machen. In diesem Vortrag wird erörtert, inwieweit die Aktivitäten in der Medizin(informatik), exemplarisch am Beispiel eines deutschen Studienzentrums in Göttingen, mit den Diskussionen und Ergebnissen der Research Data Alliance (RDA) in Einklang stehen. Dies soll dazu beitragen, die fachübergreifenden Konzepte der RDA um Sichtweisen aus der medizinischen Forschung zu erweitern. Der Vortrag ist nach folgenden Fragestellungen gegliedert:

1. Was sind spezifische Anforderungen an die Forschungsdaten-Infrastruktur in der medizinischen Forschung? Sind medizinische Forschungsdaten anders als Daten anderer Fachbereiche?
2. Welche Fachverbände und Förderungsinitiativen der medizinischen Forschung sind im Bereich des Forschungsdaten-Managements aktiv (beispielhafte Darstellung)?
3. Welche Lösungsansätze/Bearbeitungsschwerpunkte werden im Institut für medizinische Informatik in Göttingen bearbeitet, um die Forschungsdaten-Infrastruktur in der medizinischen Forschung voranzubringen?
4. Wie gestaltet sich die konkrete Erarbeitung von Maßnahmen zur Optimierung der Forschungsdaten-Infrastruktur bei der Förderungsform "SFB-INF-Projekt"?
 - a) Was sind fachübergreifende Herausforderungen (aus Sicht aller INF-Projekte)?
 - b) Was sind Herausforderungen im medizinischen Kontext (am Beispiel des klinischen SFB 1002)?
5. Welche Bearbeitungsschwerpunkte der RDA-Initiativen sind für die medizinische Forschung wichtig? Welche dieser Schwerpunkte werden bereits (in der Göttinger Medizininformatik) erarbeitet?

