

# Graphische Aufarbeitung von Workflow- und Data-Provenance

**Hintergrund** Die Nachvollziehbarkeit und Reproduzierbarkeit von Forschungsergebnissen ist eine Grundvoraussetzung für qualitativ gesicherte Wissenschaft. In allen Bereichen, die mit komplexen Daten- und Prozessflüssen arbeiten — von Labor-Experimenten bis hin zu datenintensiven Analyse-Workflows — gilt **Provenance** als unverzichtbare Informationseinheit. Sie beschreibt die Herkunft der Daten, Verarbeitungsschritte, beteiligte Akteure und eingesetzte Werkzeuge. Klassischerweise werden zwei Arten von Provenance unterschieden:

- **Workflow Provenance:** strukturelle Abfolge von Aktivitäten und Forschungsobjekten, üblicherweise als Graph dargestellt.
- **Data Provenance:** Provenance auf Daten-Ebene (Zeugenbasen, Provenance-Polynome).

Eine einheitliche, graphische Darstellung, die beide Provenance-Arten verknüpft, fehlt bislang. Ein solcher Ansatz würde Transparenz, Interoperabilität und die Grundlage für weitere Analysen deutlich verbessern.

**Aufgabenstellung** Ziel der Bachelorarbeit ist die Weiterentwicklung eines konzeptionellen Modells (basierend auf Auge et al. (Biomedical)) und eines prototypischen Tools zur graphischen Aufarbeitung von Workflow- und Data-Provenance-Informationen. Die Arbeit verfolgt dabei drei Hauptziele:

1. Weiterentwicklung des gegebenen Modells zur gemeinsamen Erfassung von Workflow- und Data-Provenance.
2. Prototypische Implementierung dieses Modells in einem RDF-basierten Knowledge-Graphen. Erstellung eines interaktiven Provenance-Graphen
3. Validierung des entwickelten Konzepts anhand eines realen Datensatzes.

Die Arbeit trägt zur Standardisierung von Provenance-Modellen bei und fördert die nachhaltige Nutzung heterogener wissenschaftlicher Datenbestände.

## Formalia

- Ansprechpartner:
  - Tanja Auge (Universität Regensburg, [tanja.auge@ur.de](mailto:tanja.auge@ur.de))
  - Sascha Genehr (Universität Rostock, [sascha.genehr@uni-rostock.de](mailto:sascha.genehr@uni-rostock.de))
  - Max Schröder (Universitätsbibliothek Rostock, [max.schroeder@uni-rostock.de](mailto:max.schroeder@uni-rostock.de))
- Voraussetzungen:
  - Interesse an der Entwicklung von Provenance-Konzepten für real world-Szenarien
  - Motivation zur eigenständigen, kreativen Lösung abstrakter Modellierungsprobleme
  - Fähigkeit zur eigenständigen Recherche und Auswertung wissenschaftlicher Literatur
  - Erste Erfahrungen mit Provenance, dem PROV-Standard und RDF (alternativ Bereitschaft zur vertieften Einarbeitung)
  - Erfahrungen mit graphenbasierten Visualisierungstools (alternativ Bereitschaft zur vertieften Einarbeitung)

## Literatur

- Auge et al.: Provenance for Lattice QCD Workflows – An Update. <https://dl.acm.org/doi/10.1145/3736229.3736268>
- Auge et al.: Provenance Levels for Biomedical In-vitro Experiments. <https://zenodo.org/records/16736028>
- Schröder et al.: Structure-based knowledge acquisition from electronic lab notebooks for research data provenance documentation. <https://doi.org/10.1186/s13326-021-00257-x>
- Genehr et al.: Modelling three dimensions of provenance for wet-lab experiments: prospective, retrospective, and evolution. <https://repository.publisso.de/resource/frl:6444990>