

# Eine Forschungsdateninfrastruktur für die Materialwissenschaften

Nico Brandt, Ephraim Schoof, Philipp Zschumme, Michael Selzer, Britta Nestler

## Motivation

- In den **Materialwissenschaften** wird das Verständnis neuer Materialien zunehmend komplexer.
- Ohne geeignete **Data Science-Methoden** werden die ständig wachsenden Datenmengen aus Simulationen und Experimenten nicht mehr erfassbar sein.
- Ein wichtiger Aspekt, um entsprechende Datenanalysen reibungslos durchführen zu können, ist die strukturierte Ablage von Forschungsdaten und entsprechender Metadaten mithilfe einer geeigneten **Forschungsdateninfrastruktur**.

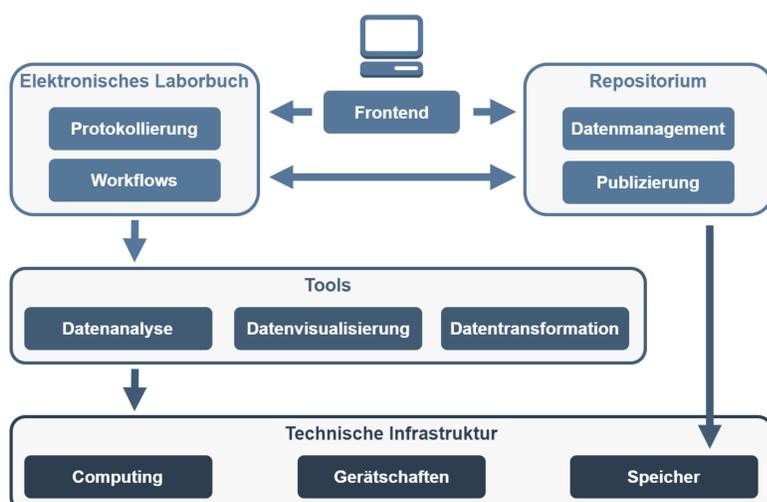
## Implementierung



- **Kadi4Mat** wird am *IAM-CMS* im Kontext mehrerer Forschungsprojekte entwickelt.
- Ein erster Demonstrator der Dateninfrastruktur ist erreichbar unter <https://kadi4mat.iam-cms.kit.edu/>
- Der bisherige Fokus lag auf dem Aspekt Datenmanagement sowie der Nutzung prototypischer Data Science-Tools.

## Forschungsdateninfrastruktur

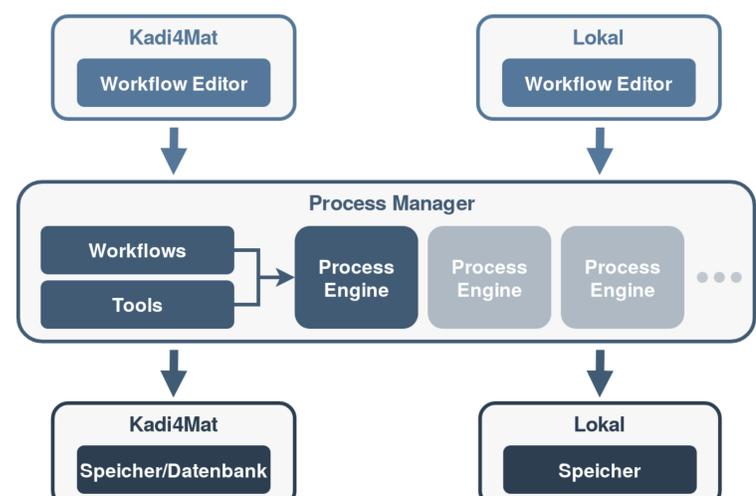
- Für die Forschungsdateninfrastruktur ist eine modulare Architektur geplant, die die beiden Komponenten **Repository** und **Elektronisches Laborbuch (ELN)** verbindet.



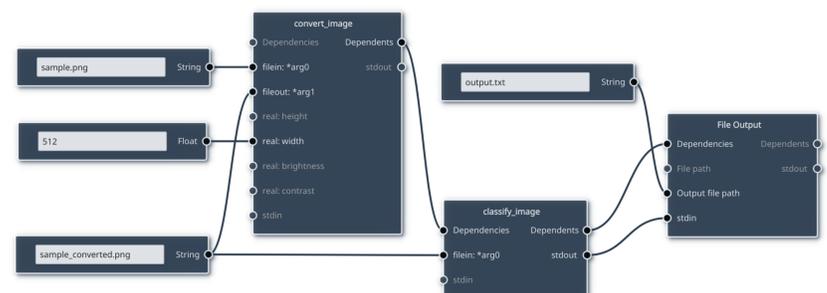
- **Die folgenden Entwicklungen stehen dabei im Fokus:**
  - Strukturierte Datenablage und webbasierter Datenaustausch
  - Publizierung und Referenzierung von Forschungsdaten
  - Ausführung und Dokumentation reproduzierbarer Workflows zur Nutzung beliebiger, anwendungsspezifischer Tools
  - Integration vorhandener Systeme sowie spezifischer Konvertierungstools für den Umgang mit heterogenen Daten aus Simulationen und Experimenten
  - Aufbau eines Kompendiums datenwissenschaftlicher Methoden, insbesondere des maschinellen Lernens
- Das Ziel ist es, etablierte Technologien und Systeme mit neuen Konzepten zu verknüpfen.

## Workflows

- Workflows in Kadi4Mat repräsentieren die automatisierte Datenerfassung und -analyse als zentraler Bestandteil des ELNs.



- Die Erstellung von Workflows kann lokal oder zukünftig direkt unter Verwendung des Webinterfaces erfolgen.
- Ein- und Ausgabedaten sowie entsprechende Metadaten lassen sich lokal oder über geeignete Schnittstellen im Repository von Kadi4Mat abspeichern.



## Zusammenfassung

- Das Forschungsdatenmanagement mit Kadi4Mat steigert die Kooperation zwischen Forschern unter Berücksichtigung der **FAIR-Datenphilosophie** (*Findable, Accessible, Interoperable, Reusable*).
- Zu den weiteren Zielen gehört es, den gesamten wissenschaftlichen Arbeitsablauf der täglichen Forschungsarbeit elektronisch zu erfassen.

### Kontakt:

Nico Brandt • [nico.brandt@kit.edu](mailto:nico.brandt@kit.edu)  
Karlsruher Institut für Technologie – IAM-CMS



Gefördert durch die *Deutsche Forschungsgemeinschaft* im Projekt *POLiS* (Förderkennzeichen 390874152) sowie durch das *Bundesministerium für Bildung und Forschung* im Projekt *FestBatt* (Förderkennzeichen 03XP0174E).