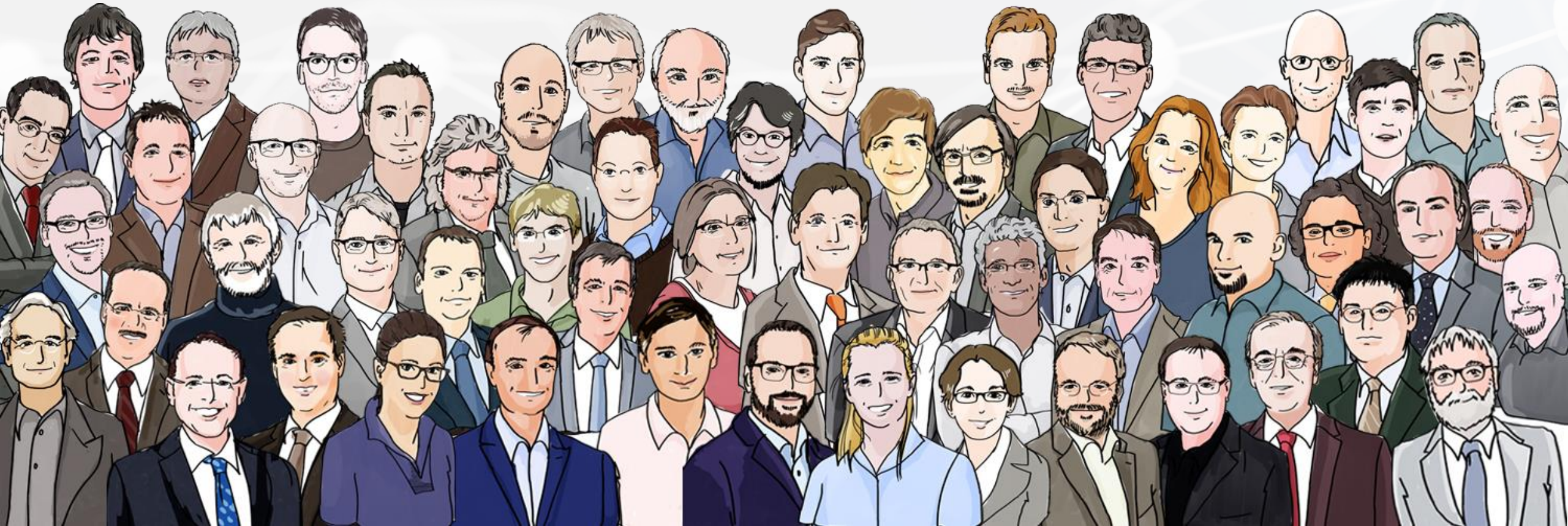
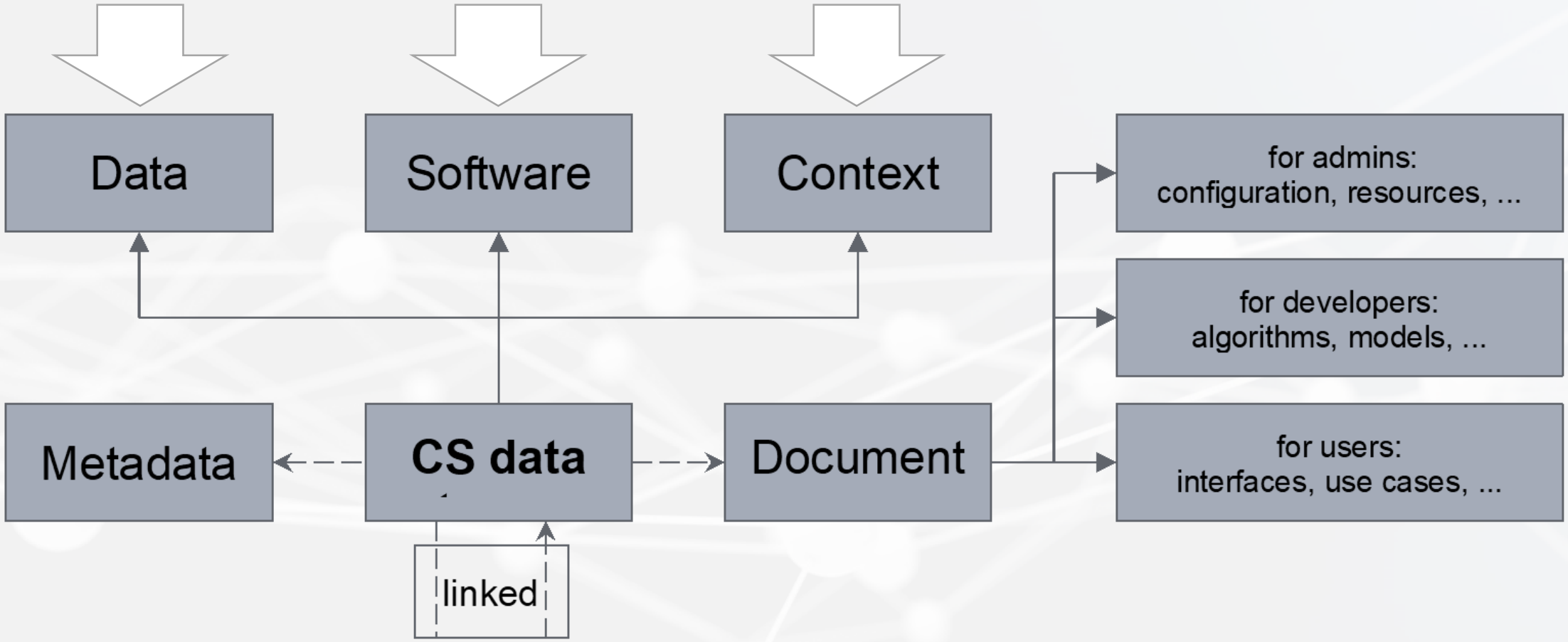
Forschungsdatenmanagement aus der Sicht der Informatik

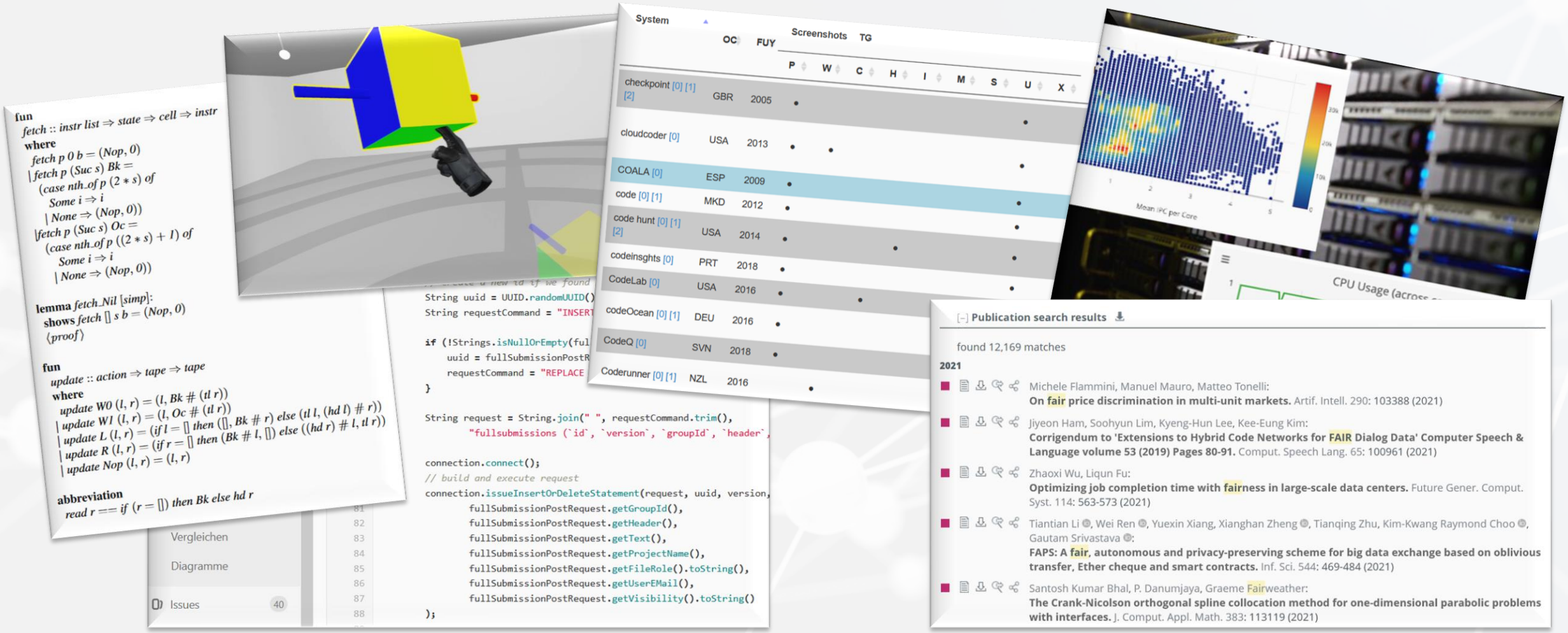
GI/GMDS Workshop „Perspektiven des Forschungsdatenmanagements“



„Die Informatik hat eigene Forschungsdaten?!?“



Beispiele für Informatik-Forschungsdaten

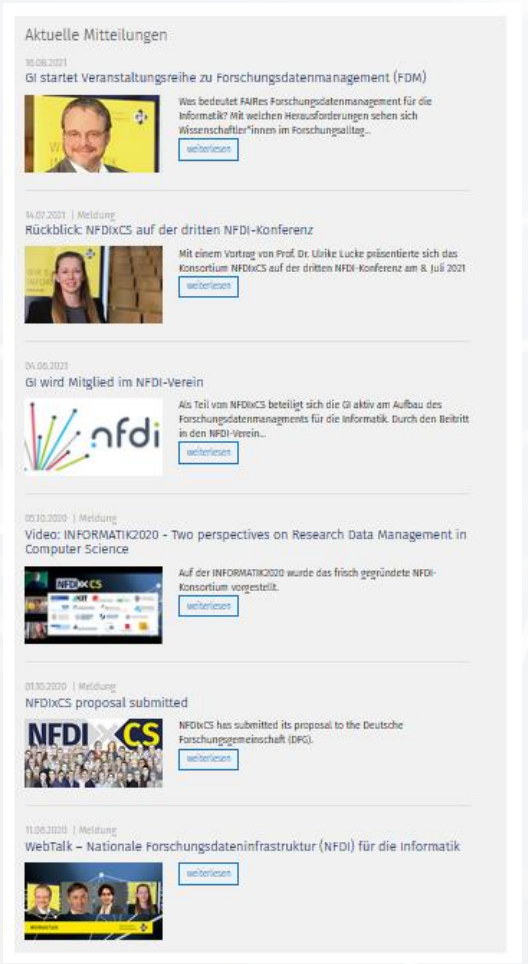


Repositories für Informatik-Forschungsdaten

Archive of Formal Proofs
arXivand
B2SHARE
Bildungsserver
Camunda Modeler
CDSTAR
Cognitive Process Designer
Cross Language Evaluation Forum
Dagstuhl Research Online Publication Server
Dataverse
dblp computer science bibliography
de.NBI
ELIXIR
EMBL-EBI
FAIRdom
GitHub
GitLab repositories
IACR Cryptology ePrint ICSE artifact evaluation track
Invenio
iRODS
koala long-term archiving
Lean mathlib
Learning Online Network
Linked Open Vocabularies
EMBL-EBI
Mizar Mathematical Library
OER repositories
Open Research Knowledge Graph
OSSIRC
Parallel Workloads Archive
PIKA
PROMISE
ReMoDD
meta-data on programming education systems
Repository of solutions to programming exercises
xAPI definitions
Score-P
SSELab
SPEC
SNIA/IOTTA
TIRA
TREC Data Archive
Virus total
xAPI
…

Genese des NFDIxCS Konsortiums

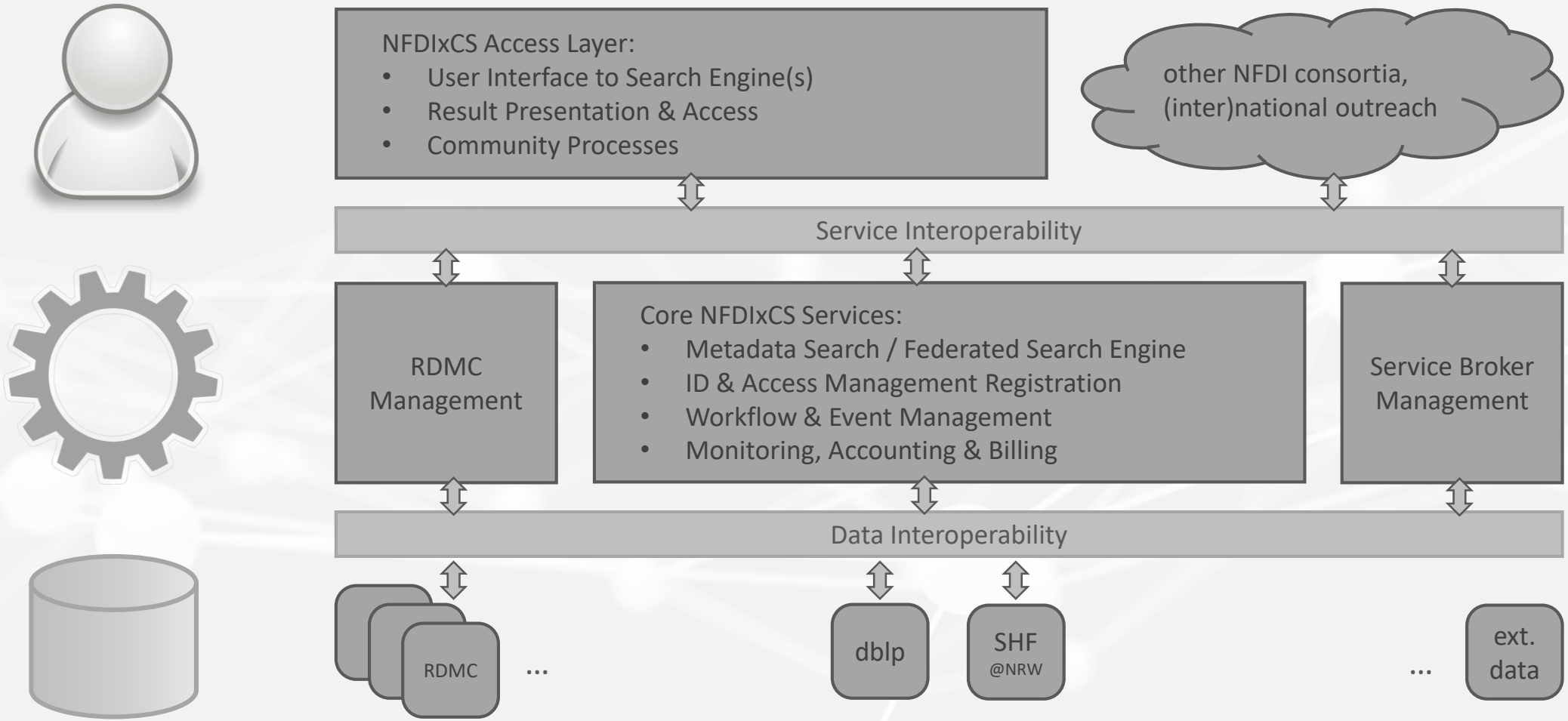
- Gesellschaft für Informatik – die Fachgesellschaft für Informatik in Deutschland (seit 1969)
- 14 Fachbereiche und über 150 Fachgruppen
  im gesamten Bereich der Informatik und ihrer Anwendungen
- Präsidiumsarbeitskreise: *E-Science* and *Data Science*
- aktive Mitwirkung im NFDI-Prozess von Beginn an (LoI 2019)
- Umfragen und Diskussionen innerhalb der GI 2019/20
- Beteiligung an der Bearbeitung von Querschnittsthemen (Berlin-Leipzig Declaration)
- Web-Talks zu Informatik-spezifischen Aspekten des Forschungsdatenmanagements
- Beitritt zum NFDI-Verein (2021)



Konsortium und Services

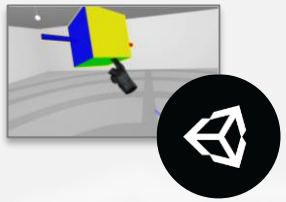
- Balance verschiedener Interessengruppen:
  - Forschung
    (schnelle Anpassung an heterogene und dynamische Bedarfe)
  - Betrieb
    (stabile, sichere, finanzierbare und wartbare Infrastruktur)
- disziplinäre Expertise + leistungsfähige Tools
  aus verschiedenen Perspektiven
- Wiederverwendung, Anpassung, Verbindung bestehender Dienste
  so weit wie möglich
- hoher Automatisierungsgrad für einen nachhaltigen Betrieb

NFDIxCS Architektur

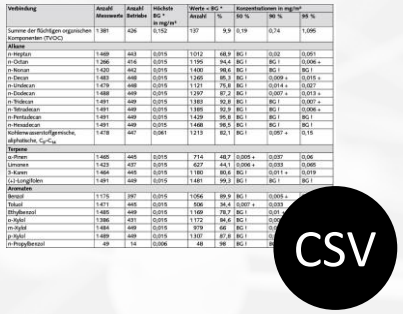


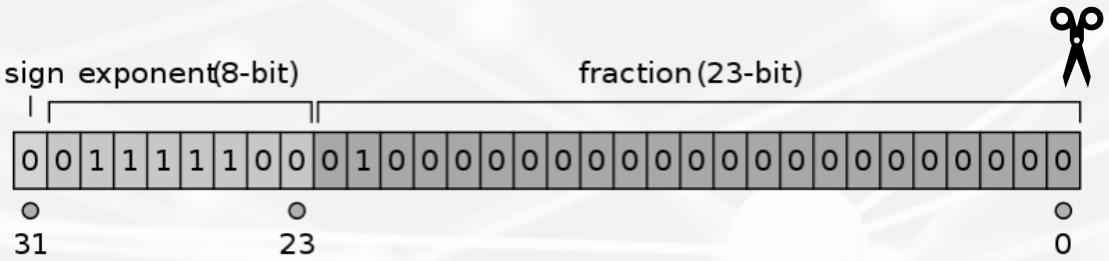
Research Data Management Container

Proprietäre Formate sind oft an bestimmte Software gebunden.



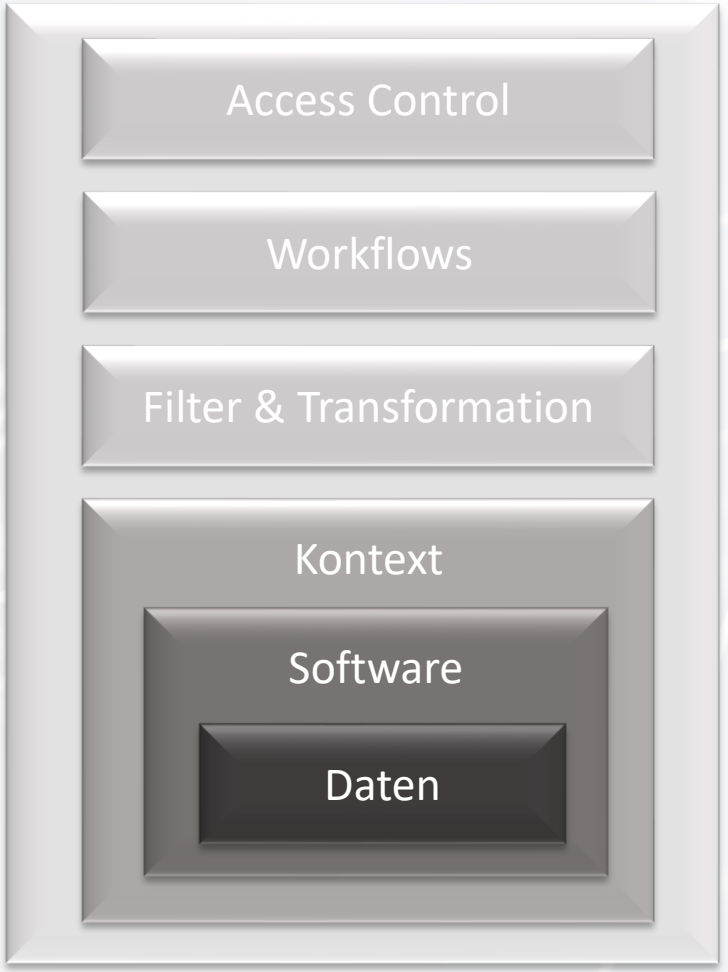
Selbst offene Formate können an bestimmte Software-Versionen gebunden sein!





Software ist an eine bestimmte Ausführungsumgebung gebunden.

→ Kapselung dieser Schichten in Research Data Management Containern (RDMC)



Weiterentwicklung der Community

Anschub durch zentrale Akteure:

- Kooperation mit Konferenzen und Verlagen
- Qualitätssicherungsmechanismen

Unterstützung für den wissenschaftlichen Nachwuchs:

- Informationskampagnen & Bewusstseinsbildung
- Schulungen und Support
- Networking
- Konferenzen & Publikationen
- Awards

Nutzung etablierter Werkzeuge:

- RISE Framework
- RMDO Web-Tool
- Data Carpentry lessons

Mission Statement

- Diskussion und Standardisierung von Formaten (Forschungsdaten und Metadaten) in der Informatik und deren Semantik
- Bereitstellung geeigneter Infrastrukturen, Werkzeuge und Dienste für Informatik-Forschungsdaten:
  - Umsetzung der FAIR Data Prinzipien für Informatik-Daten und Software vorantreiben
  - Zitierbarkeit von Software und Informatik-Daten erleichtern
  - Publikationsprozesse und –kultur in der Informatik und ihren Anwendungen modernisieren
- Unterstützung für alle Teildisziplinen im Umgang mit deren Forschungsdaten
- Einbezug von Daten anderer Fächer zur Weiterentwicklung von Informatik-Forschungsmethoden
- Austausch von Erfahrung und Wissen rund um Informatik (Systemarchitektur, Prozesse, Standards für Interoperabilität, Kommunikationssysteme usw.) mit anderen Konsortien

Mehr Informationen

NFDIxCS.org

Prof. Dr. Michael Goedicke (Universität Duisburg-Essen) – Sprecher
michael.goedicke@paluno.uni-due.de

Prof. Dr. Ulrike Lucke (Universität Potsdam) – stellvertretende Sprecherin
ulrike.lucke@uni-potsdam.de