# Information Infrastructures in the Era of AI: Opportunities and Challenges

## Workshop

by the Leibniz Science Campus

***Digital Transformation of Research* (DiTraRe)**

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

Karlsruhe Institute of Technology

funded by Leibniz Competition

E-Science-Tage, 12-14 March 2025, Heidelberg

# Agenda

Leibniz ScienceCampus
Digital Transformation of Research

- **Ice-breaker** session
- Introduction of the Leibniz Science Campus **DiTraRe**
- Presentations of **experts from different disciplines**
- **Fishbowl** discussion
- Workshop **summary**

Our main questions to be addressed in the topics of **information infrastructures** and **Artificial Intelligence**:

1. What are the opportunities?
2. What are the challenges?

Harald Sack
*DiTraRe Spokesperson*

Anna Jacyszyn
*DiTraRe Coordinator*

Felix Bach
*DiTraRe Coordinator*

# Ice-breaker session

Leibniz ScienceCampus
Digital Transformation of Research

Let's meet! Who of you…

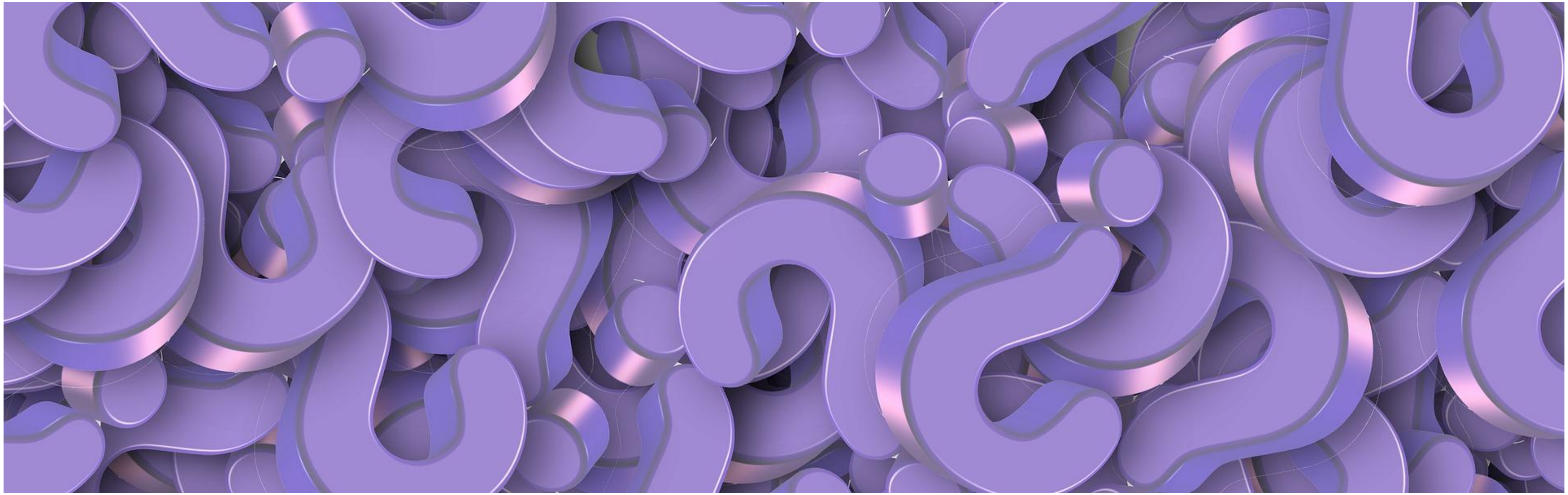


*image by The Digital Artist licensed under the [pixabay content license]*

# Ice-breaker session

Leibniz ScienceCampus
Digital Transformation of Research

Thought Experiment:

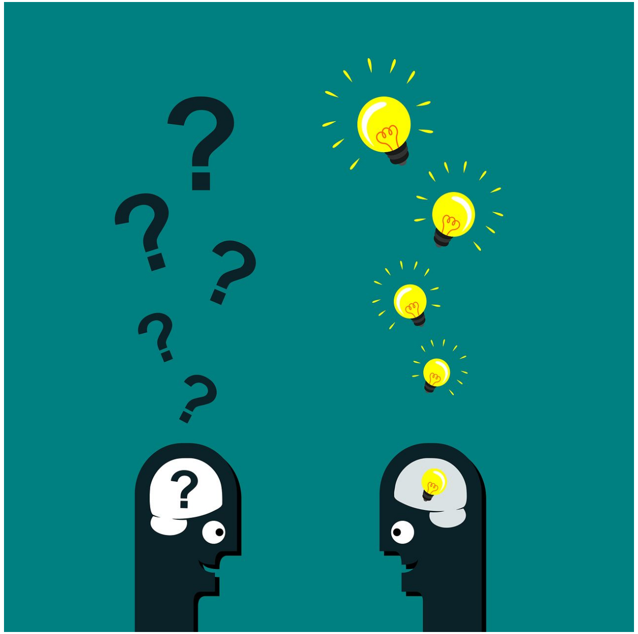


*image by: nugroho dwi hartawan licensed under the [pixabay content license]*

# Ice-breaker session

Leibniz ScienceCampus
Digital Transformation of Research

## Thought Experiment:

Let’s see your answers!

urlr.me/!ditrare-menti

# Introduction of DiTraRe

# DiTraRe

## Leibniz Science Campus *Digital Transformation of Research*

Leibniz ScienceCampus
Digital Transformation
of Research

- Growth core to **establish new research branch**.

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

+

KIT
Karlsruhe Institute of Technology

- Planned as a **4+4 years** project (start: September 2023).
- Funded by the Leibniz Association + FIZ KA + KIT.
- Analyse the process of **digitalisation of research**.
- **Multilevel interdisciplinary approach**. → We start with 4 specific use cases.

# DiTraRe Use Cases

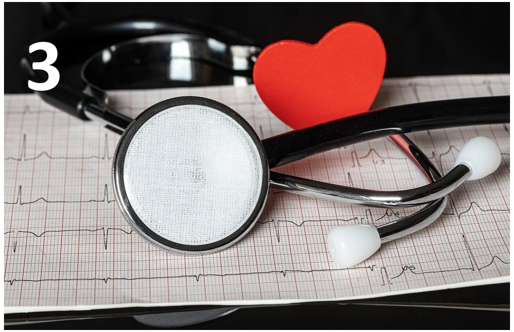
Leibniz ScienceCampus
Digital Transformation
of Research


Sensitive Data in Sports Science

*KIT Institute of* ***Sports and Sports Research***


Chemotion Electronic Lab Notebook

*KIT Institute of* ***Biological and Chemical Systems***




AI in Biomedical Engineering

*KIT Institute of* ***Biomedical Engineering***


Publication of Large Datasets

*KIT Institutes of* ***Meteorology and Climate Research***

# DiTraRe dimensions

Leibniz ScienceCampus
Digital Transformation of Research

**A. Reflection and Resonance**
a dialogue between research and society, interactive process

**B. Exploration and Knowledge Organisation**
applied AI: represent, organise, and manage domain specific and procedural knowledge

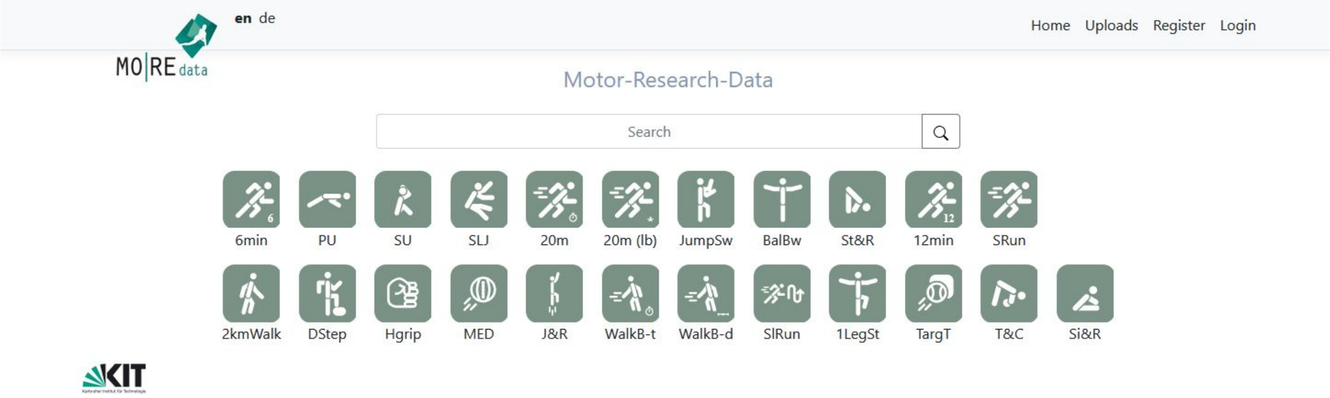
**C. Legal and Ethical Challenges**
data ethics, data protection, copyright and data law

**D. Tools and Processes**
digital tools tailored precisely to the needs of researchers

# Experts Flash Presentations

# Use case: Sensitive data in sports science

## *KIT Institute of Sports and Sports Science*




1

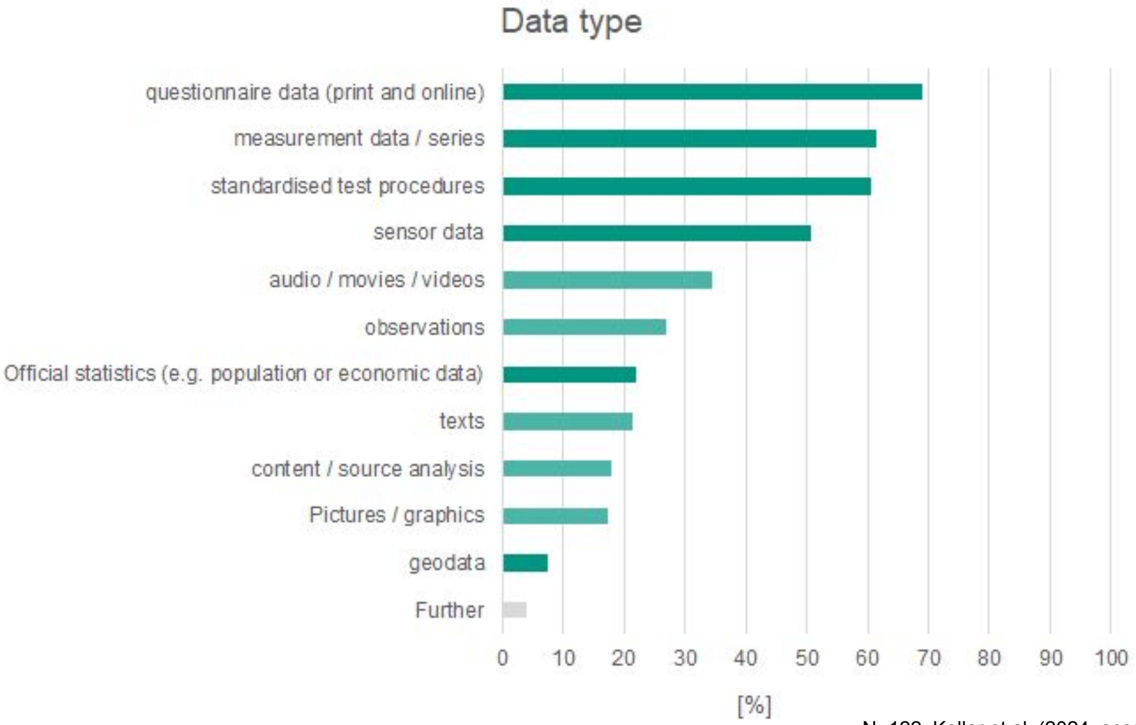
Hannah Zimmermann

- MO|RE data = first repository in sports science
- Functions: Archiving, Publication (with DOI), Download/Re-Use (after registration)
- Today over 68,000 raw data available
- BUT:
  - only motor research data
  - only anonymised data

# Use case: Sensitive data in sports science

*KIT Institute of Sports and Sports Science*

1




N=122, Keller et al. (2024, accepted)

Hannah Zimmermann

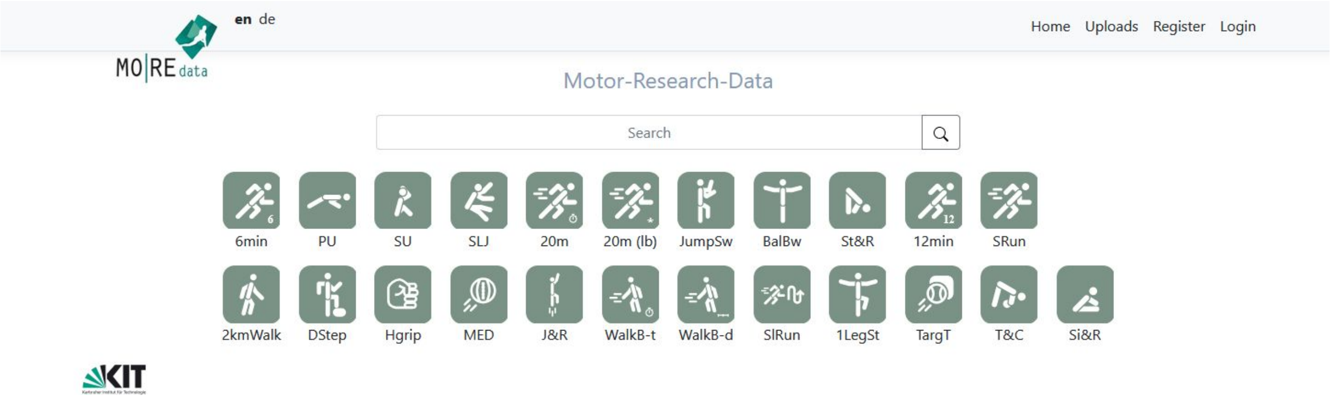
Great variance of data:

- Quantitative and qualitative
- Different formats / software
- Different levels of sensitivity
- Range from small to large data sets
- …

# Use case: Sensitive data in sports science

*KIT Institute of Sports and Sports Science*

1




Hannah Zimmermann

- Expand for further data types
- Implementation of new access paths for sensitive data
- AI usage: Building an ontology for motor performance data → Knowledge Graph

# Use case: Chemotion Electronic Lab Notebook

*KIT Institute of Biological and Chemical Systems*

2


Electronic Lab Notebooks offer an enormous potential to accelerate scientific work: but the acceptance of ELNs and connected workflows is rather low.

## Use case: Chemotion Electronic Lab Notebook

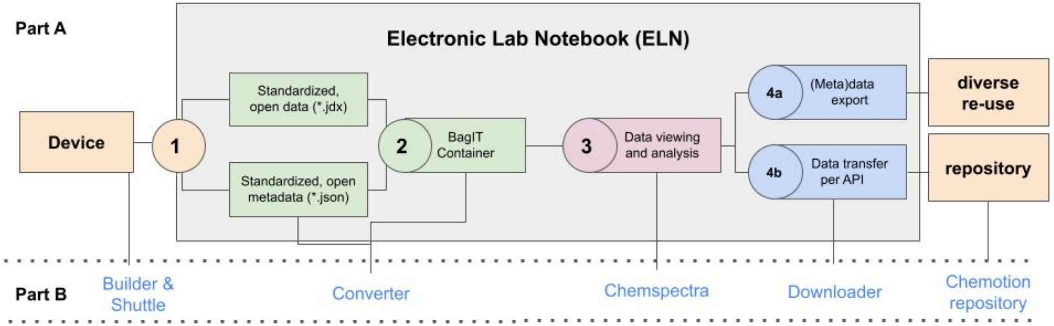
2

*KIT Institute of Biological and Chemical Systems*

Nicole Jung

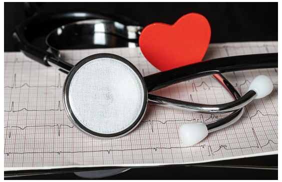
### Measures within the use case

- Improve the suitability and application area of data conversion tools
- Implement standards for data handling




- Improve data curation workflows:
  - implementation of rules and validation, LLM based review of submissions

# Use case: AI for Biomedical Applications



*KIT Institute of Biomedical Engineering*

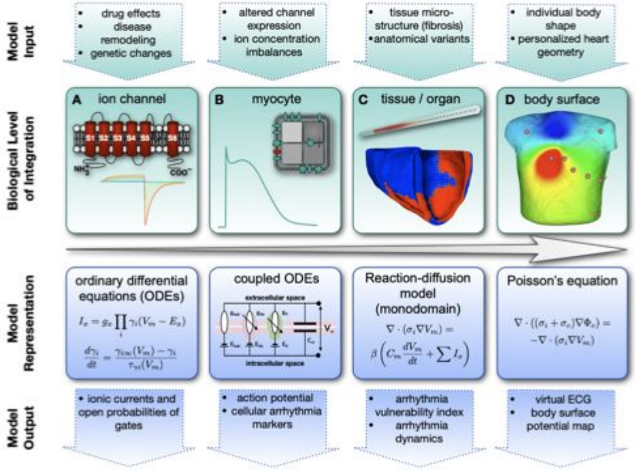
Big data and machine learning hold great potential to
- Speed up diagnosis & make it more accurate
- Personalize treatment
- Reduce costs & redistribute human effort

Silvia Becker

Axel Loewe

Challenge: “big data but small data” paradox
- Data access restrictions
- Insufficient label quality and data curation

Potential solution: synthetic data




Loewe, Hunter, Kohl: “Computational modelling of biological systems now and then”, 2025, doi:10.48550/arXiv.2501.13142

# Use case: Publication of large datasets

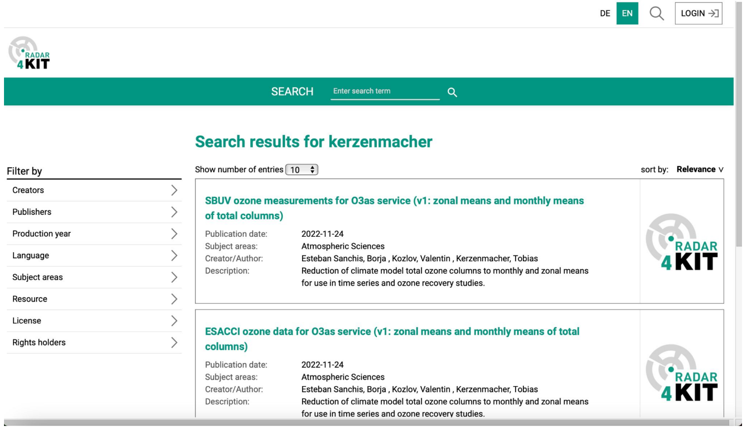
## *KIT Institutes of Meteorology and Climate Research*

Challenges & Innovations in Data Accessibility and Standardisation:

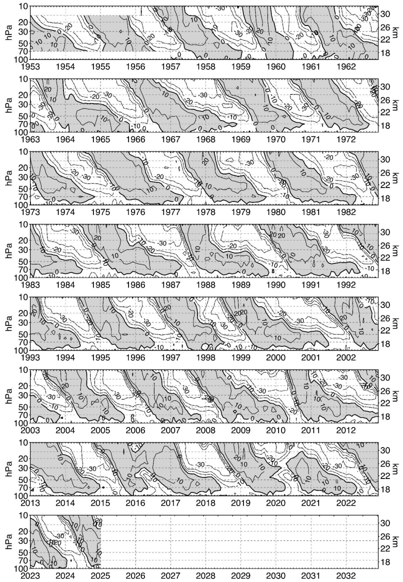
- Collaboration: KIT-IMK & RADAR Repository
- Key Challenge: Managing & publishing large atmospheric datasets

Tobias Kerzenmacher

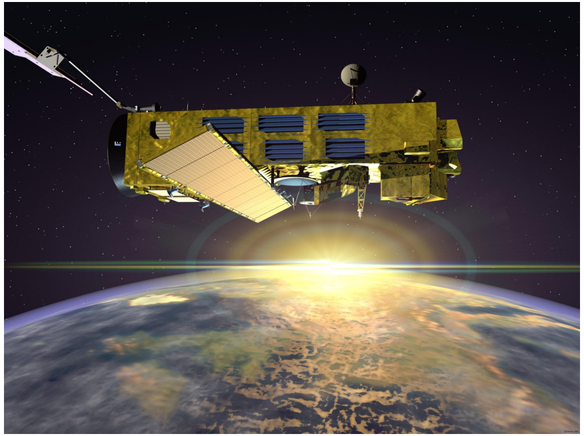
*Data repository: [radar.kit.edu](radar.kit.edu)*




*Tropical stratospheric zonal wind in Singapore*




*MIPAS on Envisat (ESA)*



# DiTraRe dimension: Reflection and Resonance

## *KIT Institute for Technology Assessment and Systems Analysis*

Leibniz ScienceCampus Digital Transformation of Research

**Generative AI as a "game changer in AI" for science with its ability to generate novel content**

- New chances for knowledge generation and processing
  - Increased efficiency and productivity in scientific working practices
  - Potentials in training small data sets, LLM-based knowledge extraction, programming tools
- But: reflections on different levels needed
  - Technical limits: Detection, authentication, "hallucinations"
  - Social aspects: AI-literacy; potential loss of competences
  - Legal issues: Data protection, limits of regulation
  - Ethical questions: Overreliance, false information, biased knowledge, misrepresentation
- Information infrastructure & AI – ensuring research integrity
  - Knowledge generation: How to deal with challenged concepts of reliability in RDM?
  - Human-AI hybrid: Which steps in the research process require human verification?
  - Options for action: Development of common standards or guidelines?

Linda Nierling

Angelina Sophie Dähms

Information infrastructures and AI: Opportunities? Challenges?

# DiTraRe dimension: Exploration & Knowledge Org.

*FIZ Karlsruhe – Leibniz Institute for Information Infrastructure*

Leibniz ScienceCampus Digital Transformation of Research

- **Generative AI**
  - *Opportunities:* supports automatisation across different disciplines
  - *Challenges:* researchers no longer navigate development of LLMs
- **Knowledge graphs and ontologies**
  - *Opportunities:* encode domain knowledge, interconnect knowledge from different disciplines, enable knowledge discovery
  - *Challenges:* difficult accessibility, costly resources (expertise needed)

*source: https://www.nextplatform.com/wp-content/uploads/2020/04/ab_network-graph-1030x438.jpg*

Information infrastructures and AI: Opportunities? Challenges?

Harald Sack

Anna Jacyszyn

# DiTraRe dimension: Legal and Ethical Challenges

*FIZ Karlsruhe – Leibniz Institute for Information Infrastructure*

Leibniz ScienceCampus
Digital Transformation of Research

- **opportunities:**
  - creation of laws that are progressive instead of reactive
  - creating a “safespace” for data flow

- **challenges:**
  - legal requirements hinder progress
  - progress ignores legal requirements
  - international discrepancies due to different legislature

dynamic/ unpredictable character of AI and data protection as it’s opponent

*graphic from Gerd Altmann licensed under the pixabay content license*

Information infrastructures and AI: Opportunities? Challenges?

Lea Sophie Singson

Ely Nova Natalia Silaban

# DiTraRe dimension: Tools and processes

*FIZ Karlsruhe – Leibniz Institute for Information Infrastructure*

Leibniz ScienceCampus
Digital Transformation of Research

Incorporation of AI into **software & research processes**

Electronic Laboratory Notebooks (**ELN**)s

Research Data **Repositories**

- Automated MD generation, expansion and indexing
- Enhancing data security
- Improving data query and performance
- Data entry, cleaning, and normalisation
- Voice-based AI assistants
- AI algorithms can process and analyse large-scale datasets

Automatic generation/attribution of data **FAIRness** score

Information infrastructures and AI: Opportunities? Challenges?

Christian Bonatto Minella

Felix Bach

# Fishbowl Discussion

# Fishbowl discussion

Leibniz ScienceCampus
Digital Transformation
of Research


urlr.me/!ditrare-hedgedoc

# Wrap-up of the discussion

Leibniz ScienceCampus
Digital Transformation of Research

- Who is the recipient of your data?
- Legal issues: Privacy and Copyright
- Re-identification, modeling and clinical data
- Repositories: ensure the data and clear licence
- Data are for public domain!?
- Field specific metadata are necessary
- AI for medical sciences
- Regulation for AI often unclear
- Data sharing hesitation
- Computer scientists often not able to keep track on data sources
- Clear licence on data re-use for training

# Join our talks!

Leibniz ScienceCampus
Digital Transformation of Research

Thursday

- 11:50-12:30 Session A2 *Data Law*
    - "**Research Data Management in Medical Data Science**: Integrating Multi-Source Data for Cardiological Simulations and Machine Learning: Challenges in Security, Privacy, and Legal Frameworks"
- 16:20-17:00 Session A6 *Digitalisation*
    - "**The Impact of Digitalisation on Science: Workflows, Outputs, and Trust**"

Friday

- 11:30-12:30 *Lightning Talks*
    - "**AI4DiTraRe: Studying Applied AI Within the Process of Digitalisation of Research**"

# Thank you for attending!

Leibniz ScienceCampus
Digital Transformation of Research

- **Symposium** on digitalisation of research: 2-3 December 2025, Karlsruhe.
    - Interdisciplinary panels, discussions with experts!
- **Interdisciplinary colloquium** on digitalisation of research: starting in May.
    - Bimonthly presentations of experts, join us online or in person!
- Stay **connected** with DiTraRe
    - Website: www.ditrare.de/en
    - Email: ditrare@fiz-karlsruhe.de
    - LinkedIn: www.linkedin.com/company/ditrare
    - Mastodon: sigmoid.social/@DiTraRe
    - Zenodo: zenodo.org/communities/ditrare

FIZ Karlsruhe
Leibniz Institute for Information Infrastructure

KIT
Karlsruhe Institute of Technology

funded by Leibniz Competition