AI4DiTraRe: Studying Applied AI across Leibniz Science Campus DiTraRe Use Cases

Anna Jacyszyn [1*] and Harald Sack [1,2]

[1]FIZ Karlsruhe – Leibniz Institute for Information Infrastructure, Eggenstein Leopoldshafen, Germany
[2]Institute of Applied Informatics and Formal Description Methods, Karlsruhe Institute of Technology, Karlsruhe, Germany
[*]Corresponding author: Anna.Jacyszyn@fiz-Karlsruhe.de.

Leibniz Science Campus DiTraRe

The Leibniz Science Campus *Digital Transformation of Research* (DiTraRe)[1] [1] is a collaboration between FIZ Karlsruhe - Leibniz Institute for Information Infrastructure (FIZ KA) and Karlsruhe Institute of Technology (KIT). The consortium investigates the effects of the broadly understood process of digitalisation of research on a multilevel scale in an interdisciplinary environment. We are not only developing practical solutions for each use case but also seeking to find generalisations valuable to the scientific community as well as society in general [2].

Our main goal is to establish a new research branch, *digitalisation of research*. In order to achieve it, we create an interdisciplinary community. DiTraRe starts with four use cases originating from different institutes of KIT:

1. Institute of Sports and Sports Science (IfSS), use case: *Sensitive Data in Sports Science*.
2. Institute of Biological and Chemical Systems (IBCS), use case: *Chemotion Electronic Lab Notebook*.
3. Institute of Biomedical Engineering (IBT), use case: *AI in Biomedical Engineering*.
4. Institute of Meteorology and Climate Research (IMK), use case: *Publication of Large Datasets*.

Each use case is being investigated by each of the four DiTraRe *dimensions*:

A *Reflection and Resonance*, KIT Institute for Technology Assessment and Systems Analysis (ITAS).

B *Exploration and Knowledge Organisation*, FIZ KA.

C *Legal and Ethical Challenges*, FIZ KA.

D *Tools and Processes*, FIZ KA.

With the current rise of AI applications across multiple fields, we can state that AI acts as an additional facet to the digital transformation of research. The DiTraRe dimension *Exploration and Knowledge Organisation*, AI4DiTraRe [3], studies methods of applied AI in each of the DiTraRe use cases. Below we describe the current state, idea and future plans of each of the working group.

Use Cases

Sensitive Data in Sports Science, partners: Sarah Rebecca Ondraszek (FIZ KA), Jörg Waitelonis (FIZ KA), Katja Keller (KIT IfSS), Claudia Niessner (KIT IfSS), Chris Rose (KIT IfSS).

Our use case partners, researchers at KIT IfSS, have developed a Motor Research repository (MO|RE)[2]. It is a platform to collect, publish, and share motor performance data [4, 5]. To make it completely interoperable and connect it to other databases, it lacks a semantic layer. The team has developed an idea of a MO|RE ontology[3], on which later a knowledge graph (KG) will be built. The

[1]DiTraRe webpage, `https://www.ditrare.de/en`

[2]MOwebpage `https://www.motor-research-data.de/`

[3]GitHub repository of the MO|RE ontology `https://github.com/ISE-FIZKarlsruhe/more-ontology`

details of the first steps are presented by Ondraszek et al. 2025 [6].

The MO|RE ontology is in its preliminary stage and the team is planning to develop it further. Since motor performance data contains private and sensitive information, special efforts will need to be put into investigating possibilities and challenges upon creating a KG constituting of these types of data.

Chemotion Electronic Lab Notebook, partners: Ebrahim Norouzi (FIZ KA), Jörg Waitelonis (FIZ KA), Nicole Jung (KIT IBCS).

Chemotion[4] is a research data management environment specifically designed for chemistry. It is tailored to the specific needs of chemists and comprises an electronic lab notebook [7] as well as a research data repository [8]. The repository complies with FAIR principles (Findable, Accessible, Interoperable, Reusable). The Chemotion repository was first developed without a semantic layer. Currently, its enhancement is under progress and the plans include creating a Chemotion KG. The working group has already presented the first steps of a Chemotion-specific ontology engineering process and Chemotion KG design (Norouzi et al. 2025 [9]).

Currently, the Chemotion KG is further developed in collaboration with NFDI4Chem Knowledge Base team, who are also involved in engineering chemistry-specific ontologies.

AI in Biomedical Engineering, partners: Genet Asefa Gesese (FIZ KA), Silvia Becker (KIT IBT), Axel Loewe (KIT IBT).

Researchers at KIT IBT develop computational models of a human heart. One of their goals is to automatise the amplified P-wave detection process. P-wave is a part of electrocardiogram (ECG) which is correlated with heart failures i.e., atrial fibrillation [10]. The partners aim at creating a machine learning (ML) model to automatically detect a P-wave and properly measure its duration also in cases of an amplified P-wave. Until now, the team has discussed applications of different ML models, including long short-term menory (LSTM) as well as utilising multimodal large language models (LLMs) as support to this task. In the current phase the team is dealing with typical issues concerning medical data, e.g. where the trained data can be stored.

In this use case also a master student will start working on an interdisciplinary thesis topic *Machine Learning-Based ICU Stay Prediction*[5]. The focus of the thesis is implementing AI methods for predicting the length of stay of patients at an intensive care unit (ICU), based on an ECG.

Publication of Large Datasets, partners: Sven Hertling (FIZ KA), Genet Asefa Gesese (FIZ KA), Gunjan Singh (FIZ KA), Etienne Posthumus (FIZ KA), Tobias Kerzenmacher (KIT IMK), Sabine Barthlott (KIT IMK), Sibylle Hassler (KIT IWU), Peer Nowack (KIT ITI).

In climate research, scientists are dealing with very large datasets and a lack of proper publication culture. Reusability and interoperability are often significantly limited, disabling a proper use and sharing. Moreover, there exist many issues with metadata, e.g. the same data is described with different metadata. This very quickly leads to chaos in climate data repositories. As described by Jacyszyn et al. 2025 [11], the team suggests developing an LLM-based tool to standardise metadata in climate reserach repositories.

Currently, the collaboration on this task is being extended to include also NFDI4Earth representants as well as RADAR[6], which is a research data management infrastructure provided by FIZ KA. An interdisciplinary topic of automatic metadata is being investigated in between the three research groups.

Summary

DiTraRe connects researchers from diverse disciplines to investigate the effects of digital transformation. A special focus is set on AI as an additional aspect to the process. We investigate multiple AI methods and apply them in DiTraRe use cases. We also study the overall perception, practices, and effects of applied AI in the interdisciplinary environment.

[4]Chemotion `https://chemotion.net/`

[5]Master student thesis topic `https://ise.aifb.kit.edu/english/119_284.php`

[6]RADAR website `https://www.fiz-karlsruhe.de/en/produkte-und-dienstleistungen/radar`

Acknowledgments

We would like to kindly thank all our DiTraRe collaborators for making this Leibniz Science Campus idea come true. It would not be possible without your efforts and time invested.

The Leibniz Science Campus "Digital Transformation of Research" (DiTraRe) is funded by the Leibniz Association (W74/2022).

References

[1] M. Razum, F. Bach, S. Brünger-Weilandt *et al.*, "Proposal for a Leibniz ScienceCampus – Digital Transformation of Research (DiTraRe)," 2023. doi: 10.5281/zenodo.11109406 ., project proposal.

[2] A. M. Jacyszyn, F. Bach, H. Sack *et al.*, "Interim Report of the Leibniz Science Campus 'Digital Transformation of Research (DiTraRe)'," *Zenodo*, 2025. doi: 10.5281/zenodo.14941634 .

[3] A. Jacyszyn, H. Sack, D.-S. Group, F. Bach, and M. Razum, "DiTraRe: AI on a Spider's Web. Interweaving Disciplines for Digitalisation," in *4th International Workshop on Scientific Knowledge: Representation, Discovery, and Assessment (Sci-K)*, vol. 3780. CEUR, 2024. [Online]. Available: https://ceur-ws.org/Vol-3780/paper5.pdf

[4] C. Niessner *et al.*, "Beyond Boundaries: Data Sharing in Motor Development Research—Best Practices, Challenges, and Opportunities," *Journal of Motor Learning and Development*, 2025.

[5] K. Klemm *et al.*, "Development and Introduction of a Disciplinary Data Repository for Sport Scientists Based on the Example MO|RE Data," *Bausteine Forschungsdatenmanagement*, vol. 1, Mar. 2024. doi: 10.17192/BFDM.2024.1.8615 .

[6] S. R. Ondraszek, J. Waitelonis, K. Keller *et al.*, "Ontologies in Motion: A BFO-Based Approach to Knowledge Graph Construction for Motor Performance Research Data in Sports Science," in *5th International Workshop on Scientific Knowledge: Representation, Discovery, and Assessment (Sci-K)*, vol. 4065. CEUR, 2024. [Online]. Available: https://ceur-ws.org/Vol-4065/paper08.pdf

[7] S. Kotov, P. Tremouilhac, N. Jung, and et al., "Chemotion-eln part 2: adaption of an embedded ketcher editor to advanced research applications." *Journal of Cheminformatics*, vol. 10, no. 38, 2018. doi: 10.1186/s13321-018-0292-9 . [Online]. Available: https://doi.org/10.1186/s13321-018-0292-9

[8] P. Tremouilhac, C.-L. Lin, and P.-C. a. a. Huang, "The repository chemotion: Infrastructure for sustainable research in chemistry," *Angewandte Chemie International Edition*, vol. 59, no. 50, pp. 22 771–22 778, 2020. doi: https://doi.org/10.1002/anie.202007702 . [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/anie.202007702

[9] E. Norouzi, N. Jung, A. M. Jacyszyn *et al.*, "AI4DiTraRe: Building the BFO-Compliant Chemotion Knowledge Graph," in *5th International Workshop on Scientific Knowledge: Representation, Discovery, and Assessment (Sci-K)*, vol. 4065. CEUR, 2024. [Online]. Available: https://ceur-ws.org/Vol-4065/paper04.pdf

[10] M. U. Rasmussen, P. Kumarathurai, A. Fabricius-Bjerre *et al.*, "P-wave indices as predictors of atrial fibrillation," *Annals of noninvasive electrocardiology : the official journal of the International Society for Holter and Noninvasive Electrocardiology, Inc*, 2021. doi: 10.1111/anec.12751 . [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC7507358/

[11] A. M. Jacyszyn, S. Jiang, G. A. Gesese *et al.*, "AI4DiTraRe: Towards LLM-Based Information Extraction for Standardising Climate Research Repositories," in *1st Bridge on Artificial Intelligence for Scholarly Communication (AI4SC)*, 2025. [Online]. Available: https://zenodo.org/records/14872359