Interim report of the Leibniz Science Campus "Digital Transformation of Research" (DiTraRe)

Project number: W74/2022

Anna Jacyszyn[1], Felix Bach[1], Harald Sack[1,2], Christiane Noe[1], Marius Albiez[3], Sabine Barthlott[4], Silvia Becker[5,6], Franziska Boehm[1], Christian Bonatto Minella[1], Angelina Sophie Dähms[3], Tanja Eberhardt[8], Nicole Jung[7], Katja Keller[8], Tobias Kerzenmacher[4], Axel Loewe[5], Claudia Niessner[8], Linda Nierling[3], Constanze Scherz[3], Ralf Schneider[3], Leonie Seng[3], Ely Nova Natalia Silaban[1], Lea Sophie Singson[1], Melanie Volkamer[2], Hannah Zimmermann[8]

[1]*FIZ Karlsruhe – Leibniz Institute for Information Infrastructure, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany*
[2]*Karlsruhe Institute of Technology, Institute of Applied Informatics and Formal Description Methods, Kaiserstr. 89, 76133 Karlsruhe, Germany*
[3]*Karlsruhe Institute of Technology, Institute for Technology Assessment and Systems Analysis, Karlstr. 11, 76133 Karlsruhe, Germany*
[4]*Karlsruhe Institute of Technology, Institute of Meteorology and Climate Research - Atmospheric Trace Gases and Remote Sensing, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany*
[5]*Karlsruhe Institute of Technology, Institute of Biomedical Engineering, Fritz-Haber-Weg 1, 76131 Karlsruhe, Germany*
[6]*University Heart Center Freiburg - Bad Krozingen, Südring 15, 79189 Bad Krozingen, Germany*
[7]*Karlsruhe Institute of Technology, Institute of Biological and Chemical Systems, Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen, Germany*
[8]*Karlsruhe Institute of Technology, Institute of Sports and Sports Science, Engler-Bunte-Ring 15, 76131 Karlsruhe, Germany*

Executive summary

The Leibniz Science Campus Digital Transformation of Research (DiTraRe) investigates the process of digitalisation of research on multiple levels in an interdisciplinary environment. It is a collaboration between FIZ Karlsruhe – Leibniz Institute for Information Infrastructure and Karlsruhe Institute of Technology (KIT). The structure of DiTraRe is based on four Use Cases, which are starting points for more general and broadly understood Research Clusters. The Use Cases and resulting Research Clusters are examined through four Dimensions, each focusing on specific aspects.

DiTraRe has primarily focused on identifying the specific challenges of the Use Cases related to the digital transformation of research and laying the groundwork for future advancements. Collaboration between Use Cases and Dimensions has been established, initial research articles have been published, and progress has been made in various ways, both in implementing potential solutions as well as planning interdisciplinary events. The insights gained so far provide a foundation for further exploration and refinement, setting the stage for substantial progress in the coming phases of the project.

1. Achievement of objectives and milestones

DiTraRe started five months after approval, in September 2023, and has been running for 16.5 months (approximately 35% of the approved duration) at the time of this report. Initially, we established the interdisciplinary DiTraRe consortium through constitutional and kick-off meetings attended by the representatives of the Research Clusters, Dimensions, and Use Cases. We then transitioned to regular meetings for the overall consortium on a quarterly basis. Our collaborative efforts have already led to the realisation of some of our milestones and deliverables, with the others well within reach in the remaining time. Furthermore, the Scientific Advisory Board for DiTraRe has been established, members have been introduced to the project and first consultations and advice have taken place.

Milestone 1 (originally planned for end of 2023): Requirements of the Use Cases were elaborated after analysis from the perspective of all four dimensions

The Use Case partners' research questions were discussed in dedicated meetings with Dimensions. Problems and challenges in the context of digitalisation were explored in detail. Further targeted meetings between Use Case partners and Dimensions led to the elaboration of Use Cases requirements, resulting in concrete work plans.

First version of the compendium published (connected to the Deliverable: "Compendium of research results on DiTraRe's website as a "living document" published and continuously updated")

We designed and published the [DiTraRe website](#) showcasing the matrix structure of our approach to addressing current challenges related to the digitalisation of research from a multidisciplinary perspective. The envisioned compendium is planned to be realised as a DiTraRe Knowledge Graph, for which the underlying ontology is under development.

First conference held

The first conference will take place on the 2nd and 3rd of December 2025 in the form of a DiTraRe Symposium on Digitalisation of Research. This event will explore the fast-evolving landscape of digital research (announcement available on the [DiTraRe website](#)).

Financial status

DiTraRe was originally planned to start in the spring of 2023 but was postponed to start in September 2023. The delay of a few months is reflected in the financial status. Overall, we have spent less than planned, but in considering the budget of each month, we have met the target. DiTraRe will continue to utilise its budget as planned in the upcoming months.

2. Activities and obstacles

We outlined and determined the legal and ethical research needs and questions of each area represented by our four Use Cases. Collaborations between the Use Cases and the four Dimensions have been established. Within these collaborations, a two dimensional approach is implemented: we aim to advance the science surrounding digitalisation of research through papers and conference contributions, as well as provide the community with hands-on materials to aid own research through Open Educational Resources such as Fact Sheets.

The Use Case partners hold regular team consultation meetings and participate in All Hands meetings of the entire consortium. They also engage in exchanges with representatives of each DiTraRe Dimension. For example, meetings of the Use Case "Sensitive Data in Sports Science" aim to create a handout specifically for sports science on what an informed consent for conducting studies must contain, as well as developing best practices for sharing sensitive and non-sensitive data in repositories. In this specific Use Case, a plan to engineer a specific ontology for the MO|RE data repository was created, reusing existing ontologies in the sports domain to enable better interoperability (Dimension "Exploration and Knowledge Organisation"). Further, MO|RE data could be improved by adapting the RADAR backend and its API, on which MO|RE data relies for data archiving and publication, according to application needs identified by the Dimension "Tools and Processes".

Within the Research Cluster "Publication Cultures" the role of generative AI in the digital transformation of the science system has been studied through a detailed literature review followed by empirical analyses. The Dimension "Reflection and Resonance" created an overview of guidelines on generative AI of scientific publishers and established a science–society interface. A series of interviews with Use Case partners was conducted by the Dimension to identify similarities and differences across all Use Cases regarding types of collected data, challenges and organisational frameworks of open data. A summary of results and identification of joint risks and challenges was prepared.

Within the Use Case "Chemotion Electronic Lab Notebook", we worked on technical solutions to further improve the efficiency and reliability of data transfer and provision processes in chemistry lab environments. We worked on the implementation of AI methods to improve and accelerate data driven curation processes for Research Data Management systems such as Electronic Lab Notebooks and repositories. We also started implementing support for Resource Description Framework files for the research data repository Chemotion.

Within the Use Case "Publication of Large Datasets", we worked on solutions to make very large datasets easily publishable, accessible and reusable. This has been done within the team working on further developing the RADAR repository for research data (Dimension "Tools and Processes") by adding new options for data ingestion and access, such as using WebDav and GitHub import. Moreover, an idea on solving the problem of standardising metadata in climate research repositories has been developed by the Dimension "Exploration and Knowledge Organisation". The team wants to utilise large language models to extract information from descriptions accompanying datasets.

To overcome the challenges posed by complex regulatory frameworks, the Dimension "Legal and Ethical Challenges" has developed a legal assessment regarding the processing of personal data for generating synthetic data. Through joint investigation with the Use Case "Artificial Intelligence in Biomedical Engineering", we analysed whether synthetic data can minimise the risk of re-identification and which legal frameworks must be considered when processing the sensitive data with a predictive AI model. We also reviewed the licensing agreement for the research data repositories of the Use Case "Sensitive Data in Sports Research" to facilitate the use of anonymised data and ensure compliance with GDPR requirements for processing sensitive data.

Obstacles encountered during the current and previous phases of the Science Campus were mostly related to changes in personnel and delays in filling open positions. These were resulting from difficulties finding candidates with specific expertise in the required fields.

Other obstacles included accessing clinical data, which is crucial for advancing medical research and improving patient care. However, accessing clinical data remains a challenge for the researchers as does ensuring compliance with regulations designed to protect sensitive data, which requires navigating complex regulatory frameworks.

3. Results and successes

DiTraRe has published 10 scientific papers across various disciplines (see bibliography in the appendix to this document), accounting for more than half of its planned deliverables. Further joint publications are in progress. Other outputs include two congress contributions and seven sports science factsheets.

The DiTraRe concept and findings have been presented at various events, including workshops such as "AI for the Research Ecosystem" (March 2024, London), "Scientific Knowledge: Representation, Discovery, and Assessment" (November 2024, Baltimore), "First AAAI Bridge on Artificial Intelligence for Scholarly Communication" (February 2025, Philadelphia). Conferences include "TA24 - Methoden für die Technikfolgenabschätzung" (June 2024, Vienna), "Forum on Philosophy, Engineering and Technology" (September 2024, Karlsruhe), and "Computing in Cardiology" (September 2024, Karlsruhe). Contributions were also made to the 6th International Motor Development Research Consortium (June 2024, Ghent), Atrial Signals (June 2024, Maastricht), "DGK Herztage" (November 2024, Hamburg), and the Network for Technology Assessment (November 2024, Berlin).

DiTraRe submitted four abstracts to the DFG call on AI in research, proposing AI-driven legal management of digital twins in biomedical engineering, a voice-based assistant for hands-free work in laboratories, knowledge graphs for improving health and fitness outcomes in youth, and AI-supported dataset management in climate research. The ideas received positive feedback.

4. Equal opportunities, career development and internationalisation

To ensure equal opportunities with regard to gender and internationalisation, we have implemented targeted initiatives in staff development and recruitment. Our recruitment processes emphasise diversity by actively encouraging applications from underrepresented groups and ensuring unbiased evaluation criteria. Gender balance is promoted through mentorship programmes, flexible working arrangements, and support for researchers balancing family and career responsibilities (e.g. working from home).

The interdisciplinary nature of DiTraRe enables all researchers involved in our cooperation to gain experience that they would not have while working solely in their original research area. The interdisciplinarity of our studies opens up new career possibilities by enabling researchers to exchange expertise, gain new competencies and work on innovative topics.

5. Structures and collaboration

DiTraRe has undergone a change in the role of its spokesperson twice. Each of the previous spokespersons had to step down due to the need to take on other responsibilities. Together with Prof. Dr. Harald Sack, who took over the spokesperson role, Dr. Anna Jacyszyn was added as a coordinator to support the DiTraRe management team. Additionally, the role of

the PI of the Dimension "Tools and Processes" as well as the Research Cluster "Smart Data Acquisition" was handed over to another project member.

A Task Force "Security" was started between the Dimension "Legal and Ethical Challenges" and the Research Cluster "Protected Data Spaces" to improve efficiency of aligning research goals. The Use Case "AI in Biomedical Engineering" collaborates with the University Heart Center Freiburg - Bad Krozingen. DiTraRe members contribute to the work of 10 different NFDI consortia. Additionally, through the DiTraRe Scientific Advisory Board, we engage with experts from institutes and universities across Germany.

6. Quality assurance

In the age of AI, we navigate data protection and privacy regulations, open science guidelines and best practices, and AI law in our research clusters to safeguard privacy, uphold research quality, and foster technology and innovation. We ensure quality assurance through adherence to established guidelines for good scientific practice, including transparency, reproducibility, and ethical research conduct. To make research results widely accessible, we prioritise open access publishing. Research outputs, including data, publications, and software, are and will be shared through open repositories, in line with FAIR (Findable, Accessible, Interoperable, Reusable) principles.

7. Outlook

There is ongoing work in the context of all Use Cases, which will lead to achieving planned results within the duration of DiTraRe. After having performed a requirement analysis, the Dimensions are currently concentrating on developing specific solutions for each Use Case. Results are regularly being published and more publications are planned. These will be promoted and advertised during various events such as conferences, workshops, symposia, and colloquia. We are also actively exploring possibilities for cooperation and synergies between different working groups (Dimensions as well as Use Cases).

At the same time, we are developing generalisations of Use Cases in Research Clusters, primarily by extending and reformulating research questions beyond individual Use Cases. Our aim is for the main research questions to reach an interdisciplinary level. One possibility is to include other Use Cases that could also be addressed by the main research questions of the Research Clusters. To reach this goal, we are focusing on extended networking during multiple events that the DiTraRe team will attend and/or organise. We hope to find more collaboration partners outside of our consortium, concentrating on a detailed exchange with various NFDI consortia.

Apart from the work described above, we have many more activities already planned, which concern our other milestones and deliverables. A knowledge graph for DiTraRe will be prepared by the Dimension "Exploration and Knowledge Organisation". Its interactive visualisation will be published on the DiTraRe website. Following positive feedback from the DFG (see Section 3), we are planning to develop each of the four abstracts into funding proposals (e.g. DFG LIS program). A science-society dialogue in different formats has already taken place and is being further developed by the Dimension "Reflection and Resonance". An interdisciplinary colloquium on "Digital Transformation of Research" is planned to start in May 2025 and will take place bi-monthly in a hybrid mode thereafter. We plan to record presentations and publish them on the DiTraRe website. A jointly supervised

project course for students was discussed during the last All-Hands meeting of the consortium and is envisaged to start in the next academic year.

In March 2025, the DiTraRe working group is organising the workshop "Information Infrastructures in the Era of Artificial Intelligence: Opportunities and Challenges" as part of the E-Science-Tage in Heidelberg, as well as presenting three talks during the event on different aspects of the LSC. In December 2025, our first conference is planned to take place in Karlsruhe. It will be a symposium to which we will invite experts from different fields of research to discuss the current state-of-the-art challenges and potentials within the main topics of each Dimension. Through events which we are planning to organise, we aim to establish a regular conference series, the "Karlsruhe Conference on Digital Transformation", which could take place bi-annually.

Appendix

Bibliography

- Becker, S., Espinosa, C. B., Unger, L., Haas, A., Luik, A., & Loewe, A. "Influence of Conduction Velocity Restitution Steepness on Atrial Fibrillation Vulnerability and Maintenance", Computing in Cardiology Conference, 2024, https://doi.org/10.22489/cinc.2024.105
- C.B. Espinosa et al., „A Cyclical Fast Iterative Method for Simulating Reentries in Cardiac Electrophysiology Using an Eikonal-Based Model", Engineering with Computers, January 2025, https://dx.doi.org/10.1007/s00366-024-02094-9, https://doi.org/10.35097/70234yr69tr09krh
- A.M. Jacyszyn, S. Jiang, G.A. Gesese, S. Hertling, T. Kerzenmacher, P. Nowack, S. Barthlott, H. Sack. "AI4DiTraRe: Towards LLM-Based Information Extraction for Standardising Climate Research Repositories", First AAAI Bridge on Artificial Intelligence for Scholarly Communication, February 2025, https://doi.org/10.5281/zenodo.14872358
- A.M. Jacyszyn, H. Sack, DiTraRe-Study Group, M. Razum, and F. Bach. „DiTraRe: AI on a Spider's Web. Interweaving Disciplines for Digitalisation", 4th International Workshop on Scientific Knowledge: Representation, Discovery, and Assessment, CEUR Workshop Proceedings, Vol. 3780, Paper 5, September 2024, https://ceur-ws.org/Vol-3780/paper5.pdf
- K. Keller, A. Pleger, C. Niessner, E. Ueding and M. Krüger. "The Status Quo of Research Data Management in the German-speaking Sports Sciences –Results of an Online Survey", 2025, https://doi.org/10.5281/zenodo.11204024
- K. Klemm, K. Bös, H. Kron, T. Eberhardt, A. Woll, and C. Niessner. „Bausteine Forschungsdatenmanagement : 2024, Development and Introduction of a Disciplinary Data Repository for Sport Scientists Based on the Example MO|RE Data", March 2024, http://doi.org/10.17192/bfdm.2024.1.8615
- A. Loewe, P.J. Hunter, and P. Kohl. "Computational modelling of biological systems now and then: revisiting tools and visions from the beginning of the century", January 2025, arXiv:2501.13142, https://doi.org/10.48550/arXiv.2501.13142
- P. Martínez Díaz, J. Sánchez, N. Fitzen, U. Ravens, O. Dössel, and A. Loewe. „The right atrium affects in silico arrhythmia vulnerability in both atria", Heart Rhythm, Bd.

21, Nr. 6, S. 799–805, June 2024, https://doi.org/10.1016/j.hrthm.2024.01.047, https://dx.doi.org/10.35097/1830

- P. Martínez Díaz et al., „Impact of effective refractory period personalization on arrhythmia vulnerability in patient-specific atrial computer models", Europace, Bd. 26, Nr. 10, S. euae215, October 2024, https://doi.org/10.1093/europace/euae215, https://doi.org/10.5281/zenodo.10726677
- L. Nierling and L. Seng. Digitale Transformationsprozesse in der Forschung - Perspektiven für Methoden in der Technikfolgenabschätzung. KSP, 2025 i.E.