

ErUM-Data-Hub - The Networking and Transfer Office serving Digital Transformation in Research on Universe and Matter in Germany

Martin Erdmann¹, Peter Fackeldey^{1,}, Benjamin Fischer¹, Stefan Fröse², Andreas Nowack¹, Judith Steinfeld¹, and Angela Warkentin¹*

¹RWTH Aachen University

²TU Dortmund University

Abstract. Research on Universe and Matter (ErUM), conducted at major infrastructures such as CERN and large observatories in collaboration with university groups, plays an important role in driving the digital transformation for the future. The German action plan "ErUM-Data" [2] promotes this transformation through the interdisciplinary networking and financial support of about 20.000 scientists. The ErUM-Data-Hub [4] serves as a central networking and transfer office to meet these ambitions. One central task is designing, organizing and performing schools and workshops for young and expert scientists in the areas of big data, deep learning, sustainable computing and many more. We present the achievements up to the first anniversary of the ErUM-Data-Hub in the German ErUM community.

1 Introduction

There is an increasing impact of the digital transformation affecting us all. It influences our social and economic future and determines the way we live, work, and especially conduct research. Basic research in physics has played a pioneering role in the progress of the digital transformation process.

A large area of fundamental research in physics investigates research questions about matter and the universe. Answering these questions requires measurement data that are typically gathered at large research infrastructures, e.g., at CERN, and subsequently analyzed. In the process, the amount of data is rapidly increasing due to ever new experiments and observational instruments as well as their simulations. Generally referred to as "Big Data" and their analysis as "Big Data Analytics", these large data sets imply significant technological challenges in many respects.

The management of these datasets and exploiting the most information out of them constitutes a key challenge in fundamental research. There is an ongoing development of new solutions for storing, handling, and analyzing big data. Synergies with other research areas and industry are explored and also taken into account.

In order to promote the digital transformation in the research fields Universe and Matter, the German Federal Ministry of Education and Research (BMBF) [1] has published the

*e-mail: peter.fackeldey@erumdatahub.de

action-plan “ErUM-Data” where the acronym ErUM refers to the Research on Universe and Matter.

2 The action plan ErUM-Data

In Germany, fundamental research on Universe and Matter unifies about 20.000 scientists from different communities: astronomy, astro-particle physics, elementary particle physics, accelerator physics, research with neutrons, research with synchrotron radiation, research with nuclear probes and ion beams as well as hadron and nuclear physics - organized in a total of eight committees. Although these ErUM scientists focus on distinct scientific questions, they are facing similar challenges concerning handling and analyzing Big Data.

For that reason, the German Federal Ministry of Education and Research (BMBF) established the action plan ErUM-Data [2] in 2020 as part of the framework program "Exploring the Universe and Matter - ErUM", published in 2017 [3].

ErUM-Data defines measures to facilitate scientific research of tomorrow by promoting a coordinated approach of all ErUM communities. The long-term goal is to pave the way for sustainable and efficient solutions regarding the digitalization of basic research in the natural sciences. Easy access to data and tools is a central pillar. Finally, the gained knowledge will play a key role regarding the digitization of basic research in the natural sciences.

In a nutshell, the key goals of the ErUM-Data plan are joint research embedded in networking, strengthening digital competencies, and communication and knowledge transfer.

3 The ErUM-Data-Hub

In order to achieve and facilitate the strategic goals of the ErUM-Data action plan effectively, the “ErUM-Data-Hub” [4], a central networking and transfer office has been founded in 2022. The ErUM-Data-Hub is funded by the German Federal Ministry of Education and Research (BMBF) and involves stakeholders from all ErUM communities. It takes up ideas from the scientific community and supports and coordinates their realization. Furthermore, it aims at providing support in the transfer of existing solutions to different research fields, which should also attract partners beyond ErUM, e.g., in the areas of energy, climate, sustainability, and health. The activities of the ErUM-Data-Hub are divided into different work areas: *digital competencies*, *transfer & communication*, and *networking & exchange*.

Based in Aachen at RWTH Aachen University, the ErUM-Data-Hub team currently consists of six interdisciplinary team members (project lead, team lead, scientific advisors, technical support, and science communication). An interdisciplinary team improves the overall quality of the organized events by adding expertise from physics, Big Data Analytics, event organization, networking, and outreach. The team works closely together before and during events (on-site) including weekly meeting reports. A close connection between all ErUM communities and the ErUM-Data-Hub team is established through bi-weekly meetings. There, the team, in particular the project lead, the team lead, and the scientific advisors, closely listens to the needs of the communities, assesses them, and distills contents for workshops and schools. Afterwards, the whole ErUM-Data-Hub team starts the planning and organization of such events and informs all ErUM communities of these events through a variety of channels. This communication and communication with the broad public (outreach) is primarily taken care of by the science communication responsible. The technical support builds and maintains knowledge databases for collecting event material, provides on-site IT support during events, and establishes communication platforms.

Furthermore, the ErUM-Data-Hub benefits from scientific advice by the community's self-organization "DIG-UM" (Digital Transformation in the Research of Universe and Matter) which was founded by all 8 ErUM-communities. This long-term and sustainable umbrella organization strengthens ErUM scientists in mastering the immense technical and scientific tasks they are faced with related to the digital transformation.

4 Digital Competencies

A central task of the ErUM-Data-Hub is fostering the digital competencies of scientists. The realization of this key goal stated in the ErUM-Data action plan is carried out by the training of scientists with schools and workshops in the areas of Big Data, deep learning, sustainable computing, and many more which are hosted by the ErUM-Data-Hub in collaboration with DIG-UM. The events serve the professional education of scientists belonging to all ErUM communities and are held at different locations in Germany. The 2023 annual program, for instance, included three deep learning schools, two deep learning "Train-the-Trainer" workshops for providing lecturers with didactic material, as well as expert workshops on "Big Data Analytics", "Sustainability in ErUM-Data", "Research Data Management", and "Next Generation Environment for Interoperable Data Analysis".

In 2023, the ErUM-Data-Hub welcomed a total of 314 ErUM scientists to participate in workshops and schools. The events were attended by students (10%), PhD students (31%), postdocs (19%), senior scientists (7%), group leaders (8%) and professors (9%). 16% of participants did not belong to any of these groups, but represented industry or attended the event in an official function of the funding body or the BMBF. The distribution of participants varied depending on the event concept. The Deep Learning Schools "Basic Concepts" and "Advanced Concepts" were mainly attended by PhD students and postdocs and the Deep Learning Train-the-Trainer Workshops and Expert Workshops by ErUM-Scientists with a level of postdoc or higher. Participants from all 8 ErUM communities and all over Germany were reached, with the distribution roughly reflecting the size of the ErUM communities: 27% from elementary physics, 14% from astroparticle physics, 10% from research with neutrons, 8% from hadron and nuclear physics, 7% from research with synchrotron radiation, 6% from the Council of German Observatories, 4% from accelerator physics and 1% from research with nuclear probes and ion beams. 23% of the participants indicated an affiliation to another research community or industry.

4.1 Deep Learning Schools

Our schools on basic concepts in deep learning are aimed at scientists who have little to no prior knowledge about deep learning. For more experienced scientists, the ErUM-Data-Hub organizes schools with advanced concepts that are currently relevant in the research of deep learning and physics. Both schools are held over four to five days with typically approximately 30-50 participants each. In basic schools, most participants are undergraduate and PhD students (78%), followed by postdocs (17%) and other (5%). The courses include three to four intense lectures hosted by expert speakers together with tutorials (90 minutes each). In addition, each school includes a group challenge, that are solved and presented by small subgroups of participants. The overall goal of these schools is to *actively* experience deep learning in addition to learning about it, enabling participants to professionally apply these methods in their research, and to expand their collaborative networks.

4.2 Train-the-Trainer Workshops

Our “Train-the-Trainer” workshops pursue the goal to educate lecturers for future teaching activities and to provide material for lectures and tutorials on deep learning. They are aimed at scientists from all ErUM communities with prior knowledge of deep learning with an academic level of postdoc or higher and are offered for basic and advanced deep learning concepts. The two-day workshop takes place at various universities with approximately 30 participants. It includes four lectures and tutorials. The number of professors, or group leaders, (43%) and postdoctoral researchers (39%) participating is relatively balanced, complemented by participants with other academic levels (18%). The overall goal of these workshops is to enable teaching deep learning with clear didactic focus and to provide material for lectures and exercises.

5 Transfer & Communication

The implementation of these events pursue another key goal of the action plan, namely transferring knowledge between the eight ErUM communities. The ErUM-Data-Hub isolates questions that are important to these communities and organizes workshops regarding them. Exemplarily, these concern topics such as big data analytics, research data management, or user interfaces. Thereby, participants share their solutions and experiences from their own communities and have (guided) discussions. These events repeatedly attracted participants who face similar questions in industry. Novel technologies developed for science in the field of big data have a great innovation potential for industry. Thus, another goal of the ErUM-Data-Hub is to set the course for on one hand a transfer of (ErUM) research into industry, and on the other to profit from industry developments. Here, key ideas related to, e.g., sustainable computing and conceptual advances in deep learning, are discussed in dedicated workshops including experts from industry. Contacts with industry have mainly been established with start-ups and smaller companies. The cooperation currently consists of contributions from industry at workshops in which specific machine learning applications are showcased or the career paths of ErUM scientists in industry are discussed. The industry contacts benefit from participating in events where knowledge about new technological findings is shared.

Apart from strengthening scientific transfer, an overall goal of the action plan is to increase the awareness of ErUM research for scientists outside our community and interested citizens. This goes in hand with another key task of the ErUM-Data-Hub, namely strengthening science communication. Topics of the digital transformation need to be communicated professionally and even more broadly: from science into society since the success of the further digitization process in Germany also depends crucially on society’s willingness to support technological change. It is important to communicate research results in a comprehensible manner tailored to the target audience, in order to manifest trust in new technologies. For example, the ErUM-Data-Hub is working on both proven communication formats, such as expanding its presence on social media channels, and on developing new innovative formats that reflect the nature of digitization. An example for such a new format is a podcast about research in ErUM currently under development.

In April 2023, the ErUM-Data-Hub represented ErUM research at the “Hannover Messe”, Germany’s largest industrial trade fair (130,000 visitors, 4,000 exhibitors from over 70 countries) at a joint exhibition stand together with the four ErUM research centers at the Large Hadron Collider (ALICE, ATLAS, CMS and LHCb) [8] and Belle II Germany [6]. Lively discussions with the broad public contributed to extend the outreach of ErUM research by raising awareness of the importance of exploring the Universe and Matter.

6 Networking & Exchange

Following the BMBF action plan, a central task of the ErUM-Data-Hub is to pick up ideas from the scientific community, create opportunities for their dissemination, and to support their realization. The promotion of the digital transformation in the research fields Universe and Matter can only be mastered in collaboration. Here, the above-mentioned close contact to the common organization DIG-UM of the eight ErUM communities plays a central role. The ErUM-Data-Hub supports synergy identification and networking between the various ErUM communities, and more generally funded ErUM-Data projects in the field of digitization. Overall, the ErUM-Data-Hub takes care of the close coordination of all relevant stakeholders that is necessary to find joint solutions for digitization, e.g., by supporting regular meetings for scientific exchange and for agreeing on common projects and standards. One of a number of examples is a workshop on sustainability in digital transformation initiated and organized by the ErUM-Data-Hub [5].

Partnerships are indispensable while expanding networks, so the ErUM-Data-Hub is collaborating with partners such as Belle II [6], “Netzwerk Teilchenwelt” [7], “LHC-ErUM-FSP” [8], “NFDI” [9], “Terascale” [10], and others and organizes regular outreach meetings. Joint workshops on research topics are also held, with, for example, the NFDI consortia and the ErUM-Data-Hub jointly organizing regular meetings on topics related to research data management.

In addition, all previously mentioned workshops and schools are designed to create a working atmosphere that encourages networking between participants. Explicitly the venue location is chosen carefully to empower this atmosphere. These events are often complemented with non-scientific activities, such as a pub quiz, sport activities, or event dinners.

7 Conclusion

The ErUM-Data-Hub helps facilitating the ErUM-Data action plan by promoting digital transformation in fundamental research on Universe and Matter. In this paper, we presented the achievements up to the first anniversary of the ErUM-Data-Hub in the German ErUM community. Functioning as a networking and transfer office, the ErUM-Data-Hub brings together scientists working in various research areas who are jointly facing challenges and opportunities of topics all around digitization with special focus on big data. The gained knowledge will be brought into research collaborations and will also play a key role in future day-to-day applications, such as AI tools or medicine. With the training of scientists, we foster digital transformation by providing young scientists with the most important competencies. Overall, we already reached about 300 participants attending our workshops. The presented achievements significantly support the realization of the strategic objectives of the ErUM-Data action plan.

8 Acknowledgements

We gratefully acknowledge funding support by the Federal Ministry of Education and Research (BMBF) in Germany through the ErUM-Data project 05D21PA1, and by the Ministry of Innovation, Science, and Research of the State of North Rhine-Westphalia.

References

- [1] German Federal Ministry of Education and Research (BMBF), https://www.bmbf.de/bmbf/de/home/home_node.html, accessed on 12.10.2023.

- [2] ErUM-Data action plan, https://www.bmbf.de/SharedDocs/Publikationen/de/bmbf/7/31640_Aktionsplan_ErUM-Data.pdf, accessed on 20.09.2023.
- [3] Research on Matter and Universe - ErUM, https://www.bmbf.de/bmbf/de/forschung/naturwissenschaften/erforschung-von-universum-und-materie/erforschung-von-universum-und-materie_node.html, accessed on 12.10.2023.
- [4] ErUM-Data-Hub, <https://erumdatahub.de>, accessed on 20.09.2023.
- [5] “Sustainability in the Digital Transformation of Basic Research on Universe & Matter” Workshop, <https://erumdatahub.de/blog/2023/06/12/sustainability-workshop/>, accessed on 11.10.2023.
- [6] Belle II Office, <https://www.belle2.de>, accessed on 20.09.2023.
- [7] Netzwerk Teilchenwelt, <https://www.teilchenwelt.de>, accessed on 20.09.2023.
- [8] ErUM FSP Office, <https://lhc-deutschland.de>, accessed on 20.09.2023.
- [9] Nationale Forschungsdaten Infrastruktur (NFDI), <https://www.nfdi.de>, accessed on 20.09.2023.
- [10] Helmholtz Alliance “Physics at the Terascale”, <https://www.terascale.de>, accessed on 20.09.2023.