HELMHOLTZ

Research for grand challenges.

Data Management & Analysis

Michael Bussmann, Volker Gülzow
Topic Speakers

Research Field Matter

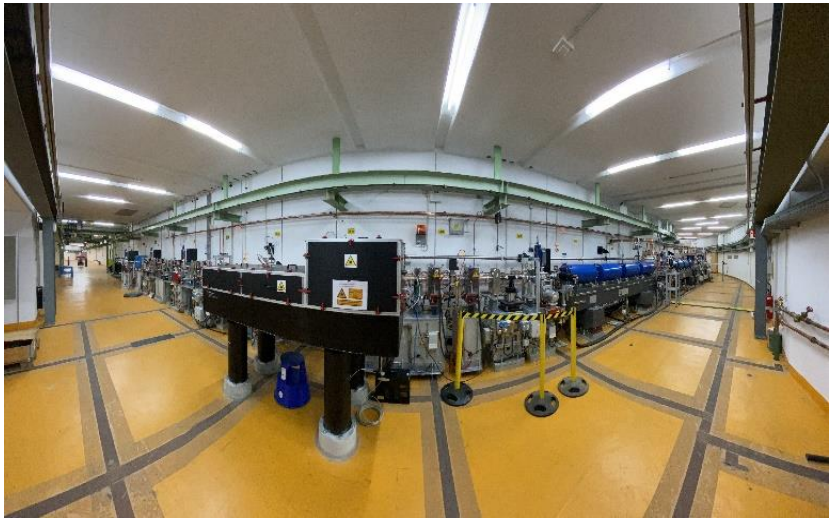
Looking forward

„Targeted challenge-driven access“ enabled by **„intelligent, autonomous, integrated facilities“**

- **Focus on data-driven science** across user communities, taking up challenges, enabling discoveries
- **Frontier science through use of frontier technologies** (AI, Exascale, Quantum, Edge, Cloud)
- **Intelligent, target-driven facility operation** enabled by AI-integration of Machines, Experiments & Simulations

Research Field Matter

What we do

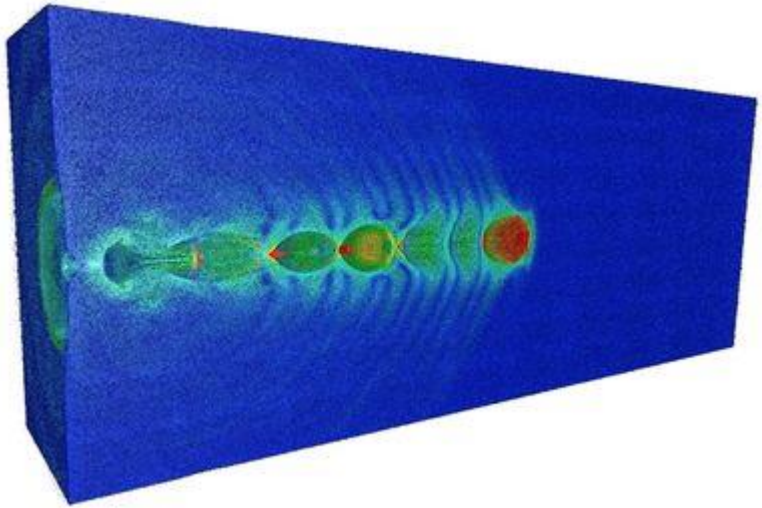


Autonomous accelerators

Prototypes towards fully autonomous, self-optimizing accelerators using machine learning.

Research Field Matter

What we do

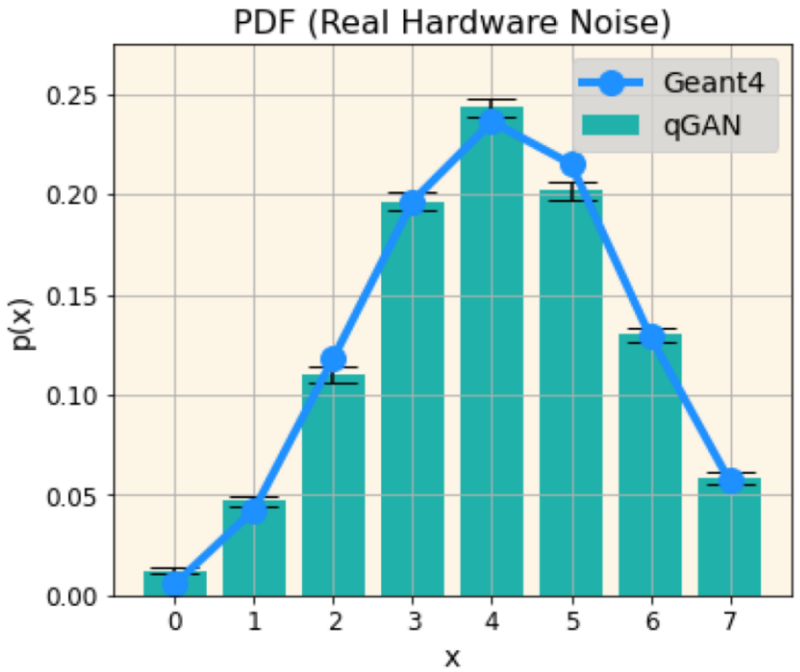


PLASMAPEPSC Center of Excellence

One of 4 European Centers of Excellence for Exascale Computing on Plasma Simulations for next generation supercomputers in Europe. Close collaboration between laser plasma and fusion community.

Research Field Matter

What we do

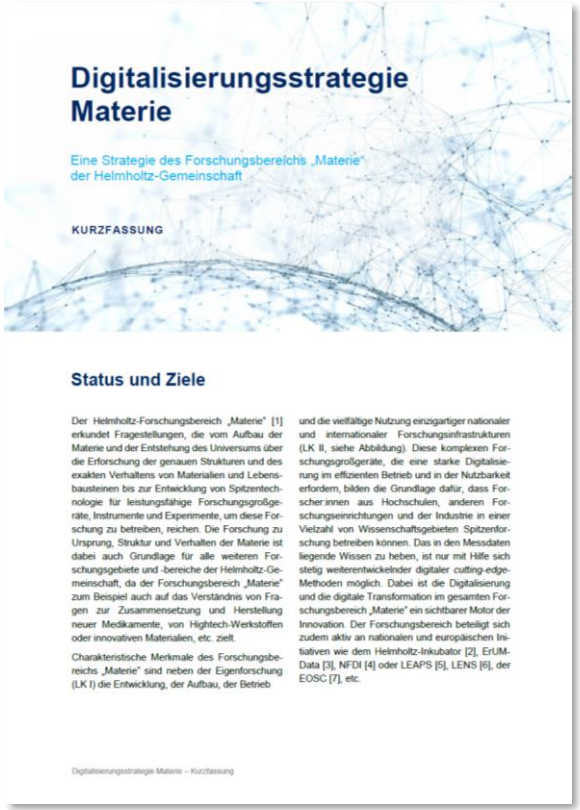


Quantum Computing for Particle Physics

Quantum Generative Adverserial Networks for fast detector simulations in Particle Physics.

Research Field Matter

Digitalisation Strategy



Digitalisierungsstrategie Materie

Eine Strategie des Forschungsbereichs „Materie" der Helmholtz-Gemeinschaft

KURZFASSUNG

Status und Ziele

Der Helmholtz-Forschungsbereich „Materie" [1] erkundet Fragestellungen, die vom Aufbau der Materie und der Entstehung des Universums über die Erforschung der genauen Strukturen und des exakten Verhaltens von Materialien und Lebensbausteinen bis zur Entwicklung von Spitzentechnologie für leistungsfähige Forschungsgroßgeräte, Instrumente und Experimente, um diese Forschung zu betreiben, reichen. Die Forschung zu Ursprung, Struktur und Verhalten der Materie ist dabei auch Grundlage für alle weiteren Forschungsgebiete und -bereiche der Helmholtz-Gemeinschaft, da der Forschungsbereich „Materie" zum Beispiel auch auf das Verständnis von Fragen zur Zusammensetzung und Herstellung neuer Medikamente, von Hightech-Werkstoffen oder innovativen Materialien, etc. zielt.

Charakteristische Merkmale des Forschungsbereichs „Materie" sind neben der Eigenforschung (LK I) die Entwicklung, der Aufbau, der Betrieb und die vielfältige Nutzung einzigartiger nationaler und internationaler Forschungsinfrastrukturen (LK II, siehe Abbildung). Diese komplexen Forschungsgroßgeräte, die eine starke Digitalisierung im effizienten Betrieb und in der Nutzbarkeit erfordern, bilden die Grundlage dafür, dass Forscher:innen aus Hochschulen, anderen Forschungseinrichtungen und der Industrie in einer Vielzahl von Wissenschaftsgebieten Spitzenforschung betreiben können. Das in den Messdaten liegende Wissen zu heben, ist nur mit Hilfe sich stetig weiterentwickelnder digitaler *cutting-edge*-Methoden möglich. Dabei ist die Digitalisierung und die digitale Transformation im gesamten Forschungsbereich „Materie" ein sichtbarer Motor der Innovation. Der Forschungsbereich beteiligt sich zudem aktiv an nationalen und europäischen Initiativen wie dem Helmholtz-Inkubator [2], ErUM-Data [3], NFDI [4] oder LEAPS [5], LENS [6], der EOSC [7], etc.

Digitalisierungsstrategie Materie – Kurzfassung

- A digitalisation strategy aligned with all Helmholtz centers in Matter for the whole Research Field.
- Set up for cross research field collaboration
- Fitting to the strategic position paper “Digitalisierungsstrategie der Helmholtzgemeinschaft”
- Concrete measures to reach the goals of the strategy

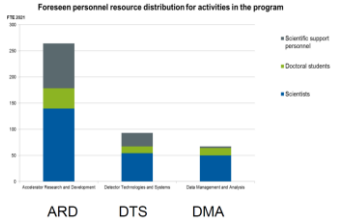
Research Field Matter

Digitalisation Strategy

1. **Enabling frontier science through federated compute, data and service infrastructures across centres and communities**
2. **Sustaining technological leadership in scientific simulation, data analytics and machine/experiment control through cutting edge technologies such as AI, Exascale and Quantum Computing**
3. **Providing an integrated support infrastructure for the diverse user communities of our research infrastructure**
4. **Supporting user communities in disseminating and using scientific data in an open, sustainable and F.A.I.R. way (Open Science!)**
5. **Developing and building large-scale research infrastructures using modern IT technologies**
6. **Developing and using intelligent and autonomous systems for maintenance/operation of machines, instruments & experiments**
7. **Using resources sustainably, from energy to materials, from data to software**
8. **Fostering the careers of young researchers in computer, information and data science**
9. **Cooperating with industry partners on applied research**
10. **Growing transfer of knowledge and technologies to industry and society both nationally and internationally**

Research Field Matter

Status 2023



In general

- DMA is smallest topic in MT.
- Many activities funded by third party funding such as Innopool (DATA-X, ACCLAIM, …), Pakt (ROCK-IT, …), Inkubator (Helmholtz.AI, HIFIS, HIP, HMC, HIDA), NFDI (PUNCH, DAPHNE, BASE, …), ErUM(-Data), EuroHPC (PlasmaPEPSC, …) and more.
- Only part of these activities will be sustainable (Inkubator, NFDI?, …)
- Many of these activities include both Helmholtz Centers in Matter and the scientific community (mainly the German communities, universities)
- Data-driven science seen as important, yet often still implemented as IT-centered topic
- Very fast pace in technologies (AI4Science, LLMs, Quantum Computing, Exascale Computing, Robotics and autonomous systems, …)
- RF Matter has clear digitalization strategy that now needs focus and prioritization with respect to resources

Research Field Matter

Status 2023

ST1:

Strengths:

- A common goal towards a unified data lifecycle approach

Weaknesses:

- Creating a strong engagement with scientific communities not yet visible
- A common strategy between centers and scientific communities not yet in place

Opportunities:

- Data is a „treasure“ of research facilities, make optimum use of it
- Taking up excellent existing projects (EOSC, NFDI, …)
- Use synergies with NFDI, ErUM and Inkubator and make RF Matter a strong partner
- Define clear interfaces between facilities and user communities
- Make research in Matter data-centric

Threats:

- Focus in ST1 (management of data vs data-driven science)
- Inhibiting further growth of data-driven science at facilities due to missing funding / talents

Research Field Matter

Status 2023

ST2:

Strengths:

- Excellency in methods (AI, Exascale, Quantum Computing) with many applications
- Sharing of solutions (ESCAPE OSSR)

Weaknesses:

- DMA enabling science must become more prominent, engagement of domain scientists

Opportunities:

- Stronger engagement of domain scientists
- Sustainability and reuse of solutions
- Access to frontier technologies

Threats:

- Software not seen as critical infrastructure the same way we recognize hardware as infrastructure
- Talents with algorithmic expertise go to industry because of opprotunities, salary and interest

Research Field Matter

Status 2023

ST3:

Strengths:

- Strong integration with experiments and targeted challenge-driven access
- Clear concept of digital twins of experiments, machines, systems
- Well-funded demonstrators and pilots

Weaknesses:

- Pace of development not in line with technology development

Opportunities:

- Autonomous facilities, intelligent steering of experiments for challenge-driven science

Threats:

- Lack of base funding support, many successes are driven by third-party funding
- Highly different pace between „small-scale experiment“ and „large-scale experiment“ communities

Research Field Matter

Towards POF V

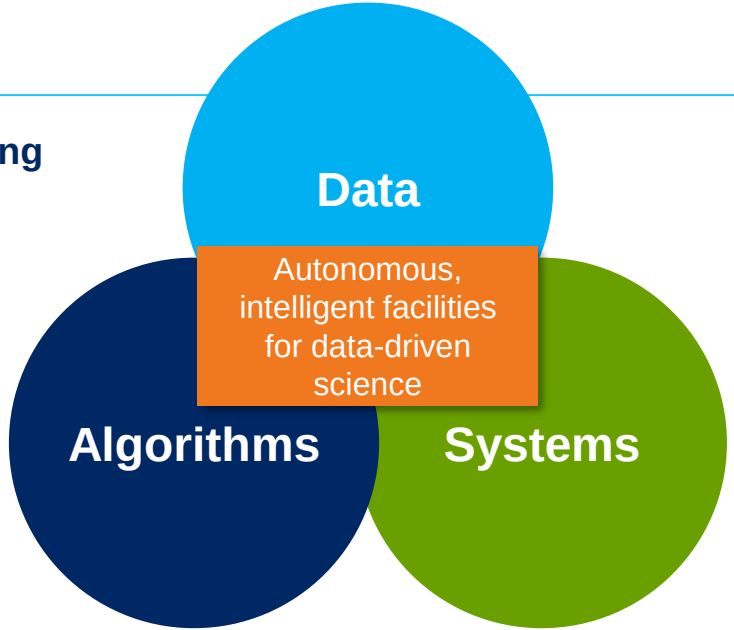
First thoughts on outcomes of April 2023 DMA strategy meeting

- The Matter digitalization strategy is a forward looking approach towards data-driven science in Matter
- Third-party funded initiatives have helped to support an otherwise underfunded topic
- Since 2019 funding of AI, Exascale computing, data-driven science has taken up pace
- DMA has to adjust to this new landscape, must define interfaces between facilities and user communities
- DMA has to leverage synergies with the third-party funded activities and invest in new technologies
- DMA has to set priorities for realising the Matter digitalisation strategy

Research Field Matter

Towards POF V

First thoughts on outcomes of April 2023 DMA strategy meeting



- **Adjust topic DMA to a radically changed landscape**
- Focus on excellent **data-driven science** that is sustainable
  - Intelligent data reduction for science
  - Integrated infrastructure from edge to data center
  - F.A.I.R., open data, synergies with communities
- **Algorithms for frontier technologies** as science drivers
  - AI, Exascale, Quantum, Edge, Cloud, federated, …
  - Intelligent algorithms (analysis, control, …)
  - Sustainable, reusable solutions
- **Autonomous, intelligent, integrated systems**
  - Combine machines, experiments, simulations and human users to steer towards discoveries
  - Enable targeted, challenge-driven access
  - Connect facilities and user communities
  - Optimize usage of facilities (discovery, impact, challenges, energy, …)