

## FAIRmat

*Sprecherin: Claudia Draxl, Humboldt-Universität zu Berlin & FAIR-DI. e.V.  
claudia.draxl@physik.hu-berlin.de*

FAIRmat repräsentiert das vielfältige und hochgradig heterogene **Forschungsgebiet** der Sektion Kondensierte Materie (SKM) der Deutschen Physikalischen Gesellschaft (DPG) mit ihren 12 Fachverbänden<sup>1</sup> sowie die chemische Physik von Materialien. Damit vertritt FAIRmat eine Vielzahl an Arbeitsgruppen an diversen Universitäten, außeruniversitären Instituten und Bundesanstalten, u.a. das Netzwerk Big-Data-Driven Materials Science der Max-Planck-Gesellschaft. Das Konsortium kooperiert mit einer Vielzahl von Forschungsnetzwerken und Großforschungsprojekten wie SFBs, Cluster-Projekten, Graduiertenkollegs, etc. Themen der Grundlagenwissenschaften, die (noch) nicht oder nicht primär durch die SKM repräsentiert werden, aber fachlich angrenzen, werden gerne aufgenommen, mit dem selbstverständlichen Ziel, dass im NFDI Gesamtvorhaben jede physikalische, materialwissenschaftliche und chemisch-physikalische Gruppe vertreten sein muss.

**Mit welchen Daten ist FAIRmat befasst?** Die Grundlagenforschung der Physik und Chemie von Materialeigenschaften umfasst die Materialsynthese und -charakterisierung ebenso wie diverse experimentelle und theoretische Untersuchungen. Die Forschungsdaten des Gebiets sind von entscheidender wissenschaftlicher, ökonomischer und sozialer Bedeutung, um dringend benötigte Technologien in den Bereichen Umwelt, Energie, Mobilität, IT, Medizin und mehr voranzutreiben. In dem Forschungsfeld von FAIRmat werden extrem heterogene Datenmengen von immensem Umfang erzeugt, die meist auf lokalen Servern gespeichert und kaum ausreichend beschrieben werden. Ohne eine effiziente Dateninfrastruktur sind diese Daten weder zugänglich noch wiederverwendbar. Um den Umfang und die Heterogenität der Daten zu verdeutlichen, denke man daran, dass ein einziges modernes Elektronenmikroskop täglich Daten im Umfang von 0.5 PB erzeugen kann. Dem gegenüber stehen z.B. auf theoretischer Seite weitaus mehr als 50 konzeptionell unterschiedliche Computer-codes (je >100.000 Zeilen lang) und komplexe Multiskalensimulationen der Polymerforschung oder umfangreiche Protokolle von Workflows der Materialsynthese. *Interoperability* und *Reusability/Repurposability* der Daten sind bis auf den Bereich Computational Materials Science (siehe NOMAD<sup>2</sup>) praktisch nicht vorhanden.

Die FAIRmat Dateninfrastruktur wird die umfassende Beschreibung, Charakterisierung und den Austausch von Daten ermöglichen, um die Grundlagenwissenschaften voranzubringen. Sie bezieht auch Methoden der Künstlichen Intelligenz mit ein. Dazu wird FAIRmat in enger Zusammenarbeit zwischen Domainspezialist\*innen sowie Informatik- und Datenexpert\*innen

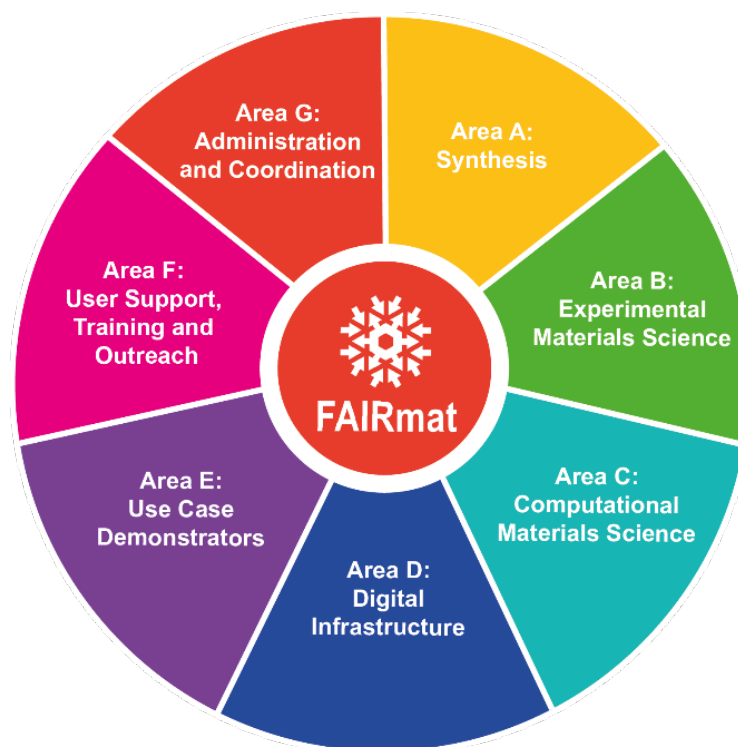
- eine dezentrale Dateninfrastruktur mit zentralem Zugang aufbauen;
- Metadaten-Schemas und Ontologien entwickeln;
- den effizienten Austausch von FAIRen Forschungsdaten ermöglichen;
- Studierenden, Wissenschaftler\*innen und Techniker\*innen (auch anderer Felder) einfachen Zugang zu den Daten ermöglichen;
- die Community mit Beratung und Schulungen unterstützen.

---

<sup>1</sup> Die SKM ist die bei weitem größte Sektion der DPG, was sich auch deutlich in der Größe ihrer Frühjahrstagung zeigt. Sie besteht aus den folgenden 12 Fachverbänden: Biologische Physik, Chemische Physik und Polymerphysik, Dünne Schichten, Dynamik und Statistische Physik, Halbleiterphysik, Kristalline Festkörper und deren Mikrostruktur, Magnetismus, Metall- und Materialphysik, Oberflächenphysik, Physik sozio-ökonomischer Systeme, Tiefe Temperaturen, Vakuumphysik und Vakuumtechnik. **Die SKM Frühjahrstagung ist nach dem March Meeting der American Physical Society (Condensed Matter) die zweitgrößte Physiktagung weltweit.**

<sup>2</sup> <https://nomad-lab.eu>

FAIRmat wird Tools zur Prozessierung, Verarbeitung, Speicherung und Wiederverwendung von Daten aus den Bereichen der Materialsynthese, der experimentellen und theoretischen Forschung an kondensierter Materie entwickeln und diese anhand von *Use Case Demonstrators* aus den unterschiedlichsten Gebieten der grundlegenden Materialwissenschaft demonstrieren. Die geplanten *Task Areas* sind in der folgenden Abbildung dargestellt:



FAIRmat basiert auf umfassender Erfahrung mit der weltweit größten Datenbank und Infrastruktur der computergetriebenen Materialwissenschaften, dem [NOMAD Laboratory](#)<sup>2</sup>, und dem Verein [FAIR-DI e.V.](#) (FAIR Data Infrastructure for Physics, Chemistry, Materials Science, and Astronomy e.V.)<sup>3</sup>. Als dessen materialwissenschaftliche Komponente ist FAIRmat nicht nur in Europa, sondern auch weltweit im Bereich der FAIRen Dateninfrastruktur für die Grundlagenwissenschaften der kondensierten Materie und angrenzende Gebiete ausgewiesen. Verschiedene Aspekte wurden bereits als *Proof-of-Concept demonstriert*. Ein Überblick dazu ist im [FAIRmat Newsletter](#)<sup>4</sup> vom Mai 2020 zu finden. Die im Juni 2020 abgehaltene [FAIR-DI Conference on a FAIR Data Infrastructure for Materials Genomics](#)<sup>5</sup> mit 539 aktiven Teilnehmer\*innen aus 33 Ländern bestätigt eindrucksvoll die nationale und internationale Vernetzung ebenso wie die Integration der jungen Generation zur Zusammenarbeit in der Gestaltung von zukünftigen Curricula (*Satellite Workshop on [NFDI@Teaching](#)*<sup>6</sup>).

FAIRmat und seine Vorreiter sind seit mehreren Jahren bereits international bestens eingebettet, beispielsweise in die Research Data Alliance (RDA), die European Open Science Cloud (EOSC) und GO FAIR und kooperieren nachweislich mit internationalen Key Players wie z. B. NIST (USA) sowie etlichen Dateninfrastruktur-Initiativen in China, Japan und Korea.

<sup>3</sup> <https://fair-di.eu>

<sup>4</sup> [https://www.fair-di.eu/uploads/documents/Newsletter/FAIRmat\\_Newsletter\\_Volume1.pdf](https://www.fair-di.eu/uploads/documents/Newsletter/FAIRmat_Newsletter_Volume1.pdf)

<sup>5</sup> <https://th.fhi-berlin.mpg.de/meetings/fairdi2020/>

<sup>6</sup> <https://th.fhi-berlin.mpg.de/meetings/fairdi2020/index.php?n=Meeting.NFDIatTeaching>

Innerhalb der NFDI ist FAIRmat mit den „benachbarten“ Konsortien im regen Austausch, um Grenzen und Schnittstellen klar zu definieren und Synergien auszuschöpfen und gemeinsam mit den anderen Konsortien die Abdeckung der gesamten Physik, Materialwissenschaft und chemischen Physik zu gewährleisten. Interoperabilität der FAIRmat-Dateninfrastruktur über Domänen hinweg ist für FAIRmat selbstverständlich, insbesondere im Abgleich mit der Chemie, mit Teilen von *Engineering and Technology*, den Geowissenschaften, der Astro- und Bio-Physik, Biologie und dem medizinischen Bereich sowie mit den R&D-Gruppen der Industrie. Die in der *Leipzig-Berlin-Erklärung zu NFDI-Querschnittsthemen der Infrastrukturentwicklung* im Detail beschriebenen Themen bilden einen guten Startpunkt für weitere Kooperationen. Dazu sind wir bereits mit dem NFDI-Direktorat im Gespräch.

**Vorgesehene Mitglieder des Konsortiums (Co-Sprecherinnen/Co-Sprecher und die weiteren, beteiligten Institutionen):**

<b>Co-Sprecher/in</b>	<b>Zugehörige Institution</b>
Martin Albrecht Leader Area A: Synthesis <a href="mailto:martin.albrecht@ikz-berlin.de">martin.albrecht@ikz-berlin.de</a>	Leibniz-Institut für Kristallzüchtung (IKZ) Max-Born-Str. 2, 12489 Berlin
Mark Greiner Leader Area B: Experimental Materials Science <a href="mailto:mark.greiner@cec.mpg.de">mark.greiner@cec.mpg.de</a>	Max-Planck-Institut für Chemische Energiekonversion (MPI-CEC) Stiftstrasse 34-36, 45470 Mülheim an der Ruhr
Matthias Scheffler Stellv. Sprecher Leader Area C: Computational Materials Science Leader Area F: User Support, Training and Outreach <a href="mailto:scheffler@fhi-berlin.mpg.de">scheffler@fhi-berlin.mpg.de</a>	Fritz-Haber-Institut (FHI) der Max-Planck-Gesellschaft (MPG) Faradayweg 4-6, 14195 Berlin  FAIR-DI e.V. c/o IRIS Adlershof Humboldt-Universität zu Berlin Zum Großen Windkanal 6, 12489 Berlin
Hans-Joachim Bungartz Leader Area D: Digital Infrastructure <a href="mailto:bungartz@tum.de">bungartz@tum.de</a>	Technische Universität München (TUM) Boltzmannstr. 3, 85748 Garching
Christof Wöll Leader Area E: Use Case Demonstrators <a href="mailto:christof.woell@kit.edu">christof.woell@kit.edu</a>	Karlsruher Institut für Technologie (KIT) Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen  FAIR-DI e.V. c/o IRIS Adlershof Humboldt-Universität zu Berlin Zum Großen Windkanal 6, 12489 Berlin