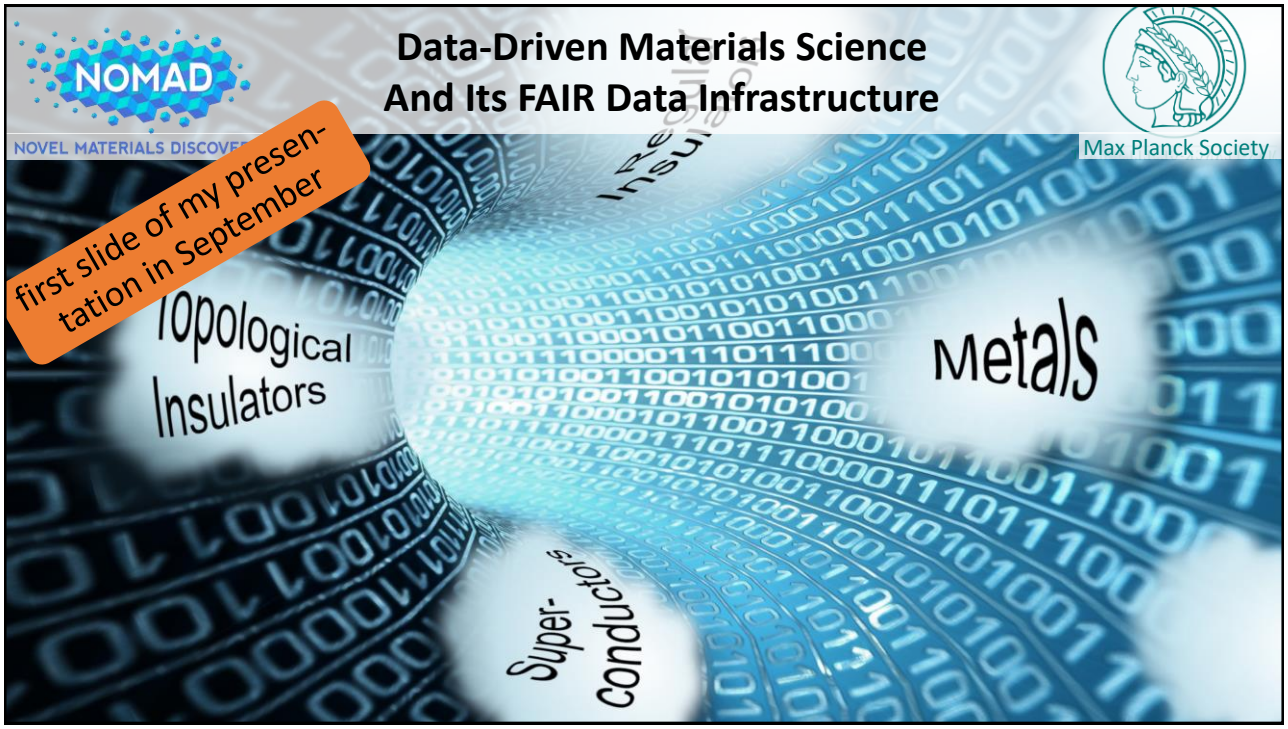
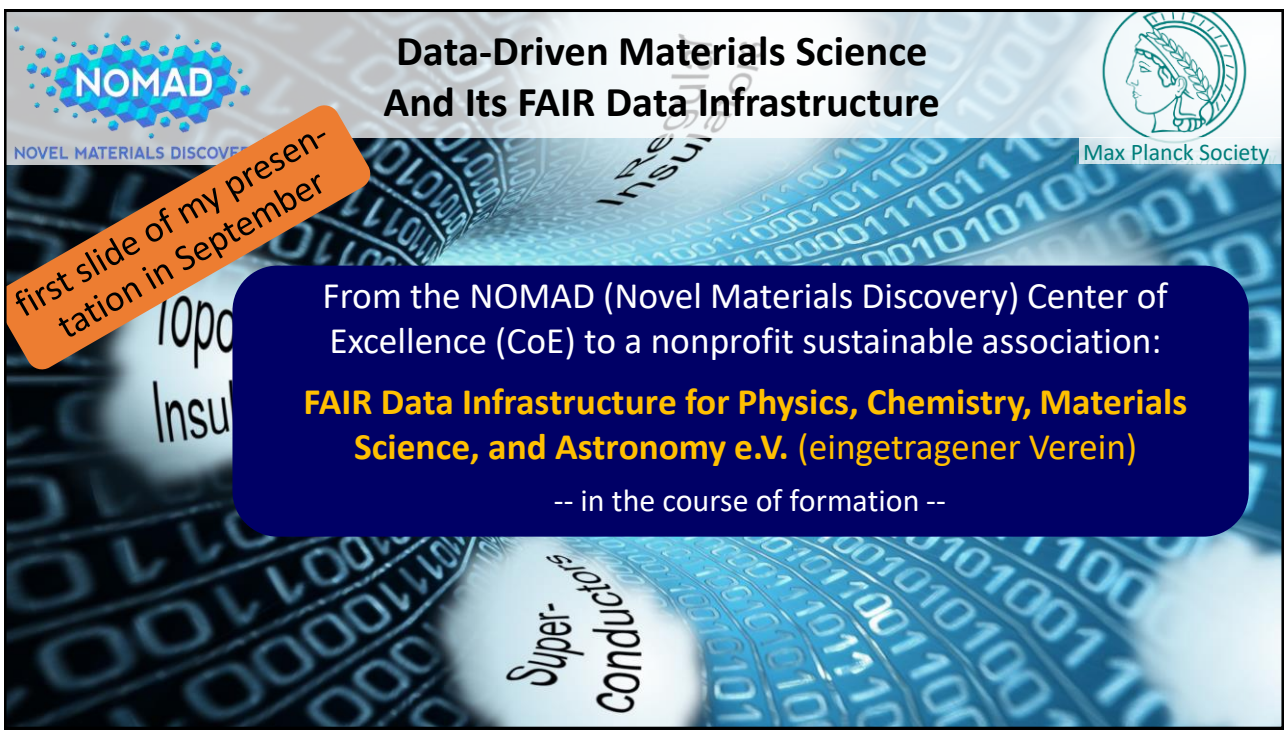

NOMAD
NOVEL MATERIALS DISCOVE
Data-Driven Materials Science
And Its FAIR Data Infrastructure
Max Planck Society
first slide of my presen-
tation in September
Topological
Insulators
Metals
Super-
conductors


NOMAD
NOVEL MATERIALS DISCOVE
Data-Driven Materials Science
And Its FAIR Data Infrastructure
Max Planck Society
first slide of my presen-
tation in September
From the NOMAD (Novel Materials Discovery) Center of Excellence (CoE) to a nonprofit sustainable association:
FAIR Data Infrastructure for Physics, Chemistry, Materials Science, and Astronomy e.V. (eingetragener Verein)
-- in the course of formation --
Super-
conductors

FAIRmat

"NOMAD CoE, FAIR-DI e.V., and a German Research-Data Infrastructure"

Germany is preparing a "Research-Data Infrastructure (NFDI); https://www.dfg.de/en/research_funding/programmes/nfdi/index.html" supported by a billion Euro over the next 10 years. The goals are somewhat similar to those of the NOMAD CoE and the FAIR-DI association/e.V. (https://fairdi.eu).

Specifically, we like to install a Consortium on "FAIR data infrastructure for materials science and related research topics". Funding will start in late fall next year and the application/preparation process is unusual and interesting. We should discuss our goals and how this NFDI may change science.
Coffee Talk [Monday 20th] 2019

FAIRmat

NFDI -- National Research Data Infrastructure

The aim of the national research data infrastructure is to systematically manage scientific and research data, provide long-term data storage, backup and accessibility, and network the data both nationally and internationally. The NFDI will bring multiple stakeholders together in *a coordinated network of consortia tasked with providing science-driven data services to research communities.*

NFDI Conference May 13-14, 2019:

NFDI -- National Research Data Infrastructure

The aim of the national research data infrastructure is to systematically manage scientific and research data, provide long-term data storage, backup and accessibility, and network the data both nationally and internationally. The NFDI will bring multiple stakeholders together in *a coordinated network of consortia tasked with providing science-driven data services to research communities.*

NFDI Conference May 13-14, 2019:

57 candidate consortia explained their goals. About 8 may be funded in the first round. We hope that FAIRmat will be one of them.

FAIRmat

FAIR data infrastructures for materials science and related research topics

Importance of the topic and my workload are significant.

From now on, and for the next ca. 2 years, my funding for students and postdocs will be mostly for the FAIRmat area (plus FHI-aims, electrical conductivity, GraFOx, ...).

FAIRmat

FAIR data infrastructures for materials science and related research topics

1. What consortia and NFDI as a whole should achieve

- Comprehensive research data management and increased efficiency throughout the scientific system
- Linking of research-oriented data services, improving interoperability
- Accepted, standardised processes and procedures in line with methodological requirements of (very) different disciplines
- A common voice for data concerns in the science-policy arena

But not

- Merely accumulate „Data“
- Collect local solutions (or repositories) waiting for future users
- One-size-fits-all
- Overly strict reglementation („juridification“)

The "executive team"

FAIRmat

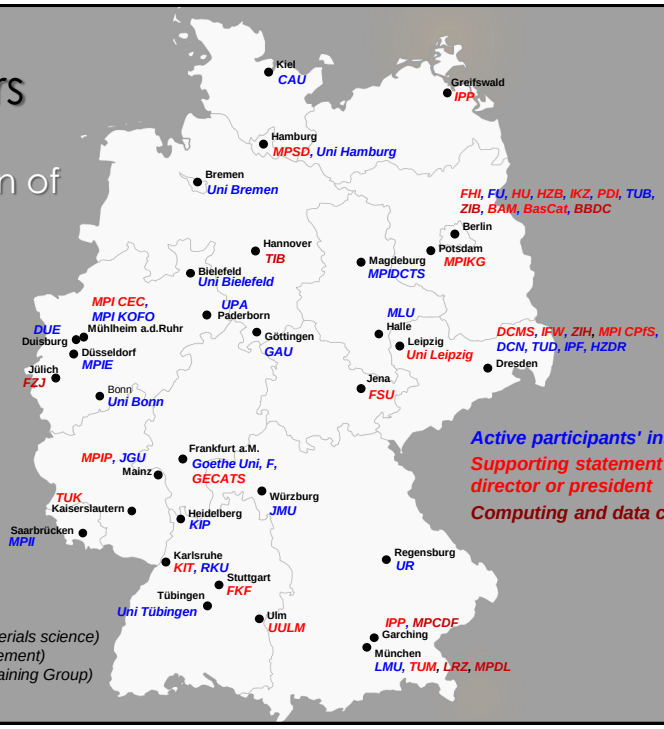
Claudia Draxl, Matthias Scheffler, Carsten Baldauf, Mark Greiner, etc.




FAIRmat members

FAIRmat

- Synergistic interaction of
  - a wide community
  - universities
  - research institutions
  - research organizations
  - computing and data centres



Cross-regional organizations and activites

FAIR-DI e.V.
Sektion kondensierte Materie der DPG
BiGmax (Max Planck Research Network on big-data driven materials science)
FDMentor (Cooperative BMBF project on research data management)
Quantum Mechanical Materials Modelling (DFG Research Training Group)

Materials – enormous diversity and impact

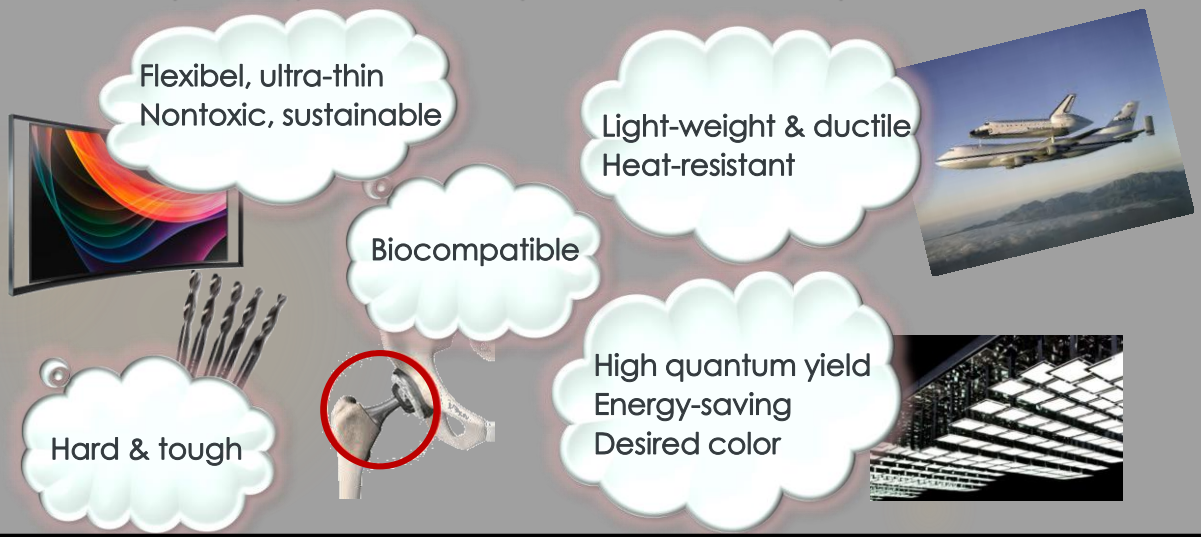
FAIRmat

- Every new product requires novel or improved materials

Materials – enormous diversity and impact

FAIRmat

- Every new product requires novel or improved materials



Materials science – an omnipresent topic

FAIRmat

- Owing to its economic and societal impact

  Recall the Materials Genome Initiative for Global Competitiveness by the White House, 2011

- Reflected in large communities, their research societies, and corrersponding conferences

  Materials Research Society / MRS Fall Meeting: 6,650 (2018)

  American Physical Society /APS March Meeting: >11,000 (2019)

  German Physical Society / DPG Spring Meeting: ~6,000 people

Data-infrastructure challenges

FAIRmat

- How to describe materials and their data?

  (Meta)data, workflows, ontologies

- Processing, storage, (preparation for) re-usage

  The 4 V's of big data – **V**olume, **V**ariety, **V**elocity, **V**eracity

FAIRmat principles

FAIRmat

- … a ***user-driven and non-bureaucratic approach to develop tools and infrastructures towards FAIR data processing, storage, curation, sharing, and future use of materials data.***

Bottom up!

Acceptance by the community

Make data ready for re-purposing and for AI tools

FAIRmat structure

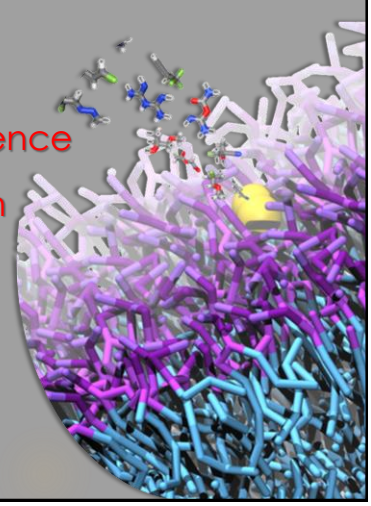
FAIRmat

- Materials science and related topics

Computation
- *Ab initio* calculations
- Classical simulations
- Excitations
- *and more*

Starting from our extensive experience

Extending existing infrastructures in terms of various methodologies, including multi-scale approaches



Extensive sharing of scientific data

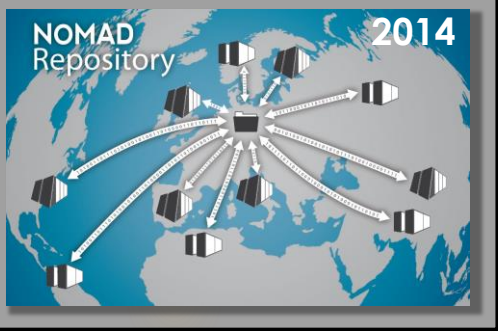
FAIRmat

- Data are **FAIR**
- Keeps data for at least 10 years
- Open access (embargo possible)
- Provides DOIs

NOMAD is an Implementation Network at GO FAIR

https://www.go-fair.org/, an international approach for the practical implementation of the European Open Science Cloud (EOSC).

Largest comp. materials data collection **worldwide**

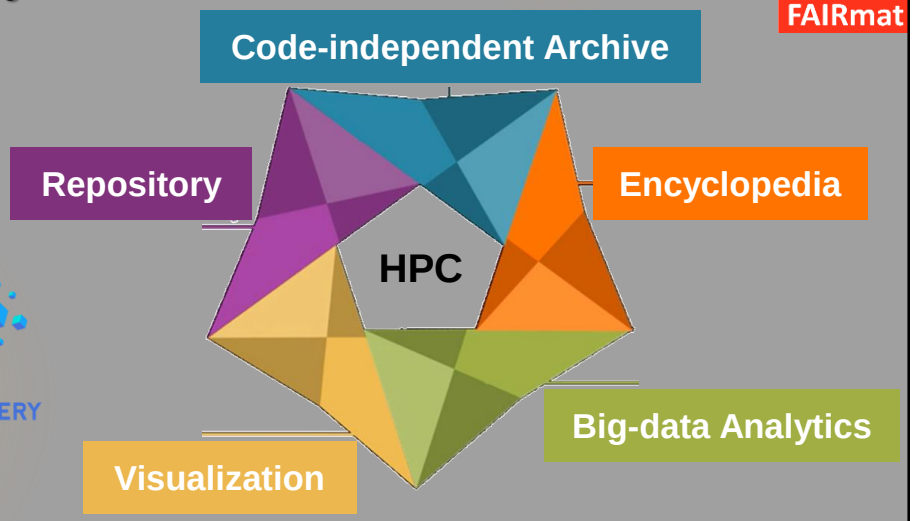


- Now part of FAIR-DI

FAIR Data Infrastructure for Physics, Chemistry, Materials Science, and Astronomy e.V.

Data services

FAIRmat



NOMAD

NOVEL MATERIALS DISCOVERY

- From raw data to artificial intelligence

FAIRmat structure

FAIRmat

- Materials science and related topics

Computation	Experiment	Synthesis	Functional materials
• *Ab initio* calculations • Classical simulations • Excitations • *and more*	• Electron microscopy • Photoemission • X-ray spectroscopy • *and more*	• From melt • From solution • From gas phase • *and more*	• Heterogeneous catalysis • Batteries • Opto-electronics • *and more*

Digital infrastructure

Selected NOMAD/FAIR-DI publications

FAIRmat

- *Towards Efficient Data Exchange and sharing for Big-Data Driven Materials Science: Metadata and Data Formats*
  L.M. Ghiringhelli, et al., npj Comput. Mater. 3, 46 (2017).
- *NOMAD: The FAIR Concept for Big-Data-Driven Materials Science*
  C. Draxl and M. Scheffler, Invited Review for Materials Research Society Bulletin (2018).
- *FAIR practices in Europe*
  M. Lautenschlager, H. Thiemann, C. Baldauf, P. Trilsbeek, and P. Wittenburg, Data Intelligence (2019).
- *Big-Data-Driven Materials Science and its FAIR Data Infrastructure*
  C. Draxl and M. Scheffler, Invited Perspective in Handbook of Materials Modeling, Springer (2019).

Embedded in community: conferences

FAIRmat

- Big Data of Materials Science - Critical Next Steps
  Lausanne, Switzerland, Nov. 30 – Dec. 4, 2015
- Towards a Common Format for Computational Materials Science Data
  Lausanne, Switzerland, Jan. 25 – Feb. 4, 2016
- Open Databases Integration for Materials Design
  Lausanne, Switzerland, Jun. 11-14, 2019
- Shared metadata and data formats for big-data driven materials science: A NOMAD - FAIR-DI workshop
  Berlin Adlershof, Germany, Jul. 8 - 12, 2019

Thank you!!

https://fairdi.eu/fairmat