Issue 1 /July 2021

NOVEL MATERIALS DISCOVERY

NOMAD Center of Excellence

Mission

Exascale computing will have a profound impact on everyday life in the coming decades. At 10^{18} operations per second, exascale supercomputers will be able to quickly analyze massive volumes of data and more realistically simulate complex processes. The goal of the NOMAD Center of Excellence is to ***bring computational materials science to the next level of supercomputing.***

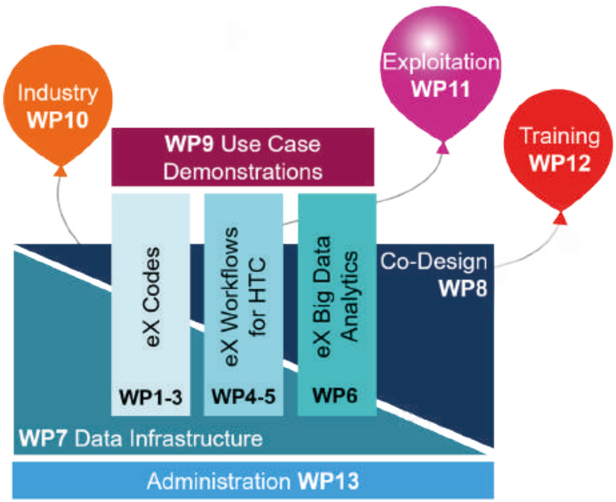
The NOMAD CoE assesses and exploits the characteristics of ***extreme-scale data and exascale computing for computational materials science***, to enable investigations of systems of higher complexity (space and time), consideration of metastable states and temperature, and all this at ***significantly higher accuracy and precision than what is possible today.***

Systematic studies and predictions of novel materials ***to solve urgent energy, environmental, and societal challenges*** require such significant methodological advancements targeting the upcoming exascale computers. Key NOMAD examples are ***catalytic water splitting for hydrogen production and the transformation of waste heat into useful electricity.***

NOMAD CoE receives funding from the European Union's Horizon 2020 research and innovation program under the grant agreement N° 951786. The materials presented and views expressed here are the responsibility of the author(s) only. The EU Commission takes no responsibility for any use made of the information set out.

Structure and goals

The NOMAD CoE is organized in three topological pillars and overall thirteen work packages (WPs).



- The goal of Pillar 1 is ***to advance ab initio computational materials science for entire code families*** to enable tackling more complex problems than what is possible today.

- Efficient use of the exascale-ready libraries and codes developed in Pillar 1 requires ***sophisticated workflows*** that are also capable of managing high-throughput computations and taking full advantage of exascale resources. This is the focus of Pillar 2.

- The overall aim of Pillar 3 is to utilize exascale technology to ***advance the existing big data tools and bring them towards near-real-time performance*** with a response time of seconds or less.

Horizontal activities support the three vertical research Pillars.

- WP7 is dedicated to ***Data Infrastructure*** and makes the NOMAD data infrastructure exascale-ready.

- The task of WP8 is the ***Co-Design*** across *ab initio* computational materials sciences (aiCMS) and HPC to deliver new inputs to HPC architects. The aim is to ensure that the hard- and software developments go hand in hand.

- We have ***Use-case Demonstrators*** (WP9) to test and demonstrate the NOMAD CoE developments in urgent energy and environmental challenges. The NOMAD examples are catalytic water splitting for hydrogen production and waste heat recovery.

- WPs10-13 add value to our research activities. The tasks include outreach to academia and industry, training, and administration.

The consortium

The NOMAD CoE represents a ***dynamic, trans-European consortium.*** It utilizes a 2-layer team structure. The first shell is represented by the core research team from 10 academic institutions and high-performance computing centers across Europe.

The second shell includes researchers from other aiCMS groups, industry participants, scientific and commercial experts and other stakeholders

The NOMAD PIs

Max Planck Society:
Fritz Haber Institute: Matthias Scheffler (Coordinator)
Max Planck Computing and Data Facility: Erwin Laure

Humboldt-Universität zu Berlin: Claudia Draxl (Deputy-Coordinator)

Aalto University: Patrick Rinke

Barcelona Supercomputing Center: José-Maria Cela

Commissariat à l'Energie Atomique et aux Energies Alternatives: Christine Menache

CSC – IT Center for Science: Kimmo Koski

Technical University of Denmark: Kristian Sommer Thygesen

Technische Universität Wien: Andreas Grüneis

Université Catholique de Louvain: Xavier Gonze

University of Cambridge: Gábor Csányi

University of Latvia: Andris Gulans

University of Warwick: James Kermode

Do you already know the NOMAD Oasis?

Your local, self-contained NOMAD server for managing your lab's data, FAIRification, and analytics by Artificial Intelligence tools

NOMAD Oasis is characterized by the following key features:

- Runs fully independent of external sources, also behind strong firewalls
- Provides full access to the data of the central NOMAD server and ensures regular synchronization
- Provides flexibility for adjustments and advancements towards users' local, domain-specific, individual needs to acquire, manage, and analyze their data
- Enables the users to contribute their data and tools to the central NOMAD server in a convenient way (if they wish to do so)
- Provides a Graphical User Interface (GUI) for characterizing materials according to all available quantities stored at the central NOMAD server and/or the Oasis
- Integrates all AI tools from the NOMAD Artificial Intelligence Toolkit

Obtaining the NOMAD Oasis

We offer two types of licenses for academic or non-academic use, respectively. More information can be found here: https://nomad-coe.eu/about-oasis

For further questions contact: victoria.coors@fairdi.eu.

NOMAD tutorials

NOMAD provides virtual tutorials and hands-on exercises covering various important aspects of the NOMAD database. The tutorial series will be reissued every year.

All videos and materials can be found online: https://th.fhi-berlin.mpg.de/meetings/nomad-tutorials/index.php?n=Meeting.Home

Repository and Archive

https://th.fhi-berlin.mpg.de/meetings/nomad-tutorials/index.php?n=Meeting.Tutorial1
The main focus is the FAIR sharing of materials science data and how to do it with NOMAD. This covers the publication of new data and the exploration and download of NOMAD's existing data; both through our browser-based interface and API.

Encyclopedia

https://th.fhi-berlin.mpg.de/meetings/nomad-tutorials/index.php?n=Meeting.Tutorial2
In addition to the NOMAD Archive, the NOMAD Encyclopedia introduces the complementary materials-oriented data organization. We cover the key difference between these two approaches, introduce the methods used to link individual calculations into specific materials, and discuss how this new layer of information brings additional value to the data.

Artificial Intelligence Toolkit

https://th.fhi-berlin.mpg.de/meetings/nomad-tutorials/index.php?n=Meeting.Tutorial3
This tutorial introduces the functionalities of the NOMAD AI Toolkit, the web-based framework for querying, filtering, and performing AI analysis on the data contained in the NOMAD Archive. The user interface is based on the Jupyter notebook environment, which is becoming increasingly popular in scientific research.

Workflow Management

https://th.fhi-berlin.mpg.de/meetings/nomad-tutorials/index.php?n=Meeting.Tutorial4Amp5
This tutorial introduces the Atomic Simulation Recipes (ASR) – an open source Python framework for working with atomistic materials simulations in an efficient and sustainable way that is ideally suited for high-throughput studies. The second part is on the FireWorks workflow manager, a free, open-source code for defining, managing, and executing workflows.

NOMAD lecture series

An online course on "Big Data and Artificial Intelligence in Materials Sciences" was offered recently. The course covers the hottest topics in artificial intelligence, including machine learning, compressed sensing, and data mining. General and specific AI concepts, in particular for the design and discovery of improved, new, and novel materials for technological applications are introduced.

All videos can be found online: *https://www.nomad-coe.eu/course-on-big-data-and-artificial-intelligence/lecture-materials*

Topics of the lectures:

- ***General Introduction to Big-Data-Driven Materials Science*** by Matthias Scheffler
- ***NOMAD Repository, Archive, Encyclopedia*** by Claudia Draxl
- ***Introduction to Artificial Intelligence and Machine-Learning Methods*** by Luca Ghiringhelli
- ***Compressed Sensing Meets Symbolic Regression: SISSO*** by Luca Ghiringhelli
- ***Decision Trees and Random Forests*** by Daniel Speckhard
- ***Regularized Regression and Kernel Methods*** by Santiago Rigamonti
- ***Unsupervised Learning*** by Luigi Sbailò
- ***Artificial Neural Networks and Deep Learning*** by Angelo Ziletti
- ***Materials Data, 4V, FAIR Principles*** by Claudia Draxl
- ***Subgroup Discovery, Rare-Phenomena Challenge, and Domain of Applicability*** by Matthias Scheffler
- ***Interpretability and Causality*** by Jilles Vreeken
- ***Applications in Real Materials*** by Rampi Ramprasad
- ***AI in Experiment*** by Christoph T. Koch
- ***Fusion of Experimental and Computational Data by AI*** by Lucas Foppa

Upcoming events

Events (co-)organized by NOMAD CoE:

September 7-8, 2021: ***Online tutorial on The Open Databases Integration for Materials Design (OPITIMADE)***
More information: https://th.fhi-berlin.mpg.de/meetings/nomad-tutorials/index.php?n=Meeting.Home

November 1-3, 2021: ***Joint CECAM - NOMAD - E-CAM Workshop: Modeling materials at realistic space and time scales via optimal exploitation of exascale computers and AI***
Location: CECAM HQ, Lausanne, Switzerland
More information and registration: https://nomad-coe.eu/events/nomad-e-cam-workshop

November 15-19, 2021: ***High-throughput workflows for materials science with the Atomic Simulation Environment and FireWorks***
Location: Technical University of Denmark, Lyngby, Denmark
More information and registration: https://nomad-coe.eu/events/ase-fireworks-workshop

June 16-21, 2022 ***NOMAD summer school: Towards exascale solutions in Green function methods and advanced DFT***
More information will be announced in due time

Visit our website to stay updated about all our events: https://nomad-coe.eu/coe_events

Recorded videos of past events can be found here: https://nomad-lab.eu/videos

Get in touch with us!

Get in touch with us and stay updated about NOMAD

www.nomad-coe.eu @NoMaDCoE YouTube The NOMAD Laboratory contact@nomad-coe.eu

Your contact persons for specific questions regarding....

General questions concerning the CoE:	contact@nomad-coe.eu	
NOMAD *ab initio* software	FHI-aims: sebastian.kokott@fhi.mpg.de	exciting: claudia.draxl@physik.hu-berlin.de
	Abinit: xavier.gonze@uclouvain.be	GPAW: thygesen@fysik.dtu.dk
Workflows:	ASE/ASR: thygesen@fysik.dtu.dk	FireWorks: geoffroy.hautier@uclouvain.be
Data analytics / artificial intelligence:	luca.ghiringhelli@fhi.mpg.de	
NOMAD data infrastructure:	markus.scheidgen@physik.hu-berlin.de	
Exascale HPC facilities:	msmit1@bsc.es	
Use-case demonstrators:	christian.carbogno@fhi.mpg.de and thygesen@fysik.dtu.dk	
NOMAD Oasis:	victoria.coors@fairdi.eu	