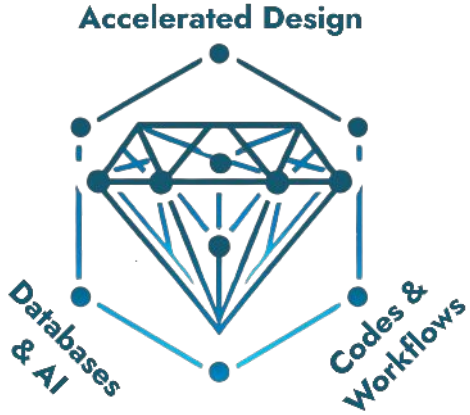
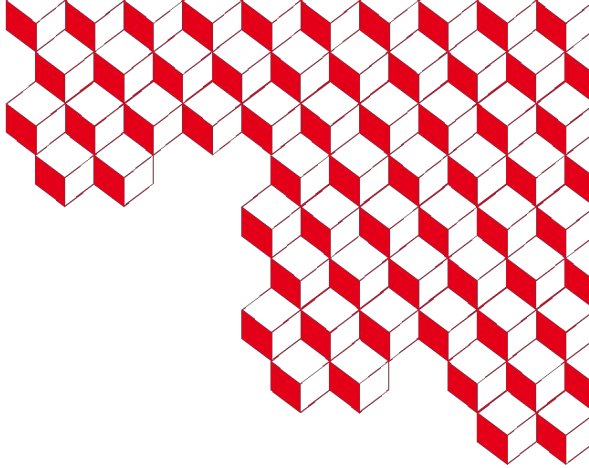
DIAMOND
ANR-22-PEXD-0015





Setting up data tools as part of the Diadem programme

Thierry Deutsch (CEA Grenoble)
V. Bergeaud, N. Jakse, D. Rodney, M. Saitta, I. Setoain, F. Willaime

PEPR* DIADEM

DIscovery Acceleration for the Deployment of Emerging Materials

Directors:

- **Frédéric Schuster, CEA**
- **Mario Maglione, CNRS**

*Priority Equipment and Research Program

Accelerated the discovery of materials

CHALLENGES

- Many technological fields are based on the discovery of materials: energy, transport, health, digital transition, etc.

- The actual implementation of new materials is delayed all the more (more than a decade of trial and error) as their complexity increases.

- The Green Deal context and sustainable growth requirements: life cycle management, energy efficiency, minimising the use of critical resources, etc

e.g. Materials Genome Initiative (USA, 2011)

Guidelines and specific features

ISSUES

- Accelerate the discovery of materials for major transitions: energy (batteries, hydrogen, nuclear power of the future, etc.) and the environment, digital technology (electronics, etc.), health, etc.
- Substitution of critical and/or toxic raw materials
- Diversity of materials and implementation processes (specificity of DIADEM compared to other initiatives of the same type)

OBJECTIVES

- Provide France with a network of platforms dedicated to the accelerated discovery of materials
- Make these platforms available to the scientific community via open calls for projects and beyond the PEPR.
- Develop nationwide synergies between materials science and artificial intelligence

The different aims of PEPR DIADEM

Platforms

- High-throughput synthesis and shaping
- High-throughput characterisation
- Digital tools and databases (DIAMOND)

Targeted demonstration projects

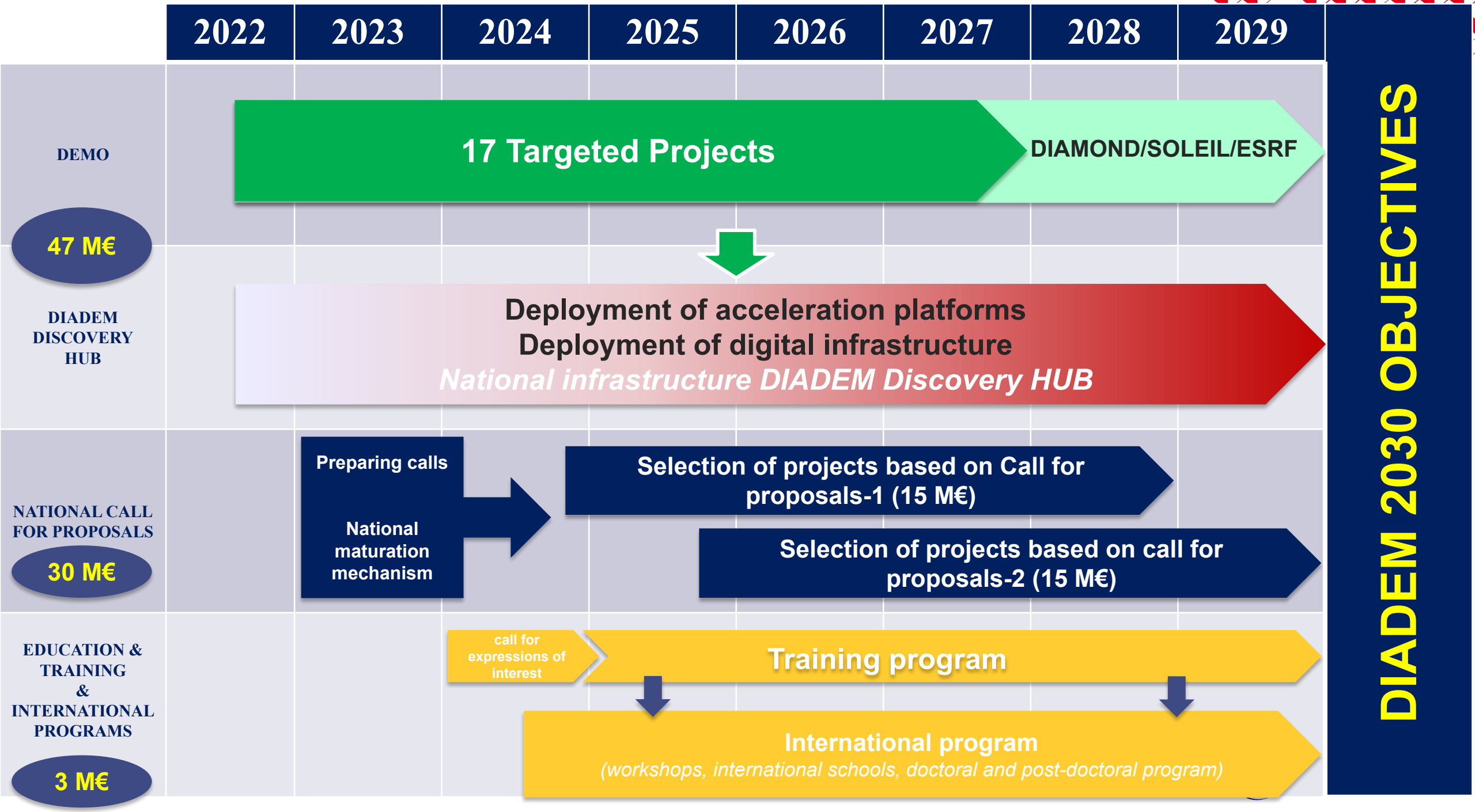
- Integrated methodological demonstrators
- Building platforms

Calls for projects (3 calls)

- 30 to 40 projects with funding of between €750k and €1m
- Open to the whole community
- Possibility of opening up internationally

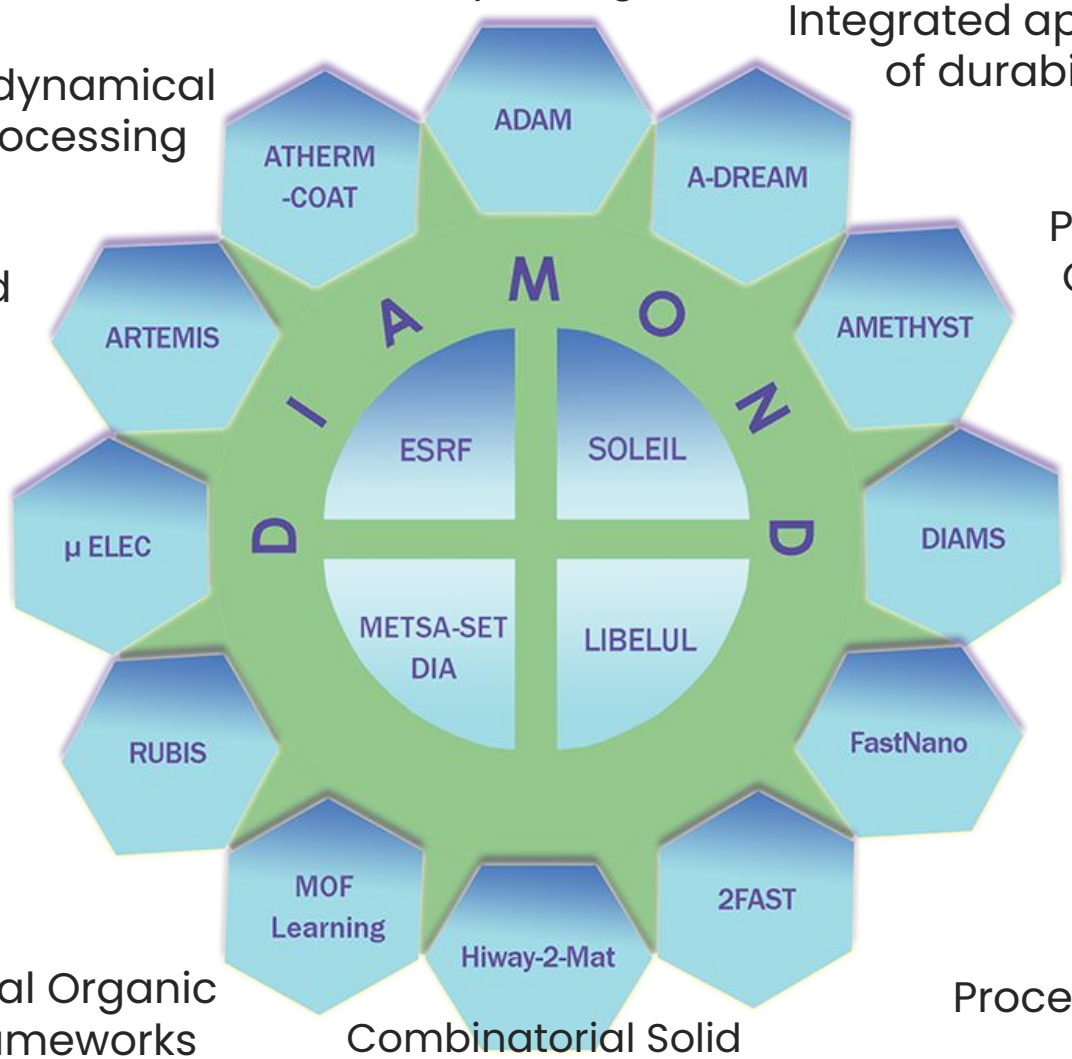
Calls for expression of interests *(5 projects)*

- Training tools
- International schools


2022
2023
2024
2025
2026
2027
2028
2029
DEMO
47 M€
17 Targeted Projects
DIAMOND/SOLEIL/ESRF
DIADEM DISCOVERY HUB
Deployment of acceleration platforms
Deployment of digital infrastructure
National infrastructure DIADEM Discovery HUB
NATIONAL CALL FOR PROPOSALS
30 M€
Preparing calls
National maturation mechanism
Selection of projects based on Call for proposals-1 (15 M€)
Selection of projects based on call for proposals-2 (15 M€)
EDUCATION & TRAINING & INTERNATIONAL PROGRAMS
3 M€
call for expressions of interest
Training program
International program
(workshops, international schools, doctoral and post-doctoral program)
DIADEM 2030 OBJECTIVES

DIADEM Discovery HUB

AI-supported open platforms (demonstration projects)



4D printing

Integrated approach of durability

Experiment on thermodynamical parameters and AI Processing

Polymers & Hybrids Organic/Inorganic

Architectured Materials

Digital Metallurgy

Microelectronics

Nanomaterials

Ceramics and Ceramics Matrix Composites

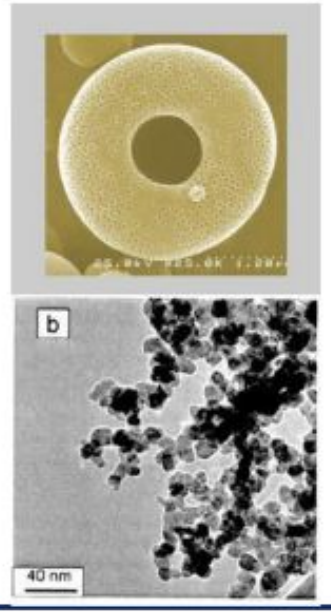
Microfluidic Processes Intensification

Metal Organic Frameworks

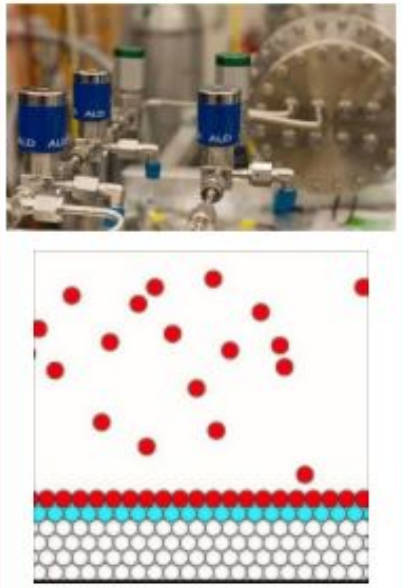
Combinatorial Solid State Chemistry

A strong effort on AI for Advanced Manufacturing

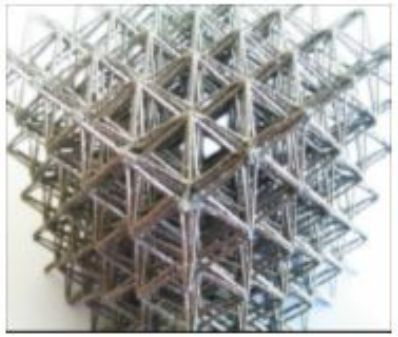
Fast Nano



A Therm Coat

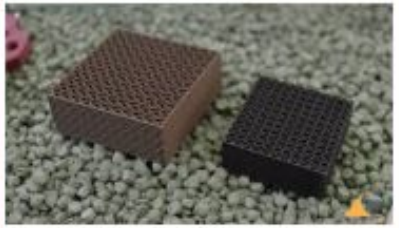


DIAMS



ARTEMIS

RUBIS

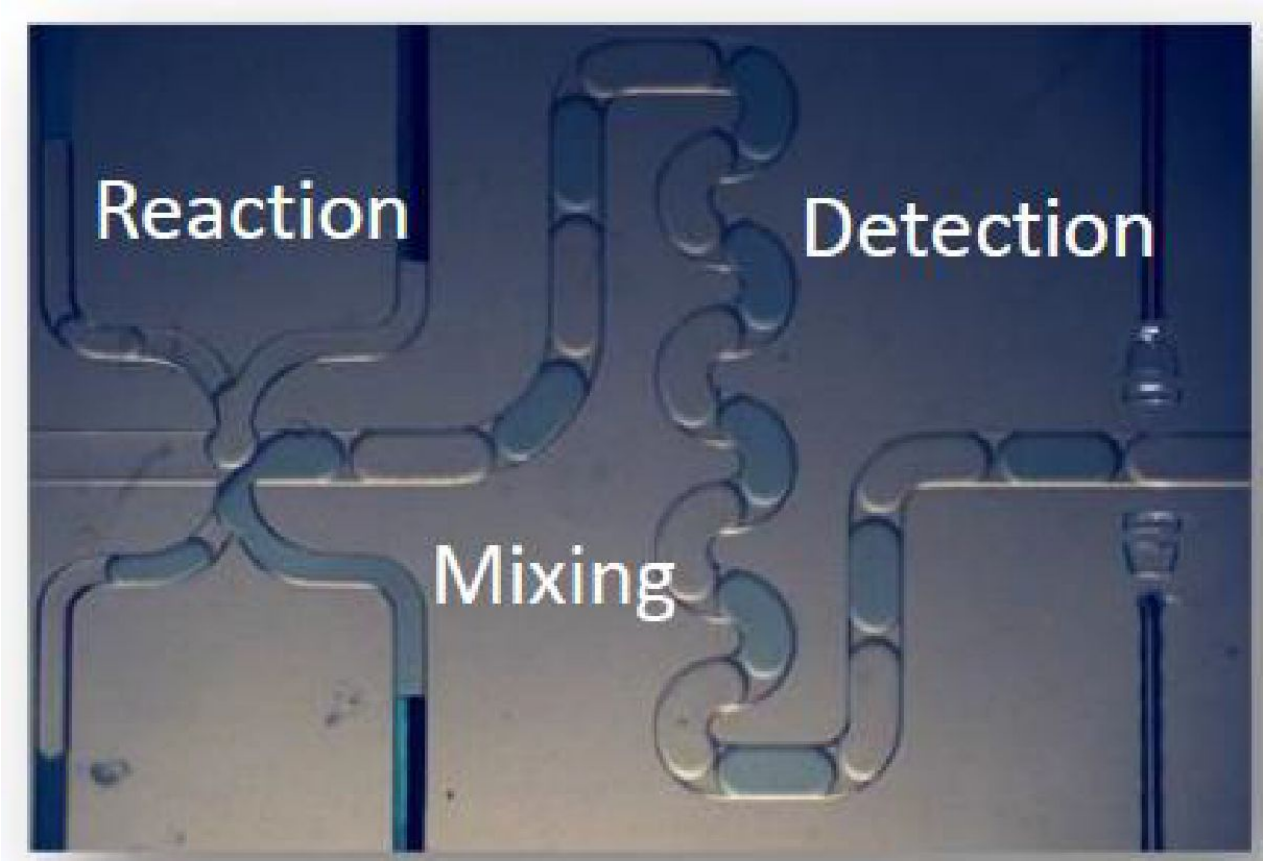


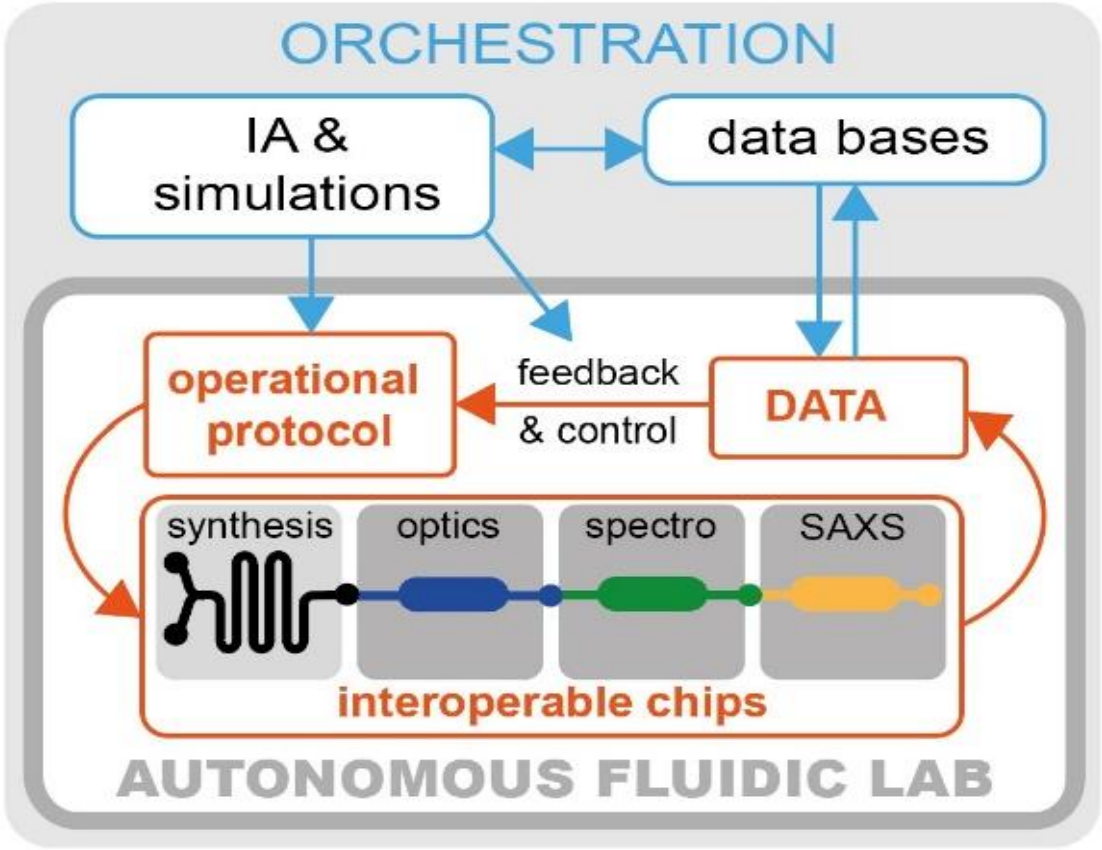
Nanomanufacturing

Surface enginering

3D & 4D printing

2FAST: Deployment of self-driving labs





Example of microfluidics process intensification (based on workflows)

Fast Nano: High throughput SAXS analysis (X-Ray)

DIADEM national calls for proposals

2023 national call – projects started in 01/2024

Accelerated discovery of materials for a Green Deal

- Metals and Alloys
- Polymers, molecular and biosourced Materials
- Non-metallic materials
- Catalysts

Accelerated discovery of sovereign and sustainable processes

- Nanoparticule synthesis
- Surface engineering
- Metallurgy

2023 call for projects: 17 projects selected

Area 1: Accelerated discovery of innovative materials for a Green Deal

- Metallic materials (4 projects)
- Molecular, polymer and bio-based materials (3 projects)
- Non-metallic materials (3 projects)
- Catalysts (3 projects)

Area 2: Accelerating the mastery of synthesis and formatting processes

- Nanoparticle synthesis (1 project)
- Surface engineering (2 projects)
- Metallurgy (1 project)

AI very present in all projects

New call for projects underway

Major axis in 2024 call for projects to be selected in 2025

ACCELERATED DISCOVERY OF MATERIALS FOR A GREEN DEAL (2/3)

ACCELERATED DISCOVERY OF SOVEREIGN AND SUSTAINABLE PROCESSES (1/3)

Focused topics

- **Thermoelectrics**
- **Spintronics**
- **Materials for the medecine of the future**

- AI for accelerated discovery is the key requirement
- Integration of DIADEM platforms is mandatory
- International cooperation and industry partners are welcome (both not funded)
- Interfaces with other PEPR programs is foreseen (Hydrogen, Batteries, Electronics, AI,…)

DIADEM International

Integrate DIADEM into the European digital infrastructure for materials discovery

PAST AND UNDERGOING EC ACTIONS

- BIG-MAP on batteries (2020-2023)
- NOMAD CoE Novel Materials Discovery
- MAX on exascale computing (2018-2026)

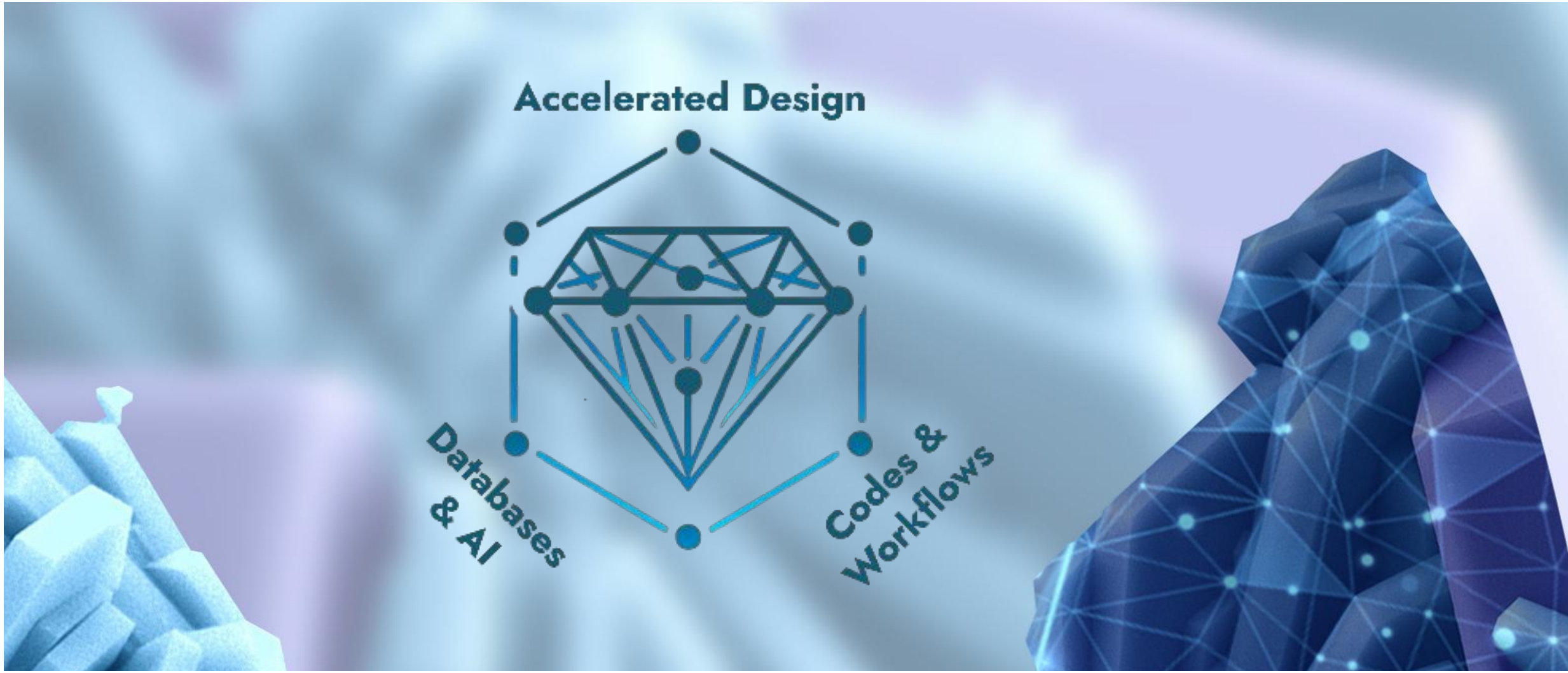
AFTER THE LAUNCH OF DIADEM

- S. Stier et al; white paper on Materials Accelerations Platforms with 28 European research institutions
- Took part to several meetings with EC Offices last one in Namur INDTECH 05/2024

FUTURE ACTIONS

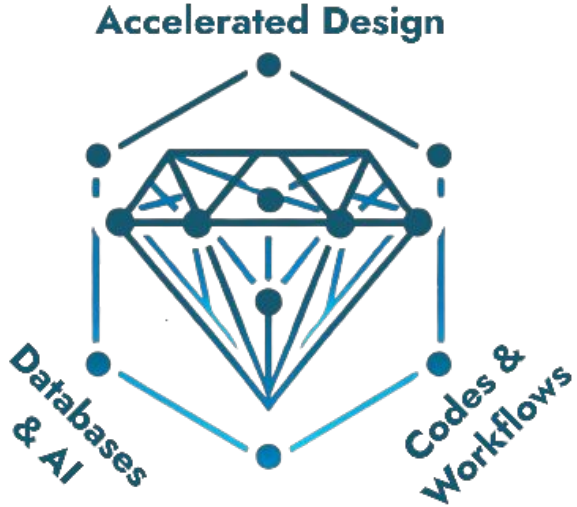
- DIADEM meeting MATERIAL DIGITAL, EMIRI and IAM4EU in Paris July 10, 2024
- The containerized codes and workflows gathered on the DIADEM DIAMOND-CW platform will be accessible to all European partners
- The DIADEM DIAMOND-DM Data Management infrastructure will be open to European partners within dedicated consortia
- DIAMOND is fully committed to work at the European level to make the data sharing easier.

DIAMOND digital platform



DIAMOND digital platforms

Coordinator: François Willaime (CEA/DES)
Vice-coordinator: Noël Jakse (SIMaP, Grenoble)

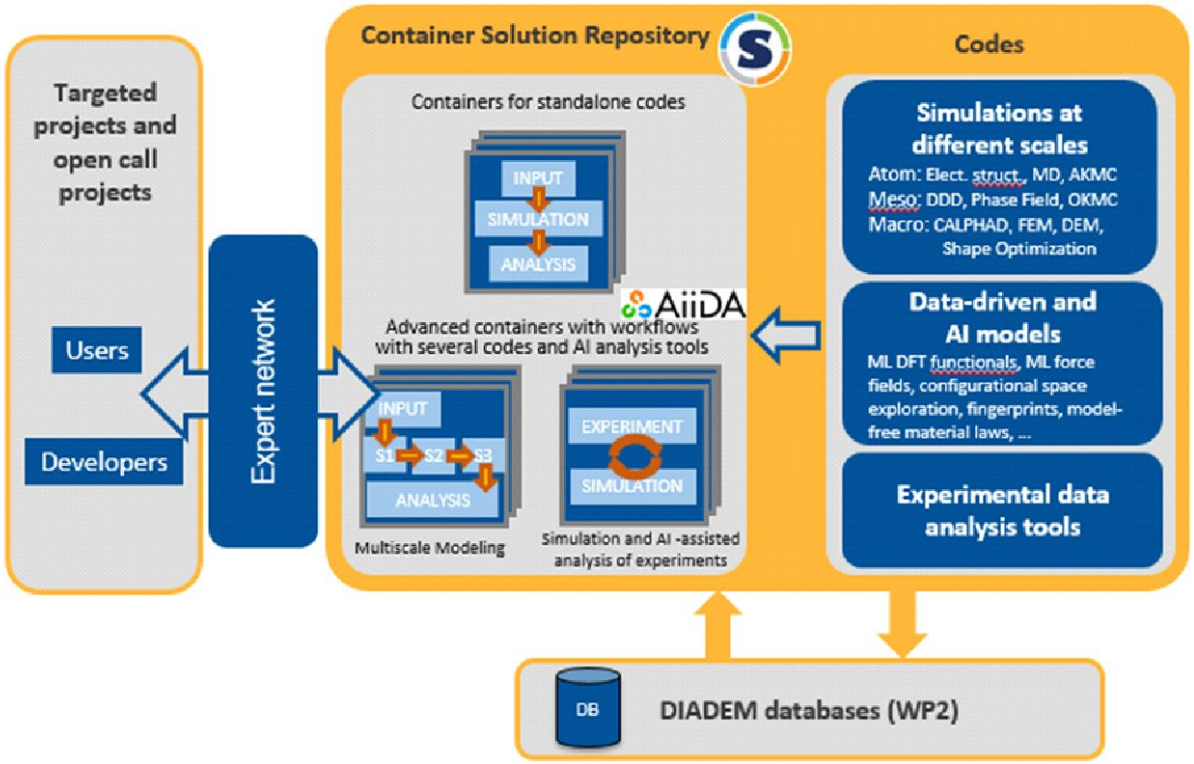


Budget : 10 M€ for 6 years
Kick-off : 11/10/2023

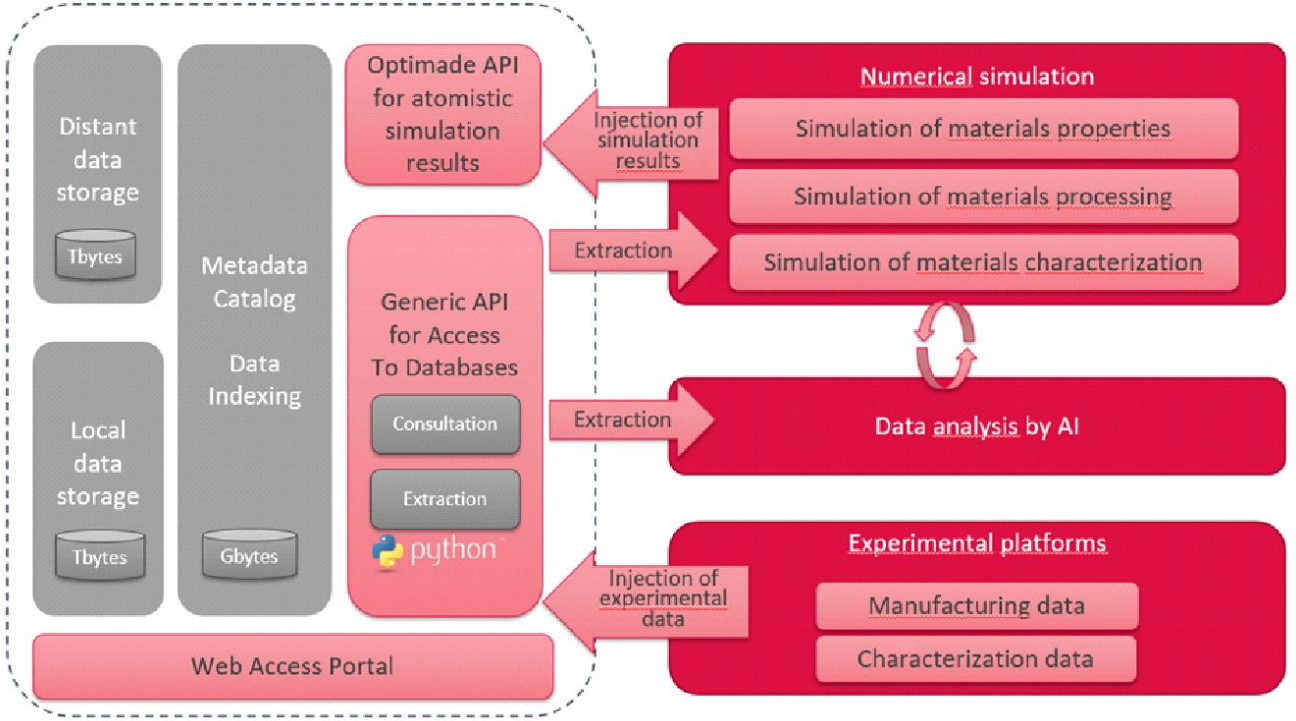
	Workpackages
WP1	Code and workflow platform (DIAMOND-CW)
WP2	Data management infrastructure (DIAMOND-DB)
WP3	Demonstrator – Machine Learning Interatomic Potentials

Create digital infrastructures

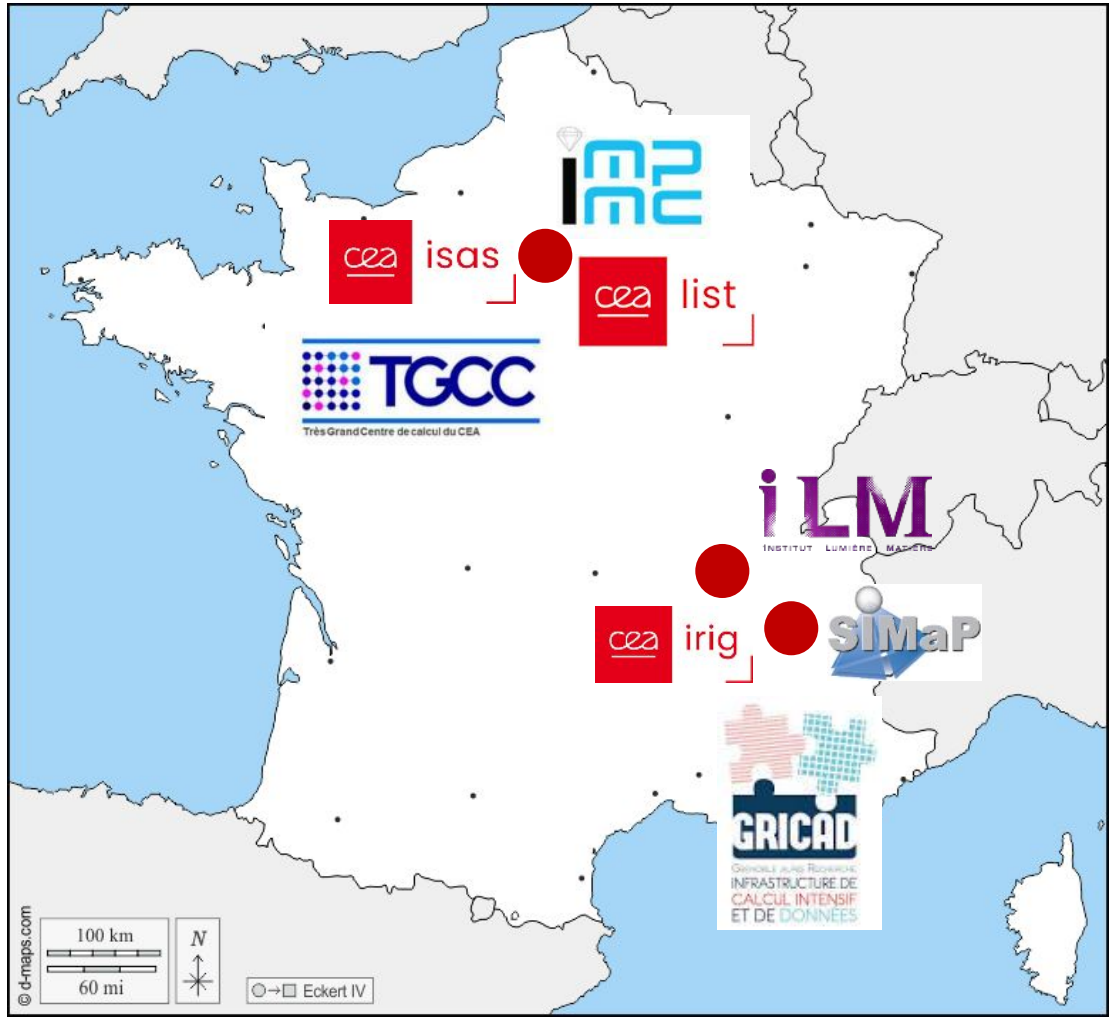
DIAMOND-CW
Codes & Workflows



DIAMOND-DB
Data Management



Teams of the consortium

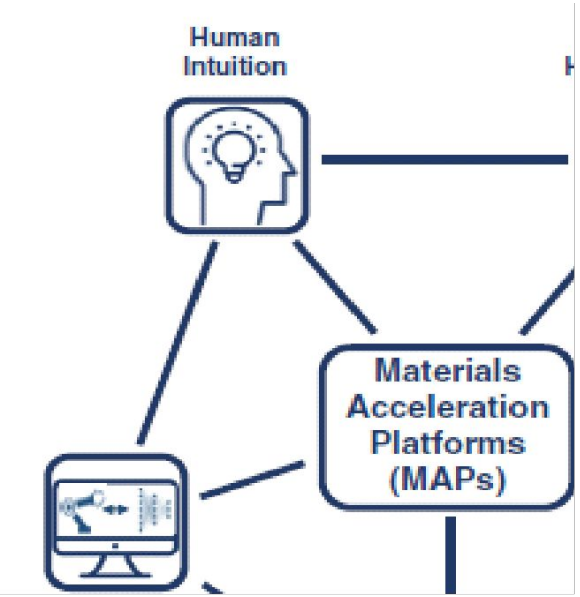


- Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie, SU
- Service de Génie Logiciel pour la Simulation, CEA
- Section de Recherches de Métallurgie Physique, CEA
- Département d'Instrumentation Numérique, CEA
- Département des sciences de la simulation et de l'information, CEA

- Institut Lumière Matière, UCBL
- Laboratoire de Science et Ingénierie des Matériaux et Procédés, INP-UGA
- Grenoble Alpes Recherche Infrastructure de Calcul Intensif et de Données, UGA
- Laboratoire de Modélisation et Exploration des Matériaux, CEA

Context and positioning

Computing



WP1

WP2

Data

- Infrastructure project
- HPC issues (GPU porting, …) addressed in PEPR NUMPEX
- No computing resources provided (see GENCI)

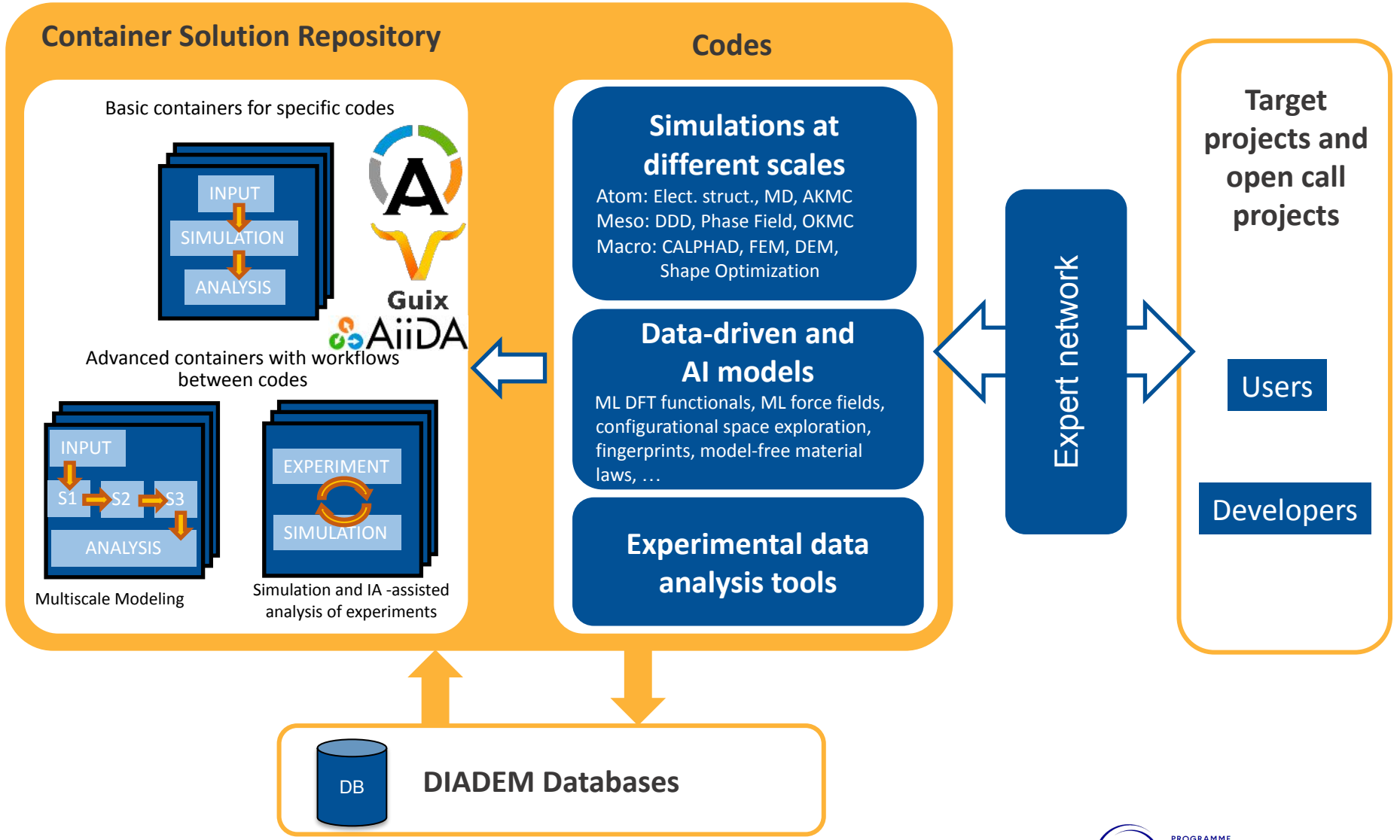
Role of Artificial Intelligence (AI)

- Simulations augmented and accelerated by AI (WP1)
- Integrated in workflows (WP1)
- AI toolbox for data analysis (WP2)
- At the heart of WP3

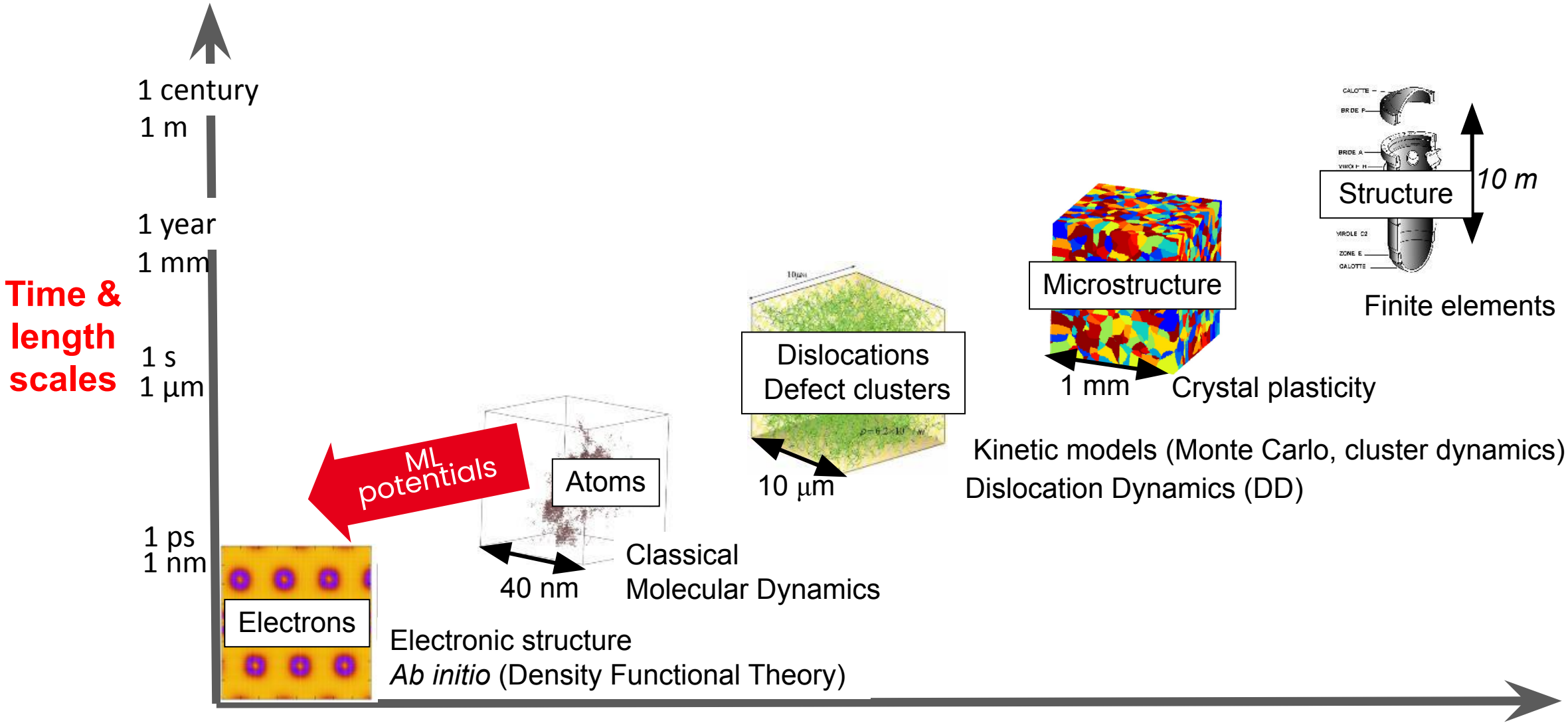
DIAMOND-CW : code and workflow platform

- **Containers:** making it easier to install, distribute and use code
- Container **repository**
- **Workflows:** enable code chaining, automation and therefore high throughput, including analysis of experimental data

Meso-center : GRICAD (Grenoble)



Multiscale modeling of materials augmented by AI



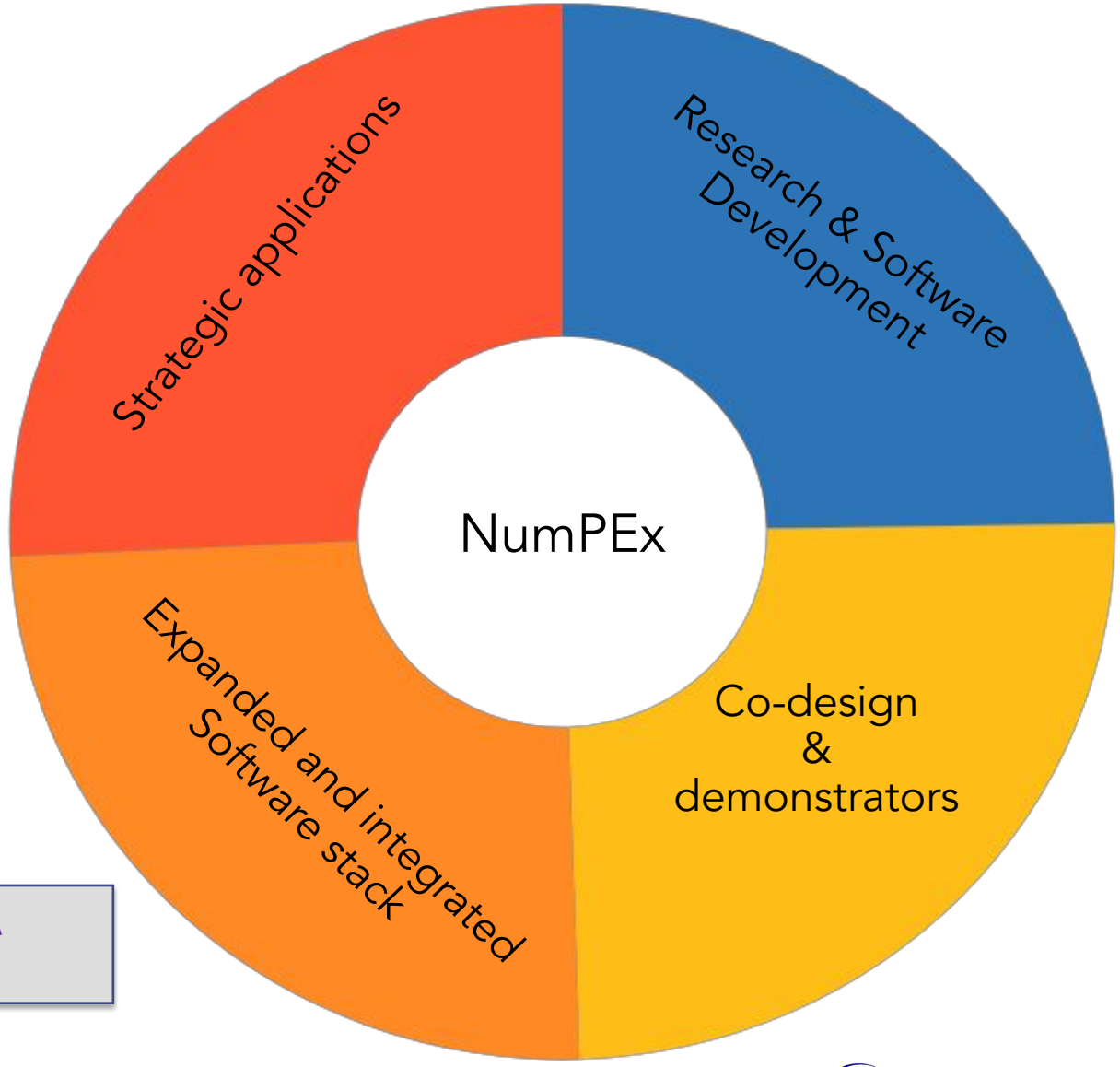
Link with NUMPEX PEPR (digital for exascale)

Contribute and accelerate the emergence of **a European sovereign exascale software stack** and **strategic applications exascale capability** in a **coherent and multi-annual framework**

Integrate and validate **co-designed** innovative methods, libraries and software stack with demonstrators of strategic applications.

Accelerate science-driven and engineering-driven developers **training and software productivity**

Aggregate the French HPC/HPDA/IA community



NUMPEX PEPR – software stack for the new exascale computer

Applications

Software integration and applications

Exa-AtoW
Continuum digital (workflow)

Exa-Dost
Data

Exa-MA
Numerical methods and solvers

Exa-Soft
Applications

NUMPEX: Exa-AToW – workflows at exascale

- Diamond is a use case for Exa-AToW
- Interoperability between workflow engines (over 100!)

Key objective: machine accessible data plan

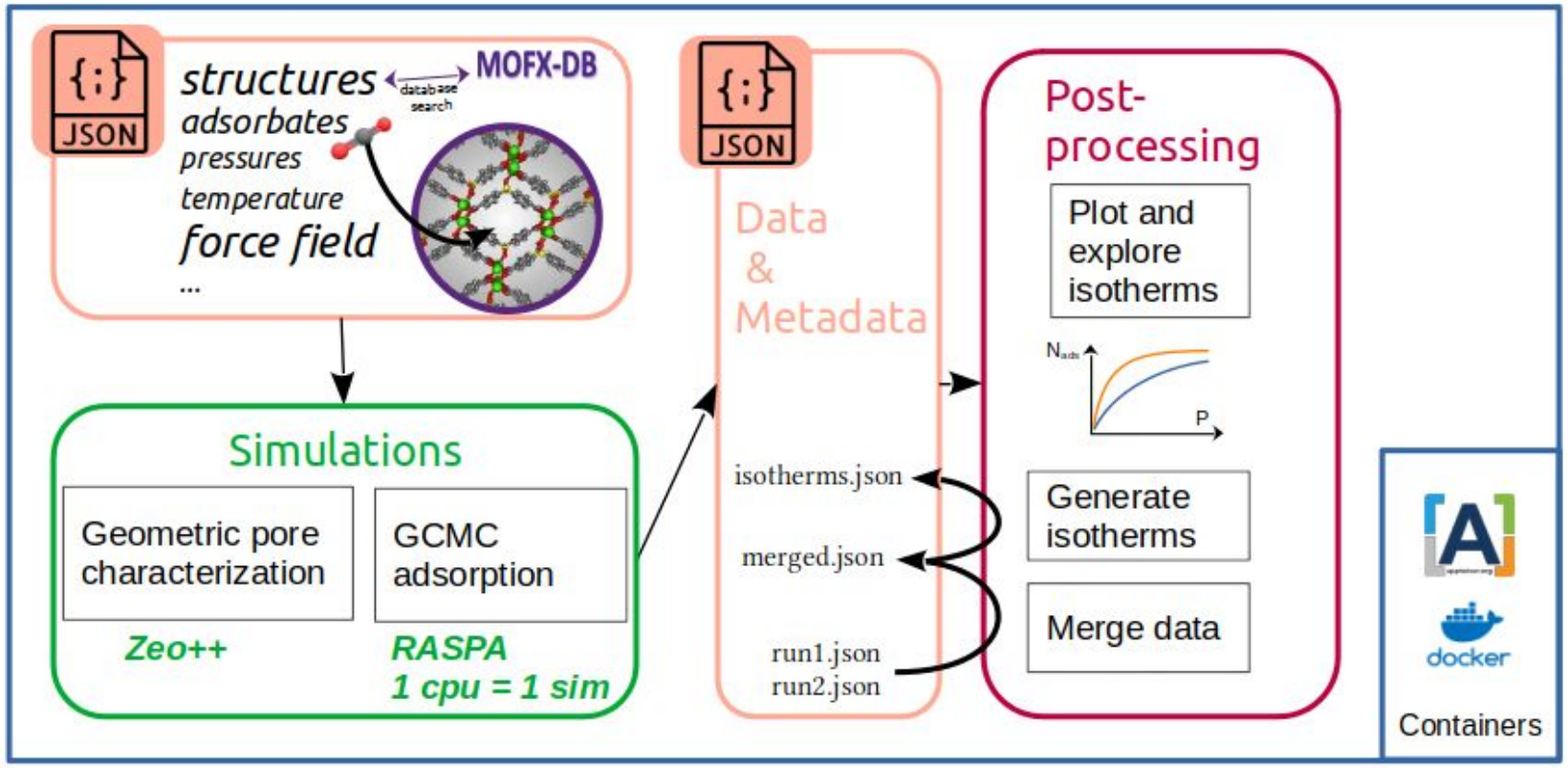
- Everything must be able to be processed by the machine, especially workflow task graphs information (metadata) to know what to do with the data and how to manipulate it

simple-adsorption-workflow (MOFLearning)

MOF = metal-organic framework

Workflow:

- Structure databases: MOFX-DB
- Calculation of gas adsorption energies with Zeo++ and RASPA
- Data generated: adsorption isotherms (number of gas molecules adsorbed as a function of pressure)



Arthur Hardiagon (IRCP, Chimie ParisTech)

DIAMOND-DM: Data management infrastructure (WP2)

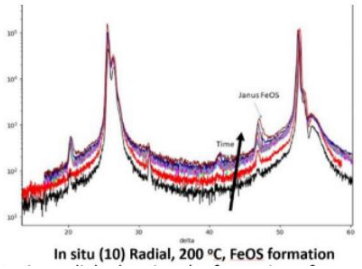
- Numerous 'producers' of experimental data: chemical and structural characterisation (MET, SAT, DRX, SAXS, Raman, LIBS, IBA, etc.), mechanics, optics,

- Some communities are very advanced when it comes to managing their data (synchrotrons).

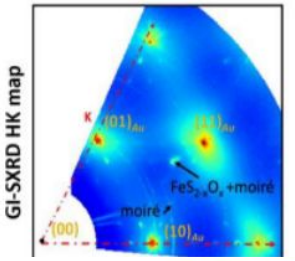
- The aim is to develop tools to help the less advanced communities build their databases.

INS2: M. Prabhu (post-Doc), G. Renaud (CEA/IRIG/MEM), J. Coraux (Inst.Néel)

FeS_2/Au(111)

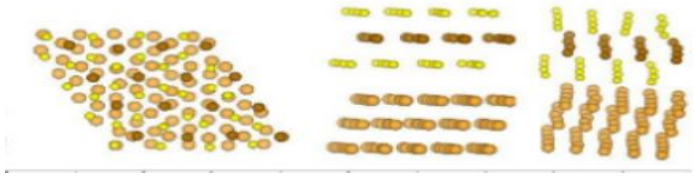


In situ growth



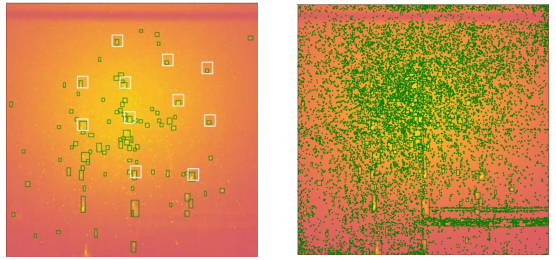
Reci. Space Map

=> Atomic structure model
up to 3-4 monolayer in substrate

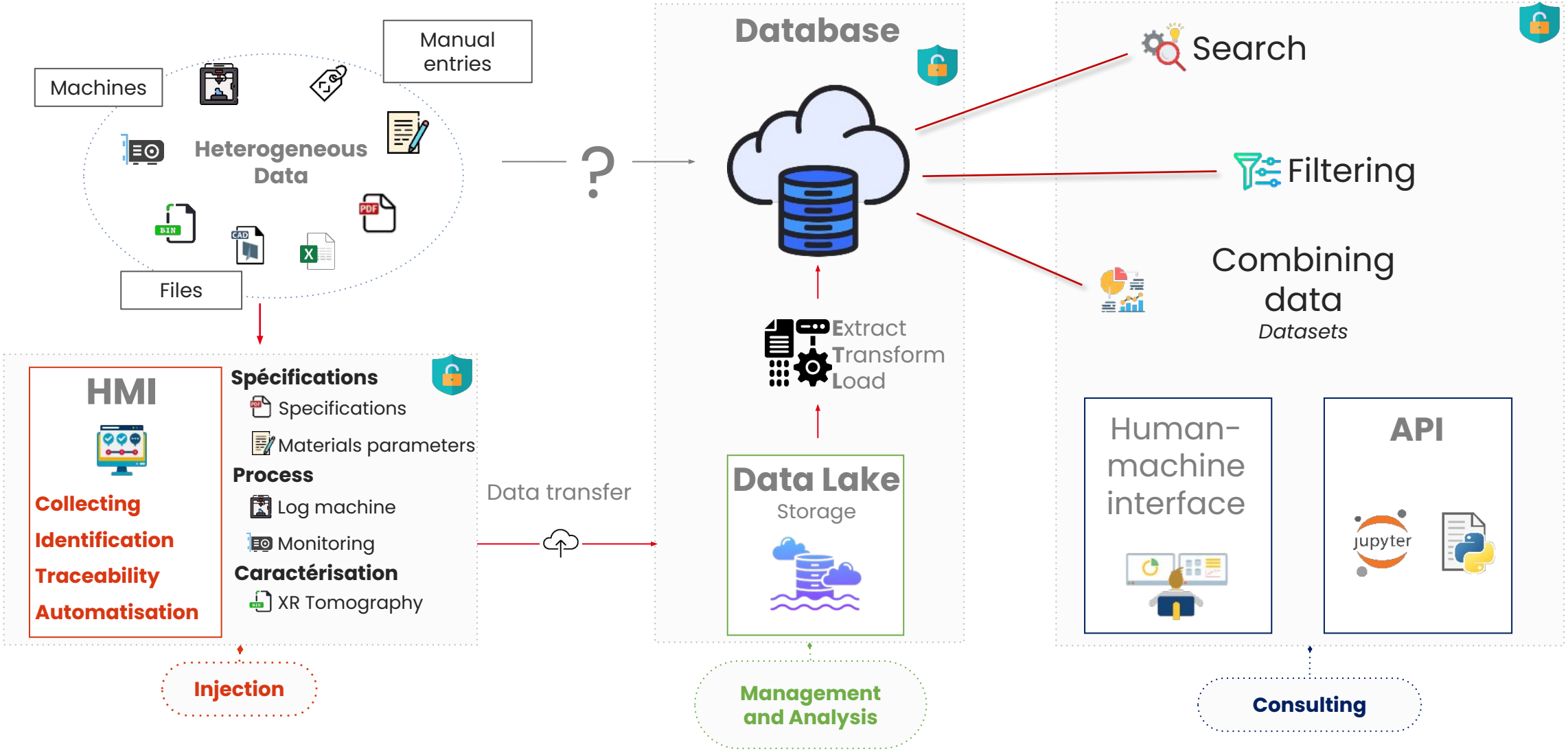


LAUE: LaueImProc (R. Richard, DIADEM)

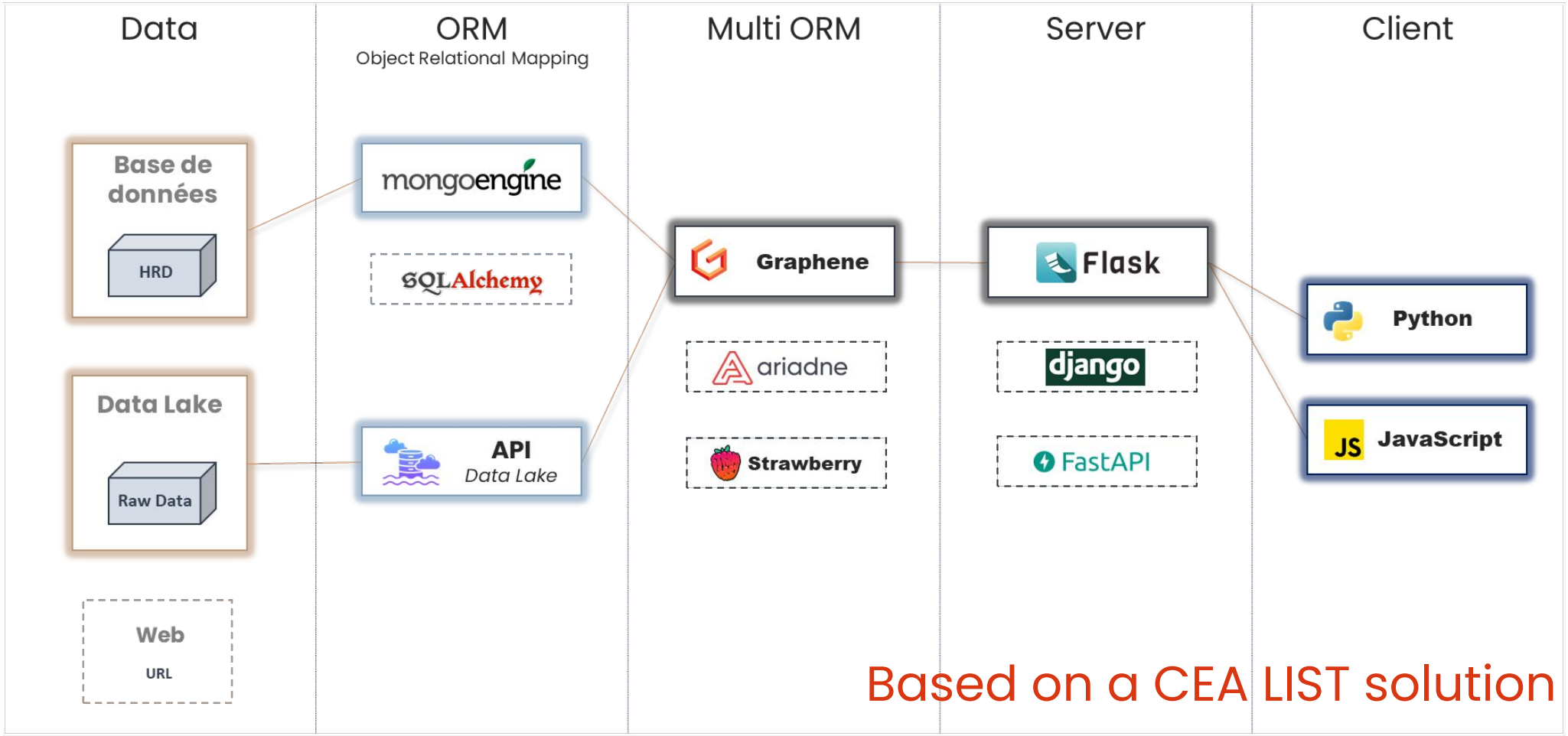
workflow for 2D cartography
// calculation, pytorch, GPU,
Laue peak filtering
notebooks @ESRF or @laptop



Diamond-DB: Architecture

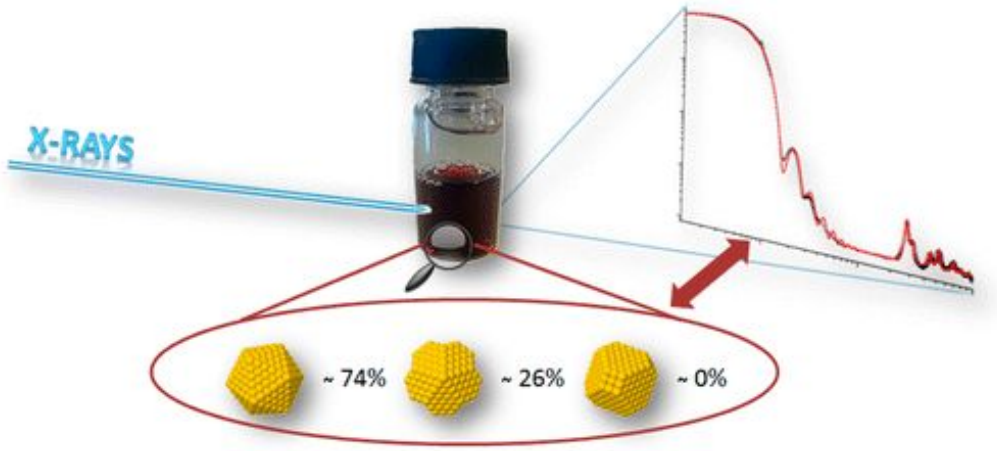


Software layer integration



DIAMOND-DB: Two planned use cases

- FastNano
  - Microfluidic synthesis of nano materials and in situ characterization via Small Angle X-Ray Scattering (SAXS)
  - Real time control of the manufacturing process assisted by machine learning



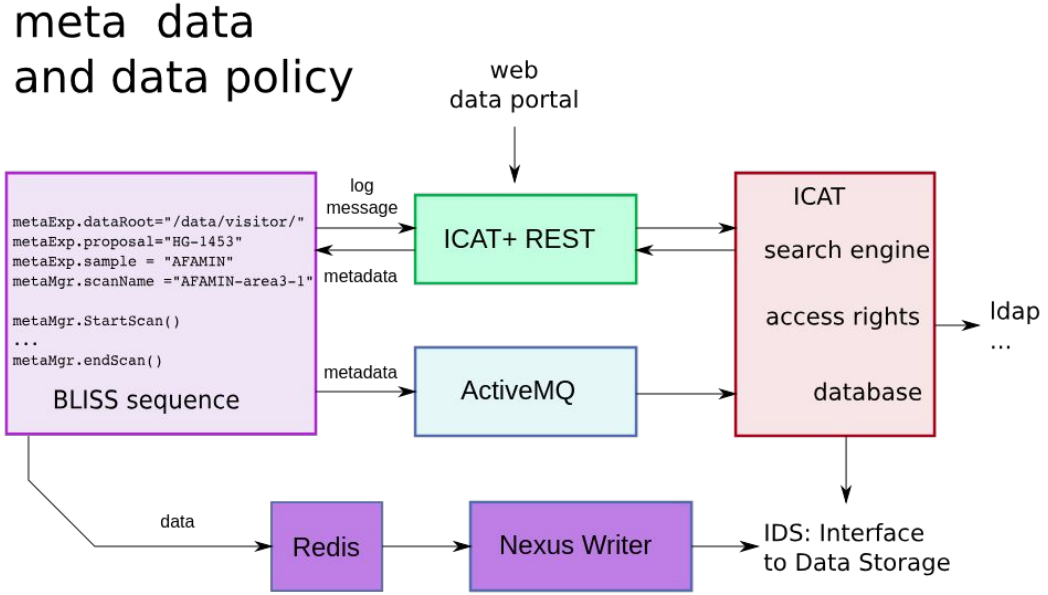
- Atomistic simulations
  - DFT and classical MD simulations
  - Optimade

OPTIMADE
Open Databases Integration for Materials Design

Diamond-DB : Links with demonstration projects

DIADEM-ESRF, DIADEM-SOLEIL, DIADEM-METSA-SETDIA

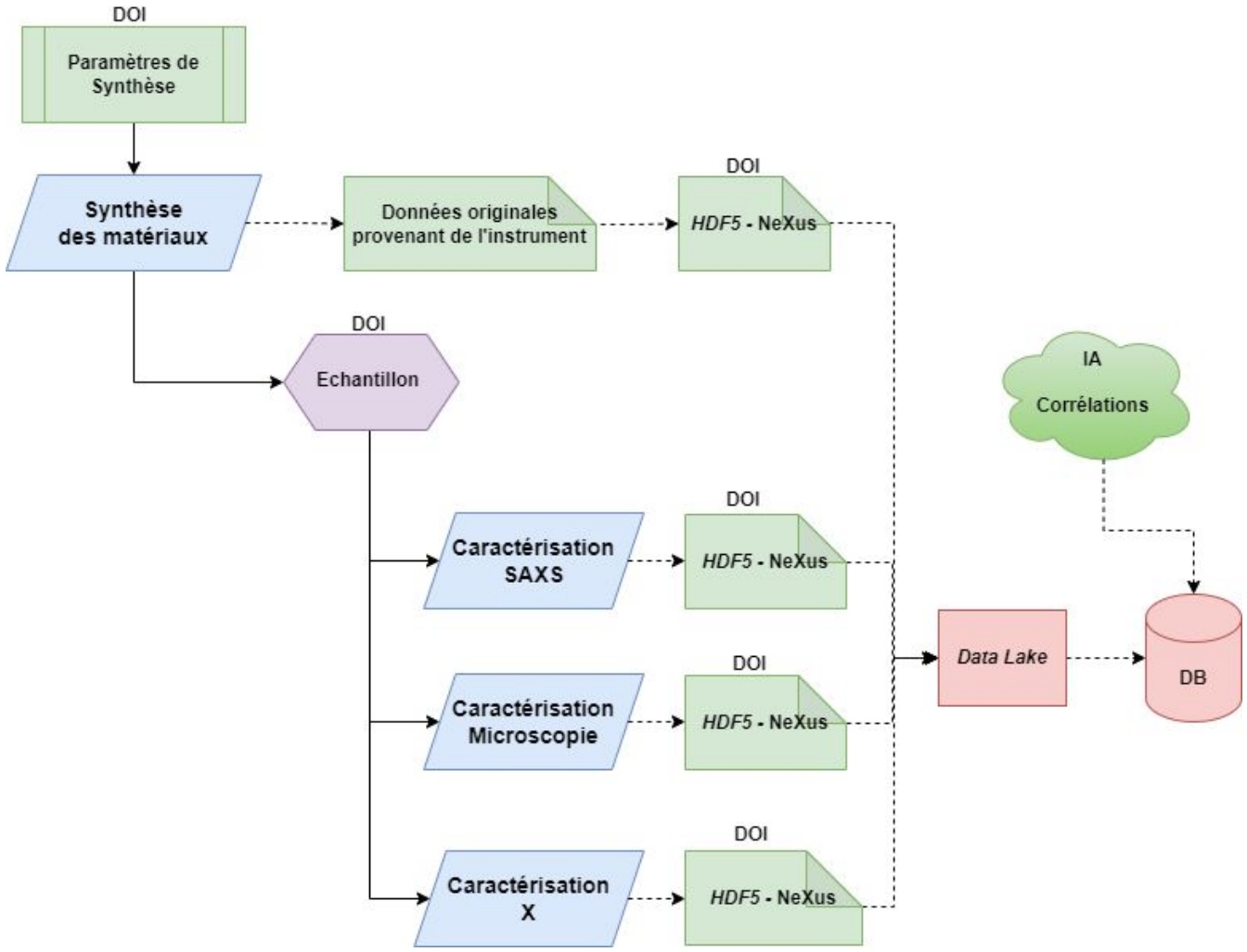
- Several discussions on requirements,
- Existing standard format based on NEXUS (HDF5)
- Use as a use case to define BoD specifications



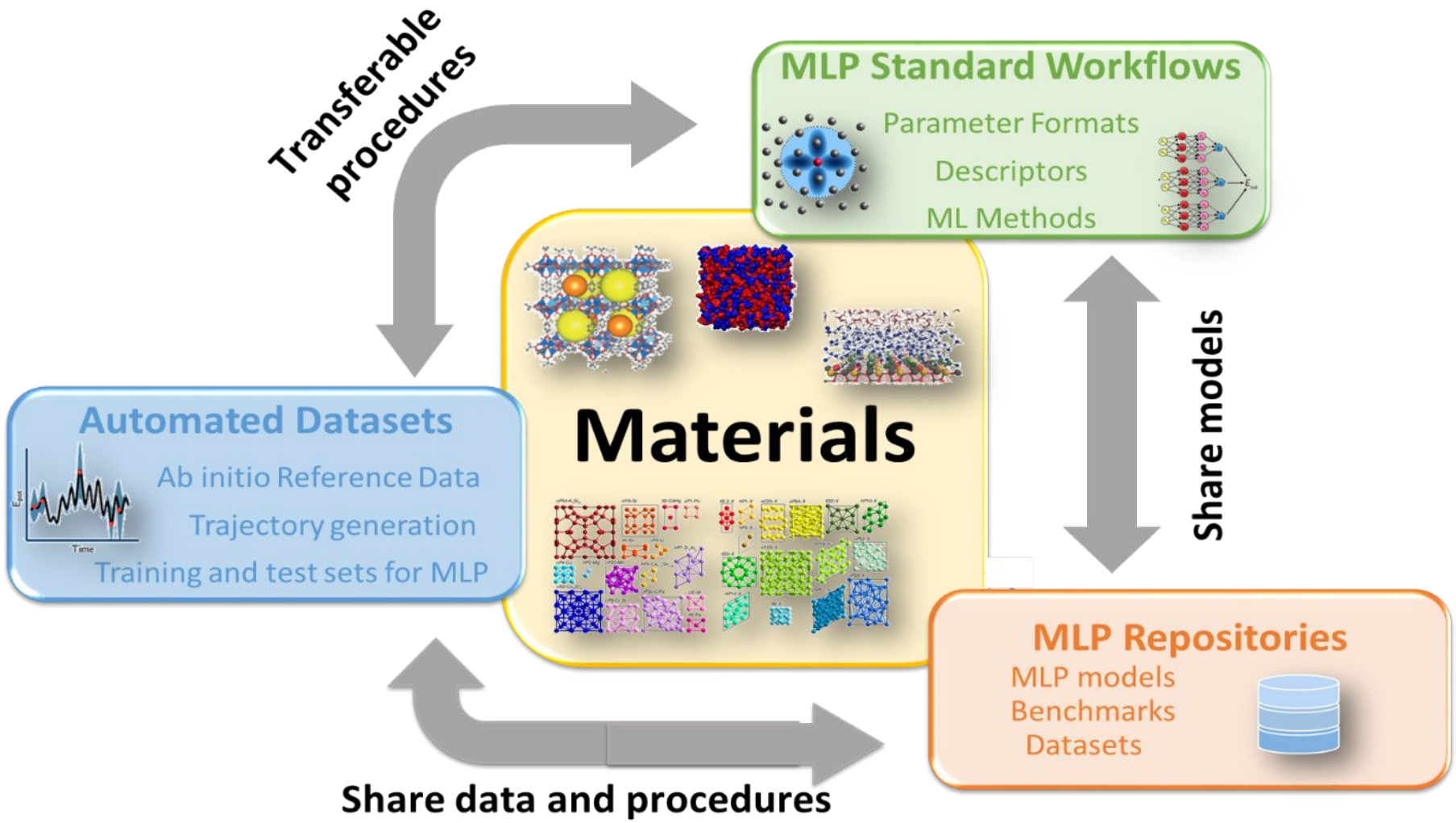
Diamond-DB: Link with demonstration projects

FAST-NANO et 2FAST

- Several discussions on requirements,
- Need to define a file standard
- (format but above all keywords)
- Need to manage raw data
- (data lake) - outside the Diamond scope



Machine Learning potential demonstrator (WP3)

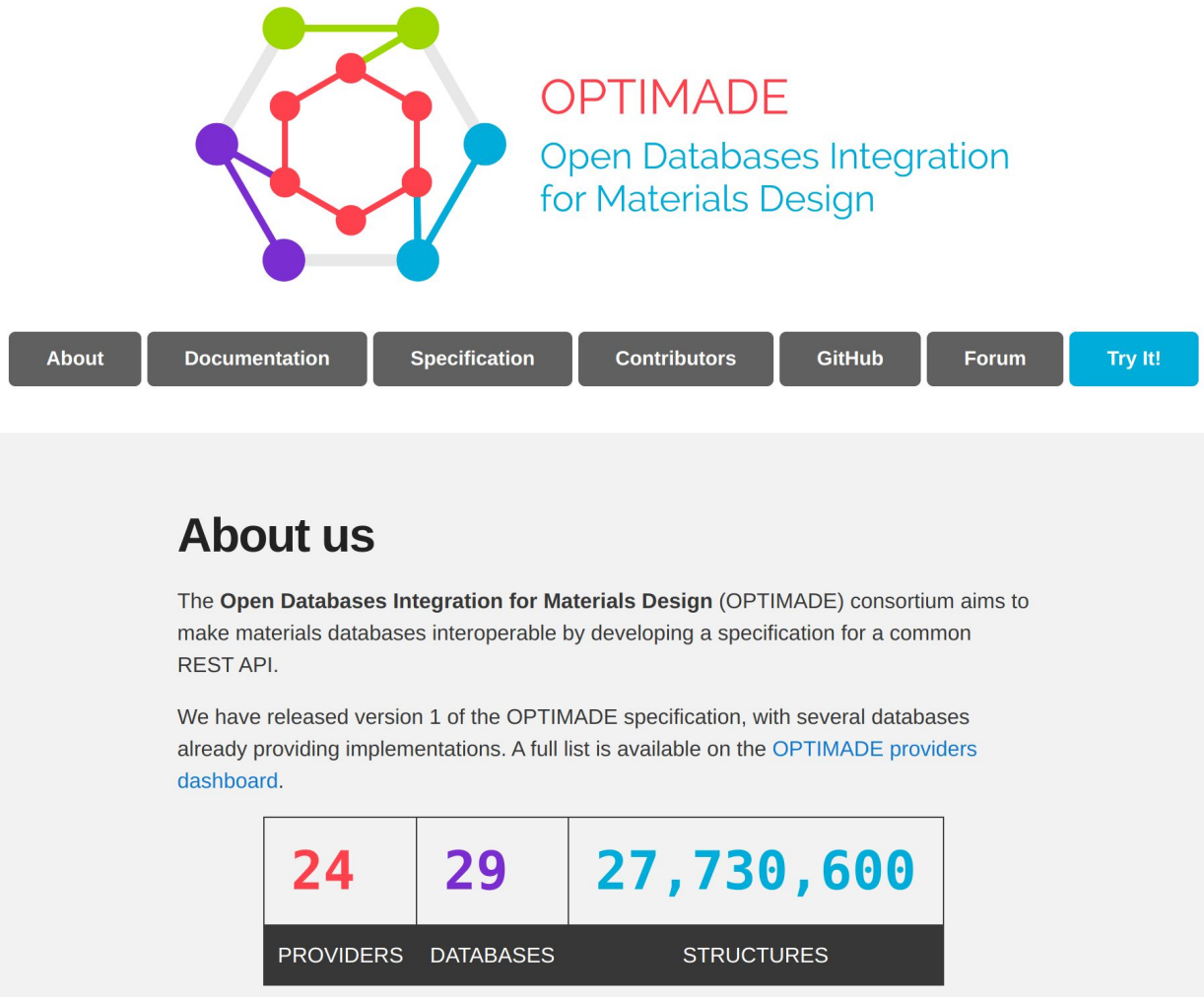


Marco Saitta, IMPMC, SU, Paris

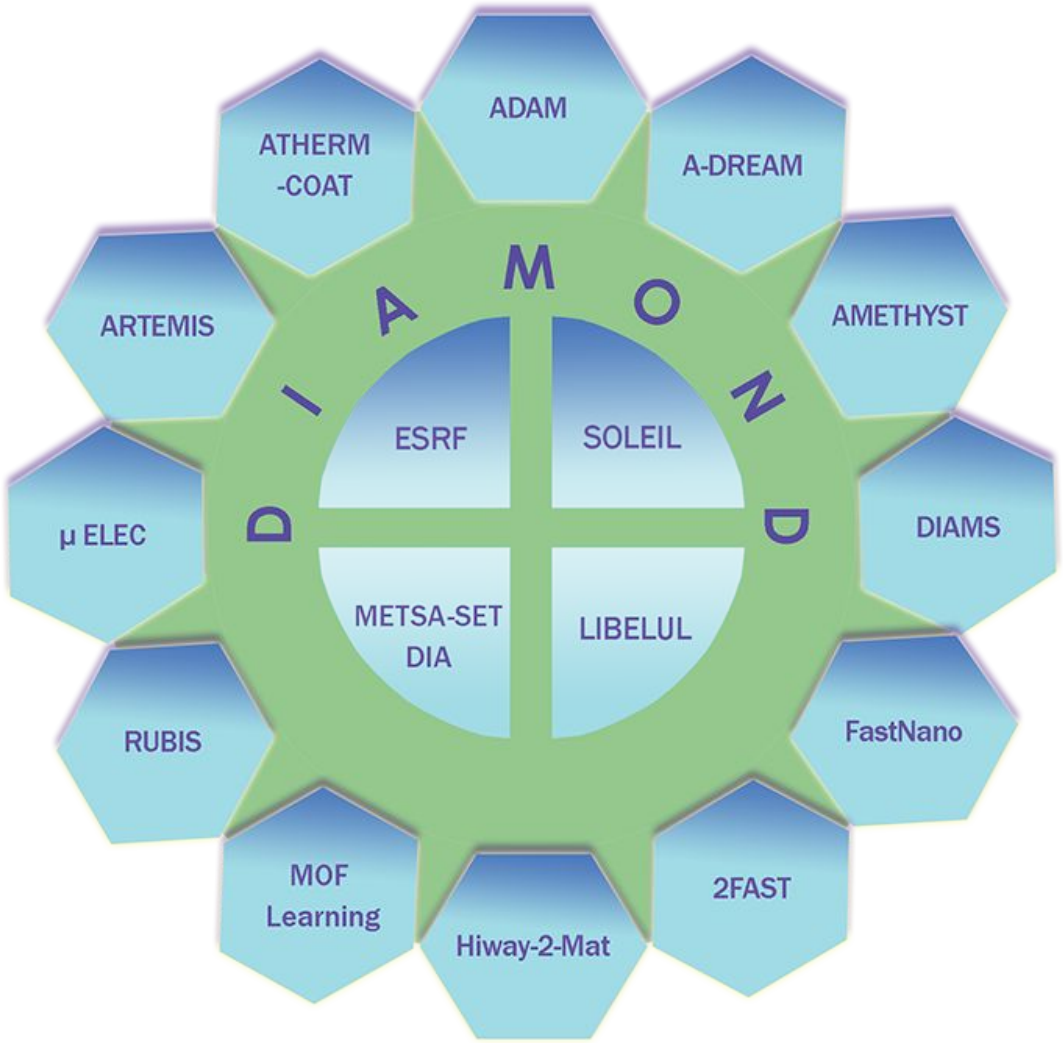
T2.2 Collect of the pre-existing ontology for simulation to specify the meta-ontology

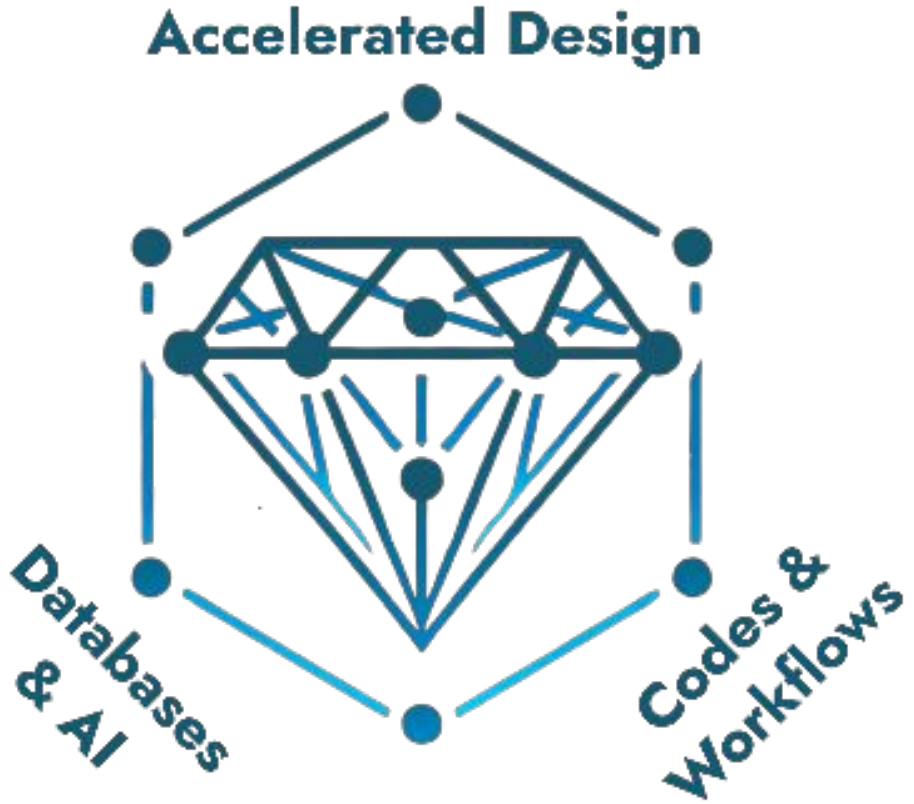
- Definition of an ontology, in conjunction with the OPTIMADE tool, for molecular dynamics-oriented simulation data

- API for querying and exploring open databases



Conclusion – Materials Acceleration Platforms





Merci