NOMAD
NOVEL MATERIALS DISCOVERY

FAIR-DI

Metadata towards FAIR data sharing for data-driven materials science: achievements and open challenges

Luca M. Ghiringhelli
FRITZ-HABER-INSTITUT
MAX-PLANCK-GESELLSCHAFT

Big Data Summer
A summer school of the BiGmax Network
Platja d'Aro, Spain, September 9 - 13, 2019

For data-driven (materials) science...

NOVEL MATERIALS DISCOVERY

Mark D. Wilkinson, Michel Dumontier, *et al.*
The FAIR Guiding Principles for scientific data management and stewardship
Scientific Data **3**, 160018 (2016)

For data-driven (materials) science...

NOVEL MATERIALS DISCOVERY

Mark D. Wilkinson, Michel Dumontier, *et al.*
The FAIR Guiding Principles for scientific data management and stewardship
Scientific Data **3**, 160018 (2016)

"Data" includes "algorithms, tools, and workflows that led to that data".

For data-driven (materials) science...

NOVEL MATERIALS DISCOVERY

Mark D. Wilkinson, Michel Dumontier, *et al.*
The FAIR Guiding Principles for scientific data management and stewardship
Scientific Data **3**, 160018 (2016)

Findable

For data-driven (materials) science...

NOVEL MATERIALS DISCOVERY

Mark D. Wilkinson, Michel Dumontier, *et al.*
The FAIR Guiding Principles for scientific data management and stewardship
Scientific Data **3**, 160018 (2016)

Findable Accessible

For data-driven (materials) science...

NOVEL MATERIALS DISCOVERY

Mark D. Wilkinson, Michel Dumontier, *et al.*
The FAIR Guiding Principles for scientific data management and stewardship
Scientific Data **3**, 160018 (2016)

Findable **A**ccessible **I**nteroperable

Fun facts on the three-gears contraption

NOVEL MATERIALS DISCOVERY

Fun facts on the three-gears contraption

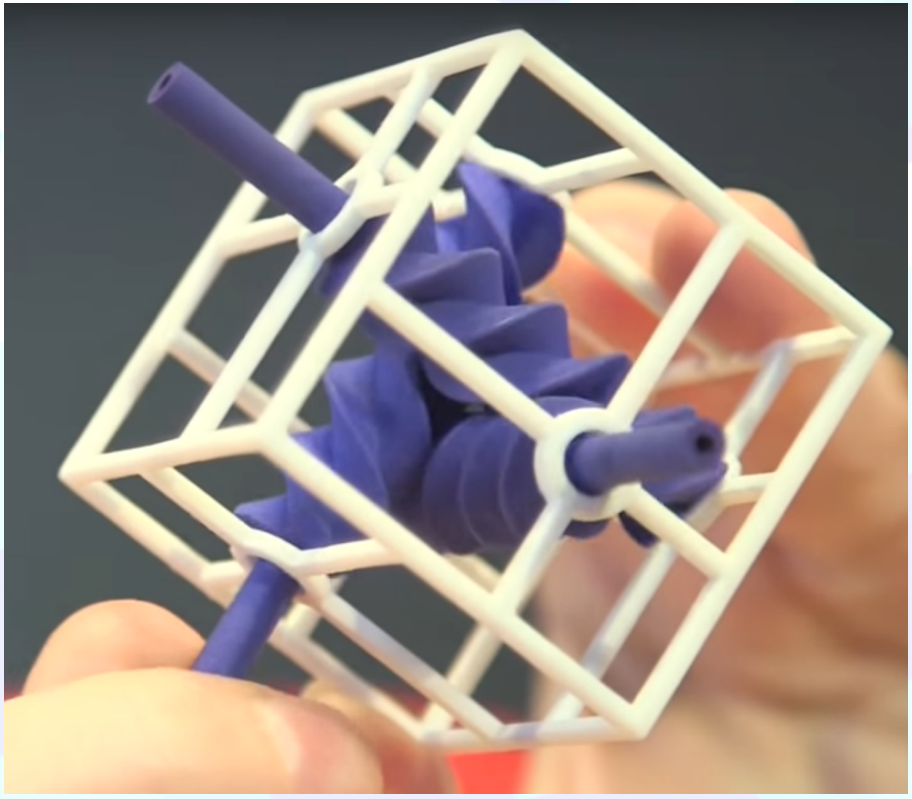
NOVEL MATERIALS DISCOVERY

Fun facts on the three-gears contraption

NOVEL MATERIALS DISCOVERY

Fun facts on the three-gears contraption

NOVEL MATERIALS DISCOVERY



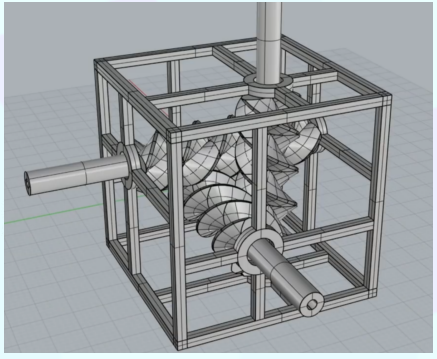
www.youtube.com/watch?v=5Mf0JpTI_gg

For data-driven (materials) science...

NOVEL MATERIALS DISCOVERY

Mark D. Wilkinson, Michel Dumontier, *et al.*
The FAIR Guiding Principles for scientific data management and stewardship
Scientific Data **3**, 160018 (2016)

Findable **A**ccessible **I**nteroperable **R**eusable

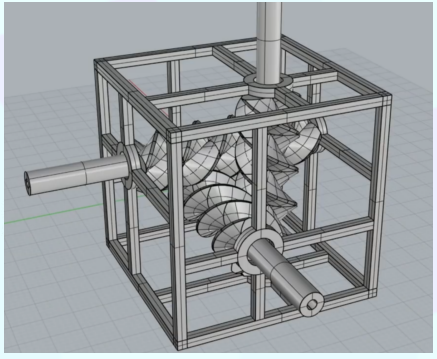


For data-driven (materials) science...

NOVEL MATERIALS DISCOVERY

Mark D. Wilkinson, Michel Dumontier, *et al.*
The FAIR Guiding Principles for scientific data management and stewardship
Scientific Data **3**, 160018 (2016)

Findable **A**ccessible **I**nteroperable **R**epurposable ecycable eusable



For data-driven (materials) science...

NOVEL MATERIALS DISCOVERY

Mark D. Wilkinson, Michel Dumontier, *et al.*
The FAIR Guiding Principles for scientific data management and stewardship
Scientific Data **3**, 160018 (2016)

Findable
Metadata are unique

Accessible
(Meta)data can be accessed somewhere

Interoperable
Useful ontologies are defined

Repurposable
ecycable
eusable
Metadata are "rich"

What are metadata?

NOVEL MATERIALS DISCOVERY

Working group at the
NOMAD-FAIRDI workshop: “Shared metadata and data formats for Big-Data Driven Materials Science.” Berlin, July 2019.

Participants:
(d*ata scientists)* Javad Chamanara, Patrick Lambrix, Tatyana Sheveleva ,
(*materials scientists*) Carsten Baldauf, Stefano Cozzini, Christoph Koch, Astrid Schneidewind, Christof Wöll.

What are metadata?

NOVEL MATERIALS DISCOVERY

Working group at the
NOMAD-FAIRDI workshop: “Shared metadata and data formats for Big-Data Driven Materials Science.” Berlin, July 2019.

Participants:
(d*ata scientists)* Javad Chamanara, Patrick Lambrix, Tatyana Sheveleva ,
(*materials scientists*) Carsten Baldauf, Stefano Cozzini, Christoph Koch, Astrid Schneidewind, Christof Wöll.

Metadata are data about data.

What are metadata?

NOVEL MATERIALS DISCOVERY

Working group at the
NOMAD-FAIRDI workshop: “Shared metadata and data formats for Big-Data Driven Materials Science.” Berlin, July 2019.

Participants:
(d*ata scientists)* Javad Chamanara, Patrick Lambrix, Tatyana Sheveleva ,
(*materials scientists*) Carsten Baldauf, Stefano Cozzini, Christoph Koch, Astrid Schneidewind, Christof Wöll.

Metadata are data about data.
Data object (information resource): a row in the data table.

UID	Structure	Method	Total energy
31415	Graphite.xyz	DFT, PBE +TS	-2718281.828 eV

Columns are attributes of the data objects.

NOMAD

NOVEL MATERIALS DISCOVERY

What are metadata?

UID	Structure	Method	Total energy	New structure
31415	Graphite.xyz	DFT, PBE +TS	-2718281.828 eV	Graphite_2.xyz

Columns are attributes of the data objects.
These attributes are **data** or **metadata** depending on **context**.
Description: who, when, where.
Provenance: how
Administrative: location, access privileges

What are metadata?

NOVEL MATERIALS DISCOVERY

UID	Structure	Method	Total energy	New structure
31415	Graphite.xyz	DFT, PBE +TS	-2718281.828 eV	Graphite_2.xyz

Columns are attributes of the data objects.
These attributes are **data** or **metadata** depending on **context**.
Description: who, when, where.
Provenance: how
Administrative: location, access privileges

Definition: The attributes that are necessary to locate, fully characterize, and – ultimately – **reproduce** other attributes that are identified as data. The metadata include a clear and unambiguous description of the data, and their full provenance.

The NOMAD Laboratory

A European Centre of Excellence

PROJECT INDUSTRY TEAM RELATED PROJECTS NEWS PRESS KIT CONTACT US

Enter Search...

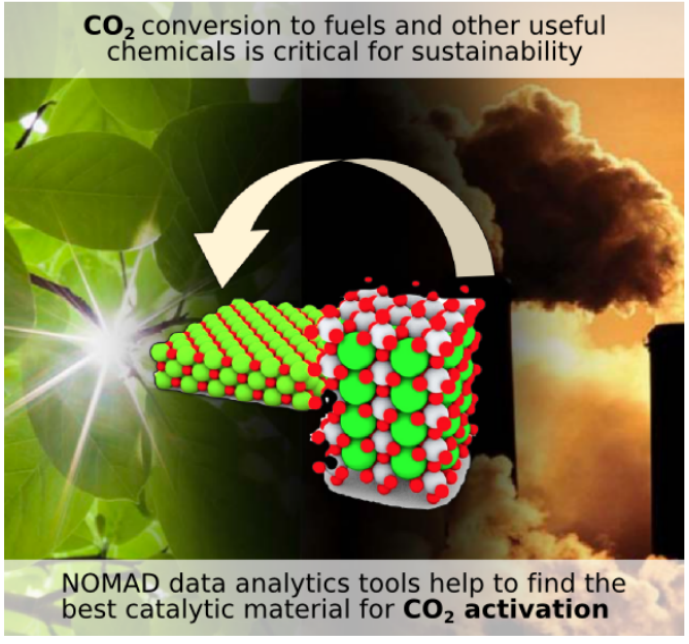
NOMAD REPOSITORY | THE ARCHIVE | ENCYCLOPEDIA | BIG-DATA ANALYTICS | ADVANCED GRAPHICS | HPC INFRASTRUCTURE | OUTREACH

https://nomad-coe.eu/
NOvel **MA**terials **D**iscovery Laboratory

NOMAD Scope and Overview

Data is a crucial raw material of the 21st century.

NOMAD Success Stories



The NOMAD Laboratory

NOVEL MATERIALS DISCOVERY

The NOMAD Laboratory

NOVEL MATERIALS DISCOVERY

Live since 12.2014

NOMAD REPOSITORY

The NOMAD Laboratory

NOVEL MATERIALS DISCOVERY

Live since 12.2014

NOMAD REPOSITORY

CONVERSION LAYER

The NOMAD Laboratory

NOVEL MATERIALS DISCOVERY

Live since 12.2014

NOMAD REPOSITORY

CONVERSION LAYER

THE ARCHIVE

The NOMAD Laboratory

NOVEL MATERIALS DISCOVERY

Live since 12.2014

NOMAD REPOSITORY

CONVERSION LAYER

THE ARCHIVE

ENCYCLOPEDIA

ADVANCED GRAPHICS

BIG-DATA ANALYTICS

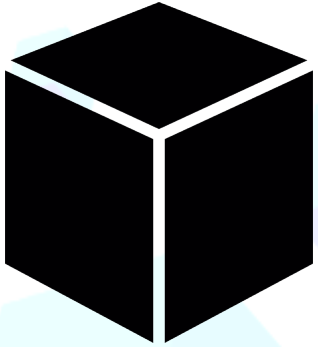
Computational materials science: basics

NOVEL MATERIALS DISCOVERY

Computational materials science: basics

NOVEL MATERIALS DISCOVERY

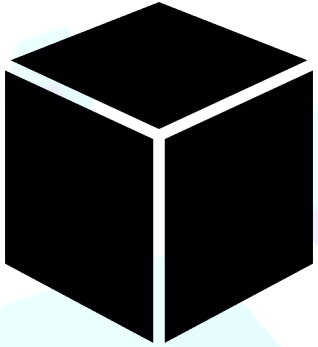
Atomistic-simulation code



Computational materials science: basics

NOVEL MATERIALS DISCOVERY

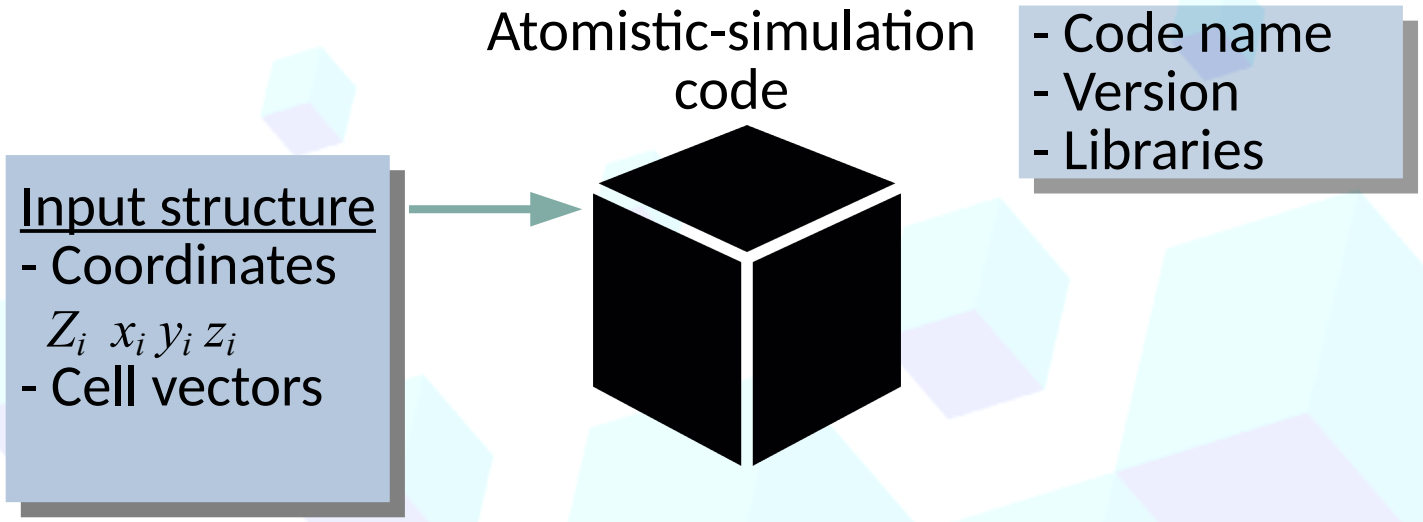
Atomistic-simulation code



- Code name
- Version
- Libraries

Computational materials science: basics

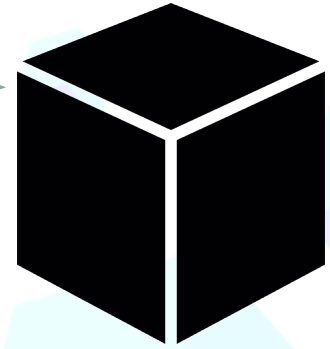
NOVEL MATERIALS DISCOVERY



Computational materials science: basics

NOVEL MATERIALS DISCOVERY

Atomistic-simulation code



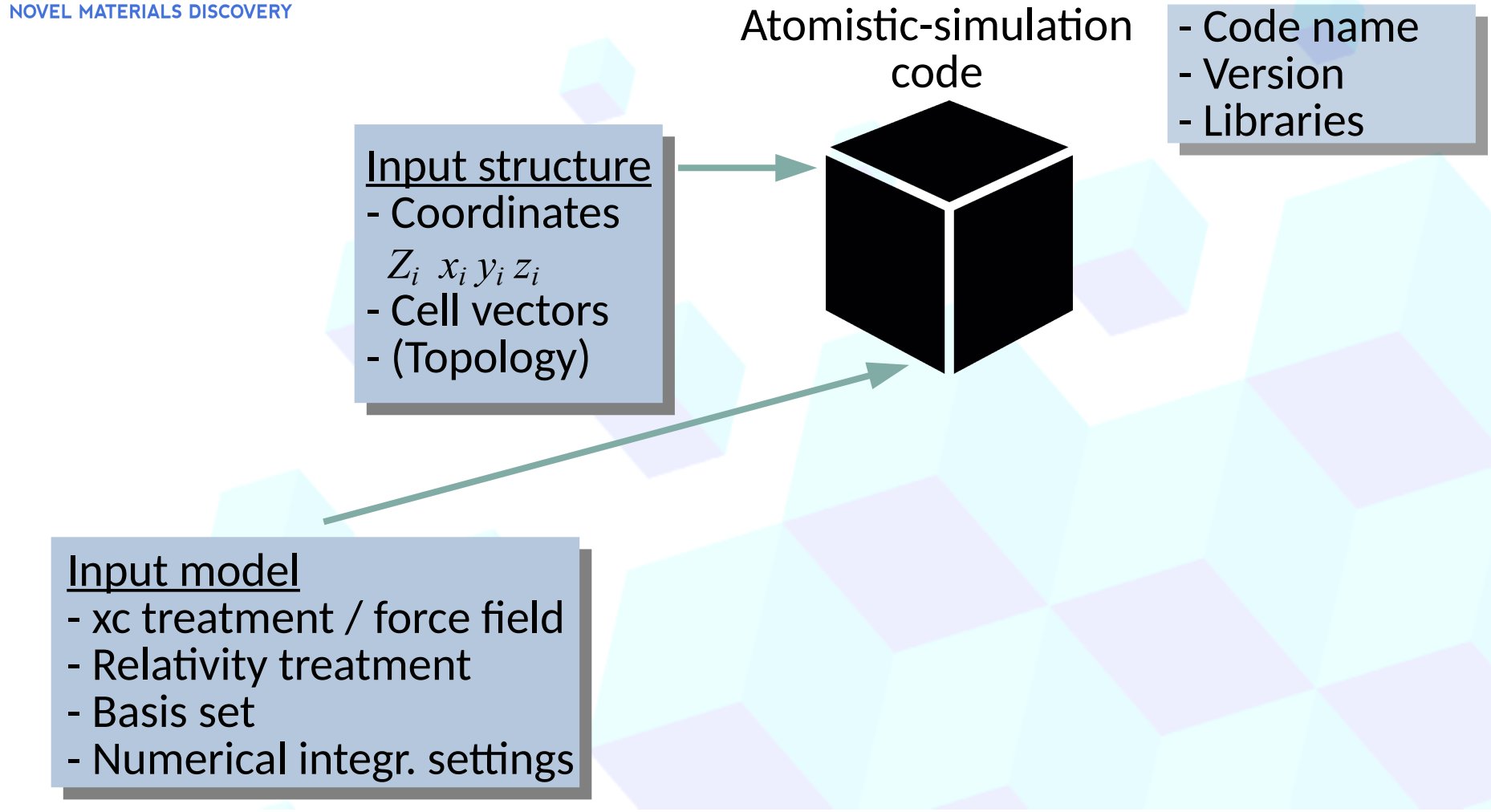
- Code name
- Version
- Libraries

Input structure
- Coordinates

  $Z_i \;\; x_i \, y_i \, z_i$
- Cell vectors
- (Topology)

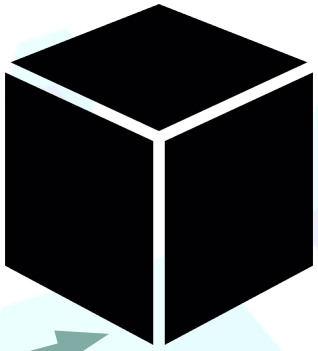
NOMAD
NOVEL MATERIALS DISCOVERY

Computational materials science: basics



Computational materials science: basics

NOMAD
NOVEL MATERIALS DISCOVERY



- Code name
- Version
- Libraries

Input structure
- Coordinates
  $Z_i \;\; x_i \, y_i \, z_i$
- Cell vectors
- (Topology)

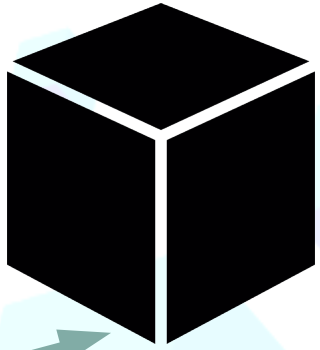
Output
- Total energy
- Forces
- ...

Input model
- xc treatment / force field
- Relativity treatment
- Basis set
- Numerical integr. settings

Computational materials science: basics

NOMAD
NOVEL MATERIALS DISCOVERY

Atomistic-simulation code



- Code name
- Version
- Libraries

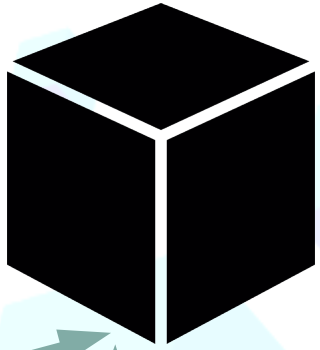
Input structure
- Coordinates
  $Z_i \;\; x_i \, y_i \, z_i$
- Cell vectors
- (Topology)

Output
- Total energy
- Forces
- Electron density
- Electrostatic pot.
- El. band structure
- Self energy
- …

Input model
- xc treatment / force field
- Relativity treatment
- Basis set
- Numerical integr. settings

Computational materials science: basics

NOMAD
NOVEL MATERIALS DISCOVERY



- Code name
- Version
- Libraries

Input structure
- Coordinates
  $Z_i \; x_i \, y_i \, z_i$
- Cell vectors
- (Topology)

Output
- Total energy
- Forces
- Electron density
- Electrostatic pot.
- El. band structure
- Self energy
- …

Input model
- xc treatment / force field
- Relativity treatment
- Basis set
- Numerical integr. settings

More complex input model
- Model_1 (scf)→ Model_2 (perturbative)
- QM/MM (space partitioning of models)
- …

Computational materials science: basics

NOMAD
NOVEL MATERIALS DISCOVERY

- Code name
- Version
- Libraries

Input structure
- Coordinates
- Cell vectors
- (Topology)

Input model
- xc treatment / force field
- Relativity treatment
- Basis set
- Numerical integr. settings

Output
- Total energy
- Forces
- Electron density
- Electrostatic potential
- Electronic band structure
- Self energy

Our solution: NOMAD MetaInfo

NOVEL MATERIALS DISCOVERY

Metadata: in general dictionaries of key:value pairs, e.g.:

UID: 3141593

Structure: Graphite128_5775defect_final.xyz

Method: DFT, PBE, TS correction, NAO basis set, tier 1, light194 settings

Our solution: NOMAD MetaInfo

NOVEL MATERIALS DISCOVERY

Metadata: in general dictionaries of key:value pairs, e.g.:

UID: 3141593
Structure: Graphite128_5775defect_final.xyz
Method: DFT, PBE, TS correction, NAO basis set, tier 1, light194 settings

NOMAD MetaInfo, basic elements:

Key:

type: section
name [section_run, section_method, ...]
description

Our solution: NOMAD MetaInfo

NOVEL MATERIALS DISCOVERY

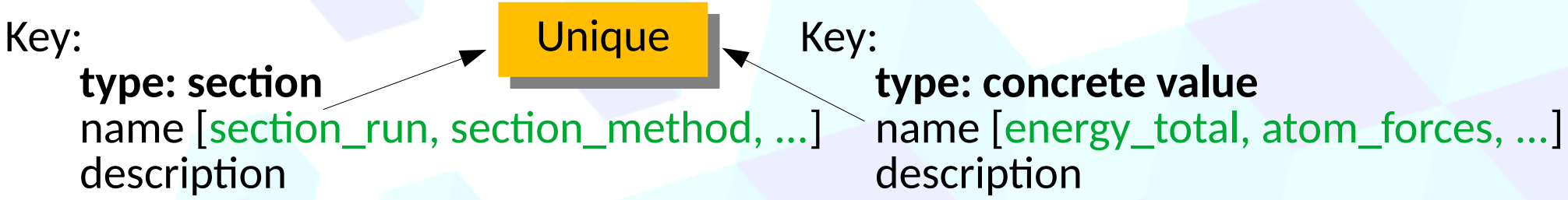
Metadata: in general dictionaries of key:value pairs, e.g.:

UID: 3141593

Structure: Graphite128_5775defect_final.xyz

Method: DFT, PBE, TS correction, NAO basis set, tier 1, light194 settings

NOMAD MetaInfo, basic elements:

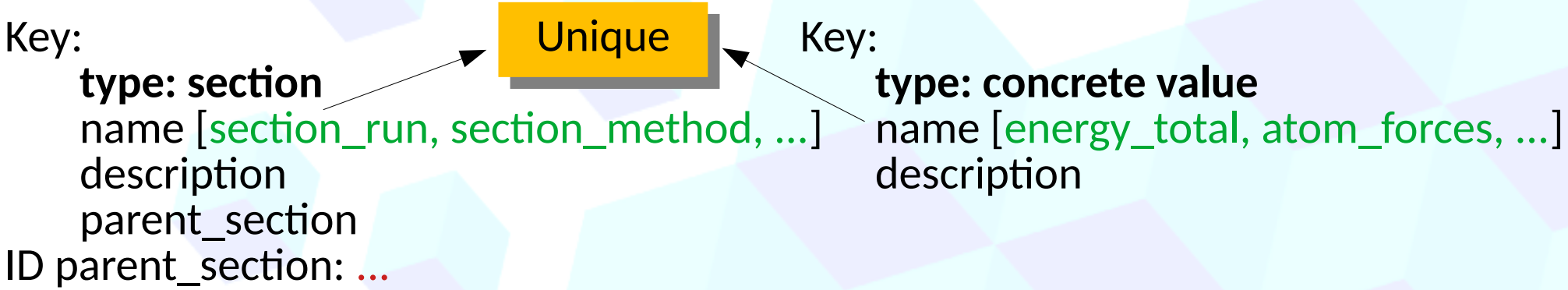


Our solution: NOMAD MetaInfo

NOVEL MATERIALS DISCOVERY

Metadata: in general dictionaries of key:value pairs, e.g.:
UID: 3141593
Structure: Graphite128_5775defect_final.xyz
Method: DFT, PBE, TS correction, NAO basis set, tier 1, light194 settings

NOMAD MetaInfo, basic elements:

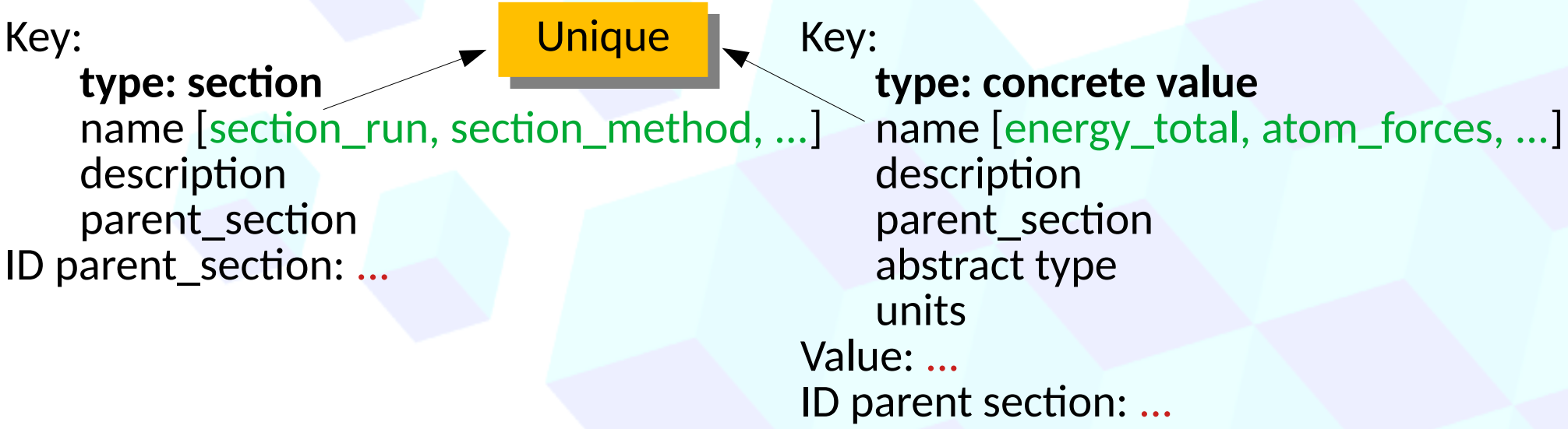


Our solution: NOMAD MetaInfo

NOVEL MATERIALS DISCOVERY

Metadata: in general dictionaries of key:value pairs, e.g.:
UID: 3141593
Structure: Graphite128_5775defect_final.xyz
Method: DFT, PBE, TS correction, NAO basis set, tier 1, light194 settings

NOMAD MetaInfo, basic elements:



Our solution: NOMAD MetaInfo

NOVEL MATERIALS DISCOVERY

Metadata: in general dictionaries of key:value pairs, e.g.:

UID: 3141593

Structure: Graphite128_5775defect_final.xyz

Method: DFT, PBE, TS correction, NAO basis set, tier 1, light194 settings

NOMAD MetaInfo, basic elements:

Key:

type: section

name [section_run, section_method, ...]

description

parent_section

ID parent_section: ...

Unique

Key:

type: concrete value

name [energy_total, atom_forces, ...]

description

parent_section

abstract type

units → SI units

Value: ...

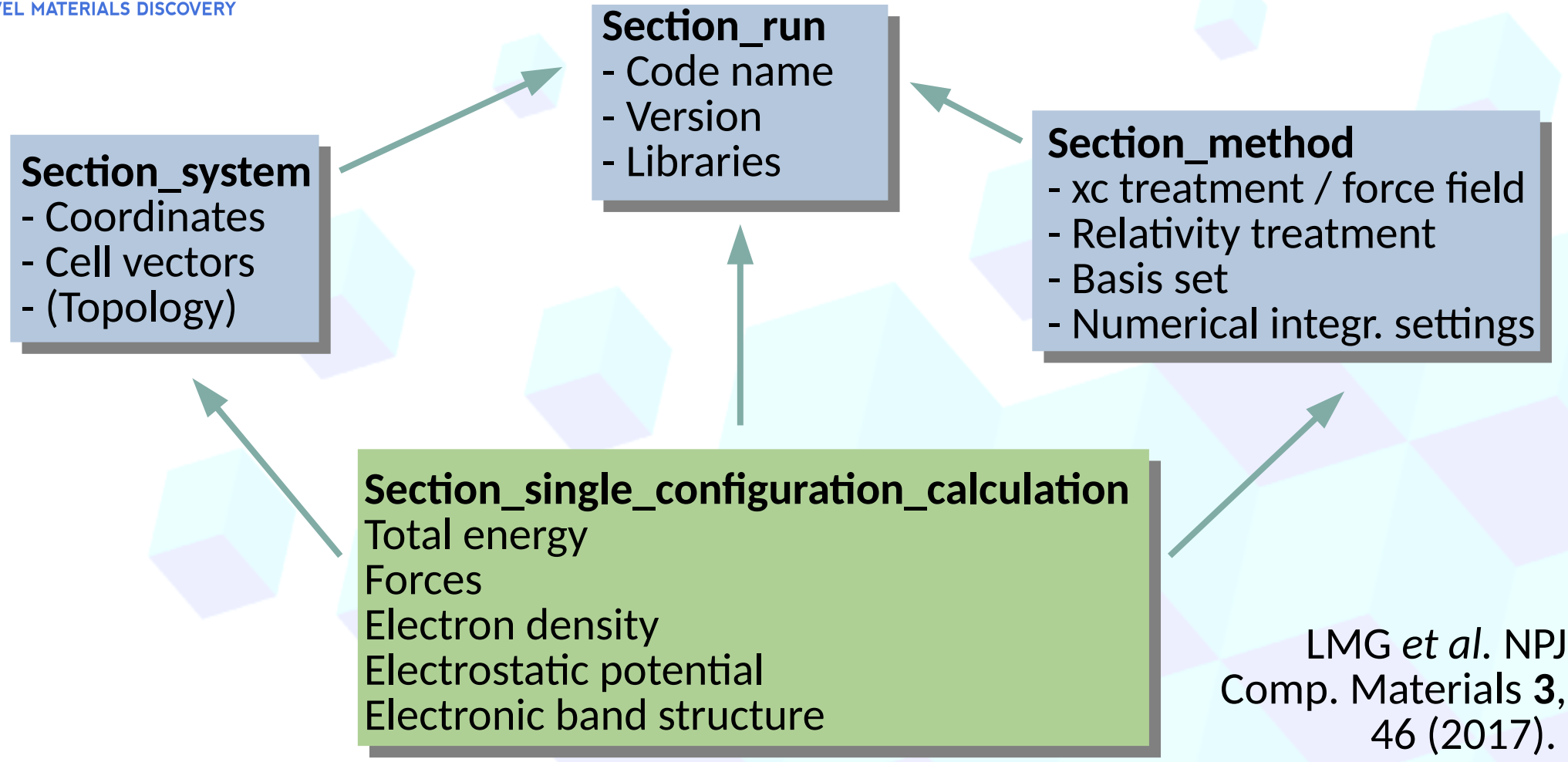
ID parent section: ...

A posteriori, hierarchical scheme

NOVEL MATERIALS DISCOVERY

NOMAD
NOVEL MATERIALS DISCOVERY

A posteriori, hierarchical scheme

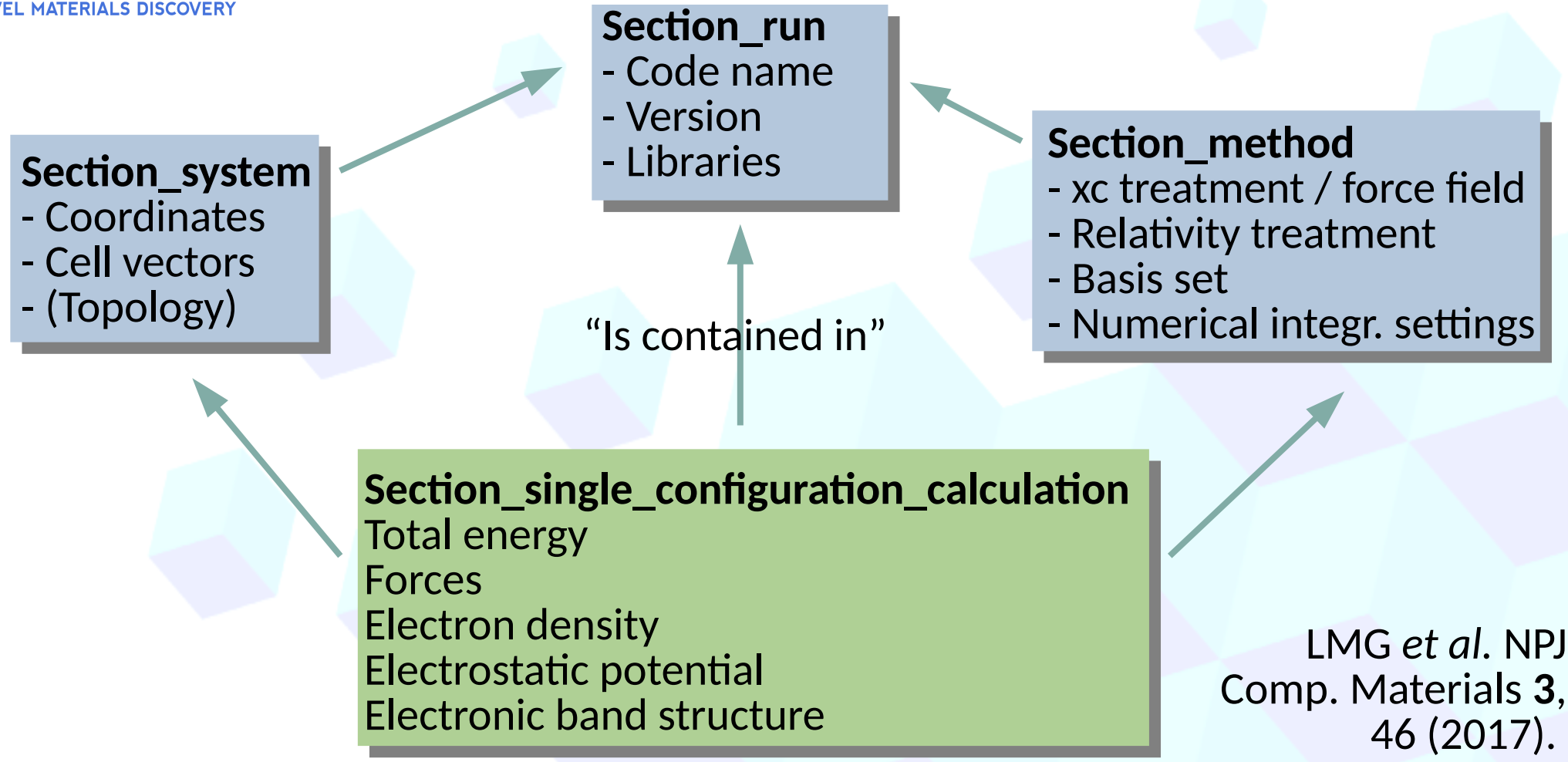


LMG *et al.* NPJ Comp. Materials **3**, 46 (2017).

NOMAD

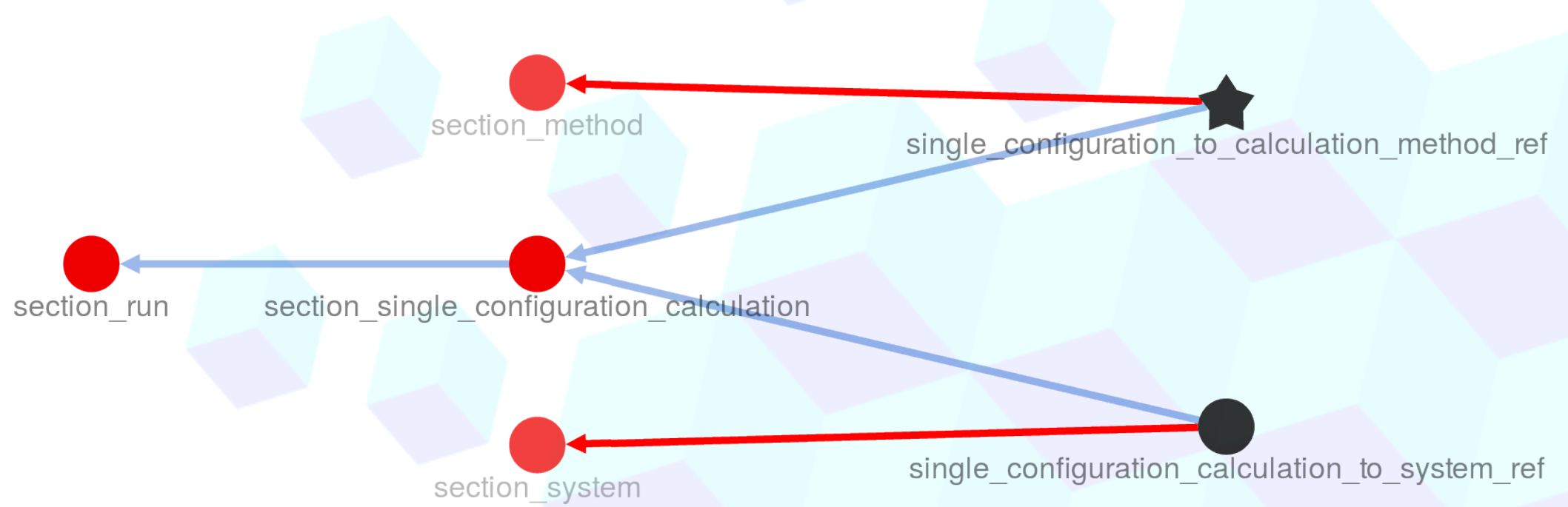
NOVEL MATERIALS DISCOVERY

A posteriori, hierarchical scheme



LMG *et al.* NPJ Comp. Materials **3**, 46 (2017).

NOMAD Meta Info

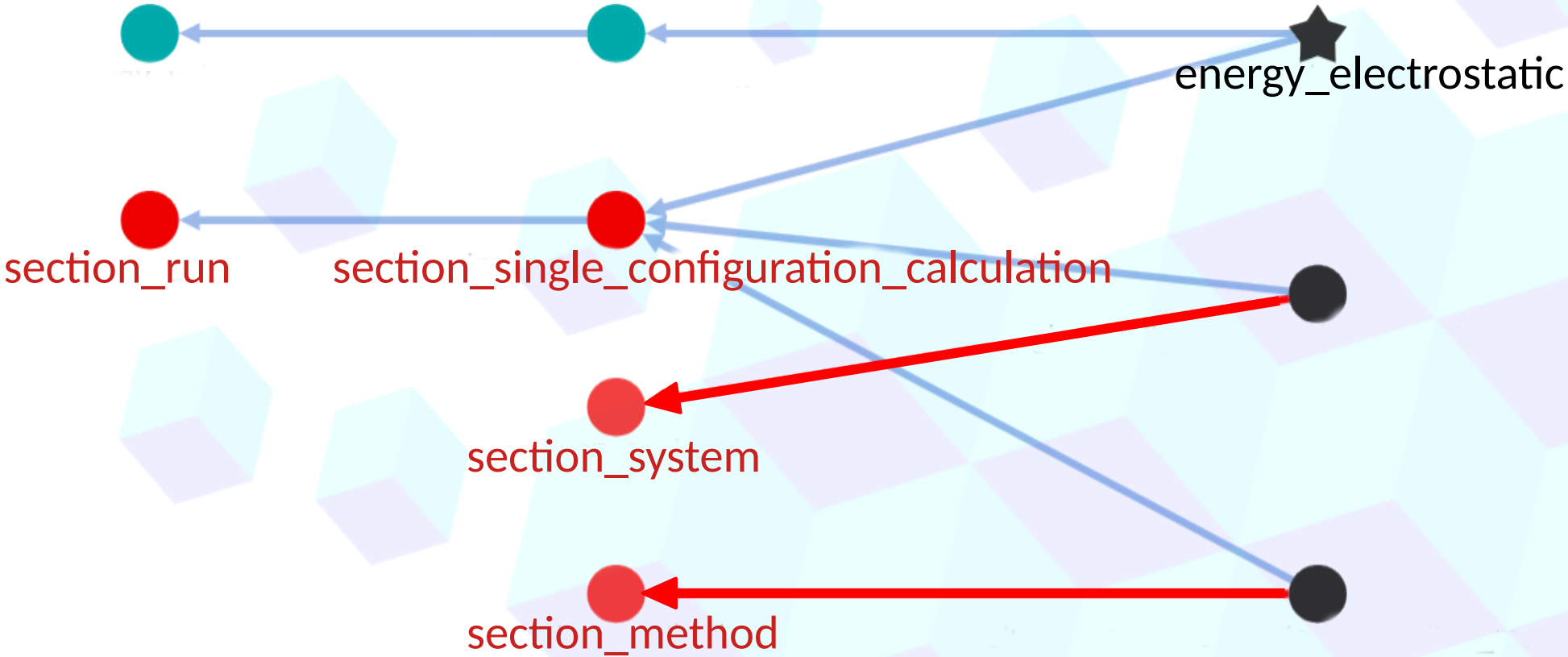


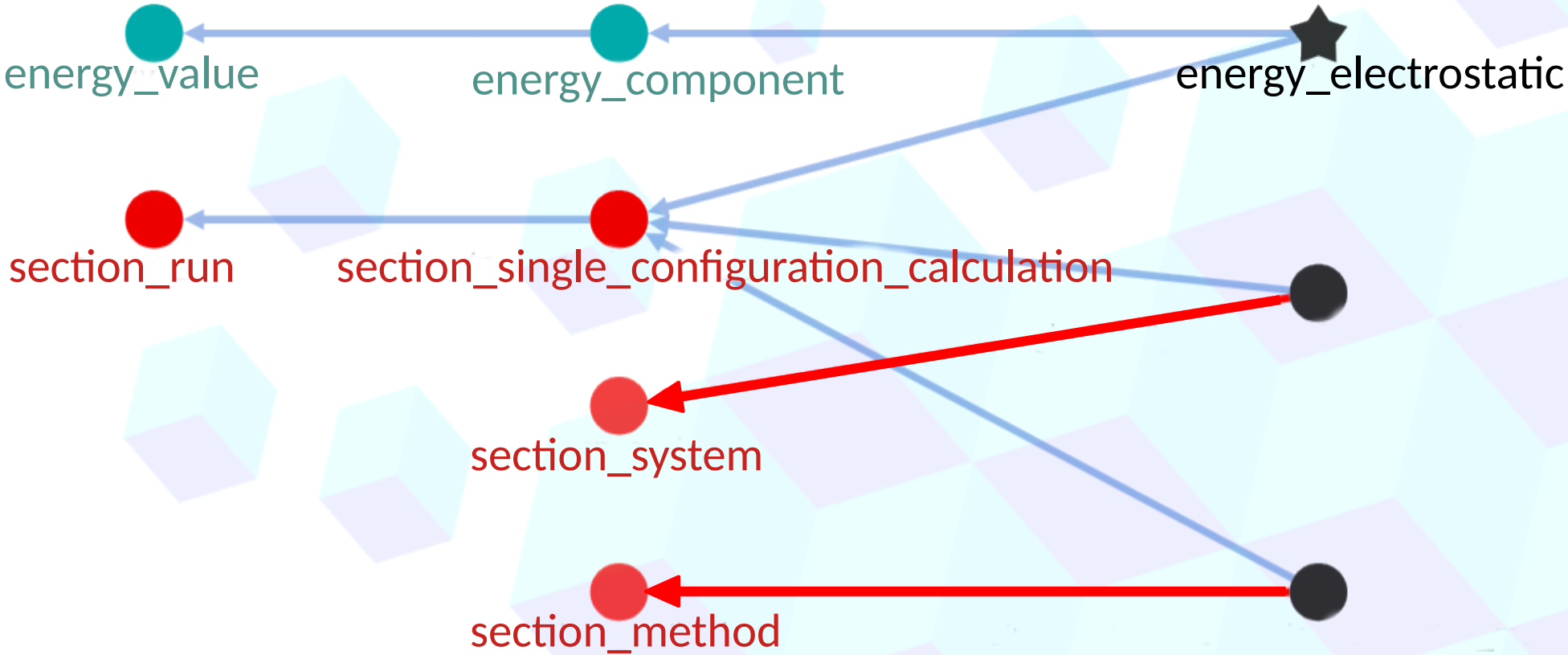
https://metainfo.nomad-coe.eu

NOVEL MATERIALS DISCOVERY

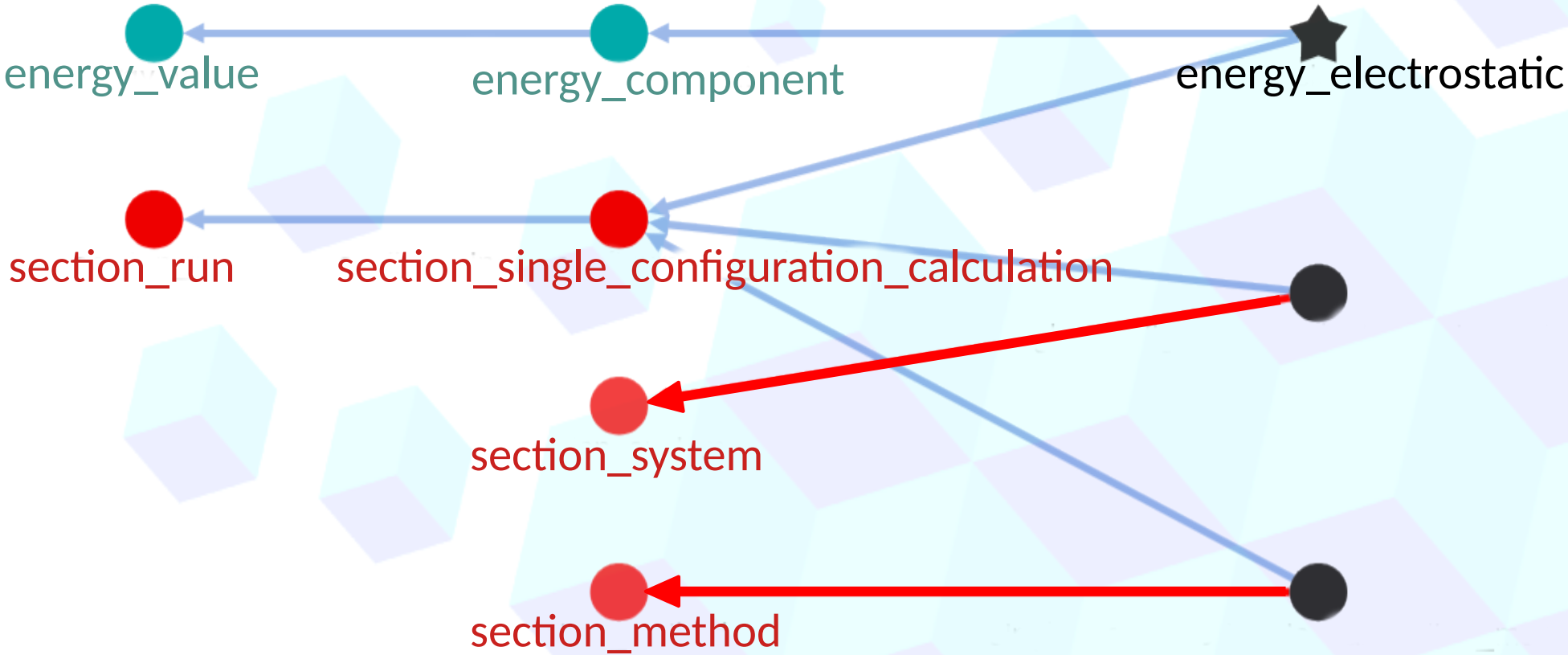
energy_electrostatic

Type: Concrete Value
Description: Total electrostatic energy (nuclei + electrons), defined consistently with calculation_method
Data Type: floating point value
Shape: []
Units: joule


energy_electrostatic
section_run
section_single_configuration_calculation
section_system
section_method


energy_value
energy_component
energy_electrostatic
section_run
section_single_configuration_calculation
section_system
section_method

Abstract type → Ontology



NOMAD MetaInfo, implementation

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.

NOMAD MetaInfo, implementation

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo

NOMAD MetaInfo, implementation

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

NOMAD MetaInfo, implementation

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

json

```
section_run
    program_name                                  FHI-aims
    program_version                               081912
    section_system
        simulation_cell                           [[1.4e-9 ...]]
        atom_positions                            [[0.0,....]...]
        atom_labels                               ["Cu",...]
    section_method
        basis_set                                 fhi_aims_tight
        XC_method                                 DFT_GGA_PBE
    section_single_configuration_calculation
        section_scf_iteration
            energy_total_scf_iteration            -1.326e-20
        section_scf_iteration
            energy_total_scf_iteration            -1.344e-20
        energy_total                              -1.344e-20
```

NOMAD MetaInfo, implementation

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

json

```
section_run
    program_name                                  FHI-aims
    program_version                               081912
  section_system
      simulation_cell                             [[1.4e-9 ...]]
      atom_positions                              [[0.0,....]...]
      atom_labels                                 ["Cu",...]
  section_method
      basis_set                                   fhi_aims_tight
      XC_method                                   DFT_GGA_PBE
  section_single_configuration_calculation
    section_scf_iteration
        energy_total_scf_iteration                -1.326e-20
    section_scf_iteration
        energy_total_scf_iteration                -1.344e-20
    energy_total                                  -1.344e-20
```

hdf5
- Binary file format
- Storing of vectors, matrices,
- Efficient for
  non-sequential reading

NOMAD MetaInfo, beyond raw data

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

NOMAD MetaInfo, beyond raw data

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

Standardization
Raw input+output → NOMAD MetaInfo “as they are” (except unit conversion)

NOMAD MetaInfo, beyond raw data

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

Standardization
Raw input+output → NOMAD MetaInfo “as they are” (except unit conversion)

Normalization
Standardized NOMAD MetaInfo → derived MetaInfo

NOMAD MetaInfo, beyond raw data

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

Standardization
Raw input+output → NOMAD MetaInfo “as they are” (except unit conversion)

Normalization
Standardized NOMAD MetaInfo → derived MetaInfo

e.g., number density = #atoms / volume

NOMAD MetaInfo, beyond raw data

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

Standardization
Raw input+output → NOMAD MetaInfo “as they are” (except unit conversion)

Normalization
Standardized NOMAD MetaInfo → derived MetaInfo

e.g., number density = #atoms / volume
Actual examples:
- band structure along path defined in W. Setyawan and S. Curtarolo, Comput. Mater. Sci. **49**, 299-312 (2010).
- space group calculated from structure via *spglib* library

NOMAD MetaInfo, beyond raw data

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

Standardization
Raw input+output → NOMAD MetaInfo "as they are" (except unit conversion)

Normalization
Standardized NOMAD MetaInfo → derived MetaInfo

e.g., number density = #atoms / volume
Actual examples:
- band structure along path defined in W. Setyawan and S. Curtarolo, Comput. Mater. Sci. **49**, 299-312 (2010).
- space group calculated from structure via *spglib* library

Useful for:
annotated storage of
"good" descriptors

NOMAD MetaInfo, community-driven

NOVEL MATERIALS DISCOVERY

Perspective paper:
L.M. Ghiringhelli, C. Carbogno, S. Levchenko, F. Mohamed, G. Huhs, M. Lueders, M. Oliveira, and M. Scheffler
Towards efficient data exchange and sharing for big-data driven materials science: Metadata and data formats.
NPJ Computational Materials **3**, 46 (2017). DOI: 10.1038/s41524-017-0048-5.

After CECAM-Psi-k workshop:
Towards a Common Format for Computational Materials Science Data
Lausanne, Switzerland, January 25 to January 27 2016

Update:
NOMAD-FAIRDI workshop:
Shared metadata and data formats for Big-Data Driven Materials Science.
Berlin, Germany, July 2019.

NOMAD MetaInfo, community-driven

NOVEL MATERIALS DISCOVERY

- A common energy zero for total energies

NOMAD MetaInfo, community-driven

NOVEL MATERIALS DISCOVERY

- A common energy zero for total energies

Atomistic approach
- build a database of isolated, spherical atom energies, for each xc functional, relativistic treatment, (pseudopotential), numerical setting

Thermodynamic approach
- build a database of elemental, (binary?) materials energies

Consider machine-learning model(s) to fill in the blanks in the database
See, e.g., Daniel Speckhard's poster

NOMAD MetaInfo, community-driven

NOVEL MATERIALS DISCOVERY

- A common energy zero for total energies
- Electronic and vibrational properties of solids

NOMAD MetaInfo, community-driven

NOVEL MATERIALS DISCOVERY

- A common energy zero for total energies
- Electronic and vibrational properties of solids
- Compact representation of scalar fields: density, wavefunction, xc potentials, etc.

  S.V. Levchenko and M. Scheffler, *Compact representation of one-particle wavefunctions and scalar fields obtained from electronic-structure calculations*. Comput. Phys. Comm. **237**, 42-46 (2019).

NOMAD MetaInfo, community-driven

NOVEL MATERIALS DISCOVERY

- A common energy zero for total energies
- Electronic and vibrational properties of solids
- Compact representation of scalar fields: density, wavefunction, xc potentials, etc.
  S.V. Levchenko and M. Scheffler, *Compact representation of one-particle wavefunctions and scalar fields obtained from electronic-structure calculations*. Comput. Phys. Comm. **237**, 42-46 (2019).
- Quantities related to excited-state calculations
  many-body perturbation theory (MBPT) calculations (GW, Bethe–Salpeter equation, etc.)

NOMAD MetaInfo, community-driven

NOVEL MATERIALS DISCOVERY

- A common energy zero for total energies
- Electronic and vibrational properties of solids
- Compact representation of scalar fields: density, wavefunction, xc potentials, etc.
  S.V. Levchenko and M. Scheffler, *Compact representation of one-particle wavefunctions and scalar fields obtained from electronic-structure calculations*. Comput. Phys. Comm. **237**, 42-46 (2019).
- Quantities related to excited-state calculations
  many-body perturbation theory (MBPT) calculations (GW, Bethe–Salpeter equation, etc.)
- Sampling of potential energy surfaces (Molecular dynamics, Monte Carlo, ...)
- Workflows

NOMAD MetaInfo, community-driven

NOVEL MATERIALS DISCOVERY

- A common energy zero for total energies
- Electronic and vibrational properties of solids
- Compact representation of scalar fields: density, wavefunction, xc potentials, etc.
  S.V. Levchenko and M. Scheffler, *Compact representation of one-particle wavefunctions and scalar fields obtained from electronic-structure calculations*. Comput. Phys. Comm. **237**, 42-46 (2019).
- Quantities related to excited-state calculations
  many-body perturbation theory (MBPT) calculations (GW, Bethe–Salpeter equation, etc.)
- Sampling of potential energy surfaces (Molecular dynamics, Monte Carlo, ...)
- Workflows
- Experiments

NOMAD MetaInfo, accessibility

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

NOMAD MetaInfo, accessibility

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

Access: via URL

https://analytics-toolkit.nomad-coe.eu/api/resolve/
N-TULHlZnc9cnbg7ihzUALIlSdyww/
C-DWGvyqvK2g_1yLyJf8nN3j_M-xd/
section_run/0c/
section_single_configuration_calculation/0c/
energy_total/0c

NOMAD MetaInfo, accessibility

NOVEL MATERIALS DISCOVERY

The NOMAD team wrote 40+ parsers.
Raw input+output → NOMAD MetaInfo → json + hdf5 file format → Archive

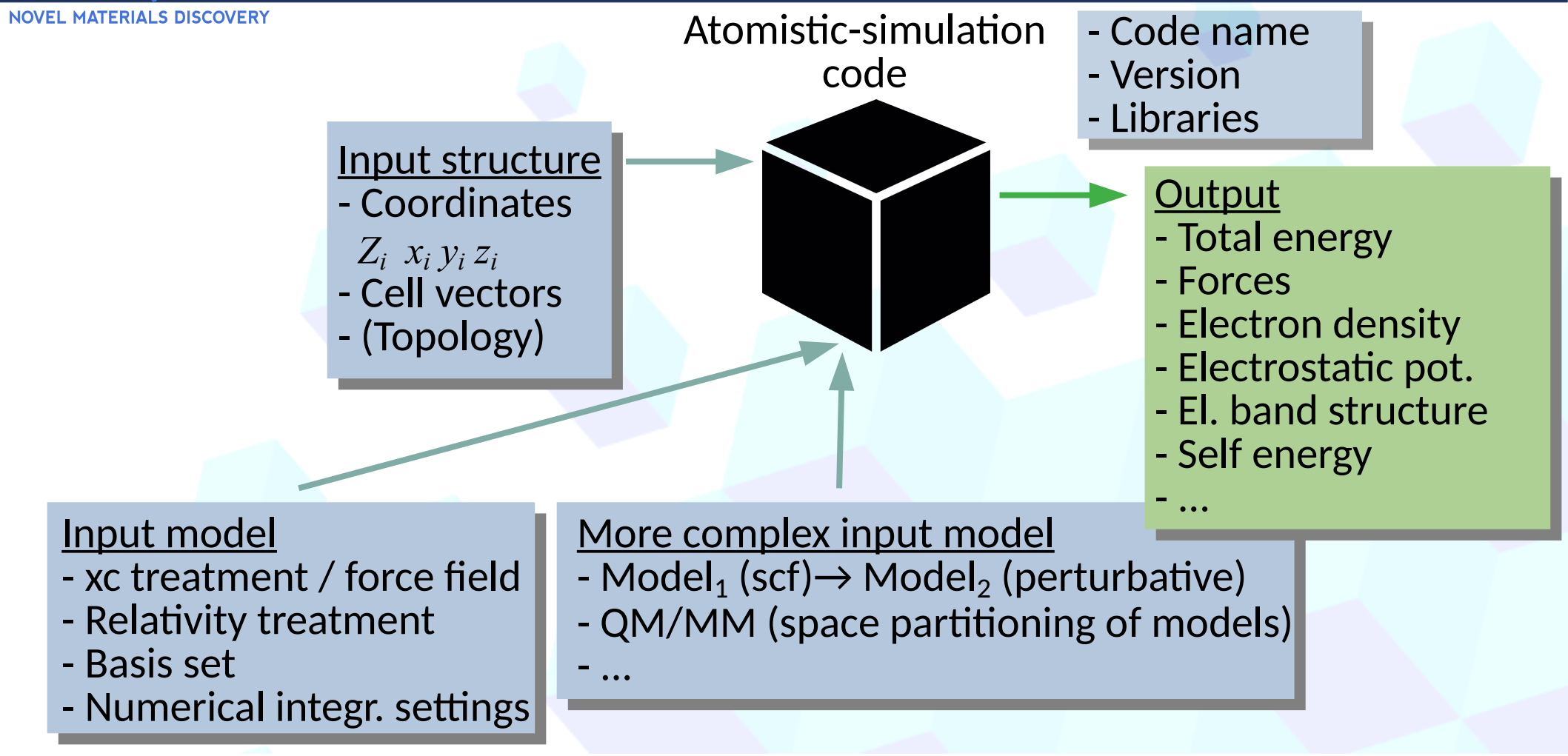
Access: via URL

https://analytics-toolkit.nomad-coe.eu/api/resolve/
N-TULHlZnc9cnbg7ihzUALIlSdyww/
C-DWGvyqvK2g_1yLyJf8nN3j_M-xd/
section_run/0c/
section_single_configuration_calculation/0c/
energy_total/0c


```
uri:        "nmd://N-TULHlZnc9cnbg7ihzUALIlSdyww/C-DWGvyqvK2g_1yLyJf8nN3j_M-xd/section_run/0c/section_single_configuration_calculation/0c/energy_total/0c"
type:       "value"
dtypeStr:   "f"
value:      -4.535646058456239e-18
```

NOMAD
NOVEL MATERIALS DISCOVERY

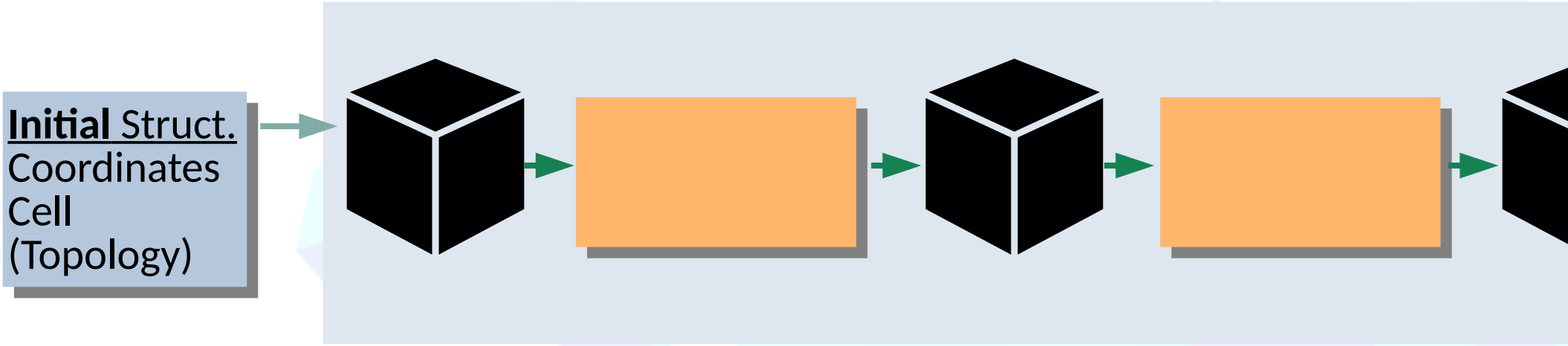
Computational MatSci: from basics...



NOMAD

NOVEL MATERIALS DISCOVERY

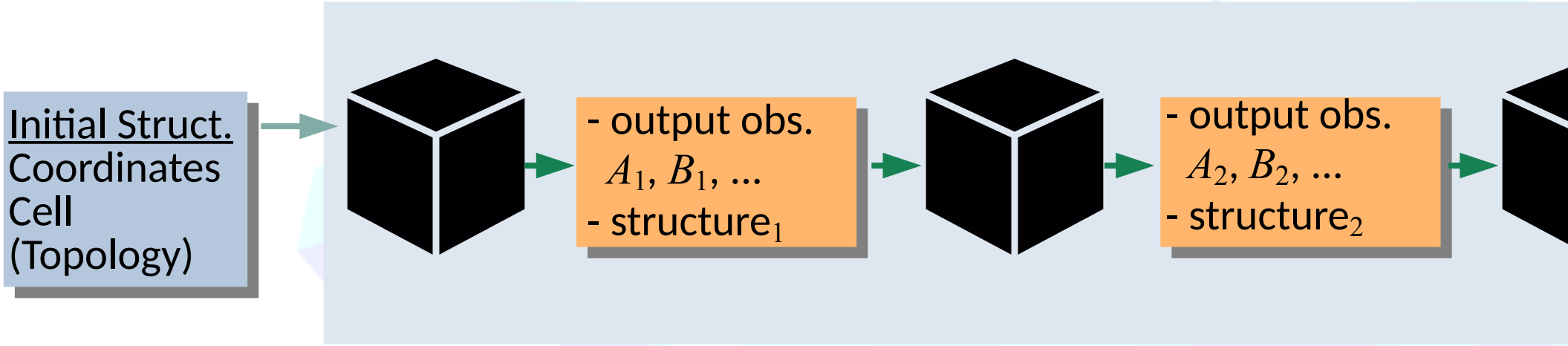
… to a more complex *workflow* …



NOMAD

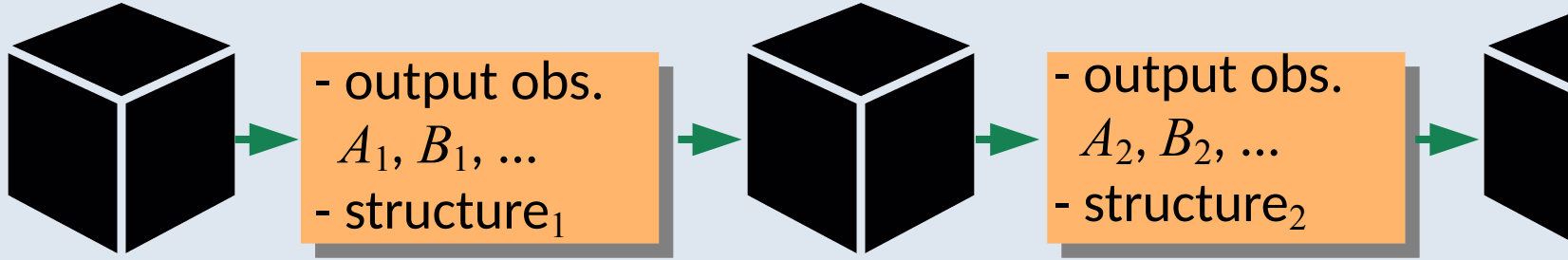
NOVEL MATERIALS DISCOVERY

… to a more complex *workflow* …



NOMAD
NOVEL MATERIALS DISCOVERY

… to a more complex *workflow* …



Initial Struct.
Coordinates
Cell
(Topology)

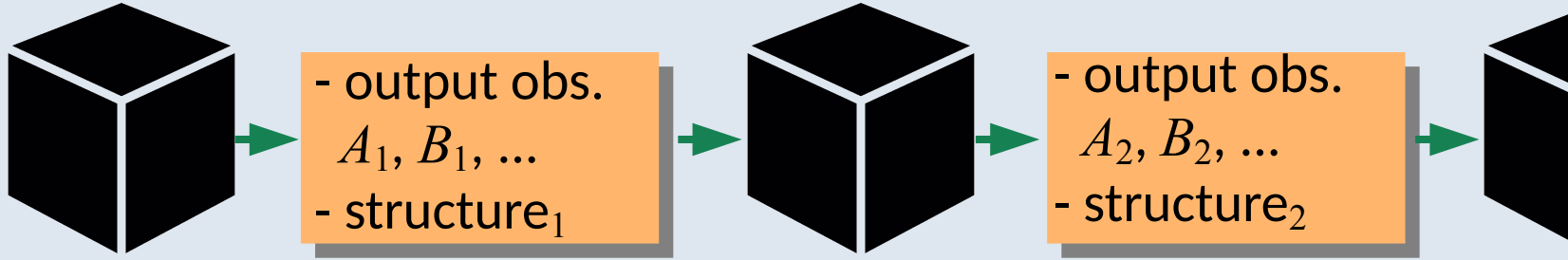
- output obs.
  A_1, B_1, …
- structure$_1$

- output obs.
  A_2, B_2, …
- structure$_2$

Type of *sampling*
- forces/stress relaxation
- molecular dynamics
- Monte Carlo
- replica exchange
- phonons
- equation of state, e.g., $E(V)$
- **“high-throughput”**

… to a more complex *workflow* …

NOVEL MATERIALS DISCOVERY



Initial Struct.
Coordinates
Cell
(Topology)

- output obs.
  $A_1, B_1, \ldots$
- structure$_1$

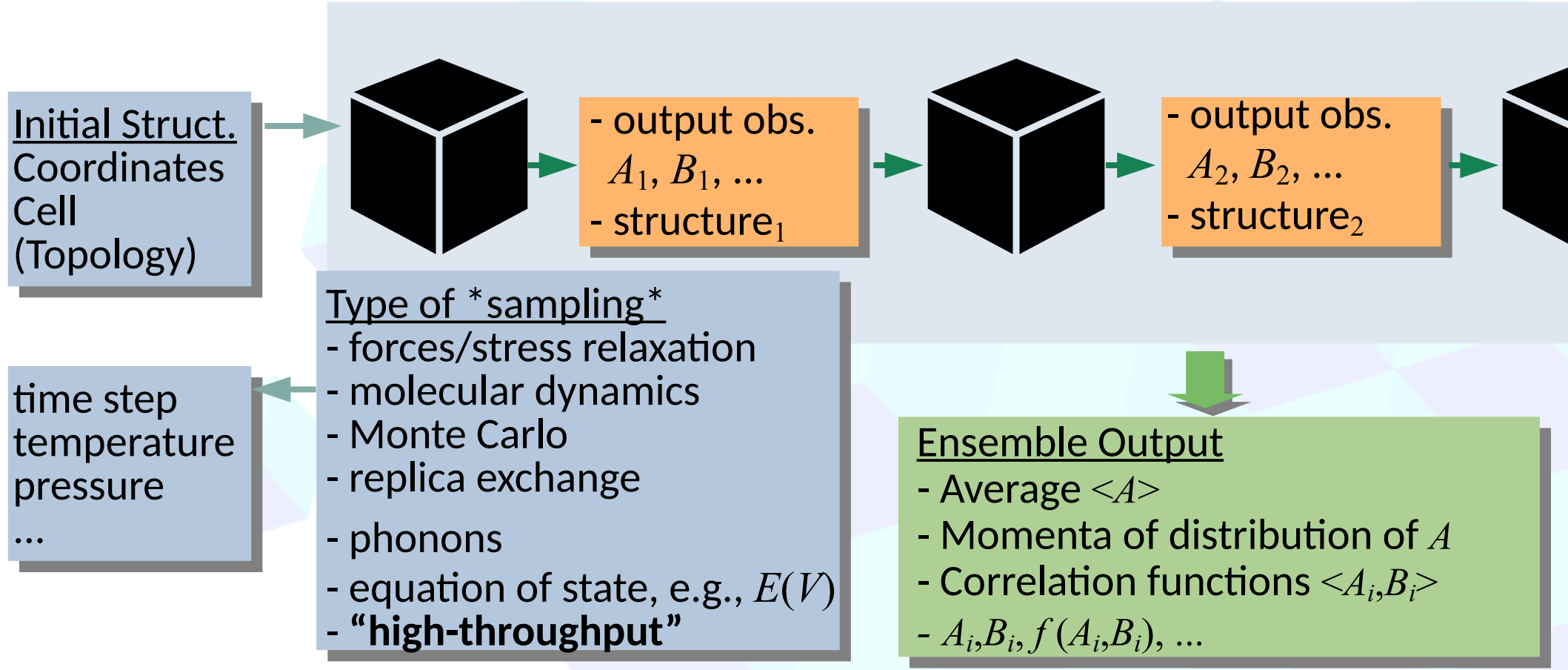
- output obs.
  $A_2, B_2, \ldots$
- structure$_2$

time step
temperature
pressure
…

Type of *sampling*
- forces/stress relaxation
- molecular dynamics
- Monte Carlo
- replica exchange
- phonons
- equation of state, e.g., $E(V)$
- **“high-throughput”**

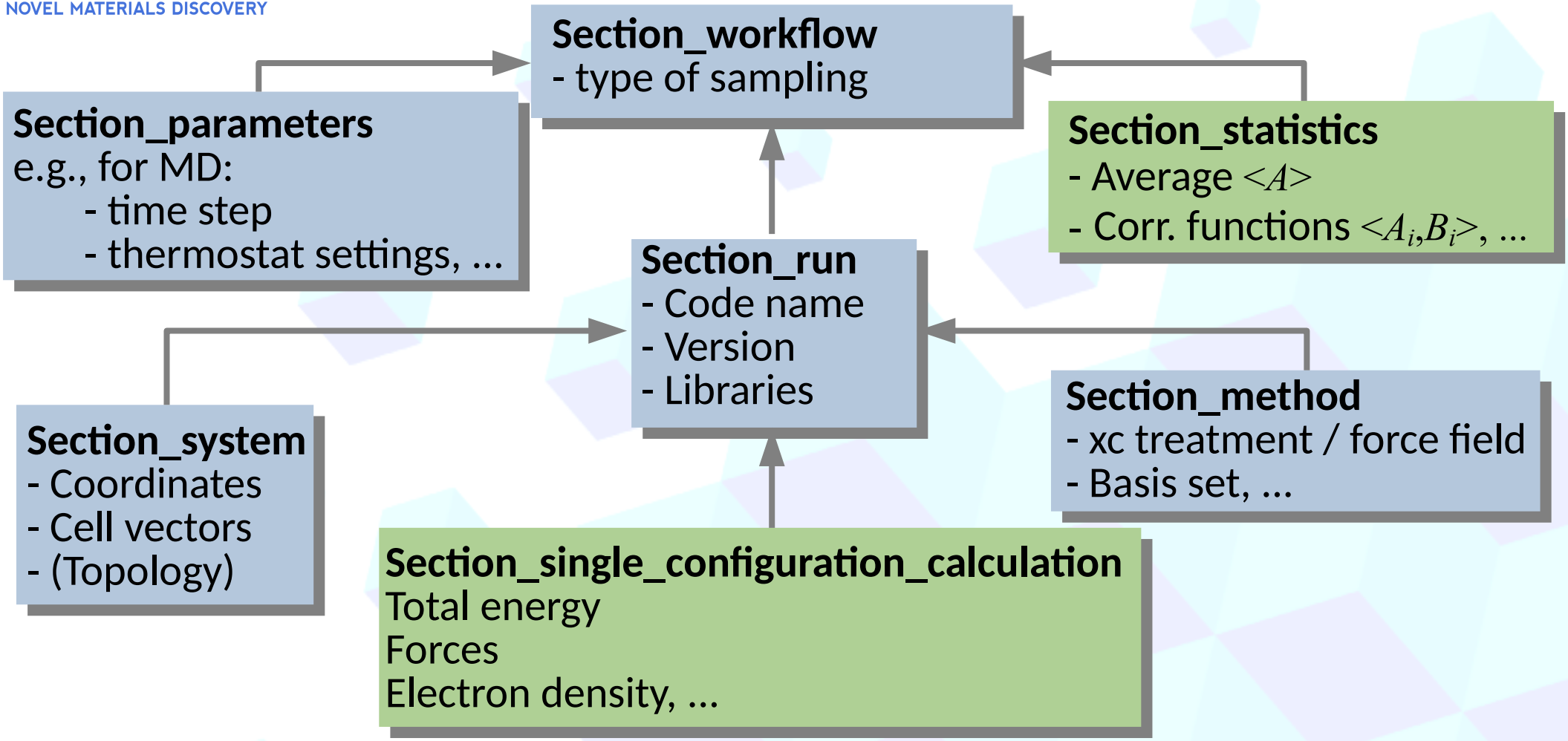
NOMAD
NOVEL MATERIALS DISCOVERY

… to a more complex *workflow* …



… and the related hierarchical scheme

NOVEL MATERIALS DISCOVERY



Keywords and Acknowledgments

NOVEL MATERIALS DISCOVERY

Metadata for FAIR scientific-data management and stewardship:

- Hierarchical (sections, concrete values, ...)
- Structured (name, description, ...)
- Extensible

Acknowledgments:

Fawzi Mohamed, Pasquale Pavone, Henning Glawe, Micael Olivera, Benjamin Regler, Bryan Goldsmith, Ádám Fekete, Markus Scheidgen.
NOMAD PIs: Matthias Scheffler, Claudia Draxl, Angel Rubio, Risto Nieminen, Francesc Illas, Daan Frenkel, Alessandro De Vita, Kristian Thygesen; Kimmo Koski, Stefan Heinzel, Jose Maria Cela, Dieter Kranzlmüller; Ciaran Clissman.

NOMAD continues via FAIR – Data Infrastructure (https://fairdi.eu/)

… and that was NOT all, folks

NOVEL MATERIALS DISCOVERY

- Extension to experimental metadata

… and that was NOT all, folks

NOVEL MATERIALS DISCOVERY

- Extension to experimental metadata
- Challenge of reproducibility: “What about artificial-intelligence models?”. AI training and learned models need to be fully annotated.

… and that was NOT all, folks

NOVEL MATERIALS DISCOVERY

- Extension to experimental metadata
- Challenge of reproducibility: “What about artificial-intelligence models?”. AI training and learned models need to be fully annotated.
- Challenge or reusability: “Is the (set of) calculation(s) I need already in some FAIR db? Is something **close** to what I need already there?”

Section_system
- Coordinates
- Cell vectors
- (Topology)

Section_method
- xc treatment / force field
- Relativity treatment
- Basis set, …

… and that was NOT all, folks

- Extension to experimental metadata
- Challenge of reproducibility: “What about artificial-intelligence models?”. AI training and learned models need to be fully annotated.
- Challenge or reusability: “Is the (set of) calculation(s) I need already in some FAIR db? Is something **close** to what I need already there?”

  How?
  We need well-calibrated **datasets** (diverse structures and quantities), recalculated with different codes, xc treatments, basis sets, etc. in order to assess the quality and “relative distance” in the stored data.