Predicting the properties of molecular materials: multiscale simulation workflows meet machine learning

DAIMON
www.daimoners.eu
Follow us! @daimoners

CNR ISMN
ISTITUTO PER LO STUDIO DI MATERIALI NANOSTRUTTURATI

Fabio Le Piane, Matteo Baldoni and Francesco Mercuri - DAIMON Team - CNR-ISMN Bologna, Italy - Email: fabio.lepiane@ismn.cnr.it

Molecular materials:
from individual molecules to aggregates and interfaces
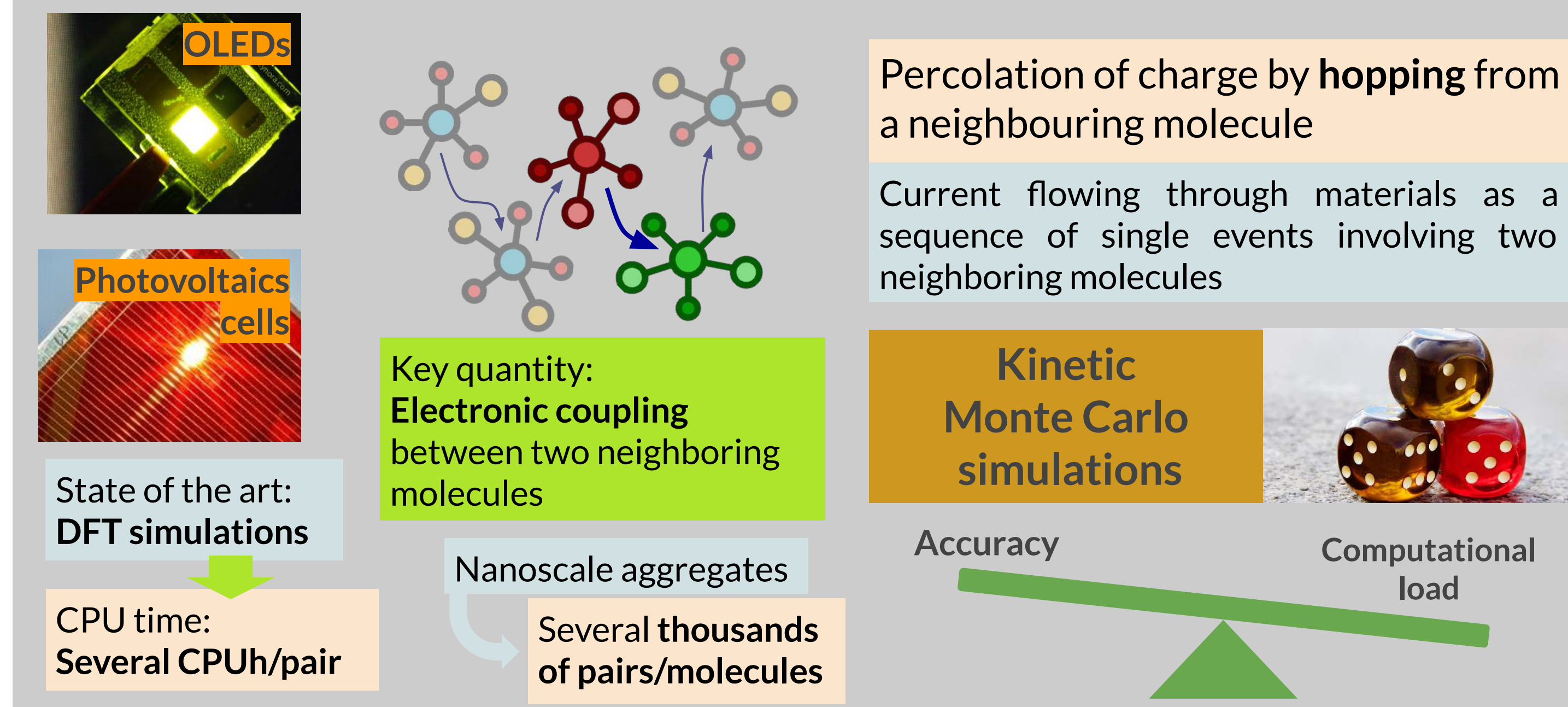
Molecule
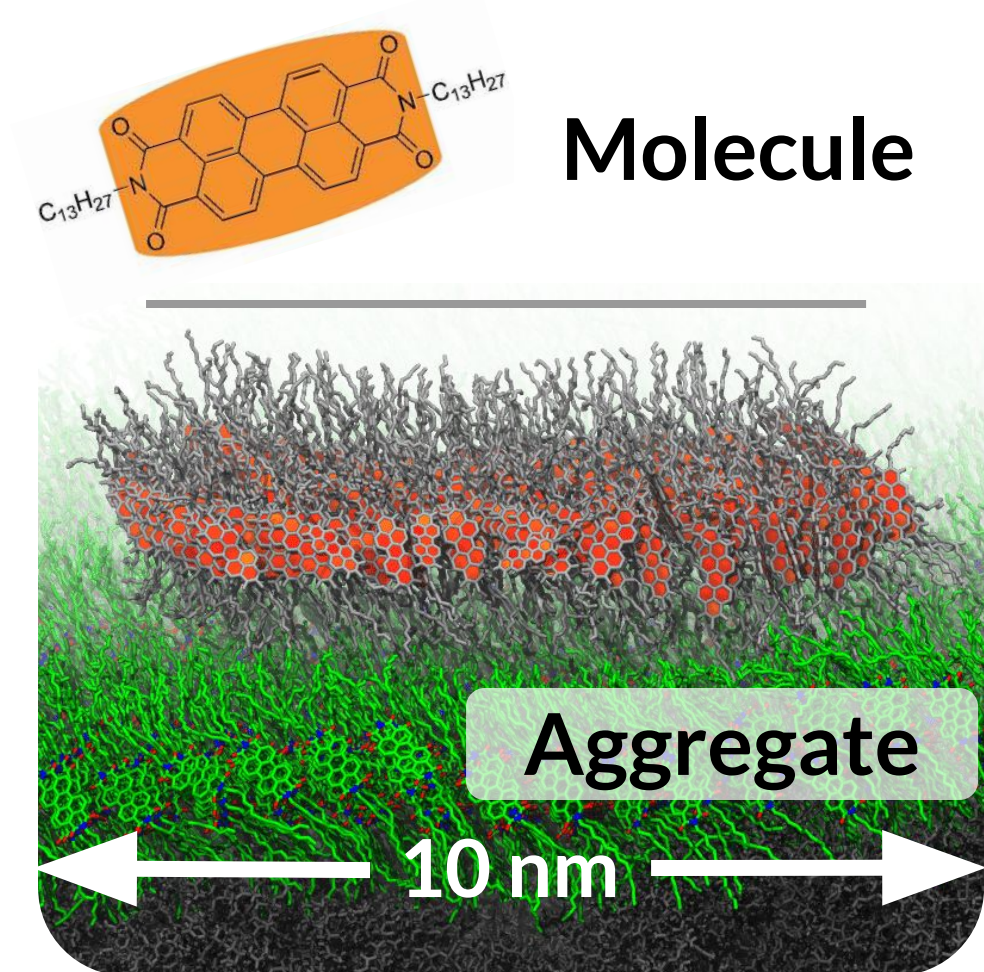
Aggregate
10 nm

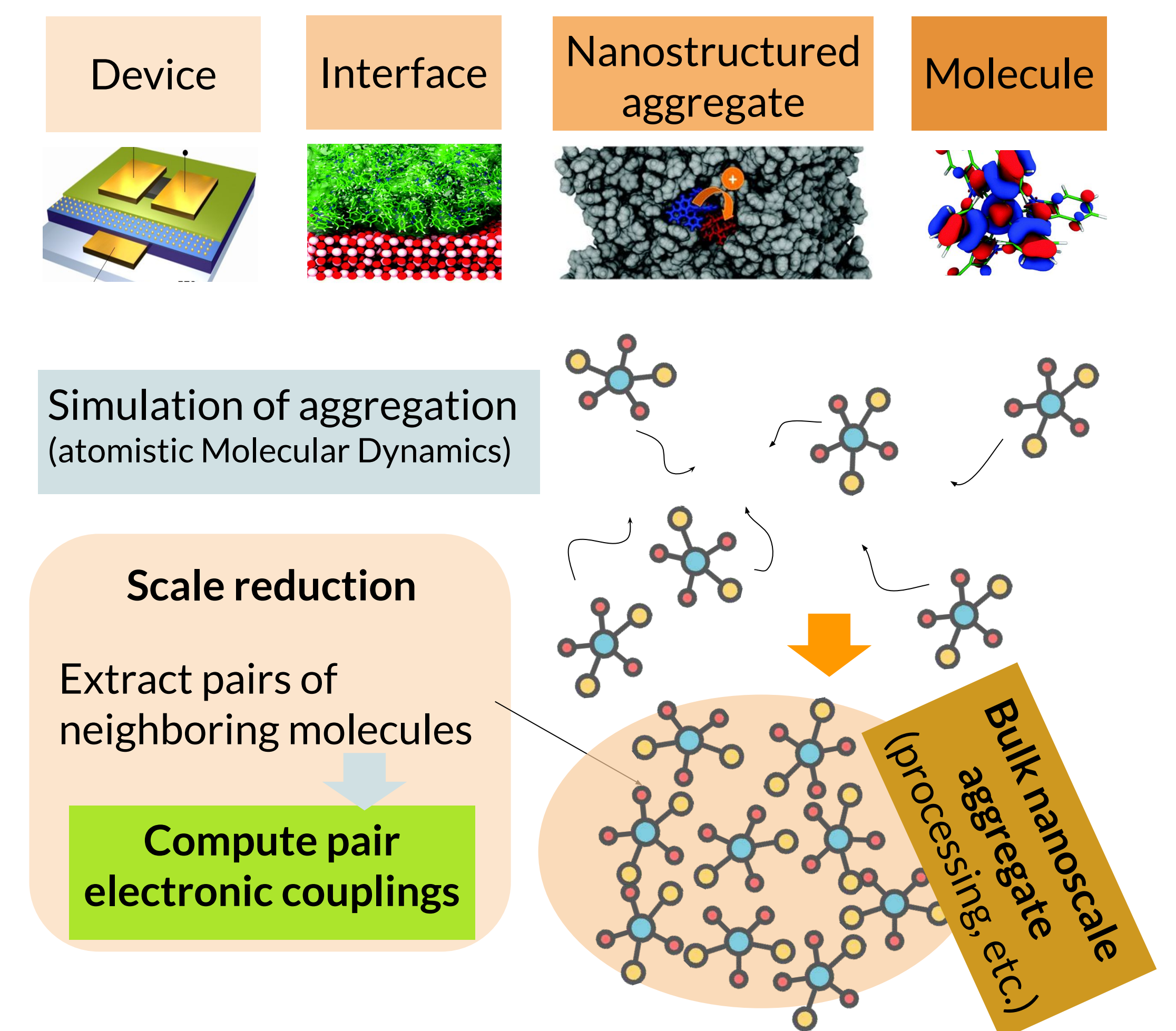### Properties of molecular materials

Properties of individual molecules

Properties of molecular aggregates

Nanoscale aggregation and morphology impact on several properties of molecular materials

Multiscale simulations
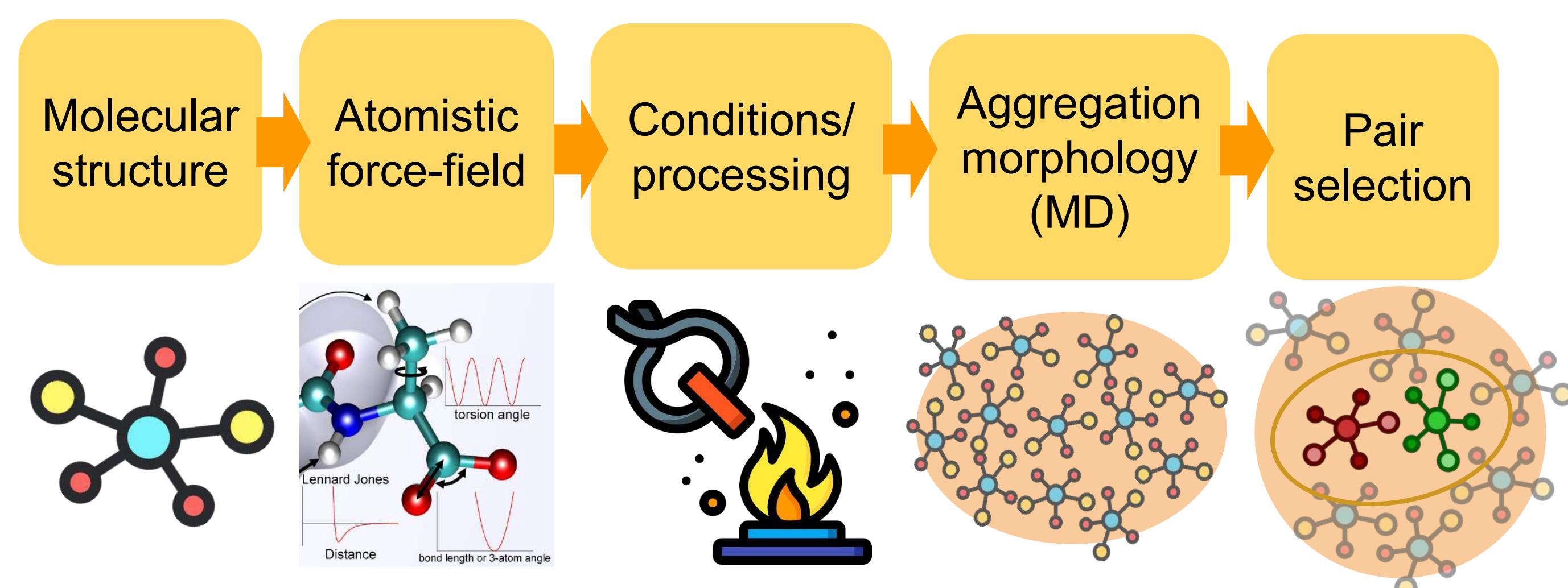Linking the properties of molecules to the properties of molecular aggregates
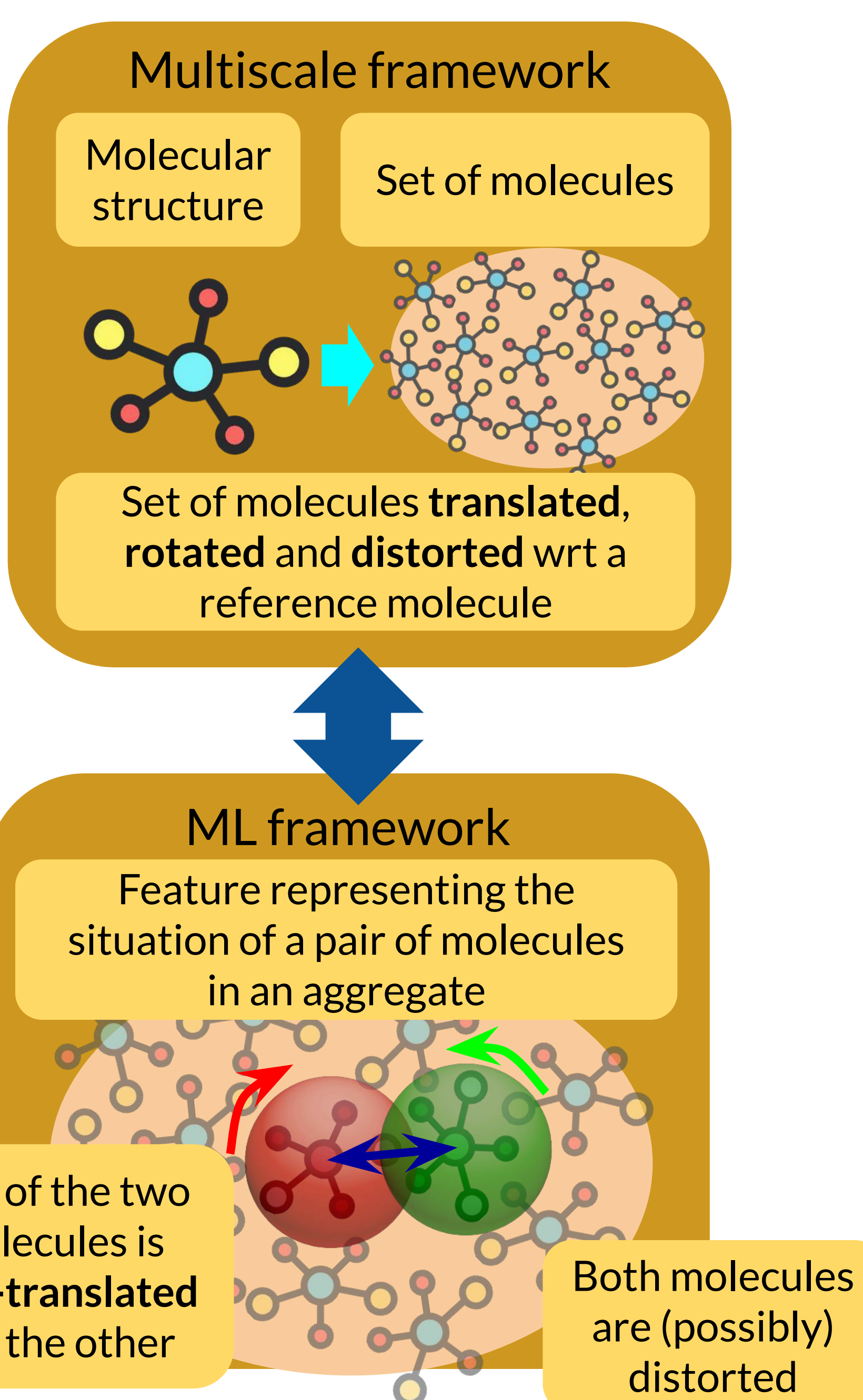
Nanoscale

Charge transport of molecular semiconductors

Electronic properties of individual molecules

Aggregation/interaction/morphology/interfaces

Charge transport in molecular semiconductors

OLEDs

Photovoltaics cells

State of the art: **DFT simulations**

CPU time: **Several CPUh/pair**

Nanoscale aggregates

Several **thousands of pairs/molecules**

Percolation of charge by **hopping** from a neighbouring molecule

Current flowing through materials as a sequence of single events involving two neighboring molecules

Key quantity: **Electronic coupling** between two neighboring molecules

Kinetic Monte Carlo simulations

Accuracy   Computational load

Multiscale top-down approach

Device   Interface   Nanostructured aggregate   Molecule

Simulation of aggregation (atomistic Molecular Dynamics)

Scale reduction

Extract pairs of neighboring molecules

Compute pair electronic couplings

Bulk nanoscale aggregate (processing, etc.)

The multiscale simulation and data workflow

Molecular structure → Atomistic force-field → Conditions/processing → Aggregation morphology (MD) → Pair selection

Learning from simulated data

Selected pairs

Predictions over pairs

Overcoming the computational bottleneck!

The Machine Learning approach

Random Forest

Kernel Ridge Regression (Gaussian) $c)^T (\mathbf{K} + \lambda$
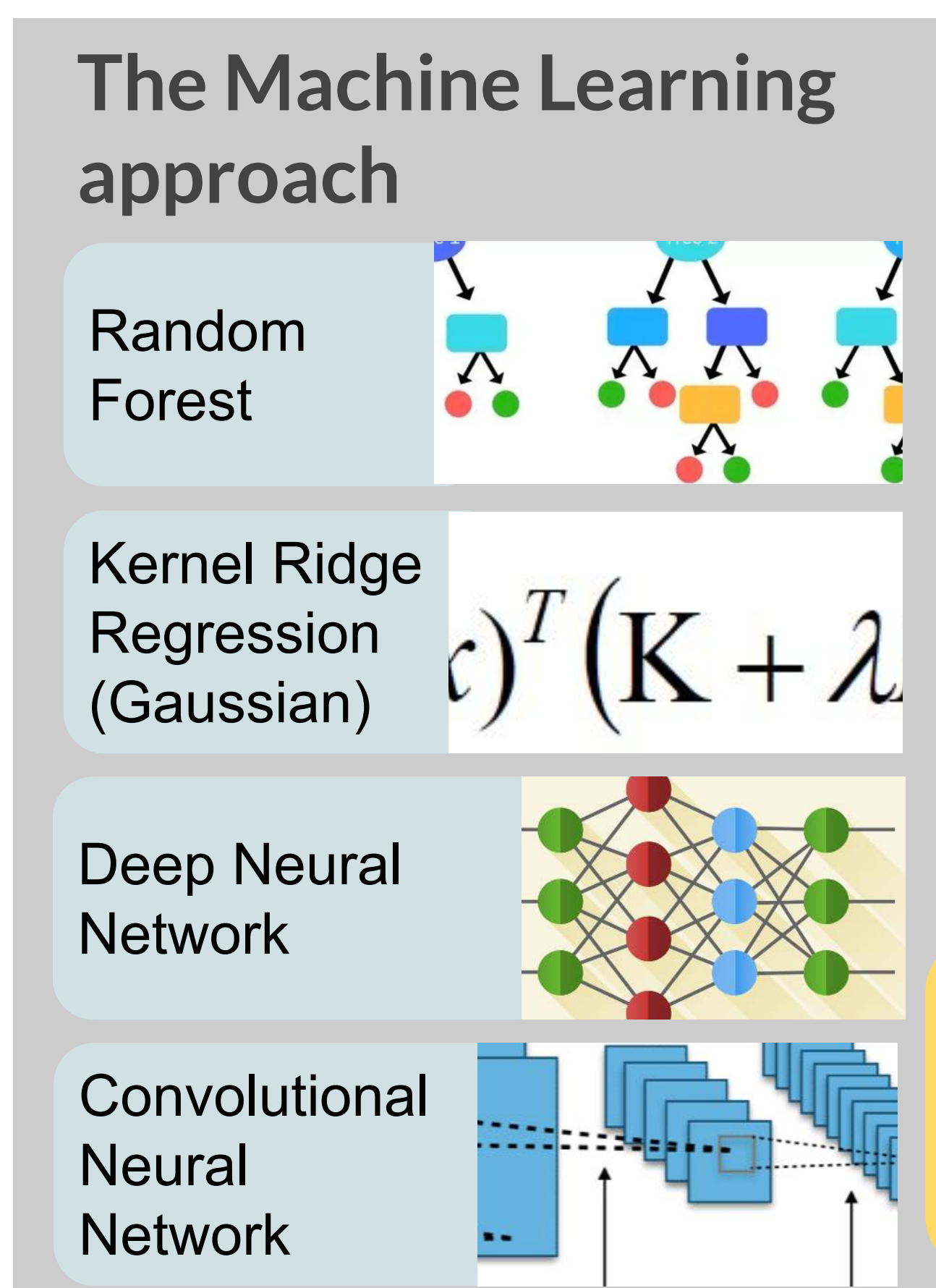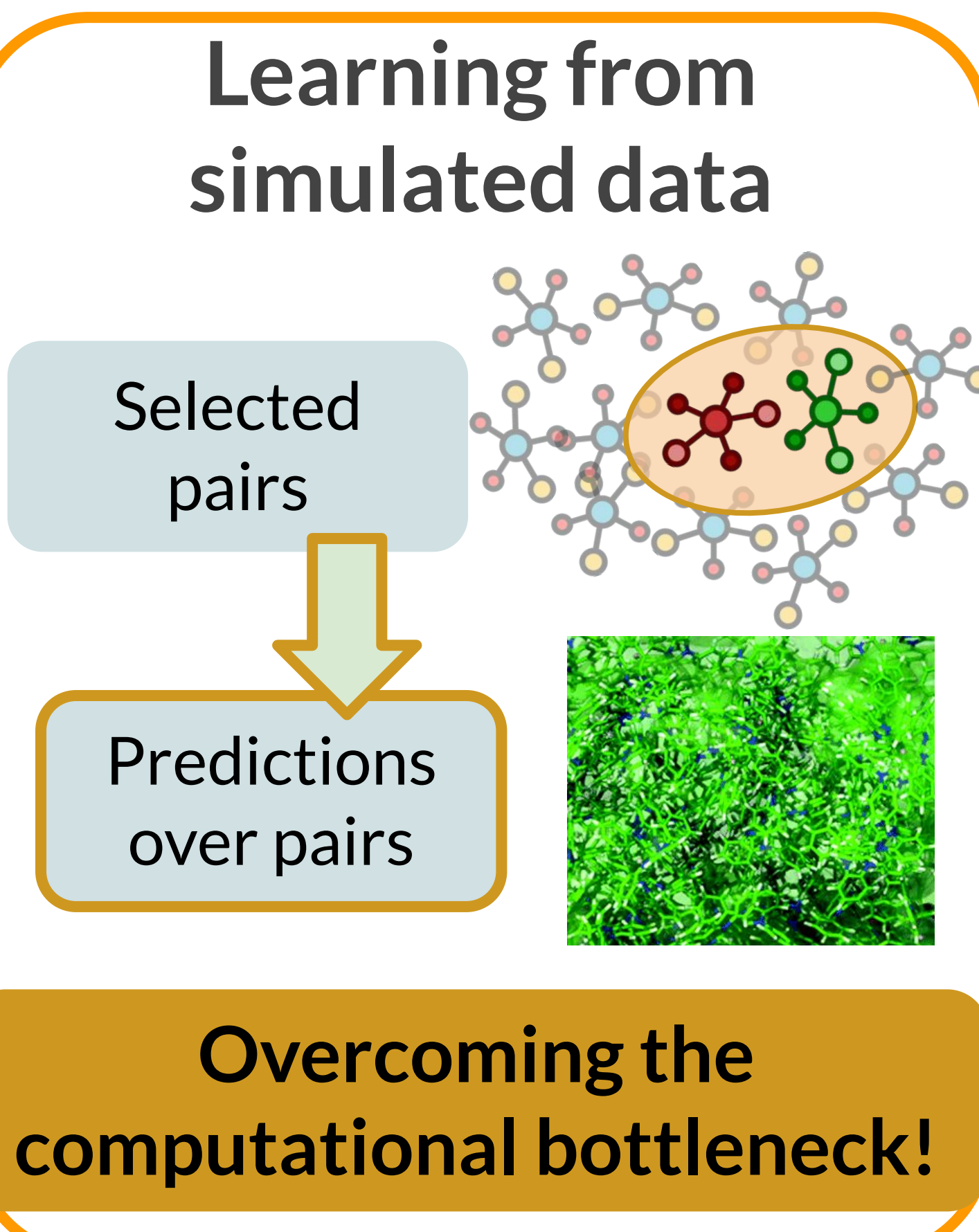
Deep Neural Network

Convolutional Neural Network

Multiscale workflow and learning

Multiscale framework

Molecular structure   Set of molecules

Set of molecules **translated**, **rotated** and **distorted** wrt a reference molecule

ML framework
Feature representing the situation of a pair of molecules in an aggregate

One of the two molecules is **roto-translated** wrt the other

Both molecules are (possibly) distorted
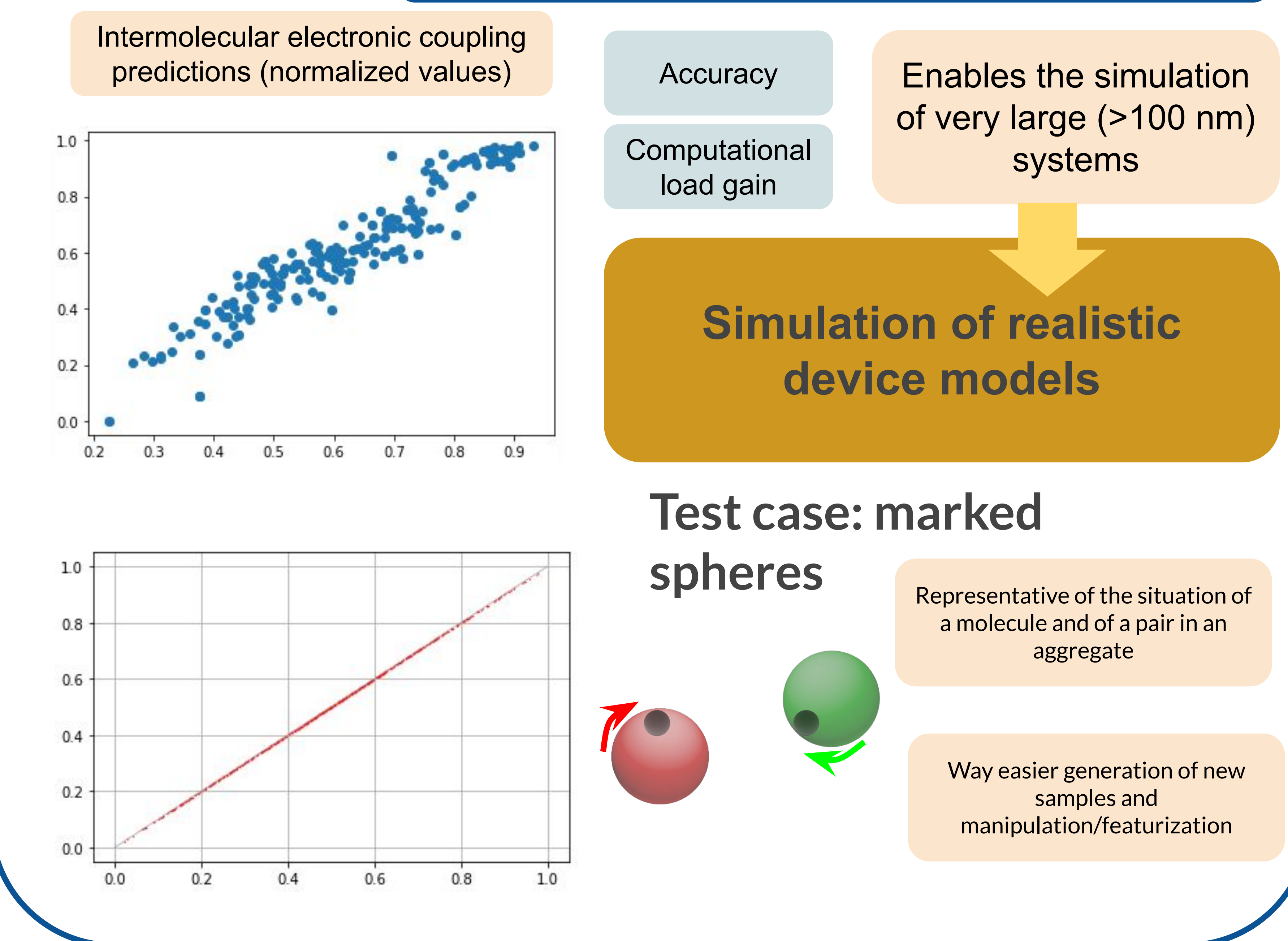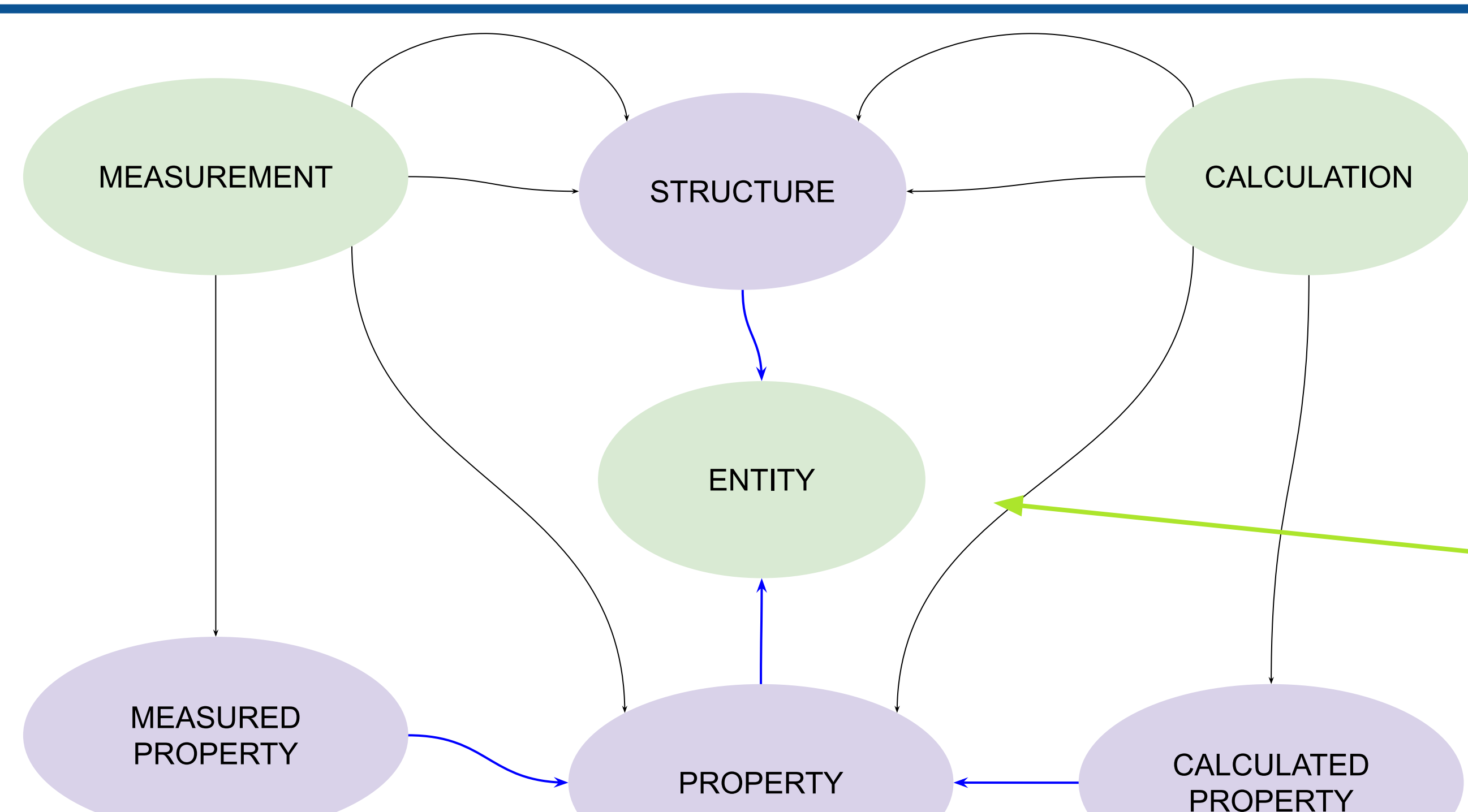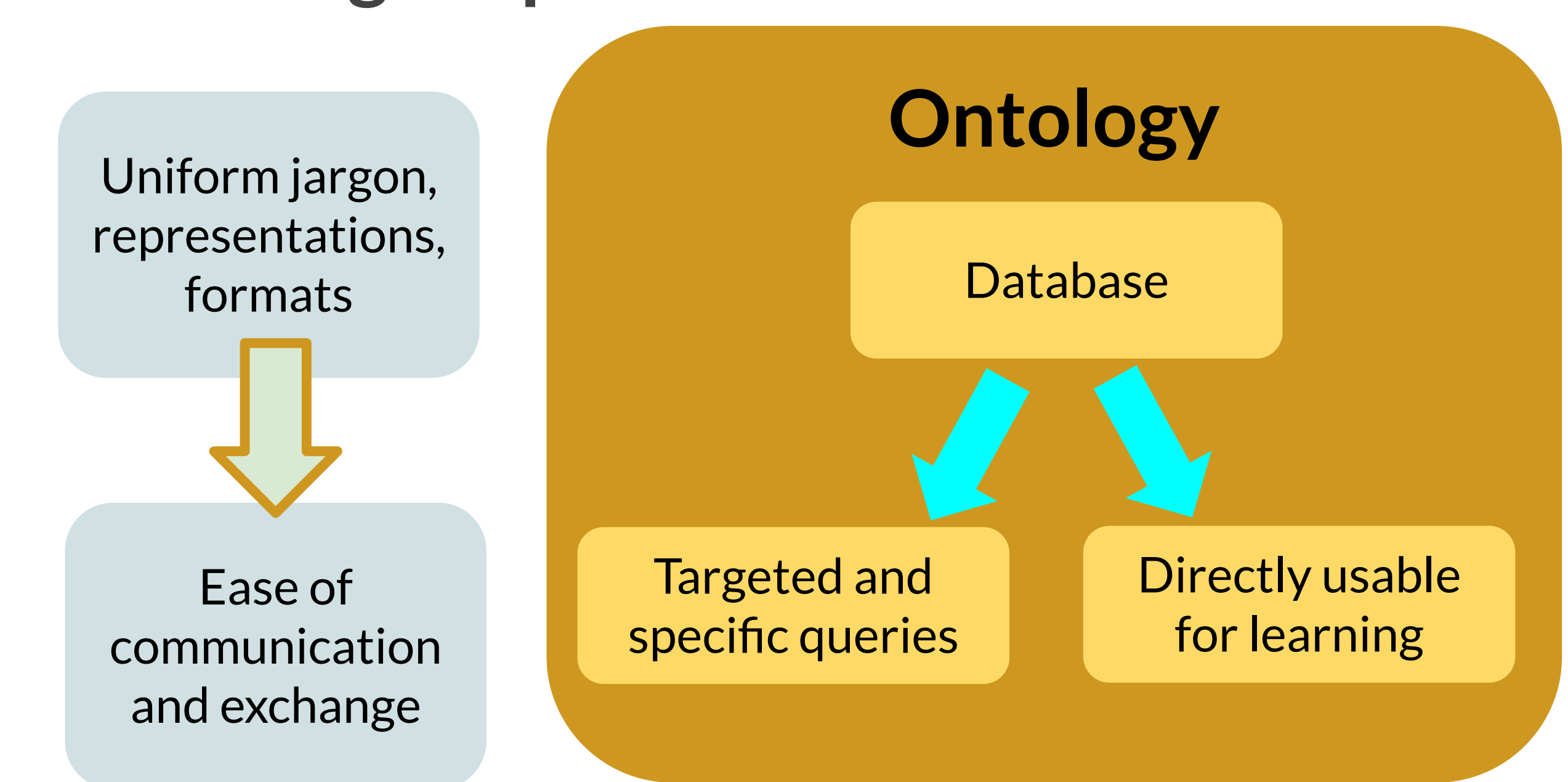
Numerical results

Case study: charge transport properties of a molecular semiconductor material used in OLEDs

Intermolecular electronic coupling predictions (normalized values)

Accuracy

Computational load gain

Enables the simulation of very large (>100 nm) systems

Simulation of realistic device models

Test case: marked spheres

Representative of the situation of a molecule and of a pair in an aggregate

Way easier generation of new samples and manipulation/featurization

Knowledge representation

Uniform jargon, representations, formats

Ease of communication and exchange

Ontology

Database

Targeted and specific queries

Directly usable for learning

MEASUREMENT   STRUCTURE   CALCULATION

ENTITY

MEASURED PROPERTY   PROPERTY   CALCULATED PROPERTY

A chemico-physical object (Entity) is represented by its structural features and properties

Focused on **molecular materials**

Linked to **pre-existent ontologies** for materials science (MDO, EMMO)

Computational and experimental workflows are separately represented but results are 100% **interoperable**