MAX DRIVING THE EXASCALE TRANSITION

Depositing data on the Materials Cloud Archive

Valeria Granata

ascale", has received funding from the European Union's Horizon 2020 project call H2020-INFRAEDI-
8

Materials Cloud

A platform to enable open and seamless sharing of resources for computational science

- Support researchers throughout the life cycle of a scientific project
- Make research output FAIR and re-producible
- Guarantee a long-term storage of raw and curated data - focused on sharing the full provenance of calculations

Materials Cloud Archive

A moderated repository for research data from computational materials science

- Long-term storage of records and associated metadata - at least 10 years
  - Runs on Virtual Machines in the OpenStack installations at the CSCS - Swiss National Supercomputing Centre
  - Data is stored on the Swift Object Store provided by CSCS, with backup replication at a different geographical location
- Findability via persistent identifiers – DOI

  DOI 10.24435/materialscloud:2019.0083/v1 - released by ETHBIB
- Accessibility via standard protocols
  - Records can be harvested using:
    - OAI-PMH protocol
    - JSON REST API

Main criteria for publication

Content

- Relevant computational materials science data
- Data quality should be as of a journal publication

Experimental work is accepted provided that

- It is the work of authors affiliated with Materials Cloud partners
- The work relates to published computational results

Files

- 5 GB
- 50 GB for AiiDA databases
- Larger datasets are accepted on request
- Should not contain the paper itself or any of its supplementary materials

Copyright

- Authors confirm to have legal right to publish
- We accept several licenses (https://spdx.org/licenses)

Materials Cloud Archive – 2nd release

First release March 2017

- Flask, Jinja, SQLAlchemy

Second release May 2020

- Built within the INVENIO framework (v3)
  - developed at CERN,
    Zenodo > 2 million records

- Flask, Jinja, SQLAlchemy, AngularJS, PostgreSQL, Elasticsearch, Celery, Redis, RabbitMQ

Materials Cloud Archive – 2nd release

Extended the Invenio framework

with customized features to address the needs of:

- Users
- Moderators
- Administrators

Materials Cloud Archive – 2nd release

Extended the Invenio framework

with customized features to address the needs of:

- Users
  - Simplified submission procedure
  - Simplified edition of published records
    - change of references and keywords after publication
  - Simplified edition of new versions of records
  - DOI registered before publication
  - Access to records in a personal work area
  - Possibility to easily search for records by keywords or free text
  - Being informed of the status of the record throughout the entire moderation process
- Moderators
- Administrators

Materials Cloud Archive – 2nd release

Extended the Invenio framework

with customized features to address the needs of:

- Users
- Moderators
  - Simplified moderators' tasks:
    - Review
      - ⇒ request changes on submitted records
      - ⇒ track email exchanges in log
    - Approve
      - ⇒ publish records
    - Reject records
- Administrators

Materials Cloud Archive – 2nd release

Extended the Invenio framework

with customized features to address the needs of:

- Users
- Moderators
- Administrators
  - Simplified specific tasks such as:
    - Retract records
    - Add links to the Explore and Discover sections

Materials Cloud Archive – Explore - Discover

materialscloud:2020.0029/v1

In silico discovery of covalent organic frameworks for carbon capture

Kathryn S. Deeg[1], Daiane Damasceno Borges[2], Daniele Ongari[3*], Nakul Rampal[4], Leopold Talirz[3], Aliaksandr V. Yakutovich[3], Johanna M. Huck[1], Berend Smit[3]

1 Department of Chemistry, University of California, Berkeley, California, 94720, United States
2 Instituto de Física, Universidade Federal de Uberlândia, Uberlândia-MG 38408-100, Brasil
3 Laboratory of Molecular Simulation (LSMO), Institut des sciences et ingénierie chimiques (ISIC), Ecole polytechnique fédérale de Lausanne (EPFL) Valais, Rue de l'Industrie 17, 1951, Sion, Switzerland
4 Adsorption and Advanced Materials Laboratory (AAML), Department of Chemical Engineering Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge, CB3 0AS, UK

* Corresponding authors emails: daniele.ongari@epfl.ch

DOI 10.24435/materialscloud:2020.0029/v1 [version v1]
Publication date: Mar 24, 2020, 00:00:00

How to cite this record

Kathryn S. Deeg, Daiane Damasceno Borges, Daniele Ongari, Nakul Rampal, Leopold Talirz, Aliaksandr V. Yakutovich, Johanna M. Huck, Berend Smit, *In silico discovery of covalent organic frameworks for carbon capture*, Materials Cloud Archive **2020.0029/v1** (2020), doi: 10.24435/materialscloud:2020.0029/v1.

Description

We screen a database of more than 69,000 hypothetical covalent organic frameworks (COFs) for carbon capture, using parasitic energy as a metric. In order to compute CO2-framework interactions in molecular simulations, we develop a genetic algorithm to tune the charge equilibration method and derive accurate framework partial charges. Nearly 400 COFs are identified with parasitic energy lower than that of an amine scrubbing process using monoethanolamine; over 70 are better performers than the best experimental COFs; and several perform similarly to Mg-MOF-74. We analyze the effect of pore topology on carbon capture performance in order to guide development of improved carbon capture materials.

Materials Cloud sections using this data

- Interactive interactive scatter plots of 67k COFs and their properties
- Browse the full AiiDA database

Export
Dublin Core JSON

Materials Cloud Archive – Explore - Discover

materialscloud:2020.0029/v1

In silico discovery of covalent organic frameworks for carbon capture

Kathryn S. Deeg[1], Daiane Damasceno Borges[2], Daniele Ongari[3*], Nakul Rampal[4], Leopold Talirz[3], Aliaksandr V. Yakutovich[3], Johanna M. Huck[1], Berend Smit[3]

1 Department of Chemistry, University of California, Berkeley, California, 94720, United States
2 Instituto de Física, Universidade Federal de Uberlândia, Uberlândia-MG 38408-100, Brasil
3 Laboratory of Molecular Simulation (LSMO), Institut des sciences et ingénierie chimiques (ISIC), Ecole polytechnique fédérale de Lausanne (EPFL) Valais, Rue de l'Industrie 17, 1951, Sion, Switzerland
4 Adsorption and Advanced Materials Laboratory (AAML), Department of Chemical Engineering Biotechnology, University of Cambridge, Philippa Fawcett Drive, Cambridge, CB3 0AS, UK

* Corresponding authors emails: daniele.ongari@epfl.ch

DOI 10.24435/materialscloud:2020.0029/v1 [version v1]
Publication date: Mar 24, 2020, 00:00:00

How to cite this record

Kathryn S. Deeg, Daiane Damasceno Borges, Daniele Ongari, Nakul Rampal, Leopold Talirz, Aliaksandr V. Yakutovich, Johanna M. Huck, Berend Smit, *In silico discovery of covalent organic frameworks for carbon capture*, Materials Cloud Archive **2020.0029/v1** (2020), doi: 10.24435/materialscloud:2020.0029/v1.

Description

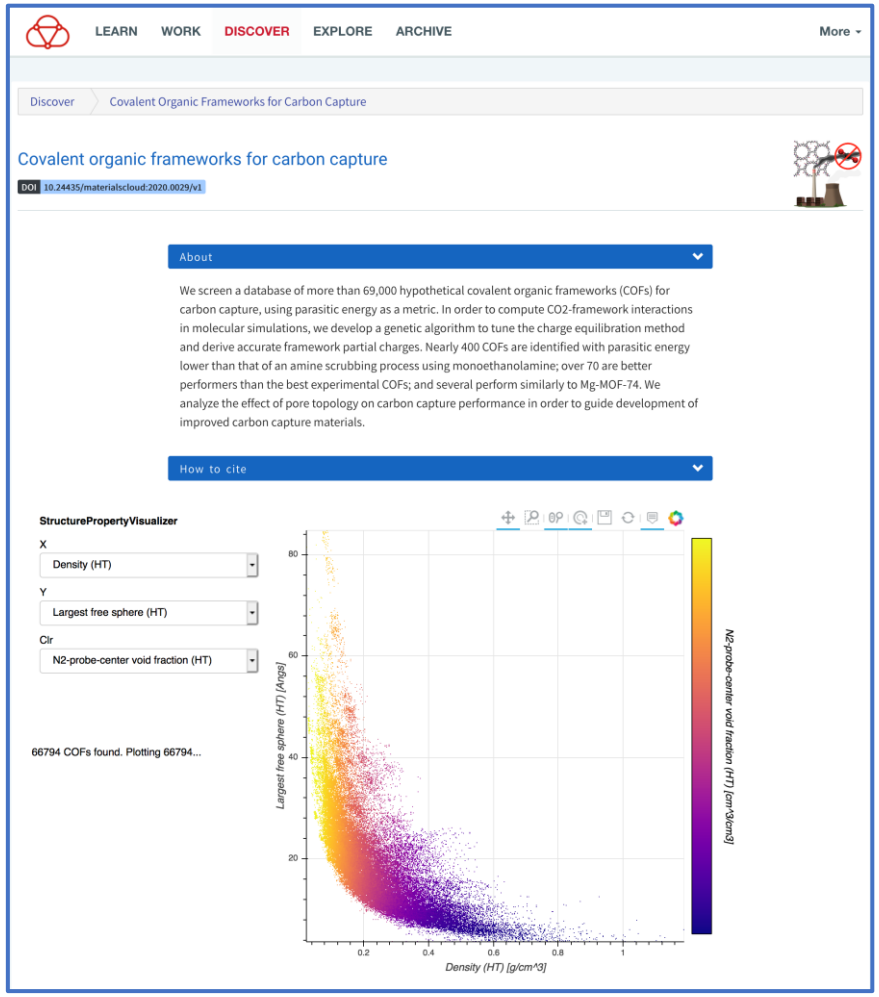
We screen a database of more than 69,000 hypothetical covalent organic frameworks (COFs) for carbon capture, using parasitic energy as a metric. In order to compute CO2-framework interactions in molecular simulations, we develop a genetic algorithm to tune the charge equilibration method and derive accurate framework partial charges. Nearly 400 COFs are identified with parasitic energy lower than that of an amine scrubbing process using monoethanolamine; over 70 are better performers than the best experimental COFs; and several perform similarly to Mg-MOF-74. We analyze the effect of pore topology on carbon capture performance in order to guide development of improved carbon capture materials.

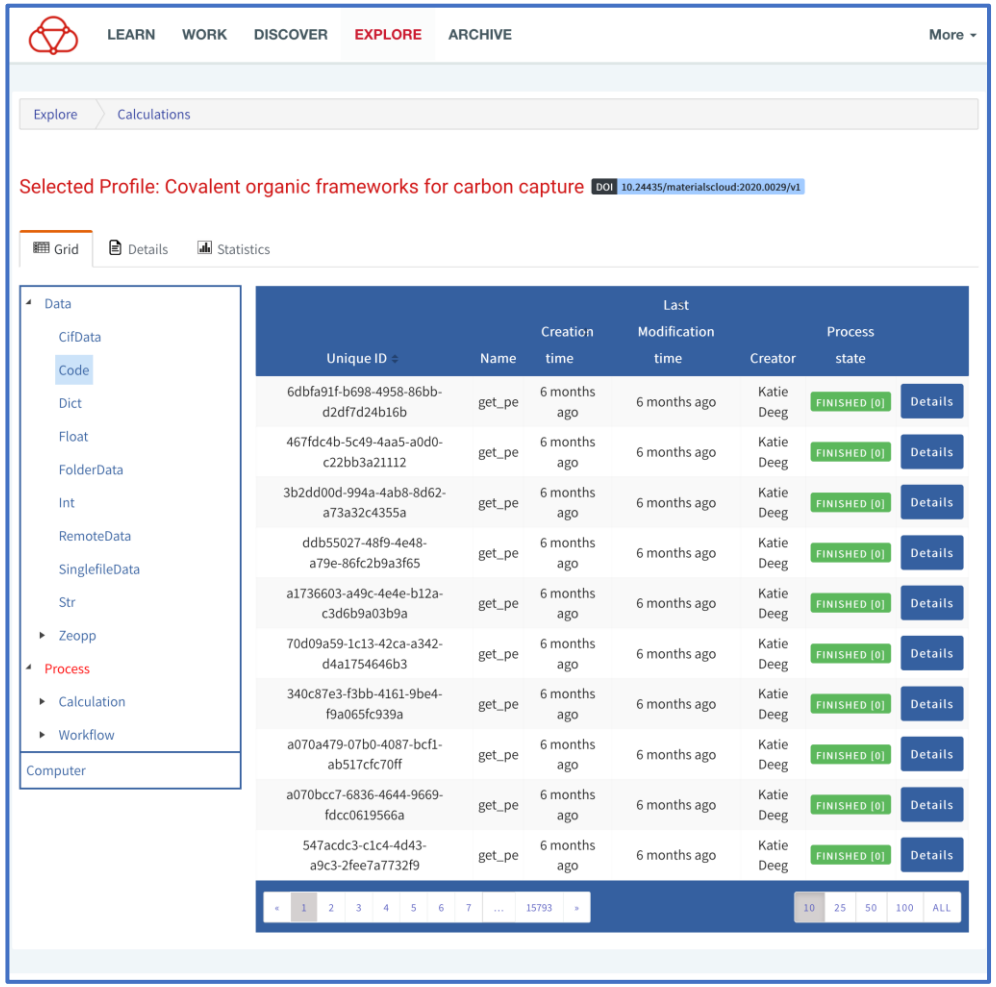
Materials Cloud sections using this data

- Interactive interactive scatter plots of 67k COFs and their properties
- Browse the full AiiDA database

Export
Dublin Core JSON

Materials Cloud Archive – Explore - Discover





Materials Cloud Archive

Video

- How to Login
- The user work space
- How to upload a record
- How to create a new version of a record
- How to change keywords and references
- How to search records
- Where to find help

Latest records

The QMspin data set: Several thousand carbene singlet and triplet state structures and vertical spin gaps computed at MRCISD+Q-F12/cc-pVDZ-F12 level of theory

DOI 10.24435/materialscloud:2020.0051/v1

Max Schwilk, Diana N. Tahchieva, O. Anatole von Lilienfeld

High-quality data sets of free carbenes have remained unavailable in the scientific literature so far. We provide approximately 5k and 8k verified carbene structures in their respective singlet or triplet state. Vertical spin gaps have been computed at higher order multireference level of theory (MRCISD+Q-F12/cc-pVDZ-F12). The carbenes presented are all derived through double hydrogen abstraction from saturated carbon centers of a subset obtained by randomly sampling the chemical space of approximately 300k carbene candidates possible within closed shell organic molecules in the QM9 data set.

Latest version: v1
Publication date: May 08, 2020

Materials Cloud Archive

SCIENTIFIC DATA

re3data.org
REGISTRY OF RESEARCH DATA REPOSITORIES
http://doi.org/10.17616/R3ZJ5W
Materials Cloud Archive

FAIRsharing.org
standards, databases, policies

- Recommended by Nature's journal Scientific Data, by FAIRSharing.org, and listed in RE3DATA.org

- Metadata indexed by Google Dataset search and EUDAT B2FIND

https://datasetsearch.research.google.com/search?query=Materials%20Cloud

http://b2find.eudat.eu/group/materialscloud

DRIVING THE EXASCALE TRANSITION

THANKS

https://archive.materialscloud.org