# Materials science with large-scale data and informatics:

[Materials science with large-scale data and informatics: 1](#_Toc26486873)

[Paper Details 1](#_Toc26486874)

[Prerequisite 1](#_Toc26486875)

[Approach 1](#_Toc26486876)

[Project Details 1](#_Toc26486877)

[Change Log 1](#_Toc26486878)

[Current Situation (As per paper in 2016) 2](#_Toc26486879)

[Computational Material Science vs Material Informatics 2](#_Toc26486880)

[Opportunity 2](#_Toc26486881)

[Importance of processing– structure–properties–performance 2](#_Toc26486882)

[Current Data Landscape 2](#_Toc26486883)

[Very Broad Materials data infrastructure 2](#_Toc26486884)

[Free Databases 3](#_Toc26486885)

[Challenges 3](#_Toc26486886)

[Other Sources 4](#_Toc26486887)

[Open Access 5](#_Toc26486888)

[Smart Manufacturing (lab to Production Scaling) 5](#_Toc26486889)

[Data Standards 5](#_Toc26486890)

## Paper Details

* Materials science with large-scale data and informatics: Unlocking new opportunities, Joanne Hill

## Prerequisite

MaterialScience.pptx

## Approach

## Project Details

Generating-Knowledge-Bases-From-Scientific-Publications

<https://github.com/epochxero/Generating-Knowledge-Bases-From-Scientific-Publications>

## Change Log

|  |  |  |  |
| --- | --- | --- | --- |
| Date | Project Team Member | Description | Version |
| 5th Dec 2019 | Amardeep Singh Sidhu | Initial Version Created | LR002\_V1\_Sci |
|  |  |  |  |

## Current Situation (As per paper in 2016)

### Computational Material Science vs Material Informatics

|  |  |  |
| --- | --- | --- |
| Feature\Field | Computational Material Science | Material Informatics  (Our Focus) |
| Physics based tools.  (No underlying governing equation such as Schrödinger equation) | Yes | No |
| Data Based | No | Yes |
| Examples | Density Functional Theory  Molecular Dynamics  Phase filed simulation  Integrated computational materials engineering (ICME) | Tomographic or high throughput X-ray diffraction. |

### 

ICME: Connects physics-based models to predictively model alloy behaviour. Systems such as atomistic simulation, dislocation modelling, thermodynamic modelling, continuum modelling are connected.

### Opportunity

Material Informatics can complement ICME by

1. Predictively supplying key materials property parameters for underlying ICME models.
2. integrating the outputs of an ICME workflow into higher-level machine learning based models of materials behaviour.

## Importance of processing– structure–properties–performance

## Current Data Landscape

### Very Broad Materials data infrastructure

1. The National Institute of Standards and Technology (NIST) Materials Data Curation System
2. Citrine Informatics’ Citrination platform

### Free Databases

|  |  |  |
| --- | --- | --- |
| Name | URL | Category |
| 3D Materials Atlas | cosmicweb.mse.iastate.edu/wiki/display/home/ | 3D Characterization |
| AFLOWLIB | aflowlib.org | Computational |
| AIST Research Information Databases | www.aist.go.jp/aist\_e/list/database/riodb | General Materials Data |
| American Mineralogist Crystal Structure Database | rruff.geo.arizona.edu/AMS/amcsd.phP | Minerals |
| CatApp | suncat.stanford.edu/catapp | Catalysts |
| Chemspider | www.chemspider.com | Chemical data |
| Citrination | citrination.com | General Materials Data |
| Computational Materials Repository | cmr.fysik.dtu.dk | Computational |
| Crystallography Open Database | http://www.crystallography.net | Crystallography |
| DOE Hydrogen Storage Materials Database | www.hydrogenmaterialssearch.govtools.us | Hydrogen Storage |
| Harvard Clean Energy Project | cepdb.molecularspace.org | Computational |
| Matbase | www.matbase.com | General Materials Data |
| Materials Project | www.materialsproject.org | Computational |
| MatNavi (NIMS) | mits.nims.go.jp/index\_en.html | General Materials Data |
| MatWeb | www.matweb.com | General Materials Data |
| Mindat | www.mindat.org | Minerals |
| NanoHUB | nanohub.org | Nanomaterials |
| Nanomaterials Registry | [www.nanomaterialregistry.org](http://www.nanomaterialregistry.org) | Nanomaterials |
| NIST Materials Data Repository (DSpace) | materialsdata.nist.gov/dspace/xmlui | General Materials Data |
| NIST Interatomic Potentials Repository | [www.ctcms.nist.gov/potentials](http://www.ctcms.nist.gov/potentials) | Computational |
| NIST Standard Reference Data | [www.nist.gov/srd/onlinelist.cfm](http://www.nist.gov/srd/onlinelist.cfm) | General Materials Data |
| NoMaD | nomad-repository.eu/cms | Computational |
| Open Knowledge Database of Interatomic Models  (Open KIM) | openkim.org | Computational |
| Open Quantum Materials Database | oqmd.org | Computational |
| PubChem | pubchem.ncbi.nlm.nih.gov | Chemical data |
| TEDesignLab | [www.tedesignlab.org](http://www.tedesignlab.org) | Thermoelectrics |
| UCSB-MRL thermoelectric database | www.mrl.ucsb.edu:8080/datamine/  thermoelectric.jsp | Thermoelectrics |

### Challenges

1. Need to check which of above data base is downloadable in bulk.
2. Lack of common data standards in these data bases. Example Crystallographic Information File (CIF Format)

Citrine Informatics is working to nucleate grassroots support for a flexible JavaScript Object Notation (JSON)-based material-data format.

Citrine Informatics, “MIF Schema,” available at <https://github.com/CitrineInformatics>

**Materials Project**

* **Pymatgen:** document-based schema-less database, and automated open-source workflow software.
* **Fireworks:** Determine structural, thermodynamic, electronic, and mechanical properties of over 65,000 inorganic compounds by means of high-throughput ab initio calculations. More compounds and properties (e.g., elastic tensors, band structures, dielectric tensors, x-ray diffraction, piezoelectric constants, etc.) are being added on a daily basis

Note: Ab initio quantum chemistry methods are computational chemistry methods based on quantum chemistry.

* **Web Application:** A series of web applications provide users with the capability to perform advanced searches and useful analyses (e.g., phase diagrams, reaction-energy computations, band-structure decomposition, novel structure prediction, Pourbaix diagrams). Calculated results are available by Material API.

**The Open Quantum Materials Database (OQMD)**

The Open Quantum Materials Database (OQMD) 24 , 49 is a high-throughput database currently consisting of ∼ 400,000 DFT total energy calculations of compounds from the ICSD and decorations of commonly occurring crystal structures. OQMD is open (without restrictions) and is online.

* Search Material by composition
* Create phase diagram
* Determine ground state composition
* Determine if equilibrium exists between two phases
* Visualize crystal structure
* Download entire data base for own use

### Other Sources

1. US Materials Genome Initiative (MGI), launched in 2011, to accelerate materials development and commercialization.
2. US Air Force Research Laboratory, NIST, and NSF launched a Materials Science and Engineering Data Challenge to encourage the use of publicly available data to discover or model new material properties.
3. <https://www.mrs.org/open-data-challenge>

### Open Access

The open-access (OA) paradigm, in which readers are able to view and (sometimes) repurpose published research at no cost.

### Smart Manufacturing (lab to Production Scaling)

### Data Standards

MatML (Promoted by NIST, XML Schema)

MatDB(XML Schema)

NMC-MatDB (XML Schema)

Material Information File (JSON)