HybriD³ Database for Hybrid Organic-Inorganic Perovskites

Connor Clayton[1], Xiaochen Du[2], Svenja M. Janke[2], Chi Liu[3], Matti Ropo[4], Volker Blum[2]

[1] Department of Materials Science and Engineering, Carnegie Mellon University, [2] Department of Mechanical Engineering and Materials Science, Duke University, [3] Department of Chemistry, Duke University, [4] Department of Physics and Astronomy, University of Turku, Finland

Motivation

Hybrid organic-inorganic perovskites (HOIPs) have been at the center of a large wave of developments in photovoltaic (PV) devices in recent years. Since their viability in such devices was first demonstrated in 2009 [1], the efficiency of these cells has soared to over 21% in 2018 [2]. This success has sparked an immense surge of research into HOIPs, especially on methylammonium-lead-iodide, or $MAPbI_3$, which is the prototypical example for HOIP-based solar cell absorber materials [3].

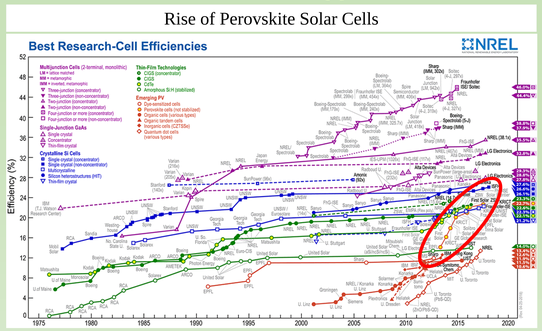


Figure 1: Perovskites are one of the newest and fastest growing types of solar cells [2].

HOIPs hold an enormous number of structural possibilities, especially when considering also lower dimensional HOIPs that allow for larger organic cations. This opens up a large material space with tunable electronic properties (e.g. band gap or level alignment between organic and inorganic) [4], and investigation into this space is producing large amounts of data, while simultaneously generating interest in having easy access to this data (e.g. for data science application). We create a database that that is designed to store the bulk of modern HOIP knowledge in a central location and to make it easily accessible, with the goal of facilitating research into next-generation materials.

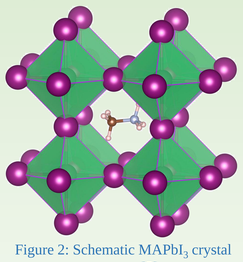


Figure 2: Schematic $MAPbI_3$ crystal structure [5].

Throughout this REU project, $MAPbI_3$ has served as an example for studying the nature of the data available on these materials to better understand how to organize its storage and how to best present it to the user.

Framework

HybriD³ is built using Django [6], a python-based web framework that employs a Model-Template-View (MTV) architecture.

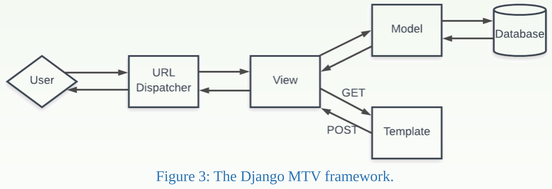


Figure 3: The Django MTV framework.

Models: explain the organization of data and are used in conjunction with the database itself to write SQL statements that dictate its structure, including all fields and connections to other models.

URL dispatcher: responsible for calling the appropriate view based on user actions.

Template: contains the HTML and CSS that is rendered and presented to the user.

View: retrieves necessary data from the model, then performs a GET request that calls the appropriate template and supplies it with this data. Then, info from the template is returned via a POST request, which tells the model to modify data as necessary.

Challenges

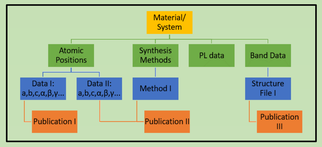


Figure 4: Original database structure — only one author associated with each publication.

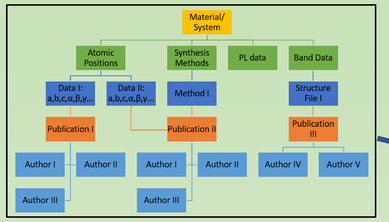


Figure 5: Intermediate solution — author data duplicated.

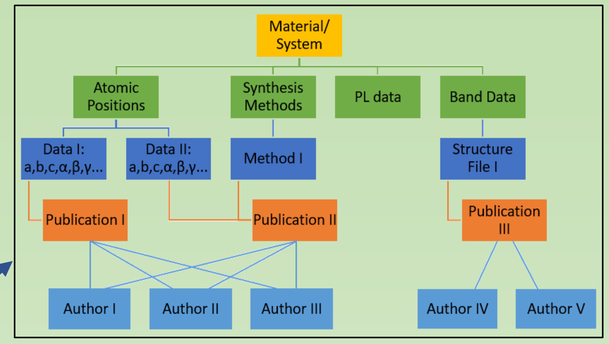


Figure 6: Solution — many-to-many relationship between authors and publications.

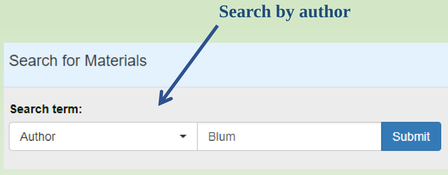
Data Accessibility

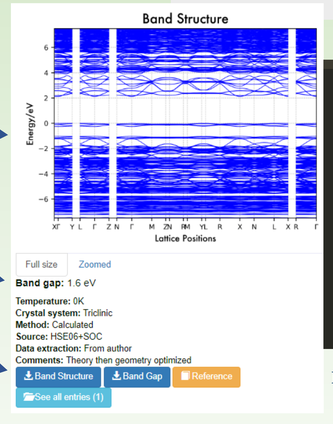
Alternate names aid searchability

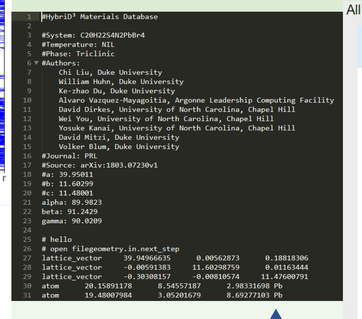
Search by author

Band structures and band gap data available and downloadable independently

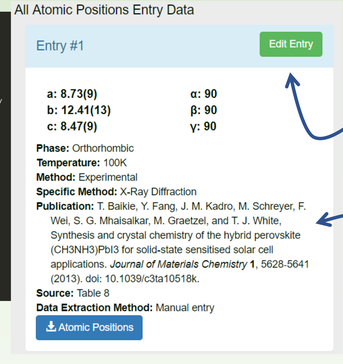
Compound Name	Chemical Formula	Alternate Names	Organic Component	Inorganic Component
MAPbI3	(CH3NH3)PbI3	MAPI (MA)PbI3 CH3NH3PbI3	CH3NH3	PbI3







Example of downloaded atomic positions file



Entries editable by their creators directly from within website

Clear, consistent reference format

Data Insertion

Figure 8: When adding information on a publication, authors are entered on a single page, with no need for the user to match against existing entries. This is an improvement over searching for every author.

Many-to-many relationship

Extract author information from form

Locate pre-existing authors in database

Reuse authors if possible

Create a new author if necessary

```
class Author(models.Model):
    publication = models.ManyToManyField(Publication)
# create and save new author objects, linking them to the saved publication
for i in range(author_count): # for each author
    data = {}
    data['first_name'] = authors_info["form-%d-first_name" % i]
    data['last_name'] = authors_info["form-%d-last_name" % i]
    data['institution'] = authors_info["form-%d-institution" % i]
    preexistingAuthors = (Author.objects.filter(
                          first_name__iexact=data['first_name']).filter(
                          last_name__iexact=data['last_name']).filter(
                          institution__iexact=data['institution']))

    if preexistingAuthors.count() > 0:
        # use the prexisting author object
        preexistingAuthors[0].publication.add(newPub)

    else: # this is a new author, so create a new object
        author_form = AddAuthor(data)

        if(not author_form.is_valid()):
            ...
        else: # author_form is valid
            form = author_form.save()
            form.publication.add(newPub)
            form.save()
```

Figure 9: Code that prevents duplication of author data. If an author already exists in the database, he/she is matched to the new publication.

HybriD³ database

materials.hybrid3.duke.edu/

Data Storage

The HybriD³ **hierarchical database structure** allows for storage of a complex network of information, as well as accessibility of data via multiple searchable keywords. For example, relevant system data can be accessed from an author and vice versa.

Figure 7: MariaDB is the database server behind HybriD³ [8].

Reuse of publication and author information **minimizes server-side storage** and improves lookup speeds. These entities share a **many-to-many relationship**, i.e. one publication can have many authors, and one author can write many publications.

Conclusion

Considerable progress was made during this phase of the Hybrid³ database project. Major improvements in data traceability were made, notably added support for variable numbers of authors and improved reference information display. The usability of the website has also been enhanced, allowing easy inclusion of author information, convenient options to edit existing data, a more robust layout describing synthesis methodology, support for band gap information, and the ability to search systems by author and alternative compound names.

In the future, the database will be extended to include more key properties relevant for prospective new HOIP materials (e.g. conductivity, stability to environmental conditions, exciton energy) while its user-friendliness will be further increased. Possible future directions include the construction of an API (applicational programming interface) for connection to related projects. We are optimistic that the database will inspire extensive flow of data and prove a useful tool for the hybrid organic-inorganic perovskite community.

References

1. A. Kojima, K. Teshima, Y. Shirai, T. Miyasaka. *Journal of the American Chemical Society*, **131**, 6050-6051, (2009). doi: 10.1021/ja809598r.
2. "Efficiency Chart." https://www.nrel.gov/pv/assets/images/efficiency-chart.png. Accessed 7/16/2018. This plot is courtesy of the National Renewable Energy Laboratory, Golden, CO.
3. P.P. Boix, K. Nonomura, N. Mathews, S.G. Mhaisalkar. *Materials Today*, **17**, 16-23, (2014). doi:10.1016/j.mattod.2013.12.002
4. B. Saparov & D.B. Mitzi. *Chem. Rev.* **116**, 4558−4596, (2016). doi:10.1021/acs.chemrev.5b00715.
5. C. Eames, J.M. Frost, P.R.F. Barnes, B.C. O'Regan, A. Walsh, M.S. Islam. *Nature Communications*, **6**, 7497, (2015). doi:10.1038/ncomms8497.
6. Django (Version 1.10) [Computer Software]. (2017). Retrieved from https://djangoproject.com.
7. MariaDB Foundation, mariadb.org/. Accessed 7/12/2018.

Acknowledgements

This work was financially supported by the NSF under award number DMR-1728921. We thank Dr. Raul Laasner and Prof. Dr. David Mitzi for fruitful discussions, as well as the REU for Meeting the Grand Challenges for the opportunity to conduct this research. Svenja Janke gratefully acknowledges funding by the German Research Foundation (DFG) JA 2843/1-1.