Plan Overview

A Data Management Plan created using DMPonline

Title: Harvesting waste heat as electrical power: Theory-led control of heat transport in thermoelectrics

Creator: Jonathan Skelton

Principal Investigator: Jonathan Skelton

Data Manager: Jonathan Skelton

Affiliation: University of Manchester

Template: University of Manchester Generic Template

ORCID iD: 0000-0002-0395-1202

Project abstract:

In May 2019 the UK declared a climate emergency, highlighting the urgent need for cleaner energy sources to alleviate global warming. 60 % of global energy is wasted as heat, making thermoelectric (TE) power vital to a successful clean energy strategy. TEs enable wasted energy from transport, industrial processes and other low-grade heat sources to be recovered as electricity. Despite a burgeoning $1bn global market by 2024, current flagship TE materials are not commercially viable due to limited conversion efficiencies and compositions based on rare or toxic elements. High-performance TEs are heavily-doped semiconductors with high electrical conductivity and low lattice thermal conductivity. Fine-tuning electrical properties by doping is a robust process, but developments in TEs are held back by poor understanding of thermal conductivity and how to control it, with recent advances relying on empirical strategies and chance discovery. The majority of heat transport in semiconductors is through lattice vibrations (phonon modes), and the limiting factor is the anharmonic processes that lead to phonon scattering and suppress the mode lifetimes. Recent advances in theoretical modelling have made it possible to simulate the lattice dynamics and thermal transport from first principles and to identify specific groups of phonon modes that transport heat or suppress transport by scattering other modes. This project will build on this work to explore the challenge of sustainable energy across chemistry, physics and materials science by developing a deeper understanding of heat transport and strategies for controlling it, applying them to improve current flagship TEs, and designing from the “bottom up” a new class of hybrid organic/inorganic TEs.

ID: 44808

Last modified: 13-01-2021

Grant number / URL: MR/T043121/1MRC

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan

as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Harvesting waste heat as electrical power: Theory-led control of heat transport in thermoelectrics

Manchester Data Management Outline

1. Will this project be reviewed by any of the following bodies (please select all that apply)?

- Funder

2. Is The University of Manchester collaborating with other institutions on this project?

- No - only institution involved

3. What data will you use in this project (please select all that apply)?

- Acquire new data

4. Where will the data be stored and backed-up during the project lifetime?

- University of Manchester Research Data Storage Service (Isilon)

5. If you will be using Research Data Storage, how much storage will you require?

- 1 - 8 TB

6. Are you going to be working with a 3rd party data provider?

- No

7. How long do you intend to keep your data for after the end of your project (in years)?

- 5 - 10 years

Questions about personal information

Personal information, also known as personal data, relates to identifiable living individuals. Special category personal data is more sensitive information such as medical records, ethnic background, religious beliefs, political opinions, sexual orientation and criminal convictions or offences information. If you are not using personal data then you can skip the rest of this section.

Please note that in line with data protection law (the General Data Protection Regulation and Data Protection Act 2018), personal information should only be stored in an identifiable form for as long as is necessary for the project; it should be pseudonymised (partially de-identified) and/or anonymised (completely de—identified) as soon as

practically possible. You must obtain the appropriate ethical approval in order to use identifiable personal data.

8. What type of personal information will you be processing (please select all that apply)?

- No sensitive or personal data

9. Please briefly outline how you plan to store, protect and ensure confidentiality of the participants' information.

N/A.

10. If you are storing personal information (including contact details) will you need to keep it beyond the end of the project?

- Not applicable

11. Will the participants' information (personal and/or sensitive) be shared with or accessed by anyone outside of the University of Manchester?

- Not applicable

12. If you will be sharing personal information outside of the University of Manchester will the individual or organisation you are sharing with be outside the EEA?

- Not applicable

13. Are you planning to use the personal information for future purposes such as research?

- No

14. Who will act as the data custodian for this study, and so be responsible for the information involved?

Jonathan Michael Skelton

15. Please provide the date on which this plan was last reviewed (dd/mm/yyyy).

2020-02-14

Project details

What is the purpose of your research project?

This is a computational modelling project targeting an improved understanding of heat transport in thermoelectric materials and engineering strategies for controlling it, with the end goal of designing a new class of high-performance thermoelectrics for recovering waste heat as electricity.
The three program work packages aim to:
(1) Model the effect of chemical substitution on the thermal conductivity of current flagship thermoelectrics, develop a machine-

learning method to speed up large calculations, and generalise design rules for selecting dopants to optimise heat transport;
(2) Develop a method for quantifying the impact of "soft-mode" dynamics on material properties, apply it to understand the effect of soft-mode dynamics in several classes of thermoelectrics, and identify engineering strategies to induce or suppress this type of anharmonicty;
(3) Apply theoretical methods to design a new class of high-performance thermoelectrics based on hybrid organic/inorganic "rattler" materials incorporating molecular guest ions.

What policies and guidelines on data management, data sharing, and data security are relevant to your research project?

The main guidelines are provided by the University of Manchester Research Data Management Policy: http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=33802%20

Other UoM policies that apply are:

- The University of Manchester Records Management Policy http://documents.manchester.ac.uk/display.aspx?DocID=14916
- The University of Manchester Publications Policy http://documents.manchester.ac.uk/display.aspx?DocID=28526
- The University of Manchester IT policies and guidelines http://www.itservices.manchester.ac.uk/aboutus/policy/

- The University of Manchester Data Protection Policy http://documents.manchester.ac.uk/display.aspx?DocID=14914

In addition to this, the Royal Society provides the following guidelines:
The Society supports science as an open enterprise, and is committed to ensuring that data outputs from research supported by the Society are made publically available in a managed and responsible manner, with as few restrictions as possible. Data outputs should be deposited in an appropriate, recognised, publically available repository, so that others can verify and build upon the data, which is of public interest. To fully realise the benefits of publically available data they should be made intelligently open by fulfilling the requirements of being discoverable, accessible, intelligible, assessable and reusable.

Responsibilities and Resources

Who will be responsible for data management?

The PI will assume overall responsitbility for data management and will delegate to other researchers (e.g. PhD students and PDRAs) working on the project when appropriate.

What resources will you require to deliver your plan?

Delivering the plan will require access to the UoM Research Data Storage Service (< 8 Tb) and the Mendeley Data (https://data.mendeley.com), GitHub (https://github.com) and Zenodo services (https://zenodo.org). No additional hardware or equipment, nor technical support or staff, will not be required.

Data Collection

What data will you collect or create?

Primary data will take the form of code implementations of theoretical methods developed during the project and data from electronic-structure and lattice-dynamics calculations. Codes will mostly be written in the Python programming language, and calculation data will be in plain-text formats (free text, XML, YAML) or standard binary formats such as HDF5.
The raw data from a complete set of electronic-structure calculations for a single material typically require around 1-2 Gb of storage space compressed, and post-processing will generate from 5-25 Gb of data depending on the analysis being performed.

How will the data be collected or created?

Codes will mostly be written in the Python programming language. Electronic-structure calculations will be performed using the Vienna Ab initio Simulation Package (VASP) software, and lattice-dynamics calculations will be set up and post-processed using the open-source Phonopy and Phono3py packages (https://atztogo.github.io/phonopy; https://atztogo.github.io/phono3py).

The quality and consistency of data collection will be controlled as follows:

1. Through preliminary benchmarking of the simulation methodology used to generate data (e.g. against available experimental data);
2. Through peer review of the methodology, generally during the publication process; and
3. By developing standardised workflows for data processing to ensure consistency.

The quality of codes written as part of the project will be ensured as follows:

1. Developing codes as open source packages hosted in public code repositories that allow other users to audit the code and submit bug reports and/or feature requests; and
2. Where appropriate, obtaining independent code reviews of release software e.g. by submission to the Journal of Open Source Software (JOSS; https://joss.theoj.org/)

Documentation and Metadata

What documentation and metadata will accompany the data?

Data will be published in online repositories alongside detailed README files detailing the contents and formats of the data and explaining how it was generated and processed. Published datasets will include input files for the generating code(s) to enable other researchers to reproduce the results.

Codes will be documented with detailed README files, including examples, to outline the methods implemented and to illustrate the intended use cases. This will be supplemented by appropriate documentation and inline commenting on the code itself.

These both follow the current informal best practices in the field, and will be revised in line with any developments that occur as the project progresses.

Ethics and Legal Compliance

How will you manage any ethical issues?

No ethical issues are applicable to this project.

How will you manage copyright and Intellectual Property Rights (IPR) issues?

Copyright and IPR issues will be managed in line with the University of Manchester Intellectual Property policy: http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=24420

Storage and backup

How will the data be stored and backed up?

Data will be stored using the Research Data Storage service provided by the University of Manchester IT Services. Datasets associated with published outputs will be archived to the Mendeley Data service and made available indefinitely with a persistent digital object identifier (DOI).

How will you manage access and security?

It is not expected that confidential data will be generated during this project. Access to data from within UoM (e.g. by staff and student collaborators) will be managed through the Research Data Service. Access to data by external collaborators will be managed through the secure sharing service provided by Dropbox Business.

Selection and Preservation

Which data should be retained, shared, and/or preserved?

Codes will be retained, shared and preserved with full version histories using the GitHub service (https://github.com), with periodic snapshots archived using Zenodo (https://zenodo.org).

Electronic-structure calculations produce large quantities of intermediate output files that consume considerable storage space and can recreated relatively straightforwardly from a smaller set of input files. Datasets comprising processed output will therefore be retained, shared and preserved alongside the input files and documentation needed to reproduce the intermediate output files if required.

Output from the lattice-dynamics calculations is generated by post-processing electronic-structure calculations. Small data files and/or large files that are non-trivial to recreate will be retained, shared and preserved, whereas large files that are relatively east to recreate will be omitted and the input files needed to recreate them, with documentation, included in their place.

What is the long-term preservation plan for the dataset?

Curated datasets made available with published results through the Mendeley Data service (https://data.mendeley.com) will be available indefinitely with a persistent digital object identifier (DOI).

Codes will be made available indefinitely, with version history, using the GitHub service (https://github.com). Periodic snapshots will be taken using Zenodo (https://zenodo.org) and archived with published results, and these will be accessible indefinitely with a persistent DOI.

Data Sharing

How will you share the data?

Complete datasets, including all input files required for other researchers to reproduce calculations, will be made available alongside published results using the Mendeley Data service (https://data.mendeley.com). This will make the data accessible, indefinitely, with a persistent digital object identifier (DOI).

Codes will be developed and released using the GitHub code-sharing service (https://github.com). The Zenodo service (https://zenodo.org) will be used to archive snapshots of the current versions of the codes when results obtained using them are published; these snapshots will be available indefinitely through a persistent DOI.

Are any restrictions on data sharing required?

The default position will be to make all data available without restriction. If new materials or processes identified during the research are deemed to have potential to generate IPR, availability of the data will be restricted with an embargo if necessary.