UKACM 2024 Conference, 10-12 April 2024, Durham University, Durham, UK
https://doi.org/10.62512/conf.ukacm2024.025

METAMATERIALS GENOME: PROGRESS TOWARDS A COMMUNITY TOOLBOX FOR AI METAMATERIALS DISCOVERY

Jacob Earnshaw[1], Nicholas Syrotiuk[1], Oliver Duncan[2], Lukasz Kaczmarczyk[3], Fabrizio Scarpa[4], Stefan Szyniszewski[1]*

[1] Department of Engineering, Durham University, Durham, DH1 3LE, UK.
stefan.t.szyniszewski@durham.ac.uk
[2] Manchester Metropolitan University
[3] University of Glasgow
[4] University of Bristol

Abstract. Understanding the limits of the design space is a key aspect in optimising complex hierarchical structures and is vital for exploring and designing novel Metamaterials. Simultaneously, abundant data (mostly text, images, and location) aggregated by multinational corporations accelerated the development of machine learning and artificial intelligence technologies. Although increasingly conceptually advanced, the origins of machine learning can be traced back to traditional statistical methods and data-centric analysis. These techniques have been used in fields where establishing relationships and using differential equations or closed-form descriptions have been challenging due to the systems' complexity. However, well-established and validated physics-based modelling tools offer direct solutions for various physical domains relevant to metamaterials. What is the right place for the emerging machine learning techniques in that context?

Key words: *metamaterials; open research commons, ai, machine learning, mechanical, optimisation*

1 Introduction

Fundamentally, numerous research groups and companies repeatedly carry out physics-based calculations, all employing the same fundamental equations such as Maxwell equations or wave or dynamic equations of motion. Therefore, one can recognise a potential for a large collection of results, which are, however, currently scattered across various research groups and frequently guarded in private data silos. In many instances, the results are discarded at the completion of specific projects or archived for a limited time. As recognised by several US federal agencies [1], [2], [3], creating an open research commons and appropriate standards and infrastructure for sharing and reusing as well as recycling existing data and computational tools can accelerate the development of novel materials, reduce effort duplication among different stakeholders, and accelerate the overall progress of the discipline as a whole.

2 Current and Future Challenges and Opportunities

In order to create knowledge graphs that can integrate multiple physical domains like mechanical, optical, electromagnetic, etc., it is necessary to establish common data standards [4]. This requires input and agreement from both academic and industrial groups. Moreover, these standards may change over time. Therefore, it is essential to establish methods for migrating and adapting data.

Shared repositories of knowledge require the trust of all stakeholders [5], sustainable funding, and a transparent governance structure. To achieve satisfactory cyber security resiliency, such a repository might require distributed storage across academic and industrial stakeholders as well as specialist cybersecurity support. Shared repositories of knowledge require the trust of all stakeholders, sustainable funding, and a transparent governance structure. Low barriers for adaptation and usage should facilitate broader benefits and adaptation of the start knowledge by the widest possible group of stakeholders. To achieve satisfactory cyber security resiliency, such a repository might require distributed storage across academic and industrial stakeholders.

Optimising the process of organising and utilising data that already exists can save time and resources for all stakeholders. Additionally, the establishment of a shared open research commons can lead to improved collaboration and knowledge mapping by linking research groups based on similarity and proximity metrics of the structures they produce. By enabling search engines for specific required properties, manufacturers and industrial stakeholders can quickly identify the key research groups and technology suppliers in the design space.

Shared data from 'adjacent' problems can be used to inform and suggest solutions for the problem being studied (also in different physical domains). Since Metamaterial problems solve a limited number of physical laws, transfer learning provides significant benefits in this area and suggests good starting points for developing Metamaterials with specific properties. This can be achieved by leveraging prior iterations and the experience of other academic groups and stakeholders. Transverse learning can be facilitated by using approximate surrogate models that are trained on prior data. Such fast-running surrogate models are also vital for optimisation and can underpin AI generative design, which could suggest promising areas for novel material discoveries.

An example of an unsolved problem in the field is the development of a compact topology parametrisation that is compatible with machine learning tools. Such encoding should be able to generate any suitable Metamaterial topologies without being restricted to a specific subclass of structures. Another challenge is to incorporate manufacturability, cost and sustainability aspects into the optimisation process and generative capabilities of the shared open research commons.

3 Advances in Methods and Techniques to Meet Challenges

The National Institute of Standards and Technologies (NIST) has developed an open-source platform called the Configurable Data Curation System (CDCS) [6] that can curate complex data structures. This platform is designed with cybersecurity in mind and is continuously updated and supported by NIST. We have used this platform to create a prototype of a shared database of metamaterials, which is accessible online at www.meta-genome.org. The database can handle complex hierarchical XML data structures, including topology information in discrete voxel formats and more parametric, vector-like descriptions such as step formats. The open-source code can also be cloned, and independent repositories can be created while still leveraging shared data standards and programmatic developments in toolkits for working and processing vast amounts of data. This is particularly relevant in the context of recent developments in the computer science community, where tools like VQGAN [7] are being used to describe topologies. These tools hold promise in establishing critical links between physical information in the Metamaterials community and rapidly developing AI toolboxes with impressive generative capabilities demonstrated for other material domains [8]. AI generative demonstrations have already been demonstrated in the context of mechanical metamaterials [9], [10].

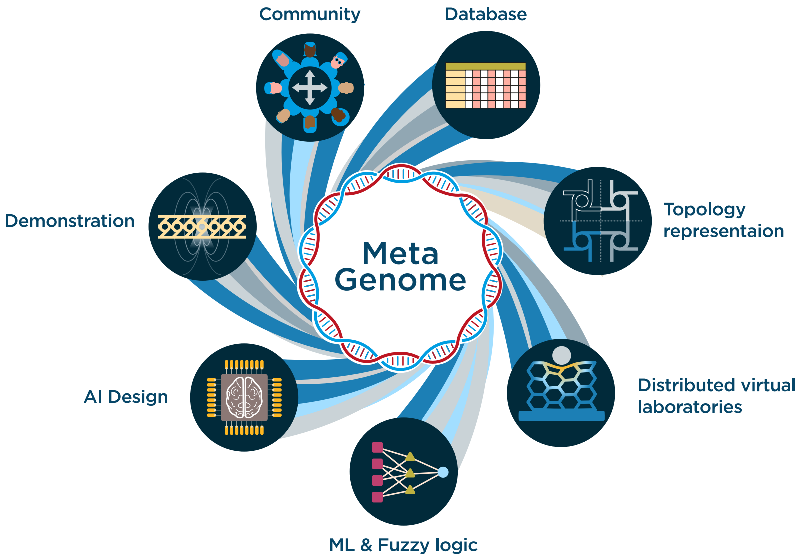

Figure 1: Synergy between physics-based simulations, data-centric analysis, community and developers is depicted as a circular concept.

4 Conclusions

Through the establishment of a shared open research commons, collaboration and knowledge mapping can be improved by utilising property-based norms and metrics. This will enable the community to quickly identify key research groups, manufacturers, and industrial stakeholders in the design space by searching for specific required properties. Secondly, it will enable researchers to focus on the most interesting and high-value activities, minimising duplication of effort and leveraging prior material discoveries. Thirdly, a standardised shared data platform lays the foundation for experimentation with and development of generative AI platforms to augment and enhance future meta-material discoveries.

Acknowledgments

We would like to acknowledge financial support from the Royce Institute for the Mechanical Meta-Materials Database (MCAP068) project, funded under the Materials Challenge Accelerator Programme (MCAP).

REFERENCES

[1] P. E. Bourne et al., Playing catch-up in building an open research commons, *Science*, vol. 377, **6603**:256-–258, Jul. 2022.

[2] A. Dima et al., ‘Informatics Infrastructure for the Materials Genome Initiative’, *JOM*, vol. 68, **8**:2053-–2064, Aug. 2016, doi: 10.1007/s11837-016-2000-4.

[3] National Science and Technology Council and Committee on Technology Subcommiitee on the Materials Genome Initiative, ‘Materials Genome Initiative Strategic Plan’, Department of Defence, National Science Foundation, Office of Science and Technology Policy and National Insti-

tute of Standards and Technology. https://www.mgi.gov/sites/default/files/documents/MGI-2021-Strategic-Plan.pdf

[4] Henry Royce Institute, 'Materials 4.0: A role for standards', Mar. 2021. Accessed: Aug. 24, 2022. https://www.royce.ac.uk/content/uploads/2021/06/Materials-4.0-A-Role-for-Standards.pdf

[5] Henry Royce Institute, 'Enabling Trust in Data Exchange for Materials 4.0', Jul. 2021. Accessed: Aug. 24, 2022. https://www.royce.ac.uk/content/uploads/2021/06/Materials-4.0-Enabling-Trust-in-Data-Exchange-Materials.pdf

[6] NIST, 'Configurable Data Curation System (CDCS)'. Accessed: Mar. 06, 2024. https://www.nist.gov/programs-projects/configurable-data-curation-system-cdcs

[7] J. Yu et al., 'Vector-quantized Image Modeling with Improved VQGAN'. *arXiv*, Jun. 04, 2022. doi: 10.48550/arXiv.2110.04627.

[8] B. A. Koscher et al., 'Autonomous, multiproperty-driven molecular discovery: From predictions to measurements and back', *Science*, vol. 382, no. 6677, p. eadi1407, Dec. 2023, doi: 10.1126/science.adi1407.

[9] A. J. Lew and M. J. Buehler, 'Single-shot forward and inverse hierarchical architected materials design for nonlinear mechanical properties using an Attention-Diffusion model', *Mater. Today*, **64**:10-–20, Apr. 2023, doi: 10.1016/j.mattod.2023.03.007.

[10] J.-H. Bastek and D. M. Kochmann, 'Inverse-design of nonlinear mechanical metamaterials via video denoising diffusion models'. *arXiv*, May 31, 2023. Accessed: Jun. 03, 2023. http://arxiv.org/abs/2305.19836