Digital Discovery

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Check for updates

Cite this: DOI: 10.1039/d5dd00061k

Large language models for material property predictions: elastic constant tensor prediction and materials design†

Siyu Liu,[ab] Tongqi Wen,*[ab] Beilin Ye,[a] Zhuoyuan Li,[ab] Han Liu,[c] Yang Ren[c] and David J. Srolovitz*[ab]

Received 12th February 2025
Accepted 19th May 2025

DOI: 10.1039/d5dd00061k

rsc.li/digitaldiscovery

Efficient and accurate prediction of material properties is critical for advancing materials design and applications. Leveraging the rapid progress of large language models (LLMs), we introduce ElaTBot, a domain-specific LLM for predicting elastic constant tensors and enabling materials discovery as a case study. The proposed ElaTBot LLM enables simultaneous prediction of elastic constant tensors, bulk modulus at finite temperatures, and the generation of new materials with targeted properties. Integrating general LLMs (GPT-4o) and Retrieval-Augmented Generation (RAG) further enhances its predictive capabilities. A specialized variant, ElaTBot-DFT, designed for 0 K elastic constant tensor prediction, reduces the prediction errors by 33.1% compared with a domain-specific, materials science LLM (Darwin) trained on the same dataset. This natural language-based approach highlights the broader potential of LLMs for material property predictions and inverse design. Their multitask capabilities lay the foundation for multimodal materials design, enabling more integrated and versatile exploration of material systems.

Property data are essential for determining the suitability of materials for specific applications. For example, flexible electronics require materials with targeted elastic stiffness,[1] thermal management systems rely on materials with sufficiently high thermal conductivity,[2] and electronic devices depend on materials with appropriate band structure.[3] Given the diverse property profile required for individual applications, deep understanding or at least robust property prediction would be a great aid to material selection and/or alloy design. While the former is a long-term goal of materials science, high-throughput material property measurements and prediction, coupled with new techniques in artificial intelligence (AI) have the potential to revolutionize materials development in the short term.

While experimental approaches for determining materials properties remains the gold standard, they are often hindered by expense and the time required to synthesize materials and measure properties (and, at times, lead to results that are either inconsistent or not sufficiently accurate), such as in the case of elastic constant measurements.[4] Recent advancements in simulation techniques and computational power have made computational modeling a critical tool for property prediction. Given the diversity of length and time scales that control material properties, multi-scale modeling has emerged as an often efficient and sufficiently accurate approach for materials property prediction. For example, atomistic simulations with quantum-mechanical accuracy can accurately predict the full elastic constant tensors and/or band gaps (using hybrid exchange–correlation functionals), while phase-field modeling and other continuum based-methods enable microstructure evolution and defect property prediction. However, challenges (*e.g.*, data transfer and error propagation) often remain significant obstacles to achieving accurate, macroscopic predictions in multi-scale modeling frameworks. The emergence of large language models (LLMs) presents a new opportunity for materials property prediction, with the potential to close the gaps between experiment data (*e.g.*, sourced from literature databases) and computational materials simulation approaches.[5]

LLMs, for example ChatGPT, have demonstrated some remarkable successes across a wide range of materials applications, including high-throughput discovery of physical laws,[6] generation of metal–organic frameworks (MOFs),[7] design of chemical reaction workflows,[8] determining crystal structure (CIF, crystallographic information file),[9] electron microscopy image analysis,[10] and guiding automated experiments.[11] LLMs achieve this by leveraging their capabilities such as rapid literature summarization,[12] prompt engineering[13] and/or

[a]Center for Structural Materials, Department of Mechanical Engineering, The University of Hong Kong, Hong Kong SAR, China. E-mail: tongqwen@hku.hk; srol@hku.hk

[b]Materials Innovation Institute for Life Sciences and Energy (MILES), HKU-SIRI, Shenzhen, China

[c]Department of Physics, JC STEM Lab of Energy and Materials Physics, City University of Hong Kong, Hong Kong SAR, China

† Electronic supplementary information (ESI) available: ESI Notes, ESI Fig. S1–S9, ESI Tables S1–S8. See DOI: https://doi.org/10.1039/d5dd00061k

© 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

integration with external tools.[14] This approach can make them superior to traditional machine learning (ML) models, particularly when dealing with complex and multitask processes at scale.[15] One major strength of LLMs is their foundation in natural language-based training, fine-tuning, and application, which lowers the barrier to entry for researchers without a strong background in computer science or coding.[16] Moreover, underlying pre-trained models encode extensive materials science knowledge, giving LLMs remarkable ability in cases where datasets are sparse (through transfer learning), an achievement that previously required highly specialized algorithms.[17]

Given the strong performance of LLMs on a wide range of low-dimensional classification and regression tasks in computer science,[18] there is growing interest in leveraging LLMs to improve numerical property prediction in materials science. Recent studies, including LLM-Prop,[19] CrysMMNet,[20] and AtomGPT,[21] illustrate two major strategies. LLM-Prop and CrysMMNet introduce architectural modifications of LLMs followed by fine-tuning, whereas AtomGPT preserves the original LLM architecture. Despite their methodological differences, all three approaches convert crystal structures into text descriptions, and fine-tune the LLMs to predict individual material properties such as the band gap, formation energy, or bulk modulus. These studies demonstrate that text-based encoding of structural information can enhance predictive accuracy. Recent studies examined the impact of prompt design on LLM property prediction performance,[22] and have benchmarked LLM-based methods against conventional models on out-of-distribution datasets.[23] These comparisons highlight the value of prompt design for optimizing LLM materials property prediction performance. Although the aforementioned works prove that LLMs can outperform traditional models in predicting certain scalar properties, there are also contrary results, especially when faced with small datasets.[24] For example, Jablonka *et al.*[25] showed that while LLMs can predict properties like HOMO–LUMO gaps, solubility, photoswitching behavior, solvation free energies, and photoconversion efficiency, the results were no better than with traditional ML models. Enhancing the quantitative prediction capabilities of LLMs, while leveraging their strengths in natural language interaction and multitasking, can significantly expand their potential in materials science applications.

In this work, we focus on predicting the elastic constant tensor as a case study of quantitative prediction of a material property. The elastic constant tensor is a fundamental property that describes the elastic response of materials to external forces[26] and serves as a indicator of the nature of intrinsic bonding within a material.[27] Mechanical (Young's modulus, Poisson's ratio,…), thermal (thermal conductivity), and acoustic (sound velocity) properties can all be derived starting from the elastic constant tensor[28] (often together with other basic material properties). Here, we introduce ElaTBot and ElaTBot-DFT (DFT is quantum mechanical density functional theory), LLMs developed through prompt engineering and knowledge fusion training. ElaTBot is designed to predict elastic constant tensors, bulk modulus at finite temperatures, and propose materials with specific elastic properties.

To our knowledge, ElaTBot is the first model capable of directly and efficiently predicting the full elastic constant tensor at finite temperatures. ElaTBot-DFT, a variant specialized for 0 K elastic constant tensor prediction, reduces prediction error by 33.1% compared to the material science LLM Darwin[29] using the same training and test sets. These results highlight the potential of LLMs for numerical materials property predictions.

Training specialized LLMs

Despite the importance of elastic constant tensors in materials science, complete elastic constant tensor data for inorganic crystals remains scarce due to experimental and computational limitations. Fig. 1(a) shows that elastic constant data is scarce in the Materials Project;[30] ∼7.9% as abundant as crystal structure data and ∼17.2% as abundant as band structure data. The elastic constants are a fourth-rank tensor with as many as 21 independent components (in static equilibrium) which is often represented as a symmetric 6 × 6 Voigt matrix C_{ij}.[31] Predicting these components is far more complex than predicting scalar properties, such as formation energy, free energy, or bulk modulus. Fig. 1(b) lists a few ML approaches for predicting elastic constant tensors. Chemical composition-based upon element descriptors were used to predict specific components or C_{ij},[32,33] but not the full elastic constant tensor. More recent models that leverage crystal structure descriptors to predict the full elastic constant tensor face challenges such as long training times and complex model architectures.[28,34] These models are often restricted to single-task predictions of elastic properties and lack the capabilities to propose new materials tailored to specific properties or learn from new data without retraining.

Fig. 1(c) presents an integrated approach, combining ML and natural language processing, for predicting material properties and identifying materials with targeted properties. Specifically, for elastic properties prediction and materials generation, we developed two domain-specific LLMs: ElaTBot and ElaTBot-DFT, which predict elastic properties such as the elastic constant tensor, bulk, shear and Young's moduli, as well as the Poisson ratio. To further improve user interaction and task handling, we implemented an AI-driven agent capable of utilizing tools and databases, and general LLMs to perform complex, multi-step tasks. This agent can process new (and unseen) data by integration of external tools and vector databases. Its responses can be fed into general LLMs (*e.g.*, GPT-4, Gemini) to further extend its capabilities and tackle more complex, multi-step tasks. Fig. 1(d) shows three capabilities of our specialized LLM ElaTBot: prediction, Retrieval-Augmented Generation (RAG)-enhanced prediction[35] without retraining, and generation.

To train the ElaTBot, we first used robocrystallographer[36] to extract structural text descriptions, then employed Pymatgen[37] to obtain compositional information. We then integrate these elements into text-form prompts, subsequently fine-tuning the general LLM Llama2-7b model to yield ElaTBot-DFT, a specialized model for predicting elastic constant tensors at 0 K.

 © 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

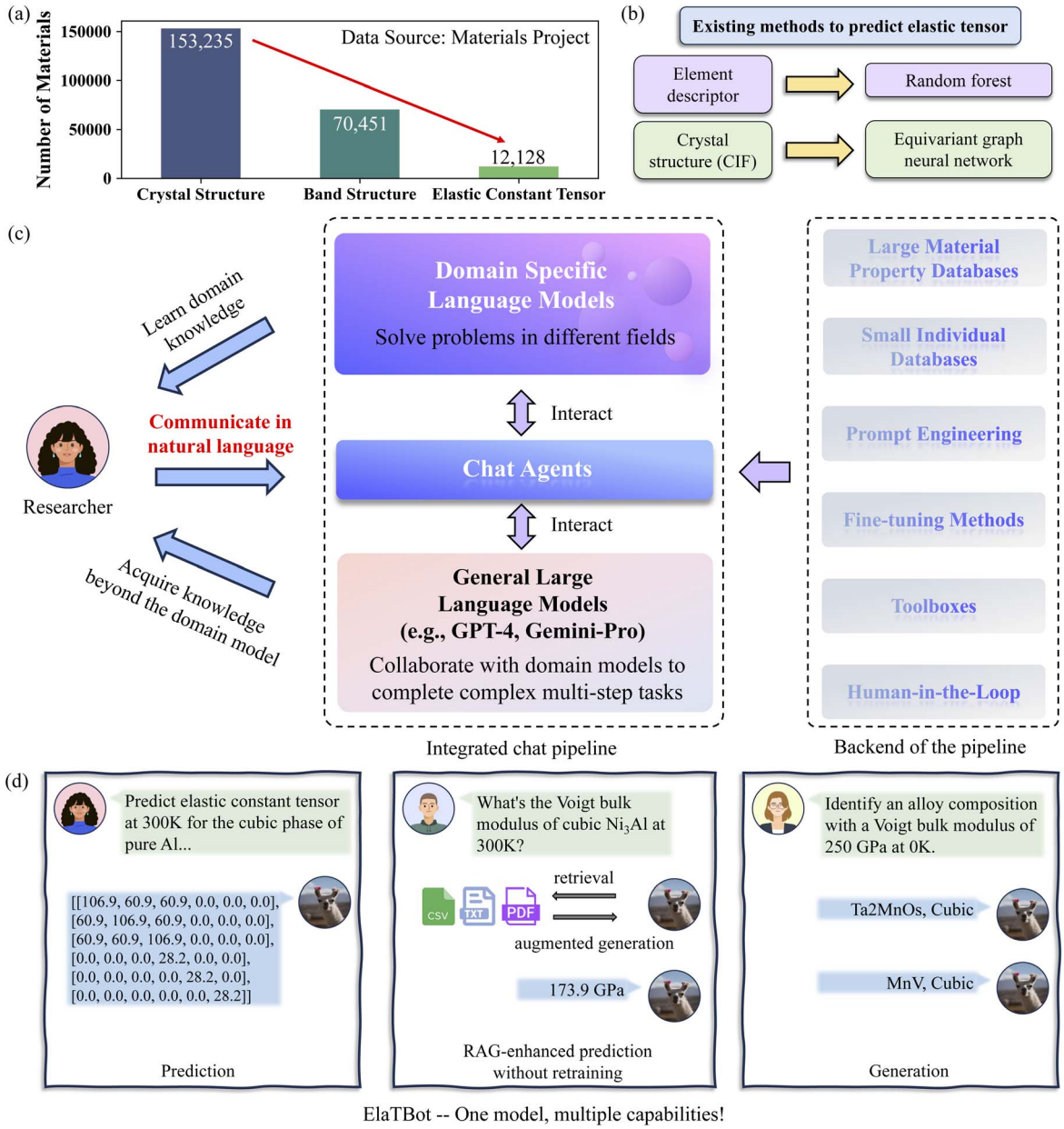

Fig. 1 Datasets and overview of the ElaTBot for predicting elastic properties. (a) Comparison of the number of materials in the Materials Project database[30] with available data on crystal structures, band structures, and elastic properties. The availability of elastic constant tensors data is significantly lower than that of crystal and band structure data. (b) Overview of existing methods used to predict elastic constant tensors, which primarily relied on element descriptors and structural features (constructed by CIF). (c) A flowchart illustrating the process of using large language models (LLMs) to acquire material knowledge. This method enables researchers to gain domain-specific insights, allowing those without extensive programming skills or theoretical expertise to conduct research, thereby lowering the entry barrier into materials science. (d) Capabilities of our specialized LLM ElaTBot. By incorporating elastic constant tensors data at finite temperatures, we develop an LLM-based agent, ElaTBot, which is capable of predicting elastic constant tensors, enhancing prediction without retraining by leveraging external tools and datasets, and generating chemical composition for materials with specific modulus.

ElaTBot-DFT serves as a benchmark for elastic constant tensor prediction, particularly since other models are limited in addressing finite-temperature predictions. Next, we employ several steps to enhance LLM performance,[38] incorporating finite-temperature data and fusing this knowledge to develop ElaTBot. Prompt engineering leads to a reduction of the prediction error of the average value of the elastic constant tensor $\overline{C_{ij}}$ (see Methods) by 33.1% for ElaTBot-DFT compared to Darwin,[29] a materials science LLM built on the same dataset. We ran the test set twice to ensure the reliability of the LLM results. Through knowledge fusion, ElaTBot accurately fits the temperature-dependent bulk modulus curves (derived from the elastic constant tensor) for new multicomponent alloys, with errors near room temperature approaching the average error for

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

the 0 K test set. RAG-enhanced[35] predictions with limited finite-temperature data further improves ElaTBot errors for the bulk modulus from 27.49% to 0.95% across nine alloys at various temperatures without retraining or fine-tuning.

We integrate ElaTBot with GPT-4o to propose/screen materials based upon bulk modulus and other requirements of targeted applications. These include materials with low corrosion rates and high biocompatibility (measured by median lethal dose) with a bulk modulus similar to that of bone for bone implantation, high bulk but low shear modulus materials suitable for exoskeletons of soft robots, corrosion-resistant materials suitable for saline environments, and materials for the protective layers of $LiCoO_2$ electrode.

Elastic constant tensor predictions

Our model uses text-based inputs to predict the elastic constant tensor; this requires carefully designed prompt templates. We develop a prompt (prompt type 4 in ESI Table S1†) that incorporates both chemical composition and crystal structure descriptions as inputs to the model (Llama2-7b, an open-source general LLM). These textual descriptions were generated by extracting relevant data from the material composition and structure and converting these into natural language.

We conducted a series of "experiments" to assess the effects of different input formats on model performance: JSON-formatted (JavaScript Object Notation) structure descriptions (prompt type 1), textual descriptions of crystal structure (prompt type 2), textual descriptions of the composition (prompt type 3), and textual description of both chemical composition and crystal structure (prompt type 4). The model was trained using a 0 K density functional theory (DFT) dataset containing 9498 materials with elastic constant tensor data from the Materials Project, with 500 materials for validation and 522 for testing. Fig. 2(a) and ESI Table S2† show that prompt type 4 achieves a mean absolute error (MAE; the average of the absolute differences between the predicted and actual values for all data points) of 2.32 GPa and R^2 of 0.965 for predicting the average elastic constant tensor component ($\overline{C_{ij}}$), outperforming other prompt types (explicit definitions of the MAE and $\overline{C_{ij}}$ are in Methods). Compared to prompt type 1 (JSON format), prompt type 4 reduces the MAE by 16.8% and increases R^2 by 1.9%. When compared to prompt type 2 (crystal structure descriptions only), prompt type 4 achieves a 5.3% reduction in MAE and a 0.8% increase in R^2. Compared to prompt type 3 (composition descriptions only), prompt type 4 yields a 13.1% reduction in MAE and a 0.9% increase in R^2. These results demonstrate that LLMs perform better when trained with natural language-like inputs, and that using both structural and compositional information improves elastic constant tensor prediction. The bulk modulus results (derived from the elastic constant tensor) in Fig. 2(b) confirm this: prompt type 4 achieves an MAE of 7.74 GPa and an R^2 of 0.963, representing a 14.4% reduction in MAE and a 1.7% increase in R^2 compared to prompt type 1. Therefore, prompt type 4 was selected for training Llama2-7b for our ElaTBot-DFT model.

We compared the performance of ElaTBot-DFT with two widely-used models for predicting the full elastic constant tensor: the random forest model, which utilizes Magpie (Materials Agnostic Platform for Informatics and Exploration) features based on composition,[39] and the MatTen model, which employs a crystal structure graph neural network.[28] As shown in Fig. 2(a, b), ESI Fig. S1† and Table S3,† when trained on the dataset, ElaTBot-DFT using prompt type 4 achieves a 30.3% reduction in MAE and a 4.4% increase in R^2 for predicting the average elastic constant tensor components ($\overline{C_{ij}}$) compared to the random forest model. Compared to the MatTen model, ElaTBot-DFT reduces MAE by 4.5% and improves R^2 by 0.2%. This demonstrates that, even with a relatively small dataset, LLMs trained with well-designed textual descriptions can outperform traditional methods, contrary to previous studies using QA-based training approaches.[25] We also examined the symmetry of the generated elastic constant tensors that result from the rigorous application of crystal symmetries; this symmetry requires certain C_{ij} components to be zero and a fixed relationship between some others. Under strict criteria (error margin of ± 2 GPa), ElaTBot-DFT achieves a symmetry accuracy of 94%, significantly outperforming MatTen (5%) and the random forest model (6%) (Fig. 2(c)). Traditional numerical models tend to produce small non-zero values due to algorithmic limitations, while the natural language-based model, ElaTBot-DFT, accurately outputs a "0" where appropriate. We also tested the elastic stability of all of the materials in the test (*i.e.*, the Born condition – the elastic constant tensor is positive definite), as shown in Fig. S9.† For the 519 materials in the test set, 518 are found to be elastically stable for predictions of both the random forest model and our ElaTBot-DFT model, whereas the predictions of MatTen failed the stability test in 32 cases.

We further compared ElaTBot-DFT predictions with those from the domain-specific materials LLM, Darwin, for elastic constant tensor prediction. Domain-specific LLMs are widely believed to outperform general LLMs on specialized problems;[40] however, as shown in Fig. 2(a, b, d, e) and ESI Table S3,† Darwin (even after fine-tuning on the same dataset) underperforms ElaTBot-DFT in predicting the $\overline{C_{ij}}$ and bulk modulus. Specifically, the MAEs of ElaTBot-DFT are 33.1% and 31.8% lower than those of Darwin for $\overline{C_{ij}}$ and bulk modulus, respectively. This suggests that integrating the reasoning abilities of a general LLM with fine-tuning on a specific dataset may yield better results for tasks requiring quantitative property predictions. Fine-tuning a model with domain-specific knowledge (like Darwin) can lead to gaps in its abilities and knowledge loss, which may reduce the effectiveness in specialized tasks.[41]

We further examined the performance of ElaTBot-DFT across different crystal systems, as summarized in Tables S6 and S7.† The model demonstrates consistently strong predictive accuracy across all crystal systems except for the triclinic system, for which there are only three/sixty data points in the test/training sets. The performance of ElaTBot-DFT is particularly strong for the very common cubic system, with $R^2 > 0.97$ for both elastic constant tensor and bulk modulus predictions. The performance is slightly lower in the orthorhombic and monoclinic systems, with $R^2 \sim 0.94$. This demonstrates that while

 © 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

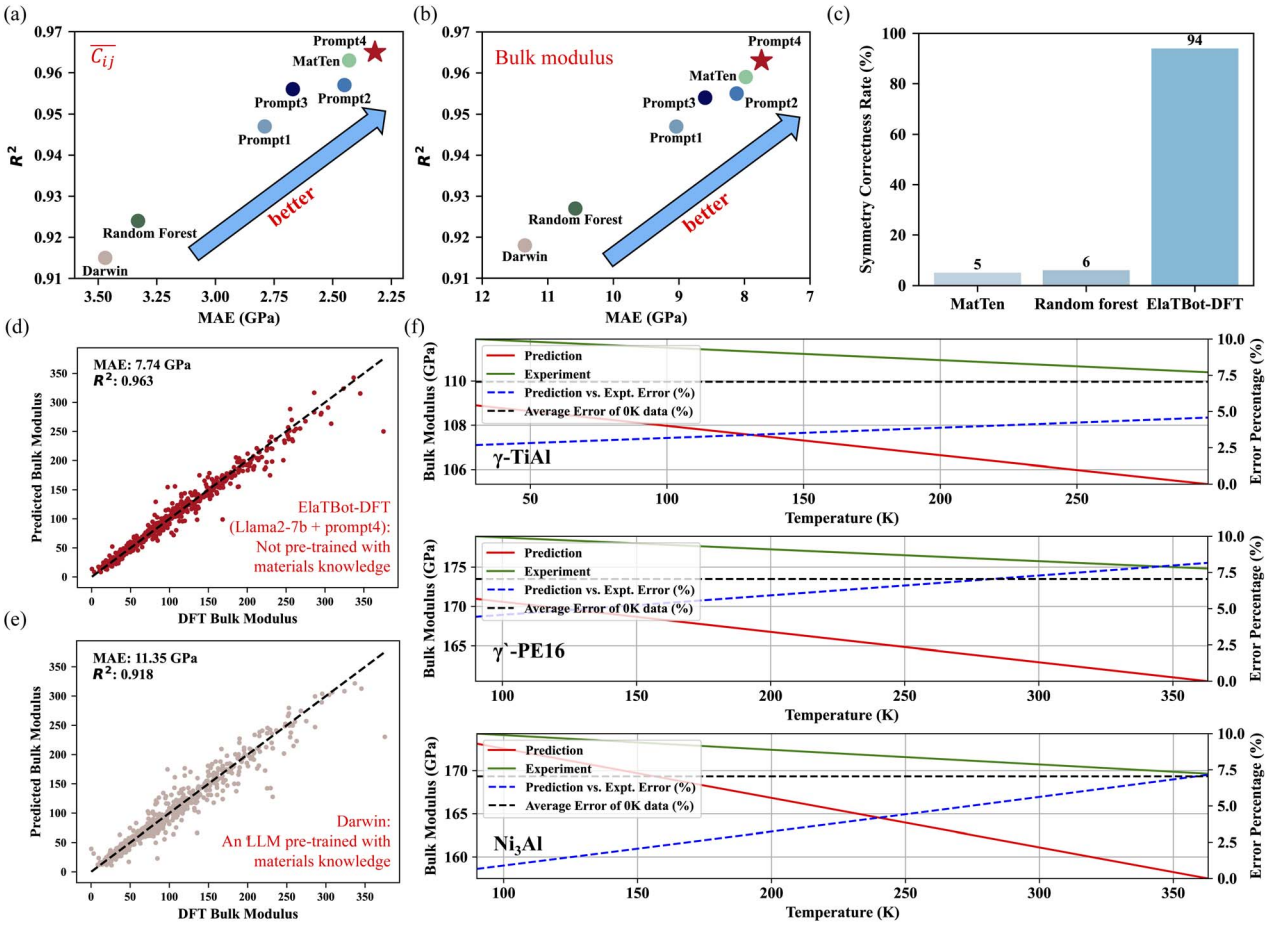

Fig. 2 Prediction abilities of ElaTBot-DFT and ElaTBot. (a and b) Performance comparison of the Llama2-7b model using different prompt types, the MatTen model, random forest model, and Darwin model in predicting $\overline{C_{ij}}$ (GPa) and bulk modulus based on MAE and R^2 on the 0 K DFT test set, all trained on the same dataset. (c) Symmetry validation for elastic constant tensors predicted by MatTen, random forest, and ElaTBot-DFT models. Symmetry correctness is defined as components within ±2 GPa where zero values are required by the Voigt format matrix. (d and e) Performance comparison of ElaTBot-DFT model against the pre-trained Llama2-7b model (Darwin, trained with a materials knowledge database) for bulk modulus prediction, using the same test set and training data with prompt type 4. (f) The capability of ElaTBot to predict finite temperature bulk modulus. The red line indicates predicted values, the green line shows experimental data,[42,43] the blue dashed line indicates the percentage error trend, and the black dashed line shows the average error (7.05%) for the 0 K temperature test set.

ElatBot-DFT is broadly effective across different crystal systems, predictive accuracy is influenced by training/test sample sizes (this is an issue only for less common crystal systems).

Finally, we integrated the finite-temperature dataset and designed four tasks (ESI Table S4†) with corresponding training text inputs (elastic constant tensor prediction, bulk modulus prediction, material generation based on bulk modulus, and text infilling) to conduct multi-task knowledge fusion training. This approach equips ElaTBot with multiple capabilities, including the ability to predict elastic constant tensors at finite temperatures. Although the text infilling task does not directly predict material properties, previous studies have shown that it improves the overall multi-task performance.[38] To test the effectiveness of ElaTBot, we selected three multicomponent alloys not in the training set (cubic Ni_3Al, γ′-PE16 ($Ni_{72.1}Al_{10.4}Fe_{3.2}Cr_{1.0}Ti_{13.3}$), and tetragonal γ-TiAl ($Ti_{44}Al_{56}$)) and predicted their bulk modulus as a function of temperature (based on the full elastic constant tensors). Given the limited finite-temperature training data-just 1266 samples-and the vast compositional space of alloys, predicting accurate values over a wide range of temperature and compositions is inherently challenging. We predicted the bulk modulus at 11 temperatures for Ni_3Al and γ′-PE16 and 15 temperatures for γ-TiAl. Fig. 2(f) shows the variation of prediction errors for three alloy systems (not in the training set) as a function of temperature; the blue dashed lines indicating the error trends. A clear increase in prediction error with temperature is observed. We note that in the original training set, there are 10 520 samples at 0 K and only 1266 entries at finite-temperature conditions. The errors are larger for the quinary γ′-PE16 ($Ni_{72.1}Al_{10.4}Fe_{3.2}Cr_{1.0}Ti_{13.3}$) alloy compared with the binary Ni_3Al alloy. The original training set had 213 times more binary than quinary data. This highlights that the model performance is less reliable for situations (alloy and temperature) where the test cases differ greatly from those in the training set. Nonetheless, the model performs remarkably well compared across a wide range of composition and temperature, especially in light of the fact that experimental data on compositionally complex materials and at high

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

temperatures is rare (and expensive to generate). Fig. 2(f) also shows that the fitted lines from the predictions of ElaTBot closely align with experimental data,[42,43] particularly at low temperatures, where ElaTBot exhibits smaller errors. This demonstrates that incorporating the 0 K DFT dataset from the Materials Project helps reduce prediction errors, highlighting the effectiveness of the multi-task knowledge fusion training approach.

RAG enhanced predictions

Retrieval-Augmented Generation (RAG)[35] provides an effective method for LLMs to access up-to-date databases, enabling the most current RAG-enhanced prediction without model retraining. RAG integrates information retrieval with generative models, enhancing the knowledge scope and accuracy of LLM[35] output. The retrieval module extracts relevant data from external sources, which is then combined with the generative model to deliver more accurate predictions or text generation (Fig. 3(a)). This approach allows LLMs to stay current with new data or literature, by incorporating knowledge rather than solely relying on pre-trained models.

Fig. 3(a) compares the ElaTBot bulk modulus predictions for γ-TiAl at 170 K with and without RAG support. Since the finite temperature data was not in the original ElaTBot training set, the model automatically queries our external database, finds bulk modulus data for γ-TiAl at similar temperatures (the 170 K data was removed from the database for comparison purposes). The predicted value (110.77 GPa) differs by only 0.1% from the true value, whereas without RAG, the error increases to 2.4%. To ensure a fair comparison, we customized the RAG prompt in order to isolate its influence from training prompts. Further testing on alloy data at various temperatures (see Fig. 3(b–d) and ESI Table S5†) demonstrates that RAG reduces the average error from 27.49% to 0.95%. RAG prediction performance can be improved by increasing the quantity and quality of data. This is demonstrated in Table S8,† where increasing the number of data points by 32% led to a 50% decrease in the error, compared with experiments. By incorporating RAG, ElaTBot achieves RAG-

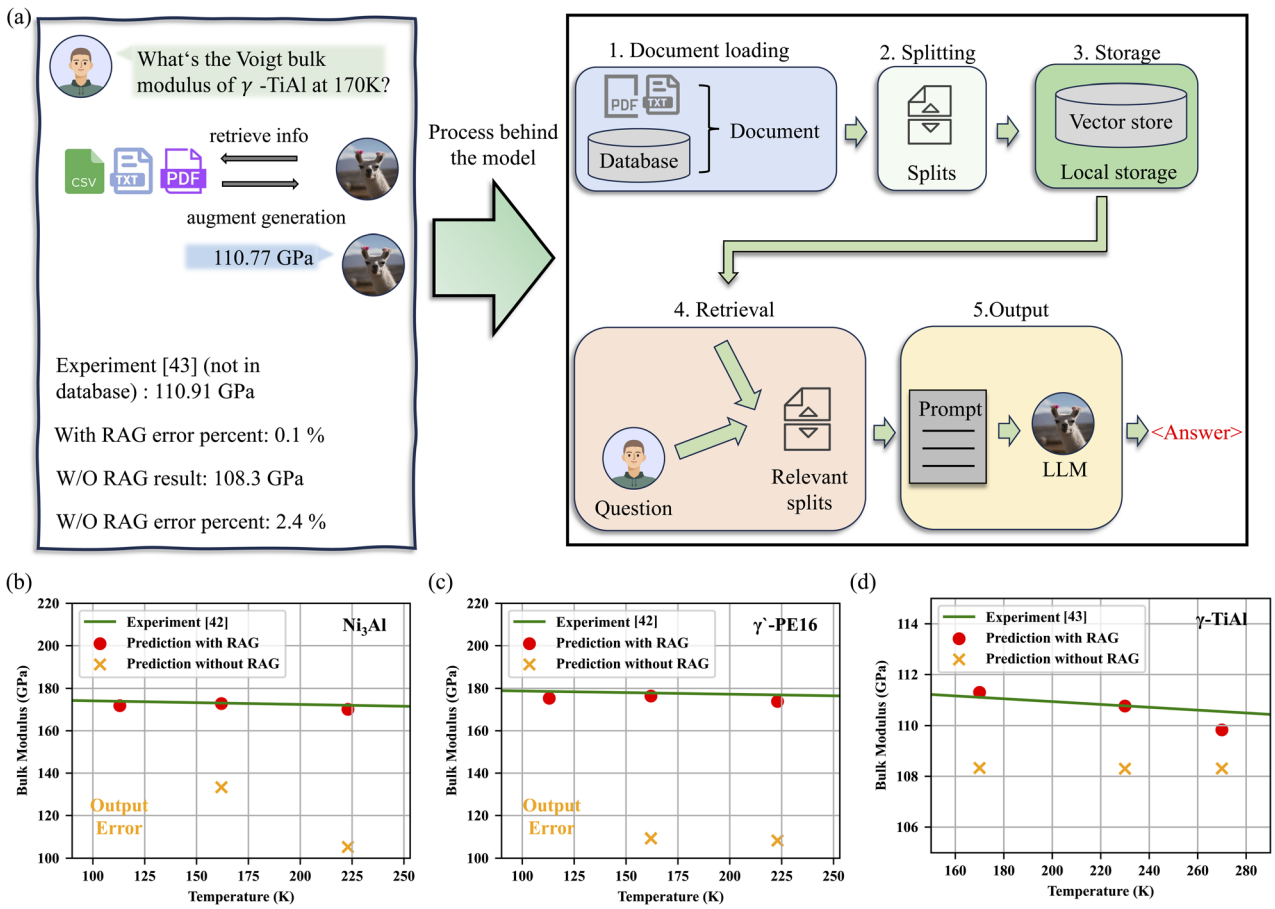

Fig. 3 Integration of Retrieval-Augmented Generation (RAG) with ElaTBot for enhanced prediction. (a) The steps involved in enabling ElaTBot to perform RAG are as follows: (1) document loading: external documents or data sources are ingested into the system for further use; (2) splitting: the documents are broken down into smaller, manageable chunks to optimize retrieval; (3) storage: these chunks are stored in an indexed format for fast and efficient searching; (4) retrieval: the system identifies and retrieves the most relevant chunks in response to the query; (5) output: ElaTBot generates a more accurate and informed response by incorporating the retrieved information. (b–d) Differences in predicted bulk modulus with and without the RAG method for Ni_3Al, γ′-PE16 ($Ni_{72.1}Al_{10.4}Fe_{3.2}Cr_{1.0}Ti_{13.3}$) and γ-TiAl ($Ti_{44}Al_{56}$) as a function of temperature. The bulk modulus error percent decreased from 27.49% to 0.95% in 9 alloys with different temperatures after using RAG.

 © 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

enhanced prediction capabilities, allowing it to reason well beyond its training set that contained minimal similar data.

Material discovery

The knowledge-fused ElaTBot may also be applied to inverse materials design. By combining the domain-specific LLM (ElaTBot) with a general LLM (GPT-4o), we can search for materials with specific bulk modulus for various applications. Fig. 4(a) shows an example where we identify orthopedic materials for bone replacement with a bulk modulus similar to that of bone (<50 GPa).[4] Using the "material generation task" prompt template from ESI Table S4,† we identify the bulk modulus target as <50 GPa and request the ElaTBot agent to generate potential material compositions. This process can be automated *via* a Python script to obtain multiple material compositions. The generated compositions are then passed to GPT-4o (dialogue record in ESI Fig. S3†) with more specific requirements, such as corrosion resistance and biocompatibility; *i.e.*, a corrosion rate <0.3 mm per year and an LD_{50} (median lethal dose) <3 g per kg body weight.[44] This results in a list of compositions that satisfy both the bulk modulus and orthopedic material criteria, along with explanations for each recommendation. Fig. 4(b–d) and ESI Fig. S4–S6† show that this process can be extended to discover materials with high bulk modulus (~250 GPa) and low shear modulus, new corrosion-resistant materials with bulk modulus similar to stainless steel (~160 GPa),[45] or materials for the protective layers for $LiCoO_2$ electrode in lithium battery that could be used to stabilize high-capacity battery electrodes by providing structural support and accommodating volume changes during charge/discharge cycles in lithium batteries. We generated protective layer materials for $LiCoO_2$ with bulk moduli ranging from 40 to 120 GPa in increments of 20 GPa, with 500 materials per modulus value. After screening, our workflow identified several promising candidates. Notably, materials such as Li_2S,[46] Li_3SbS_4,[47] CaF_2,[48] Mg_2SiO_4,[49] $BaTiO_3$,[50] Ti_2AlC,[51] and $LiNbO_3$ [52] fall within or near the generated bulk modulus range and have been experimentally validated as effective protective layers for electrodes. These applications demonstrate the potential of integrating domain-specific and general LLMs to accelerate new material discovery and design.

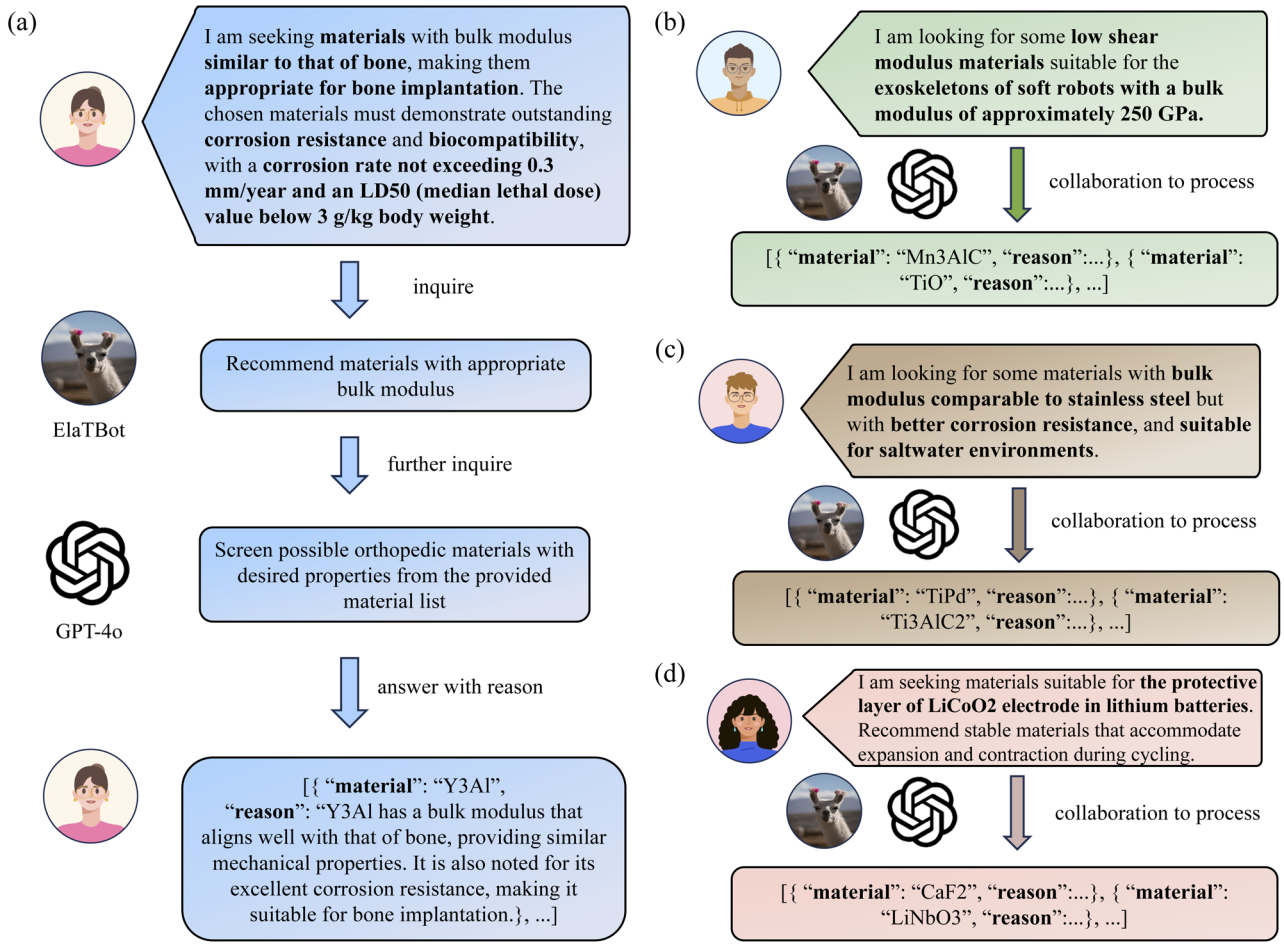

Fig. 4 Integration of a domain-specific LLM (ElaTBot) with a general LLM (GPT-4o) for materials discovery. The process begins by requesting ElaTBot to generate material compositions with a target bulk modulus. Next, GPT-4o refines the results to identify compositions that meet specific application requirements. Examples of applications include: (a) generating orthopedic materials with bulk modulus similar to bone, (b) discovering materials with high bulk modulus and low shear modulus, (c) finding new corrosion-resistant materials, and (d) identifying materials for the protective layers of $LiCoO_2$ electrode in lithium battery.

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Based on the prediction and generation abilities above, we developed a multi-function agent interface (Fig. S2†) that allows researchers to predict, generate, or engage in RAG-enhanced prediction through natural language dialogue, without the heavy load of coding.

Discussion

We demonstrated the potential of LLMs in predicting elastic constant tensors and discovering materials with targeted elastic properties. By introducing ElaTBot-DFT (to accurately predict elastic constants at 0 K) and ElaTBot (extention to finite temperatures), we showcase the ability of LLMs to predict material properties and discover materials with specified properties. Our results show that LLMs, even when trained on relatively small datasets, can outperform traditional ML approaches with carefully curated textual descriptions. The success of our LLM-based approach can be attributed to several key features. First, the Transformer architecture incorporates mechanisms such as LayerNorm, residual connections, and dropout to enhance model generalization capabilities. Second, representing crystal structures as text tokens rather than numerical features enables more accurate predictions, particularly for exact zero values that traditional numerical models struggle to represent precisely. Third, the textual representation naturally facilitates multimodal integration of crystal structures and chemical compositions. Finally, the text-based approach simplifies the development of multi-functional models that can perform numerous tasks traditionally requiring separate numerical models. Furthermore, the combination of domain-specific and general LLMs opens new avenues for materials discovery, while the incorporation of RAG enhances real-time learning and improves the scope and accuracy of LLM predictions.

Despite these promising results, several challenges remain, particularly in ensuring the stability of continuous quantitative predictions. For example, minor variations in temperature, such as between 500.12 K and 500.13 K, may lead to inconsistencies in property predictions like the bulk modulus. To address these issues, future work will focus on generating larger datasets,[53] developing multi-agent systems for incremental task-solving,[54] exploring novel digital encoding methods for LLMs,[55] and guiding LLMs to learn materials laws (such as the general trend of decreasing elastic constant tensor values with increasing temperature). These improvements, along with the addition of constraints or regularization techniques, may enhance the stability of numerical predictions.

Our work presents a fresh perspective on using LLMs for the quantitative prediction of material properties and facilitating inverse material design. A key benefit of domain-specific LLMs is the ability to interact with and generate results through natural language, without requiring users to have extensive knowledge of the underlying ML techniques. This lowers the barrier to entry for computational materials design and fosters broader participation in the field. The integration of domain-specific and general-purpose LLMs allows for access to broader research data, enhancing the synergy between materials science and AI. These advancements have the potential to revolutionize both fields by accelerating innovation, discovery, and application.

Methods

Data acquisition and processing

We trained ElaTBot-DFT, Darwin, the random forest model and the MatTen model using material data containing elastic constant tensors from the Materials Project (MP). Initially, 12 128 materials with elastic constant tensor data calculated by DFT were available from MP (Fig. 1(a)). After filtering out unreasonable entries, 10 520 valid samples remained. From this set, we allocated 9498 to the training set, 500 to the validation set used during training, and 522 to the test set (521 for MatTen because it does not support structures with element Ne). The partitioned dataset includes the material_id for each entry, ensuring consistency across all methods with identical training/validation/test sets. The partitioning methodology involved extracting 5% of the materials from each crystal system as the test set, with the remaining data forming a combined pool for training and validation sets. This combined pool was first shuffled and then 5% of these data were randomly selected to form the validation set with a fixed random seed. The distribution of materials with different crystal systems is shown in Fig. 5(a). We compared the performance of all models on the test set using the mean absolute error (MAE) and the coefficient of determination (R^2). To prepare the data for model input, we followed procedures appropriate for each model. For ElaTBot, we constructed textual descriptions based on the scheme in Table S1,† using pymatgen[37] and robocrystallographer[36] to convert composition and structural information into textual descriptions (see ESI Fig. S7†). Darwin was trained with the same prompt type 4 used for ElaTBot. The random forest model was trained using Magpie feature vectors, which were derived from the elemental composition with pymatgen and matminer.[56] For the MatTen model, we constructed crystal structure graph neural networks following the original settings.[28]

In addition to the 10 520 data points for elastic constant tensors at 0 K, we manually extracted 1266 experimental elastic constant tensor data points at finite temperatures from ref. 57. The distribution of elastic constant tensor data at different temperatures for this dataset is shown in Fig. 5(b). To enable multitasking in ElaTBot, we designed four tasks (Fig. 5(c)) and converted material composition and structural information into textual descriptions as outlined in ESI Table S4.† Given the limited availability of finite-temperature data, we did not create a separate test set for this subset. Instead, we evaluated predictive performance on unseen alloy compositions, including cubic phase Ni_3Al, γ'-PE16, and tetragonal phase γ-TiAl.

Model training and evaluation

We trained the ELaTBot-DFT and ElaTBot models using four NVIDIA V100, both based on the Llama2-7b pre-trained model. Training the ElaTBot model required approximately 24 hours

 © 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

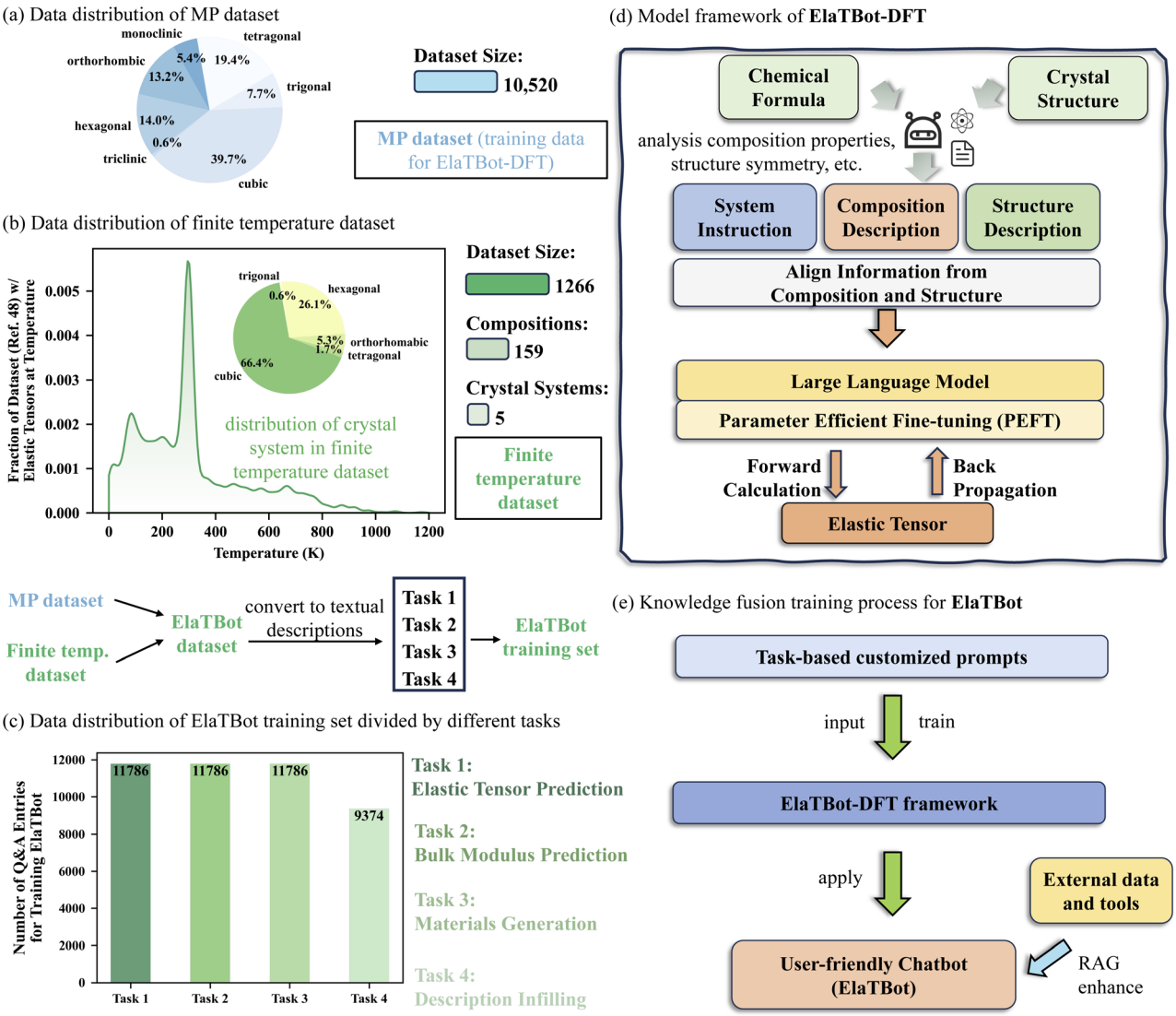

Fig. 5 Dataset and model architecture for ElaTBot and ElaTBot-DFT. (a and b) Data distribution for Materials Project (MP) 0 K DFT dataset and finite temperature dataset. For ElaTBot-DFT, we used data from the MP dataset, converting it to prompt type 4 (shown in Table S1†). The transformed textual descriptions were separated to the training set, validation set, and test set and then used for training ElaTBot-DFT. For ElaTBot, as shown in the lower part of (b), we combined data from the MP dataset and the finite temperature dataset, then converted it into question and answer (Q&A, each Q&A pair is a task and we designed four tasks) as input and output for training ElaTBot, enabling ElaTBot to acquire multiple capabilities. (c) Number of Q&A entries used for training ElatBot, categorized by the specific tasks in the training. The elastic constant tensor prediction task involves training the ElatBot to predict elastic constant tensors based on textual descriptions of materials. The bulk modulus prediction task requires the ElatBot to predict the bulk modulus from material textual descriptions. The material generation task aims to enable the ElatBot to generate material chemical formulas based on given bulk modulus and temperature. The description infilling task, given a description of the chemical formula and compositions, masks the formula with [MASK], and the ElaTBot is then expected to fill in [MASK] with the correct chemical formula. (d) Model architecture for training ElaTBot-DFT. (e) Knowledge fusion training workflow for ElaTBot, detailing how external data and tools are integrated to enhance model capabilities.

on four NVIDIA V100 GPUs (this incurred a total cost of 331 RMB ≈45 USD on the Bohrium platform). On the other hand, training the ElaTBot-DFT model (only based on DFT data at 0 K) on the same platform required 5 hours (cost: 88 RMB ≈12 USD). These costs reflect the computational resource efficiency of our approach. The Darwin model, also built on Llama2-7b, was fine-tuned using a material science literature database. For comparison, we implemented the random forest model using scikit-learn[58] and employed a crystal structure graph neural network for the MatTen model. The training process for ElaTBot-DFT is shown in Fig. 5(d), and that of ElaTBot is shown in Fig. 5(e). To ensure a fair comparison, we standardized both the training duration and the number of samples processed across all models. Techniques such as early stopping were used to finalize the model when performance gains stagnated. Hyperparameters for model training are detailed in the ESI.†

For fine-tuning, we applied LoRA+,[59] a parameter-efficient adaptation technique that extends basic LoRA.[60] LORA+ allows the adapter matrices to be fine-tuned at different learning rates, reducing GPU memory usage by approximately half without

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

compromising data input capacity, thereby accelerating training. A detailed comparison between LoRA and LoRA+ is provided in Fig. S8 of ESI.†

The LLMs (ElaTBot-DFT, ElaTBot, and Darwin) were managed using the Llama-factory[61] framework, which facilitates model loading and parameter tuning. The architectural advantages of the Transformer architecture of Llama-2, including LayerNorm, residual connections, and dropout mechanisms limit overfitting on small datasets. We implemented warm-up strategies (a neural network training technique where the learning rate is gradually increased to the initial learning rate during the first few training epochs[62]) and a cosine learning rate scheduler (that gradually reduces the learning rate during training[63]) to ensure smooth gradient updates when training on small datasets. The random forest and MatTen models were trained directly using Python and PyTorch. The LLMs were optimized by calculating cross-entropy loss, $\mathscr{L}_{\mathrm{CE}} = -\sum_{i=1}^{N} y_i \log(p_i)$, where y_i is the true label and p_i is the predicted probability. The random forest model used squared error loss, $\mathscr{L}_{\mathrm{SE}} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$, where y_i is the true value and $\hat{y}_i$ is the predicted value. The MatTen model employed mean squared error loss $\mathscr{L}_{\mathrm{MSE}} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$ for optimization, consistent with the method specified in ref. 28. These losses guide the models in learning to accurately predict the elastic constant tensor.

The predicted elastic constant tensor is expressed in Voigt form:

$$C_{ij} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{21} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{31} & C_{32} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{41} & C_{42} & C_{43} & C_{44} & C_{45} & C_{46} \\ C_{51} & C_{52} & C_{53} & C_{54} & C_{55} & C_{56} \\ C_{61} & C_{62} & C_{63} & C_{64} & C_{65} & C_{66} \end{bmatrix}$$

The average value of elastic constant tensor C_{ij} are calculated as follows:

$$\overline{C_{ij}} = \frac{1}{36}\sum_{i=1}^{6}\sum_{j=1}^{6} C_{ij}$$

The bulk modulus K are calculated by pymatgen[37] as follows:

$$K_{\mathrm{voigt}} = \frac{1}{9}\sum_{i=1}^{3}\sum_{j=1}^{3} C_{ij}$$

The mean absolute error (MAE) and coefficient of determination (R^2) are used to evaluate the model performance. The MAE is calculated as follows:

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|$$

R^2 is

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2},$$

where y_i represents the labeled average value of the elastic constant tensor or bulk modulus from the dataset for each data point i. $\hat{y}_i$ is the predicted average value of the elastic constant tensor or bulk modulus from the models for each data point i. The symbol $\bar{y}$ represents the mean of y_i across all data points, and n is the total number of data points. The symmetry of the elastic constant tensor is checked by comparing the predicted tensors with the Voigt format matrix. Different crystal symmetries imply that certain components of C_{ij} are zero and specific relations exist between C_{ij} components (*e.g.*, see ref. 31).

For finite-temperature predictions, ElaTBot generated elastic constant tensors for Ni_3Al and γ'-PE16 at $T = 90$, 113, 142, 162, 192, 223, 253, 283, 303, 333, and 363 K, and for γ-TiAl at $T = 30$, 50, 70, 90, 110, 130, 150, 170, 190, 210, 230, 250, 270, 290, and 298 K. Due to the limited training data and the discrete nature of the prediction of ElaTBot, we did a linear fit to the predicted values and used this to evaluate the deviation from experimental data and analyze error trends.

Materials generation and RAG-enhanced prediction

We used gradio[64] to build a user-friendly chat interface for interacting with ElaTBot. This interface allows users to predict, generate, and perform RAG-enhanced prediction tasks through natural language input without the need to engage directly with code. While a default prompt, used during training, is preloaded into the interface, users can modify it as needed for specific tasks.

The RAG-enhanced prediction ability of ElaTBot was enabled through the integration of RAG, which allows the model to perform real-time learning without requiring retraining. The knowledge base consists of finite temperature, experimentally-measured, elastic constant tensor data for three materials from the literature: Ni_3Al at $T = 90$, 113, 142, 162, 192, 223, 253, 283, 300, 303, 333, 363, 400, 500, 600, 700, 800, 900, 1000 and 1100 K, γ'-PE16 at $T = 90$, 113, 142, 162, 192, 223, 253, 283, 300, 303, 333, and 363 K, and γ-TiAl at $T = 30$, 50, 70, 90, 110, 130, 150, 170, 190, 210, 230, 250, 270, 290, and 298 K.[42,43,65] To ensure a rigorous, unbiased evaluation of our RAG-based system, the database accessed by the RAG system excluded 18, randomly selected data points from the full knowledge base; these 18 points formed the test-set shown in ESI Table S5.† For materials absent from the constructed knowledge base (specifically those not related to the three example alloy systems investigated in this study), the RAG was not activated. To ensure reliability in cases of uncertainty, the prompt included the explicit instruction: 'If you don't know the answer, just say that you don't know.' In instances where the language model responded with 'don't know' even after RAG was applied, the agent system reverts to using the base LLM without retrieval assistance. The RAG module was implemented using langchain,[66] and follows

 © 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM. This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

a multi-step process including document loading, splitting, storage, retrieval, and output generation. This process enhances the ability of ElaTBot to update its knowledge and handle new data efficiently, as outlined in Fig. 3(a).

Code availability

All codes used in the paper are publicly accessible on GitHub (**https://github.com/Grenzlinie/ElaTBot**).

Data availability

The method section provides the models and algorithms employed in this study, while specific parameter implementations are available in ESI.† The data, codes, and processing scripts used in this study can be found at Figshare (DOI: **https://doi.org/10.6084/m9.figshare.28399757.v1**).

Conflicts of interest

The authors declare no competing interests.

Acknowledgements

The Authors would like to thank for startup funding from Materials Innovation Institute for Life Sciences and Energy (MILES), HKU-SIRI in Shenzhen for support of this manuscript. This work is supported by Research Grants Council, Hong Kong SAR through the General Research Fund (17210723, 17200424). T. W. acknowledges additional support by The University of Hong Kong (HKU) *via* seed funds (2201100392, 2409100597) and Guangdong Natural Science Fund (2025A1515012129). We also thank Xiaoguo Gong and Huaiyi Liu for their help in finding the finite temperature dataset.

References

1 L. Hu, P. L. Chee, S. Sugiarto, Y. Yu, C. Shi, R. Yan, *et al.*, Hydrogel-Based Flexible Electronics, *Adv. Mater.*, 2023, **35**(14), 2205326, DOI: **10.1002/adma.202205326**.

2 J. Luo, D. Zou, Y. Wang, S. Wang and L. Huang, Battery thermal management systems (BTMs) based on phase change material (PCM): A comprehensive review, *Chem. Eng. J.*, 2022, **430**, 132741. available from: **https://www.sciencedirect.com/science/article/pii/S1385894721043199**.

3 A. Chaves, J. G. Azadani, H. Alsalman, D. R. da Costa, R. Frisenda, A. J. Chaves, *et al.*, Bandgap engineering of two-dimensional semiconductor materials, *npj 2D Mater. Appl.*, 2020, **4**(1), 29, DOI: **10.1038/s41699-020-00162-4**.

4 C. T. Wu, H. T. Chang, C. Y. Wu, S. W. Chen, S. Y. Huang, M. Huang, *et al.*, Machine learning recommends affordable new Ti alloy with bone-like modulus, *Mater. Today*, 2020, **34**, 41–50. available from: **https://www.sciencedirect.com/science/article/pii/S136970211930759X**.

5 M. Fan, T. Wen, S. Chen, Y. Dong and C. A. Wang, Perspectives Toward Damage-Tolerant Nanostructure Ceramics, *Adv. Sci.*, 2024, **11**(24), 2309834, DOI: **10.1002/advs.202309834**.

6 B. Hu, S. Liu, B. Ye, Y. Hao and T. Wen, A Multi-agent Framework for Materials Laws Discovery, *arXiv*, 2024, preprint, arXiv:241116416, DOI: **10.48550/arXiv.2411.16416**, available from: **https://arxiv.org/abs/2411.16416**.

7 Z. Zheng, O. Zhang, H. L. Nguyen, N. Rampal, A. H. Alawadhi, Z. Rong, *et al.*, ChatGPT Research Group for Optimizing the Crystallinity of MOFs and COFs, *ACS Cent. Sci.*, 2023, **9**(11), 2161–2170, DOI: **10.1021/acscentsci.3c01087**.

8 A. M Bran, S. Cox, O. Schilter, C. Baldassari, A. D. White and P. Schwaller, Augmenting large language models with chemistry tools, *Nat. Mach. Intell.*, 2024, **6**(5), 525–535, DOI: **10.1038/s42256-024-00832-8**.

9 L. M. Antunes, K. T. Butler and R. Grau-Crespo, Crystal Structure Generation with Autoregressive Large Language Modeling, *arXiv*, 2024, preprint, arXiv:230704340, DOI: **10.48550/arXiv.2307.04340**, available from: **https://arxiv.org/abs/2307.04340**.

10 P. Verma, M. H. Van and X. Wu. Beyond Human Vision: The Role of Large Vision Language Models in Microscope Image Analysis, *arXiv*, 2024, preprint, arXiv:240500876, DOI: **10.48550/arXiv.2405.00876**, available from: **https://arxiv.org/abs/2405.00876**.

11 Z. Ren, Z. Ren, Z. Zhang, T. Buonassisi and J. Li, Autonomous experiments using active learning and AI, *Nat. Rev. Mater.*, 2023, **8**(9), 563–564, DOI: **10.1038/s41578-023-00588-4**.

12 J. Dagdelen, A. Dunn, S. Lee, N. Walker, A. S. Rosen, G. Ceder, *et al.*, Structured information extraction from scientific text with large language models, *Nat. Commun.*, 2024, **15**(1), 1418, DOI: **10.1038/s41467-024-45563-x**.

13 S. Liu, T. Wen, A. S. L. S. Pattamatta and D. J. Srolovitz, A prompt-engineered large language model, deep learning workflow for materials classification, *Mater. Today*, 2024, 240–249, DOI: **10.1016/j.mattod.2024.08.028**.

14 B. Ni and M. J. Buehler, MechAgents: Large language model multi-agent collaborations can solve mechanics problems, generate new data, and integrate knowledge, *Extreme Mech. Lett.*, 2024, **67**, 102131. available from: **https://www.sciencedirect.com/science/article/pii/S2352431624000117**.

15 Q. Zhang, K. Ding, T. Lyv, X. Wang, Q. Yin, Y. Zhang, *et al.*, Scientific large language models: A survey on biological & chemical domains, *arXiv*, 2024, preprint, arXiv:240114656, DOI: **10.48550/arXiv.2401.14656**, available from: **https://arxiv.org/abs/2401.14656**.

16 S. Ouyang, Z. Zhang, B. Yan, X. Liu, Y. Choi, J. Han, *et al.*, Structured Chemistry Reasoning with Large Language Models, *arXiv*, 2024, preprint, arXiv:231109656, DOI: **10.48550/arXiv.2311.09656**, available from: **https://arxiv.org/abs/2311.09656**.

17 S. Yu, N. Ran and J. Liu, Large-language models: The game-changers for materials science research, *Artif. Intell. Chem.*, 2024, **2**(2), 100076. available from: **https://**

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

www.sciencedirect.com/science/article/pii/S2949747724000344.

18 T. Dinh, Y. Zeng, R. Zhang, Z. Lin, M. Gira, S. Rajput, *et al.*, LIFT: Language-Interfaced Fine-Tuning for Non-Language Machine Learning Tasks, *arXiv*, 2022, preprint, arXiv:220606565, DOI: 10.48550/arXiv.2206.06565, available from: https://arxiv.org/abs/2206.06565.

19 A. N. Rubungo, C. Arnold, B. P. Rand and A. B. Dieng, LLM-Prop: Predicting Physical And Electronic Properties Of Crystalline Solids From Their Text Descriptions, *arXiv*, 2023, preprint, arXiv:231014029, DOI: 10.48550/arXiv.2310.14029, available from: https://arxiv.org/abs/2310.14029.

20 K. Das, P. Goyal, S. C. Lee, S. Bhattacharjee and N. Ganguly, CrysMMNet: Multimodal Representation for Crystal Property Prediction, *arXiv*, 2023, preprint, arXiv:230705390, DOI: 10.48550/arXiv.2307.05390, available from: https://arxiv.org/abs/2307.05390.

21 K. Choudhary, AtomGPT: Atomistic Generative Pretrained Transformer for Forward and Inverse Materials Design, *J. Phys. Chem. Lett.*, 2024, **15**(27), 6909–6917, DOI: 10.1021/acs.jpclett.4c01126.

22 N. Alampara, S. Miret and K. M. Jablonka. MatText: Do Language Models Need More than Text & Scale for Materials Modeling?, *arXiv*, 2024, preprint, arXiv:240617295, DOI: 10.48550/arXiv.2406.17295, available from: https://arxiv.org/abs/2406.17295.

23 K. Li, A. N. Rubungo, X. Lei, D. Persaud, K. Choudhary, B. DeCost, *et al.*, Probing out-of-distribution generalization in machine learning for materials, *Commun. Mater.*, 2025, **6**(1), 9. available from: https://www.nature.com/articles/s43246-024-00731-w#citeas.

24 R. Jacobs, M. P. Polak, L. E. Schultz, H. Mahdavi, V. Honavar and D. Morgan, Regression with Large Language Models for Materials and Molecular Property Prediction, *arXiv*, 2024, preprint, arXiv:240906080, DOI: 10.48550/arXiv.2409.06080, available from: https://arxiv.org/abs/2409.06080.

25 K. M. Jablonka, P. Schwaller, A. Ortega-Guerrero and B. Smit, Leveraging large language models for predictive chemistry, *Nat. Mach. Intell.*, 2024, **6**(2), 161–169, DOI: 10.1038/s42256-023-00788-1.

26 P. R. Spackman, A. Grosjean, S. P. Thomas, D. P. Karothu, P. Naumov and M. A. Spackman, Quantifying Mechanical Properties of Molecular Crystals: A Critical Overview of Experimental Elastic Tensors, *Angew. Chem., Int. Ed.*, 2022, **61**(6), e202110716, DOI: 10.1002/anie.202110716.

27 J. B. Levine, J. B. Betts, J. D. Garrett, S. Q. Guo, J. T. Eng, A. Migliori, *et al.*, Full elastic tensor of a crystal of the superhard compound ReB2, *Acta Mater.*, 2010, **58**(5), 1530–1535. available from: https://www.sciencedirect.com/science/article/pii/S1359645409007630.

28 M. Wen, M. K. Horton, J. M. Munro, P. Huck and K. A. Persson, An equivariant graph neural network for the elasticity tensors of all seven crystal systems, *Digit. Discov.*, 2024, **3**, 869–882, DOI: 10.1039/D3DD00233K.

29 T. Xie, Y. Wan, W. Huang, Z. Yin, Y. Liu, S. Wang, *et al.*, DARWIN Series: Domain Specific Large Language Models for Natural Science, *arXiv*, 2023, preprint, arXiv:230813565, DOI: 10.48550/arXiv.2308.13565, available from: https://arxiv.org/abs/2308.13565.

30 A. Jain, S. P. Ong, G. Hautier, W. Chen, W. D. Richards, S. Dacek, *et al.*, Commentary: The Materials Project: A materials genome approach to accelerating materials innovation, *APL Mater.*, 2013, **1**(1), 011002.

31 Y. Li, L. Vočadlo and J. P. Brodholt, ElasT: A toolkit for thermoelastic calculations, *Comput. Phys. Commun.*, 2022, **273**, 108280. available from: https://www.sciencedirect.com/science/article/pii/S0010465521003921.

32 V. Revi, S. Kasodariya, A. Talapatra, G. Pilania and A. Alankar, Machine learning elastic constants of multi-component alloys, *Comput. Mater. Sci.*, 2021, **198**, 110671. available from: https://www.sciencedirect.com/science/article/pii/S0927025621003980.

33 G. Vazquez, P. Singh, D. Sauceda, R. Couperthwaite, N. Britt, K. Youssef, *et al.*, Efficient machine-learning model for fast assessment of elastic properties of high-entropy alloys, *Acta Mater.*, 2022, **232**, 117924. available from: https://www.sciencedirect.com/science/article/pii/S1359645422003068.

34 T. Pakornchote, A. Ektarawong and T. Chotibut, StrainTensorNet: Predicting crystal structure elastic properties using SE(3)-equivariant graph neural networks, *Phys. Rev. Res.*, 2023, **5**, 043198, DOI: 10.1103/PhysRevResearch.5.043198.

35 P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, *et al.*, Retrieval-augmented generation for knowledge-intensive nlp tasks, *Adv. Neural Inf. Process. Syst.*, 2020, **33**, 9459–9474, DOI: 10.5555/3495724.3496517.

36 A. M. Ganose and A. Jain, Robocrystallographer: automated crystal structure text descriptions and analysis, *MRS Commun.*, 2019, **9**(3), 874–881, DOI: 10.1557/mrc.2019.94.

37 S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, *et al.*, Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis, *Comput. Mater. Sci.*, 2013, **68**, 314–319. available from: https://www.sciencedirect.com/science/article/pii/S0927025612006295.

38 N. Gruver, A. Sriram, A. Madotto, A. G. Wilson, C. L. Zitnick and Z. Ulissi, Fine-Tuned Language Models Generate Stable Inorganic Materials as Text, *arXiv*, 2024, preprint, arXiv:240204379, DOI: 10.48550/arXiv.2402.04379, available from: https://arxiv.org/abs/2402.04379.

39 L. Ward, A. Agrawal, A. Choudhary and C. Wolverton, A general-purpose machine learning framework for predicting properties of inorganic materials, *npj Comput. Mater.*, 2016, **2**(1), 16028, DOI: 10.1038/npjcompumats.2016.28.

40 T. Xie, Y. Wan, W. Huang, Y. Zhou, Y. Liu, Q. Linghu, *et al.*, Large language models as master key: unlocking the secrets of materials science with GPT, *arXiv*, 2023, arXiv:230402213, DOI: 10.48550/arXiv.2304.02213, available from: https://arxiv.org/abs/2304.02213.

 © 2025 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

41 C. A. Li and H. Y. Lee, Examining Forgetting in Continual Pre-training of Aligned Large Language Models, *arXiv*, 2024, arXiv:240103129, DOI: 10.48550/arXiv.2401.03129, available from: https://arxiv.org/abs/2401.03129.

42 F. Wallow, G. Neite, W. Schröer and E. Nembach, Stiffness constants, dislocation line energies, and tensions of Ni3Al and of the gamma'-phases of NIMONIC 105 and of NIMONIC PE16, *Phys. Status Solidi A*, 1987, **99**(2), 483–490, DOI: 10.1002/pssa.2210990218.

43 K. Tanaka, Single-crystal elastic constants of gamma-TiAl, *Philos. Mag. Lett.*, 1996, **73**(2), 71–78, DOI: 10.1080/095008396181019.

44 R. Gorejová, L. Haverová, R. Oriňaková, A. Oriňak and M. Oriňak, Recent advancements in Fe-based biodegradable materials for bone repair, *J. Mater. Sci.*, 2019, **54**(3), 1913–1947, DOI: 10.1007/s10853-018-3011-z.

45 ToolBox TE, Metals and Alloys - Bulk Modulus Elasticity, 2008, accessed: 26 September 2024, available from: https://www.engineeringtoolbox.com/bulk-modulus-metals-d_1351.html.

46 A. Sakuda, A. Hayashi, T. Ohtomo, S. Hama and M. Tatsumisago, All-solid-state lithium secondary batteries using LiCoO2 particles with pulsed laser deposition coatings of Li2S–P2S5 solid electrolytes, *J. Power Sources*, 2011, **196**(16), 6735–6741.

47 R. Matsuda, H. Muto and A. Matsuda, Air-Stable Li3SbS4–LiI Electrolytes Synthesized via an Aqueous Ion-Exchange Process and the Unique Temperature Dependence of Conductivity, *ACS Appl. Mater. Interfaces*, 2022, **14**(46), 52440–52447.

48 M. Li, H. Wang, L. Zhao, F. Zhang and D. He, The combined effect of CaF2 and graphite two-layer coatings on improving the electrochemical performance of Li-rich layer oxide material, *J. Solid State Chem.*, 2019, **272**, 38–46. available from: https://www.sciencedirect.com/science/article/pii/S0022459619300349.

49 C. Bian, R. Fu, Z. Shi, J. Ji, J. Zhang, W. Chen, *et al.*, Mg2SiO4/Si-Coated Disproportionated SiO Composite Anodes with High Initial Coulombic Efficiency for Lithium Ion Batteries, *ACS Appl. Mater. Interfaces*, 2022, **14**(13), 15337–15345.

50 C. Wang, M. Liu, M. Thijs, F. G. B. Ooms, S. Ganapathy and M. Wagemaker, High dielectric barium titanate porous scaffold for efficient Li metal cycling in anode-free cells, *Nat. Commun.*, 2021, **12**(1), 6536.

51 L. X. Yang, Y. B. Mu, R. J. Liu, H. J. Liu, L. Zeng, H. Y. Li, *et al.*, A facile preparation of submicro-sized Ti2AlC precursor toward Ti2CTx MXene for lithium storage, *Electrochim. Acta*, 2022, **432**, 141152. available from: https://www.sciencedirect.com/science/article/pii/S0013468622013093.

52 C. Wei, C. Yu, S. Chen, S. Chen, L. Peng, Y. Wu, *et al.*, Unraveling the LiNbO3 coating layer on battery performances of lithium argyrodite-based all-solid-state batteries under different cut-off voltages, *Electrochim. Acta*, 2023, **438**, 141545. available from: https://www.sciencedirect.com/science/article/pii/S0013468622017029.

53 J. Kaplan, S. McCandlish, T. Henighan, T. B. Brown, B. Chess, R. Child, *et al.*, Scaling Laws for Neural Language Models, *arXiv*, 2020, preprint, arXiv:200108361, DOI: 10.48550/arXiv.2001.08361, available from: https://arxiv.org/abs/2001.08361.

54 A. Ghafarollahi and M. J. Buehler, AtomAgents: Alloy design and discovery through physics-aware multi-modal multi-agent artificial intelligence, *arXiv*, 2024, preprint, arXiv:240710022, DOI: 10.48550/arXiv.2407.10022, available from: https://arxiv.org/abs/2407.10022.

55 S. Golkar, M. Pettee, M. Eickenberg, A. Bietti, M. Cranmer, G. Krawezik, *et al.*, xVal: A Continuous Number Encoding for Large Language Models, *arXiv*, 2023, preprint, arXiv:230902989, DOI: 10.48550/arXiv.2310.02989, available from: https://arxiv.org/abs/2310.02989.

56 L. Ward, A. Dunn, A. Faghaninia, N. E. R. Zimmermann, S. Bajaj, Q. Wang, *et al.*, Matminer: An open source toolkit for materials data mining, *Comput. Mater. Sci.*, 2018, **152**, 60–69. available from: https://www.sciencedirect.com/science/article/pii/S0927025618303252.

57 G. Simmons and H. Wang, *Single Crystal Elastic Constants and Calculated Aggregate Properties: A Handbook*, M. I. T. Press, Cambridge, Mass, 1971.

58 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, *et al.*, Scikit-learn: Machine Learning in Python, *arXiv*, 2018, preprint, arXiv:12010490, DOI: 10.48550/arXiv.1201.0490, available from: https://arxiv.org/abs/1201.0490.

59 S. Hayou, N. Ghosh and B. Yu, LoRA+: Efficient Low Rank Adaptation of Large Models, *arXiv*, 2024, preprint, arXiv:240212354, DOI: 10.48550/arXiv.2402.12354, available from: https://arxiv.org/abs/2402.12354.

60 E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, *et al.*, LoRA: Low-Rank Adaptation of Large Language Models, *arXiv*, 2021, preprint, arXiv:210609685, DOI: 10.48550/arXiv.2106.09685, available from: https://arxiv.org/abs/2106.09685.

61 Y. Zheng, R. Zhang, J. Zhang, Y. Ye and Z. Luo, Llamafactory: Unified efficient fine-tuning of 100+ language models, *arXiv*, 2024, preprint, arXiv:240313372, DOI: 10.48550/arXiv.2403.13372, available from: http://arxiv.org/abs/2403.13372.

62 D. S. Kalra and M. Barkeshli, Why Warmup the Learning Rate? Underlying Mechanisms and Improvements, *arXiv*, 2024, preprint, arXiv:240609405, DOI: 10.48550/arXiv.2406.09405, available from: https://arxiv.org/abs/2406.09405.

63 I. Loshchilov and F. Hutter, SGDR: Stochastic Gradient Descent with Warm Restarts, *arXiv*, 2017, preprint, arXiv:160803983, DOI: 10.48550/arXiv.1608.03983, available from: https://arxiv.org/abs/1608.03983.

64 A. Abid, A. Abdalla, A. Abid, D. Khan, A. Alfozan and J. Zou, Gradio: Hassle-Free Sharing and Testing of ML Models in the Wild, *arXiv*, 2019, preprint, arXiv:190602569, DOI:

10.48550/arXiv.1906.02569, available from: https://arxiv.org/abs/1906.02569.

65 S. V. Prikhodko, H. Yang, A. J. Ardell, J. D. Carnes and D. G. Isaak, Temperature and composition dependence of the elastic constants of Ni3Al, *Metall. Mater. Trans. A*, 1999, **30**(9), 2403–2408, DOI: 10.1007/s11661-999-0248-9#Abs1.

66 K. Pandya and M. Holia, Automating Customer Service using LangChain: Building custom open-source GPT Chatbot for organizations, *arXiv*, 2023, preprint, arXiv:231005421, DOI: 10.48550/arXiv.2310.05421, available from: https://arxiv.org/abs/2310.05421.

Open Access Article. Published on 20 May 2025. Downloaded on 6/10/2025 1:46:26 AM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

© 2025 The Author(s). Published by the Royal Society of Chemistry