ChatGPT Material Explorer: Design and Implementation of a Custom GPT Assistant for Materials Science Applications

Kamal Choudhary[1,2*]

[1]Department of Materials Science and Engineering, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD 21218, USA.
[2]Department of Electrical and Computer Engineering, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD 21218, USA.

Corresponding author(s). E-mail(s): kchoudh2@jhu.edu ;

Abstract

Custom Generative pre-trained transformers (GPTs) are transforming domain-specific problem-solving across scientific fields. However, materials science lacks a dedicated, accessible GPT-based assistant tailored to its unique challenges. To address this gap, I present ChatGPT Material Explorer 1.0 (CME), a custom GPT that integrates large language models with graph neural networks (GNN) models and other domain-specific APIs to enhance materials design workflows. The assistant offers core functionalities such as: (1) intelligent exploration of molecular and materials databases, (2) GNN-driven prediction of materials properties, (3) efficiently searching arXiv for papers, and (4) interactive, conversational support for scientific analysis and writing. I evaluate its performance against general-purpose models such as OpenAI GPT-4o highlighting its superior ability to retrieve accurate data, reduce hallucination, and provide context-aware responses. By combining natural language understanding with real-time data access and physics-informed modeling, CME aims to democratize computational tools and accelerate decision-making in materials research. More details about CME are available at: https://github.com/AtomGPTLab/chatgpt_material_explorer to facilitate broader community adoption.

1 Introduction

As large language models (LLMs) such as OpenAI's ChatGPT [1], Llama [2] and Gemini [3] continue to improve, their utility is being recognized across an expanding array of scientific disciplines such as materials science [4, 5]. Although these LLMs offer impressive general-purpose capabilities, their output can lack domain-specific scientific rigor [6]. In addition, traditional LLMs often hallucinate data, fail to properly source references, or miss critical nuances to material discovery and design.

Given the number of materials can be reach up to 10^{100}, no amount of LLM training is sufficient to solve the materials combinatorics problem. Databases and AI models for materials have been key advancements in addressing these issues. Nevertheless, despite the progress made by computational materials platforms and databases such as the Joint Automated Repository for Various Integrated Simulations (JARVIS) [7], the Materials Project [8], and Alexandria [9], researchers often face steep learning curves due to the fragmented nature of materials informatics. While several efforts have been made to integrate AI with materials science tools, few offer a publicly accessible, unified GPT-based conversational interface that combines real-time data retrieval, predictive modeling, and technical content generation in a single workflow [10–13]. Moreover, maintaining an online resource can be costly, and leveraging a robust infrastructure like ChatGPT reduces the developer's burden of system maintenance.

In response to these challenges, ChatGPT Material Explorer (CME) 1.0 was developed as a domain-specific custom GPT built on the OpenAI GPT-4o architecture. With the release of ChatGPT-4o, OpenAI introduced personalized GPT instances and customizable versions of the base model tailored to meet domain-specific requirements. These custom GPTs are capable of adapting their behavior and output generation based on predefined instructions and user-defined goals.

CME is developed with several key objectives in mind: (1) lowering the barrier to entry for AI-assisted materials research, (2) integrating experimental and computational data pipelines, and (3) reducing the risk of hallucination by interfacing with real-time databases and using physics-aware modeling approaches with many more potential additions in the future. The ChatGPT Material Explorer tool is publicly available at https://chatgpt.com/g/g-67420c4cc6f48191a8876b173c93d62a-material-explorer-1-0 . A number of example prompts for CME are available on the GitHub repo: https://github.com/AtomGPTLab/chatgpt_material_explorer.

2 Building ChatGPT Material Explorer Assistant

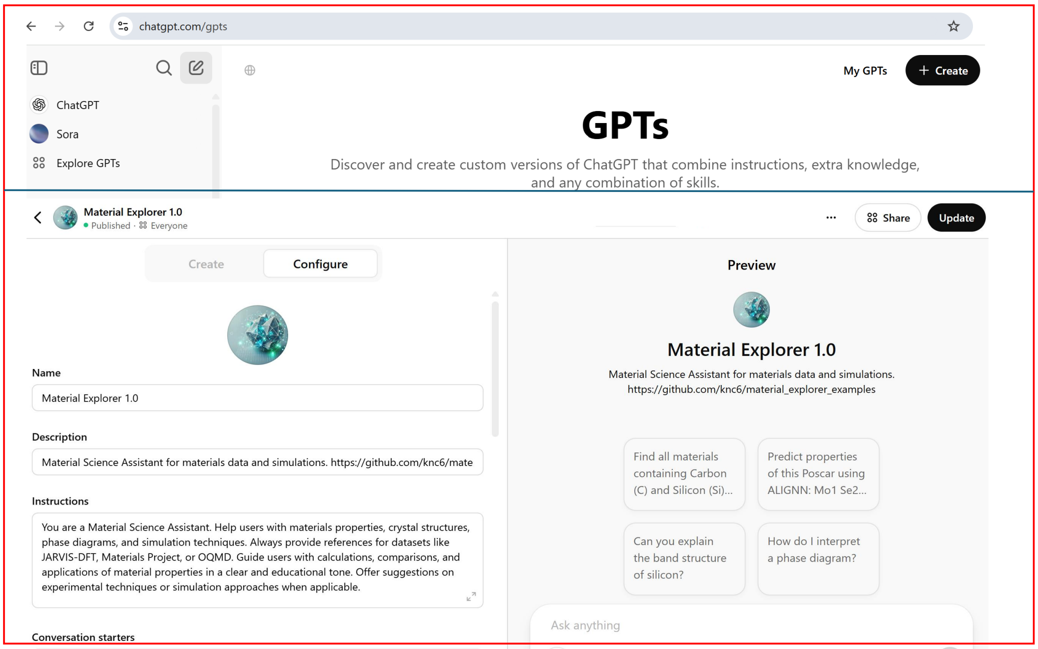

Fig. 1: Screenshot of the GPT-4o "Create" interface used to develop the custom AI assistant, ChatGPT Material Explorer.

To develop an AI assistant capable of supporting rigorous materials science research, the GPT-4o "Explore GPT" to "Create" interface was utilized. This platform enables users to construct customized versions of ChatGPT, designed to perform domain-specific tasks through a combination of natural language configuration and feature definition.

The GPT Builder consists of two primary components: the "create" interface, which facilitates dynamic conversations with the AI to establish foundational behavior, and the "configure" panel, where technical capabilities, tool integrations, and custom instructions are defined. A snapshot of the builder is shown in Fig. 1. The development process began with high-level prompts that describe the intended function of the assistant. Through iterative dialogue, the assistant's behavior was refined by specifying critical parameters such as preferred response formats (e.g., database and GNN model integrations), terminology sensitivity, and expected interaction styles. The initial configuration included a suite of capabilities such as querying real-time databases through APIs such as NIST-JARVIS [14], Atomistic Line Graph Neural Networks (ALIGNN) [15] and ALIGNN-ForceField (ALIGNN-FF) [16], Open Databases Integration for Materials Design (OPTIMADE) [17], NIH-CACTUS [18] and arXiv API [19], with support for material property analysis workflows. Once the core logic and response patterns were established, the assistant's performance was evaluated within the 'playground' environment of the GPT Builder. As new models for materials, such as AtomGPT [20], DiffractGPT [21], and MicroscopyGPT [22], are developed, they can be integrated into CME to make the LLM responses more domain-aware in materials science.

3 Usability, Advantages, and Limitations

3.1 Domain specific database queries

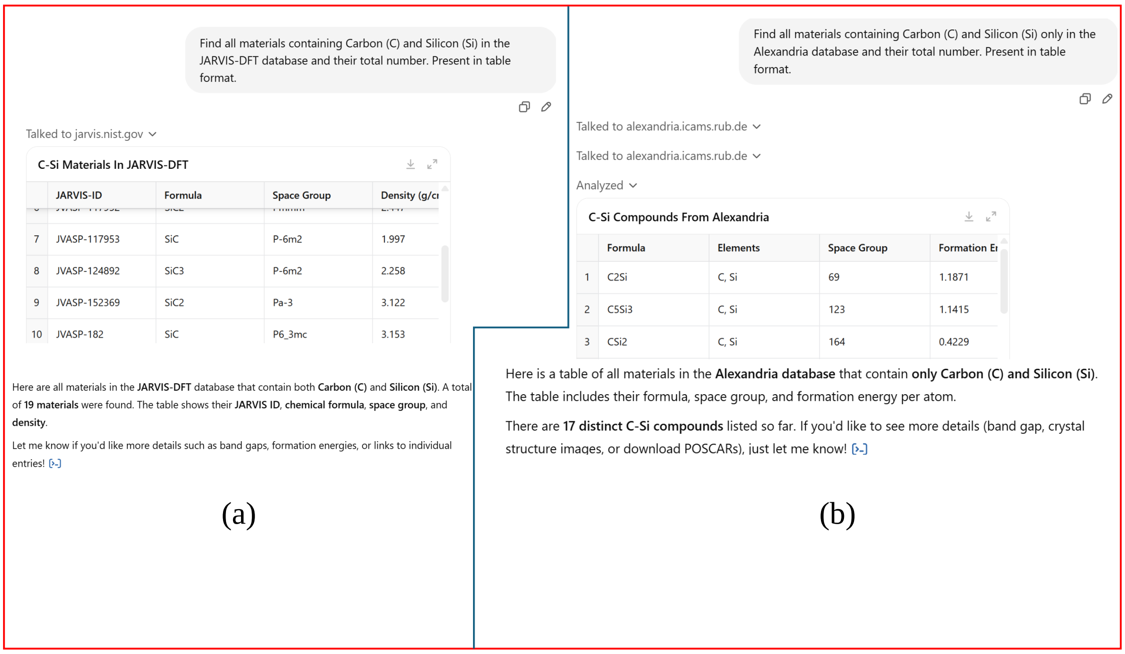

Fig. 2: Screenshots demonstrating the retrieval of materials data containing silicon (Si) and carbon (C) elements from two example computational materials science databases: (a) JARVIS-DFT and (b) Alexandria.

A key innovation of CME is its integration with domain-specific databases and tools, enabling the retrieval, processing, and interpretation of materials data from sources such as NIH-CACTUS [18], JARVIS-DFT [7], Materials Project [8], Automatic FLOW for Materials Discovery (AFLOW) [23], Open Quantum Materials Database (OQMD) [24], and Alexandria [9]. It utilizes API queries to interact with these databases. For example, if a user requests all available materials containing silicon and carbon from JARVIS-DFT, Material Explorer retrieves and interprets the dataset, presenting the results in an explainable manner, a feature particularly valuable for non-experts. While general-purpose ChatGPT can also respond to queries such as "materials with Si and C," it typically yields generic responses that may be incomplete or misleading.

Examples of such applications are shown in Fig. 2a for JARVIS-DFT and Fig. 2b for the Alexandria database. Similar examples of querying molecular information from NIH-CACTUS are presented in Table 1. I compare the responses from GPT-4o, ChemCrow, and our ChatGPT Material Explorer (CME) framework. Highlighted cells indicate correct or reasonably accurate responses that include the final answer or essential components. While Material Explorer is capable of effectively interacting with databases and retrieving accurate information, other frameworks may struggle to do so, as demonstrated in Table 1. The table presents a comparative analysis of the responses generated by the three AI systems-GPT-4o, ChemCrow, and CME in response to a series of domain-specific materials science questions. Highlighted cells identify responses that are correct or reasonably accurate to contain correct information.

Table 1: Comparison between GPT-4o and ChatGPT Material Explorer responses to a few materials science queries. The highlighted cells indicate correct or reasonably accurate response.

Question	**GPT-4o Response**	**ChemCrow**	**ChatGPT Material Explorer Response**
1. Get IUPAC name for CC(=O)OC1=CC=CC=C1C(=O)O	2-(4-acetylphenyl)acetic acid However, this molecule is more commonly known as acetylsalicylic acid, which is the chemical name for aspirin	Aspirin	This is the official IUPAC name for acetylsalicylic acid (aspirin)
2. Get me the molecular formula for aspirin	C9H8O4	C9H8O4	C9H8O4
3. Look up the SMILES notation for ibuprofen	CC(C)CC1=CC=C(C=C1)C(C)C(=O)O	CC(C)Cc1ccc(C(C)C(=O)O)cc1	CC(C)Cc1ccc(cc1)C(C)C(O)=O
4. Find all materials containing chemical formula MgB2 in the JARVIS-DFT database and their total number. Present in table format.	The JARVIS-DFT database includes materials containing the chemical formula MgB2, notably the well-known superconductor magnesium diboride (MgB2) and related compounds. The primary entry for MgB2 in the database is identified by the JARVIS-ID JVASP-1151. Additionally, the database features related materials such as Mg2B4N2, which ...	I'm sorry, but I don't have access to JARVIS-DFT database. I recommend using the JARVIS-DFT...	A total of 4 MgB2 materials were found in the JARVIS-DFT database JVASP-1151, JVASP-19821, JVASP-135405.... The table above summarizes their key properties...
5. Predict properties of this Poscar using ALIGNN: MoSe2 1.0 1.661759 -2.878250 0.000000 1.661759 2.878250 0.000000 0.000000 0.000000 35.451423 Mo Se 1 2 direct 0.666667 0.333333 0.326886 Mo 0.333333 0.666667 0.374080 Se 0.333333 0.666667 0.279691 Se	unning ALIGNN Predictions Locally To predict properties using ALIGNN: Install JARVIS-Tools: Ensure you have Python installed, then run: pip install jarvis-tools Prepare Your POSCAR File: Save your structure in a file named POSCAR. Run ALIGNN Prediction: Use the following Python script: ...	I'm sorry, but I don't have the capability to analyze Poscar file or use ALIGNN to predict properties. Please use a tool or service that supports these features.	Here are the ALIGNN-predicted properties for the given MoSe2 POSCAR structure (from JARVIS-DFT), Formation energy per atom: -0.634 eV, Total energy per atom: -4.52 eV, Bandgap (OptB88vdW functional): 0.359 eV. These values suggest a moderately stable material with a small bandgap, consistent with known semiconducting behavior of MoSe2...

Question	GPT-4o Response	ChemCrow	Material Explorer Response
6. Find papers on MgB2 in arXiv. State how many results did you find and show top 10 recent papers only.	I found several papers on MgB2 (magnesium diboride) in arXiv. Here are the top 10 most recent ones: Kinetic Inductance and Nonlinearity of MgB2 Films at 4K Authors: J. Greenfield, C. Bell...	Note: This tool can take some time to complete execution!! Tool encountered an error... RuntimeError: Failed to avoid a service limit across 5 retries.	I found 905 results on arXiv related to MgB2. Here are the 10 most recent papers: High-pressure high-temperature solution growth, structural, and superconducting properties of Fe-substituted MgB2 single crystals...
7. How do I interpret a phase diagram?	Interpreting a phase diagram-especially a binary phase diagram-is crucial in materials science, metallurgy, and chemistry for understanding material stability, melting, solidification, and transformation. Here's a step-by-step guide:...	A phase diagram is a type of chart used in fields such as physical chemistry, engineering, mineralogy, and materials science. It shows the conditions, such as pressure and temperature, at which thermodynamically distinct phases (like solid, liquid, or gaseous states) ...	Interpreting a phase diagram involves understanding the stability and transitions of different phases (solid, liquid, gas, or multiple solid phases) of a material as a function of variables like temperature, pressure, or composition...
8. Can you explain the band structure of silicon?	Yes, the band structure of silicon is fundamental to understanding its role as a semiconductor. Here's a detailed explanation:1. What is a Band Structure? The band structure of a solid describes the allowed energy levels (bands) that electrons can occupy, as a function of momentum (k-vector)...	The band structure of silicon refers to the ranges of energy that an electron within the silicon may have (called energy bands, energy levels, or simply bands) and ranges of energy that it may not have (called band gaps or simply gaps)...	Crystal Structure and Brillouin Zone Crystal structure: Silicon crystallizes in the diamond cubic structure.Brillouin zone: The symmetry points typically discussed in the band structure of silicon include: Γ (Gamma): center of Brillouin zone X, L, and K: high-symmetry points...

3.2 Domain specific AI models

The Processing-Structure-Property-Performance paradigm is fundamental in materials science. Although conventional techniques such as density functional theory (DFT) and force-field models establish structure-property relationships, these methods are often computationally expensive and time-consuming. Machine learning models, particularly graph neural networks (GNNs) and large language models (LLMs), significantly accelerate property predictions, achieving results thousands of times faster while maintaining high accuracy.

Currently, CME is integrated with the Atomistic Line Graph Neural Network (ALIGNN) model [15] to predict formation energy, bandgap, and total energy. Future updates will support a broader range of GNN-based models and unified force fields, ensuring wider applicability and reducing model bias.

An example of this application, inputting the atomic structure of molybdenum selenide and obtaining its predicted properties using ALIGNN, is shown in Fig. 3. As illustrated in Table 1, GPT-4o and ChemCrow are currently unable to provide such domain-specific information.

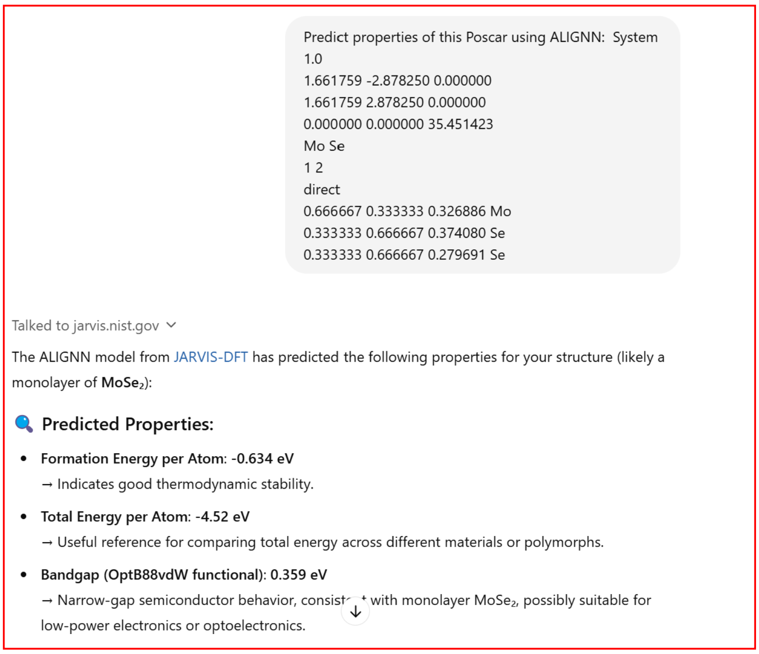

Fig. 3: Screenshot showcasing the use of graph neural network (pretrained ALIGNN models) from an atomic structure provided in POSCAR format. The custom GPT assistant interprets the POSCAR and returns quantitative predictions alongside contextual explanations.

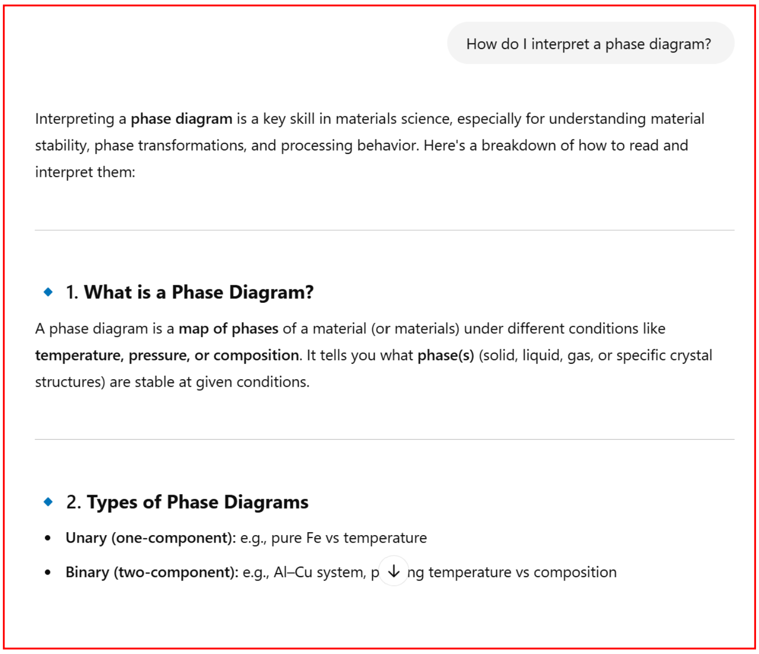

Fig. 4: Screenshot illustrating a domain-specific interaction with the custom GPT assistant.

3.3 Technical writing and question-answering

A core feature of CME is its ability to support technical writing and structured question-answering within the domain of materials science-capabilities that are currently lacking in infrastructures such as JARVIS, Materials Project, AFLOW, and others. Traditional scientific writing-whether for research papers, proposals, or technical reports-requires not only subject matter expertise but also precise language, contextual formatting, and domain-specific terminology. CME bridges this gap by enabling users to generate, refine, and structure scientific content through interactive dialogue that incorporates real-time data retrieval, LaTeX/Markdown rendering, and references to trusted materials databases.

The assistant is trained to interpret a variety of prompts, ranging from concise questions (e.g., "How do I interpret a phase diagram?" as shown in Fig. 4) to more elaborate tasks (e.g., "Generate a paragraph explaining how vacancy defects affect thermal conductivity in ceramics"). Human oversight and context-aware prompting strategies remain essential for reviewing and validating AI-generated content, ensuring adherence to scientific rigor and ethical standards. To enhance the reliability and relevance of its responses, CME is also equipped with real-time web search capabilities, enabling it to retrieve the most current information from institutional repositories, preprint servers, and official materials databases.

3.4 Limitations

While Material Explorer represents a significant advancement in AI-assistant development for materials research, several limitations warrant consideration. The current framework relies on OpenAI's infrastructure, which reduces the maintenance burden on individual developers. However, this reliance introduces risks related to vendor lock-in and potential platform changes. To address this, I am actively developing an alternative platform that offers greater control and long-term stability. A prototype is available at https://atomgpt.org. Furthermore, I plan to implement version control and provide comprehensive documentation to support the reproducibility of results as the underlying models and databases evolve.

Additionally, the assistant's responses depend heavily on the accuracy and completeness of external databases and pre-trained models, which may introduce bias or propagate existing data quality issues. The reliance on structured APIs may also limit functionality and introduce latency, particularly when accessing non-standardized or proprietary datasets.

Moreover, ethical concerns remain such as the use of non-open access GPT-4o mode, issues of data provenance, interpretability of AI-generated predictions, and the potential for hallucination (though partially mitigated by API integration). These areas require ongoing oversight and future development. Material Explorer currently surfaces multiple sources with their provenance and includes disclaimers when discrepancies are detected. The assistant does not override or average values; instead, it flags conflicts and encourages users to explore them further via source links. Future versions will integrate rule-based prioritization and confidence scoring.

4 Future Directions

Looking forward, CME is designed for continuous expansion. Planned capabilities include protein sequence-to-structure prediction, enabling researchers to explore bio-inspired materials and hybrid systems. Integration with powerful atomistic manipulation tools such as JARVIS-Tools and Atomic Simulation Environment (ASE)[25] will support the automated workflows such as creation of vacancies, surfaces, and heterointerfaces [26], facilitating studies of defect engineering and interfacial phenomena. Experimental data interpretation is also a priority, with future support for analyzing X-ray diffraction (XRD) patterns and spectroscopic signatures through guided workflows using tools such as DiffractGPT [21] and MicroscopyGPT [22]. In addition, high-performance computing (HPC) workflows will be incorporated to allow structure submissions and remote execution of simulations via APIs. Finally, its automated literature review capabilities will retrieve and summarize relevant scientific papers such as those identified using the ChemNLP tool [7] for chemistry-specific information, thereby expediting research workflows and reducing the time spent on manual searches.

5 Conflict of interest

The author declares that there are no conflicts of interest.

References

[1] Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F.L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., et al.: Gpt-4 technical report. arXiv preprint arXiv:2303.08774 (2023)

[2] Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., et al.: Llama: Open and efficient foundation language models. arXiv preprint arXiv:2302.13971 (2023)

[3] Team, G., Anil, R., Borgeaud, S., Alayrac, J.-B., Yu, J., Soricut, R., Schalkwyk, J., Dai, A.M., Hauth, A., Millican, K., et al.: Gemini: a family of highly capable multimodal models. arXiv preprint arXiv:2312.11805 (2023)

[4] Choudhary, K., DeCost, B., Chen, C., Jain, A., Tavazza, F., Cohn, R., Park, C.W., Choudhary, A., Agrawal, A., Billinge, S.J., *et al.*: Recent advances and applications of deep learning methods in materials science. npj Computational Materials **8**(1), 59 (2022)

[5] Jablonka, K.M., Ai, Q., Al-Feghali, A., Badhwar, S., Bocarsly, J.D., Bran, A.M., Bringuier, S., Brinson, L.C., Choudhary, K., Circi, D., *et al.*: 14 examples of how llms can transform materials science and chemistry: a reflection on a large language model hackathon. Digital discovery **2**(5), 1233–1250 (2023)

[6] Shen, Y., Heacock, L., Elias, J., Hentel, K.D., Reig, B., Shih, G., Moy, L.: ChatGPT and other large language models are double-edged swords. Radiological Society of North America (2023)

[7] Wines, D., Gurunathan, R., Garrity, K.F., DeCost, B., Biacchi, A.J., Tavazza, F., Choudhary, K.: Recent progress in the jarvis infrastructure for next-generation data-driven materials design. Applied Physics Reviews **10**(4) (2023)

[8] Jain, A., Ong, S.P., Hautier, G., Chen, W., Richards, W.D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., et al.: Commentary: The materials project: A materials genome approach to accelerating materials innovation. APL materials **1**(1) (2013)

[9] Schmidt, J., Cerqueira, T.F., Romero, A.H., Loew, A., Jäger, F., Wang, H.-C., Botti, S., Marques, M.A.: Improving machine-learning models in materials science through large datasets. Materials Today Physics **48**, 101560 (2024)

[10] Kabir, A., Shah, S., Haddad, A., Raper, D.M.: Introducing our custom gpt: An example of the potential impact of personalized gpt builders on scientific writing. World Neurosurgery **193**, 461–468 (2025)

[11] Mishra, V., Singh, S., Ahlawat, D., Zaki, M., Bihani, V., Grover, H.S., Mishra, B., Miret, S., Krishnan, N., et al.: Foundational large language models for materials research. arXiv preprint arXiv:2412.09560 (2024)

[12] M. Bran, A., Cox, S., Schilter, O., Baldassari, C., White, A.D., Schwaller, P.: Augmenting large language models with chemistry tools. Nature Machine Intelligence **6**(5), 525–535 (2024)

[13] Chiang, Y., Hsieh, E., Chou, C.-H., Riebesell, J.: Llamp: Large language model made powerful for high-fidelity materials knowledge retrieval and distillation. arXiv preprint arXiv:2401.17244 (2024)

[14] Choudhary, K., Garrity, K.F., Reid, A.C., DeCost, B., Biacchi, A.J., Hight Walker, A.R., Trautt, Z., Hattrick-Simpers, J., Kusne, A.G., Centrone, A., *et al.*: The joint automated repository for various integrated simulations (jarvis) for data-driven materials design. npj computational materials **6**(1), 173 (2020)

[15] Choudhary, K., DeCost, B.: Atomistic line graph neural network for improved materials property predictions. npj Computational Materials **7**(1), 185 (2021)

[16] Choudhary, K., DeCost, B., Major, L., Butler, K., Thiyagalingam, J., Tavazza, F.: Unified graph neural network force-field for the periodic table: solid state applications. Digital Discovery **2**(2), 346–355 (2023)

[17] Evans, M.L., Bergsma, J., Merkys, A., Andersen, C.W., Andersson, O.B., Beltrán, D., Blokhin, E., Boland, T.M., Balderas, R.C., Choudhary, K., *et al.*: Developments and applications of the optimade api for materials discovery, design, and data exchange. Digital Discovery **3**(8), 1509–1533 (2024)

[18] Sitzmann, M., Filippov, I., Nicklaus, M.: Internet resources integrating many small-molecule databases1. SAR and QSAR in Environmental Research **19**(1-2), 1–9 (2008)

[19] arXiv API. https://arxiv.org/help/api. Accessed: 2025-06-03

[20] Choudhary, K.: Atomgpt: Atomistic generative pretrained transformer for forward and inverse materials design. The Journal of Physical Chemistry Letters **15**(27), 6909–6917 (2024)

[21] Choudhary, K.: Diffractgpt: Atomic structure determination from x-ray diffraction patterns using a generative pretrained transformer. The Journal of Physical Chemistry Letters **16**, 2110–2119 (2024)

[22] Choudhary, K.: MicroscopyGPT: Generating atomic-structure captions from microscopy images of 2D materials with vision-language transformers (2025)

[23] Curtarolo, S., Setyawan, W., Hart, G.L., Jahnatek, M., Chepulskii, R.V., Taylor, R.H., Wang, S., Xue, J., Yang, K., Levy, O., *et al.*: Aflow: An automatic framework for high-throughput materials discovery. Computational Materials Science **58**, 218–226 (2012)

[24] Saal, J.E., Kirklin, S., Aykol, M., Meredig, B., Wolverton, C.: Materials design and discovery with high-throughput density functional theory: the open quantum materials database (oqmd). Jom **65**, 1501–1509 (2013)

[25] Larsen, A.H., Mortensen, J.J., Blomqvist, J., Castelli, I.E., Christensen, R., Dułak, M., Friis, J., Groves, M.N., Hammer, B., Hargus, C., *et al.*: The atomic simulation environment—a python library for working with atoms. Journal of Physics: Condensed Matter **29**(27), 273002 (2017)

[26] Choudhary, K., Garrity, K.F.: Intermat: accelerating band offset prediction in semiconductor interfaces with dft and deep learning. Digital Discovery **3**(7), 1365–1377 (2024)