# Smart Material AI: artificial intelligence platform for the discovery of novel materials

**Anonymous Author(s)**
Affiliation
Address
email

## Abstract

Artificial intelligence (AI) has become an essential tool in the field of materials science, however material scientists often face friction when attempting to integrate AI into their research practices. This paper presents Smart Material (SM), an innovative platform designed to accelerate and optimize the discovery of new materials, leveraging deep learning techniques to facilitate more efficient and precise analysis. SM includes capabilities for querying and learning from large experimental and theoretical datasets, generating entirely novel material prototypes, and predicting the properties of these new materials before experimental synthesis and characterization. Currently, SM is powering projects ranging from the naval industry, through the simulation of corrosion in vessel materials, to the energy industry, with the identification of more efficient materials for lithium-ion batteries. The results obtained so far with this platform have been promising, leading to the discovery of a promising lithium-ion battery cathode. This paper will discuss the platform's methodology in detail and the findings achieved thus far, along with the potential of this technology to revolutionize the field of materials discovery.

## 1 Introduction

### 1.1 General Context

We interact daily with thousands of specialized materials that are key components of advanced technology and infrastructure. The discovery of new materials is a cornerstone of innovation across a wide range of industries and domains, from information technology and energy to medicine and aerospace engineering. Despite its importance, material discovery remains heavily reliant on experimentation and serendipity. Since the development of new materials currently involves studies spanning 10 to 20 years with very high costs, finding suitable materials is the bottleneck in the transition to a clean energy future (1; 2).

Recently, interest in applying Artificial Intelligence (AI) techniques and its subclasses to generate new materials with desired properties has greatly increased (3; 4; 5; 6). This set of AI methods is often inspired by their outstanding success in processing images and texts, such as DALL-E 2 (7) or GPT-4 (8), to name a few examples, and is expanding into the field of materials science and engineering.

Therefore, implementing AI in material discovery not only promises to accelerate the pace of innovation and reduce costs but also opens new frontiers of knowledge and application for materials science and engineering (9).

Submitted to 38th Conference on Neural Information Processing Systems (NeurIPS 2024). Do not distribute.

### 1.2 Problem

The primary challenge faced by researchers and engineers in the field of materials science is the high time and financial investment required by the traditional material discovery process. Even with the most advanced experimental synthesis techniques, characterizing new materials is a laborious process that can take years, if not decades, of systematic exploration.

Moreover, material discovery is not limited to identifying new materials; it also involves optimizing their properties for specific applications. This requires understanding and mastering the relationship between a material's structure, its properties, and its performance in specific applications, adding another layer of complexity to the problem.

In this context, Smart Material AI platform (SM, https://smartmaterial.hi-iberia.es) can have a significant impact. By utilizing machine learning and other AI techniques, it is possible to analyze and learn from large experimental and theoretical datasets, identify hidden patterns and correlations, reduce the search space for new materials, and accurately infer their properties before synthesis. Additionally, by combining AI with material simulation techniques, it is possible to systematically and efficiently explore the vast space of material compositions and configurations. This not only speeds up the discovery process but also helps optimize material performance for specific applications.

### 1.3 Objectives

The main objective of this paper is to present the various services, functions, and technologies of the SM platform (see Section 1.5 and 2), as well as to showcase its efficacy and performance in the discovery and optimization of new materials (see Section 3). Our purpose is to demonstrate how this platform can overcome the inherent limitations of traditional methods, significantly accelerating the material discovery and optimization process by accurately predicting the properties of potential materials before synthesis. Ultimately, the goal of SM is to foster innovation in the field of materials science and contribute to the advancement of various industries that rely on the continuous evolution and improvement of materials.

### 1.4 Application in Energy Industry

The application of AI has the potential to revolutionize a wide range of industries. One industry where SM has already demonstrated a remarkable impact is the energy industry.

The energy industry is undergoing a crucial transformation in its efforts to achieve sustainability and climate neutrality, with batteries playing a pivotal role in this shift. However, the industry faces significant challenges, such as the need to improve energy density, reliability, safety, and sustainability of batteries while maintaining cost-effectiveness on a large scale (10; 11). In response to these challenges, the LiOn-HD project was conceived with the aim of significantly improving lithium-ion batteries (LIBs) through the investigation of advanced active and inactive materials and their synergistic combinations for different components of the electrochemical cell.

In the field of cathode materials, several classes of materials have stood out for their performance, each with its own advantages and limitations. Nickel, manganese, and cobalt oxides (NMC) are known for their high energy density and overall performance, although they present challenges in terms of cost and sustainability due to the use of cobalt (12). Lithium cobalt oxide (LCO) is another popular option with even higher energy density, but it raises similar concerns regarding cost, sustainability, and safety (13). Lithium manganese oxide (LMO) offers a more economical and safer option but at the expense of lower energy density (14). Lastly, lithium iron phosphate (LFP) is notable for its excellent safety and stability, as well as its greater sustainability due to the use of iron instead of more expensive and less sustainable metals (15).

During the development of the LiOn-HD project, the use of SM enabled the identification of a new cathode material with great potential to maximize the performance of lithium-ion batteries while minimizing their cost and environmental impact (Section 3).

### 1.5 Smart Material AI Platform

SM platform features an intuitive and user-friendly web interface that integrates the latest AI techniques and computational methods into a suite of specific and powerful services (Figure 1). Each

of these services focuses on providing the user with the best experience based on their interests and objectives.

In this way, SM offers the following main functionalities:

- **Databases Access**. Query a comprehensive, ever-growing database of material properties from experiments and simulations.
- **Analytic Calculations Methods**. Evaluate material behavior for specific applications using non-AI algorithms.
- **AI Property Prediction**. Utilize predictive models to accurately forecast the properties of various materials based on extensive data.
- **Synthesis Procedures**. Search for existing material synthesis protocols to inspire new material synthesis.
- **Insight Explorer**. Access the latest research conducted by the SM team directly.
- **Density Functional Theory**. Optimize material energy using theoretical computational methods.
- **Autonomous Research Agent**. Employ Large Language Models (LLMs) to enhance material discovery with insights from scientific literature.
- **AI Materials Discovery**. Integrate all the previous methods and algorithms in a workflow designed to optimize the identification and discovery of new materials for various applications.

# 2 Methods

All of the services offered by SM can be utilized individually or seamlessly integrated with the SM Materials Discovery service, providing a flexible and comprehensive solution for material research. Here, the workflow for discovering new materials is summarized in Figure 2:

- **Initial Phase**: Define a set of materials to be studied. These may come from popular databases (see Section 2.1) or be novel materials generated by AI models (Section 2.2).
- **Property Prediction**: Calculate and predict the most relevant properties for the application using various computational methods and AI models (Section 2.3).
- **Selection for Characterization**: Assign virtual properties to each material and select the most promising ones for experimental verification, using thousands of synthesis protocols in the SM database for guidance (Sections 2.4 and 2.5).
- **Optimization**: Use efficient optimization algorithms (closed-loop optimization) to iteratively identify materials that best meet the desired properties, without human intervention.

Users can explore results intuitively at each phase:

- **Initial Exploration**: SM offers 3D visualization of material structures and various filtering tools.
- **Detailed Analysis**: SM provides options ranging from 2D plots for comparing numerical properties to specific visualization panels for complex properties.
- **Material Selection**: SM features a flexible tool for interactively weighting properties to generate an objective function, helping users identify materials that maximize this function in each optimization iteration.

## 2.1 Databases

In SM, material databases such as the Materials Project (16), Open Quantum Materials Database (17), Crystallography Open Database (18), 2-D Perovskite Database (19), Screening Platform for Solid Electrolytes (20) and GNoME (21) play a pivotal role in enhancing novel materials discovery. These databases provide a wealth of information on known materials, including their crystal structures, chemical compositions, and physical properties. By leveraging these databases, users can access a vast repository of materials data to inform their research and guide the selection of materials for further study (Figure 3).


Figure 1: Smart Material AI platform homepage.

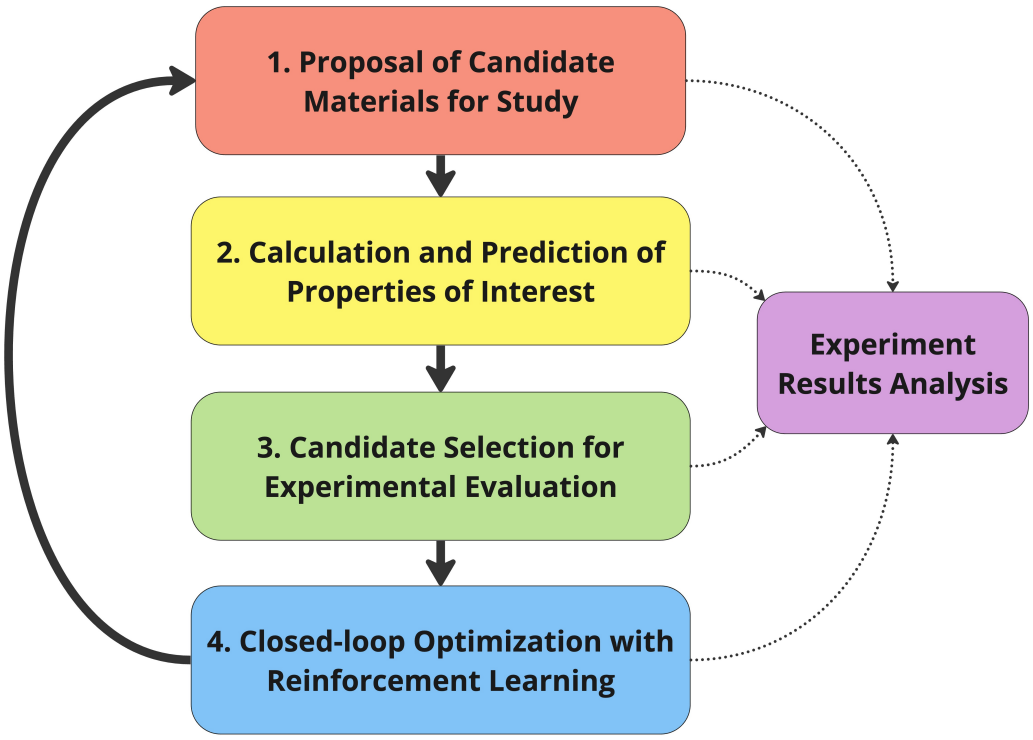


Figure 2: Workflow for the discovery of new materials in SM.

### 2.2 Generative AI

In addition to leveraging material databases, SM harnesses advanced AI methods such as FTCP (22) and CDVAE (23) to generate novel materials. FTCP is a variational autoencoder that learns the underlying distribution of a dataset of inorganic crystals and can generate new samples with similar characteristics. By training FTCP on vast repositories of materials data, SM can effectively explore the latent space of material compositions and configurations, generating novel materials with desired properties. Similarly, CDVAE is a diffusion VAE architecture which models the dynamics of crystal data generation processes, offering another avenue for generating novel materials by simulating the sequential evolution of material structures. By employing these cutting-edge methods, SM can systematically explore the vast space of potential materials, accelerating the discovery of novel compounds with tailored properties for various applications.

### 2.3 Predictive AI and Analytical Calculations

SM employs a combination of predictive AI methods and analytical calculations to evaluate both novel and existing materials comprehensively. Utilizing state-of-the-art graph-based algorithms such as Crystal Graph Convolutional Neural Networks (24), Materials Graph Network (25), and Atomic Line Graph Neural Networks (26), SM can accurately predict various material properties based on their structural features and chemical compositions. These AI models are trained on large datasets of known materials to learn intricate relationships between their structures and properties, enabling SM to efficiently assess the performance of novel and existing materials.

Additionally, SM also utilizes analytical calculations to evaluate material behaviour. Though slower, these algorithms also provide valuable insights for certain applications. As an example, methods such as Crystal Analysis by Voronoi Decomposition (27) or Bond Valence Site Energies (28) have been extensively used in the search of better lithium-ion battery cathodes. By integrating predictive AI methods with analytical calculations, SM offers a comprehensive framework for evaluating materials, facilitating informed decision-making in material design and discovery processes.

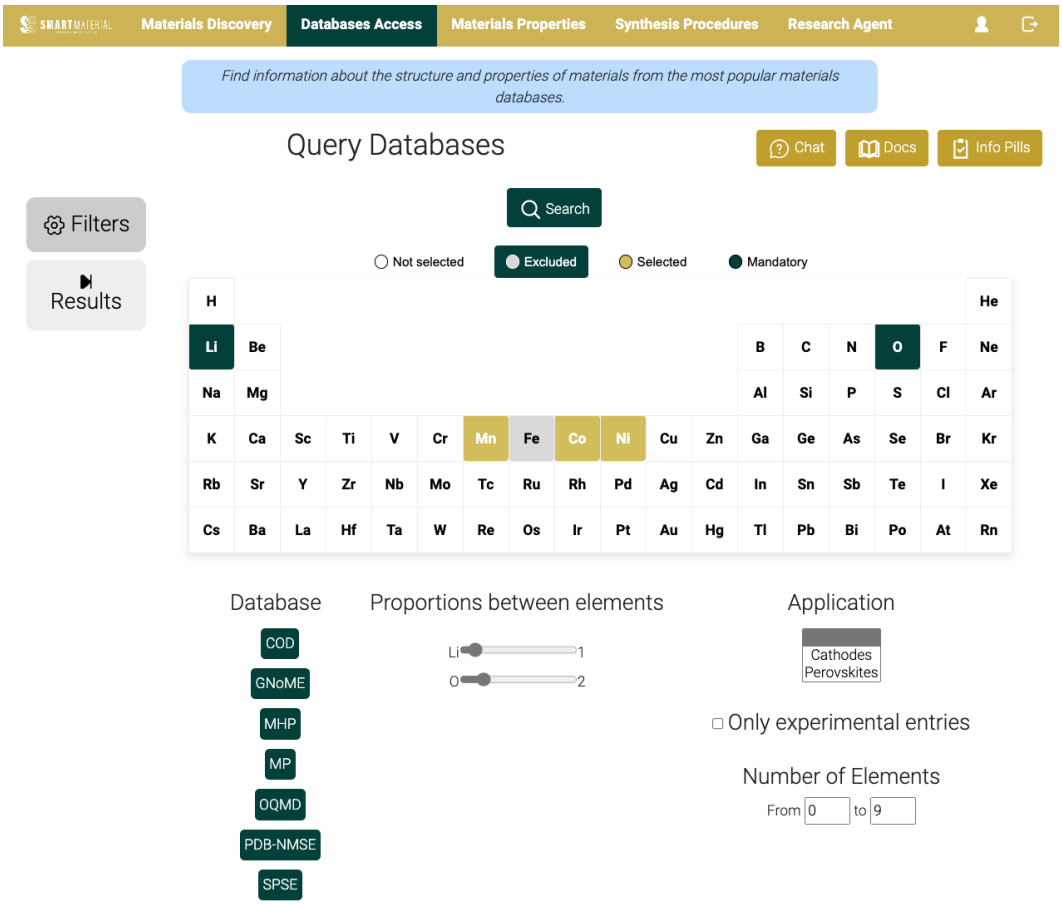


Figure 3: Example of database search for the discovery of new cathodes. In this experiment, all lithium oxides and at least one of the chemical elements highlighted in yellow were studied.

### 2.4 Autonomous Research

Incorporating LLM-based autonomous agents, SM enhances the process of material discovery by providing material scientists with unprecedented access to insights from scientific literature. These advanced language models are trained on vast corpora of texts, enabling them to comprehend and extract valuable information regarding the properties, synthesis methods, and applications of various materials (29; 30). Leveraging this capability, SM's autonomous agents continuously scan and analyze the latest research publications, extracting key insights and identifying promising materials that align with specific research objectives. Through the seamless integration of LLM-based autonomous agents, SM enhances the efficiency and effectiveness of material scientists, facilitating faster and more informed decision-making in the pursuit of novel materials with tailored properties.

### 2.5 Synthesis Procedures

SM harnesses a comprehensive database of synthesis protocols for material synthesis, facilitating the last steps of the material discovery process (31; 32). This extensive repository contains a wealth of information on experimental procedures, reaction conditions, and synthesis techniques for a wide range of materials. Leveraging this database, SM enables material scientists to access a vast array of proven synthesis methods, accelerating the transition from theoretical predictions to practical implementation.

Additionally, SM utilizes LLMs to infer synthesis protocols for novel materials that have not yet been synthesized. By analyzing the structural and compositional characteristics of these materials and drawing insights from existing synthesis protocols, LLMs can generate tailored recommendations

for experimental procedures. This integration of synthesis databases with LLM-driven synthesis inference empowers material scientists to explore novel materials with confidence, streamlining the final stages of material discovery and expediting the translation of theoretical concepts into tangible materials.

### 2.6 Compute Resources

SM AI platform is deployed on an on-premises Kubernetes cluster using k3s, which consists of 1 master node and 11 worker nodes. Collectively, these nodes provide a total of 218 CPU cores, 1200 GB of RAM, and 15 GPUs to ensure high computational performance and scalability. The cluster integrates with a Network-Attached Storage (NAS) solution, offering 60TB of storage capacity for datasets, models, and other resources.

A typical material discovery experiment on the platform utilizes 10 CPU cores and less than 10 GB of RAM, completing in under 3 minutes. The efficient deployment of models within the cluster ensures that computational resources are distributed effectively, preventing any single experiment from monopolizing the CPU or memory. SM platform is designed to maximize resource utilization and minimize latency during AI inference tasks, enabling rapid execution and high throughput of experiments.

## 3 Results

SM has demonstrated exceptional performance in the discovery and optimization of materials for battery cathodes. This section describes some of the most significant achievements to date. SM's ability to analyze extensive datasets, identify hidden patterns and correlations, and accurately predict the properties of potential materials has led to significant discoveries and the optimization of a cathode that outperforms existing ones in terms of performance and sustainability. These advancements are laying the groundwork for a new generation of more efficient, durable, and sustainable batteries.

In cooperation with a team of materials science experts and utilizing SM platform, we have discovered a high-performance cathode material that surpasses current limitations in terms of energy density and sustainability. The platform, with its capability to process and learn from massive datasets and its powerful AI-based prediction engine, was able to identify combinations of elements and structures that had not been previously considered. This new cathode material, whose details remain confidential for intellectual property reasons, has demonstrated an energy capacity of 290 mAh/g in laboratory tests, which is 45% higher than the best lithium batteries with NMC cathodes, whose maximum achieved capacity is 200 mAh/g (Figure 4). Moreover, it is composed of naturally abundant and low-cost elements, enhancing its feasibility for large-scale production and reducing its environmental impact by avoiding the need for nickel, cobalt, and manganese found in current cathodes.

The promising results of the LiOn-HD project demonstrate the potential of AI techniques to revolutionize materials science, overcoming the limitations of traditional experimental approaches and optimizing resource usage. AI not only accelerates the material discovery process but also enables innovation and new technological advancements.

## 4 Discussion

### 4.1 Next Steps

Despite the success achieved, SM continues to evolve, equipping itself with cutting-edge features and technologies. Looking ahead, we have identified three primary areas for enhancement that promise to significantly advance the capabilities of the SM platform:

- **Structural Relaxation**: Implementing AI-driven algorithms for structural relaxation will streamline the optimization process of material structures. This enhancement will allow researchers to achieve optimal configurations more quickly and easily, significantly reducing the time and effort required for material development (33).
- **Molecular Dynamics Simulation**: Integrating molecular dynamics simulations into the SM platform will enable users to gain a deeper understanding of how different materials interact

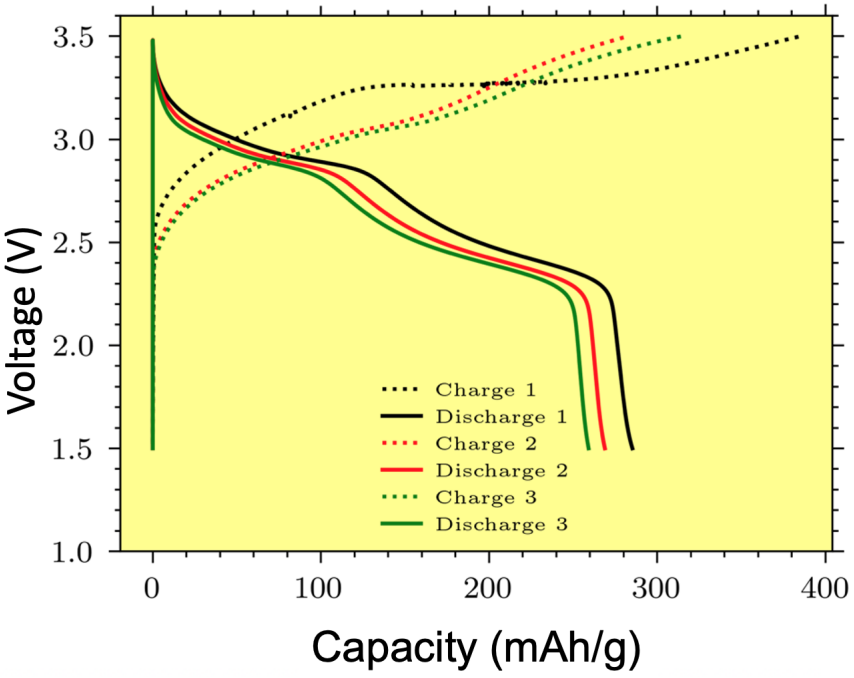


Figure 4: Galvanostatic cycling of a lithium-ion button cell with a cathode made of the discovered material and a metallic lithium reference anode. The cycling rate was C/40.

within complex systems. This addition will provide researchers with critical insights into the behavior of materials under various conditions, facilitating the design of more robust and efficient materials (34).

- **Multi-Scale Modeling**: Developing multi-scale modeling capabilities that integrate atomic-level simulations with macroscopic property predictions could provide a more comprehensive understanding of material behavior. This approach would enable researchers to bridge the gap between microscopic phenomena and macroscopic applications (35).
- **LLM-based Agents**: Incorporating LLM-based agents into the SM platform will enhance the overall user experience. These intelligent agents will make SM services more efficient and user-friendly, thereby boosting the effectiveness of material discovery efforts. By simplifying interactions and automating complex tasks, LLM-based agents will empower researchers to focus more on innovation and less on routine processes.

### 4.2 Current Projects

Building on the success of the LiOn-HD project, the NanomatIA initiative has emerged as a significant continuation, broadening SM's focus to encompass additional components of lithium batteries. This project aims to leverage AI, nano and biotechnology to enhance the overall performance and safety of lithium batteries. By applying advanced machine learning techniques, NanomatIA seeks to identify novel material compositions and configurations that can improve battery efficiency, longevity, and sustainability.

The naval industry stands to gain substantial benefits from the SM platform, particularly in addressing the critical challenge of marine corrosion. Corrosion is a pervasive issue that impacts the longevity and safety of ships and their components, leading to high maintenance and repair costs and potentially compromising the structural integrity of vessels. The hulls of ships and internal piping systems are especially vulnerable to degradation due to constant exposure to harsh marine environments.

Given the severity of this issue, SM is being employed in Kimiko project to predict and model corrosion processes in vessel steels. By utilizing AI to analyze and forecast these processes, the platform can identify materials with enhanced corrosion resistance. This knowledge enables the development of new materials that are more durable in marine conditions. Additionally, SM can sug-

gest contingency measures, such as the inclusion or replacement of specific parts, or recommending optimal steel thicknesses to mitigate corrosion risks.

Through these advancements, SM not only helps in extending the lifespan of naval vessels but also ensures operational safety by maintaining the structural integrity of the ships. The integration of AI in predicting and combating corrosion represents a transformative step in the naval industry, leading to reduced maintenance costs and improved safety standards.

These initiatives illustrate the versatile application of the SM platform across different industries, showcasing its potential to drive innovation and efficiency in both energy storage and maritime technology. As SM continues to evolve, its impact on material science and engineering is set to expand, offering robust solutions to complex industrial challenges.

### 4.3 Impact Statement

SM has the potential to transform the way materials discovery is conducted by incorporating AI into every stage of the process. With this AI platform, we are overcoming the inherent limitations of conventional experiment-based approaches, significantly reducing the time and resources needed for the exploration and optimization of materials.

The impact of SM can be seen not only in the speed and efficiency of materials discovery but also in the quality and novelty of the materials discovered. This platform has demonstrated its ability to predict materials with unique and optimized properties, opening the door to new applications and technological advancements. Furthermore, by providing researchers with the ability to virtually explore the materials space, we can stimulate creativity and innovation, expanding the boundaries of what is considered possible in materials science.

Lastly, SM also democratizes the materials discovery process. By offering an intuitive and user-friendly interface, it promotes a more collaborative style of research in which a larger percentage of stakeholders can participate. This includes materials scientists who are accustomed to using computational tools, as well as those with less training in this area, including students and novice researchers, and professionals from other related scientific fields.

## 5 Conclusions

Smart Material has proven to be a powerful tool in materials development, successfully identifying a new cathode for lithium batteries that promises higher energy density than current solutions while reducing environmental and economic impact. This achievement highlights SM's potential to revolutionize materials engineering across various industries, including energy and naval sectors. By integrating a comprehensive suite of services —including but not limited to generative AI, predictive AI and analytical calculations, and LLM-based autonomous agents— SM enables the efficient development of high-quality materials. The platform's ability to leverage vast material databases, advanced optimization algorithms, and extensive synthesis protocols ensures a streamlined and effective discovery process. By addressing the urgent technological needs of today, SM paves the way for innovative solutions and significant advancements in material science.

## 6 References

## References

[1] Goodenough, J. B. and Kim, Y. Challenges for Rechargeable Li Batteries. Chemistry of Materials, 22(3):587–603, February 2010. ISSN 0897-4756, 1520-5002. doi: 10. 1021/cm901452z. URL https://pubs.acs.org/ doi/10.1021/cm901452z.

[2] Buriak, J. M. and Toro, C. Rising to the Challenge: John B. Goodenough and Youngsik Kim, and "Challenges for Rechargeable Li Batteries": Chemistry of Materials ' 1k Club. Chemistry of Materials, 27(15):5149–5150, August 2015. ISSN 0897-4756, 1520-5002. doi: 10.1021/acs.chemmater.5b02863. URL https://pubs.acs. org/doi/10.1021/acs.chemmater.5b02863.

[3] Liu, Y., Zhao, T., Ju, W., and Shi, S. Materials discovery and design using machine learning. Journal of Materiomics, 3(3):159–177, September 2017. ISSN 23528478. doi: 10.1016/j.jmat.2017.08.002. URL https://linkinghub.elsevier. com/retrieve/pii/S2352847817300515.

[4] Sendek, A. D., Yang, Q., Cubuk, E. D., Duerloo, K.- A. N., Cui, Y., and Reed, E. J. Holistic computational structure screening of more than 12 000 candidates for solid lithium-ion conductor materials. Energy Environmental Science, 10(1):306–320, 2017. ISSN 1754- 5692, 1754-5706. doi: 10.1039/C6EE02697D. URL http://xlink.rsc.org/?DOI=C6EE02697D.

[5] Batra, R., Song, L., and Ramprasad, R. Emerging materials intelligence ecosystems propelled by machine learning. Nature Reviews Materials, 6(8):655–678, August 2021. ISSN 2058-8437. doi: 10.1038/ s41578-020-00255-y. URL https://www.nature. com/articles/s41578-020-00255-y.

[6] Lombardo, T., Duquesnoy, M., El-Bouysidy, H., A re´n, F., Gallo-Bueno, A., Jørgensen, P. B., Bhowmik, A., Demorti‘ ere, A., Ayerbe, E., Alcaide, F., Reynaud, M., Carrasco, J., Grimaud, A., Zhang, C., Vegge, T., Johansson, P., and Franco, A. A. Artificial Intelligence Applied to Battery Research: Hype or Reality? Chemical Reviews, 122(12):10899–10969, June 2022. ISSN 0009-2665, 1520-6890. doi: 10.1021/acs.chemrev. 1c00108. URL https://pubs.acs.org/doi/10.1021/acs.chemrev.1c00108.

[7] Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., and Chen, M. Hierarchical Text-Conditional Image Generation with CLIP Latents, April 2022. URL http://arxiv. org/abs/2204.06125. arXiv:2204.06125 [cs].

[8] OpenAI, Achiam, J., Adler, S., Agarwal, S., Ahmad, L., Akkaya, I., Aleman, F. L., Almeida, D., Altenschmidt, J., Altman, S., Anadkat, S., Avila, R., Babuschkin, I., Balaji, S., Balcom, V., Baltescu, P., Bao, H., Bavarian, M., Belgum, J., Bello, I., Berdine, J., Bernadett-Shapiro, G., Berner, C., Bogdonoff, L., Boiko, O., Boyd, M., Brakman, A.-L., Brockman, G., Brooks, T., Brundage, M., Button, K., Cai, T., Campbell, R., Cann, A., Carey, B., Carlson, C., Carmichael, R., Chan, B., Chang, C., Chantzis, F., Chen, D., Chen, S., Chen, R., Chen, J., Chen, M., Chess, B., Cho, C., Chu, C., Chung, H. W., Cummings, D., Currier, J., Dai, Y., Decareaux, C., Degry, T., Deutsch, N., Deville, D., Dhar, A., Dohan, D., Dowling, S., Dunning, S., Ecoffet, A., Eleti, A., Eloundou, T., Farhi, D., Fedus, L., Felix, N., Fishman, S. P., Forte, J., Fulford, I., Gao, L., Georges, E., Gibson, C., Goel, V., Gogineni, T., Goh, G., Gontijo-Lopes, R., Gordon, J., Grafstein, M., Gray, S., Greene, R., Gross, J., Gu, S. S., Guo, Y., Hallacy, C., Han, J., Harris, J., He, Y., Heaton, M., Heidecke, J., Hesse, C., Hickey, A., Hickey,W., Hoeschele, P., Houghton, B., Hsu, K., Hu, S., Hu, X., Huizinga, J., Jain, S., Jain, S., Jang, J., Jiang, A., Jiang, R., Jin, H., Jin, D., Jomoto, S., Jonn, B., Jun, H., Kaftan, T., Kaiser, , Kamali, A., Kanitscheider, I., Keskar, N. S., Khan, T., Kilpatrick, L., Kim, J. W., Kim, C., Kim, Y., Kirchner, J. H., Kiros, J., Knight, M., Kokotajlo, D., Kondraciuk, , Kondrich, A., Konstantinidis, A., Kosic, K., Krueger, G., Kuo, V., Lampe, M., Lan, I., Lee, T., Leike, J., Leung, J., Levy, D., Li, C. M., Lim, R., Lin, M., Lin, S., Litwin, M., Lopez, T., Lowe, R., Lue, P., Makanju, A., Malfacini, K., Manning, S., Markov, T., Markovski, Y., Martin, B., Mayer, K., Mayne, A., Mc- Grew, B., McKinney, S. M., McLeavey, C., McMillan, P., McNeil, J., Medina, D., Mehta, A., Menick, J., Metz, L., Mishchenko, A., Mishkin, P., Monaco, V., Morikawa, E., Mossing, D., Mu, T., Murati, M., Murk, O., M´ely, D., Nair, A., Nakano, R., Nayak, R., Neelakantan, A., Ngo, R., Noh, H., Ouyang, L., O'Keefe, C., Pachocki, J., Paino, A., Palermo, J., Pantuliano, A., Parascandolo, G., Parish, J., Parparita, E., Passos, A., Pavlov, M., Peng, A., Perelman, A., Peres, F. d. A. B., Petrov, M., Pinto, H. P. d. O., Michael, Pokorny, Pokrass, M., Pong, V. H., Powell, T., Power, A., Power, B., Proehl, E., Puri, R., Radford, A., Rae, J., Ramesh, A., Raymond, C., Real, F., Rimbach, K., Ross, C., Rotsted, B., Roussez, H., Ryder, N., Saltarelli, M., Sanders, T., Santurkar, S., Sastry, G., Schmidt, H., Schnurr, D., Schulman, J., Selsam, D., Sheppard, K., Sherbakov, T., Shieh, J., Shoker, S., Shyam, P., Sidor, S., Sigler, E., Simens, M., Sitkin, J., Slama, K., Sohl, I., Sokolowsky, B., Song, Y., Staudacher, N., Such, F. P., Summers, N., Sutskever, I., Tang, J., Tezak, N., Thompson, M. B., Tillet, P., Tootoonchian, A., Tseng, E., Tuggle, P., Turley, N., Tworek, J., Uribe, J. F. C., Vallone, A., Vijayvergiya, A., Voss, C., Wainwright, C., Wang, J. J., Wang, A., Wang, B., Ward, J., Wei, J., Weinmann, C. J., Welihinda, A., Welinder, P., Weng, J., Weng, L., Wiethoff, M., Willner, D., Winter, C., Wolrich, S., Wong, H., Workman, L., Wu, S., Wu, J., Wu, M., Xiao, K., Xu, T., Yoo, S., Yu, K., Yuan, Q., Zaremba, W., Zellers, R., Zhang, C., Zhang, M., Zhao, S., Zheng, T., Zhuang, J., Zhuk, W., and Zoph, B. GPT-4 Technical Report, March 2024. URL http://arxiv.org/abs/ 2303.08774. arXiv:2303.08774 [cs].

[9] Li, J., Lim, K., Yang, H., Ren, Z., Raghavan, S., Chen, P.-Y., Buonassisi, T., and Wang, X. AI Applications through the Whole Life Cycle of Material Discovery. Matter, 3(2):393–432, August 2020. ISSN 25902385. doi: 10.1016/j.matt.2020. 06.011. URL https://linkinghub.elsevier.com/retrieve/pii/S2590238520303015.

[10] Manthiram, A. Materials Challenges and Opportunities of Lithium Ion Batteries. The Journal of Physical Chemistry Letters, 2(3):176–184, February 2011. doi: 10.1021/jz1015422. URL https://doi.org/10. 1021/jz1015422. Publisher: American Chemical Society.

[11] Masias, A., Marcicki, J., and Paxton, W. A. Opportunities and Challenges of Lithium Ion Batteries in Automotive Applications. ACS Energy Letters, 6(2): 621–630, February 2021. doi: 10.1021/acsenergylett. 0c02584. URL https://doi.org/10.1021/ acsenergylett.0c02584. Publisher: American Chemical Society

[12] Zhao, W., Zheng, J., Zou, L., Jia, H., Liu, B., Wang, H., Engelhard, M. H., Wang, C., Xu, W., Yang, Y., and Zhang, J.-G. High Voltage Operation of Ni-Rich NMC Cathodes Enabled by Stable Electrode/Electrolyte Interphases. Advanced Energy Materials, 8(19):1800297, 2018. ISSN 1614-6840. doi: 10.1002/aenm.201800297. URL https://onlinelibrary.wiley.com/ doi/abs/10.1002/aenm.201800297. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/aenm.201800297.

[13] Mizushima, K., Jones, P. C., Wiseman, P. J., and Goodenough, J. B. LixCoO2 (0<x<-1): A new cathode material for batteries of high energy density. Materials Research Bulletin, 15(6):783–789, June 1980. ISSN 0025-5408. doi: 10.1016/0025-5408(80)90012-4. URL https://www.sciencedirect.com/ science/article/pii/0025540880900124.

[14] Thackeray, M. M., Johnson, P. J., de Picciotto, L. A., Bruce, P. G., and Goodenough, J. B. Electrochemical extraction of lithium from LiMn2O4. Materials Research Bulletin, 19(2):179–187, February 1984. ISSN 0025-5408. doi: 10.1016/0025-5408(84)90088-6. URL https://www.sciencedirect.com/ science/article/pii/0025540884900886.

[15] Padhi, A. K., Nanjundaswamy, K. S., and Goodenough, J. B. Phospho-olivines as Positive-Electrode Materials for Rechargeable Lithium Batteries. Journal of The Electrochemical Society, 144(4):1188, April 1997. ISSN 1945-7111. doi: 10.1149/1. 1837571. URL https://iopscience.iop.org/ article/10.1149/1.1837571/meta. Publisher: IOP Publishing.

[16] Jain, A., Ong, S. P., Hautier, G., Chen, W., Richards, W. D., Dacek, S., Cholia, S., Gunter, D., Skinner, D., Ceder, G., and Persson, K. A. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. APL Materials, 1(1):011002, July 2013. ISSN 2166-532X. doi: 10.1063/1.4812323. URL https: //pubs.aip.org/aip/apm/article/119685.

[17] Kirklin, S., Saal, J. E., Meredig, B., Thompson, A., Doak, J. W., Aykol, M., R¨uhl, S., and Wolverton, C. The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies. npj Computational Materials, 1(1):1–15, December 2015. ISSN 2057-3960. doi: 10.1038/npjcompumats. 2015.10. URL https://www.nature.com/ articles/npjcompumats201510.

[18] Gra˘zulis, S., Chateigner, D., Downs, R. T., Yokochi, A. F. T., Quir´os, M., Lutterotti, L., Manakova, E., Butkus, J., Moeck, P., and Le Bail, A. Crystallography Open Database – an open-access collection of crystal structures. Journal of Applied Crystallography, 42(4):726–729, August 2009. ISSN 0021-8898. doi: 10.1107/S0021889809016690. URL https://scripts.iucr.org/cgi-bin/paper?S0021889809016690.

[19] Marchenko, E. I., Fateev, S. A., Petrov, A. A., Korolev, V. V., Mitrofanov, A., Petrov, A. V., Goodilin, E. A., and Tarasov, A. B. Database of Two-Dimensional Hybrid Perovskite Materials: Open-Access Collection of Crystal Structures, Band Gaps, and Atomic Partial Charges Predicted by Machine Learning. Chemistry of Materials, 32(17):7383–7388, September 2020. ISSN 0897-4756, 1520-5002. doi: 10.1021/acs.chemmater. 0c02290. URL https://pubs.acs.org/doi/10.1021/acs.chemmater.0c02290.

[20] He, B., Chi, S., Ye, A., Mi, P., Zhang, L., Pu, B., Zou, Z., Ran, Y., Zhao, Q., Wang, D., Zhang, W., Zhao, J., Adams, S., Avdeev, M., and Shi, S. High-throughput screening platform for solid

electrolytes combining hierarchical ion-transport prediction algorithms. Scientific Data, 7(1): 151, December 2020a. ISSN 2052-4463. doi: 10.1038/ s41597-020-0474-y. URL http://www.nature.com/articles/s41597-020-0474-y.

[21] Merchant, A., Batzner, S., Schoenholz, S. S., Aykol, M., Cheon, G., and Cubuk, E. D. Scaling deep learning for materials discovery. Nature, 624(7990):80–85, December 2023. ISSN 0028-0836, 1476-4687. doi: 10.1038/ s41586-023-06735-9. URL https://www.nature.com/articles/s41586-023-06735-9.

[22] Ren, Z., Tian, S. I. P., Noh, J., Oviedo, F., Xing, G., Li, J., Liang, Q., Zhu, R., Aberle, A. G., Sun, S., Wang, X., Liu, Y., Li, Q., Jayavelu, S., Hippalgaonkar, K., Jung, Y., and Buonassisi, T. An invertible crystallographic representation for general inverse design of inorganic crystals with targeted properties. Matter, 5(1):314–335, January 2022. ISSN 25902385. doi: 10.1016/j.matt.2021.11.032. URL https://linkinghub.elsevier. com/retrieve/pii/S2590238521006251.

[23] Xie, T., Fu, X., Ganea, O.-E., Barzilay, R., and Jaakkola, T. Crystal Diffusion Variational Autoencoder for Periodic Material Generation. Technical Report arXiv:2110.06197, arXiv, March 2022. URL http://arxiv.org/ abs/2110.06197. arXiv:2110.06197 [cond-mat, physics:physics] type: article.

[24] Xie, T. and Grossman, J. C. Crystal Graph Convolutional Neural Networks for an Accurate and Interpretable Prediction of Material Properties. Physical Review Letters, 120 (14):145301, April 2018. ISSN 0031-9007, 1079-7114. doi: 10.1103/PhysRevLett.120.145301. URL http://arxiv.org/abs/1710.10324. arXiv:1710.10324 [cond-mat].

[25] Chen, C., Ye, W., Zuo, Y., Zheng, C., and Ong, S. P. Graph Networks as a Universal Machine Learning Framework for Molecules and Crystals. Chemistry of Materials, 31(9):3564–3572, May 2019. ISSN 0897-4756, 1520-5002. doi: 10.1021/acs.chemmater.9b01294. URL https://pubs.acs.org/doi/10. 1021/acs.chemmater.9b01294.

[26] Choudhary, K. and DeCost, B. Atomistic Line Graph Neural Network for improved materials property predictions. npj Computational Materials, 7(1):185, December 2021. ISSN 2057-3960. doi: 10.1038/ s41524-021-00650-1. URL https://www.nature. com/articles/s41524-021-00650-1.

[27] He, B., Ye, A., Chi, S., Mi, P., Ran, Y., Zhang, L., Zou, X., Pu, B., Zhao, Q., Zou, Z., Wang, D., Zhang, W., Zhao, J., Avdeev, M., and Shi, S. CAVD, towards better characterization of void space for ionic transport analysis. Scientific Data, 7(1):153, December 2020b. ISSN 2052-4463. doi: 10.1038/ s41597-020-0491-x. URL http://www.nature. com/articles/s41597-020-0491-x.

[28] Adams, S. and Rao, R. P. Understanding Ionic Conduction and Energy Storage Materials with Bond-Valence- Based Methods. In Brown, I. D. and Poeppelmeier, K. R. (eds.), Bond Valences, pp. 129–159. Springer, Berlin, Heidelberg, 2014. ISBN 978-3-642-54968-7. doi: 10.1007/430 2013 137. URL https://doi.org/10.1007/430_2013_137.

[29] Bran, A. M., Cox, S., White, A. D., and Schwaller, P. Chem- Crow: Augmenting large-language models with chemistry tools, June 2023. URL http://arxiv.org/ abs/2304.05376. arXiv:2304.05376 [physics, stat].

[30] Thway, M., Y. Low, A. K., Khetan, S., Dai, H., Recatala- Gomez, J., Paul Chen, A., and Hippalgaonkar, K. Harnessing GPT-3.5 for text parsing in solid-state synthesis – case study of ternary chalcogenides. Digital Discovery, 3(2):328–336, 2024. doi: 10.1039/D3DD00202K. URL https://pubs.rsc.org/en/content/ articlelanding/2024/dd/d3dd00202k. Publisher: Royal Society of Chemistry.

[31] Kononova, O., Huo, H., He, T., Rong, Z., Botari, T., Sun, W., Tshitoyan, V., and Ceder, G. Author Correction: Text-mined dataset of inorganic materials synthesis recipes. Scientific Data, 6(1):273, November 2019. ISSN 2052-4463. doi: 10.1038/ s41597-019-0297-x. URL https://www.nature. com/articles/s41597-019-0297-x. Publisher: Nature Publishing Group.

[32] Karpovich, C., Pan, E., Jensen, Z., and Olivetti, E. Interpretable Machine Learning Enabled Inorganic Reaction Classification and Synthesis Condition Prediction. Chemistry of Materials, 35(3):1062–1079, February 2023. ISSN 0897-4756, 1520-5002. doi: 10.1021/acs.chemmater.2c03010. URL https://pubs.acs.org/doi/10.1021/acs.chemmater.2c03010.

[33] Deng, B., Zhong, P., Jun, K., Riebesell, J., Han, K., Bartel, C. J., and Ceder, G. CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling. Nature Machine Intelligence, 5(9):1031– 1041, September 2023. ISSN 2522-5839. doi: 10.1038/s42256-023-00716-3. URL https://www.nature. com/articles/s42256-023-00716-3.

[34] Batatia, I., Benner, P., Chiang, Y., Elena, A. M., Kov´acs, D. P., Riebesell, J., Advincula, X. R., Asta, M., Avaylon, M., Baldwin, W. J., Berger, F., Bernstein, N., Bhowmik, A., Blau, S. M., C˘arare, V., Darby, J. P., De, S., Della Pia, F., Deringer, V. L., Elijo˘sius, R., El-Machachi, Z., Falcioni, F., Fako, E., Ferrari, A. C., Genreith-Schriever, A., George, J., Goodall, R. E. A., Grey, C. P., Grigorev, P., Han, S., Handley, W., Heenen, H. H., Hermansson, K., Holm, C., Jaafar, J., Hofmann, S., Jakob, K. S., Jung, H., Kapil, V., Kaplan, A. D., Karimitari, N., Kermode, J. R., Kroupa, N., Kullgren, J., Kuner, M. C., Kuryla, D., Liepuoniute, G., Margraf, J. T., Magd˘au, I.-B., Michaelides, A., Moore, J. H., Naik, A. A., Niblett, S. P., Norwood, S. W., O'Neill, N., Ortner, C., Persson, K. A., Reuter, K., Rosen, A. S., Schaaf, L. L., Schran, C., Shi, B. X., Sivonxay, E., Stenczel, T. K., Svahn, V., Sutton, C., Swinburne, T. D., Tilly, J., van der Oord, C., Varga-Umbrich, E., Vegge, T., Vondr´ak, M., Wang, Y., Witt, W. C., Zills, F., and Cs´anyi, G. A foundation model for atomistic materials chemistry, March 2024. URL http://arxiv.org/abs/2401.00096. arXiv:2401.00096 [condmat, physics:physics].

[35] Giessen, Erik van der, Peter A. Schultz, Nicolas Bertin, Vasily V. Bulatov, Wei Cai, Gábor Csányi, Stephen M. Foiles, et al. «Roadmap on Multiscale Materials Modeling». Modelling and Simulation in Materials Science and Engineering 28, n.º 4 (marzo de 2020): 043001. https://doi.org/10.1088/1361-651X/ab7150.

# NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introduction clearly state the paper's contributions, including the integration of advanced technologies like Materials Discovery and Artificial Intelligence, and accurately reflect the scope of the work as presented throughout the paper.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: The paper does include section 4.1 discussing the limitations of the proposed AI platform.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory Assumptions and Proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [NA]

Justification: The paper does not include any theoretical results that require formal assumptions.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental Result Reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [No]

Justification: Experimental results obtained in this work are confidential.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [No]

Justification: Code for this AI platform is not open source but data in Smart Material Databases can be freely accessed.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental Setting/Details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [NA]

Justification: The paper does not include experiments that require specification of training or test details, such as data splits, hyperparameters, or optimizer types, as it primarily focuses on delivering free access to State-of-the-Art AI methodologies which have been already validated in the corresponding references.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment Statistical Significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [No]

Justification: There are no reported error bars, confidence intervals, or statistical significance tests related to experimental results.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments Compute Resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: Smart Material AI platform runs on the computer resources specified in section 2.6.

   Guidelines:

   - The answer NA means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code Of Ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: The research aligns with the NeurIPS Code of Ethics, adhering to all guidelines related to transparency, privacy, data integrity, fairness, and social responsibility.

   Guidelines:

   - The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader Impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [Yes]

    Justification: The paper does dedicate section to discussing broader impacts.

    Guidelines:

    - The answer NA means that there is no societal impact of the work performed.

- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

    Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

    Answer: [NA]

    Justification: The paper poses no such risks.

    Guidelines:

    - The answer NA means that the paper poses no such risks.
    - Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
    - Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
    - We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

    Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

    Answer: [Yes]

    Justification: Original paper that produced each of the code packages or datasets used for this AI platform are properly cited in this work.

    Guidelines:

    - The answer NA means that the paper does not use existing assets.
    - The authors should cite the original paper that produced the code package or dataset.
    - The authors should state which version of the asset is used and, if possible, include a URL.
    - The name of the license (e.g., CC-BY 4.0) should be included for each asset.
    - For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.

- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New Assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [NA]

    Justification: The paper does not release new assets.

    Guidelines:

    - The answer NA means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and Research with Human Subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional Review Board (IRB) Approvals or Equivalent for Research with Human Subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [NA]

    Justification: The paper does not involve crowdsourcing nor research with human subjects.

    Guidelines:

    - The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
    - Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.

- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.