Generative AI in Biotechnology: Innovations and Applications

Dr. Ranjith Gopalan[1], Mr. Gopikrishna Kalpinagarajarao[2], Mrs. Rathi Ramakrishnan[3], Mr. Arjun Bharadwaj Thyagarajan[4]

[1]PhD, Principal Consultant, Cognizant
[2]MS Product Engineer, Cardinal Health
[3]MPhil, MSc Lecturer
[4]MCA, Principal Consultant, Cognizant

Abstract
This paper talks about the importance of biotechnology in contemporary research cannot be overstated, particularly as we witness rapid advancements in generative AI technologies. The intersection of artificial intelligence (AI) and biotechnology represents a transformative frontier in the life sciences, where the integration of computational power and biological data is paving the way for innovative solutions. Researchers are increasingly leveraging generative AI to enhance various aspects of biotechnology, from drug discovery to the engineering of complex biological systems. This paper walks through potential applications of Generative AI in biotechnology, advances in Protein Structure Prediction, AI driven bioprocess Design, Data sources and modeling techniques, Ethical Implications of AI in Biotechnology, Global security concerns, future trends, innovations, and Recommendation on better usage of building generative AI models for research,

Introduction

Generative AI has surfaced as a pivotal technology in biotechnology, presenting innovative solutions with the potential to revolutionize the field. It is being developed to engineer biological products with specific traits and to function as research assistants through conversational interfaces.
Language-only models face inherent limitations in representation, prediction, and causal inference, even when trained on domain-specific data. Specialized generative models for biotechnology applications, such as protein and metabolic pathway design, may lack the ability to provide exposed reasoning and support conversational queries. This review delves into the present and prospective applications of generative AI in biotechnology, emphasizing its considerable impact on pharmaceutical research and development and its capacity to shape the future. Generative AI embodies a groundbreaking method in biotechnology, utilizing sophisticated algorithms to forge unprecedented solutions. Central to generative AI are machine learning models, especially generative adversarial networks (GANs) and variational autoencoders (VAEs), which generate new data from existing datasets. These models have shown extraordinary proficiency in drug discovery, protein structure prediction, and personalized medicine, propelling substantial progress in the biotechnology sector.

Methodology

The paper discusses methodologies and processes in the potential applications of Generative AI in biotechnology, including advances in protein structure prediction, AI-driven bioprocess design, data sources and modeling techniques, and the ethical implications of AI in biotechnology.

What is biotechnology?

Biotechnology involves utilizing biological systems to achieve specific engineering goals. This can include the use of existing biological systems, their modification, or the creation of new biological systems to accomplish a task. Historically, biotechnology has been used in the production of bread, cheese, yogurt, beer, and wine through the natural fermentation of food by microorganisms such as yeast, and in herbal medicines like willow bark, which contains a precursor to aspirin. Subsequent developments included the advent of vaccination and the discovery of antibiotics from bread mold to combat harmful bacteria. Modern biotechnology encompasses the intentional design of drug-like molecules and vaccines, the employment of microorganisms to generate valuable substances for food, fuel, and industrial raw materials, and the modification of microorganisms to break down oil spills and other detrimental agents like environmental toxins or biological weapons. Ambitious current projects in biotechnology involve the development of replacement human tissues or organs, the cultivation of lab-grown meat that is both ethical and healthy, the application of bioproduction in manufacturing and construction, among other diverse applications.

Proteins can be tailored to function as enzymes that facilitate chemical reactions or as materials for structural or functional purposes. Metabolic pathways can be constructed through a method called retro-biosynthesis to biologically produce a target chemical compound in a series of steps from cost-effective starting materials, all within engineered microorganisms, as opposed to traditional chemical synthesis. Furthermore, cells and tissues can be designed to exhibit specific characteristics, as well as complex phenotypes that necessitate numerous cellular and biochemical processes to manifest.

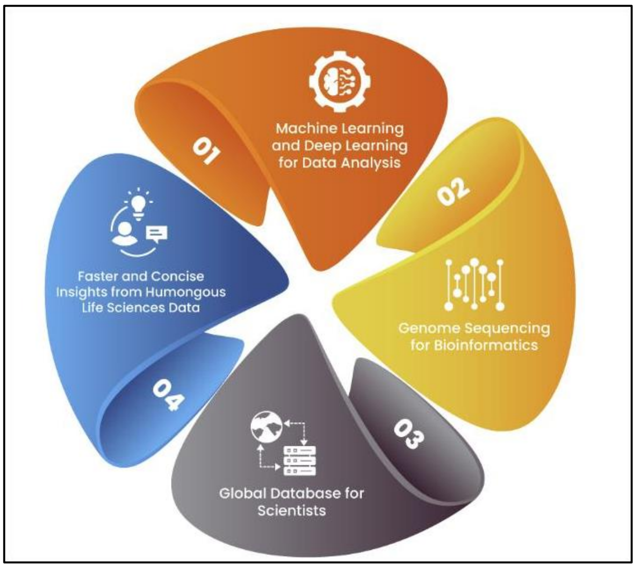

Fig1 shows Subset of AI Process in Biotechnology

Potential applications of Generative AI in biotechnology

Paper discusses about potential applications of Generative AI models in the progress of biotechnology process.

A. Drug Discovery and Development

Drug discovery and development is a complex and multifaceted process that involves the identification of potential therapeutic agents, followed by extensive testing and optimization to ensure safety and efficacy. In recent years, the integration of generative artificial intelligence (AI) has begun to transform this landscape. By leveraging machine learning algorithms, researchers can now analyze vast datasets, predict molecular interactions, and design novel compounds with enhanced properties. This constructive collaboration not only accelerates the drug discovery timeline but

also reduces costs associated with traditional methodologies.
One key application of generative AI in drug discovery is the generation of novel molecular structures.
Machine learning (ML) algorithms have significantly transformed the drug discovery,
development, and testing processes, making them more efficient and cost-effective. Here are some keyways ML can enhance these stages:

1. Drug Discovery

Target Identification and Validation: ML models can analyze vast biological data to identify potential drug targets and validate their relevance. Tools like DTI-CNN,
DeepCPI and DeepDTA are used for predicting drug-target interactions

DTI-CNN

The convolutional neural network (CNN) is a class of NN, commonly used to analyze visual imagery. DTI-CN is a simple DL-based drug–target interaction
prediction tool that is said to outperform the existing state-of-the-art methods by the intelligent interaction of three components namely, (1) heterogeneous-
network-based feature extractor, (2) denoising-auto encoder-based feature
selector, (3) CNN based interaction predictor As the model is based on random walk with restart (RWR) and denoising auto encoder (DAE) model, it is capable of coping up with low-dimensional feature vectors and noisy incomplete and high-dimensional features from heterogeneous data sources, including drug,
protein, side-effects and diseases information. The general workflow of DTI- CNN-based DTI prediction is shown below.

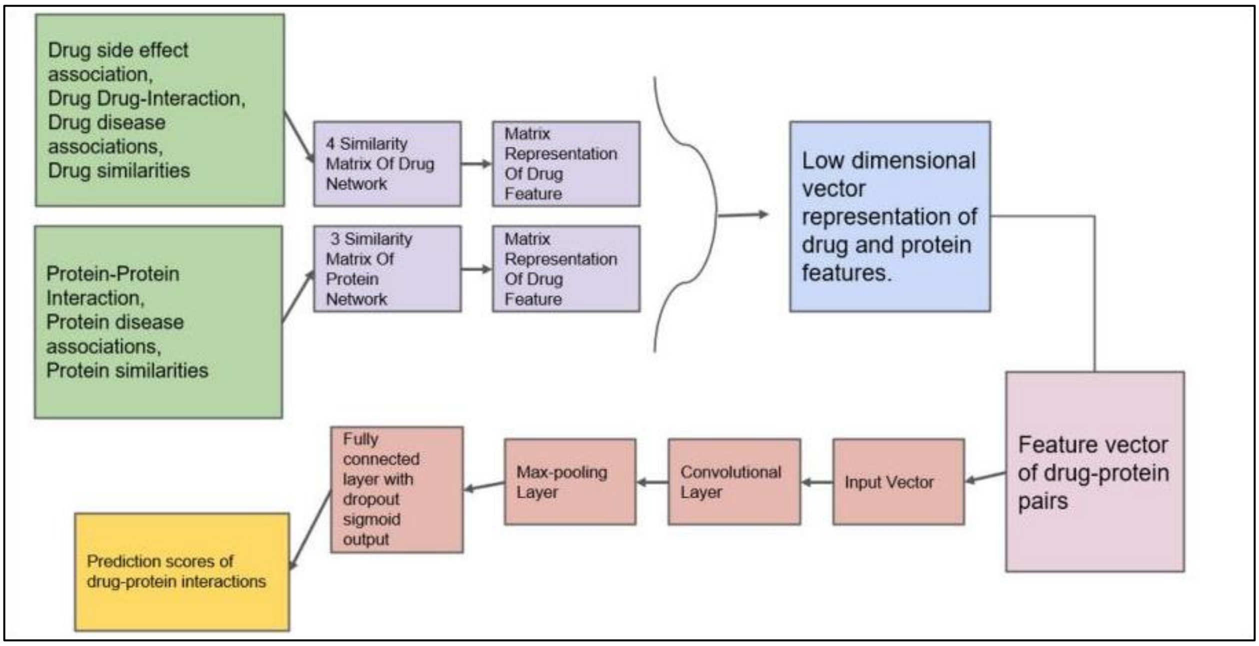

Fig 2: Flowchart for prediction of drug–protein interaction using DTI-CNN

Deep CPI

DeepCPI is a computational framework that accurately predicts chemical-protein interactions using deep learning. It integrates relevant data sources and deploys a Multiview deep neural network architecture to capture the underlying
relationships between chemical compounds and protein sequences. Schematic of the DeepCPI workflow First, motivated by the current NLP techniques, the
unsupervised representation learning strategies (including latent semantic analysis and Word2vec) are used to obtain low-dimensional representations of compound and protein features from massive unlabeled data.

Subsequently, these extracted low-dimensional feature representations of compounds and proteins are fed to a multimodal DNN to make the prediction. NLP, natural language processing; DNN, deep neural network.

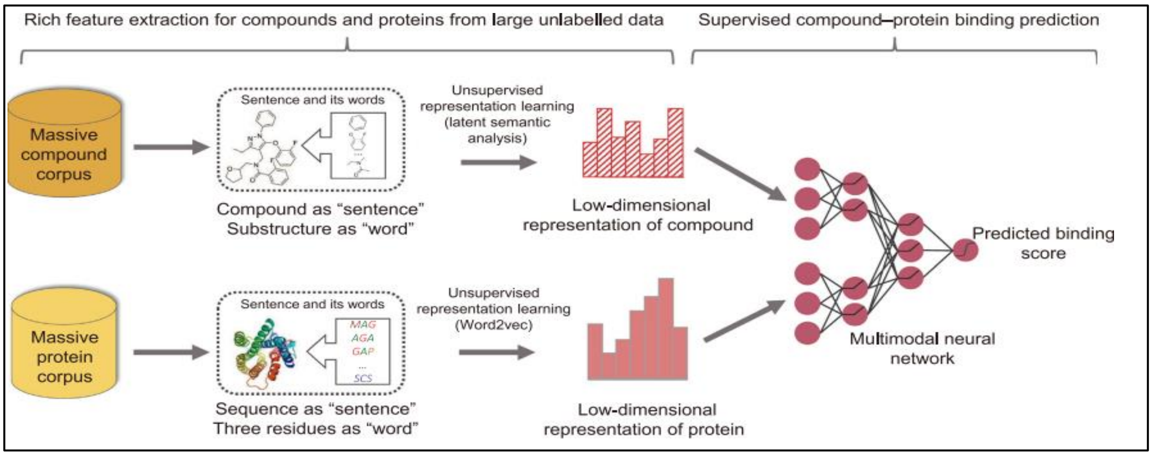

Fig3 shows Schematic of the DeepCPI workflow.

Hit Identification: Virtual screening of large chemical libraries using ML can quickly identify potential hit compounds. This reduces the time and cost compared to traditional high-throughput screening.

De Novo Drug Design: Generative models and reinforcement learning can design new drug molecules with desired properties, speeding up the discovery of novel therapeutics.

2. Drug Development

- **Optimization of Lead Compounds**: ML algorithms can predict the properties of lead compounds and suggest modifications to improve efficacy and reduce toxicity. This iterative process is faster and more precise than traditional methods.
- **ADMET Prediction:** ML models can predict the absorption, distribution, metabolism, excretion, and toxicity (ADMET) profiles of compounds, helping to identify potential issues early in the development process.

3. Clinical Trials and Testing

- **Patient Stratification:** ML can analyze patient data to identify subgroups that are more likely to respond to a treatment, leading to more targeted and effective clinical trials2.
- **Predictive Modeling:** ML models can predict clinical trial outcomes, helping to design better trials and reduce the risk of failure.
- **Real-Time Monitoring:** During trials, ML can be used to monitor patient data in real-time, identifying adverse effects or efficacy signals more quickly1.

B. Genetic Engineering

Generative AI models are also transforming the field of genetic engineering, enabling the design and optimization of novel genetic constructs. Genetic engineering has emerged as a transformative force within biotechnology, enabling researchers to manipulate the genetic material of organisms with unprecedented precision. This capability has been significantly enhanced by the advent of generative AI, which offers innovative approaches to designing and optimizing genetic modifications. By leveraging machine learning algorithms, scientists can predict the effects of specific genetic alterations, facilitating the development of tailored organisms for various applications. This intersection of genetic engineering and generative AI fosters a new era of research, where hypotheses can be evaluated virtually, significantly accelerating the pace of discovery.

Genetic engineering and generative AI are a powerful combination, driving significant advancements in biotechnology. Here's how generative AI is enhancing genetic engineering:

1. **Enhanced Precision and Efficiency**

Generative AI models, such as protein language models, can analyze vast amounts of biological data to design and optimize genetic modifications with high precision. These models can predict the effects of genetic changes, reducing the trial-and-error process

traditionally involved in genetic engineering1.

2. **Innovative Gene Editing Tools**

Generative AI has been used to create new gene-editing tools that are more versatile and efficient than naturally occurring systems. in AI-designed CRISPR systems can target a broader range of genetic sequences and perform more precise edits. This expands the potential applications of gene editing in medicine, agriculture, and other fields.

3. **Accelerated Drug Discovery**

In drug discovery, generative AI can design novel compounds and predict their interactions with biological targets. This accelerates the identification of potential drug candidates and reduces the time and cost associated with bringing new drugs to market2.

4. **Personalized Medicine**

Generative AI can analyze individual genomic data to design personalized gene therapies.

By predicting how genetic modifications will affect an individual's health, AI can help develop treatments tailored to specific genetic profiles, improving efficacy and safety3.

5. **Synthetic Biology**

AI is also playing a crucial role in synthetic biology, where it helps design synthetic organisms or biological systems. For instance, AI can engineer bacteria to produce useful compounds, such as biofuels or biodegradable plastics, by optimizing their genetic pathways4.

Advances in Protein Structure Prediction

Protein language models (PLMs) are specialized deep learning models designed to understand and generate protein sequences. These models are inspired by natural language processing (NLP) techniques and have shown great promise in various applications, from protein design to

predicting protein-protein interactions. Here are some notable protein language models:

A. MSA based structure prediction:

Multiple Sequence Alignment (MSA) based structure prediction models leverage the evolutionary information contained in MSAs to predict protein structures with high accuracy.

Although deep learning methods that use MSAs have clearly been successful in predicting protein structures, the need for high-quality MSAs can be problematic. Although far fewer than there were, there still exist so-called orphan sequences or "lineage-specific genes" with no or very few homologs in current sequence databases. Such genes have been theorized to have originated from several sources, including evolution from non-coding sequences in specific species, or through gene neofunctionalization, whereby gene duplication events give rise to novel functions. MSA-based methods typically do not produce good predictions for these sequences.

More broadly, there is also the argument that one should not need to resort to using a family of related sequences to infer the structure of one member of that family;, a protein chain folding in vivo has no knowledge of its evolutionary history. From a computational perspective, single sequence methods also offer advantages. MSA-based methods rely on time-consuming database searching, and the quality of their predictions depends on the availability and completeness of homologous sequence.

B. Single sequence-based structure prediction methods:

The development of single-sequence-based protein structure prediction methods has been a major goal in the field for some time. These methods aim to predict protein structures from a single sequence, without needing to first identify homologs and construct multiple sequence alignments. Most recently, transformers and other types of pre-trained language models trained on large protein sequence datasets have shown significant promise in this direction.

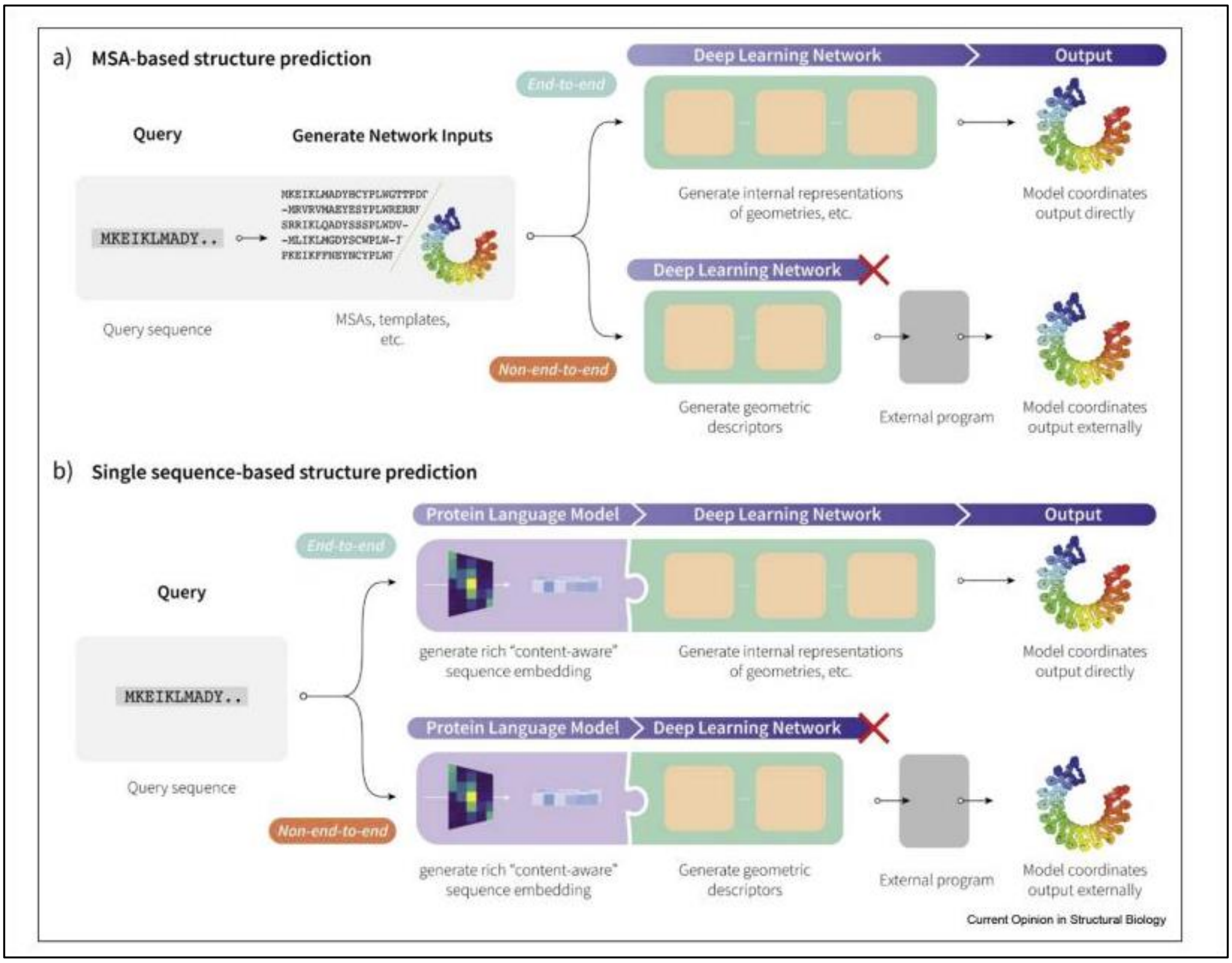

Fig 4 shows MSA based structure and Single sequence-based structure.

C. Other protein prediction models like

1. ESM Model, this is created by Meta, ESM models are among the most comprehensive AI models used in biological research. They have demonstrated their effectiveness in synthesizing novel fluorescent proteins and enhancing gene-editing technologies. The main uses of these models include predicting protein structures, designing new proteins, and providing functional annotations.

2. ProGen, Salesforce Research's ProGen is a generative model trained on an extensive collection of protein sequences. It excels in generating unique protein sequences designed to exhibit specific properties. Its applications are diverse, ranging from the novel design of proteins and enzyme engineering to the synthesis of proteins for therapeutic purposes.

3. ProLLaMA, this is a recent language model developed by Anthropic that is like GPT-4 but focused on proteins and highlights advances in large language models for protein applications like prediction, generation, and classification. The ProLLaMA model represents a significant advancement in protein language processing, demonstrating proficiency in generating protein sequences unconditionally or according to specific parameters, as well as in predicting protein attributes. Its applications are diverse, encompassing the design of proteins with tailored functions, the prediction of protein properties, and the generation of protein sequences.

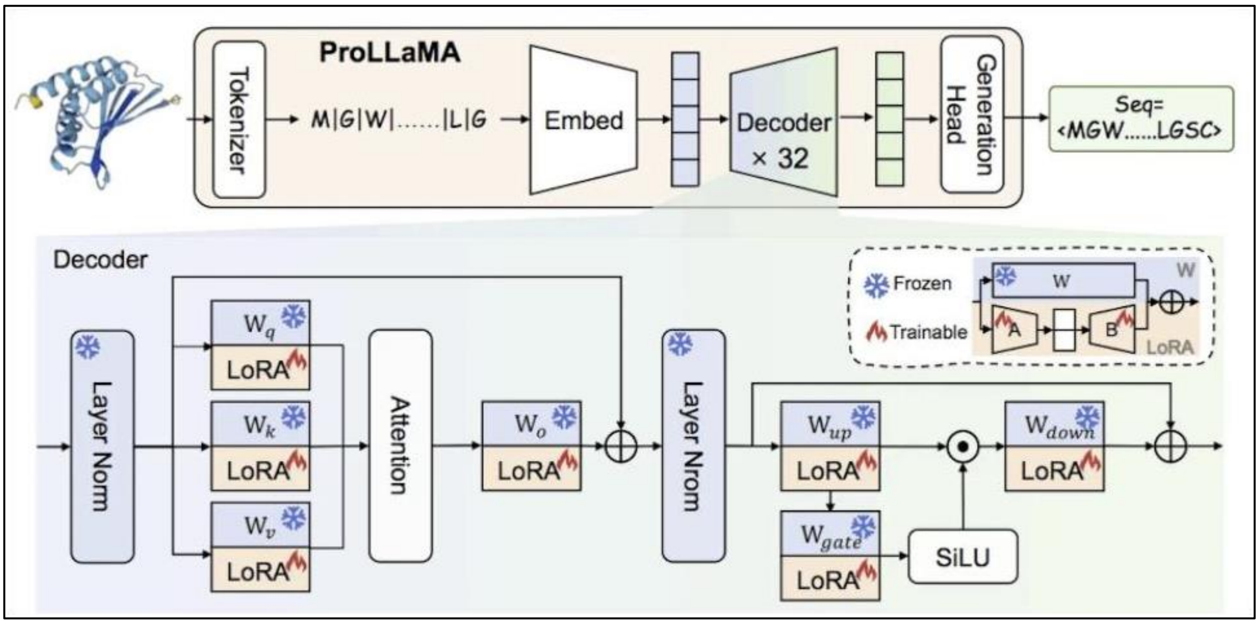

Fig 5 shows process diagram of ProLLaMA , protein structure prediction model

AI-Driven Bioprocess Design

AI-Driven Bioprocess Design marks a revolutionary shift in biotechnology, utilizing artificial intelligence to enhance and refine bioprocesses. Where traditional bioprocess design has depended on empirical methods and expert insight, often constrained by the intricate complexity of biological systems and time-intensive practices, AI integration represents a change in basic assumptions. It empowers researchers to exploit extensive datasets for predicting outcomes, pinpointing optimal conditions, and refining processes, which translates to heightened efficiency and cost reductions in biomanufacturing.

A prime example of AI's impact in bioprocess design is evident in protein engineering.

Sophisticated machine learning algorithms scrutinize protein structures to forecast the effects of alterations on their functionality and stability. This foresight expedites the creation of new biocatalysts and therapeutic proteins, facilitating swift iterations and hypothesis testing. As a result, researchers can engineer proteins to meet precise performance standards, propelling industrial and therapeutic innovations in biotechnology.

Furthermore, in synthetic biology and metabolic engineering, AI-driven design expedites the assembly of genetic circuits and pathways. AI's predictive modeling of genetic interplays forecasts the dynamics of intricate biological systems prior to laboratory implementation. Such predictive prowess allows for the design of more effective biosynthetic routes for high-value products like pharmaceuticals and biofuels, reducing undesirable byproducts and boosting yields. Preemptive pathway optimization can drastically curtail the experimental validation phase, saving time and resources.

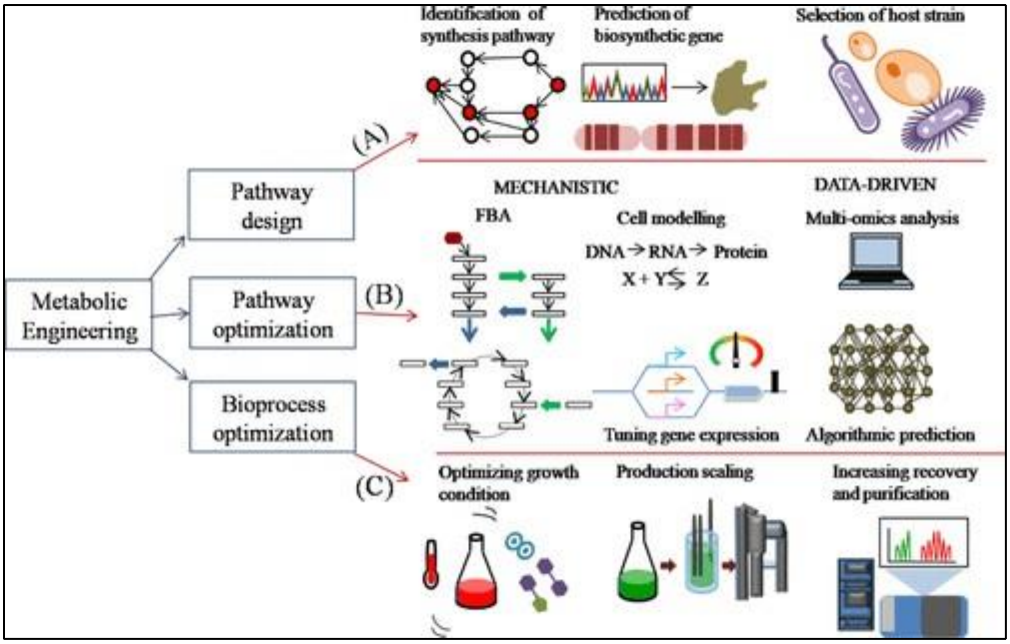

Fig 6: AI-Driven Bioprocess Design using machine learning and deep learning.

Results and Discussion

Paper expresses by adopting generative AI in biotechnology, data sources and effectiveness of generative AI model are important and highlighting some of the importance of data sources and modeling techniques. Model and its data sources are part of result analysis in every generative AI model implementation.

Data sources and modeling techniques

The efficacy of generative AI in biotechnology hinges on several pivotal factors. Firstly, the presence of extensive protein sequence datasets has facilitated the development of robust language models capable of discerning protein sequence patterns and semantics, leading to precise predictions and generative functions.

Secondly, the evolution of deep learning, especially transformer architectures, has been vital. Models such as ESM, ProGen, and ProLLaMA have shown superior performance in protein structure prediction, novel protein design, and functional annotation, with their success due to their proficiency in capturing complex dependencies within protein sequences.

In synthetic biology and metabolic engineering, modeling transcends protein design, covering an organism's entire metabolic network. Utilizing constraint-based optimization and dynamic modeling, researchers can emulate metabolic fluxes and foresee genetic modification outcomes, optimizing biomanufacturing by pinpointing bottlenecks and designing strains for increased production.

Generative AI also propels medical imaging and diagnostics forward by refining image data analysis. Training convolutional neural networks on vast medical image datasets enhances disease diagnosis and prognosis accuracy. Additionally, generative models can create high- quality images to supplement datasets, tackling issues related to data scarcity.

Finally, the integration of generative AI in bioinformatics and data analysis has the potential to revolutionize the way researchers approach environmental biotechnology and bioremediation. The use of generative models to predict microbial interactions and environmental responses can lead to the identification of novel bioremediation strategies. Moreover, the ability to model complex ecological systems allows for a more nuanced understanding of the interactions between engineered organisms and their environments. As researchers continue to explore these innovative applications, the constructive collaboration between data sources and modeling techniques will undoubtedly drive the future of biotechnology, fostering discoveries that were previously unimaginable.

Discussion

Paper discusses about ethical implications of AI in biotechnology and regulatory challenges. These aspects needed to consider in detail as this gives accuracy and great extendibility in the growth of biotechnology.

Responsible implementation of generative AI in biotechnology requires carefully navigating ethical and regulatory considerations.

Ethical Implications of AI in Biotechnology

The fusion of artificial intelligence with biotechnology introduces numerous ethical challenges that must be carefully considered by researchers and scientists. As generative AI technologies evolve, the sophistication with which biological systems can be manipulated increases,

prompting vital questions about scientists' moral obligations to use these tools safely and ethically. Issues of consent, accountability, and the risk of misuse are critical as researchers explore applications capable of altering genetic structures, engineering proteins, or synthesizing new organisms.

A significant ethical issue is the risk of unforeseen outcomes from genetic engineering. The capacity to tailor proteins or generate synthetic life forms brings into question the effects on ecological equilibrium and biodiversity over time. The prospect of such organisms outperforming natural counterparts or unsettling ecosystems requires a rigorous ethical approach that assesses not just the immediate advantages of biotechnological advancements but also their wider environmental consequences and the dangers they pose.

Moreover, employing AI for predictive modeling in disease forecasting and diagnostics introduces ethical quandaries concerning the privacy and security of data. Collecting and processing sensitive health information is vital for advancing medical imaging and diagnostic

tools. Nonetheless, the possibility of misusing personal health data or creating biased algorithms presents considerable hazards. Researchers must ensure their AI models are transparent and equitable to prevent exacerbating health inequalities or infringing on privacy. The development of strong ethical standards and regulatory measures is crucial to maintain public trust and uphold accountability in these technological pursuits.

Regulatory Challenges

The integration of generative AI into biotechnology presents complex regulatory challenges, stemming from the convergence of swiftly evolving technology and established legal frameworks. As scientists delve into the transformative potential of generative AI for protein engineering, synthetic biology, and biomanufacturing, they encounter intricate regulatory terrains that oversee biotechnological advancements. This becomes particularly intricate as generative AI has the capability to create new biological entities and processes, prompting considerations regarding intellectual property, safety evaluations, and adherence to current biotech regulations.

A key regulatory issue pertains to the categorization of biological products derived from AI.

Traditional regulatory systems typically classify products based on origin—natural or synthetic. However, generative AI disrupts these distinctions by generating molecular structures that may not align with predefined categories. For example, AI-driven protein engineering may yield proteins that are neither fully natural nor synthetic, thus complicating their regulatory sanctioning. This situation calls for a reassessment of regulatory definitions and classifications of biotech innovations to facilitate effective safety and efficacy evaluations without impeding progress.

Moreover, the risk of unintended effects from AI-crafted biotechnological applications poses substantial regulatory challenges. In fields like synthetic biology and metabolic engineering, genetic alterations could have unpredictable ecological repercussions. Regulators are tasked with evaluating the hazards linked to these new organisms. While generative AI can enhance metabolic pathways in microbes, the environmental consequences of introducing such modified organisms remain a concern.

Future Trends and Innovations

The integration of generative artificial intelligence promises transformative advancements in biotechnology. Researchers are recognizing the potential applications of generative AI across diverse fields, leading to innovative solutions that enhance research efficiency, accuracy, and speed. As generative AI evolves, its ability to analyze large datasets and generate predictive models will facilitate breakthroughs in areas like protein engineering and design.

In protein engineering, generative AI is set to revolutionize the design process by enabling the rapid generation of novel protein structures with desired functionalities. By leveraging algorithms that learn from existing protein databases, researchers can explore an expansive design space that would be infeasible through traditional methods. This innovation accelerates the discovery of new proteins and enhances the optimization of existing ones for applications such as drug discovery and enzyme development. The ability of generative AI to predict protein folding and interactions can significantly reduce experimental trial times, expediting the pathway from conceptualization to application.

Synthetic biology and metabolic engineering also stand to benefit from advancements in generative AI. The ability to model complex biological systems and predict the outcomes of genetic modifications can lead to more effective and sustainable manufacturing processes.

Generative AI can assist in designing synthetic biological circuits that optimize metabolic pathways for biofuel production or valuable biochemical synthesis. By simulating genetic interventions, researchers can identify the most promising strategies to enhance yield and efficiency, paving the way for environmentally sustainable bioproduction.

In medical imaging and diagnostics, generative AI presents opportunities for improved accuracy and efficiency. Deep learning techniques can analyze imaging data to detect disease-indicative anomalies at earlier stages than traditional methods. Furthermore, predictive modeling for disease outbreaks can leverage generative AI to analyze epidemiological data and anticipate potential health crises, enabling initiative-taking measures and resource allocation. This capability is particularly relevant as global health challenges evolve, necessitating agile responses informed by data-driven insights.

The integration of generative AI into biomanufacturing and process optimization is expected to streamline operations, reduce costs, and enhance product quality. Automated laboratory research and experimentation will also benefit from AI-driven automation, which can optimize experimental design and execution. As bioinformatics and data analysis become increasingly data-intensive, generative AI will enable researchers to harness complex datasets, providing deeper insights into biological processes and improving decision-making capabilities.

Conclusion

This paper summarizes key findings, and explores opportunities for emerging generative AI tools in biotechnology applications, including predictive design of useful biological systems that scale from molecules to organisms, Trustworthy and explainable results of AI outputs, including exposed chains of reasoning and references to scientific source literature or other forms of evidence, The paper explores the myriad applications of Generative AI in biotechnology, delving into the latest advancements in Protein Structure Prediction and AI-driven bioprocess design. It examines various data sources and modeling techniques, addresses the ethical implications of AI in biotechnology, and discusses global security concerns. The paper also looks ahead to future trends and innovations, offering recommendations for the enhanced utilization of generative AI models in research, drug discovery, and biomanufacturing.

This paper explains detailed process of for prediction of drug–protein interaction, Schematic of the DeepCPI workflow, Overview of deep learning protein structure prediction methods and AI- Driven Bioprocess Design using machine learning and deep learning. It discusses the opportunities, challenges and key regulatory considerations for widespread adoption of generative AI in biotechnology.

This paper also discusses a key challenge for the integration of generative AI in biotechnology, namely the need for a flexible and responsive regulatory framework. As the boundaries between natural and artificial processes become increasingly blurred, traditional regulatory systems must evolve to accommodate the complexities and nuances of AI-driven biotechnological innovations.

Sure, that this research paper will help students and professionals in the field of biotechnology to understand the current trends and future potential of generative AI in this domain.

Reference

1. (HEG), H E & G. (2021, June 27). Ethics and governance of artificial intelligence for health. https://www.who.int/publications/i/item/9789240029200
2. Averesch, N., & Krömer, J O. (2018, March 26). Metabolic Engineering of the Shikimate Pathway for Production of Aromatics and Derived Compounds—Present and Future Strain Construction Strategies. Frontiers Media, 6. https://doi.org/10.3389/fbioe.2018.00032
3. Banerjee, D., Eng, T., Lau, A K., Wang, B., Sasaki, Y., Herbert, R A., Chen, Y., Liu, Y., Prahl, J., Singan, V., Tanjore, D., Petzold, C J., Keasling, J D., & Mukhopadhyay, A. (2020, February 24). Genome-scale metabolic rewiring to achieve predictable titers rates and yield of a non-native product at scale. Cold Spring Harbor Laboratory. https://doi.org/10.1101/2020.02.21.954792
4. Beardall, W A V., Stan, G., & Dunlop, M J. (2022, August 1). Deep Learning Concepts and Applications for Synthetic Biology. Mary Ann Liebert, Inc., 1(4), 360-371. https://doi.org/10.1089/genbio.2022.0017
5. Brandes, N., Goldman, G., Wang, C H., Ye, C., & Ntranos, V. (2023, August 10). Genome-wide prediction of disease variant effects with a deep protein language model. Nature Portfolio, 55(9), 1512-1522. https://doi.org/10.1038/s41588-023-01465-0
6. Challen, R., Denny, J C., Pitt, M., Gompels, L., Edwards, T L., & Tsaneva-Atanasova, K. (2019, January 12). Artificial intelligence, bias, and clinical safety. BMJ, 28(3), 231-237. https://doi.org/10.1136/bmjqs-2018-008370
7. Crigger, E., Reinbold, K., Hanson, C., Kao, A., Blake, K., & Irons, M. (2022, January 12). Trustworthy Augmented Intelligence in Health Care. Springer Science+Business Media, 46(2). https://doi.org/10.1007/s10916-021-01790-z
8. Ding, Y., Wu, F., & Tan, C. (2014, December 19). Synthetic Biology: A Bridge between Artificial and Natural Cells. Multidisciplinary Digital Publishing Institute, 4(4), 1092- 1116. https://doi.org/10.3390/life4041092
9. Dorr, D A., Adams, L., & Embí, P J. (2023, March 27). Harnessing the Promise of Artificial Intelligence Responsibly. American Medical Association, 329(16), 1347-1347. https://doi.org/10.1001/jama.2023.2771
10. Gelman, S., Fahlberg, S A., Heinzelman, P., Romero, P A., & Gitter, A. (2021, November 23). Neural networks to learn protein sequence–function relationships from deep mutational scanning data. National Academy of Sciences, 118(48). https://doi.org/10.1073/pnas.2104878118
11. Generating 'smarter' biotechnology. (2023, February 2). Nature Portfolio. https://doi.org/10.1038/s41587-023-01695-x
12. Gibson, K M D R U J W M C M L L. (2020, October 4). Artificial intelligence for good health: a scoping review of the ethics literature. https://bmcmedethics.biomedcentral.com/track/pdf/10.1186/s12910-021-00577-8
13. Gomes, C P C., Cho, J., Hood, L., Franco, O L., Pereira, R W., & Wang, K. (2013, January 1). A Review of Computational Tools in microRNA Discovery. Frontiers Media, 4.

https://doi.org/10.3389/fgene.2013.00081

14. Graham, S., Depp, C A., Lee, E., Nebeker, C., Tu, X., Kim, H., & Jeste, D V. (2019, November 1). Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. Springer Science+Business Media, 21(11). https://doi.org/10.1007/s11920- 019-1094-0
15. Harrer, S., Shah, P., Antony, B., & Hu, J. (2019, July 17). Artificial Intelligence for Clinical Trial Design. Elsevier BV, 40(8), 577-591. https://doi.org/10.1016/j.tips.2019.05.005
16. Hasibuan, Z A. (2020, October 17). Towards Using Universal Big Data in Artificial Intelligence Research and Development to Gain Meaningful Insights and Automation Systems. https://doi.org/10.1109/iwbis50925.2020.9255497
17. Heßler, G., & Baringhaus, K. (2018, October 2). Artificial Intelligence in Drug Design. Multidisciplinary Digital Publishing Institute, 23(10), 2520-2520. https://doi.org/10.3390/molecules23102520
18. Hilton, C B., Milinovich, A., Felix, C., Vakharia, N., Crone, T J., Donovan, C., Proctor, A., & Nazha, A. (2020, April 3). Personalized predictions of patient outcomes during and after hospitalization using artificial intelligence. Nature Portfolio, 3(1). https://doi.org/10.1038/s41746-020-0249-z
19. Ho, D., Quake, S R., McCabe, E R., Chng, W J., Chow, E K., Ding, X., Gelb, B D., Ginsburg, G S., Hassenstab, J., Ho, C., Mobley, W C., Nolan, G P., Rosen, S T., Tan, P., Yen, Y., & Zarrinpar, A. (2020, January 21). Enabling Technologies for Personalized and Precision Medicine. Elsevier BV, 38(5), 497-518. https://doi.org/10.1016/j.tibtech.2019.12.021
20. Hodgman, C E., & Jewett, M C. (2011, September 18). Cell-free synthetic biology: Thinking outside the cell. Elsevier BV, 14(3), 261-269. https://doi.org/10.1016/j.ymben.2011.09.002
21. Jiang, Y., Oron, T., Clark, W T., Bankapur, A., D'Andrea, D., Lepore, R., Funk, C S., Kahanda, I., Verspoor, K., Ben-Hur, A., Koo, D C E., Penfold-Brown, D., Shasha, D., Youngs, N., Bonneau, R., Lin, A J., Sahraeian, S M E., Martelli, P L., Profiti, G., . . . Giollo, M. (2016, September 7). An expanded evaluation of protein function prediction methods shows an improvement in accuracy. BioMed Central, 17(1). https://doi.org/10.1186/s13059-016-1037-6
22. Kamm, R D., Bashir, R., Arora, N., Dar, R D., Gillette, M U., Griffith, L G., Kemp, M L., Kinlaw, K., Levin, M., Martin, A C., McDevitt, T C., Nerem, R M., Powers, M J., Saif, T., Sharpe, J., Takayama, S., Takeuchi, S., Weiss, R., Ye, K., . . . Zaman, M H. (2018, October 11). Perspective: The promise of multi-cellular engineered living systems. AIP Publishing, 2(4). https://doi.org/10.1063/1.5038337
23. Kang, A., Mendez-Perez, D., Goh, E., Baidoo, E E K., Benites, V T., Beller, H R., Keasling, J D., Adams, P D., Mukhopadhyay, A., & Lee, T S. (2019, September 6). Optimization of the IPP-bypass mevalonate pathway and fed-batch fermentation to produce isoprenol in Escherichia coli. Elsevier BV, 56, 85-96. https://doi.org/10.1016/j.ymben.2019.09.003
24. Kolluri, S., Lin, J., Liu, R., Zhang, Y., & Zhang, W. (2022, January 1). Machine Learning and Artificial Intelligence in Pharmaceutical Research and Development: a Review. Springer Science+Business Media, 24(1). https://doi.org/10.1208/s12248-021-00644-3
25. Kumar, A., Sun, Y., Ng, C Y., & Maranas, C D. (2018, January 8). Pathway design using de novo steps through uncharted biochemical spaces. Nature Portfolio, 9(1). https://doi.org/10.1038/s41467-017-02362-x

26. Lekadir, K., Feragen, A., Fofanah, A J., Frangi, A F., Buyx, A., Emelie, A., Lara, A., Porras, A R., Chan, A., Navarro, A., Glocker, B., Botwe, B O., Khanal, B., Beger, B., Wu, C C., Cintas, C., Langlotz, C P., Rueckert, D., Mzurikwao, D., . . . Colantonio, S. (2023, January 1). FUTURE-AI: International consensus guideline for trustworthy and deployable artificial intelligence in healthcare. Cornell University. https://doi.org/10.48550/arXiv.2309.
27. Lin, Z., Akin, H., Rao, R., Hie, B., Zhu, Z., Lu, W., Smetanin, N., Verkuil, R., Kabeli, O., Shmueli, Y., Costa, A D S., Fazel-Zarandi, M., Sercu, T., Candido, S., & Rives, A. (2023, March 16). Evolutionary-scale prediction of atomic-level protein structure with a language model. American Association for the Advancement of Science, 379(6637), 1123-1130. https://doi.org/10.1126/science.ade2574
28. Liu, X. (2017, January 1). Deep Recurrent Neural Network for Protein Function Prediction from Sequence. Cornell University. https://doi.org/10.48550/arXiv.1701.
29. Lu, Y. (2017, March 1). Cell-free synthetic biology: Engineering in an open world. Elsevier BV, 2(1), 23-27. https://doi.org/10.1016/j.synbio.2017.02.003
30. Markus, A F., Kors, J A., & Rijnbeek, P R. (2020, December 10). The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies. Elsevier BV, 113, 103655-103655. https://doi.org/10.1016/j.jbi.2020.103655
31. McCradden, M D., Baba, A., Saha, A., Ahmad, S., Boparai, K., Fadaiefard, P., & Cusimano, M D. (2020, January 1). Ethical concerns around use of artificial intelligence in health care research from the perspective of patients with meningioma, caregivers, and health care providers: a qualitative study. Canadian Medical Association, 8(1), E90-E95. https://doi.org/10.9778/cmajo.20190151
32. Neupane, S., Fernández, I., Mittal, S., & Rahimi, S. (2023, January 1). Impacts and Risk of Generative AI Technology on Cyber Defense. Cornell University. https://doi.org/10.48550/arXiv.2306.
33. Pasricha, S. (2023, June 30). AI Ethics in Smart Healthcare. https://export.arxiv.org/pdf/2211.06346v1.pdf
34. Perveen, S., Shahbaz, M., Keshavjee, K., & Guergachi, A. (2019, September 24). Prognostic Modeling and Prevention of Diabetes Using Machine Learning Technique. Nature Portfolio, 9(1). https://doi.org/10.1038/s41598-019-49563-6
35. Ruffle, J K., Farmer, A D., & Aziz, Q. (2018, October 12). Artificial Intelligence-Assisted Gastroenterology— Promises and Pitfalls. Lippincott Williams & Wilkins, 114(3), 422- 428. https://doi.org/10.1038/s41395-018-0268-4
36. Selma, S., Ntelkis, N., Nguyen, T H., & Goossens, A. (2023, February 17). Engineering the plant metabolic system by exploiting metabolic regulation. Wiley, 114(5), 1149-1163. https://doi.org/10.1111/tpj.16157
37. Sethi, T., Kalia, A., Sharma, A., & Nagori, A. (2020, January 1). Interpretable artificial intelligence: Closing the adoption gap in healthcare. Elsevier BV, 3-29. https://doi.org/10.1016/b978-0-12-817133-2.00001-x
38. Soy, S., Prabha, R., & Nigam, V K. (2020, February 5). Potential of Biocatalysis in Pharmaceuticals. IntechOpen. https://doi.org/10.5772/intechopen.90459
39. Srivastava, R K., Akhtar, N., Verma, M., & Imandi, S B. (2020, May 21). Primary metabolites from overproducing microbial system using sustainable substrates. Wiley, 67(6), 852-874. https://doi.org/10.1002/bab.1927
40. Szostak, J W. (2017, November 8). The Origin of Life on Earth and the Design of Alternative Life

Forms. World Scientific, 01(02), 121-131. https://doi.org/10.1142/s2529732517400132

41. Trinh, C T., & Mendoza, B. (2016, July 28). Modular cell design for rapid, efficient strain engineering toward industrialization of biology. Elsevier BV, 14, 18-25. https://doi.org/10.1016/j.coche.2016.07.005
42. Verweij, J J., & Stensvold, C R. (2014, April 1). Molecular Testing for Clinical Diagnosis and Epidemiological Investigations of Intestinal Parasitic Infections. American Society for Microbiology, 27(2), 371-418. https://doi.org/10.1128/cmr.00122-13
43. Voigt, C A. (2020, December 11). Synthetic biology 2020–2030: six commercially- available products that are changing our world. Nature Portfolio, 11(1). https://doi.org/10.1038/s41467-020-20122-2
44. Wang, K., Shi, Q., Sun, C., Liu, W., Yau, V., Xu, C., Liu, H., Sun, C., Yin, C., Wei, X., Li, W., & Rong, L. (2023, March 27). A machine learning model for visualization and dynamic clinical prediction of stroke recurrence in acute ischemic stroke patients: A real- world retrospective study. Frontiers Media, 17. https://doi.org/10.3389/fnins.2023.1130831
45. Wang, Q., Feng, Y., Huang, J., Lv, Y., Xie, Z., & Gao, X. (2023, January 1). Large-scale Generative Simulation Artificial Intelligence: the Next Hotspot in Generative AI. Cornell University. https://doi.org/10.48550/arXiv.2308.
46. Xu, C., Hu, S., & Chen, X. (2016, April 5). Artificial cells: from basic science to applications. Elsevier BV, 19(9), 516-532. https://doi.org/10.1016/j.mattod.2016.02.020
47. Zargar, A., Valencia, L E., Wang, J., Lal, R., Chang, S., Werts, M., Wong, A R., Hernández, A C., Benites, V T., Baidoo, E E K., Katz, L., & Keasling, J D. (2020, August 6). A bimodular PKS platform that expands the biological design space. Elsevier BV, 61, 389-396. https://doi.org/10.1016/j.ymben.2020.07.001
48. Zenil, H., Tegnér, J., Abrahão, F S., Lavin, A., Kumar, V., Frey, J G., Weller, A., Soldatova, L., Bundy, A., Jennings, N R., TAKAHASHI, K., Hunter, L., Džeroski, S., Briggs, A., Gregory, F D., Gomes, C P., Williams, C K I., Rowe, J., Evans, J A., . . . King, R D. (2023, January 1). The Future of Fundamental Science Led by Generative Closed- Loop Artificial Intelligence. Cornell University. https://doi.org/10.48550/arxiv.2307.07522