Chairs Farry and Pennycuick, and members of the committees on Communications & Technology and Institutional Sustainability & Innovation: thank you for inviting me.

My name is Jacob Gardner. I am an Assistant Professor of Computer and Information Science at the University of Pennsylvania. I have spent a decade at the intersection of AI and the natural sciences, publishing across both technical AI conferences and life-science journals.

My goal today is to explain the transformative impact AI will have on drug discovery, and what challenges require your attention.

## Background

We are all familiar with the recent rapid advances in Generative AI. We see it in tools that write clear English or generate convincing images.

The same technology that powers ChatGPT has the potential to revolutionize biology.

- The technology used to generate art—Diffusion models—is now used to figure out how proteins fold in 3D space.
- The technology used to write essays—Transformers—is now used to model the sequences of DNA, RNA, and small molecules.

But today's AI is doing something new. It is becoming more generalist.

In the past, if we wanted to predict if a drug would cause liver injury, we needed a large, hand-curated dataset of known "liver-safe" and "liver-dangerous" molecules.

The next generation of models will instead use data from **every source imaginable**. They will "read" scientific literature, datasheets, and clinical reports to learn chemical properties directly from the text.

But it goes further. By using reasoning and digital tools, these models can now actively plan and collaborate with us to reach a goal. We are watching AI evolve from a calculator into a research partner.

## The Promise of AI in Drug Discovery

What difference does this make? It comes down to **Time** and **Money**.

*1. Time.* We are moving towards personalized medicine. Many diseases, like cancer, are highly individualized. Traditional chemotherapy is a blunt instrument. We need "bespoke" therapies that target a specific patient's mutations.

But biology is slow. If a patient is diagnosed with aggressive cancer, we cannot take years to develop a drug. We need a personalized treatment in **weeks**.

AI is the only tool that can scale this. It can rapidly scan billions of biological examples to find the right target and design the right delivery mechanism to save that individual life.

*2. Money.* Drug discovery is risky. Most drugs fail in expensive clinical trials, often due to safety concerns.

If an AI predicts these failures before we spend millions on human trials, the cost of R&D drops. Increasing the success rate lowers the cost of medicine at the pharmacy counter—especially if AI leads to more competing drugs.

## The Pitfalls of AI In Drug Discovery

However, these benefits come with challenges. I want to highlight two specific areas for this committee.

*1. Scientific Data Access* This is the most critical point. To make progress, AI models must be trained. The next generation of models will be trained by reading millions of scientific papers.

Currently, much of this data is locked behind paywalls. My own lab is currently working with Penn Libraries just to understand the licensing terms for the 27 million papers we want to analyze. Restricting access will stall progress.

Recently, the federal government mandated that federally funded papers be open access.

The Commonwealth should consider mirroring this, by mandating that research funded by Pennsylvania be deposited into an existing open-access repository. This costs the state nothing, mirrors federal policy, and ensures PA-funded science is accessible to PA-based innovators.

*2. Privacy* For general drugs-like antibiotics-AI doesn't need patient data. It learns from chemistry, which is impersonal.

However, for personalized medicine, we do need individual genetic data. We need regulations that strictly limit that data usage solely to the production of the treatment, ensuring it is never monetized or used for profiling.

## Conclusion

In conclusion: AI is a tool that can read the sum of biological knowledge to help us cure disease faster. But to work, it needs data access and safety regulations.

Pennsylvania is uniquely positioned to lead this. We are home to major institutions in both biotech and AI. Our region produced the first FDA-approved CAR-T and mRNA therapeutics.

Combined with our pharmaceutical density and regulatory environment, Pennsylvania has the perfect ecosystem to lead the world in this work, should we decide that we want to.

Thank you.