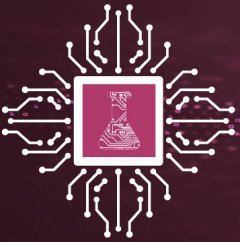


# Industry Insights

## AI IN PHARMA DISCOVERY

October 4th | Boston, MA



**Ryan Emerson**  
VP of Data Science  
A-Alpha Bio

**What have been the most influential AI and data-science based developments in R&D within the past three years?**

I think the advent and popularization of self-supervised deep learning models borrowed from Natural Language Processing is the biggest AI-based development in

R&D recently. Limitations on data volume, data quality and useful labels have all been persistently problematic when applying AI/ML to problems in biology, and models like AntiBERTy and ESM-2 go a long way toward pre-learning useful general features of biological sequences.

These foundation models are allowing for much better performance on a variety of tasks than we've ever seen before, especially on problems for which the amount of task-specific labeled data is much lower than would be required to train an effective model from

scratch. Other areas that have seen important progress recently are protein folding models – having a structure, even an imperfect one, on demand makes a wide variety of research questions more tractable – and the general trend of bringing modern data engineering sensibilities into the lab to capture and utilize as much in vitro data as possible.

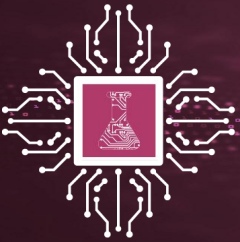
**What traditional bottlenecks in R&D are being alleviated due to these recent developments?**

I'm most interested in the design and engineering of

# Industry Insights

## AI IN PHARMA DISCOVERY

October 4th | Boston, MA



biologics, and AI/ML is having a huge impact in this space. I think that soon, rapid discovery (and eventually design) of antibody variable fragments with arbitrary specificity will be realized, and that will both enable rapid, commoditized discovery of uncomplicated antibody therapeutics, and remove a major bottleneck in the development of exotic antibody-based therapeutics like ADCs, multivalents, Fv-directed cell therapies, etc. I am also very excited by all the effort going into upstream prediction of toxicity, immunogenicity, manufacturability, etc. for biologics – but I think it's

too soon to fully know how well these strategies are working to reduce late-stage failures.

### **How does A-Alpha Bio use AI and ML to accelerate the design and optimisation of new therapeutics?**

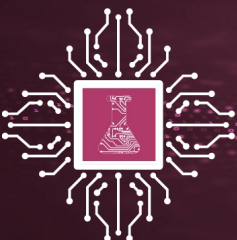
The total search space when optimizing a therapeutic is hopelessly large for in vitro screening approaches alone, but the AI/ML methods that could help us navigate this space paradoxically require large amounts of high-quality training data. At A-Alpha, we integrate high-throughput data generation (our

AlphaSeq platform) with AI/ML (our AlphaBind platform) to accelerate the discovery process while keeping the candidate funnel as wide as possible for as long as possible. A typical project for us might consist of: using AlphaSeq to generate large quantities of quantitative protein-protein affinity data for a diverse candidate library vs. a panel of relevant on- and off-target proteins; using AI/ML models pre-trained on our PPI database plus that campaign-specific data to train problem-specific models that suggest new

# Industry Insights

## AI IN PHARMA DISCOVERY

October 4th | Boston, MA



candidates to test, iterating as necessary to deliver hundreds to thousands of candidate sequences with optimized affinity and specificity profiles; and then using AI/ML predictions of biodevelopability to prioritize molecules for downstream characterization. This workflow is especially relevant when there are major specificity or cross-reactivity design goals, in which case the ability to generate and train on affinity data from multiple binding partners simultaneously gives us a huge advantage in training

relevant AI/ML models.

### **What are you most looking forward to at the AI in Pharma: Discovery Summit?**

I am most excited by the exceptional opportunity to engage with talented and passionate colleagues working on similar problems, with a wide variety of backgrounds and approaches. Integrating AI/ML into drug discovery in a way that accelerates the delivery of lifesaving medicines to patients in need is a goal that demands cutting-edge and multi-disciplinary approaches, and I'm very excited to share the

progress that the ML team at A-Alpha has made on these problems, and hear about progress from peer groups, in order to help drive the field forward as fast as possible. I am also particularly looking forward to the panel discussions, since they address issues that I think are pervasive in the field, e.g. how can we build a data-driven culture to best support AI/ML efforts? How can we effectively measure and advocate for the impact of AI/ML?

**Hear More From Ryan Emerson at AI in Pharma: Discovery on October 4th**

[www.aiinpharma-discovery.com](http://www.aiinpharma-discovery.com)