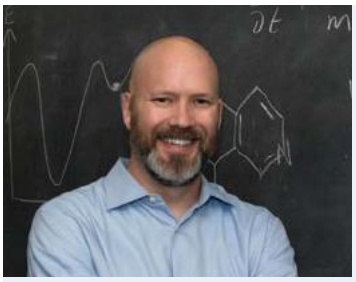
An Interview with Brandon Allgood, CTO and Co-Founder, Numerate

ARTIFICIAL INTELLIGENCE
TRANSFORMING PHARMA R&D
Boston, USA
February 21-22, 2018



Brandon Allgood is the CTO and Co-Founder of Numerate, a privately-held computational drug design company that is transforming the discovery of new medicines to fill significant therapeutic gaps by harnessing the ever-increasing amounts of drug discovery data and applying proprietary artificial intelligence algorithms at very large scale.

Numerate's drug design platform combines advances in computer science and artificial intelligence with traditional medicinal chemistry to overcome major challenges in small molecule drug discovery and significantly accelerate candidate selection and optimization. Brandon is talking at the Artificial Intelligence: Transforming Pharma R&D Congress and spoke about the pharmaceutical industry and AI ahead of the congress.

What are you excited about when it comes to applying artificial intelligence in drug discovery and what are the challenges?

The thing I'm really excited about, and this is what I'm going to be speaking about at the congress, is the impact AI can have in translating drug programs into successful clinical programs. The use of AI to drive drug discovery can improve the number of pre-clinical programs that become clinical programs, as well as, reduce the failure rate in the clinic, through modelling of key points of attrition like ADME and toxicity today, and, eventually clinical efficacy and safety.

One of the challenges in drug discovery is typically the dataset size. Unlike in other industries and applications of artificial intelligence medicinal chemistry data is more "medium data" and not "big data." This is not surprisingly due large to the relatively large cost of acquiring new data. There is no Mechanical Turk for med chem. What this means is that algorithms developed for big data problems need to be adapted and cannot be applied to our field without modification, often dramatic modification. The point is that there's still a lot of work to be done in applying artificial intelligence to low data environments which is where we're focused, especially when dealing with emerging biology.

It's not one algorithm to rule them all, it's not Lord of the Rings. There's still a lot of work to be done. Beyond the algorithms there are a lot of issues that still need to be sorted through.

Another issue is the industry has historically been data rich and algorithm poor. The industry has large databases full of failed molecules and some successful molecules but still, quite often, relies on medicinal chemists' memory to guide later stage discovery programs. They are forced to remember some poor chemical decision that killed a mouse from a 1980s anti-infective program and apply that to their hot IO program. This illustrates why drug discovery remains a funnel and the only way to appreciably change that is to apply artificial intelligence and have it inform every compound design decision based on all historical success and failure. In essence, we want to create a learning loop.

Algorithms development is being held back by the lack of data around ADME and toxicity. Numerate has now launched a data partnership program and we're now getting access to larger pharmaceuticals data sets. Large single assay, single source data sets are now being shared, creating an ecosystem where artificial intelligence can be more usefully applied within the context of pharma R&D.

What are the next steps and how will AI develop?

It's not one algorithm to rule them all, it's not Lord of the Rings. There's still a lot of work to be done. Beyond the algorithms there are a lot of issues that still need to be sorted through. There's a lot of ground-work to be done on the data, cultural and business side. Numerate has been hard at work on these issues.

Applying AI can't be seen as just an IT problem.

Some of the fundamental things that need to happen are IT based such as making data searchable and discoverable, but once that's happened then it's really a research problem, not an IT problem. As such, the R&D of applying AI to the drug discovery process needs to receive a real R&D budget.

AI is not your grandma computer-aided drug-design tool. CADD tools were designed and are used to help chemists explore their ideas and build a mental model of the problem they are trying to solve. AI algorithms can ingest more data than a human and build far more complex models that a human's mental model. For this reason, the role of AI and how it is applied cannot be the same as CADD tools. New processes need to be develop so that the AI and the human can excel at what they are best at.

Those with the data need to collaborate with those with the AI talent and technologies. There is currently a gap between those with the data and those with the algorithms and we need to bridge the gap between companies with data and those with algorithms through novel partnership and collaboration models. The encouraging things is that **companies are now starting to share the crown jewels within pharma which is the data.** We are slowly now building trust and sharing data and algorithms. I'm pleased that there is a willingness to explore business models and interaction models.

GSK and Takeda, for example, are working out how they can work with companies, such as, Numerate and to work towards a win-win business model and interaction model. I think we'll see some real payoffs from truly innovative partnership like this in the next twelve months.

Over the next twelve months certain interaction models will rise to the top. It's critical that it happens this year. Before the AI hype goes away, you want to be left with something substantive to move forward with.

What do data scientists need to do to help implement AI in pharma R&D?

I think what really needs to happen is that the average biologist and average medicinal chemist needs to get up to speed with how AI can help them and how they can interact with AI. They make the decisions and if they don't trust or understand AI then things will remain the same. **It is about a cultural shift within pharma but also in almost every other industry;** everyone will have to understand artificial intelligence on some level in every industry.

It's vital that those people working in those roles understand what AI can and can't do. It's about getting these groups to interact with these systems. As to training data scientists there's enough information out on the web that if pharmaceutical companies are serious about making the most of AI then it's about investing meaningful capital in their teams. Within the context of the overall pharma R&D budget, the investment would be rather small.

What is Numerate doing in relation to using AI in pharma R&D now?

We have been deploying and developing novel AI technologies used in small molecule drug design and this has been the focus of the company for more than 10 years. **We are kind of the grandfathers in this space.** Our focus for the last seven years has been on modelling ADME and toxicity and reducing the sources of attrition so that the entire drug discovery process can be data driven. The technology is being applied to our own internal pipeline and with partners, such as, Takeda.

What we've seen from applying our technology to over 25 different drug development programs is that our AI technology can not only make drug discovery cheaper and faster, but it can have a non-linear effect on translation. The real strength of what we have developed is its ability to unlock biology by virtualizing low-throughput / high-content biological assays, those for which biologists and doctors have higher trust and the ability to really attack the primary sources of attrition in discovery, ADME and toxicity.

A huge thank you to Brandon Allgood for sharing his thoughts on artificial intelligence in the pharma industry. If you enjoyed reading about Brandon's experiences, there will also be an opportunity to hear, and talk to him at the congress. For more information, visit: https://artificialintelligence-pharma.com/ or **download the agenda**.