Episode 4 – Drug Discovery

People

Samborne Bush, host

Jo Bates, Professor of Data and Society at University of Sheffield. Patterns in Practice lead

Dr Raminderpal Singh, founder of Hitchhiker's AI

Transcript

Musical intro

Samborne: Hello and welcome back to Patterns in Practice, Cultures of AI. Throughout this series, we're looking into how our beliefs, values and feelings shape the way we engage with AI.

Today, we're going to take a look at the drug discovery industry.

Later on in the episode, I'll be joined by Dr Raminderpal Singh to talk about his experience with AI in this sector and the Hitchhiker's AI community he founded. But to kick things off today, we're going to learn a little bit more about the drug discovery industry and how AI can be used within it. To do this, I spoke to someone who knows a thing or two about this area...

Musical interlude

So, to start things off, can you please introduce yourself?

Jo: Yeah, so I'm Jo Bates. I'm Professor of Data and Society at the University of Sheffield. And I've been leading the Patterns in Practice project for the last two or three years now.

Samborne: And the core of the Patterns in Practice research centres on three different industries, the creative arts, higher education, and of course, the theme of this episode, drug discovery. But what do we mean when we say drug discovery?

Jo: Well, I guess, you know, we're all familiar with taking medication for a range of different ailments that we suffer from, some more serious than others. While there's a lot of drugs certainly on the market, there's still room for a lot more, particularly for diseases where there's not great treatments or maybe the drugs that are available have a lot of side effects and things like that.

So drug discovery essentially is the process that scientists engage in to try and find those new drugs. It involves both working in the lab. So I guess like, you know, thinking, what you're used to doing chemistry class or something like that with all the different chemicals and so on.

But also it involves a lot of computational work where people are just sort of sat at their computers, analysing data and things like that.

Samborne: And so as part of the research, you've spoken to several people working in this industry. What sort of roles do they hold? Do they hold that either working in the lab or they're computational scientists?

Jo: Yeah, we spoke to practitioners or skilled professionals essentially in all sorts of different roles. So the medicinal chemists that do the work in the labs, the computational chemists and biologists that are working with the data and algorithms and so on to analyse the compounds.

We also spoke to managers that are looking after all of these processes and all of these people. So we really tried to get a rich understanding from all sorts of different perspectives as to what was going on within that space.

Samborne: From what you've learned, how is AI being applied in this industry?

Jo: Well, you know, there's a lot of hype about AI at the moment, but it's actually got a really long history of being used. So there's, you know, data mining techniques, which might now be described as like an early form of AI of some sort being used back in the 1950s.

And then there's been waves of interest and interest growing and then waning over the years. But throughout all this time, what they've been trying to use these techniques to do… is for prediction.

So to predict interactions between compounds, to predict toxicity of compounds and things like that. And I think, you know, with more recent computational advances around things such as deep learning and so on, we've got this new wave of interest in the sector around using what we're now calling AI.

And ultimately what they're trying to aim for is to speed up the process and make it more efficient. So there's a really low success rate within the drug discovery pipeline. A lot of investment and effort gets put into working on things that in the end don't work as drugs. So the aim is to try and use AI to predict whether something is likely to work or not and to speed up that whole process, so there's less wastage, I guess, in the drug discovery process and drugs get to market faster.

So that's the ultimate aim of why they're trying to apply AI around these sort of predictive processes.

Samborne: As you mentioned, computational models and tools have been used in this industry for quite a while now. But what do you think practitioners working in this industry's feelings are towards these newer methods, this deep learning and these AI tools?

Jo: Yeah, it's a really interesting question. I think it's very much mixed feelings. So some people are really quite excited by some of the new things that they're seeing in terms of AI techniques and so on that are available but others are more sceptical, particularly people I guess that have seen a lot of hype cycles over the years, you know they've got a lot of questions to ask about what's going on.

I think most people are sceptical of the hype in general, you know the sort of hype that we might see in the media that about sort of like these really speedy, amazing advances in drug discovery that AI is going to deliver.

I think there's a lot of scepticism about some of those claims and, but people, you know, might still be hopeful that there could still be progress in drug discovery using some of these techniques, but it's going to be more of a challenge and it's going to take longer than some of the hype would have us believe.

But then there's those people as well, that I think, you know, really believe that this whole moment of AI in drug discovery is overhyped entirely. And whilst they see that some of these techniques can be useful for certain tasks within the drug discovery pipeline, they have a lot of questions around whether the AI has the creativity to really fuel novel advances within drug discovery rather than just sort of picking away at problems that people would have been able to do anyway, maybe just a little bit slower or less efficiently in some cases.

So it's really a mixed picture. And I think it depends on what type of role people have, you know, whether they're more computational, whether they're more based in the lab and how much experience people have in the sector.

You know, if they've seen one of these hype cycles come through before, or if this is something that they're experiencing for the first time. So yeah, really it's a mixed picture of how people are responding to it.

Samborne: From your research and time spent talking to people working in this industry, did anything surprise you?

Jo: I think one of the things that a lot of people that we spoke to talked about was the data challenge in terms of trying to adopt some of the newer AI techniques such as deep learning within parts of the pharma sector.

So I expected that the pharma sector would be data rich. You know, you imagine it as an industry that has like tons and tons and tons of data, but actually it's not, at least for the types of data that people need to train these new types of deep learning models, often the data is not there.

One example that people talked about was data about what they call bad compounds. So these are compounds that may be unstable or they're never gonna work in drug discovery. And over the years, companies have collected a lot of data about the *good* compounds, the ones that do work and have some promise and potential and so on, but not the bad ones, they've just been discarded.

But to train the models, these new types of models, then you need both to be able to do that. So that's really a big barrier to adoption of some types of AI in some parts of the drug discovery sector. And one that I wasn't expecting to hear about. So that was nice to get a perspective like that.

Samborne: And since conducting the research, you've actually been back and ran workshops with drug discovery practitioners. What did these entail and what sort of reaction did you get?

Jo: Yeah, so actually we were exploring what we call the bad data problem with the practitioners. So we used an approach called 'theory of change', which is really commonly used in sort of policy making circles and social organizations and things like that.

And it helps practitioners to define an objective and then think about the barriers to achieving that objective and how they might be addressed within that local context of the company or the organization or whatever. We ran two workshops with different teams exploring how they might address this bad data problem in a way that everybody could buy into because some of the ways that they tried to address it previously you know it not worked out so well for everybody within the teams and things like that.

So this is what we were exploring and between the two groups they came up with a whole load of suggestions for how to tackle the problem going forward and these really range from you know culture change type suggestions to more technical things like data standardization and so on.

It was interesting the teams reported a real significant increase in confidence from pre-workshop, whether we can address this objective, to post-workshop, whether we can address this objective.

So it was nice to see that they increased their confidence with thinking about how to tackle it going forward.

Musical interlude

Samborne: I hope that conversation provided some insight into what drug discovery is, how AI can be used within it and the debates and conversations that are currently happening in the sector.

The next part of this episode is going to pick up where Jo left off – talking about change. As we heard, AI and computational techniques can be extremely useful in the drug discovery industry, but just how important are they?

Well, the next guest on today's episode is a self-confessed AI evangelist, has a vast experience working with drug discovery companies, and can help answer our question...

Raminderpal: So my name is Raminderpal Singh. I live in Leicester, I work for US company, trying to crack open the whole AI, Gen AI world in drug discovery and life sciences.

I also, parallel to that, about a year ago, I kicked off a non-profit called Hitchhiker's AI, which is again on that same topic, although that's a little bit more focused in the scope of the drug discovery work, but still trying to unlock the whole AI, Gen AI, popularize it, democratize it, make it understandable and used.

Samborne: You're a self-confessed AI evangelist. I wonder what this means to you?

Raminderpal: There is this barrier and it's this is real barrier around the adoption of AI and gen AI in life sciences. It's a *massive* barrier. And it's somewhat to do with the fact that it's human data and it's human disease. And these are complex biological issues.

It's somewhat to do with the fact that there's not that much data available to build models and things and whatever, but it's also cultural.

You know, people are just very resistive because they don't understand or they cut, they don't know how to take their first baby steps. And so evangelism for me is to get people across the line. It's actually about making a shift, you know, moving the dial on people's adoption, the actual company's adoption.

And for small companies, mid-sized companies, they have a major problem here, a major problem on this topic. The big companies- there are some mid-sized companies who are AI first.

But for the rest of the folks, let's say the 80%, this is a side topic which they know they need to look at, they kind of understand the value of because lab costs are so high that it'd be nice to do some computational work.

But they don't know how to get started. And they're really, they've seen a lot of people burnt on this topic, a lot of money spent. They see VCs not really that interested in anything but wet lab data.

And that creates a lot of, there's a lot of different skills involved. They don't have those skills. They've come from more of a lab environment or a clinical environment.

So where to go with all this? And that's a big problem. So it's very interesting to go after that.

Samborne: Just before we move on, we might have listeners who might not know much about the drug discovery industry and also might not know about how AI could be used in that industry. I wonder if you could tell us a little bit about how AI is being used in pharmaceuticals.

Raminderpal: Yeah, that's a very good question.

And first of all, let's just…say what drug discovery industry is because it's actually the drug industry, these big biopharmaceuticals, the whole way it works is actually four different, I wouldn't to call them industries, there are four different *phases*, independent phases of a supply chain.

The first phase is come up with a good drug candidate. That's called early drug discovery. That's like with test tubes and little, you know, Petri dishes and what they call it in vitro.

The second phase is go test that on animals.

The third phase is going to test that in humans.

And the fourth phase is once you've got, after the third phase, you get approval from the regulatory bodies. Then after you go sell it, manufacture and sell it.

And when you go from the first to the fourth, it gets bigger and bigger. The costs get bigger and bigger. The companies' investments get bigger and bigger.

So with all that going on, my work so far has been on the earlier phases, the more biology related phases, or chemistry related phases. People in white coats with goggles going around trying to make it up. And the problem there is that they're spit balling, they're guessing what's going on, they're trying different things out.

And even though the cost is nothing compared to the big late-stage clinical trials, which I mentioned earlier, the phase threes, et cetera, a lot of what happens there determines the rest of what's gonna happen in the whole program. And so, if they mess up and they get the wrong compounds and they get the wrong insights and the wrong justification and what they call efficacy, which means, works and then toxicity, safety, whether it's dangerous.

You worry 'bout two things, does it work? And is it dangerous?

Those are two things you worry about, right? If they don't get that kind of understood as good as possible early on, well, by the time it gets to further phase, you see it very clearly in humans, but you can close it down, but you've spent many years and a lot of money getting that far.

So computational techniques early are there to try to do variety of things. There are the obvious things which are speed up processes, the obvious things which is replace lab experiments. If I'm going to pay somebody to do it, what can I do on a computer and just do it quicker and just guess it sort of with a model. But there's also the case of getting more information on the table to make better decisions about these efficacy and toxicity type challenges in a situation where I'm kind of guessing because I don't even know if it's going to work in humans.

All these things become quite hard. So you need methods, even if they are spurious in what they do, so is what you're doing anyway in the lab.

But it's very hard to get your head around that if you've never, if you haven't come from a computer world.

There's a whole series of these biotechs who are trying to grow with this skill in mind. You could grow with computational first type thinking and it's very exciting. There's a, there's a lot that can be done and there's different things they do. They can predict toxicity. They can do pathway analysis, protein pathways to look at, you know, different protein strategies, where to target the target ID and many things going on there.

So it's not a single thing, but the challenge is that it's hard to get these things accepted in a system which is funded and respected based upon the data in the lab.

So bottom line is if I did nothing and I walked into a lab and I could pull it off with a data point on an assay, I've got this array of stuff going on in the lab and one of them works, because I can show the lab data works, that's worth gold.

If I can show it works on a computer, it's worth the lead.

That kind of fundamental cultural problem there, right? It doesn't mean that they're wrong. It just means that there's a problem to be solved.

Samborne: But why is it important to use AI?

Raminderpal: It's actually the motivation is less about why is it good to use AI. It's more like...we're going to lose a large amount of these companies if they don't get with the game.

Cause we know, we know if there's a curve, a bell curve, the bigger companies on the left of that curve, the smaller companies tail off on the right.

We know that the bigger companies will invest. So what are small companies gonna do? What are the mid-size companies gonna do? They're gonna die. They're just gonna die in the vine. So you're gonna have this overhaul in the way the market, you know, massive consolidation.

So this I'm fighting for the small guy. I know that sounds a bit as a bit precocious, but the point is that there is a market's fragmentation and then markets reshaping, which is going to occur.

And like we see in other industries, let's take social media, for example, how many small social media companies exist? Well, we knew it was gonna happen. When GDPR occurs, what did we know? The big guys will comply, the small guys will fail, right? Doesn't make GDPR wrong, but we knew it was gonna happen, right?

So what happens here is that we know the big guys will get through it. The small guys will struggle.

Why do I care?

The small guys actually are the innovation arm of the whole industry. The industry over the last 20, 30 years has evolved from big biopharma do their own R&D, their own invention, to small guys provide a huge amount of the innovation, invention in the ecosystem and the big guys just license, in-license it, buy them. That's just the way the industry operates. We all know that.

Well, how do you maintain that if the small guys are factored out, or, you know, squeezed out because they don't have the know-how, the experience, the confidence, the money, the VC, you know, buy-in, because the VCs have to buy-in. You see small guys, VCs, they buy-in to all these methods and the data these methods create.

So there is a problem and it's not going to shake out well for those who are not board. I fundamentally believe that everything's about science when you get there. It's not about computational techniques. The computational techniques are to help the scientists, the scientists are the centre of the universe.

And you've got some really brilliant scientists coming out who either come out of postdoc or they're coming out of, they're spinning out of a big company and they've really got something good they want to, they're going to take to the market. That's what we want, disease solved. That's what we want, right? And so if those small guys don't have the chance to stay on top of it, they won't have the chance to go innovate and therefore disease solved won't happen.

Now the big guys pay half attention to this, don't get me wrong, they know this, they need this, but at the same time, hey, if the competition's squashed, then that's good, right?

So at some point it's, I'm not trying to fight the big guys, I'm trying to help the big guys, but at some point, you know, I really believe that innovation and invention needs to be energised and needs to be protected. And that's more than throwing money at things.

So there's the UK is really good with innovate UK and grant money, but it's more than money.

It's enabling these guys and their skillsets and their confidence and giving them the clarity so that they can compete.

If it takes me two years or three years and $3 million to come up with a drug in an early phase, and it takes my buddy down the road one year and one million because she or he's got a computational technique, I'm not gonna compete.

That's a problem.

I know I've got a lot of data, and I know how to put them in Excel spreadsheet, and I know how to churn them, but I don't have these techniques.

And even if I knew where these techniques were, I don't know where to take them. I don't know how to journey map using them and what risks I'm putting myself up for and how to convince funders that this is good data coming out of these tools.

Samborne: And I think that brings us on nicely to Hitchhiker's AI and the work you've been doing with that organization. Could you speak about this and its purposes?

Raminderpal: Yeah, so as I mentioned earlier, Hitchhiker's AI is a community. It's a community of these smaller guys. It's actually what we call a platform business model, which means you have multiple parts of the community coming in with different angles of requirements.

Like you have tool vendors who need something, you have biotechs who need something. You also have academics and you have big companies and things like that. And what you're trying to do is you're trying to create an environment where the value of the innovations on the vendor side and the data science side and the consultants, so there are a lot of brilliant consultants out there, is being appreciated in the AI language, in the AI, so AI is the conversation, right, is being appreciated and seriously looked at by the scientists on the other side in these biotechs.

And the reason you want to do that is because today there's a block on that. The biotechs, the scientists, the CSOs, the CEOs, they're not really paying attention because of the reasons we talked earlier. There's a kind of fog, I talk about this fog. So they've got this, they've turned off.

How do you turn them back on? How do you make that easy without making it a sales pitch, without making it about marketing? How do you make it about problems? How do you make it about solving problems?

Samborne: So I wanted to talk a little bit about, say in 10 years time, everything went perfectly according to your ambitions and Hitchhikers' ambitions...What would the drug discovery industry look like?

Raminderpal: In five years time, as opposed to 10 years time, we're trying to get the industry to find it much easier and much more acceptable to use computational techniques in drug discovery.

And when we say the industry, I don't mean the big guys. They've got whole groups of computational. I mean, the folks I've been talking about, the people who are not enabled, they've not come from this, from this background, and they are smaller to mid-sized companies. And they're the 80% of the companies out there.

And so we want an industry which is starting to self-form, snowball itself on that. You don't need to have the industry change. You just have the industry going in that direction.

Musical interlude

Samborne: Thank you for listening to this episode of Patterns in Practice, Cultures of AI. All of today's music and sounds were produced by Craig Scott. We have an episode where we talk to him about his work with AI and music. Do go and check that one out.

I'd also like to say a big thank you to our guests today, Jo and Raminderpal. You can view a transcript for today's episode alongside a link to Hitchhiker's AI on our website. Just see the link in the description. My name is Samborne Bush. We will be back soon.