Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery

AI3SD Interview with Dr Martin-Immanuel Bittner
29/01/2021
Online Interview

Michelle Pauli
Michelle Pauli Ltd

29/09/2021

Humans-of-AI3SD:Interview-10

AI3SD Interview with Dr Martin-Immanuel Bittner
Humans-of-AI3SD:Interview-10
29/09/2021
DOI: 10.5258/SOTON/P0154
Published by University of Southampton

Network: Artificial Intelligence and Augmented Intelligence for Automated Investigations for Scientific Discovery
This Network+ is EPSRC Funded under Grant No: EP/S000356/1

Principal Investigator: *Professor Jeremy Frey*
Co-Investigator: *Professor Mahesan Niranjan*
Network+ Coordinator: *Dr Samantha Kanza*

Contents

1 Interview Details

Title	AI3SD Interview with Dr Martin-Immanuel Bittner
Interviewer	MP: Michelle Pauli - MichellePauli Ltd
Interviewee	MIB: Dr Martin-Immanuel Bittner - Arctoris
Interview Location	Online Interview
Dates	29/01/2021

2 Biography

Figure 1: Dr Martin-Immanuel Bittner

Dr Martin-Immanuel Bittner: ‘Combine human creativity and ingenuity with the raw power of automation and AI and we can achieve real breakthroughs”

Martin-Immanuel Bittner MD DPhil FRSA is the Chief Executive Officer of Arctoris, the world’s first fully automated drug discovery platform that he co-founded in Oxford in 2016. He graduated as a medical doctor from the University of Freiburg in Germany, followed by his DPhil in Oncology as a Rhodes scholar at the University of Oxford. Martin has extensive research experience covering both clinical trials and preclinical drug discovery and is an active member of several leading cancer research organisations.

In this Humans of AI3SD interview he discusses how automation can provide fully reproducible, robust and reliable research data, the rise of FAIR data and why experienced postdocs should not be spending eight hours a day pipetting small amounts of liquids.

3 Interview

MP: What's been your path to where you are today?

MIB: Our focus as an organisation is in generating the best possible data to make better decisions in drug discovery – in our own labs, and with our partners in biotech companies, pharma companies, and academic centres worldwide. The background is that, currently, reproducibility and quality in research are huge issues. We know, thanks to reports issued by Bayer and Amgen, that 80% to 90% of all drug discovery research findings cannot be reproduced by independent third parties. It is a massive problem. At the same time, more and more people are starting to look for ways to address this but many of the steps just don't go far enough. And so, we started Arctoris to rethink the way we generate drug discovery data – using automation and robotics to achieve fully reproducible, reliable research data. Our laboratory in Oxford is end-to-end automated and can run a very broad range of different assays and experiments in cellular biology, molecular biology and biochemistry/ biophysics using robotics and data science. That means we can generate data not only 24/7, but also in a fully structured, standardised and reproducible fashion.

MP: How did you get to this point? Why did you start caring about this problem?

MIB: My personal background is in medicine. I qualified as a medical doctor, and practiced as an oncologist in Germany. Within oncology, it's absolutely obvious that we need new and better treatments and therapies. I came to the UK to complete my DPhil (PhD) at Oxford to learn more about drug discovery and about where those new drugs that we desperately need actually come from. I was very surprised to find that, when I was entering the laboratories and doing my PhD with no previous lab experience, that so much time was spent by highly qualified PhD students and postdocs and even junior professors manually performing experiments: spending six, seven, eight hours every day pipetting small amounts of liquids from one vial to another, instead of spending the time thinking, reading, hypothesising and discussing results with their colleagues. It seemed to me that this work mode is very much out of sync with how other industries have been transformed over the past decades by automation, robotics, data science, the cloud and AI. That was really the starting point, seeing this very antiquated work mode, and then reading about these massive issues with reproducibility and research quality. Those two combined led us to then start the company in 2016 and build it to a globally operating biotech company over the past five years.

MP: How did you go from seeing that problem to thinking, "right. I'm going to start a company"?

MIB: The motivation came from seeing these interlinked issues with the work mode and data quality. At the same time, Oxford is an amazing environment not only for scientific research, but also for encouraging people to think outside the box and to think about the impact they can have on other people's lives. What for me became very clear is that what motivates me is the idea of having positive impact. As a doctor, it's quite straightforward because you treat patients and you can directly see the positive impact on another human being. As a researcher on the other hand, you can ideally work on finding one new drug, which could help tens of thousands of patients. And at the same time, when you look at all of these scientists, not just in the UK but globally, being held back by a lack of access to the right equipment, right infrastructure and right possibilities, you can really see that by augmenting scientists, by giving them access to the right type of research opportunities, you can transform the way we approach drug discovery – maximising the potential for positive impact.

I also met my co-founder at Oxford, who is a chemist by background, who at that time was also working in early-stage drug discovery, including both the university and large pharma. Together, we spoke with people in the industry, built a circle of advisors and mentors and said, "There is an opportunity for us to address these issues by harnessing the power of robotics and data science to bring these benefits to the research ecosystem." We started out with our very first R&D grant which we used to develop the first intellectual property (IP) around our core technology. Based on these first insights, and based on this first IP, we then raised funding from investors and built the company from an idea into something which now has its headquarters in Oxford, operations in Singapore, opening a Boston office by the end of the year, biotech and pharma partners on four continents, an active pipeline of drug discovery assets, and a rapidly growing number of experienced scientists, engineers and strategists. .

MP: How does it work in practice?

MIB: Arctoris is a platform company – we both progress our own discovery programmes, and we partner with researchers around the world, giving them access to our platform. We like to build long-term relationships and partnerships with researchers. In many cases, we have researchers and companies we've been partnering with for one or two years already. For example, a young biotech company is working on exciting new potential targets, let's say, for cancer therapy and they approach us and what they need is a rapid, robust way to validate these new targets. In this case, we combine our experienced in-house team of seasoned ex-biotech and ex-pharma scientists with the unique data generation capabilities that our robotic system enables. Together with our biotech partner we then design the project, execute it using our platform, and generate the data the company needs to make a decision about whether to move ahead on this target or switch to a different target. Depending on the size and complexity of a project we often talk about three months to six months that a project can last. We conduct projects along the entire value chain in drug discovery, ie we find and validate new targets, we generate hits as the first starting point to develop new chemical matter, we progress them into leads, and finally candidates, rapidly moving towards the clinic. We work across all of the processes that make up drug discovery, and we provide our own researchers and our partners with the right tools to move as quickly as possible from idea to the clinic.

MP: And then what happens to the data that's generated?

MIB: The data generated on the robotic system is captured and collected in our databases, which are accessible via a secure online portal. All of our partners have access to this online portal with their login credentials, and this is where they can then find all of their data being uploaded and updated in real time. In addition to the primary results data, all experiments conducted on our system also generate what is commonly referred to as metadata. Metadata is something that in the life sciences context is only starting to be regarded as an important contributing factor to overall data capture and ultimately reproducibility. Metadata describes in detail how the experiment is being conducted. It describes, for example, the individual steps of the experiment. It describes the temperature, the humidity and the sequence of operations. It describes all the experimental parameters that went into generating that particular piece of results data. That is absolutely critical when it comes to reproducibility because, right now, scientists around the world rely on ambiguous protocols and incomplete data capture in their research. The protocols currently in use could, for example, say 'mix the sample'. But of course, you can mix a sample in many different ways. You can shake it. You can stir it. You can use a vortexing machine. There are so many different ways of doing it, and all of them will lead to slightly different biological outcomes – with all of these variations adding up to noticeable differences in results, which undermines our ability to make unambiguous and clear decisions, thereby delaying research projects and, even worse, increasing failure rates. In contrast, our approach captures all of the metadata, that would for example describe this mixing procedure as performed by steering at 30 rotations per minute for two minutes at 20°C. So you're getting an exact description of what actually happened in that experiment – and you can reuse this script to generate more data. The metadata we capture, and the fully defined protocols or scripts we use, is the key to ensuring that data can be reproduced from one run to another, from one scientist to another, from one project to another. This is what our platform enables us to do, and what we use to accelerate and improve research processes.

MP: When the research is published, who has access to that metadata?

MIB: 90% of the organisations we work with are biotech and pharma companies and they don't normally make their results available in open access. They use the results for their internal research purposes to progress new drugs into the clinic. Academic researchers, on the other hand, are usually encouraged to share their data once a project has been completed with the rest of the scientific world. We, as an organisation, are huge proponents of the open science and open data movement and we always encourage sharing of data with others. But of course, it is not our data, it's the researchers' data so they have to decide on what they do with it.

MP: What is the potential in this area?

MIB: There are three areas where we see challenges but also a lot of potential for the next two to four years.

First, there is reproducibility. We know that reproducibility is something we as an industry need to address urgently. We can't continue having 90% of all research not being reproducible. It sets research on very shaky foundations. It can lead and does lead to many projects never coming to fruition because the data doesn't support certain conclusions. So we need to work on reproducibility, and automation with all its benefits is a powerful way to do that.

Second, there's also an increasing trend towards open science, open data, FAIR data. It's a very big topic in research right now – making data findable, accessible, interoperable, and reusable. And that is something where, again, much of the conventionally generated data is very hard to bring into a FAIR format, whereas something that was generated by a machine and which comes with all of the associated metadata is far easier to put into an open science, open data context and then make it reusable for other researchers.

The third area where there's huge potential is the use of artificial intelligence. Obviously, more and more research is supported by AI and machine learning. However, the key to machine learning is the quality of the input data. The input data determines how good any machine learning algorithm, any machine learning insight can be. This is, again, something where data generated using robotics comes with the structure, the standards and the annotation that makes it really usable, useful and valuable for machine learning applications.

MP: Are there any pitfalls, anything to be wary of?

MIB: When we speak with others in the industry, one of the questions often asked is, “So, what about the researchers? Won’t we lose valuable information and valuable insights that the researchers themselves have?” We take this concern very seriously, and we observe this very closely. Yet, in fact, in all of our partnerships it’s not about having fewer researchers and projects, it’s about having these same researchers doing more valuable tasks and pursuing more projects than before. If you have someone who has a PhD and 10 years of postdoc experience, it simply is a waste of potential to have that person then spend their days pipetting liquids into 96 well plates.

This is where more and more people realise that it really is about augmenting scientists, giving them more and better opportunities to think through an experiment, properly plan a project, direct the research, and be in control but not be the ones actually doing the lab work. This is very much a mindset shift where people need to think it through and see it’s just a different way of approaching research that puts the human scientist in charge by making sure that the repetitive menial tasks are taken over by a machine.

MP: How has Covid-19 affected what you do and how are you having an input into the Covid response?

MIB: First, Covid-19 has forced the world to think about remote work and suddenly many professions that we thought had to happen in an office or in a hospital we now see can happen from home. Trends like telehealth have seen massive acceleration over the past 12-18 months as people see the opportunities that come with enabling remote work. We’re seeing the exact same thing happening in the biomedical context as well, where up until now scientists had to go to the labs. If the scientists weren’t in the lab, they couldn’t get any work done. Thanks to the platform that we built and thanks to the use of automation, we were able to conduct experiments throughout the whole pandemic without any impairment. That means that researchers who were forced to be at home and who couldn’t go to the labs were able to continue their life-saving research by working with the platform that we provide. That has helped increase resilience and ensure research continuity.

Second, we’re quite proud that we were able to support several biotech companies that were doing Covid-19 drug discovery, conducting experiments and assays on our platform. We established these dedicated assays, and they then enabled research into which drugs might potentially be used for Covid-19 patients. In that context, two aspects were critical. One is providing access. Many of these campaigns would not have been possible without our support, and we provided these researchers with access to the necessary assay capabilities. And second, automation also means that experiments can happen faster. We completed projects within 24 hours where we received, for example, a new set of molecules from a pharma or biotech company on a Monday morning, and by Tuesday morning, we had profiled them, characterised them, accumulated all data points and sent back a full report on how these molecules interact with the target of interest. In a conventional setting these profiling exercises take at least a week, sometimes even two weeks.

MP: What advice would you give to early-career researchers in your field?

MIB: This is an exciting time to be in research. In my view, biomedical research is really being recognised, especially during these challenging times, as one of the most important pillars of public health and of securing the health of our populations. I think it is wonderful to have a career in biomedical research and to become an inspiring biomedical researcher. At the same time, I think there's a few things that are changing now that it's good to be aware of. For example, aspects like data quality, FAIR data practices, data stewardship are becoming more important, and more and more funding bodies also require researchers to, for example, have data management plans in place. I always encourage researchers not to see this as an afterthought or as a burden, but instead to see data management plans and the idea of making their data reusable as a way to ensure that all the research money (taxpayers' money, charity money etc.) is, first of all, spent more effectively, and at the same time, that their research can really be of benefit to themselves and to other researchers in the future. It is a bit unfortunate that sometimes the "publish or perish" culture in research can incentivise researchers to hold back data and to build a data trove of their own that they then see as theirs. The important realisation is that progress is accelerated when we all work together, especially in an area as essential as biomedical research. In other words, we should really make sure that we share data with one another and that we generate new insights together as opposed to working against one another.

MP: Does interdisciplinarity have a role to play in that as well?

MIB: Absolutely. I think we're seeing more and more of these projects where, for example, data scientists and wet lab scientists work together. That is extremely promising, because one of the core insights of the whole AI space over those past 20 or 30 years has been that machines on their own and humans on their own both are inferior to humans and machines working together, be it chess computers or AI-driven drug discovery. Once we combine the two, the human creativity and ingenuity with the raw power of automation and AI, we can achieve real breakthroughs.