

Transforming Life Sciences Data to Drive Insight and Innovation.



Atos



We spoke with Dr Katharina Lauer, Scientific Lead of BY-COVID at ELIXIR, about the role of technology in science, how the pandemic has helped shape the assimilation of scientific knowledge, and the challenge of managing the increasing complexity of life science data.

ELIXIR is an intergovernmental organization that supports the research and development of new drugs. It brings together life science resources – databases, software tools, training materials, cloud storage and supercomputers – from across Europe. ELIXIR manages these resources so that they form a single infrastructure that makes it easier for scientists and researchers to find and share data, exchange expertise, and agree on best practices. Founded in 2013 and having implemented its first scientific program in 2014, it now includes 22 member countries and more than 220 research organizations across 23 countries.

So, why does life sciences need an organization like ELIXIR?

As we know, the challenge with data today isn't generating it. It's how we manage it so that it's useable – and more than that, so that it gives us insight and is able to drive innovation in whatever field.

At ELIXIR, we see ourselves as an umbrella organization for bioinformatics – a safeguard of all the different types of life science data. So everything that's involved along

the data value chain from data production to innovation – the underlying infrastructure, databases for software analysis, cloud compute power. And not just the tools, but the training to be able to use them.

It's also a platform for collaboration – a place to exchange knowledge. By bringing all these aspects of data together, we make the technology piece of the puzzle easier for researchers, allowing them to focus on their specialist biotechnology or pharmaceutical fields.

Tell us about what's been your focus for the last couple of years – what's the BY-COVID project?

Prior to COVID-19, ELIXIR didn't have a specific infectious disease focus. When the pandemic hit, we wanted to find areas of infrastructure and technology we could repurpose for the COVID use case, which is how BY-COVID came about.

It's essentially a one-stop shop for COVID-19 research. I like to describe it as a kind of discovery portal. So you can access proteomics information and genomics information, but the scope is also a lot broader – going



beyond the biomolecular aspect to incorporate social science, clinical data, epidemiological studies. What we aim to do is help answer the bigger research questions around COVID-19. Make all the data available in one place, and make it easy to find, share and analyze.

The pandemic has changed many aspects of our lives – do you think it's also changed how we approach the assimilation of scientific knowledge?

I think it's made the industry, data generators, governments, policymakers – all parties really – have to think about data sharing in a new way. It's also accelerated some aspects of data openness. For example, as part of ELIXIR's response to the COVID-19 pandemic, we had to rapidly coordinate open deposition of datasets from different countries, institutions, and disciplines. There had to be fast agreement to move forward.

And when you look at how quickly we as a society needed to respond to the pandemic, the other critical factor, especially when we were setting up BY-COVID, was sustainability. Ensuring the infrastructure we were creating could be applied to any area of research, infectious disease, emerging threats – anything.

I'd like to think what we've learnt about bringing together knowledge, data sharing, openness of scientific data – that these attitudes and approaches will endure and help us deal with future situations.

Can you summarize your thoughts on the role of technology in the life sciences field?

It's interesting because the way technology has evolved, we don't really think about it on a daily basis. It's there and we take it for granted to a large extent. But with the work we do at ELIXIR, technology takes on a completely different level of importance.

Today, without this technology, research is impossible. When I look back to my time working in the lab and trying to analyze data on some outdated hardware, I remember how frustrating it was. All you want to do is align three sequences and you just can't do it.

The technology we have today is just miles removed from that. It no longer limits the work researchers and scientists want to do – it enables it and makes the time to insight even faster. Cloud has changed everything for most fields and industries but in life sciences particularly. The impact it can have in real life – it's a game changer.

Do you think the life sciences industry is really taking advantage of all the data that's being generated?

Of course, we can always do better. And this comes back to your first question – why do we need an organization like ELIXIR? Because however well we're doing, however much progress we've made, we can always do better.

So I guess the other question is, where can we do better?

Personally, I think we'll see a lot more automation, which will help us derive more value from data even faster. Another key area that can help us progress in life sciences, is paying closer attention to longevity of data. Not just how you, say, as a data generator want to use it, but how other people will be able to pick it up and run with it, or apply it to other problems.

Data is probably the most valuable resource of modern times. I think if we're smarter in how we use it at all stages, we'll see amazing results.