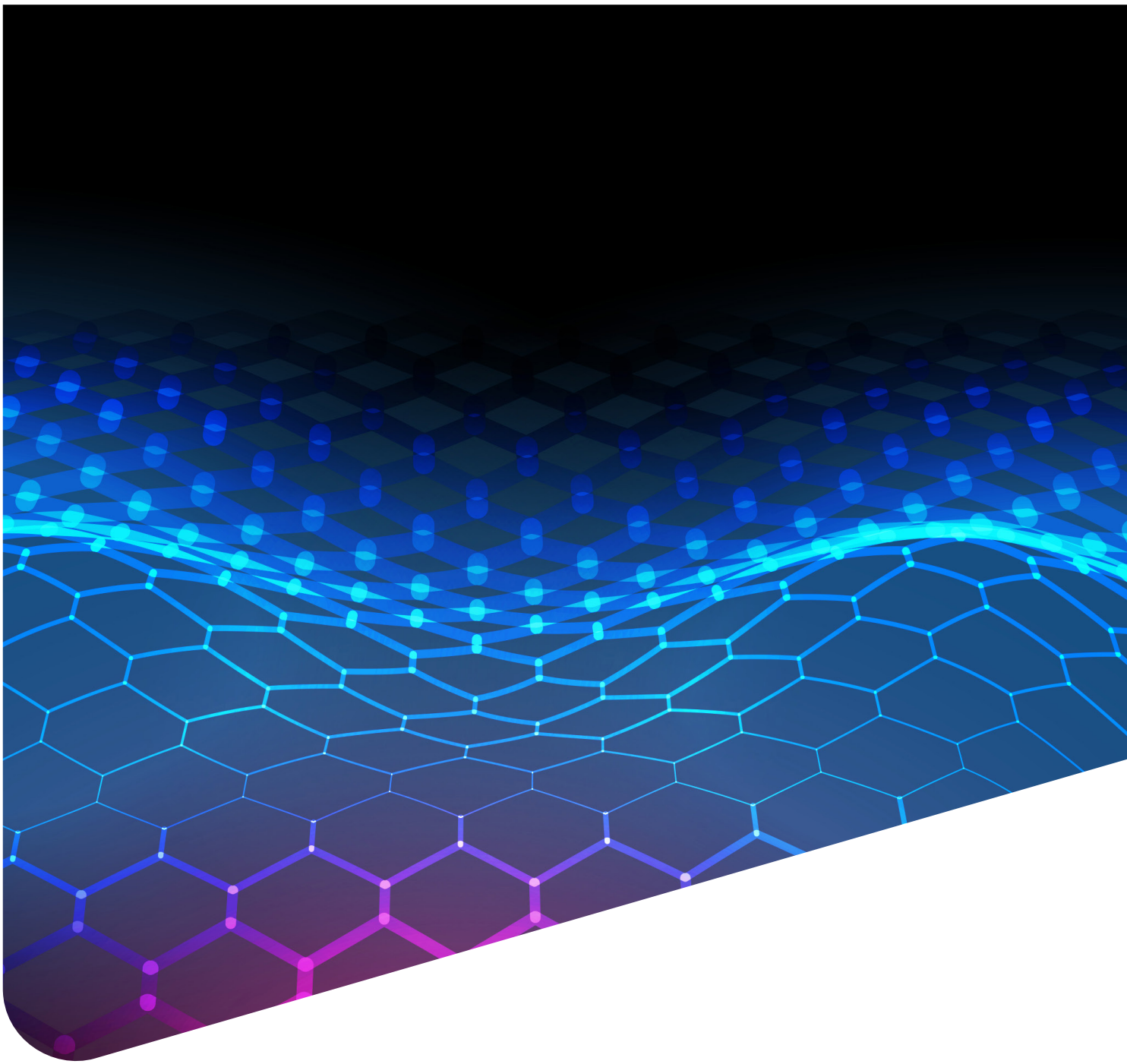
SPRINGER NATURE

springernature.com



REPORT

The Digital Transformation of R&D: Navigating the Digital Lab and Solutions for Efficiency

ADVANCING DISCOVERY

The increasing complexity of scientific research is driving a revolution in the way life science laboratories work. The exponential growth of data, and the need for more efficient and effective ways to conduct experiments, analyse data and collaborate with colleagues, can no longer be handled using traditional methods. A digital transformation of the sector is needed to reduce costs, reduce time to market, reduce waste, and increase reproducibility.

This change is being driven by a huge need to improve the returns on research and development (R&D); particularly in the biopharma sector where the average cost of developing a new drug has reached an astonishing $2.8 billion and only 12% of the drugs which enter clinical trials are approved. With this in mind it's hardly surprising that companies are looking to rethink their approach.

A revolution in life science research

So what does digital transformation look like in practice? Dr. Robin Padilla, Director of Product Management at Springer Nature Digital Life Science Solutions sees it a three-step process moving from digitization through digitalization to full digital transformation.

DIGITIZATION

The process of making information available and accessible in a digital format.

DIGITALIZATION

The act of making process more automated through the use of digital.

DIGITAL TRANSFORMATION

The process of devising new business applications that integrate all the digitized data and digitalized applications.

The journey began in the early 2000s with the adoption of technologies such as laboratory information management systems (LIMS), electronic lab notebooks (ELNs), imaging analysis software and laboratory automation systems. The next decade moved into the era of digitalization with the integration of cloud computing, bioinformatics and increasing use of virtual and augmented reality (VR/AR). This was followed by the implementation of artificial intelligence (AI) and machine learning (ML). Moving forwards technologies such as digital twins, blockchain and the integration of the Internet of Things (IoT) are set to further revolutionize the way life science research is done.

Padilla's team are responsible for the Springer Nature Experiments solution which allows researchers to find, evaluate, use, and adapt protocols and methods. But, he says, it's no longer enough to supply these as PDFs for a human to read. "We need to make things available in machine readable formats so that they can be taken and read by ELNs, LIMS and instruments. Doing this will cut down on errors, improve reproducibility and allow experiments to be automated." Padilla says his team have a number of collaborations in this space to understand how best to develop services for the future.

Moving from data to prediction

The real digital transformation comes from being able to do things that were never possible in a paper-based world, such as using text and data mining (TDM) to analyze data and come up with hypotheses. Another example is looking at how generative AI can be used to augment existing datasets leading to more robust statistical analysis, optimize experimental design and even assist in the generation of novel molecular structures for drug discovery and design.

For labs trying to achieve a full digital transformation, one of the biggest challenges is dealing with huge volumes of data and the lack of heterogeneity of that data. The cost of poor data is significant. According to the European Commission and PWC not having FAIR research data could cost the European economy as much as €26 billion a year.

Springer Nature believes it can be part of the solution. "We've got 6 million machine readable FAIR data sets," says Prathik Roy, Product Director Data Solutions and Strategy, "that we make available using industry standard application programming interfaces (APIs), allowing customers to look for relationships, for instance, between symptoms and diseases" [EB1] [IC2]. Until recently Roy had seen his customers wanting small, structured data sets but he says that AI has changed this with customers wanting as much full text as possible to train their large language models.

Life science companies are also looking to combine these data sets with their own assets to identify which drugs to develop or which existing drug they might be able to repurpose for a specific disease. Innovations such as tagging the text with chemical formulae in the simplified molecular-input line entry system (SMILES) format or properties such as connectivity value or toxicity data levels are seen as being able to move the customers further into the sphere of predictive analysis, not just in pharma but also in materials science. This is staggering and has the potential to completely replace experimentation, says Roy. Developments such as curated datasets – of, for instance, dosage data – provided to a customer is just one way in which Roy sees his team as being a key player in moving the value chain from understanding semantic relationships to predictive analysis.

Time and reproducibility

Data isn't the only challenge facing life science researchers. Time and reproducibility are also key issues for organizations looking to develop new products. Dr. Lenny Teytelman, founder of protocols.io, describes spending 18 months of his post-Doc research trying to correct just one step in a three-day protocol. This wasn't developing a new method, he emphasises, just tweaking something to make it work.

In Teytelman's case his frustration led him leaving research to become a tech entrepreneur, building a secure platform for developing and sharing reproducible methods. Although science itself is moving incredibly fast, he says, the way that it is done moves very slowly and he wants to help change that. Researchers have their protocols in a variety of forms including word documents, ELNs and even on paper but these aren't shareable and can't easily be reproduced. Teytelman's vision was to create a database of protocols which could be shared or held securely as required. Since protocols.io's acquisition by Springer Nature in summer 2023, it has been able to think strategically about its use of generative AI to help users upload historical protocols into the database. Large language models are also good at comparing text so a user can start comparing the protocols in the database to find those which match their needs all the while understanding the modifications which have been made along the way.

The value of protocols.io is endorsed by gene sequencing company Girhlet Inc who wanted a way of sharing protocols across its multi-location teams as well as with customers. Co-founder Anitha Jakaprakash says that having a shared workspace is indispensable in its work in profiling T cell receptors. "We use it daily to help with reproducibility and collaboration across our multiple locations."

What does the digitally transformed lab look like?

A fully digital life science lab would integrate advanced technologies to digitize and automate every aspect of the research process, from experimental design to data analysis, enabling faster, more efficient, and more collaborative scientific discovery.

So how close is this to being reality? In Padilla's view the fully digitized lab is already a reality. Certainly, there are companies out there who deliver a number of the 8 facets of a fully digitized lab. Emerald Cloud Lab offers access to more than 200 instruments, all fully software controlled and automated, which allows scientists to design, execute and analyse experiments remotely from anywhere on earth. Benchling ticks rather different boxes with its claim to be the only biology-first platform for scientific data, collaboration, and insights, which it says helps scientists accelerate the full R&D lifecycle. Padilla's view that fully digital labs are the present not the future is endorsed by The Guardian's Tom Ireland in his 2022 article `Cloud labs and remote research aren't the future of science – they're here.'

So what are the next steps in the digital transformation of life science research? Robotic labs are clearly becoming an established thing offering their customers the

opportunity to significantly speed up research at a greatly reduced cost. But what about the other elements of the digital lab? In its 2022 Global Life Sciences Outlook, Deloitte reports that companies are prioritizing investments in AI and cloud. While consultancy company McKinsey asserts that by broadening the technologies they invest in life science, companies could dramatically improve their bottom line.

Springer Nature is investing in its businesses which support the digital transformation of life science, research with solutions which help organizations benefit from the data they work with. Traditional content businesses are now looking to use technology, such as AI to assist the research process, to integrate that content with tools and services. Padilla, Roy and Teytelman all see the benefits they can offer users by working together to exploit the opportunities offered by a digital lab. “Instead of a lab where physical equipment slows the pace down, we can look at generative AI, recognise its power and see how it can be applied to save time and waste and deliver transparency”, says Padilla. Its early days, he adds, but very much an emergent strategic space.

Interested in seeing how Springer Nature can help you transform your lab? Contact **rd@springernature.com** or scan the code.