How data is driving the next generation of medicine

AstraZeneca has been a global leader in pharmaceutical research and development for decades. Today, the company's scientists and clinicians are using AI to help redefine medical science in the quest for new and better ways to discover, test, and accelerate potential medicines.

BRINGING NEW MEDICINES to today's market can be…complex and complicated.

In the biopharmaceutical industry, it can take up to 15 years to go from an idea for a drug to a medicine that's approved for patients—a process that requires (and produces) a staggering amount of data. Even for a global biopharmaceutical leader such as AstraZeneca, research and development have traditionally been lengthy endeavors plagued by arduously manual processes. And because of recent advances in understanding the human genome and how it intersects with disease biology, the company faces an influx of data as it looks to identify new drug targets and develop next-generation therapeutics.

Back in 2019, Anna Berg Åsberg, Global Vice President of R&D IT, began conversations with R&D leaders about where her department could apply technology to automate and augment R&D processes as well as help make the best decisions on which research projects to advance through the pipeline.

"The hypothesis was, if we can automate routine tasks, that would make a massive difference," Åsberg said. "That's where we teamed up with our internal data science, analytics, and artificial intelligence experts and started to build solutions on AWS."

Amazon Web Services (AWS) and AstraZeneca were already well-acquainted. An industry-

leading cloud provider, AWS had been instrumental in AstraZeneca's internal infrastructure as a service (IaaS) strategy—helping the organization learn how to rapidly spin up new systems and environments. But as AWS' product and service portfolio expanded across data, analytics, and machine learning (ML), AstraZeneca saw an opportunity to reach beyond simply evolving its infrastructure and IT processes—it could support a transformational shift in data strategy across its entire organization and re-envision data driven patient experiences around the globe at the same time.

"Growth through innovation"

AstraZeneca was no stranger to data, analytics, AI, and ML; it had long since implemented tailored solutions into select processes, albeit mostly siloed and scattered across the organization. But in early 2019, a more centralized vision for ML implementation—anchored on top of a cohesive, company-wide data strategy—began to take shape: AstraZeneca was launching its "Growth Through Innovation" corporate strategy, which included the need for investment in data and AI. Corporate readiness met market reality.

Fittingly, much of the company's initial investment into data and AI took place within R&D, where legacy processes dominated and potential impacts to patients were clearly visible.

"My team was working very closely with R&D to try and merit out what kind of technology we can bring to the table," said Åsberg. "Is it prediction? Is it automated decision making? Do we have the data? Is the output actionable? Who will use the data?"

While the R&D team was working through use cases and implementation, Growth Through Innovation's wide footprint began simultaneously igniting and unifying investments in data and AI across the entire organization. The early investments on the R&D side of the business turned out to be broadly applicable; use cases in cybersecurity, infrastructure, and data migration multiplied rapidly. Predictive hardware maintenance has significantly reduced machine downtime within AstraZeneca's colossal pharmaceutical manufacturing operation. A natural language processing (NLP) machine learning algorithm was even implemented to assess free text for key terms that needed to be loaded into coding tools—drastically cutting hours from lengthy manual reviews.

By combining clearly stated "top-down" organization needs with "bottom-up" engineering ideation, the company could identify and focus on uses cases with clear business and patient impact.

"Our end goal is to empower all employees with data and AI. We always try to identify the business problem, what's the right tool to solve it with, and will it make the boat move faster," said Jeff Haskill, VP of Enterprise Technology Services. "We're also a company with massive amounts of data—how can we use that data? If it's not being used, is it really worth anything to us?"

To make full use of that data, it had to be processed and harnessed correctly. Even in their initial stages with AI, Åsberg and team were focused on building scalable and reusable solutions that went beyond simply solving temporal or isolated problems. A key piece of AstraZeneca's development puzzle was its alignment to the FAIR data principles, a framework originally published in 2016 by the journal *Scientific Data*. By relying on FAIR data—data that's **F**indable, **A**ccessible, **I**nteroperable, and **R**eusable—AstraZeneca could ensure that its solutions were as broadly applicable and modular as possible, even beyond its four walls.

"While continuing to adhere to all required privacy and regulatory requirements, we ultimately want to get to the point where we can share data with research centers, universities and partners where their use of data will benefit our patients," Haskill said. "FAIR does that for us."

In just a few years, AstraZeneca had completely transformed significant portions of its R&D organization—including both the automating or augmenting of key processes. And a key contributor to this effort was AWS, simultaneously serving as both strategic thinker and reliable platform provider.

"AWS is a key partner in our R&D data strategy, because it allows us to move at speed," Åsberg said. "It's critical to helping our R&D colleagues test new ideas and scale—and bring value."

Theory into practice

One of AstraZeneca's core AI implementations was something called AI Bench. Originating in the early days of the company's data transformation, it utilized Amazon SageMaker (AWS' fully managed machine learning service) and Amazon EMR (next generation cloud to leverage enormous amounts of data) to provide data scientists with a unified data science workstation. Whereas data scientists embedded in teams scattered across the organization had previously operated in siloed environments, AI Bench brought their work into a central system—meaning individuals could tap into and borrow from the models and innovations of their peers while still protecting sensitive data.

AI Bench's 1.0 iteration was valuable. Its 2.0 iteration, which incorporated AWS cloud technologies to help expedite and automate software and application releases, was invaluable. It also carried a GxP validation, an FDA distinction that ensures pharmaceutical products are safe and in alignment with key processes. That meant data scientists could use AI Bench far later in the R&D process and on increasingly complex problems.

AstraZeneca has even found success in AI-assisted literature surveillance, turning to natural language processing—also powered by SageMaker—to review hundreds of thousands of abstracts each year for key signals. This means potential data trends or patterns in how medicines are performing in-market are identified faster and more accurately than ever before. That's the kind of speed and efficiency that translates to tangible patient impact: in some cases, AstraZeneca projects might be able to reduce as much a year from certain product development timelines.

The list of implementations goes on—and, more importantly, will continue to grow. By leveraging AWS tools and with the Growth Through Innovation corporate strategy as a north star, AstraZeneca is continually expanding its investment in data, AI, and ML—surgically identifying use cases, then building scalable solutions that are designed to outgrow them. It's modern innovation, transforming one of the world's most impactful biopharmaceutical companies.

"We're thinking big, starting small, and scaling fast," said Åsberg. "It's exciting to see how AI can enable our scientists to push the boundaries of science to deliver life-changing medicines."

This story was produced by WIRED Brand Lab for AWS.