Microsoft

Hygiene technologies leader Ecolab brings data science to production with Microsoft Azure and Iguazio

With a century-long track record of making the world cleaner, safer, and healthier, Ecolab, the global leader in water, hygiene, and infection prevention solutions, wanted to add to its customer offerings by developing sophisticated predictive risk models for water systems, industrial machinery, and other applications. To do so, it turned to Microsoft Azure and MLOps expert Iguazio, dramatically reducing its machine learning development and deployment times, cutting costs, and enabling its geographically dispersed team to collaborate seamlessly on accelerating the rollout of AI applications.

Customer
Ecolab
Website:
https://www.ecolab.com/
Country: United States
Industry: Chemicals
Customer size: Corporate (10,000+ employees)

Customer profile
Headquartered in St. Paul, Minnesota, Ecolab is a major chemicals manufacturer that specializes in water treatment and safety, with operations in more than 170 countries.

Software and services
Microsoft Azure
Azure DevOps
Azure Machine Learning
Azure Pipelines

> "Azure, given all its agility, felt like the natural fit. The tight integration with Iguazio put us over the top—enabling us to drastically accelerate deployment of AI."

Craig Senese,
Senior Director of Advanced
Analytics,
Ecolab

Headquartered in St. Paul, Minnesota, Ecolab is driven by a mission to make the world cleaner, safer, and healthier, helping customers succeed while protecting people and vital resources. The company employs 44,000 people in more than 170 countries and serves nearly 3 million customer locations in the food, healthcare, hospitality, and industrial markets.

Striving to better serve its customers with deeper, more holistic, long-term risk insights, Ecolab's machine learning journey began in 2016. What took off as a project to develop bacterial growth risk models using existing sensor data eventually became a much larger, organization-wide machine learning team.

Going digital after 100 years

"Ecolab has been a leader in water, hygiene, and infection prevention solutions for nearly a hundred years," begins Craig Senese, Senior Director of Advanced Analytics at Ecolab. "We have substantial digital ambitions and are focused on combining a century's worth of experience in products and services with new advances in digital technology to solve specific customer challenges. We wanted to develop sophisticated predictive risk models for our customers, using sensor data, to enhance our existing product offering."

To be successful, Ecolab needed a data science collaboration platform that would bring together its large, geographically dispersed. and diverse team, all while efficiently using cloud computing resources. "Having maintained a strategic relationship with Microsoft for quite a while, we were very much invested in Microsoft products. Azure, for all its agility, also felt like the natural fit, The tight integration with Iguazio put us over the top—enabling us to drastically accelerate deployment of AI," explains Senese.

Cutting through complexity

At the same time, the Ecolab team recognized the inherent complexity of building, deploying, and maintaining machine learning models in production—and the challenge of attempting to scale data science across an enterprise. According to Gregory Hayes, Data Science Director at Ecolab, "Prior to 2019, commercializing machine learning models at Ecolab followed a 'rewrite-and-deploy' pattern, where model development occurred independent of the application developers." This approach led to deployment timelines exceeding 12 months on average. "In essence, we wanted to simplify the process of provisioning secure, scalable, cost-effective compute resources. But we also needed to reduce the timeline from model 'proof of concept' to production deployment."

The company also facilitated greater collaboration both within, and across, the advanced analytics function by adopting a machine learning operations (MLOps) process. "It enables a code-first approach to model development," says Hayes, "leveraging the best practices from software engineering, adapted to meet the needs of data science."

To make its machine learning vision a reality, a recommendation from Microsoft pointed the Ecolab team in the direction of technology firm Iguazio. As experts in MLOps, the Iguazio team were able to provide a solution that enables a faster and easier deployment-to-production process through robust feature engineering, scaling, and MLOps automation capabilities. The Iguazio solution is built on an open-source architecture that includes Kubeflow, which Ecolab was already using, so it was readily adopted to extend and simplify the process. "Ecolab reached out to us for a trial, during which team members working in-platform were able to work collaboratively and significantly reduce the time to deploy new AI applications, at a fraction of the effort previously required," comments Yaron Haviv, Co-Founder and CTO at Iguazio. "Soon, data engineers, data scientists, application developers, and product owners were all working together to automate their machine learning pipelines, continuously developing, integrating, and deploying data and AI products, scaling their workloads horizontally while significantly reducing execution time. It was exciting to be able to help Ecolab realize its ambitious plans for new AI applications." Ecolab used Iguazio in combination with Git and Azure DevOps to develop the feature engineering pipelines and machine learning applications. "The enhanced visibility provided clarity on expectations, roles, and responsibilities," he adds.

"It was a huge win," says Hayes, "taking advantage of Azure DevOps and Iguazio, and a continuous integration/continuous deployment workflow, we were able to integrate the creation of machine

Microsoft

> "Prior to implementing this solution, model deployment times exceeded 12 months. Thanks to Iguazio and Microsoft Azure, by 2020, these had been reduced to between 30 and 90 days."

**Gregory Hayes,
Data Science Director,
Ecolab**

learning models into the agile development process. This allowed our data scientists and machine learning engineers to work side by side with the product development teams, which improved cross-functional collaboration. Ultimately, these gains enabled our teams to more quickly build and deploy models at scale to meet business requirements."

Rising to a new challenge

The deployment of the Iguazio platform coincided with the eruption of the COVID-19 health crisis, fueling a dramatic increase in demand to incorporate data science models into Ecolab's broader digital offerings. "Going into 2020, our expectation was to have no more than 20 users on the platform, and maybe one to three machine learning solutions," notes Hayes. "In essence, we expected just a handful of solutions running in the environment, as we really wanted to give ourselves time to scale up. But with the pandemic, our customers had to pivot, which led to a tremendous increase in demand for data science development."

By the end of 2020, Ecolab had more than 40 users—more than double what it had anticipated. And three to four times more models running than originally planned. However, the benefits were worth it, particularly in terms of efficiency, notes Hayes. "With the adoption of the Iguazio platform, deployment times for highly secure, scalable compute resources decreased from days to minutes. We were also able to reduce our costs because we're running in a shared compute environment now, rather than ad hoc compute environments. Our model development time dropped dramatically too. Prior to implementing this solution, model deployment times exceeded 12 months. Thanks to Iguazio and Microsoft Azure, by 2020, initial deployment timelines had been reduced to between 30 and 90 days."

The solution also brings direct visibility into the code that is running during production. "When we create another release and update, we tag it using a standard Python deployment pattern," says Hayes. "And that tag gets released directly to production, so we know which version of the code is running at any given point in time. That's enabled through Azure DevOps, Git, and Azure Pipelines."

From a use case perspective, all the technology used has enabled the development of sophisticated, commercial risk models. "Whether it's a water system that is at risk of colonization by bacteria or the nuances of heat exchanger maintenance, what we've done is build models that can tell us not only if there is an issue with a system, but also whether there's going to be one, and how to best address it when it happens using data science," concludes Hayes.

To find more stories like this, visit the website: **customers.microsoft.com**

Partner

iguazio

Iguazio
https://www.iguazio.com/
Iguazio, a data science and MLOps company, focuses on automating and accelerating the development, deployment, and management of AI applications in production.

Digital transformation to:

- Empower employees
- Engage customers
- Optimize operations
- Transform products