Confessions of a startup data scientist

Unlock the potential of your data to accelerate innovation and grow your startup

Table of contents

Introduction

If the twenty-first century has taught us anything, it's that data is our most valuable resource. Using insight-driven innovation, data is helping to solve huge challenges across almost every industry.

As a startup, your challenge is to make critical decisions quickly and intelligently—while maintaining the speed and agility needed to remain competitive. By assuming an integrated data strategy from day one, your startup can gain tangible and measurable value from your data in ways that accelerate momentum and company growth.

An end-to-end data strategy also creates a foundation for the use of artificial intelligence (AI) and machine learning (ML). These tools give you the speed, automation, intelligence, and predictive capabilities needed to fully leverage your data and unlock innovations that move your startup forward.

Thanks to the elastic, on-demand nature of the cloud—as well as the growing availability of no-code/low-code tools and other fully managed services—virtually any developer or data scientist can take advantage of AI and ML. And virtually any business, including startups, can reap the rewards of these transformative technologies.

With access to fully managed ML services, such as **Amazon SageMaker**, to build, train, and deploy models; a full suite of pretrained AI services; and purpose-built chips for ML workloads and optimizations for ML frameworks, startups use Amazon Web Services (AWS) to deploy models faster, improve productivity, reduce costs, and much more.

More ML happens on AWS than anywhere else, with more than a hundred thousand customers choosing AWS for its powerful compute, high-speed networking, scalable storage options, and breadth and depth of services for any ML project or application.

In this eBook, we will explore how a wide range of startups are creating and instituting end-to-end data strategies with AWS. Decision makers from these startups will reveal how their teams are leveraging flexible, cost-effective technology, frameworks, and infrastructure to streamline their AI and ML processes on AWS. Read on to discover their secrets—and learn how you can unlock the potential of your data to accelerate innovation and grow your startup.

INTRODUCTION

What constitutes an end-to-end data strategy?

Most founders agree on the importance of data, but many startups are challenged with extracting and leveraging meaningful value from their own data. According to a survey from **NewVantage Partners**, just 24 percent of businesses felt they had successfully created a data-driven organization. Motivated by an urgent need to help their teams inform better, faster decisions—so they can fuel ideas and drive higher growth—today's leaders are embracing two fundamental truths:

1. A successful journey to continuous innovation begins with data
2. Becoming a data-driven organization requires an end-to-end data strategy

To inspire your data journey, we've collected stories of early and later-stage startups that have successfully deployed end-to-end data strategies on AWS. Learn how these companies are achieving smarter, faster business outcomes with strategies that maximize the ability to turn information into actionable insights.

AWS is helping startups across the world capture the value of their data by offering flexible, scalable, and secure data analytics ML solutions on modern data infrastructures. By implementing an end-to-end data strategy for your startup, you can meet your data needs now and in the future.

Comprehensive

Integrated

Governed

An end-to-end data strategy on AWS:

- **Comprehensive:** Equipped with the right tools, with the optimal price performance for any user, type of data, and use case
- **Integrated:** The ability to integrate data that is stored and analyzed in different tools and systems to gain a better understanding of your business and predict what will happen
- **Governed:** Governance of all your data to securely give data access when and where your users need it to speed innovation

Startup success stories

STARTUP SUCCESS STORIES

Leadinfo

Helping B2B customers target sales prospects with artificial intelligence-driven insights

Netherlands-based Leadinfo provides software as a service (SaaS) that enables its customers to learn more about who is visiting their company websites. Leadinfo says its goal is to help salespeople "transform anonymous website visitors to customers."

Providing a better customer experience with Amazon Aurora

Data science is at the core of the Leadinfo service, which allows businesses to target sales prospects by analyzing information about website visitors. Leveraging **Amazon Aurora**, a relational database built for the cloud, led to fast improvements for the startup's customers. "They have noticed that our software is much more stable, [and] that the data is always there when they need it to be," said Quirijn Kleppe, head of product and growth at Leadinfo.

Building a foundation for continued business growth

Just three years after the startup was founded, Leadinfo realized it needed to ensure its technology kept up with its ambitions. Leadinfo met with a team of AWS Database Specialists, AWS Solutions Architects, and other AWS experts to develop a new data approach. The parties agreed that Leadinfo needed to unify its data by breaking down silos and consolidating its databases. This enabled Leadinfo to better manage its rapidly expanding customer base, fast-rising IT costs, and growing demands on its developers.

Leadinfo's goals

Using the **Amazon Working Backwards** process, AWS helped Leadinfo determine its goals, develop the right plan to achieve them, and put that plan into action. "It became pretty clear that we needed a system that is easy to use, easy to learn, and easy to work with," Kleppe said.

The answer? Migrating to a database and analytics system built using Amazon Aurora. Leadinfo would later add **Amazon Athena**, a query service for analyzing data, and **AWS Glue**, a serverless data integration service.

"In the end, it really brought our company 10 steps forward," Kleppe said.

Leadinfo

"Before, we had so many different systems and data platforms, it took our developers a long time to learn what works and what doesn't work. Now, we just have one system, Amazon Aurora, and we say, 'Hey, here are the keys to the ML kingdom.'

In general, loading times for a larger customer were one minute before the AWS migration. Now, they wait 0.5 seconds."

Quirijn Kleppe, Head of Product & Growth, Leadinfo

STARTUP SUCCESS STORIES

Leadinfo: Working with AWS

Improving developer efficiency by 50% and rolling out products faster

By migrating and consolidating its four existing databases to Amazon Aurora, Leadinfo improved developer efficiency by 50 percent—and cut the amount of time required to onboard new developers in half. This has helped the startup accelerate its rollout of new service features, allowing its customers to generate sales leads faster and with greater ease.

Decreasing operational costs by 30%

Moving all its customer data to the new system took about a month and a half. After that, "we saw an enormous increase in efficiency and the cost average decreased by 30 percent," Kleppe said. Kleppe expects those costs to drop even more, reaching an eventual reduction of 60 percent.

Creating an exceptional customer experience

Leadinfo has improved experiences for its customers by empowering them to make smarter, faster data-driven decisions. With customer data now consolidated into a single database, Leadinfo was able to add many more filtering options to its software. This lets customers fine-tune information about who is visiting their websites so they can target sales prospects more precisely. Filtered results also load more quickly than before.

Read the full success story ›

STARTUP SUCCESS STORIES

Gourmeat

Boosting efficiency and productivity by 40 % on Amazon Lightsail

Gourmeat, a meat boutique founded in Uruguay and with US operations in Miami, Florida, was using Microsoft Excel spreadsheets to manage isolated inventories manually, but this time-consuming process was hindering its growth. Looking to modernize its data solution, Gourmeat turned to AWS and AWS Select Consulting Partner **Nub8**. Within weeks, the company had a new inventory management system hosted on **Amazon Lightsail**, a simple-to-use virtual private server that offers everything needed to build a web application.

Moving from slow, manual processes to simultaneous data access

Before turning to AWS and Nub8, Gourmeat managed inventory through manual reports from its vendors. Gourmeat's AWS-based integrated inventory application reduces time spent on inventory management, maintains reliable and secure data, and enables key decision makers in the company to have simultaneous data access.

Delivering intuitive visibility with machine learning

For intuitive visualizations of its newly integrated data, Gourmeat uses **Amazon QuickSight**, a ML–powered business intelligence service built for the cloud. The startup's teams gain data visibility from the Amazon QuickSight reporting dashboard, enabling them to collect all the information they need in one place in real time.

Gourmeat's goals

With simultaneous access to reliable inventory data that's updated in real time, Gourmeat's decision makers can make better decisions more quickly, resulting in better-timed shipments to stores and an increased shelf life for meat products. Gourmeat can also adapt its ecommerce site as needed.

"Once we started using the solution built on Amazon Lightsail, inventory management and report creation went from approximately four hours a week to less than 20 minutes.

AWS services are designed to be effective, reliable, secure, and cost efficient; my advice for other small businesses is to go for it."

Lucia Albanell, CEO & Co-Founder, Gourmeat

STARTUP SUCCESS STORIES

Gourmeat: Working with AWS

Modernizing integrated inventory management

Gourmeat took Nub8's recommendation to modernize its testing and production environments on Amazon Lightsail. Over a six-week period ending in October 2020, Nub8 built the customized application, consulting with the Gourmeat team weekly. During one-week sprints, the team could use Amazon Lightsail to test new features and launch them without affecting the production environment.

Fostering breakthrough productivity and improved decision making

The new inventory management system has yielded meaningful benefits for Gourmeat, including increased time optimization for its staff. Time required for inventory management and report creation went from approximately four hours a week to less than 20 minutes, increasing overall productivity by at least 40 percent.

With simultaneous access to reliable inventory data that's updated in real time, Gourmeat's team can make better decisions more quickly, resulting in better-timed shipments to stores and an increased shelf life for its meat products. Gourmeat can also adapt its ecommerce site, as needed.

Unlocking the predictive potential of modern data

Gourmeat's success shows that startups have plenty to gain from modernization on AWS. Going forward, Gourmeat sees many opportunities to unlock the potential of its newly integrated data, including using intelligence to analyze and predict customer behavior.

Read the full success story ›

STARTUP SUCCESS STORIES

Proscia

Accelerating cancer pathology's digital transformation

Proscia is a software startup company that seeks to transform cancer pathology into a data-driven discipline. Its Concentriq digital pathology software and AI applications use data to advance the 150-year-old standard of research and diagnosis.

Modernizing pathology with digital solutions

As workloads expand globally, the pathology field has continued using manual research procedures while struggling to address a shortage of pathologists and laboratory staff. Proscia recognized the opportunity within this industry to innovate and reimagine processes with AI and ML—increasing efficiency, reducing costs, and improving patient outcomes.

"Advanced software—coupled with flexible, scalable, and secure information technology infrastructure—is required to build the future of pathology," said Coleman Stavish, co-founder and chief technology officer of Proscia. By building on AWS, "Proscia . . . delivers solutions that help our customers realize this vision."

A simple and unified solution for patient data

Built and hosted on AWS, Concentriq provides researchers and pathologists with a single interface for reading images, rendering diagnoses, managing imaging data, and collaborating instantly and remotely with experts around the world.

AWS enables greater scalability for Concentriq, helping it support the most complex pathology operations. Concentriq also serves as the foundation for labs looking to apply computational pathology—which analyzes digitized pathology images with AI instead of subjective observation—into their workflows.

Proscia's goals

Proscia's mission is to perfect cancer diagnosis with intelligent software that changes how disease is detected, managed, and treated with technology. Scalable, secure, out-of-the-box solutions and cloud-native services from AWS enable Proscia to accelerate toward this goal—while improving experiences for providers and patients.

PROSCIA

"Proscia's mission is to perfect cancer diagnosis with intelligent software that changes the way that the world practices pathology. AWS provides [us] an efficient solution right out of the box."

Coleman Stavish, Co-Founder & CTO, Proscia

STARTUP SUCCESS STORIES

Proscia: Working with AWS

Scaling resources based on changing demands

To support the highly variable demands of digital pathology, Concentriq runs on **Amazon RDS**, which makes it simple to set up, operate, and scale a relational database in the cloud.

"AWS provides customers with an efficient solution right out of the box," Stavish said.

Consistent availability and world-class reliability

Maintaining application availability for the vast and varying workloads of its customers is critical to Proscia's success. To help achieve this, the startup uses **Amazon Elastic Compute Cloud** (Amazon EC2), which provides secure, resizable compute capacity, and **Amazon EC2 Auto Scaling**, which makes it easy to automatically add or remove Amazon EC2 instances.

Balancing security with ease of access

Leveraging a deep knowledge base of well-documented best practices and the vast experience of trusted AWS Solutions Architects, AWS allows Proscia to handle data with care and integrity—providing a secure and supported solution to its customers.

For one customer that delivers pathology consultation for the US government, Proscia uses **AWS GovCloud** (US), which provides government organizations and their partners with the flexibility to architect secure cloud solutions that align with compliance regimes.

Read the full success story ›

Key business outcomes

With AWS services and solutions for data, your startup can achieve tangible, measurable business results, including:

1. Making better, faster decisions with data
2. Improving customer experience and loyalty
3. Handling customer data with care and integrity
4. Staying ahead of the competition
5. Preparing for the future or the unknown

Learn more ›

STARTUP SUCCESS STORIES

Tangent Works

Putting machine learning modeling into business users' hands

Belgium-based Tangent Works provides businesses with a fast, affordable way to derive value from its data. Its technology helps customers automate the ML modeling process, so they can easily perform complex analysis to drive smarter decision making.

Cutting time spent on infrastructure management

As a young startup company with a small IT team, Tangent Works turned to AWS to provide an efficient way to launch its services, manage its clients' compute and storage needs, and support its rapid growth. Using AWS, the startup's team can focus on product development rather than infrastructure maintenance. Plus, Tangent Works dynamically and cost-effectively scales resources to meet variable customer demand.

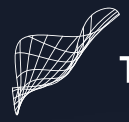
Creating artificial intelligence models in seconds instead of weeks

The main challenge Tangent Works set out to solve stemmed from many companies struggling to realize benefits from the information they hold about their operations and customers. The shortage of data scientists who have the skills to analyze data to get useful insights makes this problem even more difficult.

Tangent Works used AWS to modernize its data infrastructure and launch its services. Today, the startup uses AWS to manage customers' compute and storage needs, support its rapid growth, put powerful ML technology into the hands of business users, and help data scientists improve their productivity by reducing repetitive modeling tasks.

Tangent Works' goals

In competitive markets, making good decisions based on insights derived from ML can be the difference between success and failure. Tangent Works is bringing these advanced analytics capabilities within reach of every organization. Using AWS, Tangent Works helps businesses realize ML benefits—such as improving customer experience and staying ahead of the competition—all without the need for a dedicated data science team.



Tangent Works

"Using AWS, we help businesses realize the benefits of machine learning. And they don't need a dedicated data science team to do it.

Our customers can create models in a few seconds or minutes—this would have taken weeks before. And it's cost-effective because we pay only for the resources we use."

Elke Van Santvliet, Machine Learning Expert, Tangent Works

STARTUP SUCCESS STORIES

Tangent Works: Working with AWS

Optimizing cost savings for customers

Tangent Works helps companies get more value from their time-series data by automating the ML modeling process. The startup's Tangent Information Modeler (TIM) technology—which provides customers with bespoke AI capabilities—delivers the accuracy of manual modeling in a fraction of the time. This means organizations save money on the staff required to create the models and the compute resources needed to run them.

A framework for rapid development

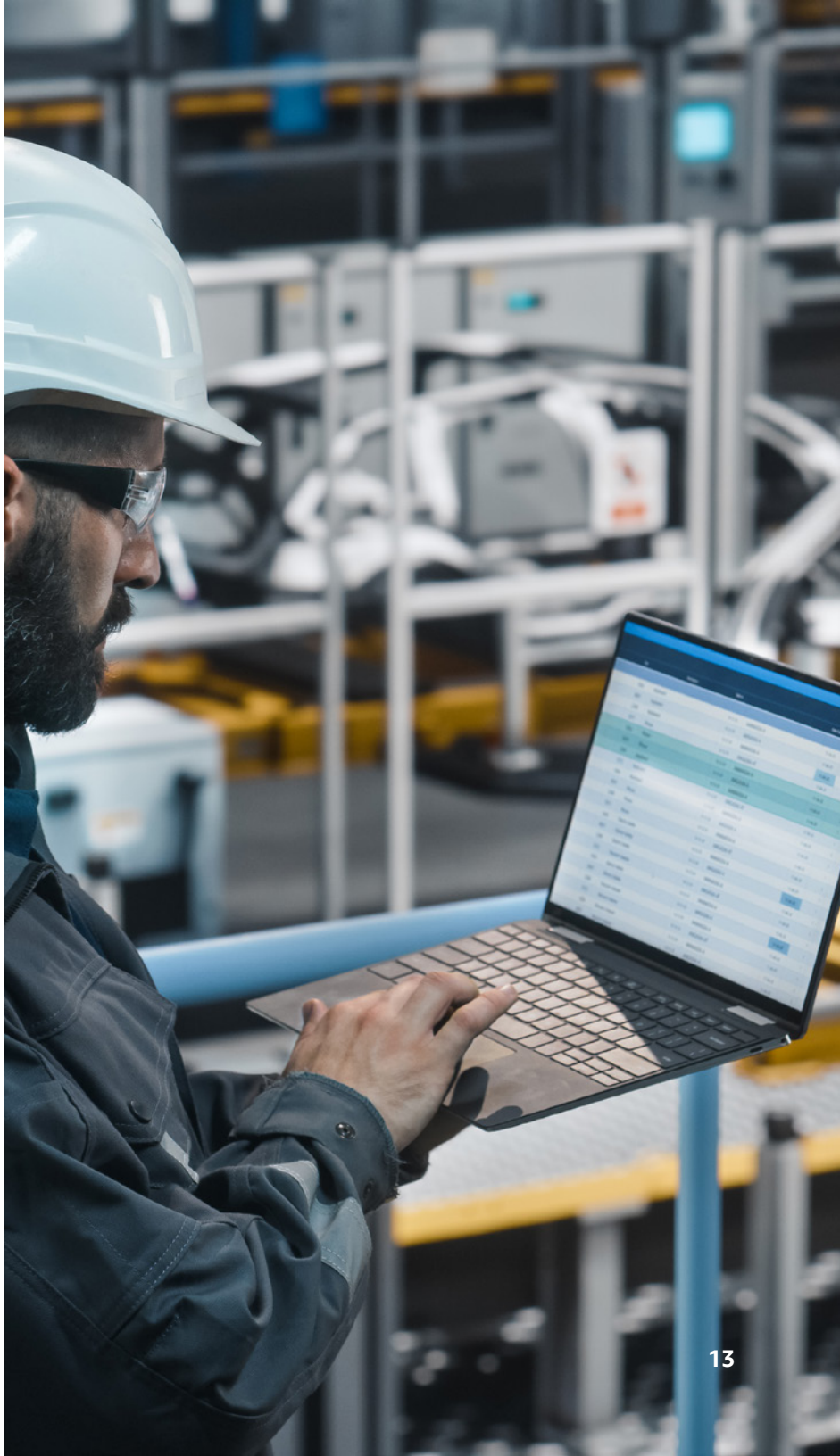
To manage customer workloads, Tangent Works uses **Amazon Elastic Kubernetes Service** (Amazon EKS), a fully managed container service for Kubernetes applications, and **AWS Fargate**, a serverless, pay-as-you-go compute engine that automates server management. Tangent Works also uses **Amazon Relational Database Service** (Amazon RDS) for PostgreSQL, which makes it easy to set up, operate, and scale PostgreSQL deployments in the cloud.

Speeding up customer onboarding

Many of Tangent Works' customers already use AWS, which simplifies onboarding and collaboration. "Getting our clients up and running quickly means we're able to grow rapidly," said Elke Van Santvliet, machine learning expert at Tangent Works.

Offering its services directly to customers on the **AWS Marketplace** has also supported growth. Tangent Works has shortened its sales cycles, winning five enterprise customers in just two months through the marketplace.

Read the full success story ›



AWS gives data scientists the artificial intelligence and machine learning tools they need

Free to try

Gain free, hands-on experience with AWS ML services through our free tier and access the Founders tier, which offers $1,000 for any bootstrapped startup.

Low cost

With AWS, you only pay for the individual features you need, for as long as you use them—without any long-term contracts or complex licensing.

Efficient

AWS services improve team productivity by up to 10 times because we automate all the heavy lifting required to manage and monitor ML infrastructure and security. Our ML services also offer collaboration tools and workflows.

Fast to deploy

You can deploy ML models with a single click, so you can start generating predictions quickly. Your ML applications are deployed onto auto-scaling Amazon ML instances across multiple Availability Zones for high redundancy.

Access to AWS experts

AWS Activate is a free program for startups and early-stage entrepreneurs that offers tools and access to expertise so you can build using AWS best practices.

Proven

Companies around the globe, such as Lyft, Snapchat, Coinbase, and Stripe, have launched and continue to grow on AWS.

Next steps

Startup success is built on an end-to-end data strategy

From democratizing ML modeling and driving sales targets with AI to providing a seamless digital banking experience and revolutionizing cancer pathology methods, data science is sparking a new wave of startup innovation and success.

While the startups in this eBook span a variety of use cases and industries, they all share two essential attributes. First, their successes are built on end-to-end data strategies, which have enabled them to transform their data into insights, innovations, and powerful business outcomes. Second, they're partnering with AWS, leveraging our comprehensive portfolio of AI services, ML infrastructure, and fully managed services, such as **Amazon SageMaker**, to grow their businesses—and emerge as pioneers in the data-driven era.

AWS is proud to be part of this next generation of successful startups, working with a diverse range of innovators to provide scalable and highly reliable solutions across different industries and verticals. With AWS delivering the power to modernize your data infrastructure, unify your data by eliminating silos, and innovate with advanced technologies such as AI and ML, just imagine how far your startup can go.

Learn more about AWS for Data for startups ›