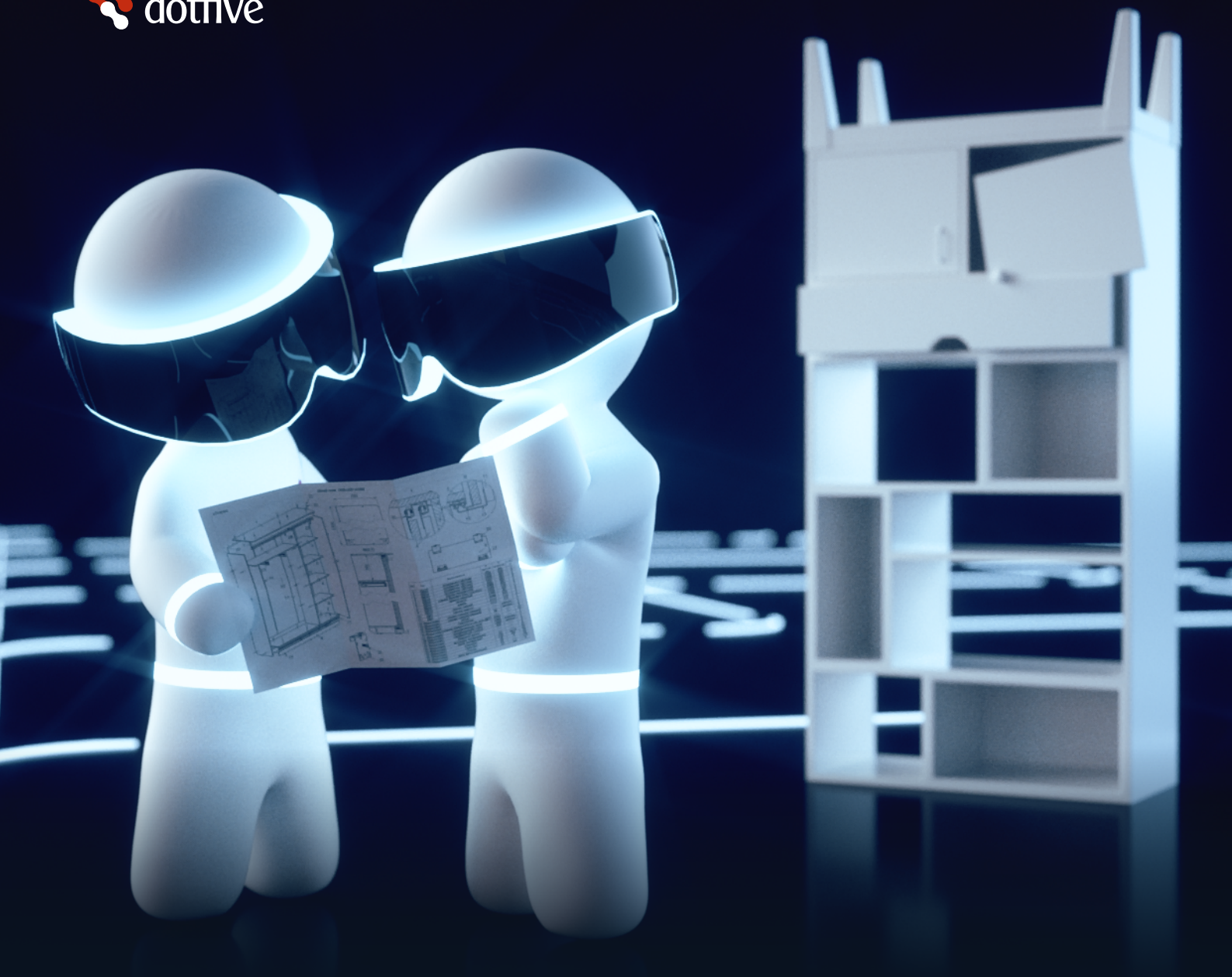
dotfive™



IKEA®

Cutting-edge, extensible data analytics platform for customer intelligence.

Cutting-edge, extensible data analytics platform for customer intelligence.

IKEA is the world's largest furniture retail company. Founded in Sweden in 1943, it currently operates over 400 stores in 52 countries across the globe.

BACKGROUND

Irish software consultancy firm NearForm approached us to partner on this project, due to our expertise in the areas of technology required.

IKEA wanted to increase their customer intelligence capabilities in order to gain additional insight into patterns and behaviours, which would help to improve the accuracy of critical decisions. Being able to analyse all of their collected data efficiently, spot trends and relationships, and model these in different ways for forecasting would enable them to focus time and money into the most beneficial areas across marketing, stock production, and new product development.

For this purpose, they needed a next-gen data analytics platform that would enable their data scientists to develop and deploy various models for analysis and prediction, running across all available data, whilst ensuring the protection of private customer information. Their scientists needed to use a wide range of approaches including machine learning, using a variety of technologies, and were not software developers by trade, so the ability to do this in a consistent and safe manner was critical.

The system needed to be user-friendly, and provide appropriate tooling to help IKEA's data scientists develop their data models, along with a web application to run the models and present the results. IKEA use Google Cloud Platform (GCP), and so the system would need to fully run on this and use appropriate GCP technology options.

This platform would be called the IKEA Model Platform for Analytics (IMPA), and the completed project would provide valuable insight into their sales, product, and catalogue data helping to improve efficiency and how they reach customers.

CHALLENGES

Data scientists are not software engineers. Although very adept at analysis, they are not intimately familiar with the complexities of cloud-based infrastructure and orchestration, or command-line software. The most critical challenge was to develop the system in such a way that it would be easy for all the data scientists to use, regardless of their background. Such a system needs to "disappear" day-to-day, to allow the users to focus on their tasks and data.

Additionally, we had the technical challenge of designing and implementing appropriate architecture to give full freedom in the development of the models, but also ensure safety when running them in the cloud, along with protecting private user data. This required establishing the right balance between guided local development and controlled cloud execution, and then designing and building the toolset and infrastructure to support this pattern.

Above all, the key objective was for all this to function with minimum technical support or maintenance, so automation in an intelligent manner was critical.

Collaboration improved, insights and knowledge gained were spread more rapidly, and decisions could be reached more quickly and with greater confidence.

SOLUTION

We established an approach for data scientists to create and test data models locally, on their own computers, running against example data. This gave them full freedom to go in any direction and use any approach they desired, with no risks. In order to ensure compatibility with the cloud environment we wrote software libraries for them to use, for aspects such as accessing data and reporting results, along with a data model template to use as a starting position. We also wrote command-line (CLI) tools, to allow them to easily create, build, and publish their models to the cloud.

We implemented automated, scalable infrastructure on GCP to run the models without any performance constraints. The privacy of personal data was enforced by allowing the cloud-based models to access the complete data set but only output the results of the analysis. This meant that data scientists could never directly access personal data, but could run their analysis against it without restriction.

Finally, we developed a user-friendly web application to manage the published models and display their output. This allowed multiple scenarios to be created and run against each model, and captured the status and output of the models, including any errors. The results of the analysis could be represented in a variety of forms, including graphs, charts, and data tables, allowing insight to be gained in whichever manner would be most appropriate.

We integrated with Active Directory to govern access to the system, and applied strict validation to all areas to ensure the reliability of the platform. The models run in containers, and the wider system uses serverless architecture, ensuring resilience and scalability.

Python was the primary language for models, but the system also allowed for other languages such as Julia to be run from within models. This makes the system future-proof, and requiring only minimal maintenance.

We are big fans of test-driven development (TDD), and this was an important factor here. We wrote unit tests around application logic and component rendering, to produce coverage analysis reports and set minimum coverage thresholds.

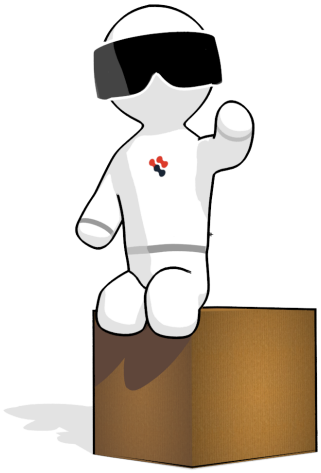
RESULTS

Close communication and use of Agile delivery processes allowed us to rapidly alter direction and adapt to feedback whilst creating the solution, ensuring that the system was delivered on time and met IKEA's requirements perfectly.

We received very positive reactions from data scientists using the platform, who found they were able to easily set up, develop, and publish their models, allowing them to very rapidly proceed through iterations of designing and testing. The ability to try out their models against real data, along with the output from the system, guided them so well they reported a significant decrease in the amount of time spent developing simulations and forecasts, along with an increase in accuracy and confidence.

Moreover, the ability to quickly and easily share these results with peers, managers, and the wider business meant that collaboration improved, insights and knowledge gained were spread more rapidly, and decisions could be reached more quickly and with greater confidence. This had a direct impact onto company costs, as budgets could be repurposed with shorter assessment periods and to newly-identified focus areas.

IKEA's data scientists always had the technical ability to create these models before, but the new system allowed such an increase in speed and safety that they were able to try new approaches and achieve more with fewer people than would otherwise have been needed.



dotfive™

www.dotfive.co.uk
hello@dotfive.co.uk
+44 (0) 8456 808 805
+1 (307) 222 4842

dotfive limited | frontell house, west coker hill, west coker, yeovil, somerset, ba22 9dg
dotfive inc. | 1908 thomas avenue, cheyenne, wyoming, 82001-3527

Technologies used:

PYTHON NODEJS REACTJS FASTIFY GCP

Python | JavaScript | NodeJS | React | Fastify | Google Cloud Platform | Cloud Functions | Kubernetes
StackDriver | Active Directory | Machine learning | Websockets | Storybook | API | CI/CD | CLI

dotfive™

www.dotfive.co.uk
hello@dotfive.co.uk
+44 (0) 8456 808 805
+1 (307) 222 4842

dotfive limited | frontell house, west coker hill, west coker, yeovil, somerset, ba22 9dg
dotfive inc. | 1908 thomas avenue, cheyenne, wyoming, 82001-3527