Democratizing data utilization with automated feature engineering

- Interview with Ryohei Fujimaki
by Shigeki Hayashi -

26 March 2025

Special Edition

Iakyara vol.397

Nomura Research Institute, Ltd.

Executive Summary

As companies drive digital transformation, leveraging data effectively is crucial to achieving their business goals. dotData has developed globally recognized technology that automates feature engineering, making data analytics more accessible and efficient. We spoke with dotData Founder and CEO Ryohei Fujimaki to explore the impact of his company's technology on data analytics.

Ryohei Fujimaki

CEO & Founder, dotData

Ryohei Fujimaki began his career at NEC in 2006 after earning a Ph.D. in machine learning and artificial intelligence from the University of Tokyo. In 2011, he was assigned to NEC Labs America, where he led research on machine learning semi-automation, heterogeneous mixture learning, and related fields. In 2015, he became NEC's youngest-ever Research Fellow, the company's highest-ranking researcher.

In February 2018, he founded dotData in Silicon Valley to commercialize feature engineering automation technology spun out from NEC.

Shigeki Hayashi

Adviser, Nomura Research Institute

Shigeki Hayashi joined Nomura Research Institute (NRI) in 1988. Initially assigned to PMS Development Department; later transferred to Insurance Systems Department. Promoted to General Manager of Project Development Department, Financial Solution Sector and then General Manager of Innovative Systems for Financial Industry Department. Seconded to Nomura Holdings in 2007. Returned to NRI in 2009. Appointed Executive Officer and Deputy Division Manager–Insurance Solution in 2012. Promoted to Division Manager–Insurance Solution Division in 2014. Appointed Senior Corporate Managing Director in 2016. Division Manager–Financial Technology Solution since 2017. Appointed Senior Executive Managing Director in 2021. Current position in April 2023.

How automated feature engineering has changed data analytics

Shigeki Hayashi: dotData has developed groundbreaking technology that automates feature engineering, allowing users to uncover insights beyond human capability. Your innovation has gained global recognition and is widely adopted by leading companies. Would you give us a brief overview of how it works?

Ryohei Fujimaki: A feature is a pattern in data that strongly correlates with a variable of interest. In the context of machine learning and statistics, features are essentially explanatory variables. Feature engineering is the process of identifying, within data, features that have predictive value vis-à-vis the problem you want to solve.

Traditionally, feature extraction has been as much an art as a science, relying heavily on a practitioner's skills, experience, and intuition. However, dotData's automated feature engineering technology enables companies to automatically extract features from their enterprise data and generate insights from them.

Hayashi: So, identifying features is an especially challenging aspect of data analytics?

Fujimaki: Yes. When extracting features, hypotheses and data patterns vary across industries. Understanding these extracted features is not only the essence of data analytics but also its most challenging aspect.

Hayashi: You first developed the technology to automate feature engineering, the key to data analytics, while working at an NEC research lab. Would you share the backstory?

Fujimaki: In 2012, image recognition technology saw a breakthrough with deep learning, which enabled the automation of feature extraction from image data using machine learning neural networks. However, the data analytics work I had been doing for clients in the financial and telecom sectors still relied heavily on the experience and intuition of data scientists. At that time, feature extraction through deep learning was essentially a "black box" and could not be applied to

business use cases where feature selection needed to be justified or explained to customers.

Around 2012, I was assigned to a project commissioned by a New Zealand prepaid wireless carrier that wanted to be able to predict which of its customers were likely to churn. Since prepaid phones are sold in convenience stores and can be used without providing identifying information to the carrier, our task was to predict churn based on available data like phone usage and payment history. A junior colleague and I spent three months manually formulating and coding around 3,000 feature hypotheses to build a model to detect signs of impending churn. During this process, I realized that data utilization could never truly progress if it continued to demand such time-consuming efforts, even from AI and machine learning experts like myself. After NEC expanded its U.S. research team and promoted me to Research Fellow in 2015, I launched a project to explore whether we could automate the feature extraction process.

Hayashi: You eventually founded dotData in Silicon Valley to commercialize the technology developed through this project.

Fujimaki: That's right. In high-tech markets, speed is critical. I believed we needed a structure that allowed us to build a business alongside product development. So, in 2017, I began negotiating a spinout deal with NEC, leading to the founding of dotData the following year.

Hayashi: Does your current technology fully automate the process of formulating and coding the 3,000 hypotheses you mentioned?

Fujimaki: It goes beyond that. Our technology can automatically generate tens of thousands—in some cases up to millions—of hypotheses, efficiently evaluate them and automatically flag the promising ones as candidates.

A platform designed to enable non-technical staff to effectively utilize data

Hayashi: I understand your automated feature engineering technology promotes democratization of data utilization by enabling even laymen without specialized skills to analyze data. How has it been received by customers?

Fujimaki: I've observed that our services are used differently in Japan compared to the U.S. In Japan, there are still relatively few data scientists who build analytical models and few data analysts who apply analytics to business use cases. While Japanese companies are beginning to establish data scientist roles, the shortage of data analysts persists. Given this gap in expertise, I believe our automation technology will play a key role in advancing data democratization in Japan.

In contrast, the U.S. already has a well-established pool of data analytics professionals. Instead of hiring more specialists, U.S. companies are leveraging automation to enhance the productivity of their existing teams. As a result, dotData is helping U.S. practitioners expand their skill sets and maximize efficiency.

Hayashi: You said your technology should drive democratization of data analytics in Japan. Does that mean you're prioritizing services that empower front-line business units rather than specialized data analytics teams?

Fujimaki: Yes, we believe data should ultimately be leveraged in day-to-day operations. Our latest data analytics platform, **dotData Insight**, is designed to help non-technical personnel identify data features that provide valuable insights.

The advent of generative AI was serendipitous for us. GenAI is good at interpreting and assigning meaning to statistical patterns in data. dotData's AI can recognize statistical patterns, however that recognition is still one step removed from interpreting patterns within the customer's business context. **dotData Insight** bridges this gap between recognition and interpretation by ascribing GenAI-generated interpretations and/or hypotheses to data features.

Hayashi: Do you have any plans to go beyond providing data analytics

technologies to companies by offering data analytics as a service?

Fujimaki: No, not currently. One key reason is that U.S. companies—even small ones—strongly prefer to keep data analytics in-house. They are reluctant to use services that require sending their data to a third party and receiving only the processed output in return. Even if we were to offer specialized services or additional support with our products, many U.S. companies would still decline, insisting they only want the software. Their mindset seems to be: *"We're only interested in your software if we can fully utilize it ourselves with minimal external support."*

Another reason is the rapid evolution of data analytics, which is expected to change dramatically over the next three to five years with the widespread adoption of agent-based AI. Unlike traditional workflows, agent-based AI autonomously determines the step-by-step analysis process, functioning like a human expert. Our technology shares a conceptual similarity with agent-based AI in that both use AI to automatically extract features from data and guide the analysis. Our vision is to empower customers to leverage data internally with AI-driven agents, rather than relying on external analytics services.

dotData success stories

Hayashi: You work with some of the world's top-tier companies. Can you share examples of how they have benefited from your services?

Fujimaki: In the U.S., dotData is most commonly used for risk management. In the consumer finance industry, for example, its business model relies on lending more to increase profits, making it crucial to minimize default risk.

One of our consumer finance customers previously spent significant staff hours manually analyzing risks. Now, with dotData, it automatically screens around 50 risk metrics each week, identifying emerging risk patterns. When a risk pattern is detected, the company takes immediate mitigative actions, such as adjusting

pricing dynamically. As a result, it has reduced its default rate by two percentage points over the past two and a half years.

Another example, one I personally find most impressive in Japan, is a trading company that has many customers and deals with an extensive assortment of products. Using dotData, the company analyzes features to determine the optimal timing, customers, and products for sales meetings. An RPA bot then checks sales personnel's calendars and automatically schedules these meetings. In essence, sales representatives receive recommendations on which customers to meet, when to meet them, and which products to discuss—based on historical patterns indicating receptiveness. This automation has led to over 150,000 meetings being scheduled annually, with approximately one-third, or 50,000, actually taking place.

Hayashi: That example should resonate widely, as companies across industries are looking to leverage data more effectively in sales.

Fujimaki: Recently, a Japanese bank approached us to trial **dotData Insight.** They wanted to help their corporate sales representatives gain performance-boosting insights from their Salesforce data. To better support this need, we're considering leveraging **dotData Insight** to create solution templates for Salesforce and other CRM platforms, making them readily usable as out-of-the-box solutions.

Hayashi: Does dotData analyze language data, such as call center conversations?

Fujimaki: Since GenAI excels at processing text, we use it to extract features from textual data. In Japan, analyzing documents like daily sales reports is still relatively uncommon, but some customers have approached us with an interest in text analysis. dotData is not equipped with voice recognition technology to convert call center voice data to text as many such transcription solutions are already available elsewhere. Instead, dotData works with transcribed textual data.

Challenges in data utilization faced by Japanese financial institutions

Hayashi: In recent years, our clients have increasingly been appointing Chief Digital Officers and intensifying their digital transformation efforts. Does dotData offer any data analytics training programs?

Fujimaki: We do but only in Japan. In Corporate America, the concept of a training budget doesn't really exist. U.S. companies basically hire people who already possess the requisite skills and lay off any employees who are no longer needed. In Japan, by contrast, the vast majority of companies that approach us about potentially adopting dotData ask whether their employees would be able to use it effectively. Additionally, many Japanese companies are committed to developing human resources as part of their culture. As a result, we offer training programs alongside our products in Japan.

Our training approach differs from the traditional Japanese model in a few key ways. First, we don't offer basic lectures that simply convey knowledge to passive attendees, as there are plenty of such programs available for free elsewhere. We place top priority on hands-on, experiential learning that involves actually working with data. We prioritize creating a successful, enjoyable experience for first-time learners so they leave feeling that data analytics is both engaging and fun. Second, our goal is for training participants to learn by directly interacting with the technology, rather than simply being taught by an instructor. Historically, data analytics trainings have often featured a consultant explaining how to interpret analytical results or answering the attendees' nuts-and-bolts questions on a certain topic. In contrast, our approach is to have trainees perform tasks like planning an analysis or interpreting results while engaging with our software's built-in GenAI engine. This approach's major advantage is that trainees continue to have access to our software and its support even after the training ends and the instructor leaves

Hayashi: Lastly, do you have any advice for Japanese financial institutions? They hire many elite STEM graduates but their data analytics teams tend to be small, staffed with few employees. Personally, I'd like to see data analytics happening at every branch.

Fujimaki: We've worked with both megabanks and regional banks, and the circumstances vary greatly from one institution to another. Large financial institutions often cite regulatory constraints as a major challenge. Even if they want to pursue data analytics, they're often hindered by strict IT security requirements. Whether these challenges can be addressed through regulatory reform will be an important issue moving forward, as data analytics relies on cloud resources and agile processes.

Meanwhile, regional banks frequently claim they lack qualified staff for data analytics, though I'm not sure that's entirely true. We recently conducted a training session for about 30 employees of a regional bank with the goal of having them develop a data utilization or analytics plan. Before the training, the organizers questioned whether their bank had employees capable of creating such a plan. However, after the session, around 80% of attendees said they found data analytics enjoyable, and many of the plans they created were impressive. Based on this experience, I believe financial institutions already have more employees capable of doing data analytics than they realize. If they put more effort into identifying these employees and gave them the opportunity to work with data openly, they could greatly enhance data utilization within their organizations.

Hayashi: When we visited our financial institution clients alongside the dotData team, they were all inspired by dotData's story. Together with dotData, we at NRI aim to help financial institutions better leverage their data. Thank you for sharing your valuable insights today.

about NRI

Founded in 1965, Nomura Research Institute (NRI) is a leading global provider of system solutions and consulting services with annual sales above $4.9 billion. NRI offers clients holistic support of all aspects of operations from back- to front-office, with NRI's research expertise and innovative solutions as well as understanding of operational challenges faced by financial services firms. The clients include broker-dealers, asset managers, banks and insurance providers. NRI has its offices globally including New York, London, Tokyo, Hong Kong and Singapore, and over 17,400 employees.

For more information, visit https://www.nri.com/en

Inquiries to : Financial Market & Digital Business Research Department
Nomura Research Institute, Ltd.
Otemachi Financial City Grand Cube,
1-9-2 Otemachi, Chiyoda-ku, Tokyo 100-0004, Japan
E-mail : kyara@nri.co.jp

https://www.nri.com/en/knowledge/publication/fis/lakyara/