How DTE Energy is steering the bus through a data-driven transformation in 2024

By Tom Martin

January 18, 2024

Many utilities feel as though their journey with data science is still just beginning. The early days of a data-driven transformation require a lot of work and determination, but they're also exciting as utilities prepare for the improved solutions that data science will bring them.

Among these utilities is DTE Energy. With 2023 now in the history books and a new year officially begun, I sat down with Yong Li, data science manager at DTE, to discuss data science at his company—past, present, and future.

Tom Martin: Can you tell us a little about your journey in data science and that of DTE?

Yong Li: I'd love to. It's a story near and dear to me. I've been working in the data-analysis field for over 30 years—in nuclear physics, in the automotive industry, as a teacher of statistics, and now in the utility sector. At DTE, I started advocating for data science back around 2015 and 2016. I even drafted the job descriptions for the first data scientists and data engineers we hired, working with human resources, customer services,

operations, and the energy forecasting team as part of our effort to bring in this exciting new capability and put it to good use. In early 2019, we hired the first data scientist for my team—I'm in electric distribution operations—and I became the first data science manager at DTE.

I'm giving you some of this history because even with all this data science experience under our belt to date and 30-plus data scientists and data engineers now at the company (supplemented by data science consultants and contractors), I still feel we're in the early days of the data-driven transformation data science can enable. The biggest impact is yet to come.

> **I still feel we're in the early days of the data-driven transformation data science can enable. The biggest impact is yet to come.**

TM: In that case, would you consider 2023 a formative year for data science?

YL: I think our data science ability at DTE is getting better each year, and that includes asking the right questions to build out the most impactful use cases. I'd like to see the data science function evolve even further at DTE so that the majority of our data scientists' time can be focused on the highest-value use cases rather than ad hoc data analysis. That said, we're a utility, which means we don't often have the luxury of time (or money) to solve big systemic problems. I do think data science will be the lever that gets us there, but in the meantime, there's a business to run and customers to serve!

I'm encouraged by our growing ranks of data scientists across our organization. Data science has taken root in electric distribution operations (my department), IT, customer service, and energy forecasting.

TM: What do you see coming in 2024 and beyond?

YL: We have a lot of opportunities ahead of us. Let me give you a couple of examples. One of the big challenges we have in distribution operations that all electric utilities face is power outages, often from storms. To the company and customers, outages are an *emergency*. The operations and communications of power restoration for storm-induced outages are pretty intense.

> **The data science team will also help emergency operation leaders with the analysis so they can make quick and effective decisions if weather conditions change or resource availability changes.**

We have started something we call the Incident Command System (ICS) to manage emergency preparedness and responses. Depending on the urgency and magnitude of storms, multiple storm-function teams are activated to restore customer power as quickly as possible. The ICS leaders have to determine the number

and types of resources, prioritize restoration activities, and communicate to all teams to quickly restore power and improve the customer experience—and do so with maximum operational efficiency.

Today, the data science team can actually develop a storm-forecast model to aid the operations leader to say, "OK, in this kind of weather condition I'll need 700 crews on the first day and 400 crews on the second day because the model is forecasting X number of outages each day, often including types of outages and geographic locations." The data science team will also help emergency operation leaders with the analysis so they can make quick and effective decisions if weather conditions change or resource availability changes.

But then it gets even more complicated. There are multiple crew shifts each day, so the question becomes, How can we optimize how many crews are on each shift given their different cost structure (where a premium is likely being paid for night work)?

I sketch this all out because this is *solvable*. This is the domain of data science. A lot of operations leaders have had to rely on their instincts to make quick estimates on how best to proceed without the benefit of a good data-analytics tool to help make better decisions. My motto is:

> Data science doesn't make decisions; it empowers business leaders to make better decisions.

And the operational benefits I've just described can have an equal benefit on customer service. I'm involved right now in projects to enhance operational efficiency around power outages, reducing outage times, etc. But we also want to extend this efficiency to how we communicate with our customers about outages, letting them know *we know* their power is out before they tell us and providing customers with more-accurate estimate restoration times. There's a ton to do here, and we're making progress.

I'll give you one more example that will resonate across utilities. We're working to harness the power of data science to help optimize our reliability investments. Again, this involves the same kind of complexity and interplay of variables as the storm-outage scenario—likely even more! The fact of the matter is we have an aging infrastructure and need to invest in it heavily over the next few years, but we need to ensure we get the maximum return on that investment, such as increased reliability and customer satisfaction.

> **We're working to harness the power of data science to help optimize our reliability investments.**

To do so, we have to model it. We need to develop a baseline model of our assets and current reliability performance and model out the impact of millions of dollars of investments in new utility poles, substation

transformers, undergrounding wires, automation devices, and vegetation management. All of this is challenging enough, but then our models will also need to consider that new poles won't be needed where lines will eventually be underground and how best to proceed in the meantime.

It's all complex but doable. I'm excited about this journey. These efforts require talented people with strong skills in data science and data engineering. We're making progress, but getting data science up to scale at a large utility while the proverbial bus is still moving means it will take some time.
