MIT Technology Review Insights

In partnership with

kinetica

Businesses can harness sensor and machine data with the generative AI speed layer.

Making sense of sensor data

Key takeaways

1. The widespread integration of IoT sensors across industries creates a steady stream of data about the physical world, enabling real-time insights and reactions.

2. As this sensor and machine data grows in complexity – often including both a time and a location – legacy data infrastructure struggles to manage and contextualize it.

3. A new design pattern – dubbed the "speed layer" – manages the complexity, speed, and costs of analyzing sensor and machine data, creating new data use cases and business value.

Consider a supply chain where delivery vehicles, shipping containers, and individual products are sensor-equipped. Real-time insights enable workers to optimize routes, reduce delays, and efficiently manage inventory. This smart orchestration boosts efficiency, minimizes waste, and lowers costs.

Many industries are rapidly integrating sensors, creating vast data streams that can be leveraged to open profound business possibilities. In energy management, growing use of sensors and drone footage promises to enable efficient energy distribution, lower costs, and reduced environmental impact. In smart cities, sensor networks can enhance urban life by monitoring traffic flow, energy consumption, safety concerns, and waste management.

These aren't glimpses of a distant future, but realities made possible today by the increasingly digitally instrumented world. Internet of Things (IoT) sensors have been rapidly integrated across industries, and now constantly track and measure properties like temperature, pressure, humidity, motion, light levels, signal strength, speed, weather events, inventory, heart rate, and traffic. The information these devices collect – sensor and machine data – provides insight into the real-time status and trends of these physical parameters. This data can then be used to make informed decisions and take action – capabilities that unlock transformative business opportunities, from streamlined supply chains to futuristic smart cities.

"The world is now instrumented in a way that is always on, always tracking, always monitoring, always listening, and always watching."

Amit Vij, President and Co-founder, Kinetica

Evolution of data types

To understand the role that sensor data plays in the broader landscape, it's necessary to examine the evolution of data types over time.

First came **transactional data** – data generated by point of sale (POS) systems, enterprise resource planning (ERP), and supply chain management activities. This early form of big data is characterized by heterogeneous sources and a structured and stable format. It demands robust data integration and governance systems.

The next stage was **interaction data**, characterized by the "three V's" – volume, velocity, and variety. This class of data, created by human-web interactions, such as clicks, social media interactions, and video game play, requires flexible data types and cost-effective storage and processing solutions.

The latest evolution is **observational data**, created through direct monitoring and measurement of real-world events and phenomena. This type of data is primarily machine-generated and characterized by extreme volumes, high velocity, and high perishability. Its readings often include both a spatial location (a longitude and latitude) and a time component. Real-time interpretation and context are critical to making the most of this type of data.

John Rydning, research vice president at IDC, projects that sensor and machine data volumes will soar over the next five years, achieving a greater than 40% compound annual growth rate through 2027. He attributes that not primarily to an increasing number of devices, as IoT devices are already quite prevalent, but rather due to more data being generated by each one as businesses learn to make use of their ability to produce real-time streaming data.

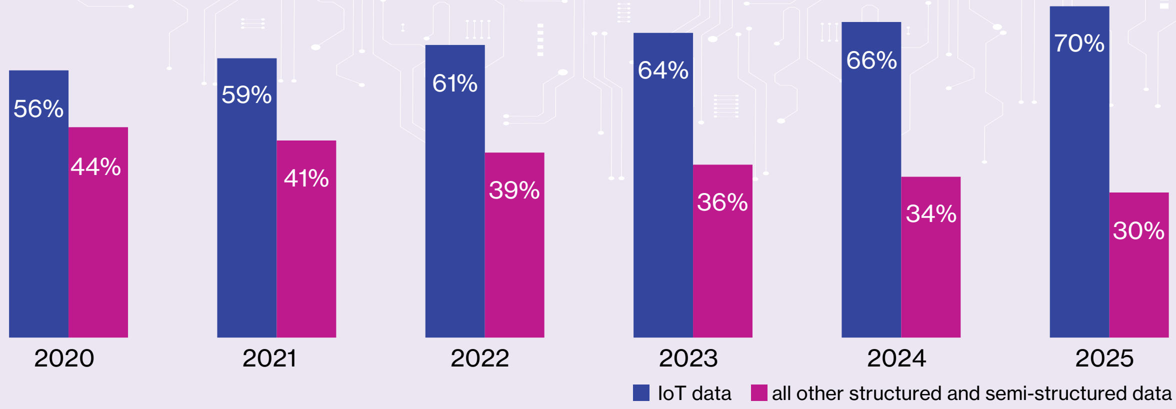
Meanwhile, sensors are growing more interconnected and sophisticated, while the data they generate increasingly includes a location in addition to a timestamp. These spatial and temporal features not only capture data changes over time, but also create intricate maps of how these shifts unfold across locations – facilitating more comprehensive insights and predictions.

But as sensor data grows more complex and voluminous, legacy data infrastructure struggles to keep pace. Continuous readings over time and space captured by sensor devices now require a new set of design patterns to unlock maximum value. While businesses have capitalized on spatial and time-series data independently for over a decade, its true potential is only realized when considered in tandem, in context, and with the capacity for real-time insights.

Understanding sensor and machine data

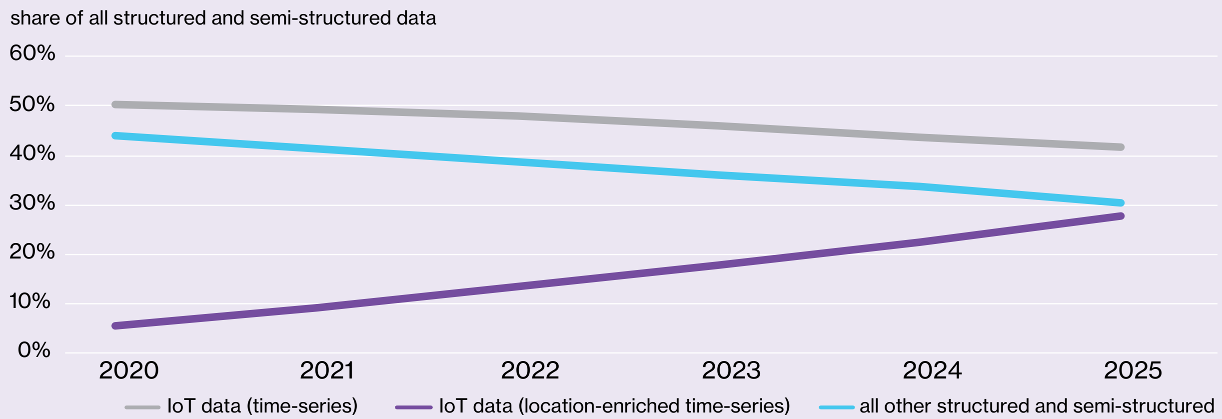
Sensor and machine data, sometimes called observational data, is a distinct class with uniquely complex characteristics. It moves faster than web data, and it's highly structured. Moreover, sensor data is often perishable – it's valuable only if acted on promptly. This often requires decision-making in near real time.

An increasing share of data will come from IoT devices ...



Source: Compiled by MIT Technology Review Insights, based on data from "**Worldwide Global DataSphere and Global StorageSphere Structured and Unstructured Data Forecast, 2023–2027**" and custom research by IDC, 2023

... And more of that data will be location-enriched



Source: Compiled by MIT Technology Review Insights, based on data from "**The rise of spatial thinking**" by Deloitte Insights and custom research by IDC, 2023

Yet, it's precisely these traits that offer exciting potential for real-time insights and responsive actions. "The world is now instrumented in a way that is always on, always tracking, always monitoring, always listening, and always watching," says Amit Vij, president and co-founder of real-time database platform Kinetica. "Historically, most of this data has hit the floor or been looked at in isolation. The software, hardware, and best practices have now caught up, and innovators are developing breakthrough applications based on sensor and machine data."

As data evolves and transforms in character, it calls for a commensurate evolution in design patterns. In the past, the move from structured to semi-structured data led to the advent of data lakes and NoSQL databases. The rise of observational data necessitates additional capabilities, such as a "speed layer" for rapid analysis. Geo-joins and temporal joins – functions that traditional analytical databases are not designed for and therefore struggle to execute at scale and speed – are at the core of these advanced data integrations.

"The old tools can work up to a point. But with the orders of magnitude and types of questions you want to ask of these systems, that's where things start to break down."

DJ Patil, General Partner, GreatPoint Ventures, and former U.S. Chief Data Scientist

It's not the first time we've seen such a shift in the data world. DJ Patil, general partner at GreatPoint Ventures and the former U.S. chief data scientist, recalls the disruption caused by the emergence of web data. Incumbent businesses were ill-equipped to handle the new data volumes. "If you were trying to run a consumer internet company off legacy systems, it worked only up to a degree," he explains. "They were great for keeping e-commerce logs, but less ideal for monitoring the state of check-out or high-speed interactions around chat, sharing, or 'like' buttons. It just wasn't optimized for these kinds of transactions."

Today, Patil sees that pattern repeating itself. For sensor and machine data, "the old tools can work up to a point," he says. "But with the orders of magnitude and the types of questions you want to ask of these systems, that's where things start to break down."

Examining sensor data use cases

The power of sensor and machine data is apparent across sectors. The automotive industry uses it to track vehicle movements and performance, as well as to improve safety and efficiency. Telecommunication companies rely on it to analyze network data and coverage patterns, ensure seamless service, and orchestrate effective infrastructure planning. In agriculture, it helps monitor soil moisture levels, weather conditions, and crop growth patterns. Logistics companies use it to keep tabs on goods in transit, optimize routes, and mitigate supply chain risk.

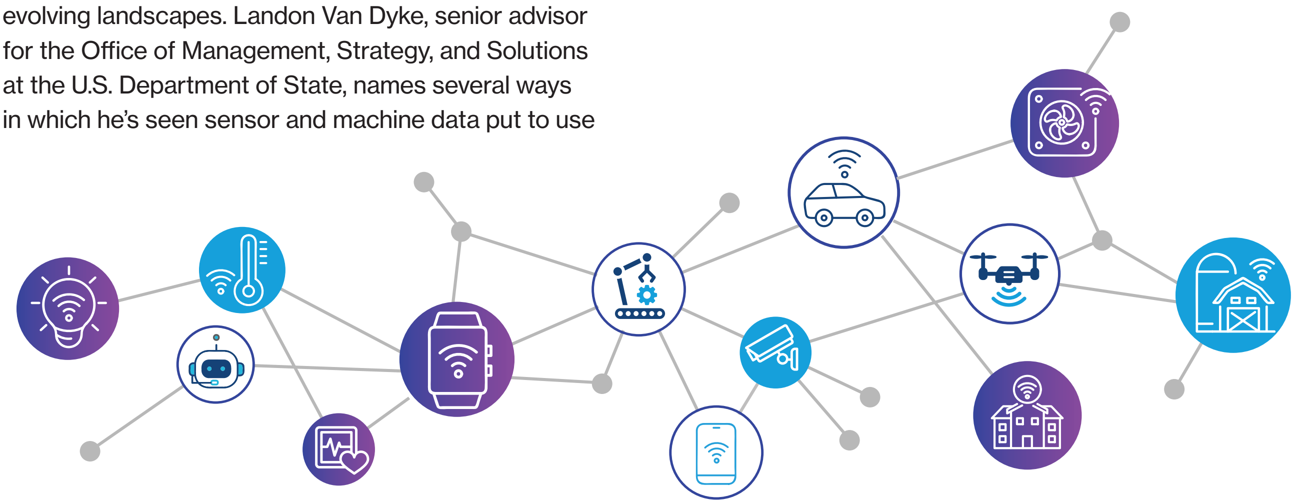
Sensor and machine data is also used in extremely high-stakes applications. Defense agencies, for instance, employ it to enhance situational awareness in ever-evolving landscapes. Landon Van Dyke, senior advisor for the Office of Management, Strategy, and Solutions at the U.S. Department of State, names several ways in which he's seen sensor and machine data put to use by the government: Smart meters can monitor building energy systems on a second-by-second – or even more frequent – basis. The State Department works with NASA to pair data from satellites with ground-based sensors to ensure the accuracy of the air quality index (AQI) at different altitudes. And traffic and route sensor data helps ensure safety for diplomats and other government VIPs, making sure their routes are secure.

Challenges of designing for sensor and machine data

Designing for sensor and machine data presents a unique set of challenges. Executing advanced time-series and spatial analytics on vast data sets can be complicated and slow. The sheer size of these data sets often requires scope reductions, either through omitting potentially valuable data or by implementing more rudimentary analytics to garner any insights at all.

Existing systems are often too limited, struggling to integrate sensor and machine data on both spatial and temporal dimensions, which leads to a lack of crucial context. While many platforms can perform basic operations – shutting down an assembly line when a specific reading is exceeded, for instance – they may find it challenging to execute more sophisticated directives.

For instance, telecommunications companies need to combine diverse data from signal strength monitors, call-drop detectors, and bandwidth-usage sensors with historical network performance data. Then advanced machine learning algorithms have a larger corpus of data to enhance network reliability, and improve the customer experience.



Existing systems are also too slow to enable timely action on opportunities and threats. An agricultural company analyzing frost alerts from weather stations and soil sensors, for instance, has only a short time to react. In logistics, real-time data from shipment trackers and vehicle sensors must be processed quickly to enable agile route changes.

Patil says traditional design patterns can't keep pace with such timely tasks: "If you have a whole bunch of cars interacting with each other, and you take that data and push it to a warehouse and then do a bunch of SQL queries and then push it back, the cars have moved," he says. "Such use cases require a different form of speed."

Moreover, current systems are too expensive. Legacy databases are inefficient for sensor fusion, leading to high compute costs. Vij notes that traditional databases face an exponential problem when new data is added into the mix: Everything has to be recalculated from the root. "We see a lot of use cases where organizations want to fuse mobile phone data with other kinds of data, and that requires a join," he explains. "That's a really computationally expensive problem." The challenge further intensifies with the need for extensive data engineering pipelines to bypass these inefficiencies – not to mention the high usage costs associated with employing scarce data scientists and programmers.

Navigating these nuances requires a strategic approach. To unlock the immense business potential of sensor and machine data, companies need to focus on real-time decisions and context-determining mechanisms, ensuring their new architectures enhance – rather than replace – their current systems. For modern businesses, patching gaps in existing data architectures is not merely advantageous, but crucial for survival in a competitive landscape.

The rise of the speed layer

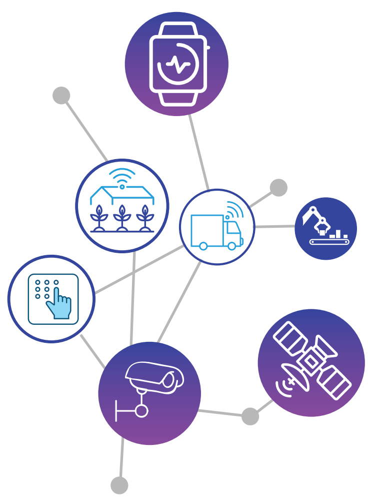
An emerging design pattern that addresses the above challenges is the speed layer. The speed layer is a real-time analytics platform designed to process and analyze streaming data in context at speed and scale. It enables organizations to ingest, integrate, analyze, and visualize large volumes of data as it's generated, allowing for rapid insights and decision-making. A key feature is the ability to join streams with other streams and historical data on temporal and spatial dimensions without sacrificing latency. The speed layer offers machine learning and generative AI integration to harness the power of data as events occur.

Van Dyke says the ability to view relationships between different data sets is particularly helpful when parsing sensor and machine data in high-stakes scenarios. "You can essentially create the relationships you want for the moment; you're not beholden to them for another query down the road. It makes things easier to understand, especially in a temporal review," he says, citing an example use case that might fall under the purview of the State Department: "Let's say we notice a glitch in the system – an embassy shut down in the middle of the night. We can look at the energy logs, and then overlay the harmonics of the energy logs with the energy. Then we can overlay the air quality at the same time. Now, we're seeing a larger picture."

The speed layer also markedly reduces two types of latency: data latency (the delay between data creation and its availability in the database) and query latency (the time taken to obtain query results). While many databases excel in reducing one type of latency, they often fall short on the other. The speed layer balances both, ensuring that even the most complex queries can be processed without delay.

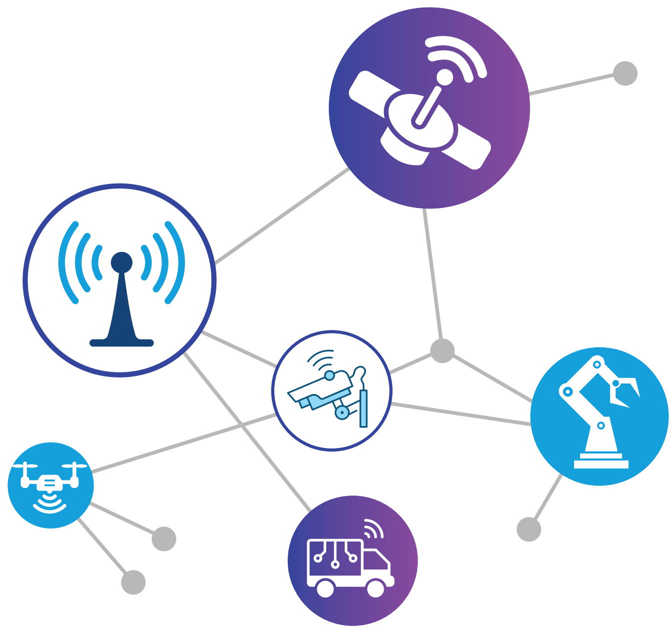
> "You can essentially create the relationships you want for the moment; you're not beholden to them for another query down the road."

Landon Van Dyke, Senior Advisor for the Office of Management, Strategy, and Solutions, U.S. Department of State



Finally, adoption of the speed layer can dramatically lower compute costs and cause a significant decrease in data engineering expenses, due to the reduced need for pipelines, indexing, and summaries. According to Vij, the efficiency gains are substantial. He says, for example, a major bank using Kinetica – a natively vectorized database that incorporates the speed layer – yielded faster results with 16 nodes than with 700 nodes on a competitor platform.

Vectorization, a method of analyzing data through matrix calculations in arrays, is the technology behind this shift. Vectorized databases enable simultaneous processing of time-series, spatial, and graph data, as well as vector search. The business applications are numerous: “A vectorized database lets organizations marry up their streaming inventory data with their fleet data, for instance, and then pair that with consumers in different regions to provide real-time replenishment,” explains Vij.



The impact of generative AI

Generative AI is proving a powerful enhancement to the speed layer. Large language models (LLMs) are capable of translating natural language to SQL, facilitating conversational interactions with complex databases.

This level of sophistication facilitates actionable insights from a sea of complex sensor and machine data. For instance, Kinetica’s recent **integration with LLMs** empowers customers to query databases using simple text prompts. This fosters a greater level of contextual awareness and makes multi-layered data interactions more accessible to more users.

Vij elaborates on how Kinetica’s LLM integration streamlines data query processes for the platform’s users: “Kinetica can compute against the entire data corpus with streaming coming in and crank out results in seconds. By leveraging LLMs like ChatGPT, you can have conversational dialogue and get results on the fly.” With this technology, useful queries – “What are the top 20 threats in the D.C. metropolitan area?” or “Which trucks in Detroit are behind schedule for the day?” for example – are only a few keystrokes away.

Generative AI, noted for its ability to create language, images, and more, can also be applied to numeric data, unveiling concealed patterns. This makes it extremely useful for discovering insights in noisy sensor and machine data. In manufacturing, employing vector similarity search on sensor data can reveal anomalies more effectively than traditional machine learning methods. By directly comparing current readings with historical patterns, this approach swiftly detects deviations, enabling proactive maintenance and process optimization, surpassing the complexity and training demands of traditional techniques.

Patil says that realizing generative AI’s potential will require developing capabilities around high-speed data. “Most of the stuff we see around LLMs today is low-speed data; it’s very static, and it hasn’t been updated,” he says. “We haven’t yet figured out the applications for AI that are going to show up with this higher speed of sensor data. That’s something I think we’re going to see develop over the next 24 months.”

And in the face of such a fast-paced evolution, Patil emphasizes the importance of a cohesive and up-to-date technology stack. “Generative AI is going to iterate unbelievably fast. Even if you’re keeping the core system in place, you have to make sure that you’re upgrading the surrounding systems,” he says. “The small, auxiliary muscle that everyone forgets about, like your core muscles protecting your spine – that’s what’s going to be the most critical.”

"Making sense of sensor data" is an executive briefing paper by MIT Technology Review Insights. We would like to thank all participants as well as the sponsor, Kinetica. MIT Technology Review Insights has collected and reported on all findings contained in this paper independently, regardless of participation or sponsorship. Teresa Elsey was the editor of this report, and Nicola Crepaldi was the publisher.

About MIT Technology Review Insights

MIT Technology Review Insights is the custom publishing division of *MIT Technology Review,* the world's longest-running technology magazine, backed by the world's foremost technology institution—producing live events and research on the leading technology and business challenges of the day. Insights conducts qualitative and quantitative research and analysis in the U.S. and abroad and publishes a wide variety of content, including articles, reports, infographics, videos, and podcasts. And through its growing MIT Technology Review **Global Insights Panel**, Insights has unparalleled access to senior-level executives, innovators, and entrepreneurs worldwide for surveys and in-depth interviews.

From the sponsor

Kinetica is the real-time database platform that leverages generative AI and vectorized processing to let you ask anything of your sensor and machine data. Many of the world's largest companies across the public sector, financial services, telecommunications, energy, health care, retail, automotive, and beyond rely on Kinetica to create new time-series and spatial solutions, including the U.S. Air Force, Citibank, the NBA, Lockheed Martin, the FAA, T-Mobile, Ford, and others. Kinetica is a privately held company, backed by leading global venture capital firms Canvas Ventures, Citi Ventures, GreatPoint Ventures, and Meritech Capital Partners. Kinetica has a rich partner ecosystem, including NVIDIA, AWS, Microsoft, Dell, Tableau, and Oracle. For more information and to try Kinetica, visit kinetica.com.

Illustrations
Illustrations assembled by Tim Huxford with elements from Shutterstock.

MIT Technology Review Insights

www.technologyreview.com

insights@technologyreview.com