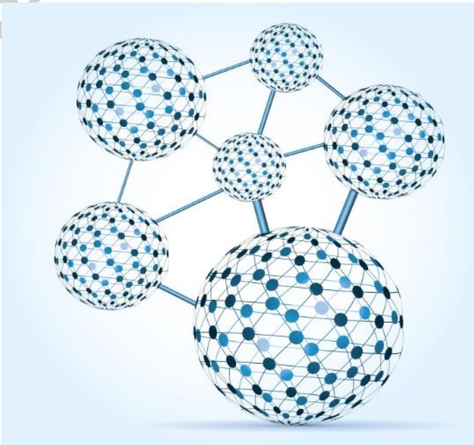
www.pipelinepub.com
Volume 19, Issue 11

AI and Analytics for IoT

By: Ken Figueredo



The rapid adoption rate for generative artificial intelligence (AI) systems, such as Bard and ChatGPT, has astonished industry analysts while galvanizing business interest and public awareness. Grandparents are experimenting with research tasks and creative writing, while young students are using these tools as study aids and to create art. In the business sector, the OECD's latest outlook on employment anticipates significant job impacts with finance and manufacturing industries highlighted as priority areas. Efforts to estimate the annual economic impact of generative AI place it in the $2.6 to $4.4 trillion range. For perspective, this is about 10-15% of the current US GDP.

Below the macro level, a new industry ecosystem is forming. On LinkedIn, freshly minted "experts" publish advice on techniques to make best use of GPT tools. Innovators are fine tuning generative AI for specific applications. To encourage adoption, other businesses are simplifying integration with legacy IT systems. For example, businesses can boost the utility of existing travel, shopping, dining, and other ecommerce services using "plug-ins" of the type offered by ChatGPT. Notwithstanding the excitement around generative AI, users should not lose sight of other techniques in the portfolio of AI and machine learning (ML) tools. Consider the automation of check-out processes in retail stores, for example. Pattern recognition applied to CCTV data and rule-based fuzzy logic are two techniques for monitoring items placed in the bag-loading area. While some customers value human assistance, many others prefer the convenience and speed of automated check-out processes. In addition to a novelty factor, retailers also see benefits from automation and a reduced headcount. To handle exceptions, one agent can supervise multiple stations. This is a typical arrangement for people-centered use cases but not one that necessarily applies to IoT scenarios.

No Human in the Loop

The scope of IoT devices encompasses connected machines in factories, moving vehicles on highways, and sensors located on farms and isolated water pumping stations. Many of these devices have no easy user interface because they are expected to work with minimal human intervention. Unlike a smart home appliance, rebooting a connected sensor by physically switching it off and then on is not always an option. Instead, these are well established industry procedures for managing and maintaining IoT end points using over the air (OTA) software update techniques.

Such differences between people-centered and constrained IoT devices have profound implications for AI. Take the example of a sensor sending a stream of real-time readings to a decision-making application. If the readings trigger an alarm, who is to say that the underlying cause is not due to a malfunctioning sensor rather than an undesirable change in the process or property being monitored? System protection for this type of operating scenario relies on built-in fault-detection capabilities that are packaged as intelligent tools for developers and system operators to use. To understand these capabilities, it helps to begin by defining a simple framework for how AI functions in an IoT system.

Data to Decisions Framework

Interrelationships between AI and IoT depend critically on data flows and their impact on the architecture of an IoT solution. The flow begins with data sourcing from connected devices and sensor end points, before passing through signal processing and machine learning processes. The latter processes extract key features from IoT data streams and label or annotate them as knowledge-like objects. As an example, a data stream might be described as “normal” while it stays within a range of pre-set values. Depending on the use case, many other labels are possible. Another description, for example, might apply to a machine’s resonant frequency pattern, just like the labels used to describe chords that are played on musical instruments.

Feature extraction and labeling are techniques for creating a digital twin of the system or process being monitored. This is an intermediate stage in the data to decision-making flow. The next stage involves the application of rules-based AI for reasoning or decision-making purposes. For example, if the resonant frequency profile changes significantly over time, this might indicate accelerated machine wear. In the case of a musical instrument, this corresponds to the instrument going out of tune. The AI processing stage might go further and trigger interventions automatically. These could range from a simple action, such as adding lubrication or scheduling a maintenance check, to stopping the machine abruptly to prevent a potentially catastrophic failure.

The sequence from data sourcing to AI interventions highlights the importance of a framework for complementary AI and IoT technologies and data provenance management. Device and data management are its foundations.

Enabling AI for IoT

In IoT technology stack terms, device management covers several functions related to network connectivity, identity management, registration, and discovery. These are common service functions, designed for interoperability and reuse, so that developers can use them time and again.

Data management is a second foundational capability because of its impact on analytical workloads and decision-making quality. Data scientists spend 80% of their time on data cleaning and exploratory analysis. There is clear justification to automate these activities by equipping developers with a tailored set of service functions. These include tools to manipulate training data sets. They also cover configurable information models for common IoT items such as actuators, meters, and switches. These are examples of AI/ML common service functions.

The interdependency between IoT and AI/ML shows up in the way that device and data management capabilities affect the AI and ML layers higher up the IoT stack. An example is the “registration” function that manages the identities on which IoT data-supplier and AI/ML data-consumer interactions are built. The registration function is also a source of data provenance information. If a pattern-recognition or causal inferencing AI application produces a machinery shut-down alert intermittently, the plant operator will want to check on the factors contributing to that diagnosis. After querying the IoT data feed, the operator might then query data about the sensor, the manufacturer's details, and information about the most recent calibration settings. These other details—a form of meta data for the sensor readings—can help an operator to figure out whether the sensor, rather than the manufacturing process, might be suspect. This is possible through tightly coupling the AI/ML layers of the technology stack with the device and data management layers.

Users will have higher performance expectations as the use of AI/ML and IoT techniques becomes more pervasive. System designers should therefore plan on data provenance checking as well as transparency of decision-making rising in importance. These developments will call for new and generalized functionality. Rather than customize solutions on a case-by-case basis, users and developers stand to gain from a common "AI4IoT" framework and standardization.

Pre-standardization Insights Into AI for IoT

Anticipating the need for new capabilities, ETSI (the European standardization agency) published a pre-standardization study on the topic of AI for IoT systems. The study explored a variety of IoT use cases employing AI and ML capabilities. Some of these applied to vertical applications for use in communications networks, road traffic management, smart parking, and to recognize patterns in data collected from social media devices. To complement these, horizontal use cases explored the application of knowledge graphs in smart buildings as well as approaches to support trustworthy and verifiable AI. This analysis aimed to study architectural implications for IoT systems. It also aimed to find AI/ML common service functions that could be candidates for standardization.

Three proof of concept applications, using IoT platforms based on the oneM2M standard, provided feasibility insights to guide future work. They also confirmed the existence of a set of recurring requirements that solution designers should be prepared to address, ideally via standardized and reusable tools.

While there will be many ways to use AI in IoT systems, most applications share common elements. If these commonalities are not exploited, the study shows that application developers risk becoming overwhelmed by the burden of maintaining and integrating a large variety of AI-based modules, data models, and data sets. There will be benefits to drawing on a library of reusable AI-enhanced components matched to IoT enabling counterparts. These would reside in a service layer between AI/ML applications and IoT devices in the form of common and callable AI-as-a-service functions. These ideas are being carried forward as part of oneM2M's standardization roadmap.