

AI at the Edge:  
The next frontier  
of the Internet  
of Things



# Table of contents

Today's case for AI at the edge	3
The power of actionable data	6
Cloud power isn't always enough	8
Take it to the edge	11
Computing capabilities require reliable power	12
Edge computing produces better results, but only if it's the right choice	15
What's next on the journey to AI	16





## Today's case for AI at the edge

Once the basis of story lines for science fiction movies, Artificial Intelligence (AI) now has practical applications that are changing the way businesses get things done. Smart developers are exploring ways to combine AI with the internet of things (IoT) to help companies in a variety of industries benefit from the data generated by connected devices. The ultimate goal is to improve production capabilities, gain efficiencies and reduce operating costs by delving into real-time data from multiple points to produce actionable insights.

The market is ready. According to Statista, by 2020, projected spending by discrete manufacturing, transportation/logistics and utilities industries for IoT platforms, systems and services will exceed \$40B for each category. In 2018, Forrester Research found that 48% of companies in North America had already invested in AI/machine learning as part of their digital strategies. The combination of AI and IoT boosts the capabilities of the devices that run these applications to contribute to improved business processes. When developing solutions, it is important to consider what infrastructure best supports the ability of AI to drive valid, real-time decision making.



BY 2018

**48%** OF NORTH AMERICAN COMPANIES had already invested in AI/machine learning

While cloud solutions get all the headlines right now, latency matters, and waiting on data centers miles away to power instant decision-making on site is not feasible for many applications.

Edge computing is the answer in many cases. Emerging advancements in the hardware and modules needed to push progress in AI at the edge fuel the possibilities. Edge devices and gateways-to-edge devices are now more powerful and enable the local collection, storage and analysis of data without waiting for value to be derived from the cloud and then passed back to the device. By combining AI and edge computing, IoT solutions are more powerful because the latency issues associated with cloud computing are eliminated.

<sup>1</sup><https://www.forbes.com/sites/louiscolombus/2017/12/10/2017-roundup-of-internet-of-things-forecasts/#330568ce1480>

<sup>2</sup><https://www.unisys.com/Style%20Library/Unisys/usi2018/reports/UnisysNetworkMicrosegmentationOpportunitySnapshot.pdf>

## The power of actionable data

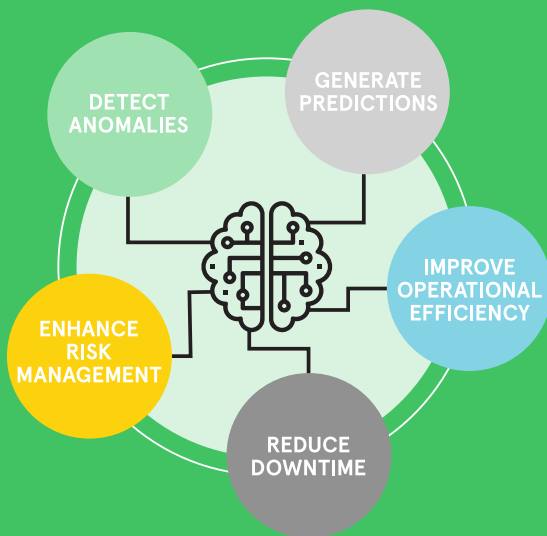
Data collected from one IoT device has limited value on its own. Plus, according to Forrester Research, 60% to 73% of data in an enterprise that could be used for analysis goes untapped. Real value is derived by combining data sets from multiple devices to uncover patterns that can be used to predict future performance.

For example, a manufacturing plant may have 10 IoT devices or sensors monitoring processes on various production machines. Rather than looking at the data from only one device or each device independently, analyzing data from across the site can provide a holistic view of what is happening at that location.

AI is the enabling technology that processes large amounts of data and recognizes patterns in the data. Using powerful algorithms, AI adjusts to new inputs and makes decisions based on what it has learned over time to provide automated, accurate feedback to guide decision-making. It's the tool that adds value to all the data collected by IoT devices. AI takes advantage of the aggregation of big data to do more than just discover what happened in the past. Rather, AI produces analyses about ways processes can be more efficient and predicts what could happen based on multiple scenarios.

**ACCORDING TO  
FORRESTER RESEARCH,  
60% TO 73% OF DATA  
IN AN ENTERPRISE  
THAT COULD BE USED  
FOR ANALYSIS GOES  
UNTAPPED.**

The collective use of the data is what drives AI's ability to conduct machine learning, an essential component of the technology. Machine learning uses algorithms with the computational ability to "learn" from data and adapt on their own over time based on additional inputs. That way, with limited human intervention, AI and machine learning turn the analysis of the data into actionable insights to help:



Where that analysis takes place in the infrastructure can have a significant impact on the value of the insights and timing of the resulting actions.



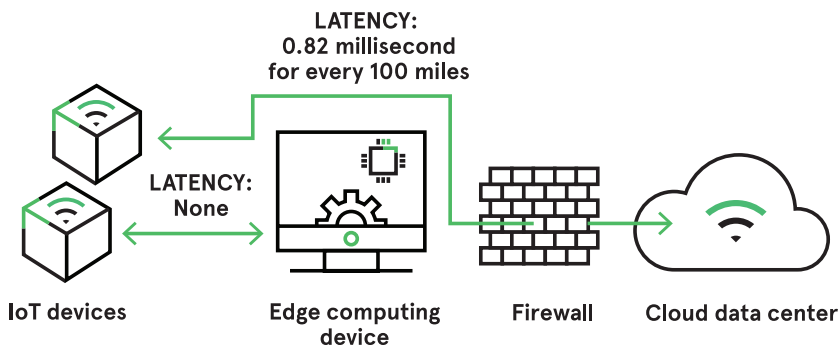
<sup>3</sup> <https://go.forrester.com/blogs/hadoop-is-datas-darling-for-a-reason/>

# Cloud power isn't always enough

The trend continues for companies to host more data in the public cloud with Amazon Web Services (AWS), Microsoft Azure, Google Cloud Platform and a variety of other service providers. Currently, most data from connected devices in the IoT ecosystem is collected and transmitted to the cloud for processing and analysis. The computing power in the cloud's data center is where data is aggregated, and AI-powered decisions are made.

While this approach has proven reliable, the amount of time it takes to complete the transfer of data to and from the cloud introduces latency issues that can affect real-time decision making. The farther away a cloud data center is geographically located, more latency is introduced. For every 100 miles data travels, it loses speed of roughly 0.82 millisecond.

Cloud computing is agile but cannot sustain the growing demands of heavy workload IoT applications for industries such as healthcare, manufacturing and transportation, to name a few.



As the number and practicality of AI-powered IoT solutions continues to grow, cloud computing will remain an important part of the IoT ecosystem for complex and historical data processing. But, to power real-time decision making, edge computing is a better and faster approach for many applications that provides computing and analytics capabilities to end devices.

For example, ABB is a leader in industrial motion, power grids, electrification products and industrial transport infrastructure. One of its biggest challenges is predicting plant emissions in advance and taking proactive measures before violations occur.

**ABB's Predictive Emission Monitoring System (PEMS)** powered by AI uses an empirical model to predict emission concentrations based on processed data. The team successfully implemented PEMS as part of a comprehensive Environmental Management System in one of the largest gas processing plants in the world. However, PEMS (inferential analyzer), cannot measure emissions directly. So, the system uses an empirical model to predict emission concentrations based on data like fuel flow, load, operating pressure and ambient air temperature. These kinds of intelligence analytics applications are driving AI adoption—and forcing companies to consider whether cloud-based or edge-based analytics are right for their particular business case.

**FOR EVERY 100 MILES DATA TRAVELS, IT LOSES SPEED OF ROUGHLY 0.82 MILLISECOND.**



<sup>4</sup> <https://www.techwalla.com/articles/network-latency-milliseconds-per-mile>.



## Take it to the edge

Operational technology is the hardware and software stack that can both detect and control changes in physical devices throughout the enterprise. AI-enabled IoT devices take the concept of OT to the next level by combining data inputs to drive intelligent real-time decision.

Edge computing shifts the collection, storage and analysis of data collected from IoT devices for real-time decisions away from the cloud. Where AI in the cloud is managed by one big processing center, AI at the edge is more like a hive architecture of small, yet powerful computing devices that work together to drive local data-informed decision making.



### REAL-TIME RESPONSE

By eliminating the need to transfer data to the cloud for processing, latency issues are eliminated that can affect the validity of real-time decision making. Real-time response is critical for applications such as manufacturing, medical imaging and autonomous driving where AI-based decisions the real-time performance of IoT machines.



### BETTER RELIABILITY OF OPERATIONS

Decisions about processes, machine health and operations are all made locally with fewer concerns about connectivity. Real-time information can ensure uninterrupted processes by preventing asset breakdown or sudden failures. The parameters that identify when predictive maintenance is warranted are integrated into the IoT solution.



### ENHANCED SECURITY

Edge computing keeps sensitive data in the local IT ecosystem, avoiding the security issues of the public cloud. AI-enabled solutions can also detect anomalies at the edge of the network if cyber attackers attempt to gain access to the network through IoT devices and quickly implement mitigation tactics. Risk analysis determines all the possible points of entry for attackers and proactive plans are created to mitigate security issues if they arise.



### LOWER COMPUTING COSTS

Because edge computing aggregates data locally versus sending it to the cloud, there is a reduced need for expensive bandwidth for connectivity.

### USE CASE: PREDICTIVE MAINTENANCE

AI-enabled IoT solutions can predict when maintenance is needed for corporate assets before breakdowns interrupt daily activities with expensive corrective repairs. Based on inputs from the IoT devices, the AI algorithms can predict the useful lifespan of components and ideal planned maintenance schedules.

This approach can also reduce costs. For example, currently service technicians might visit office buildings on a pre-set schedule to inspect elevators. Instead, IoT monitoring devices can instead provide 24x7 scrutiny of how the elevators perform and send alerts when maintenance is required and predict when parts need replacement. This model reduces the number of trips service technicians need to make.

#### Predictive maintenance benefits include:

- Reduced operational risk
- Increased rate of return on assets
- Better controls on maintenance costs
- Lower component/parts inventory costs



## Computing capabilities require reliable power

AI-enabled solutions at the edge of the network need reliable computing capabilities to guide real-time decision making. Currently, most IoT devices are powered by batteries which cannot reliably support the ongoing collection and analysis of data. Over time, companies will need to transition to new AI-enabled IoT devices that get their power from the wired infrastructure.

As new edge devices are introduced that have more processing power, the need for reliable power increases. For example, in healthcare, IoT sensors can collect patient data 24x7. This capability is especially useful for remote care scenarios that enable real-time monitoring of patients with chronic illnesses to assess their need for proactive treatments. It is critical that IoT gateways that transmit the data have a dependable power source.

Soon, new chipsets will be available that require more power to enable their computational capabilities. Intel recently announced a new neuromorphic chip that incorporates “digital circuits that mimic the brain’s basic operation” to enable efficient machine learning.

---

### USE CASE: SUPPLY CHAIN OPTIMIZATION

Supply chain challenges can drain limited cash reserves and result in too much or not enough stock on hand. AI-enabled IoT solutions can help optimize inventory by predicting customer demand and guiding decisions about demand-driven forecasting. People using smart apps on mobile devices can contribute to the real-time flow of actionable data more efficiently with edge computing.

#### Supply chain optimization benefits include:

- Reduced working capital requirements
- Streamlines supply chain processes
- Better control of just-in-time delivery of materials

<sup>6</sup> <https://newsroom.intel.com/news/intel-advances-quantum-neuromorphic-computing-research/>

# Edge computing produces better results, but only if it's the right choice

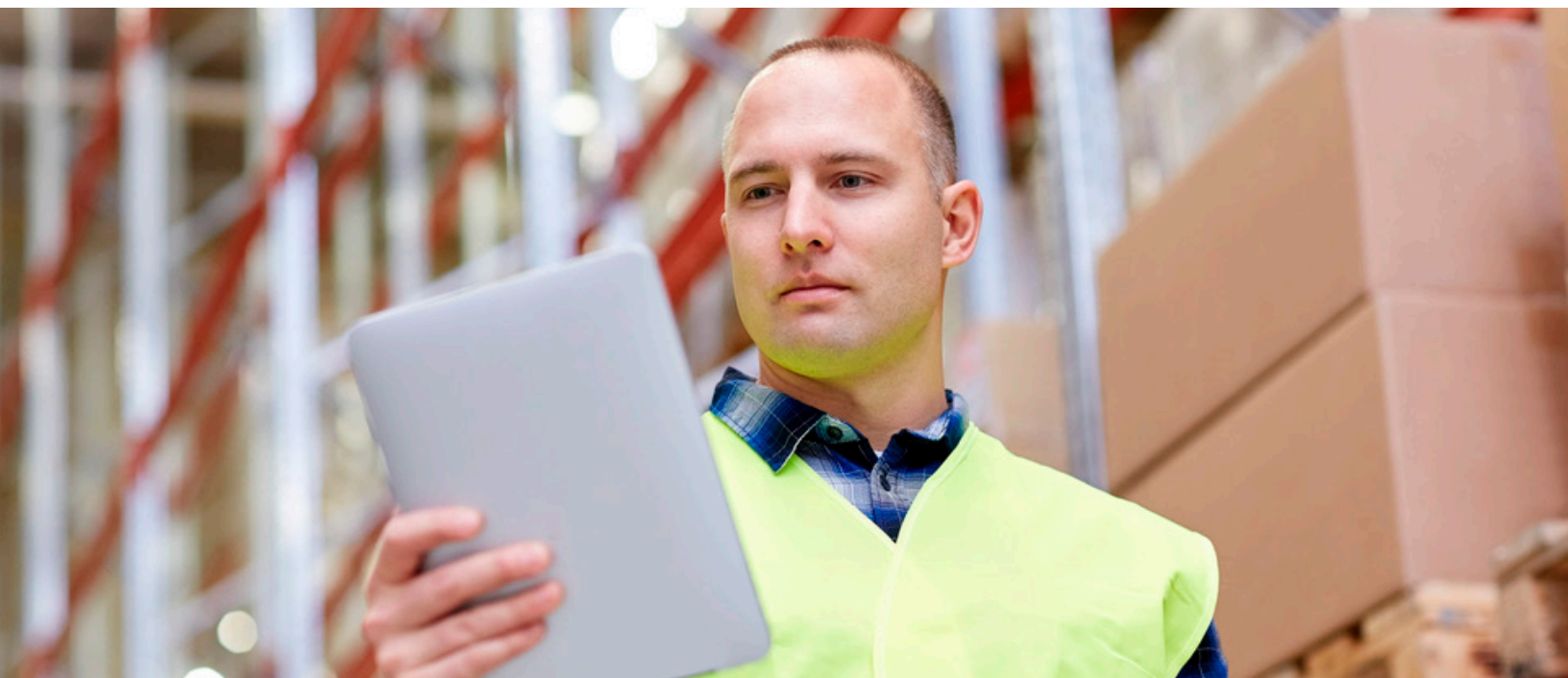
The benefits AI can provide are compelling. Organizations that are eager to make digital transformations may jump right in seeking vendors to help them understand implementation options.

As more IoT devices are deployed, the need for AI-enabled solutions supported by edge computing grows exponentially. It will no longer be feasible to rely on the cloud to process and analyze data to drive real-time decision making. The capabilities of edge computing to handle AI algorithms and machine learning locally without the inherent latency of cloud computing will provide insights to more efficiently drive operations and increase productivity.

By understanding the value of AI at the edge, business decision makers can ask informed questions about what infrastructure best support their operational goals. Is a cloud-based model the right choice, or is edge computing the best path forward?

## CHECKLIST: CLOUD OR EDGE?

	AI SUPPORTED BY THE CLOUD	AI AT THE EDGE
<b>Collected data</b>	<input type="checkbox"/> Does not drive real-time decisions	<input type="checkbox"/> Need to collect and analyze data to drive real-time decisions
<b>Latency</b>	<input type="checkbox"/> Slower transmission of collected data is acceptable	<input type="checkbox"/> Near instantaneous data transmission is optimal
<b>Bandwidth</b>	<input type="checkbox"/> Can support bandwidth requirements to send large amounts of data to the cloud for analysis	<input type="checkbox"/> Local network support for data collection is preferred
<b>Autonomy</b>	<input type="checkbox"/> Reliance on the cloud for connectivity is acceptable	<input type="checkbox"/> Network downtime affects productivity of connected devices





# What's next on the journey to AI

Where a company is in its exploration of the benefits of AI greatly affects what steps it should take next to understand how to leverage the technology. Enterprises generally fall into one of four categories in their understanding of how to apply AI.

## LEVELS OF AI ADOPTION

**PASSIVES:** Has not adopted AI and has a limited understanding of the technology

**EXPERIMENTERS:** Are piloting or adopting AI without a deep understanding of the technology

**INVESTIGATORS:** Understands AI but are not deploying it beyond the pilot stage

**PIONEERS:** Understands and has adopted AI

## STAGES OF AI UNDERSTANDING AND APPLICATION

Honestly evaluating what stage a company is in in its understanding and application of AI is useful to help determining what to do next. Companies that fall in the Passives, Experimenters and Investigators categories do well when working with a partner that can help them kick-start their AI journey by following this five-step process.



## A PARTNER YOU CAN TRUST

The right strategic partner creates a roadmap that addresses four key areas:

### Technology:

Determines what platform is ideal for the use case and what infrastructure best supports the solution

### Data:

Determines what data is needed and creates the model that drives decision making

### People:

Provides needed ongoing support and provides training for internal staff

### Compliance:

Recommends solutions that are compliant in the regions in which they operate



Essentially, the goal of the strategic partner like Avnet is to eliminate the unknowns to quickly create a path to operationalize AI in a way that best supports business objectives—whether that involves implementing AI in the cloud, in the edge, or in both over time.



NEED A GUIDE FOR THE NEXT STEPS OF YOUR AI JOURNEY?  
GO TO [AVNET.COM/AI](https://www.avnet.com/ai) TO LEARN MORE.