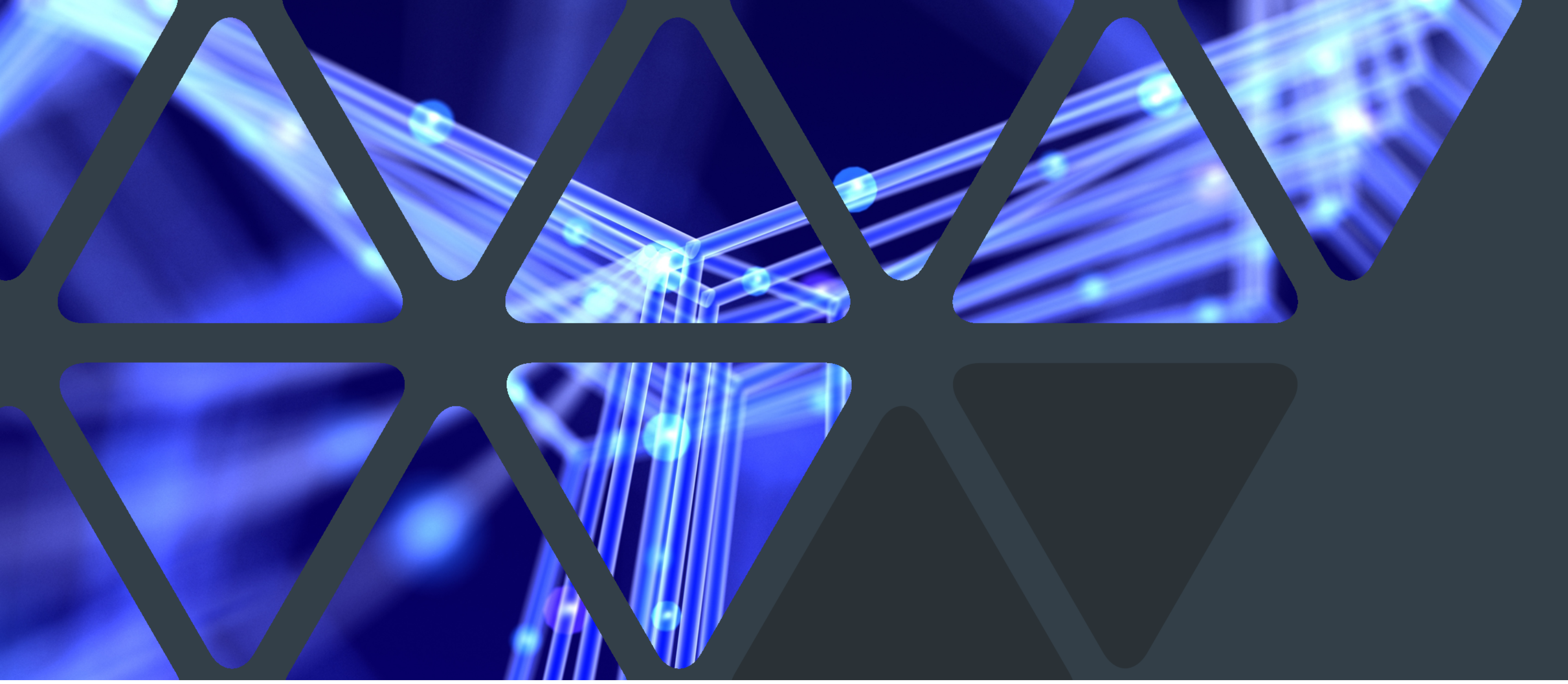
# FEDERATED LEARNING: DECENTRALIZED AI

Federated Learning enables Artificial Intelligence (AI) to leverage data models created in decentralized environments.  Aggregating these local models from various sources allows Internet of Things (IoT) to learn from each other.

The benefits realized from Federated Learning include reduced bandwidth consumption, localized personalization of the model, and increased data security, to name a few. Federated Learning has become possible recently due to more powerful devices at the edge, ranging from mobile phones and IoT devices to small form factor PCs and servers.

# TABLE OF CONTENTS



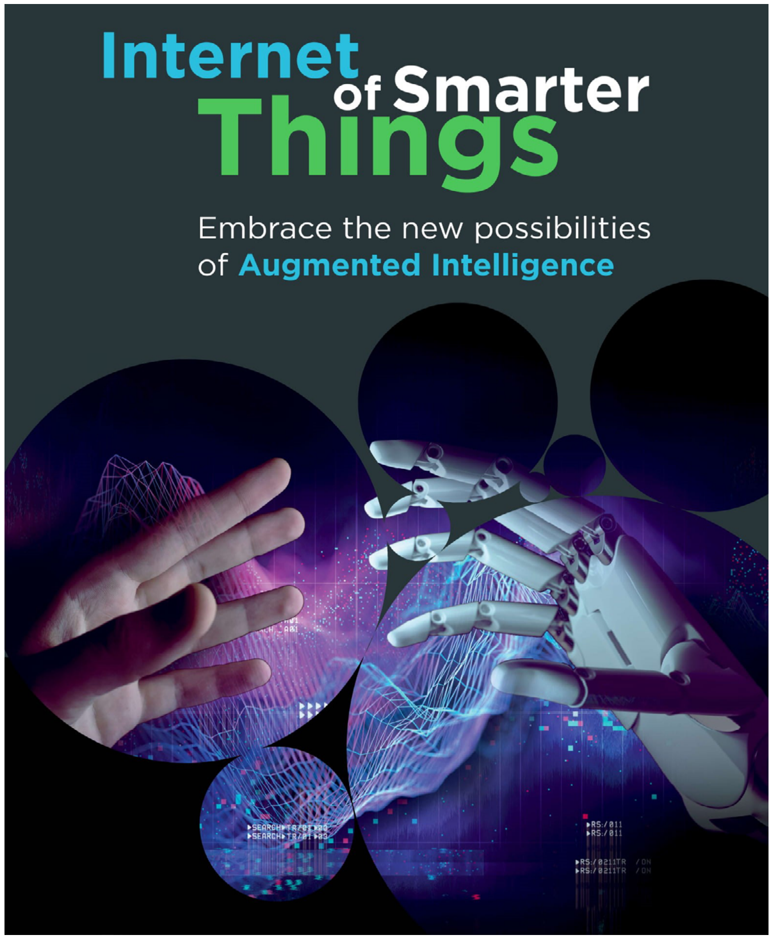
## INTRODUCTION

Edge devices such as IoT sensors and mobile phones have been limited by compute power to drive insights from the significant amount of data they produce. With increasing compute power at the edge, we are seeing a trend of moving inferencing workloads from the data center to the edge devices. This is also contributing to the development of new architectures that enable edge training, such as Federated Learning.

Although Federated Learning has very promising benefits, it is still an early stage architecture. CB Insights has listed Federated Learning in the Experimental quadrant of its NExTT Trend framework (CB Insights Artificial Intelligent Trends 2019). Much of the experimentation and production grade use-cases that have seen promising results focus on mobile devices, since there is already a significant user base and immediate benefits that can be realized from Federated Learning.

Although, some of our experimentation has focused on mobile device use-cases, we also believe there is a significant opportunity in industrial and business use-cases, specifically in processing IoT data due to the bandwidth and latency limitations of most industrial private networks.

byteLAKE and Lenovo have partnered together on various AI initiatives from computer vision to intelligent manufacturing. The goal is to combine both companies' diverse set of data science and AI expertise with deep industry knowledge and provide real world solutions for industry segments, including healthcare, life science, energy and manufacturing.




We recently demonstrated a holistic approach to AI by promoting the "Internet of Smarter Things" concept. It builds on a principle that the AI workloads (partially or in full) are moved from the Cloud or Data Center directly onto the devices. This became possible mainly due to the progress in dedicated hardware acceleration developments. During the ISC (International Supercomputing Conference) event we jointly showcased computer vision algorithms (objects recognition / video analytics) working on small, embedded devices as well as running in client-server based architectures, utilizing the best of both worlds: client devices and powerful resources from Lenovo's AI-Ready ThinkSystem servers.

During the AI Summit event, which took place in San Francisco in September 2018, byteLAKE and Lenovo presented their Federated Learning solution for IoT. They showcased how IoT (devices / sensors) can learn from each other and how local models (information) can get aggregated from various sources.

# 1 / WHAT IS FEDERATED LEARNING?

Federated Learning is new approach in Machine Learning for distributed environments. The typical use case for Federated Learning consists of a significant number of devices and/or computers located outside of the data center.

Each edge device collects data and trains a component of the Machine Learning model. The model is consolidated at a central location, typically in the cloud or data center.

Primary Benefits of Federated Learning:

### 1. Reduction of Network Bandwidth

In a typical machine learning architecture, all data is transferred to a central data lake. This can cause significant network constraints when transferring vast amounts of IoT Data or image/video files. This transfer can be challenging for devices operating outside of the traditional networks, oftentimes connected through slow mobile connections. With Federated Learning, only model parameters are transferred to and from the device and central server.

### 2. Model Personalization

With Federated Learning, you can control the weightage of the local model vs. the centralized model to balance between accuracy and personalization. This differs from traditional machine learning architectures, where a single model may be deployed on all devices.

### 3. Data Privacy

Federated Learning allows each party to keep their data private. It allows multiple users or companies to build and benefit from one shared model without handling the whole data to one location.  Transferring sensitive data from one location to the other creates a vulnerability that can be exploited, even within your internal network.

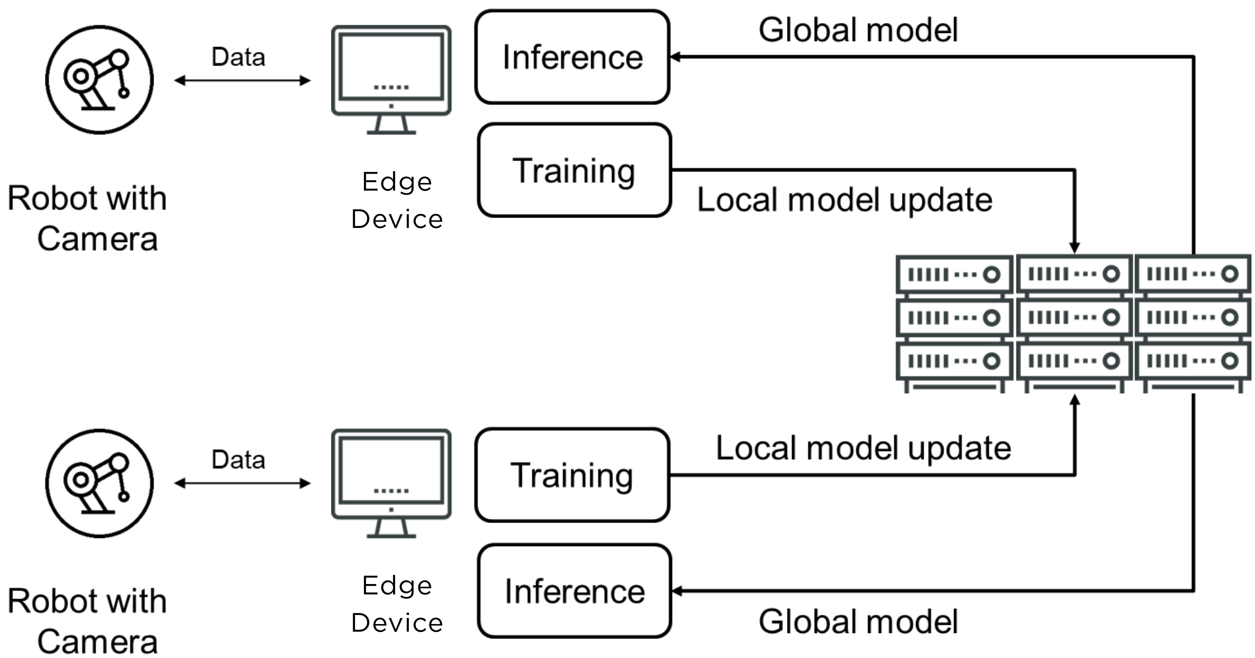
### 4. Cost Savings

By reducing the demand on the network, organizations may realize cost savings by utilizing existing network infrastructure. Network and backbone upgrades can be costly and time consuming. If organizations can realize the benefit of emerging technology without significant upgrades to existing system, it can greatly reduce the barrier to entry for adopting technology such as AI.

### 5. Improved Accuracy

Training at the edge can result in faster model development time and higher predictive accuracy than conventional batch processed training models provide. By training more often, we learn faster, and we therefore can adapt our predictions to the data in a more timely manner.

**How federated learning works:**

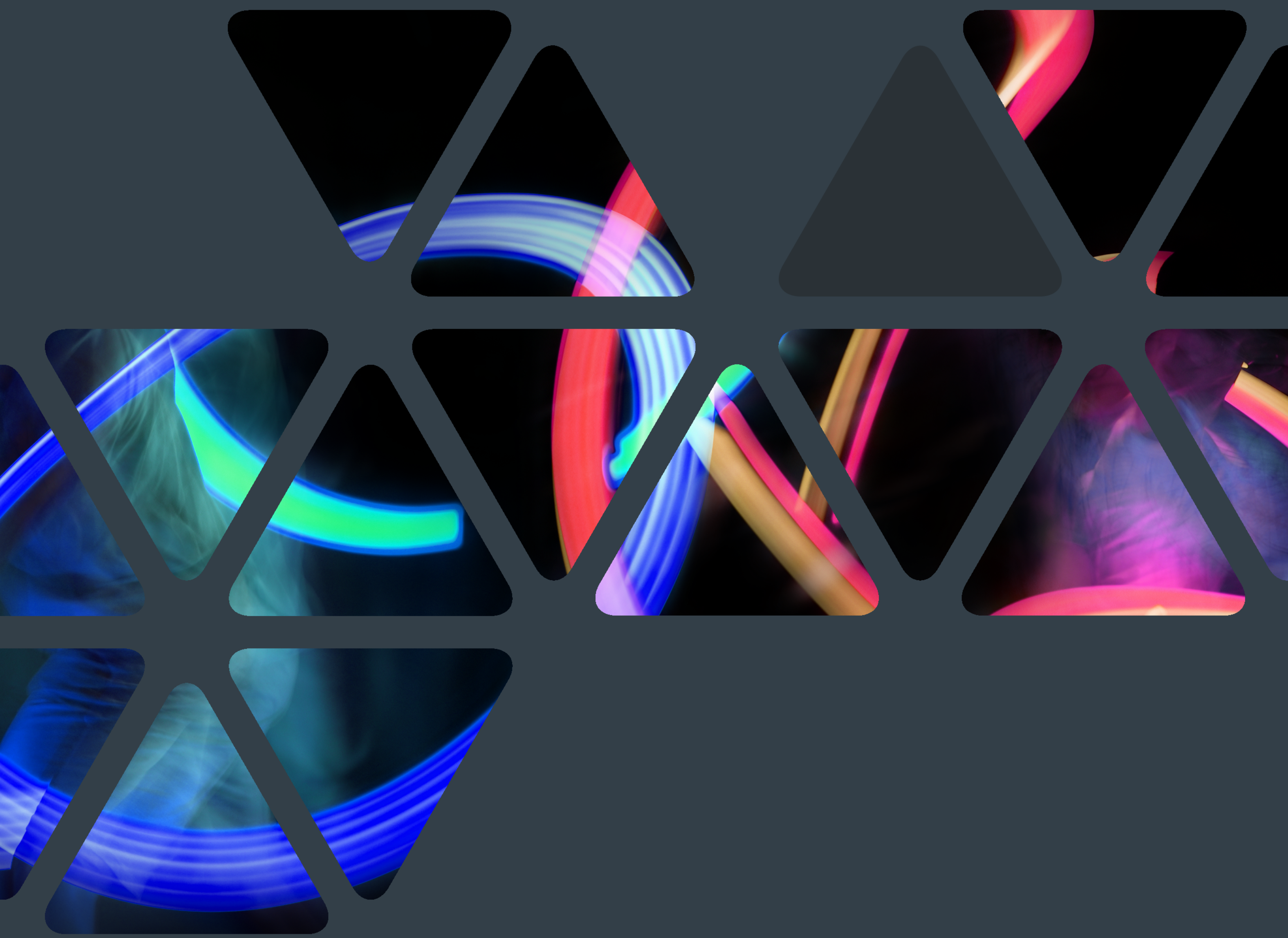



In this example, each "Robot with Camera" is taking pictures for an AI/ML application to make a prediction. The prediction is done on the Lenovo ThinkStation P330 Tiny Workstation, the edge device in this example. The edge device stores the picture and will use new data to update the model.

Only the updated model is transferred to a central location. All of the data remains local. This means this approach is using limited network capacity and only during low traffic hours.

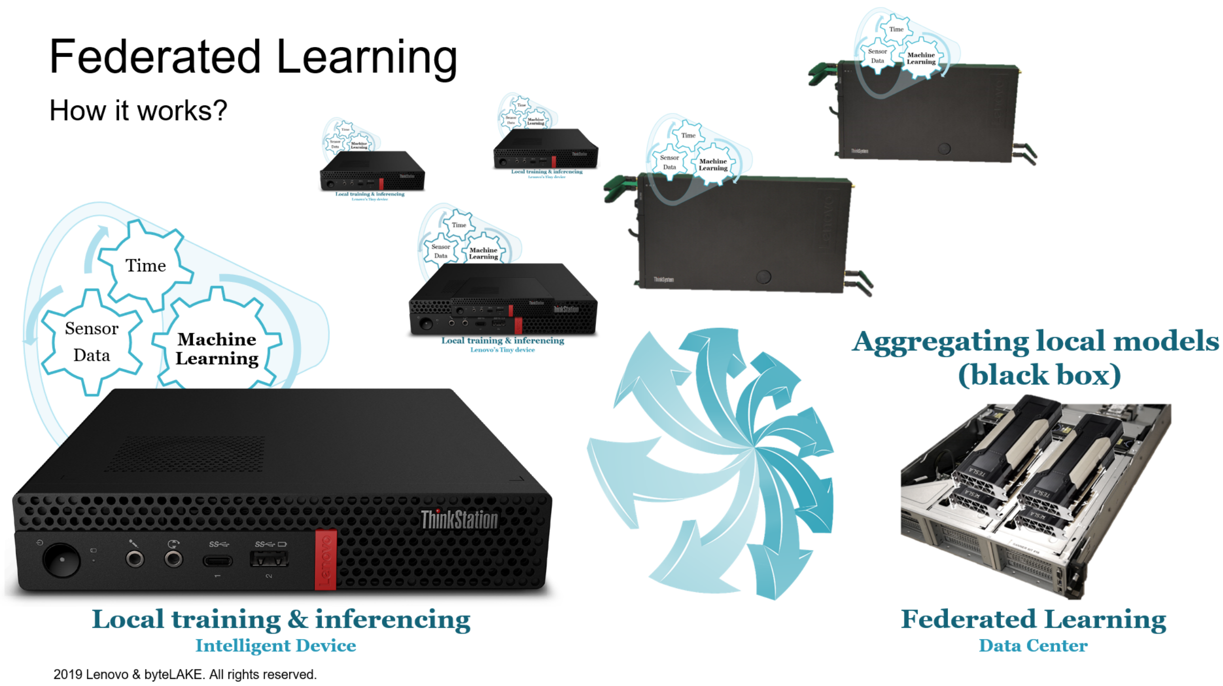
The central server consumes the locally updated models from all devices and combines them into one global model. The global model is abstracted from the actual data, protecting privacy and maintaining anonymity. Each device downloads the new global model and bases the next round of predictions on it, so all devices can benefit.

# 2 / FEDERATED LEARNING FOR MANUFACTURING USE-CASES



Manufacturing is one of the leading verticals adopting AI with promising results in predictive maintenance use-cases. The key constraints impacting manufacturing and Industrial IoT (IIoT) AI adoption are bandwidth, latency & security. The benefits of Federated Learning could significantly lower the barrier to entry for AI adoption in manufacturing and other predominate IoT use-cases.

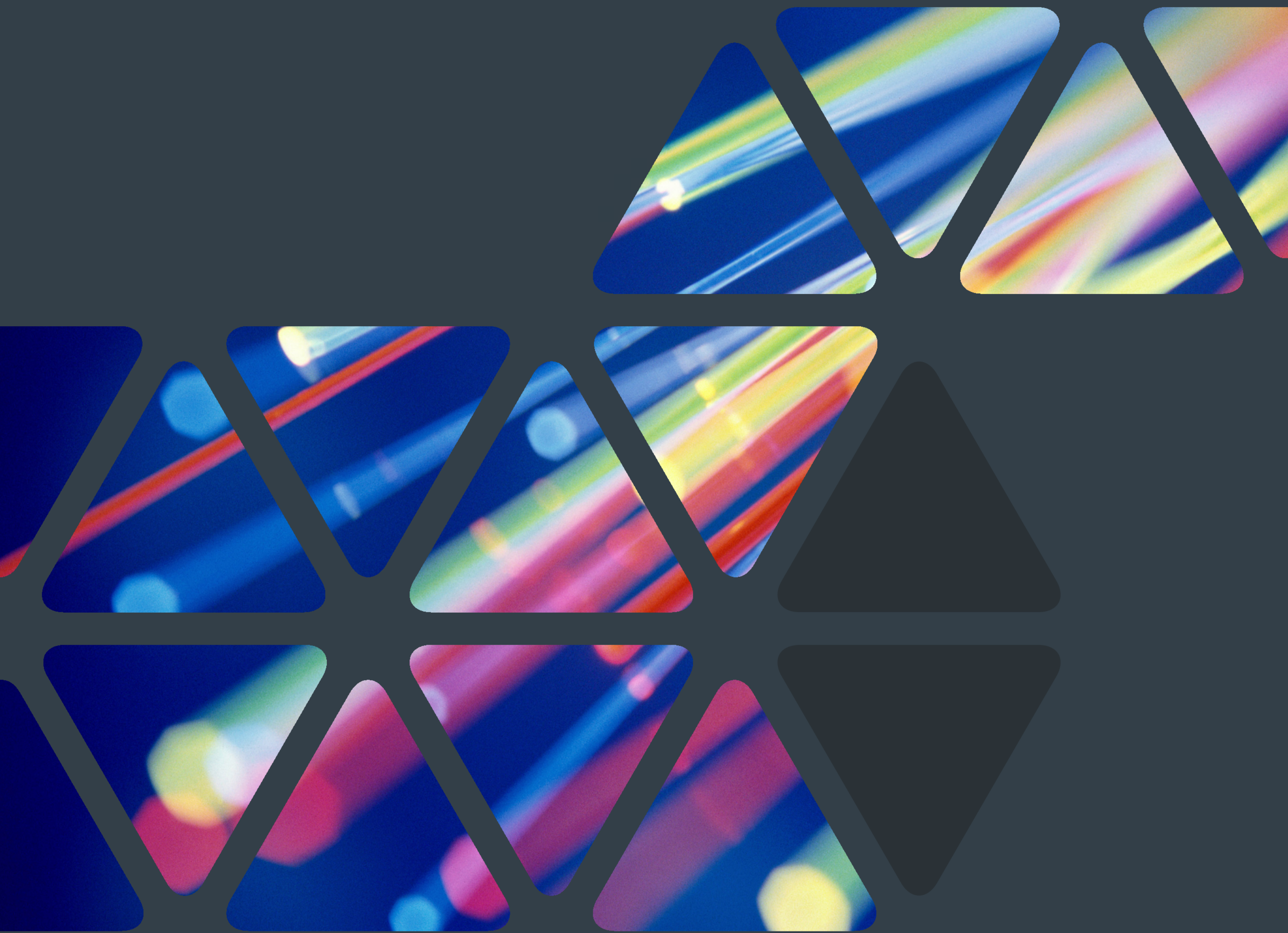
In our proof-of-concept, each edge device independently works to learn the patterns in its data and creates a local model to make a prediction. The local models are aggregated by a central server into a global model that is then distributed back out to the edge devices. Consequently, all edge devices can benefit from the information gathered and processed from various angles and in different scenarios. Therefore, we can say that the Federated Learning enables IoT (devices / sensors) to learn from each other.




We utilized the Lenovo ThinkStation P330 Tiny for our edge device for this PoC due to the small form factor, rugged mil spec testing and the higher performance processing capabilities for CPU or GPU based training. The edge device runs local training, inference and transfers the model (not data) back to a central server in the data center. The AI-Ready Lenovo ThinkSystem™ SR670 server aggregates all of the edge models into one Federated Model, transfers the Federated Model back to the edge devices.

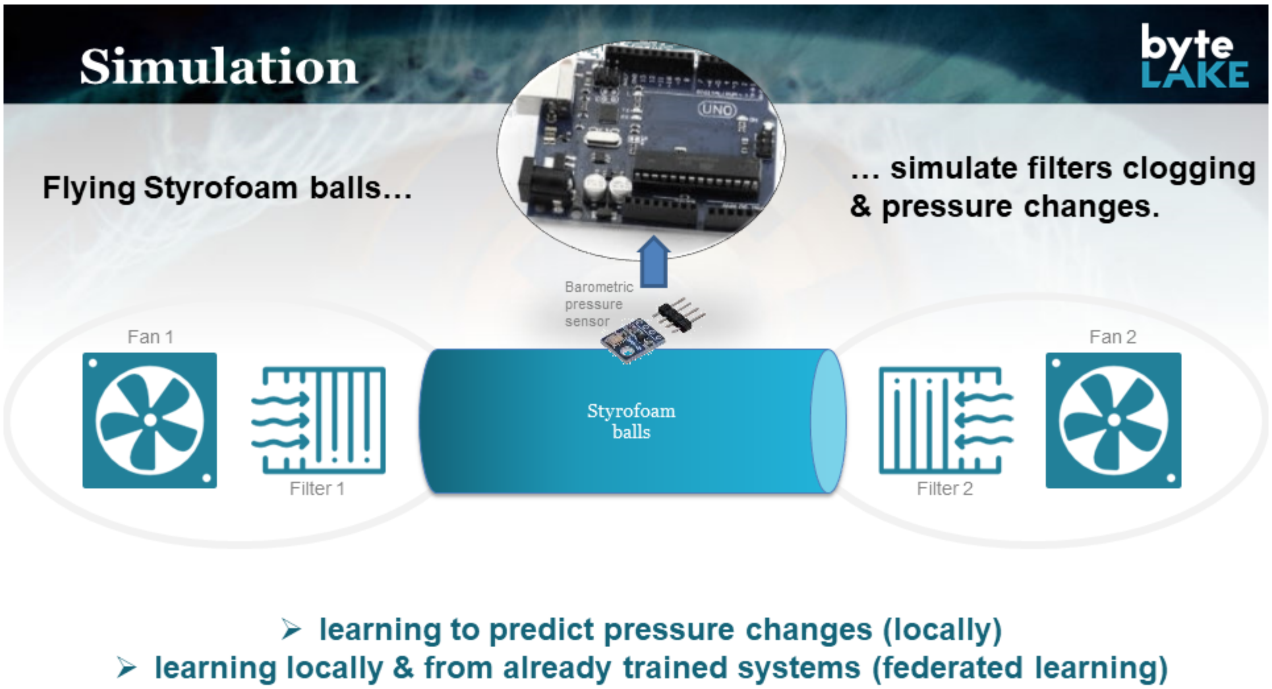
**The benefits of Federated Learning could significantly lower the barrier to entry for AI adoption in manufacturing and other predominate IoT use-cases.**

# 3 / IoT SIMULATION



For the purpose of the demonstration, we utilized a piece of hardware, simulating industrial processes in factories. Although, Federated Learning can be applied to many use-cases and industries for predicting trends hidden in the data, we picked manufacturing to start with.

To simulate an industrial process, we built a physical device with a transparent tube and fans attached to both of its ends. Styrofoam pellets placed inside the pipe are sequentially pushed by the air from one side to another, simulating the filters on pipe's ends getting gradually clogged. Also, the pressure inside the pipe is continuously changing due to the fans speed and Styrofoam balls creating a blockage (among other factors). As one picture is worth 1000 words, have a look at the graphics below to better visualize the device.




**Live Demonstration**

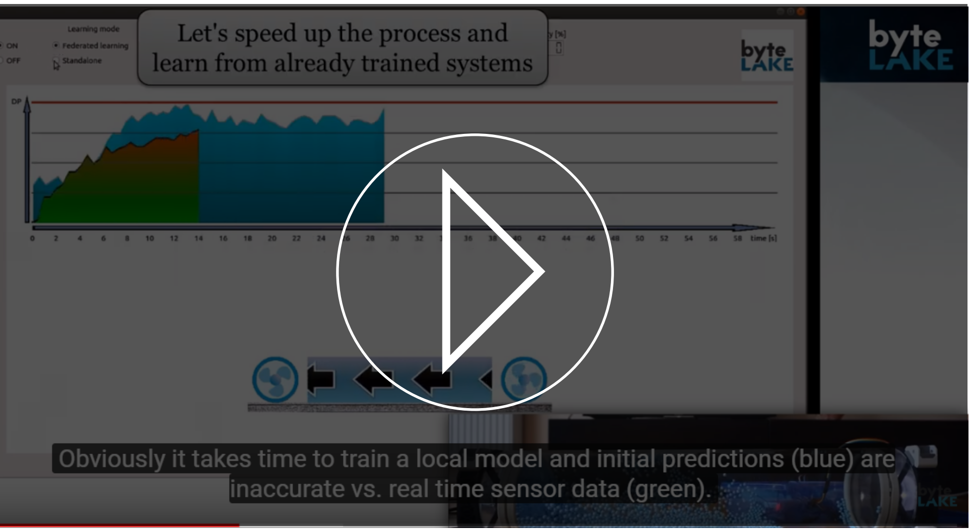
**Here is the final implementation of the PoC running live.**

We decided that in our case one the most interesting phenomenon was the constantly changing air pressure inside the pipe. We measured its real time values (time series data) with a help of a little sensor and used a Machine Learning model to predict how the barometric pressure may change over time. The Machine Learning model was used to analyze the historic data while simultaneously building a model to find patterns and predict trends.

The simulation ran virtually on many devices, measuring actual values of the air pressure over time & building a local model on the edge device. Over time, we had many local models on multiple devices which were transferred back to Lenovo's AI Innovation Center data center to be aggregated into a Federated Model.

The global model was distributed across the edge devices as an update for the local models. Consequently, all the local models benefited from each other's "knowledge" (training results) that is shared across the edge devices.

Ideas in action [pull quote] During the AI Summit event we showcased live how the local models improved over time, producing more and more accurate predictions. We also demonstrated how Federated Learning  accelerated the training process and more importantly it significantly improved the accuracy of the predictions.

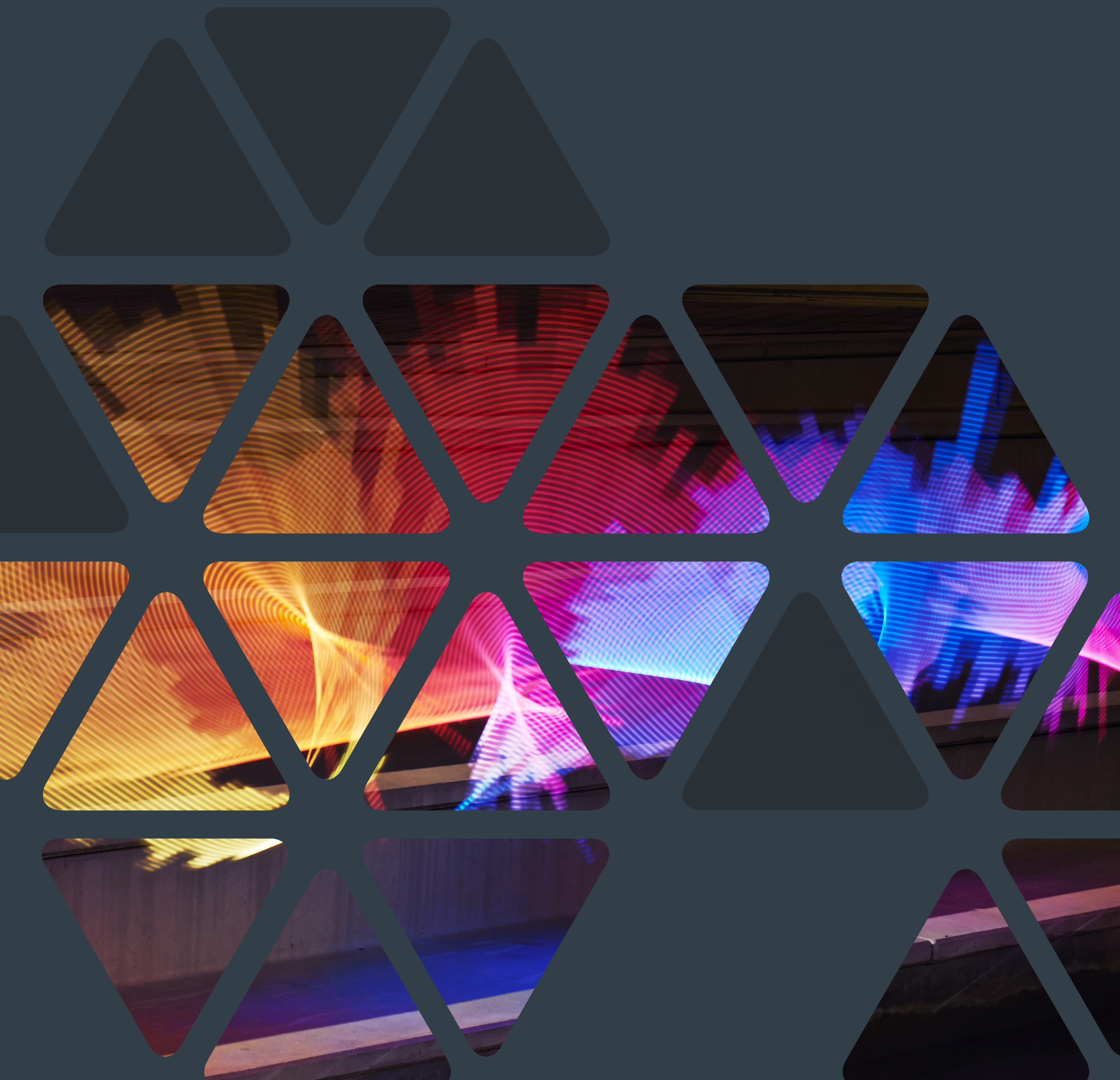



http://www.bytelake.com/en/AISummit2018-video

### Ideas in Action:

During the AI Summit event we showcased live how the local models improved over time, producing more and more accurate predictions. We also demonstrated how Federated Learning  accelerated the training process and more importantly it significantly improved the accuracy of the predictions.

# 4 / TECHNICAL DETAILS



## Procedure

The built model simulates filter clogging inside the pipe and allows us to predict the pressure changes depending on air stream.

The process inside the pipe:

- Run the first fan on one end;
- Wait until the pellets clog the filter on the other side;
- Switch the fans.

During the process we collect the following conditions:

- Air pressure changes inside the pipe;
- The distance between the pipe and ground;
- The current power for each fan.

The goal is to predict the air pressure changes across the entire process by generating a pressure profile and predict the future increases and decreases of the air pressure inside the pipe. The time of each profile depends on the power of fan and is dynamically updated as the power changes. The pressure profile is a prediction of the current simulation (filter clogging by one side airflow, i.e. from left to right or right to left).

## Used Methods

The barometric sensor connects to an Arduino, which transmits the measurements (air pressure and the distance between the pipe and the ground) to the Lenovo Tiny edge device where the time series data is accumulated and visualized. Concurrently, a Machine Learning model based on the pressure profiles is generated. The user can modify the conditions inside the pipe by changing the power of fans or switch them through the interface running on the edge device.

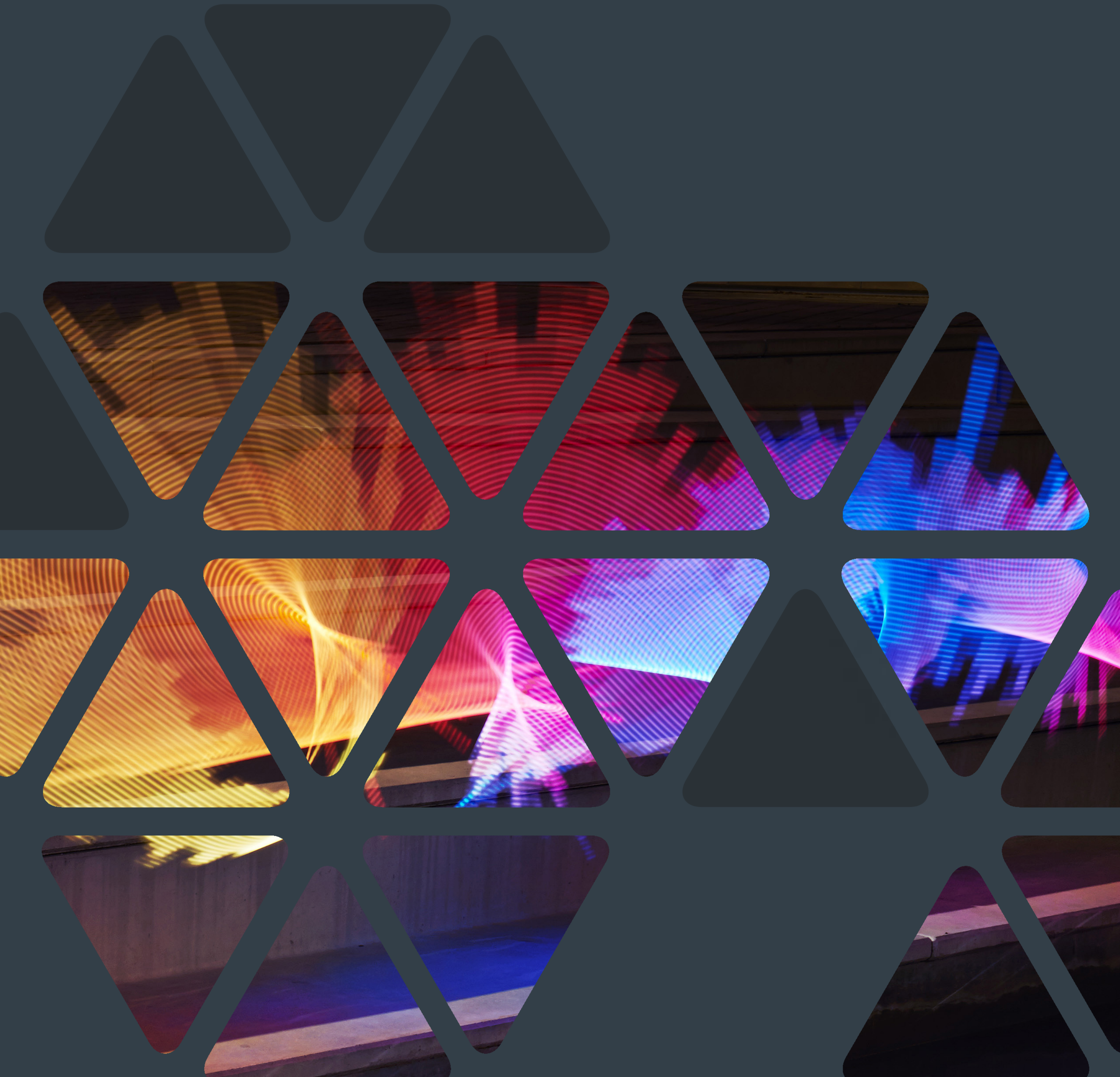
The airflow generates a new pressure profile which can be used in the future predictions. The prediction of the pressure profile for the simulation is based on the regression method that selects the most fitted profile to the actual measurements. After each simulation, the system returns the accuracy of the prediction by calculating a squared correlation coefficient (R2).

R2 equals the square of the Pearson correlation coefficient between the observed and modeled (predicted) data values.

After a interval of time (30 minutes by default), the local Machine Learning model is sent to the Lenovo AI Innovation Center data center. There, the model is aggregated with the other models from other edge devices, appropriately adapted to a single device, and distributed across all the devices. After this procedure, each device contains a fully trained model from all the connected devices.

The process of adaptation of the Machine Learning model is guarded against excessive growth since it selects the most similar pressure profiles based on R2 estimator and reduces the model accordingly. Through this refinement process, we end up with a model containing the most unique profiles, which are able to predict the very dynamic environments.

# 5 / ADDITIONAL BENEFITS



AI is already delivering real value to businesses across many industries and use-cases, enhancing the ability to make more accurate and timely decisions. We expect to see exponential in adoption over the next 24 months, as use-cases are moving from experimentation to implementation,  with a significant growth of edge use-cases.

We see more and more categories of devices integrated with some level of AI technology, from intelligent smart home devices improving with every generation to any flavor of chat bot assistant and of course industrial solutions bringing automation across various businesses.

Before any of these technologies hit the shelves, architects & product owners need to make some not so easy choices. Some obvious decisions range from picking the right hardware, acceleration (and we have a blog post on that here) and software stacks.  However, what opens a plethora of opportunities is making these devices actually share that intelligence, and that's exactly what Federated Learning enables.

## Moreover, it gives the following benefits:

- **Enables scalability**
  Decentralized AI enables IoT/devices to learn from each other
- **Solves low-throughput and high-latency challenges**
  local AI models reduce latencies, lower power consumption
- **Improves accuracy**
  have smarter models via aggregation of many local models
- **Reduces training time**
  benefit from local training and already trained models in the neighborhood
- **Lowers the cost of training**
  bringing data from all devices is expensive
- **Ensures privacy**
  sensitive data stays local

**Demonstrating the power of AI Federated Learning**

**Moving AI to the edge**

## Start your AI journey with Lenovo & byteLAKE

We understand that not one single company can provide comprehensive, end-to-end solutions for every use-case. That's why we're bringing together the right ecosystem of partners to deliver AI solutions. Lenovo has recently launch the AI Innovation Centers around the world and we're working on a number of use-cases with ByteLAKE, including some very promising work in Federated Learning, image recognition and more.

Interested in learning more or exploring your own AI Proof-of-concept? Reach out to your local Lenovo or ByteLAKE representative to schedule a briefing or a LAUNCH:AI Workshop and work directly with data scientists and researchers to explore and embrace the full potential AI for your organization.

Learn more

## About Lenovo

Lenovo is a $46 billion global Fortune 500 company providing innovative consumer, commercial, and data center technology. Lenovo delivers compute, storage, networking, and services including converged, cloud, Big Data, HPC, AI, IoT and hyperscale solutions for the enterprise market.

www.lenovo.com/AI

## About byteLAKE

byteLAKE is a company that specializes in building highly optimized AI and HPC software solutions. byteLAKE delivers these for Data Center / Cloud infrastructures (CPU, GPU, FPGA - based) as well as embedded, edge devices (IoT).

ByteLAKE.com | welcome@byteLAKE.com

## Authors


**Marcin Rojek**

byteLAKE | Co-Founder

**Robert Daigle**

Lenovo | AI Innovation & Business Leader

**Boris Tvaroska**

Lenovo | Global AI Solutions Lead

**Jonathan Quick**

Lenovo | AI Strategy Director


Lenovo™