AI Solutions for Automotive

Launching Customer-centric AI for Autonomous Vehicles and Smart Cars

appen

Table of Contents
AI Solutions for Automotive

appen

Introduction

The automotive industry is one of the best examples where, when it comes to artificial intelligence applications, reliable training data and brilliant engineering are critical for creating amazing driver, passenger, and self-driving experiences. It's an industry where there is little margin for error, and the stakes are high. At Appen, we've been fortunate to work with the top seven out of 10 automotive manufacturers. It's given us a chance to truly understand just how many opportunities there are for innovation with artificial intelligence (AI) and how vital collecting and labeling training data at scale is for the industry. As these auto giants look to get more creative and find a competitive edge, it's clear that ecosystem partners are also working to do the same in order to keep pace with the industry.

Even beyond the automotive industry, many AI projects never reach deployment – meaning companies spend a lot of time and money working on projects that never make it off the ground beyond the pilot phase. So how can you not only make AI a strategic investment for your business but also do so effectively?

We've created this guide to help the automotive industry identify and successfully scale AI initiatives to production by focusing on the consumer's experience.

Competitive Edge with AI

The number of enterprises implementing AI has grown by over 270% since 2016. – Gartner

The amount of heavy investment (read: <u>billions</u> of dollars) into AI initiatives signals a market shift, with companies looking to find a competitive edge, and feeling confident that AI can get them there.

In fact, companies that have already adopted AI report that it has directly impacted their customer satisfaction and ultimately boosted their bottom-line results.

So how do you tap into finding business value with AI? Or, if you're already experimenting, how do you move from pilot to production to core business?

The first key to success is remembering that some time is needed before ROI for AI can be seen. Before allowing time for your models to expand and prove its value, you'll need stakeholder buy-in for the benefit of AI. To do this, you'll need to identify the key business and consumer use cases, the data you have, and the data you need. We discuss this further in our AI Center of Excellence ebook.

Other critical elements for successfully deploying AI initiatives? Moving beyond stakeholder buy-in to include executives by establishing a clear business objective, building an AI Center of Excellence team and architecture, and launching your initiative with a flywheel – which we'll discuss in more detail.

When implemented correctly, AI and machine learning (ML) will deliver wide-sweeping value to businesses across many industries. Our research and experience have found one of the easiest ways to move AI pilots to scaled deployments with tangible profits is to focus on one key objective. We find that most companies have early success by building AI that positively impacts the consumer experience – whether it be inside the cabin as a passenger or driver, or outside the car to improve safety and autonomy.

How AI Will Reshape the Auto Industry in an Experience-first World

Use Cases

Market research: understand what people do while driving

Localization: multi-lingual infotainment systems

Chatbots & in-car virtual assistants: enable conversational commands within the cabin

LiDAR: autonomous vehicles
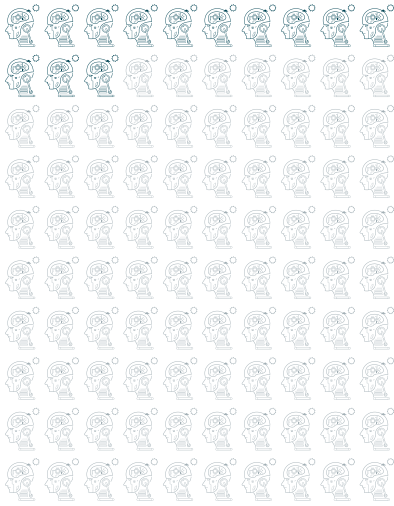
Facial recognition: driver profiles

Safety data: improve features and functionality

The auto industry is undergoing a profound transformation, with AI and electric changing the way companies build cars and influencing the type of consumer that will ultimately buy or use those cars.

The future of transportation will be built with world-class AI, ultra-fast connectivity, and environmental impact in mind. Because of this, the scope of potential use cases for AI is massive. And while business use cases for AI and ML are becoming more varied (ranging from supply chain and manufacturing to autonomous vehicles and mobility-as-a-service), consumer experience-centric applications continue to be the most common and successful to deploy at scale. This is because both in-cabin and out-of-car experiences are directly tied to clear KPIs, and many automotive companies have large amounts of **untapped data that can be leveraged to improve these experiences.**

For companies heavily investing in self-driving technology and the future of the connected car, teams are busy building a fully autonomous vehicle, improving driver assistance features, or any solution in-between.

To do this, they often have to work with multiple vendors and applications to collect, label, prepare, and converge all that data to train their AI models effectively.

But building the future of transportation is complicated enough without having to connect several dozen different data pipeline components and integrate a plethora of APIs. In order for a car to "see," "hear," "understand," "talk," and "think," it needs video, image, audio, text, LiDAR, and sensor data to be correctly collected, structured, and understood by its machine learning models.

In the case of autonomous vehicles, like with healthcare or other use cases that might risk human lives, training data needs to be annotated and verified by humans at scale, so machines deliver 100% accuracy every time. Combine that with the fact that cars not only need to abide by strict national and regional regulations, but also have to understand hundreds of languages and dialects, and it becomes an exponential challenge. This is how automotive organizations fail to operationalize their models – and without fully operational models, ROI and business value remains absent.

To bring pilots to production, turn to a data partner for support. Their expertise will enable you to put together these demanding components for the consumer experience, paving the way for success and scale.

87%

of data science projects never make it into production

– VentureBeat AI
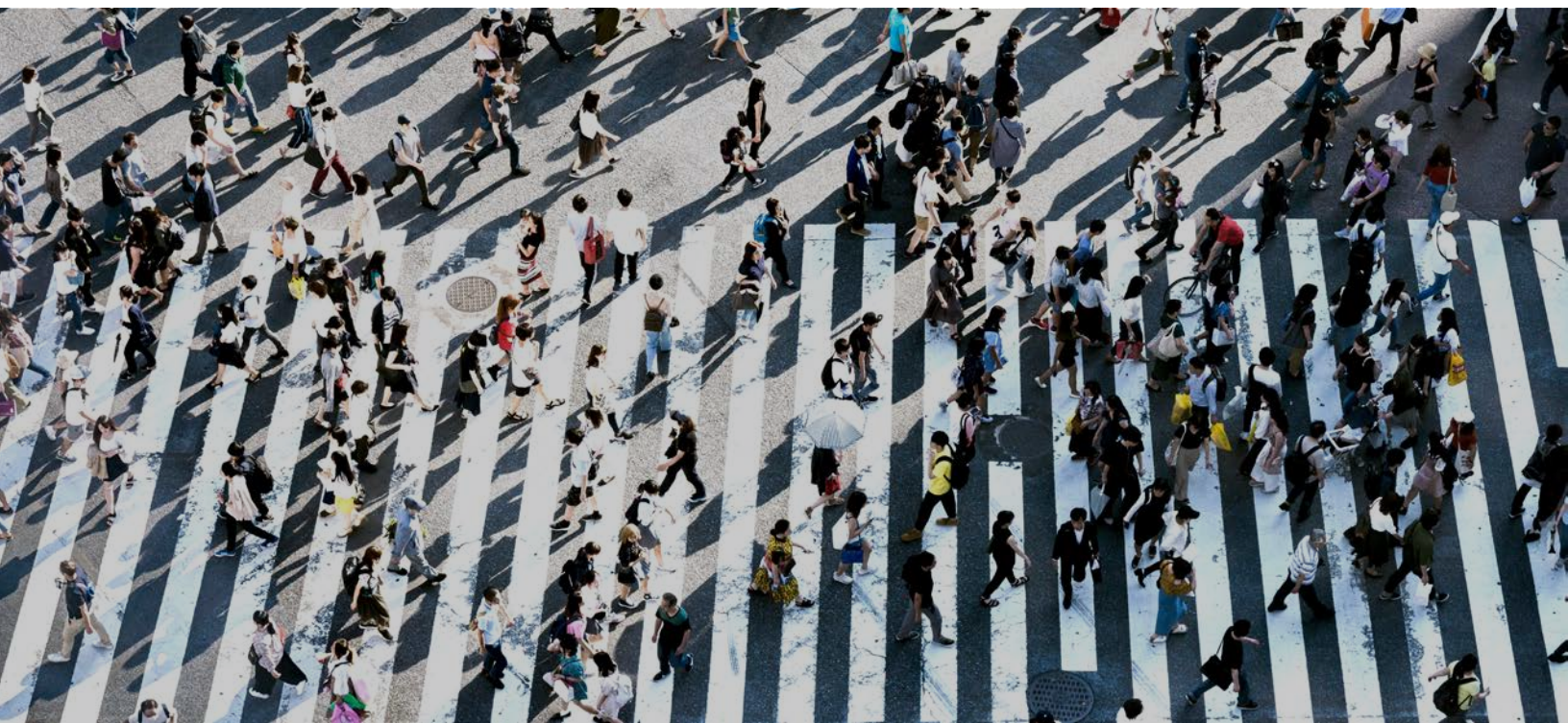
Where to Focus AI Investments

As the race to a fully autonomous vehicle continues, a standard has been put into place to define six levels of automation so that automakers, suppliers, and policymakers can discuss and compare systems. These six levels of automation correlate to different consumer experiences, **with a pivotal shift occurring between levels 2 and 3**. This is when the responsibility for monitoring the vehicle shifts from the driver to the system. Because of the different levels of autonomy, zeroing in on consumer experiences creates quick wins for both in-cabin and out-of-car experiences, making them susceptible to successful scalability.

	No Automation	Driver Assistance	Partial Automation	Conditional Automation	High Automation	Full Automation
	L0	L1	L2	L3	L4	L5
Driver						
Vehicle						
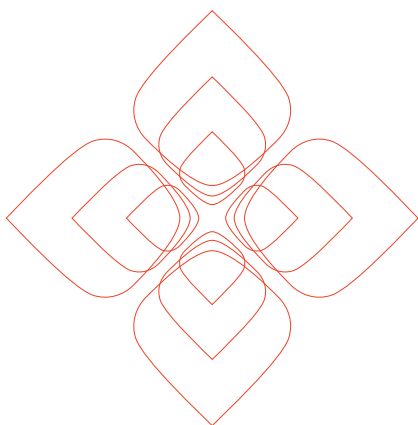
In-cabin experiences:
This is often described as the AI-powered cockpit but goes beyond just the driver's experience. In-car experiences encompass the entire user experience – including the driver and all passengers, working to make the overall in-cabin experience more intelligent and enjoyable. This might include AI uses for intelligent driver assistance programs that improve safety or infotainment systems that can provide directions for the driver while giving content recommendations for those in the backseat.

Out-of-car experiences:
Autonomous vehicles continue to dominate the conversation when it comes to out-of-car experiences. While the push is to reach level 5 autonomy, there are many gradual progressions of how AI impacts the out-of-car experience on the road to get there. Smart cars powered by AI require higher-levels of computer vision and computing power – with sensors from radars and cameras delivering massive amounts of data every second to process things like hazardous road conditions, objects in the road, and road signs.

The best part about both in-car and out-of-car experiences? Both are ripe for AI adoption and scalable deployments with the help of training data.

Where to Focus AI Investments

AI-Powered Cabin

The AI-powered cabin has become synonymous with many companies' brands.

Auto manufacturers are partnering, or looking to partner, with relevant ecosystem providers to create more value for their customers. Benefits of the AI-powered cabin include improved driver experience and safety, as well as intuitive in-car assistants.

Tractica forecasts that the market for automotive AI hardware, software, and services will reach **$26.5 billion by 2025, up from $1.2 billion in 2017**
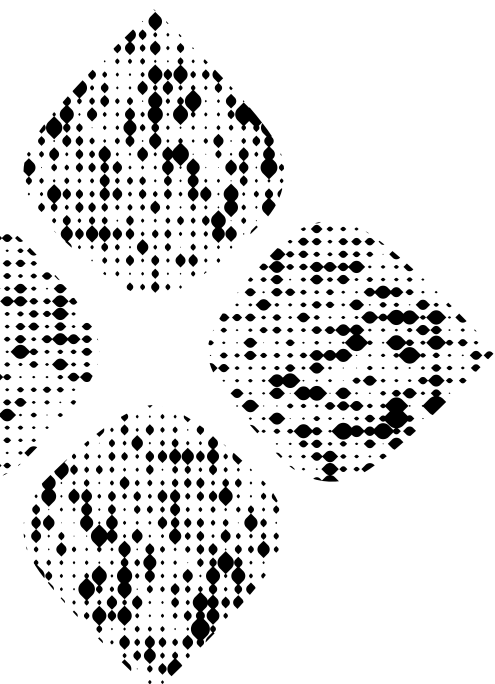
Staple elements of a smart cockpit consist of object classification, scene understanding, location services, natural language understanding, face detection/ fingerprint recognition, sensor processing and fusion, voice/noise cancellation, and voice/speech recognition. These come together to create AI-powered in-cabin experiences like:

Conversational Assistance: Enables consumers to interact with their vehicle in a 21st century way – through speech. Your data partner should provide comprehensive training for conversational AI, including speech collection, speech annotation, transcription for the creation of Automated Speech Recognition models, lexicon building, and more.

This extensive training is critical when factoring bias. For example, a native English-speaking male driving a car manufactured for the U.S. market will likely experience a much higher level of speech recognition success than a woman or non-native speaker. In simple terms, speech recognition systems that rely primarily on data collected and annotated based on native English male voices will not work as well for other voices. Bringing in natural language understanding to tune models for over 180 languages and dialects will improve the AI experience for everyone.

Driver & Passenger Monitoring System: Used for instances such as alerting a driver when they are no longer focused on the road. Your data partner should support your tracking models through a vast set of multimedia data such as facial and gesture annotation to ensure AI-models remain unbiased. For example, If the training data for a driver monitoring system is based on data collected in an environment with one quiet passenger, the system will respond poorly when monitoring a car full of a large family and their favorite pet.

Ensuring enough unbiased training data for multimodal and multimedia visual and speech recognition systems requires an enormous number of diverse annotators representing a wide range of geographies, cultures, genders, and languages. And all this data must be collected, annotated by human experts, and used to train and improve ML models efficiently, with speed, and at scale. Accessing quality training data across the spectrum is a difficult, but essential, component for the automotive industry in the race for an autonomous vehicle.
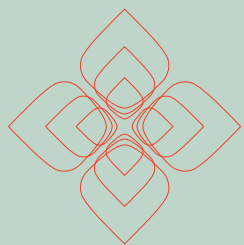
Bringing in natural language understanding to tune models for over 180 languages and dialects will improve the AI experience for everyone.

❖ appen

Where to Focus AI Investments

Out-of-Car Experiences

Autonomous vehicles need not only to understand their passengers and drivers but to be able to navigate a complex world. This is a mission–critical use case for AI, where there's very little room for error. While the process of full automation has been gradual, the slower process helps build trust with consumers as automotive companies work through the different level classifications of autonomy. Thanks to recent research in computer vision ML models, the AI–powered self-driving opportunities are heavily focused on computer vision with LiDAR, video object tracking, and sensor data. These help cars "see" and "think" when driving from point A to point B. Data annotation services that help train models to perform include:

Point Cloud Labeling (LiDAR, Radar): Understand the scene in front of and around the car by identifying and tracking the objects in the scene. Merge point cloud data and video streams into one scene to be annotated. Point cloud data helps your model understand the world around the vehicle.

2D Labeling including Semantic Segmentation: Help your model get a fine–tuned understanding of the input from its visual light cameras. Find a data partner that can help with scalable bounding boxes or highly–detailed pixel masks created for your custom ontology.

Video Object and Event Tracking: Your model has to understand how objects move through time, and your data partner should assist in labeling temporal events. Track objects in your ontology (like other cars and pedestrians) as they enter and exit the area of interest over many frames of videos and LiDAR scenes. It's critical to maintain a consistent understanding of the object's identity through the entire video, no matter how often they drop in and out of sight.

Companies used to have to lean on multiple vendors and applications to collect, prepare, and converge all data in order to effectively train their AI models.

Until now. Whether you are building autonomous solutions for Level 1 or Level 5, improving driver assistance features, or something in–between, a reliable collection and annotation partner provides a unified offering to train and test your AI systems in one place.

Accelerating AI from Pilot to Production with a Data Partner

"Machine learning has been transformational for the automotive industry. cnvrg.io has worked with world renowned car innovators, building AI features we never thought possible from autonomous driving, to lane sensors and predictive maintenance alerts. As autonomous cars become reality, we see the importance of combining machine and human power to have scalable ML pipelines that ingest human annotated data. Setting a mechanism for live automatic retraining and refitting of models to adjust to the new reality. There is a tremendous need for machine learning applications that can adjust to live data while maintaining human in the loop."

Yochay Ettun
CEO and Co-founder
at cnvrg.io

Many AI projects begin with data collection of what is immediately available and then try to understand how to use it. Successful approaches to scaling models beyond projects will avoid general data (collected from public sources and the web and dirty/dark data), and instead will focus on gathering specific data related to realistic goals and use cases. To be successful, this data will be reliable, clean, and sufficiently annotated, and teams will commit to data maintenance and outsource where they lack expertise.

The amount of time spent on **preparing training data is up to 80%**. **To ensure your AI pilots do more than look good on paper, invest in your training data.** When it comes to actually adopting the strategy of a pilot model and providing ROI, many projects fall short on delivering meaningful results. This leads to push back from senior management, failure to impress CIOs, and unrealized value killing the pilot. Consequently, management will struggle to justify projects, and often won't want to invest in scaling future pilots.

Through 2022, only 20% of analytic insights will deliver business outcomes, according to Gartner

Launch your world-class AI initiatives by turning to data partners who provide reliable, high-quality training data that enables you to scale through each of the five key phases:

01 **Pilot**

02 **Data Annotation**

03 **Test & Validation**

04 **Scaled Deployment to Production**

05 **Retraining**

Accelerating AI from Pilot to Production with a Data Partner

01 Pilot

Working with a vendor to supply you with reliable training data during your larger pilots will help ensure your models can move quickly through the remaining phases and get you to scale sooner. Your data partner can also help with annotating data that yields low confidence or with annotating edge-case scenarios.

02 Data Annotation

The heavy-lift of training data is often needed after a small pilot. This is when you want to train your model with massive data sets to make sure it works in every scenario, isn't biased, and works the way the model is intended. Further, this data needs to be accurate, or else your model won't be trained properly and will immediately tank the business problems being solved for, resulting in low buy-in for a scaled deployment from stakeholders.

Here is where many companies think they can use the general data they have but instead should turn to experts in the data annotation and collection space. This will reduce the time spent on getting data as well as ensure the highest accuracy possible.

03 Test & Validation

Once you've trained your model, it's important to validate it by using a set of data it was not trained on in order to tune the model. Working with a data partner during the validation phase can help test that the data was properly labeled with the right intent and ensure that the model isn't leading to any biases or failing due to edge-cases. Once your model is dialed in, you'll need an additional set of new data from your data partner that your model was not trained or tuned with to get an unbiased estimate of the skill of the final tuned model.

Accelerating AI from Pilot to Production with a Data Partner

04 Scaled Deployment to Production

If your model has been successful in the test and validation phases, it's time to scale your deployment. Companies may turn to their data partner to further evaluate and validate low-confidence answers but should feel confident taking their pilot to scale.

05 Retraining

Great! You moved to scale – but how long will your model accurately perform at full deployment? Regular retraining is mission-critical to avoid model drift and account for use-case transformation.

Bonus:

Building a Flywheel

To turn AI into a core part of your business, you'll need to build the flywheel to scale your training data-creation and model-building processes across other facets of your business to deliver AI from experiment to production to core product.

This AI flywheel should be a larger part of your organization's AI & data center of excellence, a topic we discuss in greater detail in our AI Center of Excellence eBook.

❖ appen

TRAINING DATA — TEST & UPDATE — PREDICTION — MODEL
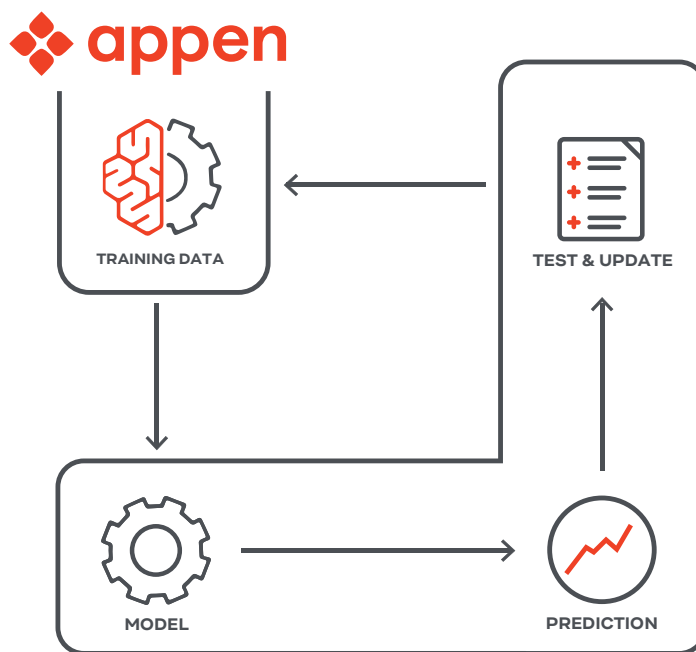
Key Challenges Ahead

While following the above five steps seems simple enough, organizations should be aware of and prepare for some key challenges. Doing so will help build in safeguards to keep forward momentum.

Fundamentals

Teams who are eager to get started often forget to pause and think about the fundamentals. Once you're ready to engage a data partner to scale your projects, details such as how to get data to your data partner or how you'll view data from your data partner can get skipped over. Ensure your data partner offers end-to-end support and can offer their expertise and guidance. Once you receive your annotated data, how will you actually view that data? How will you ensure it meets your quality standards? For example – do you know what program you need to view annotated LiDAR data? If you cannot view this data, how do you ensure it was done correctly, and the project was properly annotated so your models can properly leverage the data? A good data partner will be able to offer support through every phase of your project from start to finish.

"The collection, management and AI-based exploitation of automotive- and transportation-relation data will determine the winners of next-generation mobility. It will require strategic vision, business acumen, and technology prowess. Appen is bringing together the necessary ingredients to help corporations succeed."
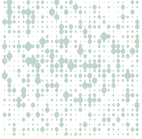
Evangelos Simoudis
Managing Director, Synapse Partners
author of "The Big Data Opportunity In Our Driverless Future" and of the forthcoming book "Transportation Transformation"

Key Challenges Ahead

Level of Complexity

Similar to the fundamentals, organizations may not be tuned-in to how the level of complexity can influence their projects. By turning to a reliable data partner, their expertise can help provide direction and insight. The larger the ontology, for example, the more complicated the project. A well-versed data partner will help identify how this leads to more time and cost and find solutions that will work for your overall business objectives, which is especially critical for factoring in images and videos.

Localization
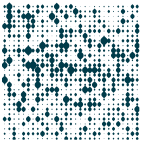
Localization is especially crucial within the automotive industry. Because automotive companies need to design models with multiple markets in mind, it's important to factor in different languages, cultures, and demographics to customize the consumer experience properly. Localization projects are great for getting in front of your data partner who can leverage teams of skilled linguists to develop things like style guides and voice personas (formal, chatty, etc.), as well as optimizing across many languages.

Key Challenges Ahead

Security
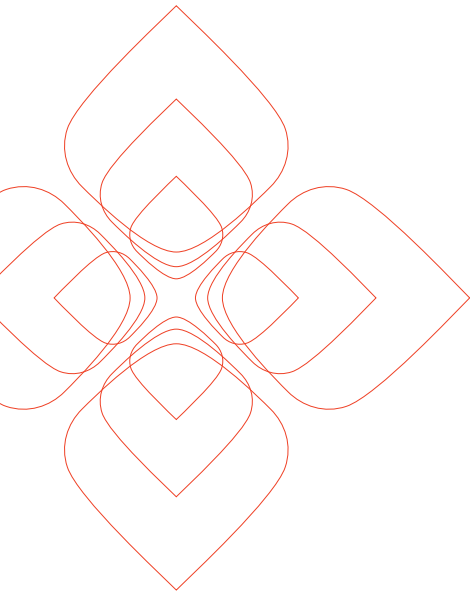
A lot of data collection in the automotive industry contains sensitive data that requires additional security measures in place. A proper data partner will not only offer a variety of security options but will also have strong security standards at even the most basic level to ensure your data is properly handled. Look for data partners who offer options such as secure data access (critical for PII and PHI), secure crowd and onsite service options, private cloud deployment, on-premise deployment, and SAML-based single sign-on.

Secure Data Access ensures all data security requirements are met for customers working with personally identifiable information (PII), protected health information (PHI), and other sophisticated compliance needs.

Secure crowd and secure onsite service options where contributors access tasks through machines that are owned/operated by the channel in a controlled and monitored physical location.

Private cloud deployment which can be hosted on your specific cloud environment or hosted and managed by us.

**On-premise deployment** deployed in your particular network either air-gapped or non-air-gapped.

SAML-based single sign-on (SSO) which gives members access to the our platform through an identity provider (IDP) of your choice.

Key Challenges Ahead

Retraining

According to McKinsey, one-third of AI products that go live need monthly updates to keep up with changing conditions, like model drift or use case transformation. Many companies skip over this critical step or put it on the back burner altogether. Still, the risk of your AI project deploying at scale and being successful long enough to prove ROI becomes increasingly limited the longer retraining is avoided. Retraining allows you to iterate on your model, making it more accurate and successful – this is best done by leveraging a data partner for relabeling data and providing support by using human evaluators to analyze low-confidence predictions.
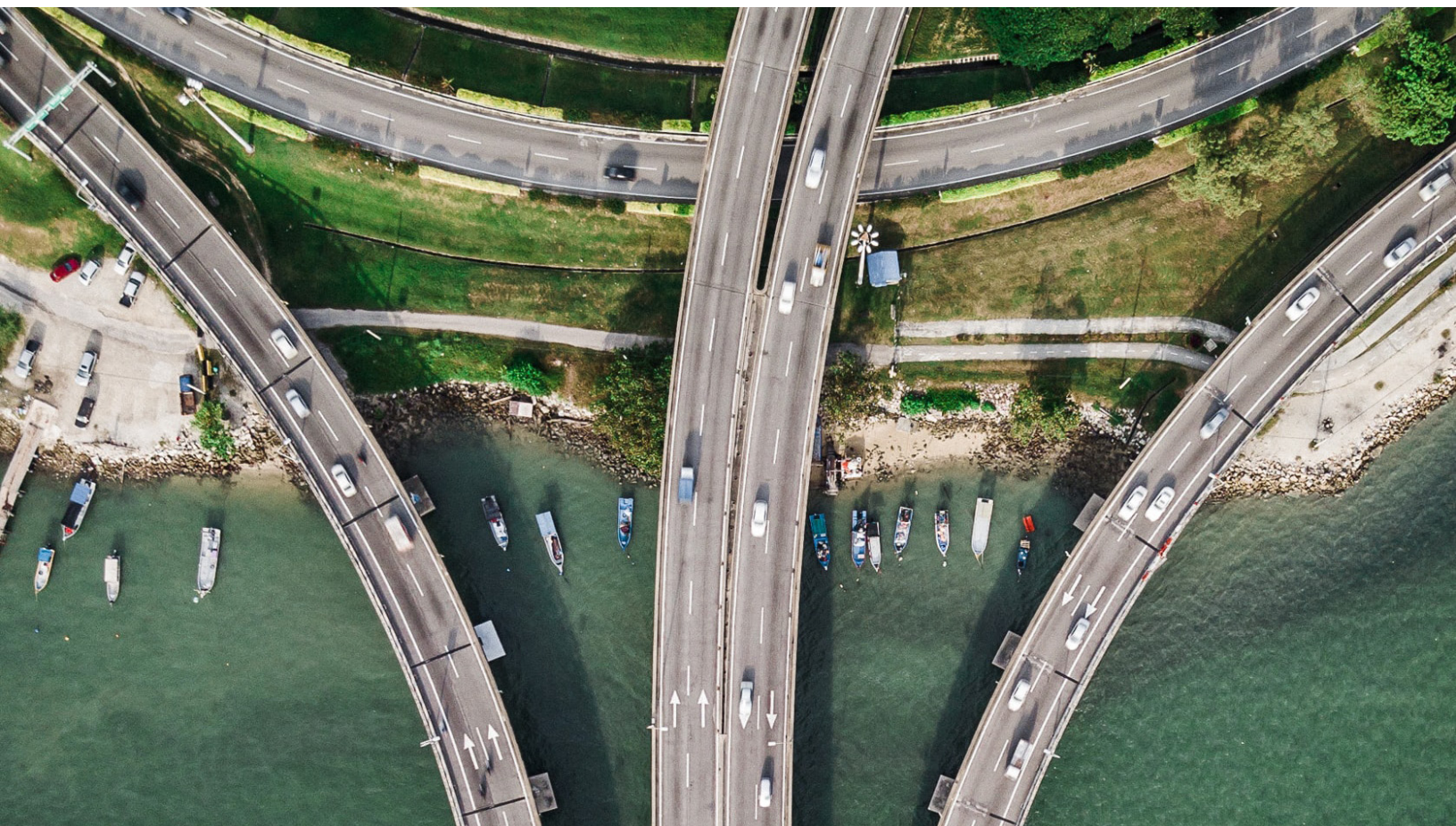
Looking Ahead

When it comes to launching world-class AI, the opportunity for the automotive industry is massive – whether you're working to build smart cars, evolve the customer buying process, or enhance the in-cabin experience. It's evident that only a fully operational model that reaches deployment will deliver any kind of business value – and the best way to beat the less-encouraging odds is to identify use cases where reliable training data (with the right data partner) can get you there.

While the path to an AI-driven automotive revolution is currently gradual, we're hopeful that more and more organizations will leverage the substantial amounts of reliable training data to get their AI projects into the real-world.

As part of that mission, it's important to recognize that **world-class AI has to work for everyone, in every market.** Leaders building self-driving cars are thinking beyond simple efficiencies, speed, and cost. Removing bias from the data is paramount so that AI recognizes everything and everyone equally. As a top OEM or Tier 1 automotive supplier, you want your consumers to be safe and understood by the cars they are using, no matter their ethnicity, gender, age or geography.

What's more, leaders are urged to consider their supply chain impact. **An ethical-first approach to data means unbiased AI, which creates a positive global impact.** This can offset some of the disruption Level 5 self-driving will bring to the world.

Our capabilities for the Automotive Industry:

Automatic Speech Recognition
Improve consumer interactions with Automatic Speech Recognition systems by training them to better understand human language.

In-car Infotainment
Ensure your in-car infotainment system provides the best user experience with high-quality training data and testing.

In-car Navigation
Train your in-car navigation system with high-quality data to ensure consumers get where they need to go.

Text-to-Speech
Improve consumer interactions with TTS systems that are fluent in every language.

Virtual Assistants and Chatbots
Train your virtual assistant or chatbot to better understand and respond to human interaction, driving higher levels of consumer satisfaction.

Images, Video, and Computer Vision
Support your autonomous vehicle efforts through simple image classification to pixel labeling semantic segmentation for a better understanding of out-of-car environments.

Machine Learning Assisted Video Object Tracking
Receive annotated video faster through our video annotation solution, which combines machine learning and human-generated training data labels to track objects moving through space and time up to 100 times faster than human-only solutions.

About Appen

Appen collects and labels images, text, speech, audio, video, and other data used to build and continuously improve the world's most innovative artificial intelligence systems. Our expertise includes having a global crowd of over 1 million skilled contractors who speak over 180 languages, and the industry's most advanced AI-assisted data annotation platform. Our high-quality training data gives leaders in technology, automotive, financial services, retail, healthcare, and governments the confidence to deploy world-class AI products. Founded in 1996, Appen has customers and offices globally.

- **With 25 years of experience,** 15 of which in Automotive. We offer a full suite of multimodal computer vision annotation tools with in-cabin vehicle collection and NLP annotation services to help with your autonomous vehicle projects.

- **1M+ crowd in over 130 countries,** speaking 180 languages and dialects

- **Experienced team based in the heart of Motor City,** Detroit lends its expertise and resources on the ground to accelerate your product development and testing workflows.

- **Data annotation expertise** ranges including conversational assistance, point cloud labeling (LiDAR, Radar), 2D labeling including semantic segmentation, and video object and event tracking.

- **Workflows:** Our simple user interface empowers teams to build and automate multi-step data annotation projects without relying heavily on technical resources. Break complex projects down into simple jobs, then automatically route data between the jobs using configurable routing rules. String multiple jobs or models together in a branching or linear configuration. Leverage machine learning in workflows to offset costs and expedite project completion.