

AUTO AI BERLIN 2017

Conference Summary And Key Takeaways By The Chairman



AD PUNCTUM

October 2017

Executive Summary

AutoAI 2017 was held in Berlin on the 28th and 29th of September. Attendees representing academia, original equipment manufacturers (OEMs), suppliers, start-ups and consultants shared ideas through presentations, peer-to-peer round table sessions and informal discussions. Attendees were asked questions using anonymous live polls throughout the event.

There were four broad topics relevant to understanding the development of artificial intelligence in the automotive sector:

- Artificial intelligence outlook including machine learning & deep learning techniques
- Computer vision, imaging & perception
- Sensor fusion & data
- Simulation, testing & validation of artificial intelligence

Key Takeaways:

- **No one seems to be advocating a rules-based approach for autonomous driving.** Although possible in concept, a set of detailed rules appears impractical in the real world because the volume of instructions needed to cover all driving conditions would be massive and debugging would likely fail to spot all errors or contradictions.
- **There needs to be a common view of the acceptable safety level for deployment of autonomous vehicles.** This requires the input of regulators, industry, academia and probably 3rd party consumer groups too. In addition to the safe level, we need to determine how to measure it and confirm compliance -- the second part is challenging without a test fleet covering tens of billions of miles.
- **Safety is related to the use case.** The more complex you make the driving conditions, the more corner cases you'll encounter. The scale of the problem can be reduced by steps such as: learning a particular area (geo-fencing) and not driving in extreme conditions or at very high speeds. Although this diminishes the value for the retail user, there are plenty of industrial applications that can operate within those constrained operating conditions.
- **There needs to be widespread collaboration on a shared framework for testing and validation of AI.** Governments, companies and academia should all be involved and it would ideally use open data that was not tied to specific simulation tools or sensor sets. The experts at the event were wary of letting driverless cars to transport their family without seeing safety data beforehand (government and manufacturer assurances weren't trusted).
- **Work needs to be done on explaining AI.** There are big differences between the capabilities non-technical people think AI has and what it is capable of -- there should be no talk of killer robots. At the same time, deep learning techniques mean that the system function cannot be explained in the same way as traditional coding. New ways to explain how the system operates are required and without them building trust will be very difficult. It could even be necessary for the AI to learn how to explain itself using natural language or other tools.
- **Virtual testing is vital.** This is for three reasons: firstly, simulation dramatically decreases real world miles; secondly, because AI techniques like reinforcement learning need crashes to take place in order for the AI to learn; and thirdly because even real-world data becomes a simulation once you interact with it in a different way to the original situation. It's better to do that virtually! For a virtual environment to be successful it must be something that can be replicated in the real world with the same results.
- **There is plenty of disagreement over the right approach to many areas.** The event live poll highlighted differences of opinion regarding how AI should act, how much information it will be capable of processing and what level of redundancy was required for safe operation. More consistent was the high burden of proof that AI systems will be faced with and a view that currently no one really knows how to convincingly do that.
- **Implementation timing remains uncertain.** In the event live polling, over a quarter of respondents believe that self-driving will be widespread by 2023 or even earlier. The majority believe that we will be waiting beyond 2025 -- a healthy difference of opinion. Health Warning: informal discussions revealed that in general the timescale comes down if the question is about level 4 specific use case vehicles on public roads (they operate on private land already) and goes further out if asked about go-anywhere level 5 vehicles.

Artificial intelligence outlook including machine learning & deep learning techniques

Key Takeaways:

- Vehicle electronics are growing in value but require further standardisation and reductions in power consumption
- Data storage is a major issue -- techniques from traditional big data do not work very well with images and video
- Image recognition is improving but research would benefit from wider availability of labelled video datasets
- Further work is required to create greater depth of scenarios and improve simulation processing times
- Realistic visualisation of simulations for humans is different to modelling the sensor inputs vehicle AI interprets
- Understanding machine learning isn't always hard... sometimes it comes up with simpler rules than we expect!

Market Growth... is forecast at 6% compound annual growth rate (CAGR) for electronics -- reaching \$1,600 of average vehicle content. For semi-conductors the figures are even more impressive -- 7.1% CAGR. The specifics of market development are less clear -- these growth figures include L1/2 systems but not full autonomy. Although there is a definite role for the technology, standardisation is a must, requiring a yet-to-be-established framework. Safety is a big challenge: without clear agreement on what safety level is acceptable, definite technical standards cannot be set. Another open issue is the degree to which the car will have to make decisions for itself versus interaction with infrastructure and other vehicles. The problem is the latency (response time) of large data sets. Finally, self-driving chipsets must consume significantly less power than current prototypes.

Researchers have gained new insights by translating real world crash data into a virtual environment... the information came from records collected by the regulators. Technology in production today sometimes makes facile errors (e.g. lane keeping recognising a bike path rather than the kerbside). Research has shown that it is possible to correlate the virtual models with real world data (for instance replicating a collision with a pedestrian) but the challenge of testing thoroughly remains substantial. Multiple different environments are needed; there are thousands of types of crash situation; and each vehicle has unique attributes. Through all of this, it is vital that the results correlate to the real world. Researchers aim to reduce modelling times from days (currently) to hours -- real time is the ultimate goal. Without improvements in processing speed, virtual sample sets are in danger of remaining too small or too slow to be usable.

The challenge of staying aware of the state of the art in data processing and artificial intelligence... large OEMs are interested in AI in the broadest sense -- self-driving, handling customer data and improving business efficiency. The first challenge is the data itself. Not only will the car become a massive source of data, but much of it does not easily fit into existing data structures -- images and video are more complicated and unstructured than traditional inputs. With images, it may be necessary to pre-process and store the resulting output, rather than the image itself, to reduce storage space and retrieval time. Image capture and object recognition as a definite area where more work is required and machine learning is already relevant, for instance recognising brands of truck trailer may help build broader recognition of what a trailer looks like. By studying a whole range of machine learning activities (a huge resource undertaking), organisations can develop an understanding of the best fit between problems, data collection methods and analysis tools.

There are different ways of obtaining image data in real time... dedicated chips can translate lidar traces (compatible with multiple lidar types) into an instantly available augmented image. This allows object identification from the raw data and for less expensive lidar units to be used. Examples showed a 16-line lidar unit being augmented for higher resolution.

Machine learning already has applications in ADAS feature sets... it has been put to use in two frequently encountered highway situations: roadworks and other drivers cutting in and out. Video and radar example data was combined with machine learning and human guidance about acceptable limits of driving behaviour. Interestingly, in both cases although the machine learning was given multiple data inputs, only a few key elements were required to provide very good accuracy in use. This reduces the sensor inputs and complexity of processing. For example, machine learning identified a high correlation between the angle of the vehicle in front and whether it was intending to cut in, in preference to more complicated rules combining relative speeds and side to side movement. Multiple sensors should be used for decision making: although a camera is better for monitoring many of the situations, its limited field of view means that radar needs to be used in camera blind spots.

The car will become one of the biggest generators of natural language data... and its proper use will enable manufacturers to create a personalised experience that the customer values. For relatively complex commands ("when X

happens, then please do Y”), contemporary techniques have 95% correct transcription of what the customer is saying and mid-80s% task completion. This is encouraging but shows further development is needed. OEMs will also have to create ecosystems that allow them to control the customer experience inside the cabin, yet are seamless with the personal assistants the customer might have on their mobile phone or home speaker system.

New techniques are improving image recognition... Using industry benchmark tests, computer image recognition is now superior to humans. In some specific use cases this already has practical uses, for example a smartphone app that assesses melanomas. However, at around 97% correct identification of a random image (versus about 95% for humans), improvement is required. Different methods are being tested, with greater progress on static images than video; partly due to difficulty but also because video has less training data: smaller libraries and fewer labelled categories. Video identification accuracy can be improved by running several different methods in parallel. One of the most promising approaches is turning the video into a set of 2D images with time as the 3rd dimension -- a technique pioneered by Deep Mind (now of Google). Combining this process with different assessment algorithms (such as analysing the first and nth frame rather than all frames), teams have achieved accuracy of nearly 90% for gesture recognition. Generally, late fusion (a longer gap between frames) gives better results than early fusion -- there is variation in what combination of processing algorithms yields the best accuracy. Progress is happening all the time. New ways of addressing machine learning problems sometimes create step changes, so improvement may not be at a linear rate.

It is hard to create different virtual environments for vehicle testing... Using video game tools and very experienced developers, near photo realistic models can be created, but this appears to be the easy part! Because the structure of computer graphical data is different to real life, models need to be adjusted to create the correct type of artificial sensor inputs. This is even more challenging with radar and lidar input data as the model must accurately simulate the “noise” -- a factor of both the sensor and the material it is detecting. Perfecting this could take several years. More useful immediately is the ability to create virtual ground truth (e.g. that is a kerb) that can serve SAE Level 2 development. Because L1/2 inputs are more binary, sophisticated sensor simulation issues are less relevant. Researchers believes that a virtual environment of 10km-15km is sufficiently to assist development of these systems, assuming the ability to impose different weather conditions (snow, heavy rain etc).

Computer vision, imaging & perception

Key Takeaways:

- Different options emerging to help machine vision check against pre-determined ground truth
- Calibration and synchronisation of multiple sensor types is a challenge
- Using hardware specific processing techniques may improve performance -- but also impairs plug & play
- Virtual testing will be valuable but requires a robust feedback loop and real-world validation

The state of the art in mapping... Mapping companies have several layers of information such as: road layout; traffic information; photographic information; and lidar traces of the road and its surroundings. At present, vehicle navigation relies on basic inputs and workable levels of accuracy (a few metres). High definition mapping allows a car to be more precise about its surroundings and relieves sensors of the need to fully determine the environment for themselves. Detailed lidar-based maps of the roadside can be used to build a “Road DNA” that autonomous systems can reference. AI can also be used to update maps. Crowd-sourced position data helps to determine where road layouts have changed (because everyone appears to be going off-road). Currently, AI sends these problems for human investigation but in future it could make decisions for itself. There may be value in collecting images from user-vehicles to update maps, both for ground truth and road sign interpretation.

An in-depth review of the StixelNet image processing technique... This method breaks images down into lines of pixels (columns in the case study) and then determines the closest pixel, allowing identification of ground truth (kerbside, people and cars) and free space. Camera data can be combined with lidar traces from the same vehicle to allow the system to train the camera recognition using the laser data. The positive of this approach is that it is continuous and scaleable -- more cars added to the fleet equals faster learning. The downside is that it is difficult to calibrate and synchronise cameras and lidar on the same vehicle to the accuracy required. It is also difficult to write the algorithms -- several processing options are available, all with different weaknesses. Studies indicate that systems trained on the camera and lidar data showed better results than stereo cameras and better than expected performance on close-up images.

Simulation and assessment of images captured by different types of camera... Measures of image quality that go beyond traditional quantitative techniques and are relevant to machine learning can be identified. Software has been developed that can simulate the images a camera will create in particular lighting conditions and using a range of filters. One of the findings is that image processing for machine vision may be able to forego some of the steps currently performed in digital photography (e.g. sRGB). This could save time and processing power. Research has a wider scope than autonomous vehicles but the application looks promising. A transparent and open source platform looks beneficial.

Implications of deep learning... Neural networks bring high purchasing costs and energy consumption. Some of these costs do not provide a proportional increase in accuracy, there is a law of diminishing returns. It may be better to reduce the cost of one part of the system and purchase additional elements, whilst retaining some of the saving. For instance, going to a single 5mm x 5mm processor rather than two 3mm x 3mm acting in series reduces the power consumption by about half.

Creating virtual environments... researchers used scenarios designed for video games to look at visual detection. The process developed creates grid information to assess the underlying physics and then layers skins on top to create realistic images. Driving characteristics and reaction of both the target vehicle and other vehicles can therefore be modelled, including the effect of collisions. The same information can also be used to create artificial data such as a lidar trace.

The importance of hardware in artificial intelligence... test vehicles can now drive autonomously without reference to lane markings because of their free space detection ability. Power consumption is reducing -- a chip commonly used in prototype AVs uses 80W but its replacement requires 30W despite an increase in processing capacity. It may be beneficial to use processing software that is designed around the hardware. Some studies indicate that properly matched chipsets and processing suites can reduce latency and improve performance. An example end-to-end research project, which has several neural layers and multiple sensors including camera and lidar still finds decision making difficult in the following areas: gridlock, extreme weather, temporary road layouts, muddy tracks and difficult turns. There are also numerous edge cases and multi-party trolley problems.

The challenges of developing the next generation of technologies... Although lane keeping can be further improved, it is necessary to develop systems with a different approach, especially because in many cases there are no clear lane markings. One potential method is training networks to detect extreme weather and tune object recognition based on that information -- for instance a camera may be best in sunny conditions but 3D data may be best where due to snow the lane markings become nearly invisible. There is also a downside to sensor improvement... for instance, as camera resolution improves, the data labelling may become unusable and need replacement (possibly manually). Power consumption of prototype chips is too high. Some concept demonstration chips draw 250W and in series production a processor needs to be below 4W.

The challenges of helping self-driving AI to learn and be tested... Driving situations have a long tail -- there are typical situations that recur with high probability and critical situations that have low probability, such as a child running into the road chasing a ball. Despite the child running into the road being low probability, it is important to replicate multiple situations (big child, small child, from the left, from the right etc). Although difficult to get perfect digital realities, it is possible to get close and then to validate against real world experiments. It is important to get feedback from the AI about why it performed a certain way during the simulated event -- this will identify bugs in the model where unrealistic situations are created, preventing mis-training of the AI. Considerable challenges to achieving this vision remain: creating the models; constructing realistic scenarios automatically; ensuring full coverage of likely circumstances; getting realistic sensor data; and creating a feedback loop to real world tests. Not to mention getting everyone involved and doing it quickly! Happily, some simulation data is already usable, despite not being perfect. When one group evaluated how ADAS systems on the market today reacted to children running into the road (physical trials using dummies), the results suggested performance could improve through combining existing analysis and decision-making techniques. Although at least one vehicle passed each test, none passed them all.

Sensor fusion & data

Key Takeaways:

- There are many opportunities to reuse knowledge on other applications but the transfer is unlikely to be seamless
- There are ways around complex problems without solving them -- for instance vehicle re-routing or copying others
- Still unclear whether self-driving AI should be a number of small modules or a centralised system
- Large evolutions in hardware capability likely to reduce the value of previously collected data

Examples of autonomous driving technology being applied across different use cases... A fully autonomous sensor set is likely to contain five cameras, five radars, lidar covering 360° and ultrasonic sensors. This creates plenty of challenges including integration problems and extreme weather conditions. Real world experience from production and research projects is useful here. The first case study was the execution of ADAS in trucks. The initial translation of passenger car technology (using the same type of camera, but mounting it higher) meant that the field of vision over short distances was reduced. In the second generation, a fish eye camera was added alongside the original camera. This provides enhanced short distance recognition whilst preserving the strengths of the existing system over longer distances. The second example was of a prototype automated tractor trailer coupling system where the vehicle lines up the trailer at the same time as detecting any objects (e.g. humans) in the way. This is done in conditions where sensor input may be impaired, for example the camera is likely to get mud on it.

Practical learnings from driverless research and production implementation... Full autonomy in all situations is a massive challenge. If it can be limited, perhaps to an urban use case such as robo taxis, then it can become much more attainable. The downside is that limiting application is likely to reduce the private vehicle take-up. There remains a substantial difference of opinion among many in the driverless development community about how important different AI capabilities are. For instance, Waymo has dedicated significant resource to understanding the hand gestures of traffic policemen, whilst others assign lower importance. There still appears to be no concrete approach to walk from conventional programming with limited machine learning to decision making AI that can deal with complex edge cases (such as Google's famous ducks being chased by woman on a wheelchair). If one does exist, it has not been submitted for peer review. Cross domain learning seems like a big part of the answer. Instead of AI trying to understand hand gestures by policemen, why not give control to a remote operator, or even re-route and avoid the problem altogether? It seems almost certain that V2V and V2G communication is necessary. Extreme weather conditions, domain shift (changes in road layout) and complex traffic may all be too difficult for a single vehicle operating alone to overcome. It is also unclear whether the right approach is a small number of very capable systems or a larger grouping of software modules with clearly defined roles. It also seems that even today's state of the art may not be good enough for the real world. Due to closure speeds, cameras rated for object identification on highways could need framerates of 100hz to 240hz to be capable -- this requires more powerful hardware. At the same time, OEMs want components that use less power. Selecting the right componentry also requires appropriate benchmarking to be developed. Camera systems cannot simply be assessed in terms of framerate and resolution; latency and power consumption are also important. Audi is undertaking extensive comparison of learning techniques, hardware and processing methods. Some correlations are appearing: hardware specific processing appears better than generic methods; using real and virtual and augmented learning seems to improve decision making, but not in all analysis models.

Lessons learned from different object recognition experiments... self-parking systems are in production, as are research fleets testing automated highway and urban driving technologies. AI's task can be divided into four elements. First is classification (recognising an object on its own). Second is object detection (spotting it within a scene). Third is scene understanding. Finally comes end-to-end (working out how to safely drive through the scene). Training is difficult due to limited availability of off-the-shelf data. There is none for ultrasonic, cameras or fish eye and only a small amount for laser scans. Experiments continue on the best way to develop scene understanding. Lidar products can detect most vehicles at a range of 150m, but how should that data be handled -- as single points or pre-clustered? Should object detection be 2D or 3D? Researchers are trying to develop processing that can spot patterns in point clouds to identify objects but is also intelligent enough to interpolate for point clouds of lower resolution (e.g. recognising objects with a 16 line lidar that were learned using a 32 line data). Determining how best to use lidar data in concert with other sensors is a subject of ongoing research. For instance, OEM opinions differ on the minimum points per metric metre requirement.

Simulation, testing & validation of artificial intelligence

Key Takeaways:

- Self-driving AI requires both innovation and functional safety -- may need to tilt today's balance more towards safety
- With a greater volume of properly annotated data, new machine learning techniques can be employed

Safety considerations for self-driving vehicles... Autonomy depends on collaboration between two very different disciplines: functional safety and artificial intelligence. Functional safety is associated with strong discipline, conformance to protocols, standardisation and making things purposefully boring. Artificial intelligence is highly disruptive, innovative and uses multiple approaches. Recent crashes by vehicles in driverless mode show that AI could be more safety conscious -- that is not to say that AI is to blame for the example accidents, but that if the vehicle had driven more cautiously the accident might have been avoided. For AI systems to be approved by regulators it is likely that they need to: lower their exposure to accidents; act in ways that other road users find highly predictable; reduce the likely severity of collisions and; increase their control over the actions of surrounding 3rd parties. Self-driving vehicle creators must be able to explain the AI to regulators and the public. AI must be robust to domain shift (e.g. either be able to drive equally well in San Francisco and New York or be prevented from doing actions it cannot complete properly). AI must act consistently take decisions that can be traced to the inputs it receives.

Research into machine learning techniques that improve image recognition... performance is improving but machines still find it very difficult to recognise objects because they do not look at pictures in the same way as humans. Machines see images as a large string of numbers, within which are combinations of numbers that form discrete objects within the picture. Learning what objects are is therefore not intuitive; it is pattern based and machines require large datasets of well annotated data in order to gain good recognition skills. Researchers have been developing a concept called "learning by association". The key innovation is that beyond a comparison of a new and unlabelled image to an existing labelled image, the associations identified by the AI are then compared to a second labelled image to determine the confidence of a match. Training in this way led to enhanced results in tests and an improvement in recognition of a brand-new dataset that was added without a learning stage.

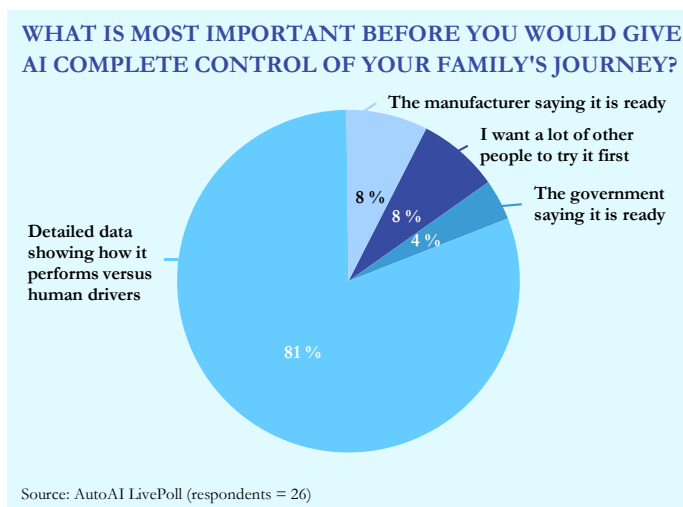
Live Poll Results

Attendees were periodically asked multiple choice questions which they answered using a smartphone app.

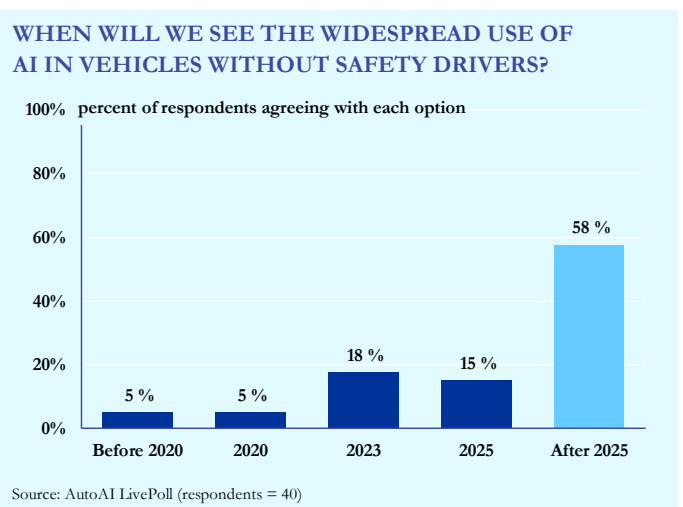
Key Takeaways:

- Most attendees want to judge a self-driving solution for themselves
- Although some attendees believe self-driving will be a reality soon, the majority think it is post 2025
- The biggest challenges were seen to be the lack of clarity from regulators and the need to test more
- Attendees overwhelmingly believe that no one knows how much real-world testing will be required
- In the face of a new and difficult situation, most attendees thought the vehicle should simply stop safely
- Attendees expressed a clear preference for object recognition by multiple sensor types over camera or lidar alone
- Only a minority see all the data being collected as constantly necessary, but most thought it was sometimes required
- There was complete agreement on the need for redundancy, but a 50:50 split between high and low capability

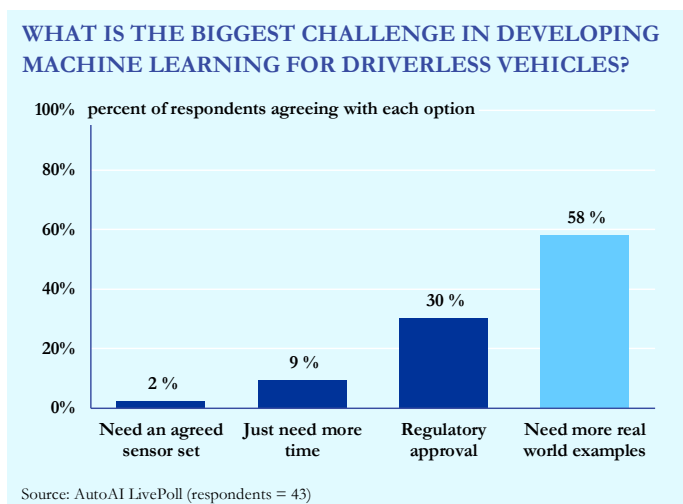
What can I do to make you love me?



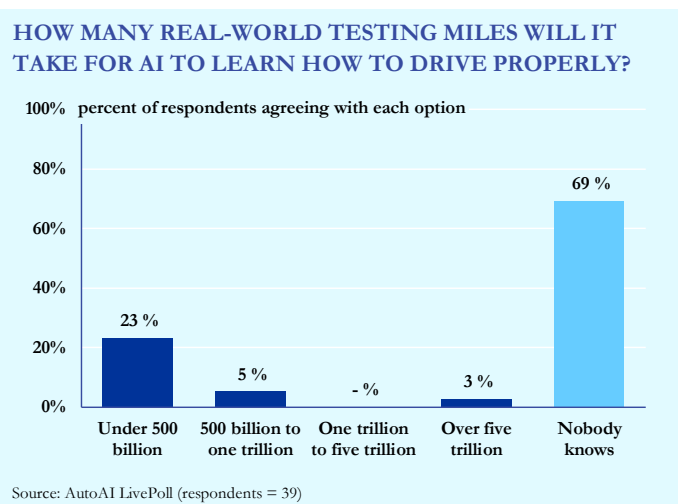
When will I be famous?



Please release me

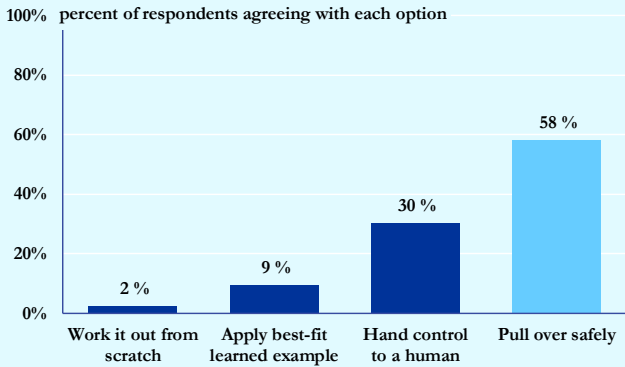


Life is a highway



Stop right now, thank you very much

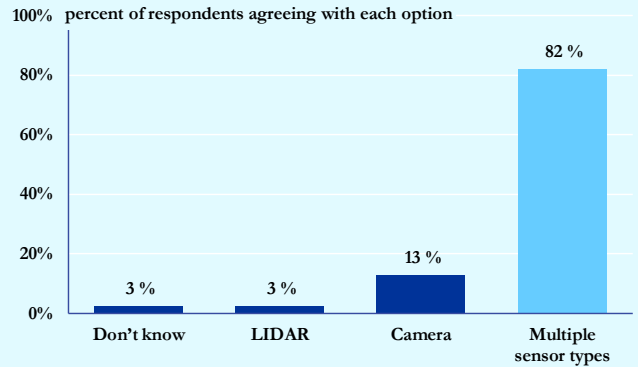
WHAT IS THE BEST THING FOR AN AI TO DO IN A NEW AND DIFFICULT SITUATION?



Source: AutoAI LivePoll (respondents = 43)

Hello! is it me you're looking for?

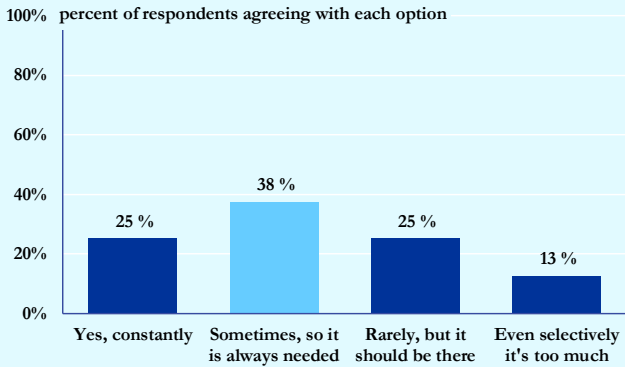
WHAT SYSTEM IS BEST FOR OBJECT RECOGNITION?



Source: AutoAI LivePoll (respondents = 39)

I want it all, and I want it now

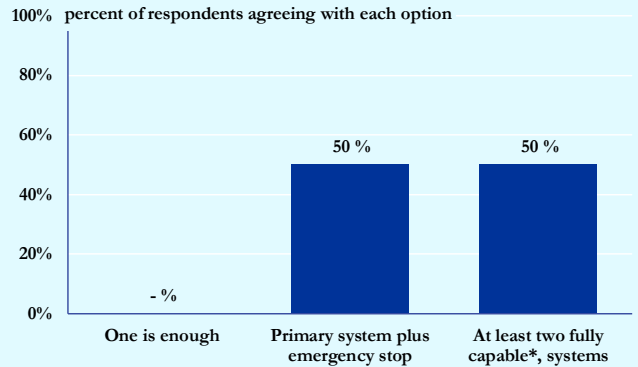
WILL AIS ACTUALLY USE ALL THE INFORMATION A COMPREHENSIVE SENSOR SUITE* CAN PROVIDE?



Source: AutoAI LivePoll (respondents = 24) *e.g. 4 x 64 line LIDAR + camera + radar

It takes two baby

HOW MANY LEVELS OF REDUNDANCY WILL AUTOMOTIVE AI NEED?



Source: AutoAI LivePoll (respondents = 26) * and independent

In Closing: A Summary Of This Report

- **No one seems to be advocating a rules-based approach for autonomous driving.** Although possible in concept, a set of detailed rules appears impractical in the real world because the volume of instructions needed to cover all driving conditions would be massive and debugging would likely fail to spot all errors or contradictions.
- **There needs to be a common view of the acceptable safety level for deployment of autonomous vehicles.** This requires the input of regulators, industry, academia and probably 3rd party consumer groups too. In addition to the safe level, we need to determine how to measure it and confirm compliance -- the second part is challenging without a test fleet covering tens of billions of miles.
- **Safety is related to the use case.** The more complex you make the driving conditions, the more corner cases you'll encounter. The scale of the problem can be reduced by steps such as: learning a particular area (geo-fencing) and not driving in extreme conditions or at very high speeds. Although this diminishes the value for the retail user, there are plenty of industrial applications that can operate within those constrained operating conditions.
- **There needs to be widespread collaboration on a shared framework for testing and validation of AI.** Governments, companies and academia should all be involved and it would ideally use open data that was not tied to specific simulation tools or sensor sets. The experts at the event were wary of letting driverless cars to transport their family without seeing safety data beforehand (government and manufacturer assurances weren't trusted).
- **Work needs to be done on explaining AI.** There are big differences between the capabilities non-technical people think AI has and what it is capable of -- there should be no talk of killer robots. At the same time, deep learning techniques mean that the system function cannot be explained in the same way as traditional coding. New ways to explain how the system operates are required and without them building trust will be very difficult. It could even be necessary for the AI to learn how to explain itself using natural language or other tools.
- **Virtual testing is vital.** This is for three reasons: firstly, simulation dramatically decreases real world miles; secondly, because AI techniques like reinforcement learning need crashes to take place in order for the AI to learn; and thirdly because even real-world data becomes a simulation once you interact with it in a different way to the original situation. It's better to do that virtually! For a virtual environment to be successful it must be something that can be replicated in the real world with the same results.
- **There is plenty of disagreement over the right approach to many areas.** The event live poll highlighted differences of opinion regarding how AI should act, how much information it will be capable of processing and what level of redundancy was required for safe operation. More consistent was the high burden of proof that AI systems will be faced with and a view that currently no one really knows how to convincingly do that.
- **Implementation timing remains uncertain.** In the event live polling, over a quarter of respondents believe that self-driving will be widespread by 2023 or even earlier. The majority believe that we will be waiting beyond 2025 -- a healthy difference of opinion.

About AutoAI 2017

Auto.AI is Europe's first platform bringing together all stakeholders who play an active role in the deep driving, imaging, computer vision, sensor fusion and perception and Level 4 automation scene. The event is run by we.CONECT Global Leaders, a young, owner-managed, medium-sized company from the heart of Berlin and subsidiary office in London. The next AutoAI USA conference runs from March 11th to 13th 2018 and the next European AutoAI conference takes place from September 16th to 18th 2018.

About Ad Punctum

Ad Punctum is a consulting and research firm founded by an ex-automotive OEM insider. We bring focused interest, an eye for the story and love of detail to research. Intellectual curiosity is at the centre of all that we do and helping companies understand their business environment better is a task that we take very seriously.

About The Author

Thomas Ridge is the founder and managing director of Ad Punctum, based in London. You may contact him by email at tridge@adpunctum.co.uk.



AD PUNCTUM

To find our latest research, please visit www.adpunctum.co.uk

Follow @Ad_Punctum on Twitter

© Ad Punctum 2017. All rights reserved.

Publication Number: 171-BRA-P000030