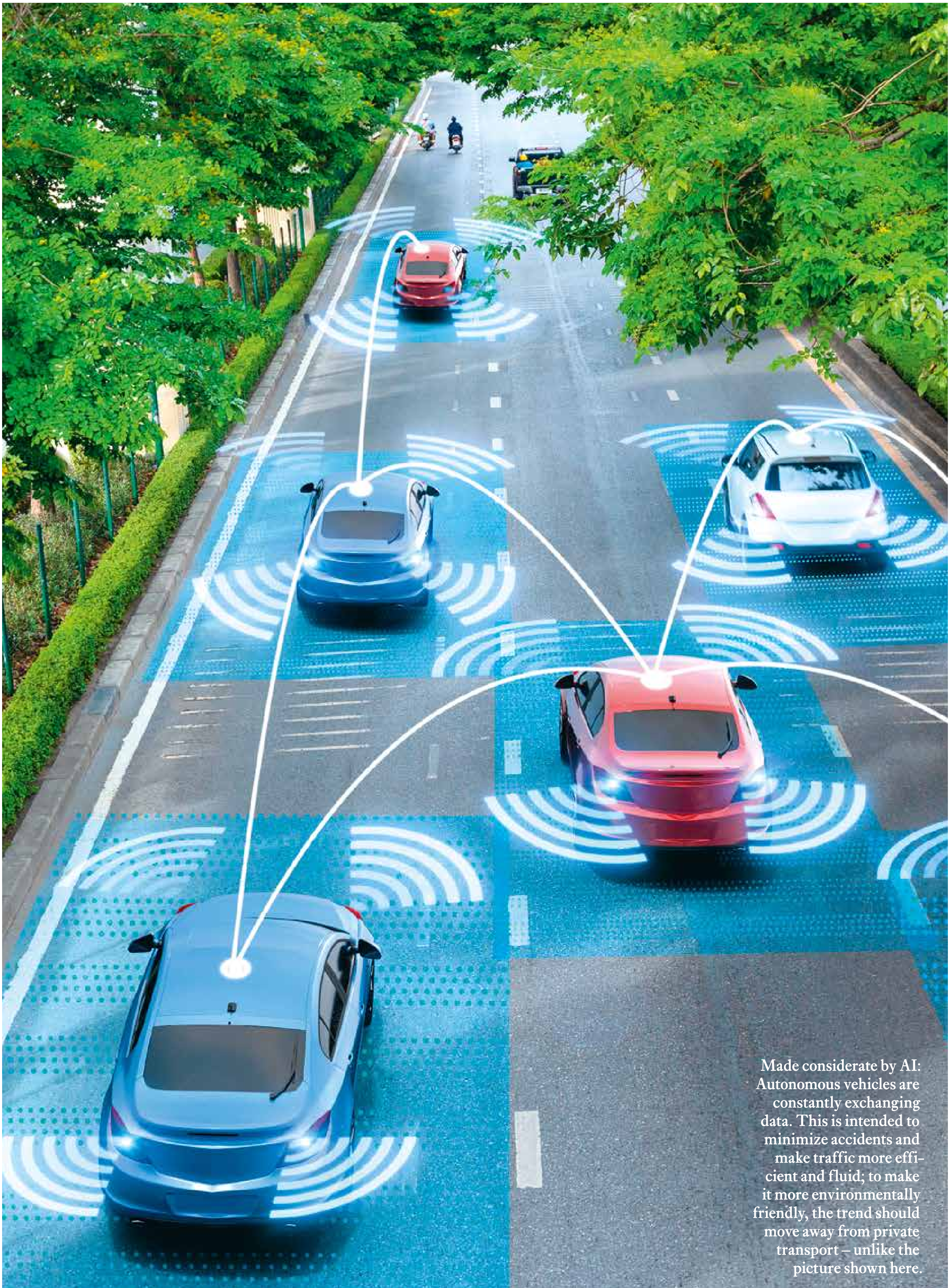


IMAGE: SHUTTERSTOCK/SUWIN66

Made considerate by AI: Autonomous vehicles are constantly exchanging data. This is intended to minimize accidents and make traffic more efficient and fluid; to make it more environmentally friendly, the trend should move away from private transport – unlike the picture shown here.

CARS WITH INTERNAL VALUES

TEXT: THOMAS BRANDSTETTER

The first self-driving cars are already on the road. Yet the technology is not yet fully developed, and certain ethical issues remain unsolved. It is also high time to think about how the new technology can promote sustainable transportation in the future.

Expectations for a mobility revolution caused by autonomous vehicles are high: fewer accidents, more sustainable transportation, and enhanced comfort are among the anticipated benefits. Autonomous vehicles are expected to improve safety and sustainability in traffic, offering convenience to those who find driving stressful and nerve-wracking. Although there are still no driverless cars on the roads in Germany, the trend seems to be continuing unabated. While an autonomous fleet from Google's sister company Waymo is undergoing practical tests in San Francisco, Hochbahn AG, the operator of Hamburg's local public transportation system, has taken up the cause. Together with the on-demand service Moia and vehicle manufacturers Holon and VW, the Hamburg-based company plans to augment its fleet of approximately 1,000 buses with a fleet of autonomously operated vehicles in the coming years.

However, there are still a few problems to be solved before robotaxis are in widespread use. The idea of an autonomously operating machine that may have to make life-or-death decisions in road traffic also raises significant ethical questions. There will likely be a need for a widespread social discussion concerning the balance between the lives of a vehicle's occupants and those of other road users. What's more, the environmental benefits of the new technology are still unclear. However, the most pressing problems right now revolve around technical shortcomings. In San Francisco, for example, there are increasing complaints about driverless cars that simply stop in the middle of the road for no apparent reason and block the path of emergency vehicles.

"When it comes to autonomous driving, it is crucial that the systems also learn to recognize complex traffic situations and then make the right decisions," says Bernt Schiele, Director at the Max Planck Institute for Informatics, whose research interests encompass computer vision. Together with his team, he has developed the AI model

→

Mobility for tomorrow: This all-electric on-demand shuttle leaves Hamburg's Elbphilharmonie concert hall behind during a test run in July 2021. As part of the Alike research project, a newly designed vehicle will take up to 15 people each to a destination of their choice starting in 2026 – and from 2030 onwards without a driver.

PHOTO: HOCHBAHN

MTR++, which can increase the efficiency of training for autonomous vehicles when it comes to assessing and responding to the behavior of other road users. The researchers have already won two first places in Waymo's Motion Prediction Challenge – a competition in which different AI systems are trained on an identical data set and their predictions are then compared. For example, if several vehicles are approaching a junction from different directions, there is a multitude of possibilities as to which one will turn in which direction and when, and exactly which path a vehicle is following. "In a situation like this, each driver has a relationship with the next and influences their behavior," says Schiele. "To train the systems, we use transformer models, which function similarly to those powering large language models."

Artificial intelligence (or more precisely: machine learning, which is what Schiele's research group uses), essentially relies on artificial neural networks. These networks, which bear similarities to biological nervous systems, optimize themselves through extensive training data until they recognize correlations and consistently produce accurate results. To achieve the recent remarkable accomplishments of ChatGPT and similar systems, this concept was expanded with the capability to efficiently correlate and relate various elements to one another. While their predecessors had to analyze sentences word by word in order to fathom the structures of human language, the transformer architecture allows the new language models to look at entire sentences as a whole. The transformer simultaneously determines the extent to which all of the words are related to one another. This allows the algorithm to learn what individual words mean in context. Invented in 2017, this method has not only revolutionized computa-

SUMMARY

The technology behind self-driving cars is not yet fully developed. Scientists like Iyad Rahwan are also researching how AI should prioritize human lives in road traffic and how trust in this technology is created.

People are more likely to trust the AI if it is clear why it makes certain decisions. In light of this, Bernt Schiele is developing neural networks that better recognize and interpret complex traffic situations.

Autonomous vehicles could make traffic safer. But they will only make it more environmentally friendly under certain conditions, such as when on-demand shuttles reduce private transportation.

tional linguistics, but has also been adopted in computer-based vision. Just as the meaning of a word depends on how it is embedded in a sentence, the behavior of road users is also determined by the overall traffic situation. In light of this, the systems analyze different traffic situations and try to understand the behavior of other vehicles. “A language model generates the next word in a sentence based on a given prompt,” explains Schiele. “And based on the input of where you want to go, our system predicts the possible trajectories of all vehicles involved at an intersection.”

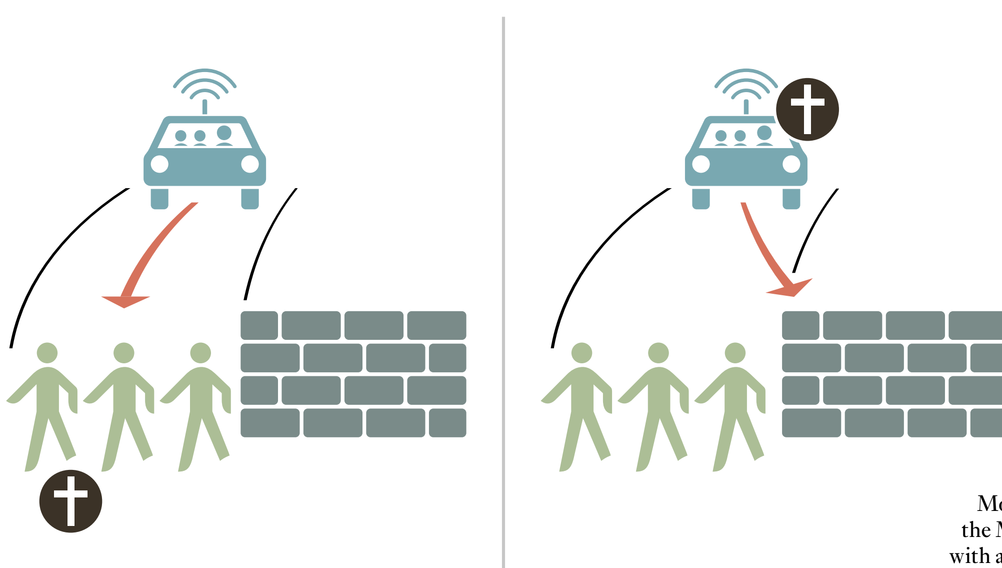
In this case, the AI processes all information about vehicles and the road at the same time. The training is therefore much more efficient compared to previous methods, in which the individual motorists in road traffic and their relationships to each other still had to be modeled separately. “When these processes occur simultaneously, it grants the network greater flexibility in its learning. And that makes the learning process much more powerful,” explains Schiele. “However, we no longer know exactly how the network has modeled the relationships.” The fact that the results of artificial neural networks cannot be fully explained is a fundamental problem. While classic computer programs allow for a line-by-line understanding of how the code operates and its outcomes, machine learning involves understanding the algorithms that train a neural network; however, the exact process by which the network, or model, arrives at its results largely remains opaque. The training data essentially provides the outcome, and the AI algorithm optimizes a highly intricate function with numerous parameters. The goal is to design the training process so skillfully that the model also performs well in practical applications.

Explainable decisions

This approach works effectively for straightforward classification tasks, where a neural network can learn to recognize what defines a bicycle, for instance, after being exposed to a sufficient variety of images of bicycles. However, occasionally something goes wrong during training and the network does not base the classification on the actual object, but on other image content that is related to it. “In one of our analyses, we discovered that sometimes the system identified bicycles only when a person was sitting on them,” says Schiele. Such errors, which are based on correlations with other image content, are difficult to detect with conventional neural networks. “The result is correct, but for the wrong reason.”

This is why Schiele and his team are developing special neural networks that also allow for troubleshooting. “Currently, networks are primarily trimmed for maximum performance and not for explaining their decisions,” says Schiele. The Max Planck Institute’s new system, on the other hand, not only delivers the result of what it thinks it recognizes in an image, but also shows exactly which pixels of the image were relevant for this decision. This allows it to explain its decisions to the user during operation. And if it highlights image content such as the person on the bicycle rather than the bicycle itself, this immediately signals a problem. The researchers also discovered another insufficiently trained network that only recognized a road

CHART: GCO ADAPTED FROM MORALMACHINE.NET



Morally unambiguous? In this example, the Moral Machine presents participants with a choice: should an AI steer a car with two children and one adult on board into a group of three adults, or is it better to drive into the wall?

as such when a car was driving on it. "When we took the car out of the picture, the network no longer saw the road," reports Bernt Schiele. Such errors can only be identified at an early stage, when networks can be interpreted. "Or at least it helps to uncover the mistake in retrospect if something went wrong, perhaps an accident has even occurred." For Iyad Rahwan, Director at the Man and Machine Research Unit at the Max Planck Institute for Human Development, this is an important prerequisite for the acceptance of autonomous vehicles. "In order to build trust in this new technology among people, it's essential that we are able to question the cars about their decisions," says the researcher, who focuses on ethical issues at the intersection of computer and behavioral sciences. "After all, it makes a difference whether a vehicle has simply not seen a cyclist, for example, or whether it has decided to take a certain risk." The former would be a purely technical problem, which, of course, needs to be resolved as quickly as possible. The latter, on the other hand, could be the result of an assessment by the AI as to whether it places the protection of its own occupants above that of other people involved.

With their platform "Moral Machine," Rahwan and his team have pushed the life-or-death decision scenarios that the AI may face in road traffic to the extreme. Anyone who is interested can take part in the online study. Using simplified, illustrative depictions of dangerous traffic situations, the researchers ask, for example, whether an autonomous vehicle that has no possibility to swerve should run over a child or two adults. Or whether the AI at the wheel should instead crash the car into a wall beforehand and thus endanger the lives of its own occupants. "Of course, such situations will only occur very, very rarely in reality," acknowledges Rahwan. "But people can't stop thinking about questions like these. And if we want them to use autonomous vehicles, we have to provide answers."

Over three million people worldwide have now taken part in Moral Machine and entered 40 million decisions into the system. If you want to evaluate a traffic scenario yourself, you can do so online. Some findings were predictable, such as the prioritization of children's lives over adults'. There was also general agreement that as few people as possible should come to harm, even if this is at the expense of the occupants. The problem with this, however, is that nobody wants to sit in such a vehicle themselves. "There is a huge tension between what people think is right and what consumers are prepared to accept," says Rahwan. This poses a societal challenge, as it's assumed that widespread adoption of autonomous vehicles would substantially reduce overall road fatalities. "It will be important to convince people to think

PHOTO: WAYMO

Waymo's autonomous vehicles have three types of sensors, each of which performs its own tasks. LiDAR laser sensors and a radar system scan the surroundings in three dimensions, whereby the radio waves from the radar can also penetrate the Hamburg fog. A 360-degree camera on the roof monitors and categorizes the surroundings visually.

not only about their own safety, but also about the common good," says Iyad Rahwan. The striking questions posed by Moral Machine may be exaggerated, but there is no question that autonomous vehicles have to make important decisions. "In reality, it will mainly be about statistical issues," says Rahwan. Programmers have to decide whether a vehicle should drive in the middle of a road in normal operation or closer to the edge. This automatically influences the likelihood of either running into oncoming traffic and causing a serious accident or perhaps instead exposing those who are cycling on the side of the road to greater danger. "Such decisions will be reflected in the accident statistics in the long term," warns Rahwan. "If we want people to accept autonomous vehicles, we must also ensure that the safety benefits they provide are distributed fairly."

Less individual traffic thanks to AI

In any case, there seems to be a consensus that autonomous vehicles would make road traffic safer as a whole. After all, the vast majority of accidents are the result of human error, caused by people driving too fast, carelessly, or even while drunk. Whether and, most importantly, in what manner autonomous driving will be adopted is a different question. "The hope among us researchers is that shared-use autonomous vehicles will gradually replace private cars," says Michael Krail, who heads the Mobility business unit at the Sustainability and Infrastructure Systems Competence Center at the Fraunhofer Institute for Systems and Innovation Research.

This could significantly reduce the contribution of road traffic to global warming, especially through ride pooling, where multiple requests from passengers heading in roughly the same direction are consolidated. "If one of these vehicles were to pick you up at your doorstep, it would have the same functionality as your own car," says Krail. "This could also be attractive in terms of pricing and could even compete with public transportation." After all, the most expensive thing about taxis or shuttles with a small number of passengers is the driver. The Alike project in Hamburg will initially have attendants in the vehicles, but by 2030 at the latest, the fleet of minibuses should be operating autonomously on the roads. According to the Karlsruhe Institute of Technology, this could potentially replace 250,000 cars. By then, it should be possible for all Hamburg residents to access public transportation within five minutes. According to Krail, well-coordinated combined solutions would also be an option. The taxi could take you to the train, for example, and a shuttle would be waiting at the destination station to continue the journey. This is the exact goal of the EcoBus system developed by a team from the Max Planck Institute for Dynamics and Self-Organization. The system is currently undergoing testing with various transport companies (see Max Planck Research 1/2019).

But how sustainable is a single autonomous vehicle on the road compared to a conventional car? It is assumed that the algorithms will facilitate predictive driving, thereby potentially saving fuel or electricity, depending on the type of propulsion. And ultimately, when all vehicles are autonomous and connected, this should improve traffic flow and lead to additional savings. On the other side, artificial neural networks, which rely on numerous sensors to perceive their surroundings and make driving decisions, consume a lot of energy themselves. And the computer hardware needed for current prototypes often fills the entire trunk. "At the moment, this additional electricity consumption is still a major problem," says Krail. However, the shift towards electromobility is also compelling manufacturers to rigorously manage power consumption in the future. Failure to do so could result in reduced vehicle range and diminished market appeal. "It is therefore reasonable to assume that future vehicles entering the market will require less power for code processing and data transfer," says Krail. What remains, however, is the concern about how the new technology will actually be used in the future. Even if autonomous vehicles turn out to be safer, more comfortable, and possibly even cheaper, this does not automatically lead to more sustainable mobility. In the worst-case scenario, privately used autonomous vehicles could simply replace cars one-to-one, dropping users off at their destinations. If they cannot find a parking space there, they may even drive around with no passengers until needed again. "In order to prevent such excesses, we should monitor developments very closely and create suitable conditions in good time," warns Michael Krail. After all, the new technology should be part of the solution and not become another problem. ←

GLOSSARY

MACHINE LEARNING
classifies large volumes of data, recognizes motifs in images, and makes predictions. People train a model by manually marking picture elements and indicating their meaning. How well the finished model independently recognizes image elements in unfamiliar images depends on the quality of the training.

NEURAL NETWORKS
are trained in deep learning, a sub-area of machine learning. The network processes data such as pixel values through many layers. The upper layers recognize simple features such as edges and shapes, while the lower layers recognize concepts such as "road" or "human." Advantage: the model independently recognizes complex features in unfamiliar images. Disadvantage: the more complex the network is, the more difficult it is to understand the results.