Waymo

**Drago**

So I was tech lead of 3D vision and post estimation for Street View and got a lot of exposure to all kinds of platforms that the team was collecting data with, including boats, snowmobiles, backpacks, and of course cars. And afterwards, I got back a bit closer to my machine learning AI roots. I worked on the team and eventually ended up leading a team that developed, even in those early years 2013 to 2015, large neural networks for understanding images. And so in those times, we had a very exciting set of developments and architectures between Vantage and we won the ImageNet challenges in 2014. That used to be a very popular competition on object understanding and detection and classification. So we won those two challenges in 2014 with the team that I lead and some collaborators from Google Brain. And we also work a lot on launching the backends that would annotate Google Photos, Google Photos was first of its kind to offer deep semantic annotation, photos enabling search, and all the photos applications that they would do, where they suggest you various events and albums and so on and organize them. That was based on backends and models that we developed at the time, and that were launched in the Google data centers. And around 2015, when I felt that deep learning has entered the level of maturity that is sufficient to potentially make autonomous driving a reality, I got back into the autonomous vehicle space, into robotics that I had been exposed to in my early lab days. And so I joined initially Zoox for two and a half years. It's an autonomous driving startup. I led the perception team there, helping them build out their, essentially 3D understanding and the machine learning around it. And then I came to Waymo, right. So that's my story more or less. I would say I'm a machine learning person. And my initial expertise had been on understanding images and 3D data, geometry. And that's generalized to other applications, since autonomous driving is a, it's a complete robotics stack, which includes all the key robotic systems, perception, planning, behavior, prediction, simulation, mapping, localization, and so on. So, by now I have this maybe more of a cross functional exposure to all of this through the lens of machine learning.

**CRAIG**

Your undergraduate was where? And then at Stanford, you were studying under Daphne Koller. And what was that work about?

**Drago**

Actually I did my undergrad at Stanford too but afterwards, I continued into a PhD that took six years. So I spent a total of 10 years at Stanford, it’s hard to get me to leave for a while. I worked closely with Sebastian Thrun because I was Daphne's first PhD student that worked on computer vision perception topics and we wanted access to robots and data from robots. And at the time, Sebastian wasn't even at Stanford. But Daphne somehow made a connection with him. And I ended up visiting his lab in Carnegie Mellon and collecting data with robots and shorting one robot. I'm a software person, unfortunately there's things I learned the hard way there. But Sebastian was always a commenter of sorts, even though my primary advisor was Daphne. I think I've always worked with Sebastian as well, and we have joint works and he was a key influence too.

**CRAIG**

Yeah. And so Waymo, when you got involved with the company, where was it in its development and where are you right now? I mean, I know you have robotaxis operating in San Francisco and Phoenix and I hear you're gonna have them operating in LA.

**Drago**

Yeah, we have San Francisco and Phoenix today. You can experience autonomous driving it exists at scale in these cities. We give 10s of 1000s of rides to users in both cities every week. And Waymo overall, has given our a million of paid rides, driving over 10 million miles autonomously as well. So that is the two main markets that we have. The Phoenix area that we cover today is around five times the size of San Francisco. And next, we're looking to expand and start actually paid rides. In the coming weeks in Los Angeles. We got a permit a couple of weeks ago, to expand to Los Angeles and charge and so we will start doing this. And we also started doing driverless rides for employees in Austin, Texas. So we're going to be expanding in this fourth market that we have announced as well. Now, my time at Waymo has been a wonderful time, seeing a transformation of the stack and getting to maturity that I think has been a great experience, and working with the talented team. I think when I joined Waymo in 2018, In the summer, Waymo had hit some great milestones. One of them was, in 2015 it gave the first fully autonomous ride in Austin actually, a city we're getting back to. The person called Steve Mahan, friend of the company, as a blind person, the first ride involved him. And so we had done that and we were working towards our first commercial deployment. And so we were evolving the stack to make sure that it will be a safe and scalable deployment. And this eventually happened in 2020, when we launched our Waymo One service in Phoenix, East Valley, around Chandler. And so that's been the journey. So, evolved the first fully autonomous service in Phoenix, East Valley, then get to San Francisco, and then significantly upgrade the stack to be able to handle the complex, dense urban environment of San Francisco. And now we continue to all these, more cities expanding in Phoenix, we will be working also on tackling more highways, which is important in large areas like Phoenix. Our area, if we don't go on highways takes almost an hour to cross on certain streets today. So, that's I think, maybe the trajectory of the last five and a half years. It's been tremendous also in terms of development of machine learning and machine learning technology. So in 2015, when I joined, oh 2018 Excuse me, when I joined Waymo there was already a lot of impressive machine learning work. Especially, there was a lot of really interesting, unique work on 3D perception using LIDAR. And this was in collaboration with Alex Krizhevsky, and the Dali Angelo**,** and other people from Google Brain, Alex Krizhevsky being of the 2012 AlexNet fame. So he took a lot of interest in autonomous driving as an application. And he worked a lot with Waymo and developed some really interesting work, including inspiration to a first paper Waymo ever did, even though we published it in 20, early 2019. It's called ChauffeurNet, it's actually a model, a train model to predict and drive by imitating others. That was really groundbreaking, unique work at the time. And then that was still relatively early in its development and maturity, but it was very exciting as a concept. So we publish that and we took all these technologies, and we worked a ton on expanding them and transforming Waymo stack. I think at this point Waymo is an AI first company. So we strive to maximize the impact of machine learning and its power across most of our applications.

**CRAIG**

Yep, I haven't spoken to anybody about autonomous driving for a while. I did an episode with Alex Kendall, at Wayve AI, but we were really talking about world models. In the in the tech stack is Waymo, from what I've read, you have detailed custom maps for navigation, you have the suite of sensors, and then you have neural nets that are kind of wired together to process that data and guide the car. How much hard coded stuff is in that stack, and how much is pure machine learning?

**Drago**

So big parts of the whole stack are machine learning. And this spans all the core systems of Waymo. This spans things like the perception system. And it also covers major models, applying machine learning to planning, behavior prediction or others and simulation, and also lots of machine learning for employed also in evaluation extracting signal from, for example, simulation or other tests we perform on the car. So at this point, I would say all major parts of Waymo have powerful, large machine learning models underpinning them.

**CRAIG**

Are all those models on the car, or what's on the car, what's in the cloud?

**Drago**

So we have some flexibility. I think, especially the whole evaluation system simulator, that is in the cloud, right, because we develop a variant of the stack, and then we need to test it. And we tested in the data centers to a large extent, even though we also tested with human safety operators. And we want to have specialized areas where we can stage certain rare situations. But testing in the cloud and simulation is key. I think generally we avoid requiring a human to drive our cars at anytime. It's possible to ask a human occasionally, specific question in some very, very rare situations when preferably the vehicle is usually, almost always, the vehicle is stopped at that time, but that is also extremely rare. And we're trying to minimize this. So the vast majority of the system that drive the car, they're on the car on the local compute, and they operate fully autonomously as advertised.

**CRAIG**

And so how many models or neural networks are involved in the driving system?

**Drago**

So there are several, right. I think the history of our space is one of consolidating models, right. And traditionally, many players in the space, and also when machine learning was in its relatively early stages and there was limited computing data, it has been beneficial to, or at least at the time, to partition the tasks and have maybe one neural net to detect bounding boxes and another to segment, say, the scene into semantic segmentation of various kinds and another maybe—. So traditionally, that was the history. And now, what's changed is we understand the power of what, the models have become a lot more powerful, they're scaling properties have become much improved, especially with the advent of transformers, we understand a lot better how to make models do well on many, many tasks at the same time. That's actually been an area of research by us and the whole community. And some of these AI trends can be leveraged to train very large models, to transfer knowledge from the internet, and fine tune on your tasks. So we have ways to make models do well and a lot more tasks today. And there is benefit then, in efficiency, in simplicity, and quality usually, to consolidate. So there has been a push to consolidate. That said, it's an ongoing process. There's some very interesting question whether you should have one model or several, there is challenges in having just one model in our case, especially if you need to iterate on it and ensure it satisfies I mean, hundreds to 1000s of requirements, right. You need some way to be able to do work on parallel problems in this domain, to ensure safety. And so there is this constraint. And there is some other constraints that make it difficult for our case, for example, compared to say, traditional robotics, maybe you talk to Peter Novig or Vincent Vanhoucke they have it on a robot, and then they have a world model and they can train it end to end. For us, it's more challenging infrastructure wise, because we have, say, over a dozen camera to two dozens, we have a few LiDAR, we have a few radar fitting all this data for all in one on the computer and scaling it is quite daunting. And that even motivates potentially training things in stages and so on, right. So I would say that the answer is we are evolving, and we are evolving, satisfying, both when the learnings in the industry and the requirements in production, and generally the trend is to consolidate.

**CRAIG**

Yeah. And I know it's a very complicated system, but can you can you sort of walk us through how the car drives? I mean, they're the sensors and they're feeding the data into a perception model, and then the perception model, I mean, just kind of walk me through the pathway.

**Drago**

So, the simplest way to think is the onboard system, right. And there's a lot of very interesting problems and questions when you build the simulator as well. But let's talk about, which is also machine learning task to me, by the way, I'm a machine learning person and ultimately, we want to solve problem with the data we have as much as possible. But the traditional autonomous vehicle stack takes data from sensors, can be camera, LiDAR, or radar. And then there is a perception stack that processes this data, and creates a representation of the world. And this representation of the world, one of the key aspects of it is it wants to be expressive, but also concise, right. And in that representation, then you train your behavior and potentially simulation models because that's a representation of the world, and then you need to in that representation of the world, you need to plan what you're doing. And you need to also potentially imagine what the world is going to do in response to your plans to validate that your class can be safe. So essentially, there's this main two parts to the stack perception produces a model of the world, and then in that model of the world in plan, you validate your plans, that's the behavior system. And the model of the world is very interesting. So there is a lot of interesting question of what it can or should be. Traditionally, we have opted for at least partially interpretable, intermediate representation. So one of the things that has been very beneficial in our space, and I think it's generally superior to treating all the sensors as a, an ordered collection of sensors is to take their data, fuse their representations in the knowledge provided to them, and create what I call a bird's eye view model of the world. So you, essentially, you center a grid, or implicit grid on the space around the vehicle, and you essentially populate this grid with everything, all the information that you require from the sensor. So that's been very popular and there is certain benefits to having interpretable representations, people understand them, they know what the problems are, they can annotate or correct them. And so, the drawback of those is you need to label data to create them typically, they're representations that we have chosen, because they satisfy requirements interpretability, testability, right. Potentially people can look at the outputs of a system and impose further constraints or corrections that's why you asked me, is all your system just a neural net, our system is hybrid even though big parts of it are neural net and we strive to solve the majority of the problem with machine In learning, we still build such a presentation and design where you can control and steer this system by experts, right, that can introspect the results and introduce additional constraints requirements on top of what the machine learning models do. And that has been very beneficial for robustness and safety.

**CRAIG**

I guess that's where Wayve is different. They, they're creating a world model and then everything happens, all the planning happens in that, the simulation and planning happens within the world model.

**Drago**

I actually don't think they’re that different, honestly. I think that it was oversold. I have a perception system, it produces a compressed version of the world, in our case, some of it is interpretable but you can also have various embeddings or quantized embeddings of tokens, trained end to end that can also be produced, right. It's a question of choices in that representation. And then given that representation, you have prediction models, and these are you can call them, they’re world models. And there is a question of which exactly world models do build? Do you build the model that just dreams all the video, all your cameras? Well, you can but on board, maybe that is overkill. So then which things do you choose to dream and autoregressive or predict concisely so that you still satisfy your requirements, that's a very heavy way. But we still have major machine learning models that are still word models, just the design of them may be different. And we also can train end to end. However, right, we have also put historically a lot of focus on having intermediate representations because we have found they're very important so far, in actually building a fully working system. Machine learning is great, but it has limitations, there's always cases where eventually you want to impose guardrails of constraints and intermediate representations are very powerful for doing it. Now, lately, there's been yet one more trend with large language models and the like. Now there is a lot, kind of language based representations. So language compatible embeddings are very, very popular for robotics because they bring a ton of knowledge from the internet and the that said, the model, the language models that they trained on, which makes it, I mean, the system understands a lot more common sense and be able to interact with humans. So now that is also very interesting to us, right. So language or language compatible representations are great. That said, there is also another learning in our domain, you really need these representations to be spatially correct. So our models, perception system needs to be very powerful in spatial reasoning because spatial reasoning, the ability to represent the world and ultimately evaluate your plans in the world is key to ensuring safety. And so these concepts are now overlapping and I think we have a lot of exciting, both experience, expertise, and of course, we're also experimenting with the best mix of these concepts.

**CRAIG**

Yeah. I mean, I just had Vincent Vanhoucke on the podcast talking about RT2. And in that system, which is amazing, and I mean, stuff is advancing so quickly, how do you—

**Drago**

We’re very much inspired by their work and follow it. And I think there's a lot of parallels to our domain. There is certain differences where we are a lot more concerned about correctly predicting and modeling the behavior of others just because in our domain, it's so important to interact correctly with all the traffic participants.

**CRAIG**

Yeah. You know, a lot of people don't think of cars as robots, but they are.

**Drago**

They're absolutely robots with well established hardware stack, yes.

**CRAIG**

Yeah. Is your system, could it be adopted, or adapted for other kinds of robots?

**Drago**

I think, definitely, it can be adapted. I think a lot of the concepts across the various robotic systems over time, I believe are converging. There are still certain peculiarities in our domain, you can think of our robot, maybe as the first manifestation, maybe the most economically- hopefully impactful, and application impactful of most robotic applications. At least I believe it, and that's why I'm in this domain. But a lot of these concepts, of you, know how to build their stock, how to test your stock, how to ensure robustness, what kind of machine learning models you do. I think we're we're becoming more and more similar across robotic, but yet, in our case safety, especially of high speed, we have a lot more requirements of those kinds. You need to need to have very fast reactions, you need to have very high bar to safety, the interacting, it sometimes in parts of second with other agents who could behave also irrationally or adversarially, we need to handle all of this. And the other part is we have very many sensors on the vehicle to see in every possible direction. And in our case, we've opted for a lot of additional safety by adding LIDAR radar, which active sensors, they ensure that even if machine learning in cameras misses something, we have other ways to capture it, right? So when we put all this together, there are certain differences. But if you build the stack for autonomous vehicles, we can build on top of the strengths of this stack, we can adapt it to many other elements.

**CRAIG**

Yeah, and as the system evolves, when you were saying that it's evolved a lot, and then there's all this new research coming out with generative AI and large language models, multimodal models, how do you—

**Drago**

It's a great time in our space. I think robotics is, I think progress in robotics, I mean, is accelerating even though we're an example of already a successful robotics product out there in the real world.

**CRAIG**

How do integrate these new technologies as they come along? Obviously you do have a whole research arm? That's the assessing, but when do you decide to bring that into the car?

**Drago**

So the great thing at Waymo is we have the data and the frameworks to test the performance of our stack. And so the history of Waymo, as you know, we've been around 15 years, and the technology in this 15 years has been dramatically evolving. And every two or three years, there's major advances in our space, right. And so we've had to continually rethink and redesign our stack to take advantage of the latest technologies. And so the core is, is not machine learning, is we have the data and the evaluation, right. We have collected, as you know, we have many hundreds of vehicles out there in the world, various sensors, we have a lot of systems that can process this data, collect the interesting cases, potentially automatically annotate things we want, if we put super rare things we can get, we have systems to annotate human guidance and so on, right. And so we have all this data and now it's a question of, okay, well, it's not that hard to try new paradigm or models and compare to the performance of your previous models. And anytime we can show a significant improvement, then we evolved the stack and this was on many times already so we continue doing this. But the data is the key, right? So because you have all this data, and you have all these testing frameworks, and safety frameworks, and very rich evaluations, the key is to be able to tell when you're better, right. And we worked hard to be able to do that.

**CRAIG**

Yeah, there's so many questions I want to ask, one is on corner cases that arise in this long tail. I read something this is for five years ago I think, of a case where an autonomous vehicle, it may have been Waymo, drove up behind one of these big tanker trucks with a highly polished back that was like a mirror and that confused the system. How do you deal with that? Do you add that to your training data or do you tweak the algorithm? I mean, how do you deal with those corner cases?

**Drago**

So, you can annotate your data, right. And you can use it either for training or for testing. And clearly, the very first thing you want to ensure when you, say very rare cases is you want a test set to ensure that you got it right. And then you look for more cases like this, and you have those two training sets to help you get it right. But I think in our case, because we have also radar, and we have LiDAR, and we have camera, we will see the truck, we want see just car reflected in the truck. So we by design have mitigated this to a large extent compared to camera on the system. For example, some people had an idea, well, let's put a TV on the back of my truck and it’ll show you the world in front of the truck, in the TV, right? Well imagine the autonomous driving car with a camera looking at that TV thinking what's in front of me, I don't know, is there a car there? Maybe it's not so relevant because it may look even far away in some cases but there can be a lot of confusing things. And that's why having more sensors, so different modalities, quite conducive to safety. That's not the only one but we collect these rare examples, we mine for them, we add them to the training and testing data, and we validate the stack can handle them. I think it's a standard engineering machine learning approach.

**CRAIG**

Yep. Can you talk about the open data set, the Waymo open data set? And as part of that, Vincent was talking about the RTX project where they collected data from all these various labs where the data has been siloed for years. And then combined it and then sent it back to the labs to train on the bigger dataset and there were improvements. Were you part of RTX? And how does the Waymo open data set fit in your stack, and then in your training?

**Drago**

So let me explain. I mean, generally we’re separate markets, I think that dataset is primarily focused at objects or robots, that manipulation is on a table, right. And ours is focused on understanding the outdoors, environment and traffic scenes. So it has somewhat different parameters. The reason why we made the Waymo open data set is based on the realizations we had in the early years, especially around 2018, when we really started the project, but even before. We did research in the space, and the data sets that were available, even though Seminole did great, there was a dataset called KITTI at the time that was made in 2011, 12 by Karlsruhe and I think European collaborators.So, we started seeing that dataset does not satisfy our needs when we build the actual stacks. And the problem there was it was typically very, very small. And when the data set is small for a machine learning person, it distorts the kind of models that win. So then that dataset, to do well, it encourages you to build models that either have a lot of built in bias, or you built in a lot of overfitting techniques, which is not at all what you're gonna do when you build the real stack. And so we understood the set of problems they covered were a bit limited, but what was even more so is the data. And so the academic community, such talented, great community— And it's not easy to get your hands on large amounts of reasonably clean, autonomous vehicle data. I mean, you need to have a vehicle you need to place all the sensors you need to instrument them, you need to align the times and calibrations in the poses is just very prohibitive, right. And so we wanted to enable the community and the best way to do this and to unlock the talent is to give them the right kind of data. And through this Waymo open dataset, this was our attempt to push what is possible in the standards, the amount of data and the quality of data that this community would get. And I would like to think that we did a good part into popularizing and enabling a lot of very strong AI work, and in addition to releasing data and enriching it with more and more labels over five years. So we constantly improve the dataset based on feedback and our understanding about the space, and we expand the things you can do with it. They expand the set of tasks you could potentially try to show state of the art in. And we also, as we introduce new tasks and labels, we create challenges. And the challenge allows our different researchers across the world to compare their solutions on criteria that are informed by experience in the space. For example, what's good 3D semantic segmentation? What is a good metric for simulated agents, right? We take the learnings we have from our space, and help them essentially create leaderboards where people can compare their contributions on definitions of problems informed by us. So this has been an ongoing five year effort and we're launching actually, next week, today's March 12, when we're taping this. We expect next week to launch or the week after, our 2024 versions of challenges. There are four challenges on exciting problems, 3D semantic segmentation, 3D occupancy prediction, and 3D flow prediction challenge. Then we have motion prediction in terms of trajectory challenge and embodying simulated agents. So these are the challenges this year, and the winners will get prizes and the best work will get invitations to present it CVPR Premier Computer Vision Conference in June, at our workshop that we're organizing with other academic collaborators. So I encourage people to check it out. Especially if you're interested in our domain or doing research in our domain. I think there's some very exciting challenges this year. And I believe this is our fourth or fifth year of the challenges. So we do our best to sustain improvements to the dataset and the amount of tasks we release over time.

**CRAIG**

Yeah, on the dataset, how is it collected?

**Drago**

It was collected by Waymo vehicles, the cities of San Francisco and Phoenix. The majority of the data was collected around 2018, or 2017, which is when we chose a lot of the run segments or the snippets of real data that we decided to release. And we released 2000 or so run segments. They're sequences of LiDAR and camera data for 20 seconds. And we've released since then 100,000 run segments, both with intermediate representation in terms of road graph and bounding boxes moving around for the various objects. But we also released LiDAR for those segments, in this series, even image embeddings. So 100,000 segments is very low, very large number. It's quite prohibitive to release all the sensor data, camera data for five plus cameras on this but we released compressed versions of the camera representations to enable the Atmore research, and to combine in camera Lidar and intermediate representation, to see how good prediction can be made. The key for us is that this imitative prediction of agents is one of the key crutches in our domain for machine learning, because we observe how everyone drives, we observe how everyone walks. But in general, you can learn a lot about driving by watching drivers and that's a key. That's one of those field domains where, you know, maybe for other robots, you never see how a certain dog robot should walk much in the real world. But for driving, you do. And so it's really, really powerful data and we have been pushing a lot of researchers. You can see some of our challenges this year of predicting various things to take advantage of this key part of the data.

**CRAIG**

Are there other contributors to the dataset, other self driving car companies, or other labs that that have data that's relevant?

**Drago**

I think people contribute other labels to our dataset and do potentially their own leaderboards or competitions on top of that. There are other companies that have released data sets as well. We're not the only one since we started, right. So this also shows that the companies themselves understand that this is an important thing to do. And as the field evolves, your understanding of the problems we want to make as challenges and the data needed to support them evolves. And so you I think we will be as a community, and I hope it Waymo we can sustain what we're doing. But as a community, I would expect that there'll be a lot more interesting releases and challenges possible.

**CRAIG**

Yeah, the data as it’s coming in from your cars, are the cars sharing data with each other in real time? Or is the data going back to a central repository?

**Drago**

So it’s collecting in a central repository. When our cars drive, they typically don't send each other's data to each other that presumes that you will have very high bandwidth connections, it's actually a lot of data, it's hundreds of millions of pixels in LiDAR scans a second. So that's why the storage even for the things we're releasing is so big, it's almost a terabyte, right? If not more. And so even when we compress it, right? So these are the challenges. But when they're in the data center, sometimes we see the scene from several vehicles at the same time, right, which can of course, help us at that time reconstruct much higher fidelity models of well, what is seen by now, a collection of Waymo vehicles, not a single vehicle, this is possible, but it's not common practice in the field.

**CRAIG**

Part of the trajectory is to make the models generalize more and more.

**Drago**

That's one of the main tasks of the machine learning problem, you make the most of the data, you have both to build the agent and to build this environment to test the agent. These are the two main problems as I see them, right. Now, of course, you can break them down into subproblems but that is at the core of it for the data you have that you're collecting, and maybe keeping how much of this can I learn with my models. The more I can learn, right, the more we tilt the system towards machine learning solutions because those just scale. Then I go to any new environment, I collect more data, it will just learn from it, right.

**CRAIG**

In different parts of the world, I remember in the early days there was a lot of talk about, you train a system in Scandinavia and then try and operate it in, I don't know, the Philippines.

**Drago**

Yeah, I go to Singapore and everyone drives on the left side, right, for example?

**CRAIG**

Yea, do you need different datasets to train models for different environments, different regions? Or do the models generalize enough that they can quickly adapt to a new region?

**Drago**

So we've seen that with large models, right, they actually benefit combining the data and this parallel with development in the generative AI field, right. They try to train them on some large internet scale datasets. Or when you train a translation model, often you're training it on multiple languages at the same time. So when size is not the constraint, there is positive feedback across different tasks and different regions. And we've seen some of the seven in the case of, say, car data helps trucks, right. So you need now a lot less drag data to do well, because you can just mix in large proportions of the car data that you have. And lo and behold the results improve, right? And the same is true across cities, and environments, and conditions for larger models, we essentially would look to pull all the data and train much larger model and then essentially, potentially distill this model, or adjust it to fine tune it to the specific cases as needed. But that seems to be the recipe today. So we aim to have one Waymo driver and as few of them as possible. Traditionally, so far, all the cities we drive is a single set of models. It's not yet fragmented, said that's really hard to maintain, and it's not scalable. That's not what you want.

**CRAIG**

Yeah, I don't know if you can talk about the Cruise debacle. We don't have to talk about the company, but what went wrong there? And how is Waymo avoiding those problems?

**Drago**

I mean, I would say that, to me, the lesson is safety comes first. It's paramount, right, and I think trust is key. And so it's very easy to lose people's trust. And ultimately, I would like to think that so far Waymo has a track record of is expanding thoughtfully and successfully in the domains where we have operated. And so we want to expand relative to what we believe is beneficial and safe. And I think we have a lot of data to show that our deployments, I think, compared to, especially to humans, in various studies is shown that crash rate is a lot less and our property damage claims, there is actually a collaboration with a Swiss insurance company showing injury claims went to zero with our vehicles. That’s based on 3.8 miles that they analyzed and property damage claims went down 76%, right. And this is track record that we want to have and continue expanding on. So we look to expand thoughtfully. And I think the way to do it thoughtfully is to have extremely robust safety methodologies and analyze what you put out there with incredible rigor. And honestly, our safety methodology is very multifaceted. So there is no one golden hammer to ensure the safety of the system and there is no one golden machine learning model simulator to prove that your system is safe. Our learning is, you need to approach what we have with a variety of methodologies and make sure that we looked at it from many angles. And so this is actually one of the most main recipes that make Waymo Waymo, is that we have been working on this for 15 years. And there is a level of maturity that honestly when I joined Waymo at the time, that was one of the things that most impressed me, is that type of work.

**CRAIG**

Yeah, so it wasn't that the data set was inferior, or the models were inferior, or there weren't enough guardrails, or hard coded rules; it's sort of a combination?

**Drago**

I mean, you want to build a system that you convinced yourself is safe to put out there, right. That's a high bar. That's not a simple question to solve, right. And I think at Waymo we tested extremely extensively in a variety of different ways. And that determines how quickly we expand, is our success in our tests, right. We do this also thoughtfully. We need to engage and work with the community. So, Phoenix, we have been delivering rides close to six years now, right. And over time, now we're at a scale which is quite noticeable and thousands of people or maybe tens of thousands take us every week; but it's not like we, were not reckless, right. It's hard for me to talk about Cruise, per se. This is more of a lesson for us. I think that ultimately the details of the incident that had them grounded, I think a lot of the details are out there and people can judge for themselves. But we need to ultimately build trust and a lot of this trust is based on the safety record you have.

**CRAIG**

Yeah, from what I understand China has more robotaxis deployed than the US. Do you think that they're, presumably you pay a lot of attention to what they've got on the road, do you think that they're ahead of you guys? Or is everybody kind of on the leading— self driving car systems on par? And then the other question I want to ask is— as I see we're gonna run out of time, there's a lot of confusion about Tesla's system. Does it approach what Waymo is doing? Or is it something different altogether?

**Drago**

So maybe let me first address the Chinese companies question. I mean, clearly, a few companies there have started fully autonomous deployments. Unfortunately, I don't have full visibility into their system, I have not had the chance to ride in them or study them. I mean, a lot of that design is proprietary. I think ultimately, one needs to experience the service and the people who are best suited to applying how we all compare technology wise, is people who experience all these different services, right. Clearly in China that’s an area of investment and maybe even government assisted domain, right. That said, they're not here currently in meaningful deployments, so we're not head-to-head with them. And I don't follow them that closely even though I understand that, you know, that is an area of rapid development. In terms of companies like Tesla that deployed more Driver Assist type technologies, I think, you know, I appreciate the machine learning work and innovations they do; I think to me in their nature, the system is not full self-driving. So they do good machine learning but as with most machine learning, I’m a machine learning person and as much as I appreciate it, right, there comes a time when you get to situations that are extremely rare, and machine learning cannot handle. And to have real deployment out there, you need to carefully think through your stack to make sure that even if machine learning does not solve everything 100%, your full self-drive product does. And that's a very big gap. I think that's not easy to do at all. And it leads potentially to rethinking core parts of your design, if you are to go there. And so I mean, all of our machine learning models are improving. To release a driverless service, you need to also answer the question of how to make sure the whole thing is fully robust. And that adds a level of design and complexity that a lot of those companies have not tackled yet.

**CRAIG**

 Yeah, would your system ever be available for consumers to buy?

**Drago**

I mean, in the beginning Waymo’s charter, Waymo’s driver, we eventually want to enable many different kinds of use cases and form factors. I think our first, kind of, flagship application is ride-hailing. But we're building a general stack. I mean, we discussed that maybe we can transfer a lot of the models or types of models that were built for autonomous vehicles to other mobility robots. But even before that, we aspire to build a stack that learns from that, and generalizes across these platforms, and enables various autonomous driving applications, preferably with as similar of a system as possible, or maybe the same system. That is possible. So I think down the line, we have the opportunity to build solutions for personal vehicles, for trucks, potentially other form factors. But I think ultimately, for a company to succeed, you need to nail your first product. I think where we are is, I mean, I personally really enjoy riding in our vehicles. I think it's a delightful experience. Of course, I'm biased, but they’re available. Everyone can try them. You can come to San Francisco, and Phoenix, and soon LA and Austin, and just ride them, judge for yourself. But it exists, it's real, it's out there, and we've proven that the product exists. Now it's on us to make sure we can have a great business doing it and expand from there.

**CRAIG 52**

Yeah, and how long do you think before every city has a robotaxi service? And then how long beyond that before they're autonomous vehicles on the highways and all over the United States? I mean, I understand it's largely a regulatory issue, but in terms of the reliability and the safety.

**Drago**

We are expanding every year, rapidly. You can think about it for now in a multiplicative phase. So every year we go many times the driverless deployments we've had before, right. And even though that's thoughtful expansion, I believe. So it will take a few years to continue this trend until people really feel it all over the place. And to me, mostly, I don't count how many cities we will be in. I want to build the stack that makes it economically great and scalable to do a city, n+1. And so there is a ton of headroom still possible with the technologies that are being developed. And so I hope we can facilitate it. I believe it's not far out. My personally great hope is that this multiplicative expansion of operational design domain, in its scope and volume, if we sustain it for a few years, people will see these cars in many more places.

**CRAIG**

But you think, I mean, you're quite a bit younger than me. But do you think in 10 years, when I'm still— got my arms, I hope?

**Drago**

I mean, sure, I'm an optimist generally. And as an optimist, I would say it should be less than 10 years and probably by some margin.

**CRAIG**

Yeah, that's exciting. And where in the development of the, either the data or the tech stack, what are you most excited about?

**Drago**

I'm excited about, I mean, all this great technology we have that allows us to generalize a lot better from the data we have. And I think we understand better and better how to land these, kind of, scaling multipliers. So the more I can learn from the data, and that they have, right, I'm excited about keeping up with improving our scaling multipliers on what we can learn from the data, the more we do this, the faster we will expect, right. Because ultimately, machine learning, a lot of the other things, So when you get to the hybrid design of the system, then when you need humans to essentially add expert expertise or guardrails. That doesn't scale very well. But machine learning is our great scaling tool. And we can use it, the faster we will, we will grow across the areas that we serve. The other part that really encourages me actually is, maybe the thing that makes it easier to do this, which is; so what we've discovered is, after you cover some of these areas, for example, Phoenix is more, parts of it are 45 miles an hour road with certain challenges. San Francisco is dense-urban with certain challenges. Maybe we cover the highway; now there’s a lot of work on harnessing the highway and starting to enable that for customers. I think what we saw is that after you have, say, Phoenix and San Francisco and you go to LA or Austin, it takes a lot less work and effort. And a lot of it is actually in validating our system, that it does well as opposed to actually making it do well. So, the learnings generalize, and that's a great kind of wind at our backs. And so as long as we design the system well, and we don't, I mean, again, partition every city for itself, now we actually benefit from San Francisco to do LA, and we benefit from Phoenix to do LA, and then they have more cities that the city plus one will benefit from that. And so, I see it also as, that's a great property, right. Scale begets scale in sense. And then, you know, having these deployments, having all these experiences, and the data from them will help us with next.

**CRAIG**

So the ambition is that there'll be fleets of robotaxis in every city and town. And will it be, from what you can see, I know that you're not on the business end of this, but will it be inexpensive enough that people will not need to own a car? You just, on your phone, you summon a robotaxi, it arrives in five minutes, you go where you're going? I mean, in much of the way that Uber operates?

**Drago**

So I will tell you now, right, we're roughly the same price, but we’re much better experience, personally. Again, I'm a Waymo representative so that's my personal bias, but you're in the vehicle by yourself, you can play the music you like, there is a certain level of, I mean, it's a hand vehicle. It's a premium experience today for the price of a normal experience, right. I think over time there is a tremendous opportunity to optimize both our models and vehicle cost, and operations to make it yet more affordable than it is. Now, there is a lot more work that needs to be done and I think it will shift from, for now we don't even have as many vehicles as is the demand, over time- hopefully. As we make it more affordable and yet better experience, then it's a beneficial look. If we can succeed with this, we'll expand the market.