



EXPERT INTERVIEW

UPDATED

03/09/2025

Sankaet Pathak, CEO of Foundation, on why humanoids win in robotics

TEAM

Jan-Erik Asplund

Co-Founder

jan@sacra.com

DISCLAIMERS

This report is for information purposes only and is not to be used or considered as an offer or the solicitation of an offer to sell or to buy or subscribe for securities or other financial instruments. Nothing in this report constitutes investment, legal, accounting or tax advice or a representation that any investment or strategy is suitable or appropriate to your individual circumstances or otherwise constitutes a personal trade recommendation to you.

This research report has been prepared solely by Sacra and should not be considered a product of any person or entity that makes such report available, if any.

Information and opinions presented in the sections of the report were obtained or derived from sources Sacra believes are reliable, but Sacra makes no representation as to their accuracy or completeness. Past performance should not be taken as an indication or guarantee of future performance, and no representation or warranty, express or implied, is made regarding future performance. Information, opinions and estimates contained in this report reflect a determination at its original date of publication by Sacra and are subject to change without notice.

Sacra accepts no liability for loss arising from the use of the material presented in this report, except that this exclusion of liability does not apply to the extent that liability arises under specific statutes or regulations applicable to Sacra. Sacra may have issued, and may in the future issue, other reports that are inconsistent with, and reach different conclusions from, the information presented in this report. Those reports reflect different assumptions, views and analytical methods of the analysts who prepared them and Sacra is under no obligation to ensure that such other reports are brought to the attention of any recipient of this report.

All rights reserved. All material presented in this report, unless specifically indicated otherwise is under copyright to Sacra. Sacra reserves any and all intellectual property rights in the report. All trademarks, service marks and logos used in this report are trademarks or service marks or registered trademarks or service marks of Sacra. Any modification, copying, displaying, distributing, transmitting, publishing, licensing, creating derivative works from, or selling any report is strictly prohibited. None of the material, nor its content, nor any copy of it, may be altered in any way, transmitted to, copied or distributed to any other party, without the prior express written permission of Sacra. Any unauthorized duplication, redistribution or disclosure of this report will result in prosecution.

Published on **Mar 09th, 2025**

Sankaet Pathak, CEO of Foundation, on why humanoids win in robotics

By Jan-Erik Asplund



Background

With step-function increases in the quality and reliability of hardware—sensors, actuators, microprocessors, cameras—and the emergence of transformer-based AI, the moment has arrived for humanoid robots. The primary battleground now—for companies like Figure, Foundation and Tesla—is building a real-world training data flywheel. For more, check out our interview with Foundation co-founder and CEO Sankaet Pathak.

Key points from our conversation via Sacra AI:

- **Circa 2022-2023, we saw 1) step-function increases in the quality and reliability of hardware—sensors, actuators, microprocessors, cameras—and 2) the emergence of transformer-based AI, making this the moment for humanoid robots that can operate in unstructured, dynamic environments without hardcoded rules. "Before 2022, we had no path to building end-to-end policies. That's**



why most of the robotics that succeeded in the last decade were robots that are very dumb from an AI perspective. If you build them the perfect environment—if you need a robot to palletize, just make sure items show up where they're supposed to, all humans are gone, etc.—they can do simple tasks.. After 2022, when we saw the autoregressive transformer model architecture scaling, it became obvious that now you can encode end-to-end policies into these models that you couldn't before."

- **Compared to traditional industrial robots like Amazon's that rely on 12-18 month retrofits, a robot with a humanoid form factor (bi-pedal, two arms) can quickly slot in alongside a human in performing the most dangerous, high attrition tasks in a factory—those with ~100% worker attrition over 1-2 years.** We're building general-purpose robots that are humanoids because the world's made for humans, and we just want to be able to assimilate in it very quickly... For instance, when you have to carry a car bumper, you just need to be able to lift it up and bring it somewhere. You need two hands... If you have a [wheeled robot], turning is going to take longer. So you need something that's bipedal... At that point, we can try really hard to make it not look like a humanoid, or we just make it look like a humanoid."
- **All of the humanoid robotics companies, from Figure (\$1.5B+ raised) to Tesla (19TB per day from each car on the road) to Foundation, are still at the starting gates of what has become the key competition—collecting real-world training data from factory floors and other deployment environments to train their AI models.** "Tesla probably has the best and most comprehensive labeled dataset of the real world. The only caveat is they only have that for roads. They don't have it for indoor autonomy. If they had it for indoor autonomy, then I think they would pretty much win, because at that point, you have all the data you need. So even Tesla is in the same boat as us. You have to collect data as you go, and then your robots get better and better... Very soon—by very soon, I mean probably a decade—you have pretty much all the data you need to build a foundation model."

For more, check out this other research from our platform:

- Scott Sanders, chief growth officer at Forterra, on the defense tech startup playbook
- Anduril, SpaceX, and the American dynamism GTM playbook



- Anduril: the \$342M/year Nintendo of American dynamism
- Shield AI (dataset)
- Ross Fubini, Managing Partner at XYZ Capital, on the defense tech opportunity
- Anduril at \$1B/yr
- Anduril (dataset)
- The biggest mistake defense startups make

Interview

In short, what is Foundation?

The way I describe Foundation is an advanced technology company with the mission of building autonomous machines that can automate all kinds of labor. For us, that means logistics, defense, manufacturing—all of the industrial use cases plus defense and security. We're less interested in making coffee or folding laundry. We're much more interested in building homes, building cars, moving stuff around, etc.

That's essentially what Foundation does. Our first product is a humanoid robot that is set for initial delivery in April and May this year.

Can you talk about where those initial deliveries are going and who those initial partners or customers are?

We're sending our first fleet to an auto manufacturing OEM. Later in the year, we might send some of our fleet to other customers.

A decade ago you co-founded Synapse, the prototypical banking-as-a-service company. What got you interested in robotics and starting Foundation?

I've been writing code since I was 12 or something. I did my graduate school in electrical engineering. In college, I worked a lot on microprocessors, ECUs, and various other things. I did more hardware work in college than software.

When I started Synapse, it's because no one wanted to fund a hardware company that I wanted to build, but people wanted to fund a software company. So I started with software. When I



was thinking about what's next that I want to do, I wanted to go back into building something more tangible.

The irony is it's easier to explain to your friends and family that you're building a robot than it is that you're building banking as a service in the cloud. They say, "I have no clue what you're talking about."

So it also makes my family conversations much easier. I show them the robot.

Can you walk us through why right now is the right time to go all-in on shipping a humanoid robot?

There's a few reasons. On the hardware side, electric motors have gotten far more commoditized and efficient. They usually don't break. ECUs, because of new cars and smart gadgets, have gotten substantially better. All microprocessors have gotten really good. Cameras, which are sensors, are cheap now and quite reliable. And compute is also quite reliable. There are many GPUs available for inference.

All of that is a very strong, potent combination for building new hardware from the ground up. It doesn't have to be a humanoid, but any new hardware is reaching stability faster and faster. It used to take decades to make a microwave efficient so it doesn't break. Now, in a year or two, you can have vehicles like EVs and gadgets that are quite efficient.

Those are big step-function changes in hardware that now give everyone confidence that we can definitely build modern hardware. The second piece that is very important is the AI side.

Before 2022, we had no path to building end-to-end policies. Now you can give them some data, they extrapolate everything that's happening, and they give you an output. Before that, you'd take a picture, do computer vision, label things, then write code to do something. But you cannot do that in an unstructured dynamic environment.

That's why most of the robotics that succeeded in the last decade were robots that are very dumb from an AI perspective. But if you build them the perfect environment—if you need a robot to palletize, just make sure items show up where they're



supposed to, all humans are gone, etc.—they can do simple tasks.

After 2022, when we saw the autoregressive transformer model architecture scaling, it became obvious that now you can encode end-to-end policies into these models that you couldn't before.

This essentially means that when the hardware is reliable and you have the right AI technology, it starts becoming not a hardware problem, not an AI problem, but just a data problem. I think that's why everyone's excited about it.

We have all the right building blocks to build general-purpose robots. We're building general-purpose robots that are humanoids because the world's made for humans, and we just want to be able to assimilate in it very quickly. But any kind of dynamic robot that can operate in an unstructured environment was not possible until 2022.

Companies like Tesla have highlighted the critical role of data (and feedback loops from existing products) in advancing their autonomous capabilities. How do you think about the importance of data for Foundation's strategy and how do you think about building a data flywheel?

Tesla probably has the best and most comprehensive labeled dataset of the real world. The only caveat is they only have that for roads. They don't have it for indoor autonomy. If they had it for indoor autonomy, then I think they would pretty much win, because at that point, you have all the data you need.

So even Tesla is in the same boat as us. You have to collect data as you go, and then your robots get better and better. Unlike cars, where you can still drive them even though they're not driving themselves and they still provide utility in some form, that's not the case with robots. If they're not autonomous, they're providing no value.

To me, the most important thing is determining the best way to acquire data if this is a data problem. The best way is by giving people something that's valuable.

Going to customers and saying, "Give us your most boring repetitive tasks where you have the highest attrition rate, and



we're going to automate it and collect data so we can do it." Then slowly, you start building more use cases like that and increase your fleet size.

Very soon—by very soon, I mean probably a decade—you have pretty much all the data you need to build a foundation model. That's the progression. Step one is figuring out how to make your current robot useful for people, which has a different AI technique in my opinion.

Then, once you have all the data, how do you generalize this? That's a second technique on the AI side that enables you to do better.

Can you talk about your approach to embodied AI?

There are two important aspects of building robotic intelligence.

One is your high-level reasoning. The ChatGPT or Grok part—being able to show them the environment and have them describe it, tell you exactly what's in it, listen to a task from you and spit out a recipe. Like, "How do I open a water bottle?" And it will literally spit out exactly how you do it.

"Where is the water bottle?" It'll show you exactly where the water bottle is. You can even ask, "How do I open this kind of water bottle?" And it's going to do that. You could be even more vague: "Bring me something healthy to eat." And it's going to think, "Oh, an apple's the most healthy thing here." So now the recipe is: pick up the apple, take the apple, place the apple on the table—table's X, Y, and Z coordinate here, apple's A, B, C coordinate here.

That's your high-level reasoning. It helps you understand instructions and then gives you a recipe of exactly what you need to do. What LLMs do not do is tell you how to do it. They will tell you what you need to do, but they cannot tell you how you would do it.

For how to do it, you have to build out another kind of policy called an action model, which essentially takes this recipe from an LLM and says, "For me to be able to pick up this apple, I need to move my right hand from position X to position Y, grip the apple, pick it up, bring it somewhere else."



The first part is getting to a place where it's pretty commoditized. You can essentially get Llama, Facebook's open model, and fine-tune that for this recipe book kind of use case I described, and it just works. So we're not spending most of our time working on that.

The action model is the second model. There's nothing off the shelf available that's any good. We're spending a lot of our time building that, and we're building it slightly differently from how others are building it. We think state-based models, which is the technique we're using, are far more efficient when you don't have a large enough dataset.

Why build humanoid robots, assuming that not every task out there across manufacturing or defense use case actually inherently requires a human shape to complete?

I can tell you my calculus as to how I arrived at the humanoid form factor. I'm really concerned about our overall industrial throughput in the next thirty years because birth rates are declining everywhere.

We're below replacement rate in almost all countries, which means we would not have enough people actually wanting to do the labor jobs, and that risks the whole civilization collapsing—a terrible outcome for everybody.

Based on that, I think we have thirty years to automate our GDP. If we do that, I think we're going to accelerate to becoming a type one civilization. If we don't do it, then we risk regressing. So that's the time window we have to work with.

Before I even started Foundation, my other cofounder and I went to tons of factories. They were all ex-customers of his, and we toured all of them to see exactly what kind of jobs people are doing, talk to the people running the factory, understand their pain points, and what they would want a robot to do to alleviate those pain points.

As we started looking at most of the tasks, it was very obvious to us that you need two hands. For instance, when you have to carry a car bumper, you just need to be able to lift it up and bring it somewhere. So it's not going to work with one hand. You need two hands.



Then there are tons of use cases in confined spaces. You need the robot to not just have two hands, but also move. You can move them on wheels, but the reality is there are many tasks in congested spaces. If you have an AMR, which is just this wheeled robot, turning is going to take longer. It's going to kill your cycle times. So you need something that's bipedal.

At that point, we can try really hard to make it not look like a humanoid, or we just make it look like a humanoid. We need two hands and two legs, so we decided to make it look like a humanoid.

The other constraint that factories had—and that I had on the thirty-year time horizon—is that it's a tall order to expect a factory to completely retool how they do manufacturing, which is what most robotic solutions require. If you go in with automated palletizing, how humans were doing it is very different from how traditional industrial robots would do it, requiring you to completely change the environment. It takes twelve to eighteen months to retool that whole thing before you start getting benefits.

Customers would much rather not do that. They'd prefer to say, "As we have attrition, instead of trying to get another human, we just replace it with a robot." Some of these companies have 100% attrition rates in twelve to twenty-four months. It's not a small number, and they're having a tough time replacing and training people.

If we take this approach, we can slowly and gradually automate everything, and that thirty-year timeline seems more feasible than if I had to rely on expecting every single factory in the world to change how they've set up their factory. I don't think you'd be able to do that in thirty years.

Do you think these two categories—consumer robots and industrial robots—will remain distinct, or are they destined to merge over time?

These two categories will converge, absolutely. 100% will converge. Would it converge at the same form factor? I don't know. Right now, the robot's about 5'6". Our next generation robot's about 5'10", 5'11". I don't think people would want those in the house.



But something smaller, maybe, something less threatening, maybe—it has to be extremely safe, has to be able to navigate a home, which is a very unstructured environment that changes with every single home. Your AI needs to get so much better. Your safety stack needs to get so much better, and your form factor is TBD, primarily because people feel differently about these things when they're in their house versus not.

So yes, I think these categories would absolutely converge. Right now, my focus is that thirty-year time clock. If in the next five years I'm thinking, "Oh, we're totally on track, we're going to get this in thirty years," then I'll contemplate doing consumer robots as well. But I could do a consumer robot right now and that might distract us.

If we don't automate labor fully and no one else does it, then we would only have those house robots roaming around and no humans. I think that would be a terrible outcome. So I believe the first focus for everyone trying to build a product in this category should be industrial.

You've talked about the idea of a "fleet coherence" model, where multiple robots can learn collectively and share insights in real time. How does that factor into your long-term vision of how these robots operate—are we talking about a single robot getting smarter over time, or an entire network of robots pooling their experience?

Yeah. I mean, in context of deploying the fleets we're deploying this year, we're just talking about deploying a few robots that, by and large, work independently. They don't really work in conjunction with each other. A path to fleet coherence is, initially, you have robots that are doing their individual jobs really well.

Then you have some kind of things like what Verity does—they do a phenomenal job of this for drones. They essentially do these drone light shows. That's like a swarm network. Every single drone is talking to each other about which position they need to be in, etc. But it's also very dumbed down. It can only work in some preset formation loop.

The fleet coherence I'm talking about is very similar to a GPU cluster coherence. Most of your training requires every single GPU inside your network to know exactly what the other GPU



is doing, and it happens at very fast clock speeds. Then you build larger and larger models that end up being more intelligent than the ones before.

In that context, fleet coherence for humanoids or robots generally means you have a larger overarching goal. Let's say build a city on Mars, and you have about a hundred thousand humanoids to ship there, and their sole job is to build a city on Mars. If one robot is finishing a hospital and you only need one hospital in that city, you want to make sure that other robots do not start building a hospital. They know what's been completed and what's left.

They also know where you don't have enough bricks and need to get some materials from the other side of town. That happens when an entire network of humanoids are able to stay in sync constantly around its state and status towards the task and reason from there how you would expect them to collaborate. Very similar to an ant colony. That's the right framework.

Tesla and Waymo have both integrated teleoperation as part of their autonomous strategies. Is teleoperation a part of Foundation's roadmap towards autonomous robots, or is it important that you design with the constraint that you don't use it?

I think you should definitely design to use teleoperation. The reason is that for the customer, you want the experience to be very seamless. If your model mispredicts, you don't want the whole line to stop. If your model mispredicts, you have to have a human behind the scenes that intervenes and brings it back to normal.

For that reason, I think the right UX for autonomous machines is obviously autonomy first. Deploy a model that has low intervention rates, but when you have to intervene, you intervene.

A human does that, which is what happens with VAML as well. After that, you label and collect that data as intervention data, and then you use that to improve your model. Intervention-based teleop is highly critical in being able to build very smart, robust models.



Now if some company says, "All I'm going to do is teleop," I just don't see them scaling to large fleets. You have to have intelligence in there as well because otherwise, you're going to have to hire one teleoperator per robot per shift. So technically, three teleoperators for one robot.

If you deploy 10,000 robots, that's 30,000 teleoperators—it's a logistical nightmare. So you do have to have the models built, but I think having an intervention-based teleop system ends up being far more productive and helpful.

Humanoid robotics seems like a capital-intensive space, and Figure has raised \$750M and is in talks to raise another \$1.5B at a \$40B valuation. From the outside looking in, an investor might think Figure has been anointed—why would that be wrong?

I think people are just bad students of history. I will give you some numbers, and they're going to blow your mind. Tesla raised \$100M before they did their first Roadster deliveries. Nikola raised \$700M and didn't deliver a single car.

How much did Fisker raise and still have no cars out? \$1.5 billion. How much did Faraday Future raise? \$2 billion. How much did Canoo raise? \$600 million.

If you look at the battlefield, Tesla is the company that took the least amount of capital and ended up being successful. And that's primarily because it forces you to prioritize. It forces you to focus. It forces you to move fast. Those things are important.

What history has taught us is that a lot of capital before going to market is usually detrimental. It kills companies. Harvard Business School did a study a few years ago where they studied the difference between capital-efficient companies and companies that are not capital efficient, and what that really looks like. If you were a capital-efficient company, your likelihood of success was 40% higher than a company that was not capital efficient.

So it's a very straightforward calculus. I think capital efficiency is the moat for deep tech businesses. It's not just the amount of capital we raise. Looking at history, anyone who raises too



much money before they ship a product ends up dying. So I'm not worried about any of them for that reason. The market's nowhere close to being done with all of this stuff. And right now, if you're burning \$22 million a month, which, I believe, what Figure's burn is, and you don't have your product out in front of customers, that just sounds like a bad idea.

SoftBank had this whole philosophy—Masa did—which is if I make it so the company is not capital starved, if I give them so much capital, they will be successful. The answer is that this philosophy has never worked in the history of business. The market's too early. We're pretty much toe to toe with Figure and Tesla, and we've raised substantially less capital to get there.

Obviously, I'll raise more capital, but I would be more proud of shipping our robots and scaling them and spending less than \$100 million, versus raising \$2 billion and then dying.

Is there meaningful divergence between the three companies on initial use cases or customers?

Tesla's going after consumer, so that's their primary focus. Figure and Foundation are head to head. We're both going after industrial use cases. Figure recently announced that they're also building a consumer bot, which is another instance of abundance of capital making you not have to force a choice.

But that's the differentiation. Tesla's consumer. Figure used to be just industrial. Now they want to do industrial and consumer. Foundation is industrial plus defense. Figure has sworn off doing anything in defense.

How do you think about defense as a market, and why did you decide to go after building for DoD?

When I was starting Foundation, I actually did not want to do industrial. I only wanted to do defense. Why I wanted to do defense is because I thought, if I need to teach our robots how to build cities on a different planet, probably the military is the best place to learn that than anywhere else. That was my big motivation for going into defense.

To me, any capital that Foundation makes is so that we can accomplish building bases and cities on other planets. So I'm



thinking, how do I finance my way through it?

As we started talking to the DoD and doing some work for them, we realized that all of the initial applications they wanted to use us for were in maintenance or logistics.

Then we thought, why are we just constrained to defense?

Those things also apply to industrial use cases. So we decided to broaden our aperture and be dual-use—do both.

Around 75% of humanoid robotics companies are based in Asia today. How do you think about the competitive dynamics, particularly with humanoid robotics companies in China?

I don't think we're in competition with China because, at least based on the current political climate, it wouldn't be acceptable to import stuff from China. So I don't think that is our near-term competition.

China is really good at manufacturing. There's no doubt. They have a thriving middle class that is very hungry to get ahead. People can work multiple shifts there. People are very intense and dedicated, and you don't get that in the US. So I do think that is somewhat of a risk for America for sure.

As it comes to Foundation, we want to move as fast as we can and be as scrappy as we can while still following the law. We can't make people do five shifts or something like that. It's probably not going to work. But I think there are other things you can do, which is build a very tight-knit working team.

So, no, I'm not that concerned about Chinese competition. I do have a lot of respect for Chinese competition because what they've been able to accomplish, even if you just look at the humanoid robotics market, is phenomenal. There are so many companies there that are building these robots. Some of them have already shipped a thousand plus robots as of last year. So clearly, what they're building is working, and America is slightly behind.

If everything goes right for Foundation over the next 5 years, what does Foundation look like and how is the world changed?



You'll probably see millions of these robots in five years. You'll probably see thousands of these robots next year, then hundreds of thousands of robots the following year, and then double-digit millions of robots by five years. People don't realize this. We're very close.

And I know people have cynicism of, "Well, we said that before with Boston Dynamics." We were never this close—by a long shot, we weren't this close.

I'd be very surprised if there are not millions of humanoids running around outside in the next five years. It's going to be more concentrated at the beginning. There are going to be a few cities where it starts, and then it grows and grows and grows. But I think we're getting pretty close.

Disclaimers

This transcript is for information purposes only and does not constitute advice of any type or trade recommendation and should not form the basis of any investment decision. Sacra accepts no liability for the transcript or for any errors, omissions or inaccuracies in respect of it. The views of the experts expressed in the transcript are those of the experts and they are not endorsed by, nor do they represent the opinion of Sacra. Sacra reserves all copyright, intellectual property rights in the transcript. Any modification, copying, displaying, distributing, transmitting, publishing, licensing, creating derivative works from, or selling any transcript is strictly prohibited.