



Episode 227 – What’s the Path Forward for AI in Space?

Speaker: Paul Lasserre, General Manager, AI for Space, Loft Orbital – 33 minutes

John Gilroy:

Welcome to Constellations, the podcast from Kratos. My name is John Gilroy, and I'll be your moderator. Today, we're speaking with Paul Lasserre, general manager, AI for Space at Loft Orbital. We will dive into how artificial intelligence is moving from Earth-based cloud systems to autonomous decision-making in orbit. We'll explore how satellites are evolving from passive observers into intelligence systems that can interpret events, prioritize information, and coordinate across constellations in real time. What a big topic, Paul. You ready to handle it?

Paul Lasserre:

I am. Hi, John. Thanks for having me.

John Gilroy:

Whoa. Wow, wow, wow. This is going to be a good one here today. So, Paul, satellites have mostly been passive observers. When does observation start to cross into intervention? And how should we think about that shift from both a technical and an ethical perspective?

Paul Lasserre:

Sure. Look, that's interesting to notice. And as you and I discussed before, John, my background has not been in space, but mostly AI and cloud systems after I started my career in the French Navy. And my first surprise when I joined this industry is that we operate today basically the same way we have for the past decade, which is if I simplify for Earth observation, pretty much about capturing data from space, from your different sensors. You need to wait for your satellite to fly over a ground station, dump all of these heavy data, and later on analyze it, manually, automatically, does not matter.

And as a matter of fact, due to these laws of physics, those observation until now has mostly led to strategic decision making, analyzing data after the fact, which has value in itself. But what we've witnessed on Earth across industries over the past decade where AI has helped in real time decision making, whether that's like autonomous driving, medicine, and any type of industries where AI today helps make decision in real time has not traveled to space yet.

And so that's really what we observe today, this fundamental shift from pixels to events. And that's pretty much what we do at Loft Orbital. We are running models at the edge that can bring intelligence to the sensors. And you can think of AI in space and go from observation to action in a way, in three way. That's what I call the three-legged AI stool if you want. One is really about this ability to do onboard in front, to run intelligence in space at the sensor. And that's really a shift, where it's the common denominator to real-time decision making in space for these laws of physics I just described, and the fact that data, otherwise information cannot travel with a traditional way.



And so this first leg is really about having the ability to define what matters to you, analyze data in real time, and to send an insight just a few kilobytes of data through real-time connectivity, high altitude satellites in Marset, for instance, to let someone know in seconds that the wildfire is starting, submarine of interest is leaving a harbor, or whatever condition is met that you've predefined. That's number one. And that's fundamentally changing how you can use satellites, not only in a strategic manner, but tactically.

The second leg of this tool, if you want, which has a lot to do with intervention as well, is intelligent tasking. How can you basically translate the experience we increasingly make on Earth with a ChatGPT-like interface to translate intent into answers or even actions with agents that now can fulfill actions and trigger systems? The same is progressively true in space, where you can basically tell in plain English, plain French, whatever you care about.

If take the same example of a wildfire and you want to prevent fires from starting in Yellowstone Park, the operator can say something like, "Just let me know, send me an alert when there is some risk of a fire starting in Yellowstone Park." And imagine the system automatically getting the latitude and longitude of the park, selecting the right sensor, having the right tip and queue approach when one satellite can have a large swath, the next one can be more precise to confirm, trigger the right sensor, around the right model, send the right alert through inter-satellite link and get that on the phone or the system of the operator.

And the third one, to remove from this observation to intervention, the last leg of this tool is openness. We don't want to limit this ability to add to what Loft can do. So we are pretty much see ourselves as a platform, like AWS for space, where others who have knowledge in wildfire prevention, intelligence, energy, surveillance and climate change can come and build this type of logic on top of our platform. So that's really how we think about evolving your observation.

John Gilroy:

Now, you mentioned AWS, that's a pretty big environment and you've had experience there. So given your work with AI in large cloud environments and now in on-orbit systems, what differences do you see and how autonomy is designed and used in each environment?

Paul Lasserre:

It's kind of like day and lights in AI when you look at what you start with. So maybe first clarifications, like AI in space does not mean at least for us and what I just described to replicate the infrastructure on Earth in space. And I know it's been a popular topic these days with Elon Musk or my former both, Jeff Bezos, and others talking about data centers in space. It's a moonshot. I'm sure given the first principle of these areas, that's something that might materialize in the fullness of time.

What we're talking about now to change observation is different. It's really about putting as little intelligence as needed next to the sensor to bring intelligence at the edge. And whatever you can run on Earth, you run on Earth. Whether you can run on AWS, GCP, your server, you run it there. The part you can't, and you have to go and run inference in plus and minus 100s, sorry, I'll talk in Celsius temperature, 100 and minus 100 temperature differences with very low connectivity, with radiations that really bring a lot of constraints on your GPUs, you want to minimize that.

And so if you have to summarize the bigger difference in terms of infrastructure with what people are used when they do AI on Earth, is one, you don't have real time connectivity, at least a high bandwidth



real time connectivity. You need to wait to be on top of your ground station, unless you just send metadata to higher altitude satellites. You can't use state of the art compute. Why? Because temperature and radiation force you to harden your compute. And so the latest like Blackwell from NVIDIA is probably not yet a thing in space because it needs to go to this cycle of space readiness. And even if you could, the other thing I discovered is that you would totally be off in terms of budget power. There is no power plant in space, and the battery, you just need to deal with whatever solar panel you have.

To give you [inaudible 00:08:14] an order of magnitude, the H100 from NVIDIA that are commonly used today in data centers, are about 700 watts. The latest Blackwater, I just talked about maybe one kilo, 1.2 kilo. We talk about 500 watt max in our satellites to do everything, the tasking, the CPUs, the GPUs, the different sensors. And so, yeah. But we still can have pretty solid compute, like Jetson Orin, for people familiar with this type of hardware, that allow us to do classification and also type of quite advanced tasks that are needed at the edge.

John Gilroy:

I wrote down the phrase you talked about, intelligence at the edge. Now there's some people who think the edge is Colorado and some people, like you, think the edge is in outer space. And so this brings all kinds of problems with it. You mentioned a few of them, but space really does force limits on power compute and bandwidth. You mentioned that. So what do these constraints reveal about AI system design that we tend to overlook in our more resource rich environment in Colorado?

Paul Lasserre:

Yeah, you need to... And that's exactly how we work internally and with our partners. You need to think hard about the trade-off on what you want to process at the edge, because the real time nature dimension is critical and you can't wait, and what you want to really process with more generous power budgets and bigger models. So for instance, you might think of a use case like looking for illegal fishing somewhere. And so the first order priority might be to detect all of the vessels that don't respond, like AIS or are where they should not be. So that's about detecting something and maybe classify as a fishing vessel. And that is a model that can run at the edge, lightweight, like CNN and so YOLO types of models for people familiar with this family of models, that are relatively low budgets in terms of power and size and number of parameters.

But if you then want to be able to have a state-of-the-art model that will tell you which type of gears we use, the nationality of this ship, maybe try to detect the name to do super resolution and all the techniques we use an imagery today, that's maybe something you can do 20 minutes later. Once you've been alerted, there is something that catches your attention. You've got the time to queue your data and to ask this data to be downlinked first once you're over this ground station, and then run these more classic high power, large number of parameters models to do things that are closer to the state of the art today.

John Gilroy:

So Paul, I know some data scientists and they debate this concept of what's meaningful, what isn't meaningful all the time. They can go for eight hours all day long. But when a satellite decides what counts as a meaningful event, so how is that criteria set and expressed in the model?



Paul Lasserre:

Today, it's relatively straightforward where these conditions are defined by the customer or the partner that creates an application for the end customer. And so you can think of conditions such as, "Hey, I want to know if anyone is crossing this border." And you might look for illegal immigration or you can look for deforestation or illegal resource, basically activities, related activities, and the condition will only always be set by a human, like the customer. And the model then is going to, based on the training it's received and the data processes in real time, to surface a risk of this condition being met.

And of course you might have some false positive. Of course, you will have false negative by nature of, again, given the laws of physics, like there are cloud, clouds about half of the time. And so you miss events, that's for sure. And even when you catch this event, sometimes you'll have false positive. But you can think of the information that you get from these models as early warnings. You basically look at vast amounts of territories or at sea, and you act as a broad net that you're casting, and the model is helping you and giving you hints on where to look more carefully. And if anything, you can treat the system as a reordering queue workflow.

Again, if you make the parallel with healthcare, we had very similar discussions in healthcare 10 years ago. Good friend of mine, Chris Mansi who started like Viz.ai, they are now like leader in AI power, like cerebral stroke detection running in most hospital networks in the US. I don't know if... I'm sure the system initially did not get it all right all the time. And what they did is basically treat their system before FDA approval as a reordering queue for physicians in hospitals where minutes matter, as you know, when you get the symptoms of a strike.

And if you, after getting a CT scan, if you can basically look at the risk, even if you have some false positive in the mix, but the risk and to basically get people at risk after the CT scan in front of a physician and a neurosurgeon first, you can save a lot of lives. And the parallel I think is valid with space where of course the systems won't have it all right at first every time, but the signal and the early warning they will send will help operators do a more effective job by prioritizing and changing their workflow versus the way they've been working in the past.

John Gilroy:

Let's play a little scenario here. Maybe our listeners can play along with us here. So let's say AI screens and prioritizes information before it ever reaches Earth. Okay. So how do we prevent model assumptions or training data gaps from creating blind spots?

Paul Lasserre:

So as I said, physics creates blind spot in the first place.

John Gilroy:

Yeah.

Paul Lasserre:

And so of course you have got day and light, you've got cloud. So for some sensors, it's a blind spot in itself. I'd say the first blind spot we've had for decades is this notion of real time. You can't really act in a timely manner if you don't get this information in seconds, or minutes in the worst case scenario. It's not been the case. And unlike the US military and when you can have a multiplication of assets, for most



people, most companies, most governments, these real time dimension was a blind spot and it's changing.

I would say the other blind spot you can think about are also, I would say temporal. I don't know if it's a mistake, but the way it's done today when you think AI on Earth observation and geospatial data, not only does it happen after the fact in many cases, but it also happened on a single snapshot. And we love to train model that can do a lot of things on a 10 by 10 kilometer image. And we've got self-supervised learning now that can learn to understand what this type of building is and this type of material is, and what is this situation. And you can ask a lot of questions and get a lot of answers. That's fantastic.

What might be missing is the change over time. Comparing this data over time and specifically with a persistent revisit, and that's where the notion of constellation gains a lot of value. Not only looking at the single image at a given point in time, but looking at the sequence of images, multiple daily passes, multi-sensor and basically analyzing that change. And the notion then of revisit time matters a lot. There is pointless, what we talked about before, to get wildfire alerts in seconds, if you need to wait hours or days before the next satellite revisit the same area. What if the satellite starts in between? So maybe this little game that would be a big blind spot. Sensor vision might be another one, where today in the literature, you mostly have mono sensor analytics, mostly imagery, a bit of SAR, radar.

John Gilroy:

Yeah.

Paul Lasserre:

Few today are training models on multi-sensor data diffusion with not only imagery and radar, but also hyperspectral radio frequency, and that's the type of model we increasingly use at Loft given we have these types of sensors in the same satellites and on constellations of satellites.

John Gilroy:

So, Paul, what does accountability look like, I guess, when an autonomous system in orbit makes a prioritization decision that affects safety, security, or geopolitical dynamics? Who's accountable?

Paul Lasserre:

It's a good question, John. First, I would say that it can't be an afterthought, accountability can't be an afterthought that you built on after launch. It need to be architected into the platform from day one. Coming from the world of cloud myself, and you mentioned AWS, the public cloud companies, and I think AWS, invented this concept of shared responsibility model, which pretty much apply to the space intelligent infrastructure we're building at Loft. Where of course we're responsible for the infrastructure and the operator is responsible for the output, the output of their mission, their payload. And so that shared responsibility model is important.

Now we are paving the way, I think, with the 10 satellites multi-sensor constellation with GPU and real time communication on board, we are launching this year. I believe, tell me if you know of any example, I think we'll be the first company operating at this scale like AI in space across a constellation. And so in a way, we want our best practices to be used to maybe define standard. And I'm sure that what we saw with UBAC software for airplane, for instance, with the FAA, we probably come to space. We don't want



to react to that, but we want to be an active participant and really share what we are learning with practice and not theory at Loft Orbital.

John Gilroy:

So what, if anything, can the space sector learn from how other industries are institutionalizing AI oversight and experimentation?

Paul Lasserre:

I talked about healthcare as good proxy. I think the financial industry, the automotive industry are also good ones that have really gone through these changes over the past decade and adapted and the rules came alongside. Medical imaging, the example I used, I think is good proxy, where there was a lot of hype claims that did not hold like 10 years ago, which eventually matured and led to more regulars validation. The FDA in the US today has a clear AI/ML pathway, and arguably space needs something equivalent as it become a topic of interest.

So I think again, same proxy the people I remember at this time 10 years ago were asking if we will still need radiologist with AI. And 10 years later, it proves out to be the wrong question. The right question is how much faster can AI make the radiologist? And that's exactly what we see. The radiologist plus AI is stronger than either alone. Similarly, we anticipate any analysts using space based system that process multi-sensor data at the edge would be able to make better decisions.

John Gilroy:

Well, this is the Constellations podcast, so we love talking about constellations. I got a constellations question for you here. So as constellations become more distributed and adaptive, at what point does coordination between satellites start to resemble a shared decision making system?

Paul Lasserre:

That's the second leg of this tool I talked about, which until now AI in Earth observation was mostly processed at the edge on the single satellites. And so this notion of coordination across assets was not really a thing. It's obviously changing, and we again paving the way with this first 10-satellite constellation that is going to continue to grow where you want to work backward from something you want to achieve, that is sometimes an objective that is sometimes expressed in plain English, and to coordinate your satellite from forward-looking sensors that can help you recognize the presence of cloud to detection of an object with a large swath or radiofrequency sensor, coordinating and tipping the next satellite to a precise high resolution imagery being taken and coordinate all of your fleets alongside this logic.

And we have a lot of proxies. You open your DoorDash or Uber app and you'll see a lot of drivers and participants to the network. And how do you use algorithm to optimize this distribution is a very well known problem in machine learning. And we use classic machine learning algorithm to do that way better than rules today, which makes it extremely hard otherwise. And so AI is in fact helping us coordinate this work of satellite within the same constellation.

The bigger vision, maybe that's constellation of constellation, as going above and beyond like the name of the podcast is. And for engineers at Loft maybe listening to this podcast, it's not an in-year item, but let's have this North Star where reaching critical mass of assets in space that talk to one another and can



act for the greater good against use case, like fire detection, climate change, pipe leaks, make a lot of sense for a lot of the participants participating to the network.

And so you can imagine this coordination not only happening at the constellation level, but really at a pulling and sharing level of resources in let's say this time and use cases that benefit everyone in the network, where everybody benefits for higher revisit time and a higher variety of sensors and the constellations of consolation.

John Gilroy:

Paul, earlier in this interview, you mentioned models and different models you're talking about. So if satellites start hosting increasingly capable models, then what's the real bottleneck here? Is it hardware, data, or maybe the willingness to share infrastructure across organizations? You kind of touched on that, huh?

Paul Lasserre:

I see maybe three types of bottleneck. One has to do with the nature of this industry. Space is complex. And I spend some time on nuclear submarines in my early days in the French Navy and I often compare the level of complexity that we witness in space with what you have to deal with when you're underwater with a nuclear submarine. And so that is not going to change. And the time to get an asset in orbit, the decision you're making today in terms of what hardware to use and sensors are probably going to materialize at best one, two, three years from now, which is already a [inaudible 00:24:54] time in the industry, and that's what we're pursuing here at Loft Orbital.

That is like one bottleneck. It's a capital intensive industry, which is another one. You can't go and launch your own satellite as easily as you would fire a server on AWS or any cloud provider. And, again, physics doesn't help we talked about this communication today to uplink a large model that you would use easy on your phone of a few gigabytes. It would probably take you longer to applying this model to the satellite than creating the model in the first place. So I would say that's one bottleneck that's not going to disappear overnight.

And the second one is talent. I discovered that you don't have that many people that work at the frontier in AI and know space well. It's hard to find any of this category independently, but people who master both are really a bottleneck today. Most of these people are in the US or in Europe, and we try to work with as many of them quite deeply because they are scarce. The third one I think is mindset, interestingly, where a lot of people in the industry are not necessarily welcoming the idea to work with insights and alert and to move from pixel to events. A lot of people in the industry are used to working with heavy, high quality imagery and data they've been working on, and there is a bit of reluctance to trust a model somewhere remotely that's going to give you some alerts or some insights.

And in a way, I think it's a matter of time and change management, and we have a role to play there. I often tell the team, it reminds me to the 19th century evolution from Navy with sails to steam-based ships, where technology was probably discovered in the late '70s, but only at the end of the '80s and at the end of the 19th century that people stopped sailing and navigating with both chimneys and sails with both systems for decades, because they would not trust...

John Gilroy:

Yeah.



Paul Lasserre:

And they would want to depend on charcoal and then on fuel and they wanted to keep the sails just in case. And I think the same is true here. There is this need to see the data and to verify it, and that's fine. I think we'll start with both the insights that give you real time information on where to look, low quality compressed data, vignette of a ship or a car or whatever event right away, that will help you reorder and change your workflow to get the high quality, heavy data to then go and analyze with the existing process. And in the fullness of time, we'll probably move to a model where people mostly use or maybe only use the metadata, the insights.

John Gilroy:

Coming from a partnerships role at AWS, Paul, do you expect collaboration between companies to increase as technologies evolve in real time or will competitive pressure push organizations to just close off their systems?

Paul Lasserre:

At least what I witnessed is that there is appetite to collaborate. Space is capital intensive, no one can do it all on their own. Right now, AI in space is a hard topic. Many companies, such as the AI Frontier Labs, or the cloud providers you name, like AWS, have appetite to work with space players like ourself who can basically provide them with a new type of infrastructure that allow this real time insight, this always on, orbit patrol type of logic.

So I foresee increased collaboration. There is a lot of appetite across the board for partnerships and to maybe bridge the gaps between the different industries and traditional AI and cloud and space. Of course, we can't ignore the forces in our current world, and geopolitics has been rapidly evolving. And if anything, I also anticipate an increase in sovereign needs. And you see companies, for instance, like ESA, who've been embracing a dual business model with shared resources where different participants are on the same network and people can buy what they need alongside sovereign customers who also want to buy and hire constellation, get full access to their data, and work themselves directly with the partners of interest.

And I think the same is true of Earth observation in general. We see more and more of these increased appetite across the board and from various technology companies, but also an increase in sovereignty demand.

John Gilroy:

Large organizations all over the world typically struggle with what they call legacy applications and moving to something like AWS. And so maybe we have some parallels here too. So where do old pre-AI space frameworks begin to strain under autonomous decision making in orbit?

Paul Lasserre:

Hopefully Loft Orbital is category defining in that space.

John Gilroy:

Yeah.



Paul Lasserre:

And it has nothing to do with me. Head architect in the company has been an advocate from day one of this vision of translating the best practices of DevOps, macro services and cloud computing to space, where you basically want to expose your resources. You want to give access to these primitives, to different participants, to create the right developer tools, and eventually unleash the creativity of higher number of participants. And that's exactly what we do, the standard we are setting how we define our SDK API documentation.

We'll announce with the commissioning of the satellites from this first multi-sensor GPU equipped and real-time communication constellation, the creation of an open marketplace with this standard that basically allow various companies to come and to build on top of it with defined standards and bringing value to the customers in their industry.

John Gilroy:

So if orbital systems gain the ability to interpret and coordinate across the planet in real time, you've mentioned that several times now, what obligations fall on the organizations deploying them? And so who has the authority to enforce those obligations?

Paul Lasserre:

If you look at what exists today, there are a number of building blocks, from NIST, like AI, IMF, the DOD Responsible AI, UAI Act in Europe. We don't start from zero. There are a few framework that apply to AI in general. Space has a lot of regulations that exist, and obviously AI does not change how you abide by them. I think there are still a few things that need to be defined. For instance, we validate every model that is being deployed before they can actually fly on our satellite and be used.

We probably need some space-specific adaptation to this framework. And again, our plan here at Loft is to share these best practices to help define these standards. I just think that we should not wait. It's important that we get out there, see what is possible with Earth observation powered with AI and help define these dos and don't based on actual experimentation and not theory.

John Gilroy:

Yeah. Yeah. Good, good, good. Well, Paul, I think you've given our listeners a great understanding of AIs from space. I'd like to thank our guest, Paul Lasserre, general manager, AI for space at Loft Orbital.

Paul Lasserre:

Thank you, John, again, for having me.