



**SDS PODCAST
EPISODE 763:
THE BEST A.I.
STARTUP
OPPORTUNITIES,
WITH VENTURE
CAPITALIST
RUDINA SESERI**

Show Notes: <http://www.superdatascience.com/763>



- Jon Krohn: 00:00:00 This is episode number 763 with Rudina Seseri, founder and managing partner at Glasswing Ventures. Today's episode is brought to you by the DataConnect Conference and by Ready Tensor, where innovation meets reproducibility.
- 00:00:17 Welcome to the Super Data Science Podcast, the most listened to podcast in the data science industry. Each week we bring you inspiring people and ideas to help you build a successful career in data science. I'm your host, Jon Krohn. Thanks for joining me today. And now let's make the complex simple.
- 00:00:48 Welcome back to the Super Data Science Podcast. Today we are graced by the presence of the incredible Rudina Seseri. Rudina is founder and managing partner of Glasswing Ventures in Boston. She's led investments and/or served on the board of directors of more than a dozen SaaS startups, many of which were acquired. She was named Startup Boston's 2022 'Investor of the Year', amongst many other formal recognitions. She's a sought-after keynote speaker on investing in AI startups. She's an executive fellow at Harvard Business School and holds an MBA from Harvard.
- 00:01:17 Today's episode will be interesting to anyone who's interested in scaling impact with AI, particularly through AI startups or investment. In this episode, Rudina details how data are used to assess venture capital investments, what makes particular AI startups so spectacularly successful, her AI palette for examining categories of machine learning models and mapping them to categories of training data. How well generative AI isn't a fad, it is still only a component of the impact that AI more broadly can make. And she also talks about the automated systems she has built for staying up to date on all of the most impactful AI developments. All right, you ready for this sensational episode? Let's go.

Show Notes: <http://www.superdatascience.com/763>



- 00:02:03 Rudina, welcome to the Super Data Science Podcast. It's been months since you and I first spoke off-air and I've been so excited since to get you on air because you have amazing insights that you share online through your AI Atlas newsletter, which people can get on LinkedIn. Such amazing posts. You're clearly one of the most in touch VCs about what's happening technically with AI. It's amazing to have you on the show. Thank you so much for coming on.
- Rudina Seseri: 00:02:29 Thank you so much Jon, and I love your show and I love what you do ahead of time. I saw all your soliciting questions and input. Love it. It's how you go about it. So delighted to be here. Thank you for having me.
- Jon Krohn: 00:02:40 Thank you. Well, I'm very lucky to be supported by an amazing team here that makes me look good. It's really all the other people behind the scenes that make it happen. Rudina, where in the world are you calling in from today?
- Rudina Seseri: 00:02:53 I'm in Boston, Massachusetts, smack on the middle of Newbury Street.
- Jon Krohn: 00:02:57 Nice. And we were introduced in an interesting way because the Data Universe conference is coming up in April, so about a month after the release of this episode, and Stephanie Christie, who works at Data Universe connected us, said that I've got to speak to you. And she was absolutely right. Our first conversation off-air was amazing and I'm delighted to now be recording your thoughts for our many listeners out there. Let's get right into the topic. So for nearly eight years now, you've been the founder and managing partner of venture capital firm, Glasswing Ventures. The name Glasswing Ventures is inspired by the glasswing butterfly. What's up with that? What's the metaphor behind the naming of the firm?



- Rudina Seseri: 00:03:47 I've been in venture capital for nearly 20 years in one firm or another. Well, two firms really, Glasswing being that second firm that we co-founded. And having the prior experience nearly a decade of being in venture capital, it was an opportunity to be reflective on what did we want this new firm, which became Glasswing to stand for. And Glasswing, the notion of a butterfly embodies the idea of transformation, much like you go from a caterpillar to a cocoon to a butterfly, much like companies, startups in particular, go through their journey of an idea, then a young startup, then a market disrupting and transforming large company. So there was that notion behind it being a butterfly.
- 00:04:37 Glasswing in particular, the species, it's a very interesting one. It's a transparent butterfly. So the idea of transparency with our founders and with our LPs. Venture capital has not always been the most transparent of asset classes, if you will, and we've taken a very different approach to that and that's embodied in our name. And lastly, with Glasswing the butterfly, you can actually see the mechanical structure of the butterfly and it's quite durable. It's clear and transparent, but high durability. So the notion of building a long-lasting firm that will have an impact.
- 00:05:20 Also-
- Jon Krohn: 00:05:20 Very nice.
- Rudina Seseri: 00:05:21 ... I had a little child who was about two years old when I and my co-founder Rick Grinnell founded the firm, and the very hungry caterpillar was a constant staple in our early reading. So that too had an impact in shaping what we went with.
- Jon Krohn: 00:05:36 Very nice. Yeah, transformation, transparency and durability. Very cool. Do you have any examples that you

Show Notes: <http://www.superdatascience.com/763>

can get into of how these principles of transformation, transparency and durability have informed an investment decision or shaped a relationship with one of your portfolio companies?

- Rudina Seseri: 00:05:55 Oh, absolutely. Well, you're Canadian. So in your honor, I'll pick a company out of Toronto, Waterloo area. It's a relatively new investment, a couple of years old called Basetwo.ai. Basetwo is leveraging a particular type of architecture and technique called PINNs, physics-informed neural nets, to solve a very large problem that has existed for decades and that is to deliver a digital twin for simulating process manufacturing, so the engineering piece of it, which is very hard to get right. Think about compounds, whether they're medications, whether it's additives in the aerospace and other verticals. Getting those compounds right at large scales that they get produced is an incredibly costly and difficult effort and undertaking.
- 00:06:53 So what Basetwo has done has leveraged a pretty cutting-edge approach to AI and ML to basically solve that problem by creating a digital twin that requires no code on the part of the process engineers to actually render what can happen in the real world at scale and in turn deliver optimization in this space in the range of 40% of cost savings, particularly driven around savings for energy and how long you run the process. Early stage, but very high-impact product and company. So transparency is the notion that we are incredibly open with each other, we know what our challenges are. I'm their first call, the founder's first call on good news and bad news, and I consider myself, and I hope, I believe that they would consider me as well as an extension to their team to basically take this very transformative idea, technology and execution that they're delivering to create value. How? Well, a number of their customers that they have yet to announce, but think of very large pharma



names are introductions and customers that we have brought to the company all the way to thinking about financing, hiring, and bringing talent across the board.

00:08:16 We have a journey and an exciting journey ahead of us, and it will be a long one. Oftentimes we are in these investments for 5 to 10 to 15 years. So from that perspective, the transparency, the ability to create value and transformation and be around and supportive for the long run. So we go early and we go long in what Glasswing name and logo embodies.

Jon Krohn: 00:08:42 Very cool. Yeah, that is something that I remember from our intro call is this idea of these long-term investment horizons. So if I remember correctly, you'll often be the lead investor in an early stage like C or series A, but then you'll participate in follow-on rounds for years and years.

Rudina Seseri: 00:09:00 In fact, we are over 90% of the time, Jon, we are the first capital in. So we're leading that seed round and taking a board seat and are becoming very, very active with the company with a goal of really shortening the horizon cycle in them delivering the first AI MVP all the way to actually getting the logos and getting the first customers and doing so in a time horizon that's much shorter for the right customer, for the right product than they otherwise would. So if you were to look at the construct of Glasswing Ventures, the firm, we have a team of 12, which for an early-stage firm is a pretty sizable team. And if you look truly around the partners and around the table, we are builders. For every function in an AI native company, we have a partner who's held that leadership role over and over again for that very goal of supporting our early-stage companies to crossing that chasm, the product market fit earlier, especially for AI native companies, which are quite different than SaaS and how they go to market.



- Jon Krohn: 00:10:10 Yeah, crossing the chasm, a key term there. These tricky points in startup growth, it sounds like you have a lot of experience. Your builders, your partners at your firm have a lot of experience with crossing the chasm, and that sounds like it would be invaluable. You hear a lot of... I guess horror story isn't the right way to describe it, but you hear a lot of stories from startup founders that VCs can be particularly at kind of big brands, they can often be just providing capital and little else.
- Rudina Seseri: 00:10:43 I'll talk about Glasswing Ventures. I think it's diplomatic. The mindset and the way we approach it is we've put the capital and now we're an extension to the management team, at least in the early days for that heavy lifting that is needed. How are we going to support them? In fact, if you were to poll the various members of the investment team, they would share with you, I hope two data points. One, when we go into what we call our portfolio company reviews, where we discuss any given company traction, what's working, what's not working. If you want to think of it as a two-page document, the first side of it or the first page of it is what I just outlined.
- 00:11:29 The second page is tracking what have we done for the company lately. And that's also that data point from day one when we make an investment in a company, as I shared, we take a board seat, but we also have principles and associates that are assigned with the partners because we are also training them eventually to become good board members and solid board members and value-add. And before we ever enter a meeting, whether it's virtually or in person, I will ask the question, what have we done for them lately? Just because we put the capital doesn't mean we have earned... We have earned the right from a letter of the law, but truly in substance we haven't earned our right to expect. That's the mindset.

- Jon Krohn: 00:12:13 Beautifully said. Yeah, yeah, yeah. And so beyond these kinds of philosophies and these approaches that you have, you aim at Glasswing to back companies that transform industries. So for our listeners out there who might be thinking about how they could use AI to make a huge impact, do you have any suggestions on industries or problem areas that our listeners could be thinking about to tackle?
- Rudina Seseri: 00:12:35 At the 30,000-foot level, AI is disrupting anything and everything. So it's going to be an all-pervasive, if you will, technology. Almost like a slice that cuts across all facets of our lives, both at the B2B and enterprise and security level and in our every days as consumers. So, I will comment it from a Glasswing lens, which is we focus on B2B and security. We're also thesis-driven. So the areas that I find most interesting right now have to do with data infrastructure and data, the data backbone by infrastructure done here, hardware. Here we have so many breakthroughs on the algorithmic and the model side. How do we get our data to be ready, cheaper, faster, and truly turn it into an asset?
- 00:13:28 If you look back to the last 10, 15 years, this whole notion of digital transformation, the impetus for that is really, or the underpinning of it is data as an asset and then digital transformation gets used as a way to actually drive transformation within the enterprise beyond data. But the data side of machine learning and AI, if you want to think of AI as the output is the what, it's a key area that there are still a lot of opportunities.
- 00:13:58 Another area is driving productivity. In a software sense, in an AI native software sense that typically lends itself to functional areas. Think about the breakthroughs we've seen around development and coding, for example, at least across our portfolio. Now mind you, our portfolio of companies are all AI native, so they might be ahead of the



curve in that regard, but still about 60-plus percent of the code is actually AI generated and it's a balance of nearly 40% that is human-driven and highly differentiated.

00:14:37 Similarly, I can think of another company, Reprise in our fund one. They're the category maker and leader in demo automation for sales. A problem that has existed forever, which is how do you perform a demo, give a demo, put a demo in the hands of your prospective buyer without it crashing or without needing sales engineer, they have cracked that problem. So there's a lot of opportunity to leverage ML techniques and algorithms around driving productivity.

00:15:09 A third area perhaps, and I'll stop at it, there are a number of other areas, but it's around vertical industries. So how do you deliver not just increased efficiency? During the cloud and SaaS era, we fundamentally brought efficiencies to market across different vertical and horizontal platforms. With AI, we have the opportunity to drive optimization. So supply chain, manufacturing has been an area of high interest for us, a vertical if you will, as has insurance tech to a more recent extent. But supply chain, manufacturing, inventory management, logistics. Those areas if you want to think of them as all the industries are ripe not just for efficiency creation by way of AI that you previously couldn't, but optimization.

Jon Krohn: 00:16:06 2 days, 50+ presenters, unlimited opportunities to connect with a global data audience. The annual DataConnect Conference brings together industry leaders, technical experts and entrepreneurs to discuss the latest trends, technologies and innovations in data, analytics, machine learning and AI, all while amplifying diverse voices in the space. Join us in Columbus Ohio July 11th and 12th to hear from incredible speakers like Carly Taylor, Megan Leiu, Olivia Gambelin and more! Can't make the trip to Ohio? Join us for the 2nd annual

Show Notes: <http://www.superdatascience.com/763>

DataConnect Conference West, hosted in Portland Oregon on May 8th where you'll hear from AI experts from the likes of Microsoft, Alteryx and Women Defining AI. Save 15% on your DataConnect Conference pass by registering with the code superdatascience.

00:16:54 Makes so much sense. Yeah, so there's opportunities for productivity improvements across the board, but there's also opportunities in specific verticals. So having a kind of niche and building AI solutions for that specific vertical in your case, a supply chain management is something that is a particular interest right now with tons of opportunities for optimization there.

Rudina Seseri: 00:17:03 And if I may build on that, niche shouldn't be understood as small opportunities. Should be understood as focused and large opportunity.

Jon Krohn: 00:17:12 Nice. Yeah, that is a great way of saying it. In your experience, are there types of problems that end up being more ideally suited to AI solutions? So you talked about AI disrupting anything and everything cutting across all industries, but are there particular types of problems? Are there ways that we can think about, oh, this is the kind of problem that's going to be great for solving with AI today or in the coming years we could build a solution for this? Whereas other kinds of problems where you think, well maybe this is something more in the distance that we could have as a vision slide in our pitch deck?

Rudina Seseri: 00:17:45 Yeah. Unfortunately, I cannot give you a catch-all answer. I will give you the cop-out, it depends, and then substantiate it with some specific areas. So what type of AI or what type of machine learning really is in the how of it? If you think about generative AI, I waited how many minutes before saying that? 20 minutes, 15 minutes? If you really think about areas like generative AI, and really if I talked more specifically about the techniques used

around transformers and attention type architectures and techniques, their application needs to be in instances where the error rate is tolerable. And what I mean by the error rate, I'm speaking to facets such as hallucinations and errors of omission. So if you want to use it to summarize content, make the interactions more human-like, go with 'God', it will be incredible.

00:18:49 If you were using it for something that requires high precision where life or death depended, would I want any kind of that bucket of techniques to be used in any way for someone to perform surgery on any living being? Goodness no, because accuracy is still a high challenge. So that's why I gave you the answer. It depends. In other instances and other techniques where accuracy is more important, there are other architectures which would be better focused and better suited for those kinds of use cases. But in general, we are at an era where it's a human in the loop or a co-piloting relationship between humans and machines or humans and algorithms. And if accuracy is needed to my second point, then the narrower the domain that it is focused on, the higher the accuracy. So you have a inverse correlation between the broadness of the use case or the breadth of the use case and the accuracy that you forego.

Jon Krohn: 00:20:01 That's a great way of framing it. That makes a huge amount of sense. I like that cop-out answer of it depends, of course with AI.

Rudina Seseri: 00:20:08 It truly does because can we take a step back if we may, Jon? We use AI as this monolithic notion and it is not. If someone said to you or me today, "I have an AI company," let's congratulate them on their entrepreneurial spirit and pose the question of what they actually do. Because again, as we talked at the beginning of the podcast, this AI cuts across the board. Furthermore, AI is the what, the intelligence that comes with the output, the how being

one of the areas being machine learning. There are other facets of AI, but usually machine learning and AI and everyday lingo are used interchangeably, albeit it's only so correct to do so. But let's roll with that. Even machine learning is not a one-dimensional effort. If you think about the last 70 years, we've had AI since the 1950s with the Turing's test, but really we've had classical machine learning till about 2006 in academia, 2010 in industry.

00:21:22 Then in 2010 we had the rise of neural nets and deep learning. Within just that facet, you have architectures like transformers and attention, you have recurrent neural net architectures which have memory. You have forward feeds, where the data moves one directionally. You have CNNs, convolutional neural nets, which are really good for vision, graphic, et cetera. So GNNs and others. And then within each of those you have different types of techniques that you can leverage. And really if you're sophisticated, you're leveraging ensemble models which combine techniques and architectures from both classical and deep learning. And then you have to figure out what training approach you're going to take and what data set is this best matched.

00:22:12 So when we talk and we wave our hand and say AI, I challenge all of us to pause for a second and say, wait a minute, there's a whole universe that needs to be understood and correlated appropriately to the data availability, use case and problem that one is solving. That's my pitch for the day on understanding AI.

Jon Krohn: 00:22:33 That was incredible. That was an amazing whirlwind tour of AI over the past 70 years now and the opportunities that lie within it. I was almost mesmerized as you were speaking, nailing everything right on the head and I was thinking to myself-

Rudina Seseri: 00:22:51 Emphasis on almost...

Show Notes: <http://www.superdatascience.com/763>



- Jon Krohn: 00:22:55 And yeah, I was thinking this sounds like the kind of venture capitalist you'd really want to be working with if you're an AI company because you clearly... I knew that coming into this, but if it wasn't already obvious to our listeners, you just made it abundantly clear that you [inaudible 00:23:09].
- Rudina Seseri: 00:23:09 Some of it we'll blame on intelligence, but some of it is hard learned and experience. My first AI native investment was in 2010 and we started Glasswing in 2016, like I said at the beginning of the show with a notion that it was all-pervasive. So I've made every mistake in the book and hopefully learn from those as well. And by the way, to the point of it's a lot to digest. If interested folks want to learn more, we actually open-sourced one of our frameworks called the Glasswing AI Palette, where it maps the major, not all, but the major architectures and techniques within those architectures and gives you a high level framework on how you map that to the right data sets and the right use cases. So if folks want to learn more, it's on glasswing.vc for venture capital, it's called the AI Palette.
- Jon Krohn: 00:24:06 Yeah, we will be absolutely linking to the Glasswing Ventures AI Palette in the show notes. I was checking it out just before we started recording and it is magnificent. So it covers across classical machine learning techniques like regression classification, clustering, deep learning approaches like the ones that you highlighted when you were speaking a few moments ago. So things like recurrent neural networks, convolutional neural networks, Raft networks and transformer architectures and attention. It talks about feedforward neural networks and ensemble methods. So yeah, a lot of what you were just describing, people can get lots of detail on that if you're not already familiar and some ones that actually we haven't mentioned yet on air but are also super powerful, like unsupervised learning for unlabeled data

Show Notes: <http://www.superdatascience.com/763>



sets, semi-supervised learning to blend supervised and unsupervised approaches. You even get into reinforcement learning a bit. This really covers the gamut.

- Rudina Seleri: 00:25:08 Yeah, it covers, because I think when we get into supervised, unsupervised or semi-supervised, we're actually talking about how you are training the algorithm. So think of it as a framework, architectures techniques, how you train them, what are the various combinations, and also zero loss minimization, and then you tie to the data sets and use cases that are generally more applicable. So, that's why I didn't use the term framework lightly. It's actually a framework that I hope we open-source to make it valuable for the ecosystem and the startup community.
- Jon Krohn: 00:25:50 Something that seems immediately useful to me to startups is you could be thinking about what kind of data do I have? What kind of moat do I have from the data that I've been collecting from my users? Refer to your AI palette. And then from there, think about, okay, by the mapping that you've already done the hard lifting on, mapping data to different kinds of machine learning techniques, you can think, okay, I have this particular kind of data that maybe nobody else has out there from my users, from my platform, or from some other approach that the startup has used to collect those proprietary data, go to your palette and see how can I be leveraging these data that I have for some kind of machine learning model that provides value? Yeah, very cool. Thank you.
- Rudina Seleri: 00:26:31 Exactly. Or you are a founder who comes from a particular industry that has identified a deep problem with a big budget if you're B2B and in need of a must-have solution that previously wasn't solvable. Now with the advances on run machine learning and AI one can, so you know what the use case would be, you don't know where to start from there. You can also start from the use

case and say, what kind of data do I need and what would be the best machine learning approaches to tackle this problem? So one can pick the threat from any possible side, whether it's the algorithmic side, the data side, or the use case side.

- Jon Krohn: 00:27:11 Very nice. Yeah, so again, we will have that AI palette in the show notes for everyone to check out. Speaking of data Rudina, from your perspective, how data-driven is the venture capital industry?
- Rudina Seseri: 00:27:26 So it depends on what you mean by data-driven. Is it performance a hundred percent data-driven I hope? Or is it how much of our decision-making is data-driven or informed? So venture capital is an interesting industry in that the later stage one is the more data you have available. If you think about a series B and C sort of almost entering the growth bucket of venture capital, you have a ton of financial and otherwise data points and information around performance from revenue metrics to the magic number to the rule of 40 to customer retention, expansion and all the various inputs. So you can very much be extremely data metric driven and perspective in that regard.
- 00:28:26 If you are early stage first-capital-in, the asymmetry of information is much bigger and your data is of a different kind. So for example, within Glasswing, we have our own ML capabilities, but it would be a big euphemism if I said that all decisions are driven by our ML capabilities in how we're assessing founders and companies. We don't have enough data points and et cetera, but it informs us. And where it's really more valuable is not in the decision making, at least not yet, of should we pursue this investment or not, or should we invest in this founder and company or not. It's much more actually in signaling. So we found it to be an extremely valuable tool as we source and giving us a signal about this founder is



potentially starting a company and/or in recruiting talent. It's also been valuable in that regard. That's why half-jokingly I said sort of which data.

00:29:33 And of course at the end of the day we manage capital on behalf of our investors that run the gamut from academic institutions, foundations, pension funds like retirees, family office, what have you, and ultimately it's how well you do on behalf of your investors by backing this exceptional founders building large companies. So that's the ultimate data and performance based on data. But between those two data points, there's a pretty long time horizon. So there are other indicators in between that one needs to leverage.

Jon Krohn: 00:30:12 Nice. So what are those indicators?

Rudina Seseri: 00:30:14 So it depends on the stage that the company is at. For example, as I talk about the underpinnings of our performance, which would be the portfolio companies, if it's in the earliest days, it's the ability to go from product development to product launch, and getting the right design partners and launch partners and ultimately early customers. So how is that tracking? Then along that journey is truly making the transition from pre-product market fit to product market fit. So then all of a sudden once you have hit that product market fit stage, you start thinking about revenue and ARR and bookings and how that will scale. So pricing models start to become very, very important as one refines those... Pricing and business models I should say more inclusively and how those start to scale and relative to how one is scaling costs, largely people initially and largely on the development side. Then in building out the go-to market engine, efforts and team, how those are tracking. And then increasingly we get into that stage of highly quantitative and metrics driven in the financial sense of the word. So for different stages there are different sets of



metrics and data points and early indicators of problems versus strengths.

- Jon Krohn: 00:31:45 Research projects in machine learning and data science are becoming increasingly complex and reproducibility is a major concern. Enter Ready Tensor, a groundbreaking platform developed specifically to meet the needs of AI researchers. With Ready Tensor, you gain more than just scalable computing, storage, model and data versioning and automated experiment tracking. You also get advanced collaboration tools to share your research conveniently and securely with other researchers and the community. See why top AI researchers are joining Ready Tensor, a platform where research innovation meets reproducibility. Discover more at readytensor.ai. That's readytensor.ai.
- 00:32:26 Nice. And I could see how you might be able to dig up some data for some of those indicators like building the go-to market engine. You could have things like the lifetime value of clients starting to work those at the customer acquisition costs and then those ratios can end up being really great. So if you know that the lifetime value of a customer is likely to be a multiple of your customer acquisition costs, then even if you're still in relatively early stages of finding that product market fit, you can have some pretty good indicators of the go-to market engines.
- Rudina Seseri: 00:32:59 Other than that Jon, if you haven't found the product market fit, it's practically impossible to really determine the lifetime value of a customer because you still don't know who's your right customer. So be careful... Not to you, but in general, one should be careful with when one makes those calls. And then in the spirit of RAI focus, it's also important to know that AI companies have inherently different challenges in the early days than the run-of-the-mill software company. And where I draw the distinction

Show Notes: <http://www.superdatascience.com/763>



is the fact that if you have a software company, let's say it's a SaaS model, you develop it, you test it, you get your beat out, it's in the hands of the early customers, you are getting feedback, you're tweaking it, but it's kind of mostly done. I mean, over time you might build or make changes to have more scalable software, et cetera, but fundamentally it's done.

00:34:02 With AI is different. With AI, the first day that an AI native product or platform is being used, it's the least valuable day because it's learning from the data. At the same time, it has to be valuable enough for the customers or launch partners to want it. So it's a really interesting balance in terms of output and what we call internally the AI MVP and the value that it delivers. And if it's valuable enough, all else equal, it should continue to improve over time as you feed the algorithms more and more data, particularly if it's the right kind of data. And there could be a whole show on what constitutes the right kind of data, but really, really important to have that distinction and to understand also from a team perspective... I went straight to the output, but an AI-native company doesn't just have the co-founder do all... stick with two of the CTO and the CEO, sort of the technical and the go-to-market leaders.

00:35:12 It also has the data scientist or the chief AI officer, whatever the title is emerging, whether that person manifests itself in the form of the CTO also having a data science background. Hard to find those, but not impossible, but more typically having a third leg to the stool of having a data scientist. And oftentimes these folks, at least historically were academics, so striving for perfection while the CEO is striving for an MVP. So how do you align everybody's priorities? It's actually a very interesting journey and distinctly different in the early days than the typical software company.

- Jon Krohn: 00:35:56 Very cool. Thank you for highlighting those specific issues that AI companies have relative to many other kinds of SaaS businesses. Specifically, I noted down here how an AI company's least valuable day can be their first day after go-to-market launch because the AI MVP at that point obviously must be good enough to get some kind of traction, but as more proprietary data are collected that AI MVP will become less and less of an MVP and more and more of a fleshed out note of an AI model.
- Rudina Seseri: 00:36:25 And the MVP for an AI company has to be more fully baked, if you will, than a typical MVP because it has to create value.
- Jon Krohn: 00:36:36 Makes a lot of sense.
- Rudina Seseri: 00:36:37 Albeit it has a long way to go on its maturity, but yeah.
- Jon Krohn: 00:36:41 Yeah, that makes a lot of sense. With our own company, Nebula, we've had to invest a lot of time to get our AI models exactly to a point where it's good enough for our customers to be using. And yeah, I can imagine how that took longer than it would've in some other circumstances where it's more about standing up a platform and it being able to run.
- Rudina Seseri: 00:37:03 And to build on your point, Jon, that comes full circle to there are hundreds of big and small decisions that are being made around what architecture, what data do we have, what training, what techniques. It goes back to the discussion which we don't need to revisit, but it's not a well-defined science. There's a lot of trial and error and art and science in combination that goes into it. So I don't think we should lose appreciation for it's not an easy thing to do. If you are truly AI native rather than you are putting what we call a shiny wrapper with a ChatGPT-like interface, that's a whole different... It's very valuable.

That's not a moat and doesn't make you an AI-native company.

- Jon Krohn: 00:37:44 Yeah, it's a nice distinction there. An AI native company being a company that's developed its own proprietary AI models as opposed to just leveraging an off-the-shelf third-party API.
- Rudina Seseri: 00:37:53 Yes and no. An AI-native company being one that's not just looking at the GenAI and large language models, but it has something that's proprietary to itself, whether it's on the algorithmic side at the foundation level or more probably if we're talking within the world of GenAI, which again is only a subset, but more its own small models on top with its own proprietary data, or at least data that's differentiated from the rest of the web that the large language models have been developed in. And beyond GenAI, of course, you have all those other techniques and architectures that can be native and proprietary. So point being, I don't have anything against algorithms that are open-source for AI, but then it's what do you do with them and how do you build modes? And that's a very different consideration than having an API call to one of the language models for your interface and user experience.
- Jon Krohn: 00:38:53 Very cool. Yeah, thank you for clarifying that. Speaking of generative AI, I know you've already alluded to that generative AI shouldn't be what people think of as everything these days, but we have had an explosion of generative AI models and applications. Especially in the last year, capabilities since the release of GPT-4 have been, in my view, staggering. I personally went from being skeptical about whatever AGI is happening in our lifetimes, but the release of GPT-4 made me think, whoa, something like whatever AGI is could maybe happen in our lifetimes. So I think there has been this explosion, I think that it relates to real-world powerful capabilities we

see in foundation models like GPT, Gemini, Dolly. Do you think that foundation models like these are passing fad or are they an integral part of the future of AI?

Rudina Seseri: 00:39:52 Oh, no, no. I think generative AI is highly transformative and the word fad shouldn't enter the vocabulary. I'll talk about the limitations. I'm going to break your heart, but I don't think generative AI is leading to AGI anytime soon, and I can at least give you my view of why that is. But let's put that as an aside for a second. Listen, what generative AI has done for us as a society is delivered a whole new way to interact with technology and software, whether it's software standalone or wrapped in a device robot, some form of hardware, whatever one might want. We fundamentally can interact with technology in a human-like manner. And as long as we have fault toleration, again, back to an earlier part of our conversation, it can do summarization, it can, I joke that you know a technology has crossed the chasm and is here to stay when grandma wants to hug the computer because ChatGPT told such a cute little joke and your 10-year-old thinks that they don't have to write any original paper, ChatGPT is doing it for them.

00:41:02 So we've crossed that chasm of, I'm using the term three times now, but of mass consumerization. But at least early breakthrough in that regard is with user interfaces, user experiences, and multimodalities for content. Hugely transformative. In fact, if a software company from here on doesn't have some GenAI facet to their interface and user experience, they will fall behind no matter what, because that will be a consumer and an enterprise expectation. AI is bigger than that, and the reason why GenAI has taken off is because what we just outlined, which is it's tangible. Everyone can interact with it and see the immediate impact. When you combine gene AI with some of the other architectures... And it's making its way into the other architecture, so it'll be interesting to

see where we are five, 10 years from now. I think it'll be impactful, but there are many, many, many, many use cases that don't involve GenAI that are going to be and have been incredibly transformative. I think OpenAI and in particular ChatGPT drew a lot of attention to the space and there are questions to be had. So I don't think it's a fad, but I have questions around what the steady state will be.

00:42:32 If you think about the large language model or foundation model providers, whether it's OpenAI, whether it's Google, whether it is Anthropic, Cohere, what have you. First of all, it bears a lot of analogy to what we saw in the cloud because even though some of these names may feel very start-upy, really behind them you have the large existing incumbents. You have Microsoft, you have Google, you have Amazon. Google is twice both as its own platform and the companies that it backs. Amazon, Salesforce, Oracle, I can go on and on. That foundational or more accurately perhaps that infrastructure layer is very capital intensive and therefore it's been backed by the incumbents where on a relative basis the capital abundance is available.

00:43:35 There are then a ton of opportunities well above sort of in that middle layer, but really more importantly in the application layer within GenAI, where you can develop both proprietary algorithms and sort of your own prompting and whatnot to be differentiated. In fact, again, I made a reference to the cloud. If you think about what we saw in the cloud era for every AWS and GCP and Azure and a few others that are farther lower on the list, you had hundreds and hundreds of companies that became SaaS applications that became multi-billion dollar companies. So important to differentiate between the two because I will at times or more times than I care get the question, "Is it game over because OpenAI will do everything?" And it's like, oh, no, no, no, no.

00:44:26 Additionally, it begs the question of where is the differentiation at this infrastructure layer between this sort of... It has oligopoly-like characteristics and open-source. I don't think Meta is a dumb player and they've gone fully open-source. I think one should watch them carefully. Will this be the infrastructure that at some point is a dumb pipe like... Will there be a higher differentiation between an OpenAI platform and a Google platform or an open-source platform? Will it be driven by performance? Will it be driven by pricing? At some point, ultimately, if the web is the entire data set, we will be trained on the same data set. So where will the differentiation come? Will they be commoditized? I think there is a high, high risk of commoditization. I could argue, albeit I don't know that that's why OpenAI is trying to enter so many adjacent areas and areas higher up on the stack and lower with what they're doing around compute and chips, et cetera, to perhaps try to build a differentiation. But I have fundamental questions around differentiation for this major infrastructure players. I'll pause at that.

Jon Krohn: 00:45:49 Eager to learn about Large Language Models and Generative A.I. but don't know where to start? Check out my comprehensive two-hour training, which is available in its entirety on YouTube; yep, that means not only is it totally free, but it's ad-free as well. It's a pure educational resource. In the training, we introduce deep learning transformer architectures and how these enable the extraordinary capabilities of state-of-the-art LLMs. And it isn't just theory. My hands-on code demos, which feature the Hugging Face and PyTorch Lightning Python libraries, guide you through the entire lifecycle of LLM development—from training to real-world deployment. Check out my “Generative AI with Large Language Models: Hands-On Training” today on YouTube. We've got a link for you in the show notes.

- 00:46:32 Very cool. Well, so I guess there is heartening news there for people who are listening and want to be creating really valuable SaaS businesses at the application layer because it sounds like there is a smaller risk of commoditization that you see in the infrastructure layer.
- Rudina Seseri: 00:46:49 Correct. Well, and very few players in that. I mean, you have to be an incredibly well-funded beyond what we've seen typically company to play at the infrastructure layer.
- Jon Krohn: 00:47:01 Yeah, yeah, yeah. Billions. Billions.
- Rudina Seseri: 00:47:04 I heard OpenAI trillions. There you have it.
- Jon Krohn: 00:47:10 Yeah, we're getting there for sure. All right, nice. So now we've covered a lot of questions for people out there who might be thinking of their own AI startup ideas or growing their own AI companies, what about for enterprises adopting AI? What kind of advice do you have for our listeners who are interested in adopting AI services that others are providing? For example, the kinds of services that companies in your own portfolio are providing?
- Rudina Seseri: 00:47:35 So it's interesting because AI adoption in the enterprise has started and there is no going back. It's actually quite exciting. I think I'll start perhaps with some forecasts and data. I think our data indicates that by 2030, the AI market size will be about 1.6 trillion, and that is equal in size to what the cloud market will be. So put differently, I think cloud adoption, the enterprise has been pretty hefty pace and AI adoption will just about double the speed of how cloud was adopted. So it's very exciting from that point of view.
- 00:48:27 I think what I would encourage enterprise consumers or enterprise users of AI and buyers of AI products is to not buy AI for the sake of AI. So every enterprise has AI as one of their top three priorities. What are we doing

around AI? And oftentimes two execs talking to each other mean very different things around AI. So getting that lexicon and understanding what exactly one is looking for is really important. But then also thinking about going to the fundamentals of what problem am I solving? Can I solve it without AI?

00:49:10 Listen, just because AI is sexy and the flavor du jour, if you can solve it in a simpler way, please do it. One should adopt AI for areas that were previously unaddressable and unsolvable where there is a real differentiation and a real impact and result as one uses or adopts an AI native product. So asking the fundamental questions impact, effort, ROI, is really important. I think at the risk of stating the obvious, starting with areas where you have more visibility, where you can tolerate more fault, more areas is important, and then doing the crawl, walk, run, and that will vary from business to business.

00:49:57 You also have to get a balance between the CIO organization, the chief information officer organization that's driving this or the part that's driving it with also the chief risk and the chief security officer that has to deal with the notion that AI widens the surface area, if you will, that needs to be protected. And so how do you balance between the, let's get this done versus the naysayers? Both groups are actually trying to do well by the enterprise, do the right thing. How does one balance those priorities? It's not a trivial exercise.

Jon Krohn: 00:50:35 Yeah, no doubt. I think you've touched on a lot of the key complexities there though in your response. Do you think that... I mean, I guess this is going to be another one of those questions where it depends, but do you have any ways for enterprises to be thinking about whether they should be using foundational models or be training their own specialized models, like maybe taking a 7 billion parameter Llama 2 model or what have you, and fine-

tuning it to their own use case and running it on their own specialized infrastructure. So given those two possibilities of using something like an Anthropic or OpenAI or Google API off the shelf versus fine-tuning your own LLM using an open-source starting point.

Rudina Seseri: 00:51:18 It all depends. I really keep giving you [inaudible 00:51:22] answers. Listen, it depends on the DNA of the business and the construct of the organization. If they're a financial institution with 40,000 developers, they can open-source it and go on their own and build from there. If you're an enterprise that doesn't have that kind of talent pool and DNA in-house, then the historical expression used to be no one got fired for hiring IBM back in the day, well, no one will get fired for using the Microsoft enterprise-ready seal of approval. So from that perspective, you might be more inclined in leveraging an incumbent set of tools and then innovating on top of those and both work just fine. It depends on your priorities. So I don't think there is a religious answer to your question of this way and no other path. It's a question of what are you solving? What is your risk appetite? What is your talent DNA, and how do you get around it?

00:52:31 The number one concern that enterprises have is they don't want OpenAI or other models to be trained on their specific data. So how do they keep the data within their firewall? How do they maintain that advantage which they perceive they have? They may or may not have fully contextualized and quantified with that advantages, but they know instinctually that it's an asset and that's why you saw the rise of vector databases quickly sort of overshadowed by existing databases giving those capabilities. So it's an interesting journey. I think you can get to a strong or bad outcome in multiple ways. There isn't one side fits all, but again, for more conservative or sort of followers in the adoption, I think enterprise-ready

seal of approval by well-respected and understood incumbent gives comfort.

- Jon Krohn: 00:53:31 Well, it was another it depends answer. You did provide a lot of great ways of thinking about the problem there. So for example, if you are a much larger enterprise with talent in-house, then it's something you can be thinking about more often, fine-tuning your own open-source model. But absolutely, I think for most listeners relying on these enterprise-ready certified APIs is a great way to go for a lot of generative AI capabilities. It is. I continue to be mind-blown on a daily basis. I say this on the podcast all the time, but on a daily basis, I'm blown away by the breadth of tasks that I can rely on the cutting-edge LLMs of today, proprietary LLMs like Claude 2, GPT-4, Gemini.
- Rudina Seseri: 00:54:17 And to build on that, ultimately with GenAI and beyond, I keep pushing you and challenging you to just look a little bit beyond GenAI, but as you see the complete picture, we're really going to enter this paradigm of ambient AI, if you will, where we're going to have our own personal co-pilot. You can think of them as agents, you can think of as a co-pilot lingo. One can use whatever terminology pleases them, but fundamentally it's going to be this symbiotic relationship where AI is all pervasive and all around and we leverage it in every facet and do so in a very cohesive and orderly manner. I think we're definitely heading for that. I don't know that there's AGI. I am going to come back to the AGI now because I remember that. All I will tell you on AGI is that just because it feels like a human and you get responses that seem to have human-like intelligence, does not AGI make.
- 00:55:25 If we wanted to go and talk down and dirty on how transformers and summarization works and vectors, you will know that there is limited ability still that will change over time, but limited ability in terms of reasoning, there is no reasoning that's going on with generative AI with

these models. Its summarization is largely... I mean it's entirely statistics driven, so if you and I trained the model, then it saw the two plus two equals four and it had never seen that five plus five equals 10, it can't extrapolate. It can write incredibly complex content but can't reason and extrapolate. And I think those are some fundamental capabilities of true intelligence and general intelligence. So I'm excited about what's coming down the pike. I'm not sure that it's AGI, at least not in the time horizon that has gotten you all excited. So here I am crashing the joy.

- Jon Krohn: 00:56:37 No, no, we don't disagree on any points. We are absolutely on the same page. I don't think that AGI is around the corner, but I found even though you're absolutely right, what's happening under the covers is not sophisticated reasoning. It is transformer architectures and it is relatively simple probabilities scaled up massively. But I didn't expect that even such a simple approach could create such astounding results. And so it made me a believer in the idea that while yes, the underlying architecture, I think it needs to be more sophisticated, that machines scaled up sufficiently will be able to do incredible things possibly AGI in the coming decades.
- Rudina Seseri: 00:57:40 Yeah. Listen, I think we're moving very fast and innovation and disruption is not linear, so not always easy to predict as a result. I am incredibly excited about what's coming. I think I just want to make sure that just because it feels human-like, we don't therefore conclude that it has the same level of reasoning, extrapolation and intelligence as the human mind, at least not just see it in the current form.
- Jon Krohn: 00:58:09 Yep. Yep. Well said. And yeah, I was going to get back to the AGI. I wasn't going to let you off the hook with that. We were going to get back there, but it's nice that you

brought it back up again. And so we were talking about the enterprise going into that bit of a gear change here, and maybe this is now not necessarily directly related to generative AI, so that might bring you some relief. But in talks you've given in the past, you've talked about model collapse, which isn't something we've talked about on the podcast before. So how concerned are you about model collapse and what can be done... I mean, first of all, what is it? How concerned are you about model collapse and what can be done to mitigate it?

- Rudina Seseri: 00:58:56 The fundamental challenge around model collapse is the notion that without human-generated training data, so if it's all synthetic data, AI systems malfunction. And at a large, large scale, this could be a real problem if the internet becomes flooded with AI-generated content that's synthetic data-driven. How worried am I? Listen, it's not a new problem. Let's start there. I mean, we've had this problem with content farms long before ChatGPT, long before GenAI, and so it's an understood problem and we have at least some prior precedence on how it has been handled. The concerns around, again, singularity, I think we can mitigate first by developing approaches and ways to differentiate between artificially generated and human-generated data.
- 01:00:06 And to that end, a very nice precedent is sort of what we've seen historically 10, 20 years ago with what Google did with Google search, where different sources of content had different rankings in that context. So we can certainly tweak and change the weights on the AI side to be able to give more weight to human-generated data versus the synthetic ones and be able to differentiate that way. So it is a challenge to be overcome. We've faced it previously in the web era and other technologies, and I think we have some models on the basis of which we can innovate to address it.

- Jon Krohn: 01:00:58 Nicely said as with everything else that you've said so far in this episode. That was a great definition of model collapse and something to be potentially concerned about. It is interesting how it is a tricky problem to be able to distinguish human-written copy from machine-generated copy. And the better that the algorithms get... There were the kinds of tools like ZeroGPT that were pretty good at distinguishing GPT-3.5 outputs from human outputs, but that then became a lot trickier with GPT-4. And so it's a-
- Rudina Seseri: 01:01:32 I mean, I'll give you an example that's not exactly addressing this notion of model collapse, but certainly determining human-generated versus AI-generated algorithms and sort of code. We have a company that I will not name because they're still in stealth. I'll come back to you when they come out of stealth and we can update it, but they come at the world from a compliance risk and security point of view, going back to the CISO has to contend with, and the head of risk and whatnot with regulations that are coming right and left around the proper use of data and proper use of AI.
- 01:02:13 So the particular company, we incubated it actually, basically goes into the enterprise and it's able to deliver an assessment of what is code that's AI generated and what's human generated and can track it. And again, they've started at it from the point of view of compliance, risk management and security, but one can quickly see how it can extend into productivity. How can it extend into starting with the first basis of assessing and knowing how much of your information is synthetically generated versus not, and then building from there in mitigating for it. So it's an area of interest for me and for us I think at Glasswing, and I suspect it'll become more front of mind for enterprises as adoption increases.
- Jon Krohn: 01:03:12 This is definitely changing gears now. I said that we might be changing gears a bit on the last question. It ended up

not really being so, and it was more generative AI than I expected it, but really changing gears now and opening up to audience questions. I announced on social media that I would be interviewing you as I often do for many of our guests that I think there might be a lot of questions for, and we got way more questions than usual for you.

- Rudina Seseri: 01:03:42 Woo-hoo! No, I'm kidding.
- Jon Krohn: 01:03:46 Yeah.
- Rudina Seseri: 01:03:46 That's not acceptable reaction on a professional podcast, you said.
- Jon Krohn: 01:03:52 I don't know what it means exactly that that happens. I actually mused on this same thing in our preceding guest episode with Lisa Cohen, episode number 761. She was responsible for Google Gemini's release to the public. And so absolutely fascinating guest, had a huge amount of reactions to the announcement that I would be having here on the show, but not many questions, whereas you got tons of questions. I don't know. I think part of this might be that there's a lot of people out there who are interested in how they could maybe themselves be growing an AI startup, getting investment, these kinds of things. And so we do have, I mean, Mohammad Raza who does audit data analytics at BDO Advantage. One of his questions is one that I was going to make sure you didn't get away from this podcast interview without answering, which is, how do you, what do you look for in potential investments in AI startups that you're looking at? So you must get hundreds of times more pitch decks thrown at you than you actually invest in. What is the difference between the companies that you invest in and that you don't?
- Rudina Seseri: 01:05:15 So some things don't change. Others do. In the categories of some things don't change, the number one thing or

facet that any venture capitalist, focus on AI or not, looks for is an incredibly strong founding team that doesn't just have a strong vision for the company they intend to build, but also unparalleled execution capabilities. The ideas are important. Execution is even more important. There is no wave of disruption, at least not to date, that has done away with that need. So I look for that every single time, and it manifests itself in different ways. If it's a first time founder, it's the hustle, is what else they have done in life and how they set XYZ mark and then they beat it. If it's a repeating founder, they have different characteristics and then other proof points, if you will. But I look for that execution ability or extra amazing execution ability in every case.

01:06:37 I look for founders that understand the problem that they're solving, very, very well. And understanding the problem well is an interesting notion. It could mean that they come from industry and from the industry that they're trying to solve the problem for, or that they are an outsider in. So they function as a disruptor to how the way things have been done in a particular space. But they also surround themselves with outsiders who are the adopters to get that perspective. Both work, both fail, but both work, but importance around understanding the space. And here's why. You could have an incredibly strong tool and platform in your view or in absolute terms, and I'm going to give you an example in a second, but it fail because it doesn't follow the workflow of the users or adopters, or it requires a mindset change or behavior change. Those are hard.

01:07:46 So one of our partners at Glasswing is Vlad Sejnoha, the former CTO of Nuance and sort of godfather really to natural language understanding and speech recognition. And he often points to the example of medical transcription where in the early days there were tools that could do away with a human transcribing the dictation

notes of doctors and doing away with that by leveraging software that was 80-plus percent accurate. And you think, "Great, 80% accurate. Oh my goodness." If you talk to Vlad, he tell you that it was a total disaster and it had to be refined because it didn't take into account how efficient these transcribers had become and in the traditional way versus needing to understand, oh, what was this word that the software or the algorithm did not understand correctly, et cetera. And in fact, the app had to become something like 90 to 95% in that range, accurate, for it outweigh performance. So that's why I mean it really matters because it's in what context?

01:09:06 In another context, 80% might be off the charts good and beyond good enough. In others it is not. So understanding the context of the application for the problem, the industry, the vertical, whatever the function, whatever the case might be, it's incredibly important. But execution and then understanding how AI works, again, it's as much an art and as a science as you're building the architectures. And what data do you have access to? So I look for modes. I look for modes around explainability. I look for modes around traceability. I look for modes around AI nativeness, and then execution.

Jon Krohn: 01:09:48 Nice. I think I've noted the key ones down. So some things that don't change and apply across AI investments or otherwise include a strong founding team, which does involve vision, which we hear about a lot. But most important to you on the founding team is execution capabilities, demonstrable execution capabilities, which differ for first-time founders relative to serial founders. And then for AI companies in particular, understanding how AI works and what kind of moat the company has. And you listed several kinds of ways that you could show a moat, for example, around explainability and data. Nice. That was a great answer. Thank you.

01:10:34 And another, maybe this is a relatively quick one from the same person, Mohammad, and this is something that I'm interested in myself, which is how do you stay up to date on the big trends and technological breakthroughs given that AI is so fast moving? I know from your AI Atlas newsletter, which I will of course have a link to in the show notes because it's brilliant. I know that you are unusually on top of what's going on in AI and you can explain it with a clarity and a depth that is indicative clearly of you really understanding what's going on there in the implications. So how do you do that? Do you have a structured process or particular kinds of resources that you follow that allow you to be right at the top of... at these kinds of abilities?

Rudina Seseeri: 01:11:28 I will answer it by saying it's a multi-pronged approach. I don't have a silver bullet. Experience in the space makes a difference. So I'm not learning everything new. Again, my first investment in an AI-native startup was in 2010. So I have the advantage of being an originalist firm in that regard and team in that we've been deeply steeped in the space. So every disruption is incremental if it's in an existing type of technique or net new but we're not starting from zero. We're starting from a solid foundation. One cannot make much out of that. So put that on the side.

01:12:09 Then it's about embedding oneself in that AI cutting-edge breakthrough innovation space and surrounding oneself with individuals and groups that are driving that innovation. So what do I mean by that? We are deeply involved with, especially the academic institutions, think MIT, Harvard, Columbia, Cornell, what have you. Even on the West Coast, Berkeley and Stanford, albeit somewhat to slightly lesser extent, but by being actually affiliated with those institutions and having connections to those super nodes of innovation and having a number of exclusive advisors to Glasswing that work with Glasswing

first and foremost and on an exclusive basis that are the leaders in this innovation. The research paper's almost coming out of academia and industry. So this is what I mean by embed ourselves. So a lot of consumption of information, a lot of surrounding oneself with those folks and then building our own internal capabilities.

01:13:22 What I mean by that is we have internally tools where we look for signaling in papers that are getting published, whether it's out of academic institutions or R&D groups at corporate. We're digesting that information and looking for both signal in terms of what area is gaining traction and keeping ourselves up to date. And then summarizing internally. We divide in congress, so one person is not the owner of all knowledge. We have knowledge sharing and tracking of different areas. And it's not perfect, but we live and breathe the space and automate as much of the information as we can and then we consume it as best we can. And then knowledge share.

Jon Krohn: 01:14:11 Very cool that you have this in-house proprietary tools for staying on top, for looking for signals. I guess I shouldn't be surprised, but it is something that is surprising to me. I wasn't expecting that in your answer, but it's a very cool thing to be doing to be leveraging AI yourself. Yeah, the dogfooding.

Rudina Seseri: 01:14:31 We've been working on it for years actually. And that's the other thing with data sets, how do you balance between the size that's needed, etc. It's produced some really interesting results in our view. Historically this will change, I think. But based on and what we have seen on the signaling, I'll give you an interesting data point, between a paper being published and for the ones that take off becoming mass market leverageable, historically it's been about three to four years, which is fascinating and sometimes more. And then more recently that's shrinking to two to two and a half years. So this is

general. There are exceptions. There are some papers that literally within three months made it from academia or the research world to industry, but those are more exceptions.

01:15:26 So I give you the data point to say that the question that was posed is there's data to actually substantiate that it's coming at a faster pace. Maybe the papers are not coming at a faster pace, slightly faster, but not that dramatic. The adoption from a paper to transform to technologies for the ones that do get transformed into technologies and products, it's shortening the timeframe.

Jon Krohn: 01:15:50 Very cool, very cool insight to share with all of us. One last question here, which you've already actually answered online, and so I know you have a great answer to it, which is Michelle Bonnard, who is the CTO and AI officer at an AI startup. She had this very specific question around ethics. So she said: "In December 2023, Amnesty International reported that VC firms should adhere to the UN Guiding Principles, UNGPs, which require investors and investees to take steps to identify and respond to the potential or actual human rights impacts of generative AI. I'm curious how you Rudina may have adjusted your investing process to include this?"

Rudina Seseri: 01:16:38 I saw that question on social and I was tagged, so it got my attention and decided to address it because it's both a fundamentally important question. And because it's addressable, we can't just raise our hands and say, "Look, there is no explainability. That's mass market, therefore it cannot be done." Now, there isn't a one size fits all, but I'm happy to share what Glasswing has done. So for us, the initial highest level framework beyond just AI, I'll come back to AI, I promise, has been to have been a signatory. In fact, we believe we're the first or one of the very first signatories in North America to the UNPRI,

which is the principles of responsible investing from our early days in 2019. So from day one, I think ethics and responsible AI has been key to us.

- 01:17:39 Then we parse it from our view and say, what can be done on the algorithmic side and what should we consider on the data side? On the algorithmic side, I think we make certain decisions in terms of the companies that we back and what they will and will not leverage. For example, we have hardwired it in our legal documents that vision usage at the facial recognition at the individual level will not be something that Glasswing will invest in. Period. Full stop. It can be used in an extreme case, perhaps it can be used in a completely dictatorial state for persecuting, discriminating, causing all sorts of damage to individuals. So that's a principle that we have hardwired and will not cross.
- 01:18:35 It's also important to us that who is developing these algorithms? Human brains and men and women or different backgrounds; our brains function differently. If the code and the algorithms are developed by a largely homogenic group of people, then the output will favor potentially certain groups versus others. So considerations around diversity have ramifications. Data is incredibly important. But before I go to data, actually one more important point, and it ties to our prior point around how do you stay cutting edge and keep techniques in mind. We're actually seeing some really interesting breakthroughs around the transparency and explainability of algorithms in AI.
- 01:19:22 In, I think it was February, about a year ago, February of 2023 out of MIT, a new framework, if you will, or approach called MILAN. It stands for mutual-information-guided linguistic annotation of neurons, that's applicable to vision in particular, actually made some very interesting headway in understanding how the different

neurons, how the different nodes in a neural net function to give us more visibility on the explainability side of it. So I think it's both being purposeful in the decision that one makes around what one funds, what one doesn't fund as with my VC hat on, and what technologies we help advance and products, therefore, as well as the breakthroughs even before then at the fundamental research level.

01:20:19 On the data side, listen, junk in, junk out. If you have discrimination that the data captures because humans have discriminated against each other, so will the model. So I think auditing data and knowing the type of data set one has is really important. From a proactive point of view, we are part of the World Benchmark Alliance and actively participate with the working groups that work with the large tech incumbents to establish the norms and the principles around the right use of data above and beyond what you see around regulations where Europe historically has led the charge, but regulations in these instances are always a little bit too late. So it's both in who we back in the breakthroughs that we are seeing and the active roles that we are taking and what we're hardwiring in our principles and furthermore in our legal docs.

Jon Krohn: 01:21:19 Great comprehensive answer and it doesn't surprise me that you have that comprehensive an answer given the transparency and the ethical kinds of concerns that Glasswing has had from the very beginning. So thank you for that.

Rudina Seseri: 01:21:34 Thank you.

Jon Krohn: 01:21:36 Well, we have gotten through my questions. We've gotten through the audience questions. The only thing left to do is to ask you the same two final questions I ask all of my guests, which is for first a book recommendation.

- Rudina Seseri: 01:21:49 Oh, God! I'll tell you what I'm reading right now or just about to finish. So I'm going through an Isaac Walterson phase and read Elon Musk. That was a little while ago when it first came out. And then from there started reading all the other books. I'm just about almost done with Benjamin Franklin. That biography, which I find utterly fascinating, just has been enlightening and an interesting thought exercise about an incredible human who was truly a human flawed but also exceptional all at once. And I am about to start the Leonardo da Vinci book, by the same author, the biography, which I did not read when it first came out a while ago. So I'm playing catch up, but I'm on a roll with his biographies.
- Jon Krohn: 01:22:50 Great recommendation. I haven't yet myself read a Walter Isaacson's book, but his biographies have come up many times on the show and he certainly has had some big ones recently like that Elon one that you mentioned. But of course lots of great biographies that he has poured out over the years. I think he's probably the foremost biographer on the planet today.
- Rudina Seseri: 01:23:09 I mean, at this point I've read Jobs, I've read Einstein, I read Musk's book. I'm just about finished with Ben Franklin and he's incredible. I think I would say probably Einstein and Ben Franklin were my favorites, which is interesting because we have two other techies. Given how much I love tech, I should have... Or maybe they're more recent, I don't know, but I found Ben Franklin for sure my number one spot right now.
- Jon Krohn: 01:23:34 Wow, super cool recommendation. Thank you, Rudina. And then final question is how should we follow you after today's episode? Obviously you're prolific on LinkedIn with your AI Atlas, we know that already. Anywhere else that we should be following you?

- Rudina Seseri: 01:23:47 I think it depends on what the interest is. I typically tend to be active on LinkedIn with the AI Atlas and beyond as well as on X, formerly known as Twitter. That's a mouthful. So my handle is Rudina 11, the number, so Rudina11. And otherwise if someone has a company that they want funded and they think it would be a good fit, please reach out and we will respond.
- Jon Krohn: 01:24:15 Awesome. Thank you so much for that offer, Rudina. Hopefully lots of listeners out there with great ideas and some may be inspired by the ideas that you gave today to get going on that AI startup. Awesome, Rudina. Thank you so much. This has been an amazing episode. Thank you for taking the time with us, your hugely valuable to be on the show. We appreciate it so much.
- Rudina Seseri: 01:24:35 Well, thank you for having me, Jon. It's been a lot of fun. Appreciate it.
- Jon Krohn: 01:24:42 How incredibly knowledgeable Rudina is and how well she communicates everything. I was enthralled throughout that conversation and was furiously taking notes. In today's episode, Rudina filled us in on how AI, but not just generative AI will disrupt anything and everything. She provided us with her AI palette for mapping data types to particular ML models. She talked about how the further along a startup is, the easier it is to use data to drive investment decisions because once product market fit is found, data like revenue, customer lifetime value and customer acquisition costs start to become calculable. She also talked about how in addition to the CEO and CTO, AI startups often benefit from a chief data scientist or a chief AI officer. That kind of role is a third co-founder. And she talked about how great AI entrepreneur opportunities lie in the application layer while the infrastructure layer becomes commoditized.



- 01:25:35 As always, you can get all the show notes including the transcript for this episode, the video recording, any materials mentioned on the show, the URLs for Rudina's social media profiles, as well as my own at superdatascience.com/763. And if you'd like to meet in person as opposed to just through social media, I will be in person at the Data Universe conference at the massive Javits Center in New York City on April 10th and 11th. I'll be giving a talk on generative AI and we'll also be walking around interviewing attendees to capture what you think of the massive conference.
- 01:26:08 All right, thanks to my colleagues at Nebula for supporting me while I create content like this Super Data Science episode for you. And thanks of course to Ivana, Mario, Natalie, Serg, Sylvia, Zara, and Kirill on the Super Data Science team for producing another sensational episode for us today. For enabling that super team to create this free podcast for you, we are also very grateful to our sponsors. Please consider supporting this show by checking out our sponsors links, which are in the show notes. And if you yourself are interested in sponsoring an episode, you can get the details on how by making your way to jonkrohn.com/podcast.
- 01:26:40 Otherwise, think about sharing the episode with people that you think might like it, review it on your favorite podcasting platform or on YouTube, subscribe if you haven't already. But most importantly, just keep on tuning in. I'm so grateful to have you listening and I hope I can continue to make episodes you love for years and years to come. Until next time, keep on rocking it out there and I'm looking forward to enjoying another round of the Super Data Science Podcast with you very soon.