

Renovata

AI Roadmaps



Transforming Product
and Tech Strategy



Live Event
Overview & Key
Takeaways

About the event



AI roadmaps have become one of the most urgent challenges in modern product and tech leadership. This event brought together a panel of experienced operators and AI-native founders to explore the strategic choices shaping the future of product development and engineering.

Together with our panel, we unpacked critical questions such as:

- How can we build AI into products without bloating the roadmap or burning out the team? We'll explore lean implementation, balancing AI as infrastructure vs. user experience, and maintaining delivery velocity.
- What capabilities are needed to lead AI transformation from within; should we hire, train, or wait? We'll tackle the build/buy/borrow dilemma and how to keep teams ahead of the curve.
- Where is the commercial upside, and how is short-term success measured? From productivity gains to net-new revenue, how can leaders know if AI initiatives are paying off?

Event Foreword

Julia Barber
Director of Platform
Engagement
Renovata & Company



We are delighted to share our overview of AI Roadmaps: Transforming Tech & Product Strategy with you. The evening offered both insights and new connections, plus the kind of open, honest discussion that can be hard to find with the current level of noise around AI.

Researching and producing this expert panel event reinforced what many of us already feel: that the pace of change in AI is relentless. For product and technology leaders, it presents a powerful duality - it's an engine for innovation, and also a source of mounting complexity. The stakes are high, and navigating what's truly worth building demands clarity and collaboration - and guts!

What also came through loud and clear when putting this event together was that talent remains a critical lever.

The ability to hire the right people - those who can think commercially, move fast, and still build with care and intelligence - is one of the hardest challenges facing companies today. Especially in this market.

If you're facing that challenge, we'd love to partner with you. Our team has worked with some of the most ambitious companies in tech to build out high-performing product and engineering functions as you can see here and we'd be glad to explore how we can support your next chapter.

Please do get in touch directly to learn more about this event, upcoming events, or just to continue the conversation and discuss how Renovata can support you.

Meet the panel

renovataplatform.com



Sara Saab
VP Product
Prolific

Sara Saab is VP of Product at Prolific. She leads their product disciplines, shaping the tools that enable AI developers to access high-quality human data - the critical foundation for building better AI models.

Drawing from her experience across Healthtech innovation and transportation technology on both sides of the Atlantic, Sara brings a unique perspective to the challenges of AI development.



Michael Smith
Founder & CEO
Sagittal AI

Michael Smith is the Founder & CEO of Sagittal AI. With over 25 years in tech, Michael has built and led teams across a wide range of domains, including search, ads, finance, mobile, machine learning, embedded systems, and video streaming.

He's held senior engineering and product roles at some of the world's leading tech companies - Qualcomm, Google, Yahoo, Amazon, and SwiftKey -and served as VP of Product or CPO at multiple startups from Series A to Series C.

Michael has a track record of bringing complex, multi-million dollar platforms to market and scaling teams from the ground up.



Mounir Mouawad
CEO & Co-founder
Portia AI

Mounir Mouawad is the CEO & Co-founder Portia AI . Mounir's career spans prominent roles at major tech companies, including Google, Amazon and Stripe as well as entrepreneurial ventures.

He led Stripe's bank-as-a-service product area in EMEA, launched and grew Google Pay to 30 markets in Europe, served on Google Payments' UK board and advised closely with the UK's Open Banking implementation entity and the European Payment Council.

Mounir also built and grew Amazon's retail deals business which today includes some of their largest events like Black Friday and Prime Day.

How product leaders are separating signal from noise to shape scalable, trustworthy AI experiences.

The noise around AI is deafening, but for product and technology leaders, the message is undeniable: it's time to move from hype to impact, and to focus on shifting from experimentation to execution.

What does a strong AI roadmap look like and how can businesses move from proof-of-concept to meaningful production?

In a recent panel discussion hosted by Renovata & Company, industry leaders came together to explore the practical realities of AI product development. Moderated by Julia Barber, Director of Platform Engagement at Renovata & Company the conversation featured insights from Sara Saab, Michael Smith, and Mounir Mouawad - each of whom brought deep experience building and operationalising AI-powered systems.

The discussion tackled everything from managing non-determinism in AI models to aligning roadmaps with real user value. Here are the key takeaways.



Cutting through the hype



01

From “vibe coding” to viral proof-of-concept demos, the AI space has no shortage of noise. But for those working in enterprise environments, the challenge is to stay grounded.

“LinkedIn is a trash fire of people posting four lines of AI-generated code that look amazing one day and collapse the next,” said Michael Smith. “That’s not maintainable software. The hard part is separating out what actually moves the metrics from what just looks shiny in a single-shot example.”

Sara Saab also cautioned against what she called “shiny orbit syndrome”

Meaning, the tendency to bolt AI features onto products without a clear user problem or performance threshold. Instead, she stressed the enduring relevance of product fundamentals:

“You don’t lob things at customers because they’re cool. AI or not, the principles don’t change. Start with a well-defined user need, scope narrowly, and put proper governance and guardrails around it.”

Why transparency matters



02

Trust and oversight emerged as dominant themes throughout the conversation. Mounir Mouawad pointed out the risks of applying generative AI to processes better served by deterministic tools like RPA:

“We’re seeing people force LLMs onto workflows that don’t need them. If your task is human-intensive and language-heavy, like triaging customer tickets or gathering supply chain data, then AI can be very helpful. But without that fit, it becomes a liability.”

Sara echoed this sentiment with sharing a practical example: an autocomplete feature in a widely-used enterprise tool that broke the user experience due to an overly complex AI inference layer.

“There was this model running in the background trying to infer similarity, and it completely broke my user experience. Why? Because it likely wasn’t grounded in a real user need, or properly scoped and thought through from the start,” she said. “AI added friction instead of value.”

For AI to scale responsibly, teams must test, observe, and improve the abilities of their customer incrementally. “It’s about evolving the ways that we interface with the world and work. What is the thing I can do to delight my customer that makes them feel like a superhero?” added Sara.

AI roadmaps require ruthless focus



03

When asked what defines a successful AI roadmap, the panel was aligned: it starts with tightly scoped, low-risk, high-impact use cases.

"The teams we've seen succeed began with use cases that tick three boxes," Mounir explained. "They're reasoning-heavy, language-intensive, and involve lots of human labour. That's where AI delivers real efficiency gains."

From that starting point, businesses can begin to scale but only when appropriate controls and interfaces are in place. "You start getting internal FOMO," Mounir said. "Suddenly everyone wants to apply it everywhere. That's when you need the right SDKs, human-agent interfaces, and guardrails to avoid overreach."

Rethinking workflows, not just features



04

AI isn't just a plug-in - it reshapes the work itself. To use it effectively, you must reimagine your product and processes from the ground up.

"The hardest part of reinventing a workflow," said Michael, "is trying to figure out how to pass context-rich information back and forth between a human and AI. That's not something you can tack on, it often it means rewriting the workflow entirely."

Rather than asking what AI can do within an existing system, the panel encouraged leaders to think at service design level. What would it look like to embed AI throughout the entire workflow? How would it reshape the product experience?

From crystal ball to practical benchmarks



05

We continued the conversation with a forward-looking question: How do you measure success when working with non-deterministic systems?

Standard KPIs often fall short, especially when dealing with probabilistic outputs. Instead, panellists advocated for custom metrics tied to user impact, reliability thresholds, and system behaviour under real-world conditions. While AI maturity is moving fast, the fundamentals remain clear: roadmap success hinges on clear value, measurable outcomes, and disciplined execution.

"You might come into work and realise everything you thought was true about your AI roadmap no longer holds," said Sara. "That's the pace we're moving at. The only constant is staying grounded in product-first principles."

From “POC purgatory” to real value



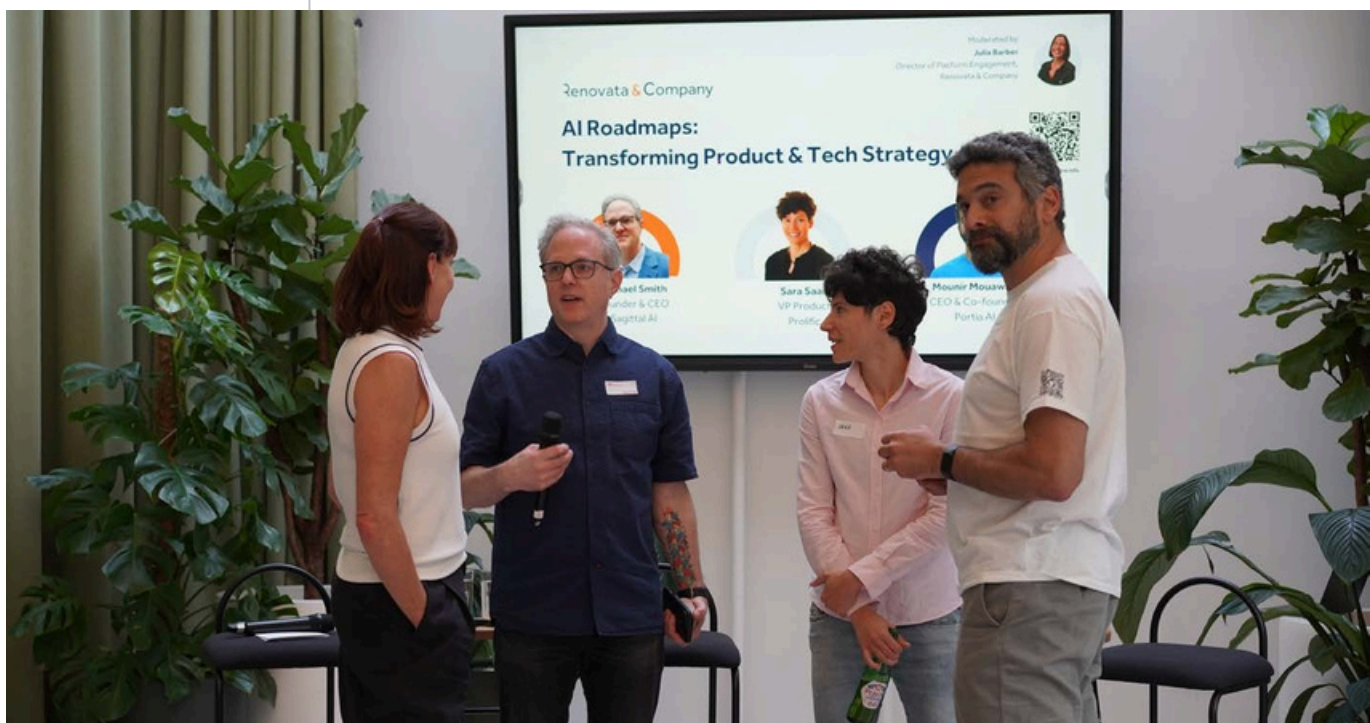
06

We then focused our attention on how organisations avoid falling into “POC purgatory” – a state where promising proofs of concept never reach production?

For Sara Saab, the answer lies in focus and first principles. “We see a lot of teams spun off into innovation pods and told to ‘implement AI,’” she said. “But without a direct connection to customer pain points, these initiatives rarely translate into lasting impact.”

The consensus was obvious: AI should not be bolted on as a novelty. Instead, successful projects start by identifying real-world needs and designing solutions that directly address them. Mounir Mouawad added:

“The use cases that succeed are reasoning-heavy, language-intensive, and human-resource dependent. That’s where AI can deliver exponential value but only if it’s well-scoped and user-aligned.”



Roadmaps under pressure: keeping pace with a moving standard



07

A vivid example of roadmap disruption came with the rise of the Model Constitution Protocol (MCP), an open standard initially spearheaded by Anthropic. “Overnight, it became the new standard,” said Mounir. “In three months and 30,000 memes, writing your own tools for common apps like GitHub became the dumbest thing you could do. It didn’t make sense anymore.”

Michael Smith, reflecting on the deeper implications for product design, added: “If your value lies in how the tool is used or encoded in your product, MCP can help with the connection layer, but it can’t

replace your business logic. You don’t want to outsource that logic to the latent space of a reasoning model, that’s just not good enough.”

The landscape had shifted dramatically, forcing product teams to fundamentally rethink their strategies. “We had to seriously rethink how we adapt our roadmap in response,” Mounir shared. “But it also reminded us that even amid hype cycles, the fundamentals don’t change: know what you’re building, who it’s for, and why it matters.”

Control, context, and the limits of abstraction



08

Trust in AI isn't just a technical hurdle - it's a design responsibility. From unexpected behaviours to performance drift, today's models present new risks that demand equally new forms of control.

"LLMs don't fail like traditional software," Sara noted. "They hallucinate, they drift. You need to move safety and observability earlier in the process, shift them left."

Michael used a striking metaphor: "Think of your customer like a human in a giant mecha suit - the AI is doing the heavy lifting, but the human is in control. That's the product experience you want."

Maintaining that control requires decomposing tasks into measurable, testable components. "Don't outsource top-level logic to a black box," he added. "Build deterministic layers where it matters."

Killing old metrics in an AI-first world



09

In a world of probabilistic outputs, traditional KPIs no longer suffice. The panel discussed the importance of outcome-based evaluation over stylistic judgment.

"In KYC workflows, the question isn't 'did the output sound right?' it's 'did the system make the right call?'" Mounir explained. "We're moving toward deterministic metrics that map to actual business outcomes."

Sara raised another long-term consideration: "things like observability, explainability, transparency, safety, and oversight become interesting spaces for innovation. I wonder, do we think these capabilities will be commoditised and solved by vendors, or will they need to be built in-house? It's a fascinating space to watch."



Preserving human judgment in an AI-native world



10

As AI tools mature, so too must our human skills. The panellists warned against over-reliance on systems that may obscure errors or erode user intuition.

"LLMs don't know when they're making things up," Michael cautioned. "Humans do. We're best in the complex space - navigating ambiguity, drawing on lived experience, and interpreting context."

The emphasis, he argued, needs to shift from production to evaluation.

"Engineers today spend 10x more time writing code than reading it. That needs to flip. We must learn to read, test, and contextualise AI outputs."

At Prolific, Sara shared a similar principle: "We care deeply about preserving human authenticity. AI should be a tool, not a replacement. We need to protect our knowledge - not discard it. I firmly believe humans will never be fully removed, but their involvement is evolving."

The future of teams, tooling, and trust



11

As AI changes how teams build and operate, it's also challenging traditional organisational structures. "Trust within teams is inseparable from how we manage risk," the panel agreed.

Michael highlighted the importance of layered design: "We structure workflows with clear planning and execution layers. Each layer has its own checks, was the tool call correct? Was the summary aligned with the task?"

Sara added: "In five years, some of this may be abstracted away. But for now, we're building the playbook in real time."

From analog to digital, and now to AI, each transformation has raised the level of abstraction at which we work. As Mounir put it:

"We've gone from optical cartridges to compiled languages to natural language. But no matter how high the abstraction, foundational skills - systems thinking, critical reasoning - remain essential."

Centering the user not the roadmap



12

The conversation underscored a simple but powerful truth: start with the user. Define the problem clearly and prioritise transparency. And never stop asking what value looks like, and how to measure it.

Building with AI is less about chasing breakthroughs and more about staying grounded where you are focused on real problems, real users, and real value.

The most effective teams aren't trying to predict the future; they're building in motion.

They're adapting to shifting standards, managing the risks of non-determinism, and rethinking how humans and machines work together. They know that success doesn't come from adding AI for the sake of it, testing with intention, and keeping the user at the centre.

There's no fixed map for what's ahead. But the companies that will lead are the ones treating the roadmap itself as a living process — practical, responsive, and built around trust.

Audience Q&A

Transforming Product
and Tech Strategy

Great panel – but the discussion felt somewhat unsatisfying, because the current complexity means we are totally in the weeds. Do you think we really have any idea what we are talking about? Do you think we will be anywhere near where we are now in 2-3 years or is the future totally unknown?



Sara: That's an incredible question to ask. Honestly – it's just so early and even very qualified people aren't ready to state clear opinions. However, you know those thought experiments where an ant can sort of tell that it's on the surface of a balloon it just doesn't realise the exact dimensions of that balloon? That's where we are.

By looking at the negative space we can somewhat understand where we are on the innovation curve and what kind of moment we are in but I don't think we currently have any idea what the future will bring.



Michael: That said, I think there are some things that aren't, won't and shouldn't change: you still need to worry about the metrics for your user, in a global and not just a localised fashion.

It's another technology that you apply to your customers' problem and your design, your opinion, your taste and your understanding of the market that you're in – that will not be outsourced to an LLM. But the capacity to deliver against it is something that you can probably hold less tight.

There are things that we know to be less and more true now and we need to get together like this and tease out which elements need to stay constant and which ones are still malleable.



Q&A

Q

There have been some great examples of how to narrow focus so you're not trying to boil the ocean – to what extent is time the really underrated aspect of this? We are obsessed with instant gratification and I think we are missing the forest for the trees. Why not get the LLM to work harder for us overnight?



Sara: So this is one of the magic moments. Anyone who has used the Deep Research model of ChatGPT? That's a goosebumps moment. It's the mixture of speed and explainability, the immediacy of response plus the texture of something that feels really human.

As an example, I haven't coded in python for 20 years but I had AI assist me in writing some code last weekend and realised I could still do it! We underestimate the immediacy.



Michael: I agree I think Deep Research is amazing but sometimes you just don't need that kind of depth. I am still learning whether I need detail or immediacy – is that kind of depth always needed? That's actually a great product question to ask and that's where our judgement call as humans with domain expertise comes in.

What comes first: reinventing workflows, or getting customer perspective on how those workflows operate? Surely if you reinvent the workflow you haven't built trust. How do you balance trust and transparency?



Mounir: There's a tension between autonomy and things going off the rails. I don't think there is a universal answer but what we see with users is the higher the risk appetite the more they are willing to delegate to the AI and have guardrails from an evaluation perspective. If you take an example like: creating a refund, I would say you need a deterministic execution hook where the customer can say I don't care how confident the AI sounds, I want to talk to a human.



Michael: I'll double down on the tension in this question and say that this is going to drive us to use AI better, and then the whole thing improves. The way we interface with AI today and the amount we trust or not will shift over time.

Q&A

Q

We have a motto at work which is 'bet on the model' - what that means is during the design phase we bet ahead to where we expect the intelligence to be in the future. We've found that we have consistently not bet far enough ahead. How would you advise we build that into your product strategy and roadmap?



Michael: I go back to that framework – what are humans good and what is AI good at. AI increasingly leans towards the complicated and complex, but isn't so good at judgement calls. So unless you are tuning it in that way you will have to understand when to surface the human connection. So you should always bet that the complexity will increase.

Since 2022 my cofounder and I have observed the models evolving and expanding, growing in capability but not a total leap beyond that in terms of reasoning etc. So perhaps that distinction will help you with your bets.

Mounir: This might be speculation but when you speak to OpenAI or Anthropic they are looking for differentiation, creating LLMs with vertical specialisms in either research or reasoning. So the one thing I would advise is flexibility, and the willingness to flip your roadmap. Don't bet the entire business model on something that will undoubtedly change.

Q

How are you considering versioning the AI eg. in traditional computer science you check in your changes, you run the all correction tests, and you see the delta. How are you thinking about changes to the AI and not just checking the delta between A and B for drift over time?



Mounir: Your question touches on the stochasticity of the AI. At Portia we have a planning agent that helps plot out your workflow, with the aim of getting to 99% reliability. Then a new model will come out and suddenly your prompt is not effective. Its unpredictable and its especially unpredictable across vendors.

The answer is that we have daily evals that run across all models for every prompt to check that is never falls below a certain threshold. We have it on a screen and it's all hands on deck if anything happens. Whenever a new model comes out we're thorough making sure that with any new feature, eg execution hooks, all the evals are triggered.

Q&A

Q

I am CTO at an AI native company where we have great product and technical conversations every day. The one thing I fear the most with AI? Linkedin. The amount of my energy I spend either explaining or moving away from a problem is a bigger worry than building the tech itself. The amount of misinformation or overhype from people not qualified in the subject is crazy. Do you have any thoughts?



Sara: We've hit the hype bubble in tech at the same time as the peak of mindless human ephemera. There is some stuff that's interesting eg peoples' early impressions of the models and felt reactions to the usability of new developments. But largely, I think we all agree that LinkedIn is a minefield!

Q

We've spoken about trust and the notion of common knowledge vs correct knowledge. I am not talking about revisionism, but one thing we are seeing with AI agents is patterns – we accept their suggestions, and then false behaviours are reinforced. What guardrails can you suggest against this?



Michael: The way we handle that with our product, is we encourage users to figure out how they want the system to behave and train the system in those standards. When you are looking at coding agents for example, you need to interrogate the quality of your code base – but maybe think about setting higher standards?



Sara: That's such an exciting question – one of the things we are trying to do is make sure AI is trained on representative sets of knowledge from the general human population. The problem is the more representative you are, the more bad stuff you are including from humanity.

Are we looking to train models on all human knowledge warts and all? It's a question we need to keep coming back to.

Thanks for a fantastic panel. You've been talking about building on frontier LLMs and foundational providers and many in the room are building on the same sets of assumptions. There is a lot of price discovery going on so how do you see that playing out in terms of business models?



Michael: We have found is that if we're doing much more curated workflow, with more deterministic programming so the more narrow your use of the LLM, the more inexpensive it gets.

If you take a zoomed out view of user problem that has the largest impact in terms of end to end metrics you can spend a handful of dollars instead of worrying about the tiny part that might not be correct and spending a fortune. When you have more time – you can spend more money. That said the inference costs are plummeting and we've started seeing people apply hardware that's specific for inferences eg. Groq.



Mounir: I'm not a hardware specialist but there have been great podcasts on this recently: between the tactical usage of GPUs and the changing architecture of data centers, I am hopeful that inferences will continue to go down over time - otherwise the companies running models themselves will bleed out. This is not just about you, me and our startups. In the early days cost per query was so much higher and hopefully it will level out.





Final Takeaways

Q&A - Final takeaways



Sara Saab
VP Product
Prolific

In the spirit of not having all the answers...what should the human / machine interface look like going forward? What does that evolution look like? There's a lot of surface area in that question for the audience.



Mounir Mouawad
CEO & Co-founder
Portia AI

My question would be - what would your product, your business and the world look like if AI had 100% reliability?



Michael Smith
Founder & CEO
Sagittal AI

And I will piggyback off that and say what can you do with 80% reliability and being able to reinvent the workflow around humans – so what does the best configuration look like and what can we achieve with that?

→ Get in touch

The Renovata & Company Events Programme exists to create value and increase engagement within our network, deepening relationships with our clients, candidates and partners.

To learn more, or to register your interest visit our website. If you would like to be involved in our events schedule going forward - as a guest, speaker or partner - please get in touch with Julia Barber directly.



Julia Barber
Director of Platform Engagement
julia.barber@renovataplatform.com