Episode 59: Breaking into AI Product Management - Continued

Speaker 3 (00:32:45):

Let's dig a little into the decision-making part, which is a pretty crucial part across many different areas in developing AI products. Now, if you've been working on kind of products that not use AI, how do you actually think about coming up with the use cases that require AI? Because a lot of times we think AI is a solution, but AI is the technology means to the solution means to the problem, and still the problem is very important. So how does one kind of think about what use cases should I think about that could leverage AI and make this even better?

Speaker 1 (00:33:34):

Yeah, I think you're completely right. I see a lot of folks saying, oh, I've been told we need to develop something. The product is AI, the product is never AI. I think there is an opportunity though to say what are some problems right now that we could maybe solve better if we have a more evolved understanding of what AI capabilities are today and where they're heading. So the key I've found for picking the best use cases in your organization is do start out with some education around everyone has the same level of understanding around particularly generative AI, if that's where you're focused, what the capabilities are and then what those limitations are. So hey, let's not go after use cases where a hundred percent accuracy and absolute applicability are a key unless we are a hundred percent aligned on, we'll take on whatever costs it takes to get to that a hundred percent and those costs may be high, right?

Speaker 1 (00:34:43):

I've seen that sort of deciding on use cases, done a lot of different ways you can start top down to start bottom up and have a combined approach where you get ideas from both sets of stakeholders. And I think dividing those ideas into three buckets has always been helpful for me, where it's looking at low hanging fruit where you're saying a lot of the time these are either replicating current human conversational experiences. So we have a call center and we want to look at replacing it with the chatbot. That's a very common one that came out earlier this year for a lot of companies we think use their understanding what is that metric we are assuming this would drive up, is it going to be faster? Are customers going to call more customers going to call less saving costs or do they love us more?

Speaker 1 (00:35:39):

So really being clear on why you would do this is key for every use case. So there's low hanging fruit like chatbots. Another one is just any kind of internal process automation where any kind of moving document content around, we're doing more and more with moving even data around and translating it from different formats. He may have seen opening AI announced an increase in the capabilities around ensuring you get JSON output from your models, et cetera, right? That's just because these are used more and more not just for English writing, but for code writing and internally any processes. So those are the low hanging fruit ones. I think the interesting ones to look at from there are anything around growth. If you have anything to do with a use case where you can any more content drives more engagement, A really great example is Spotify's move to translate all of their podcasts into a hundred languages.

Speaker 1 (00:36:37):

So they're automatically translating this. Each podcaster now has a hundred x the audience, each audience member now has a hundred x. The selection is driving engagement, which will drive longer listening times, which will drive ads, which will drive revenue. And you can see how looking just a little differently on how can we go not just that incremental step, but something much larger in that case has really helped them drive themselves forward. And then the last one is like, is there a moonshot here where we could transform the industry, transform how everyone thinks about how this job is done at all? Right? My favorite so far has been looking at Khan Academy's approach to education where instead of teachers teaching students, we can scale them by having LLMs kind of be the teacher and teachers overseeing and correcting or discussing anything interesting that comes up. But in this way, students instead of each having a highly skilled teacher, having a personalized assistant bought there to tutor them, personalized tutor can really scale education in ways that we could have never afforded to before and changes the whole paradigm for how teaching is done.

Speaker 1 (00:37:51):

So I think it helps to encourage teens to think of things held separate brainstorming sessions for each of those buckets and figure out which of these are we willing to invest in. One of the biggest challenges, of course, people would traditionally say, oh, well then you determine the ROI for these projects and pick the hardest ROI. AI projects can be very difficult to determine the ROI when you don't necessarily know how long, how much effort's going to take. So I do think for particularly growth and moonshot capabilities, you may need to get buy-in from executive leadership on let's invest in this as a capability. And even if we can't achieve this initiative in six months or a year or 18 months, we hope we'll be so much closer and have built so much more machinery around experimentation with AI around understandings limitations. We'll be so much further ahead that it will have been worth it over the longer term. Some organizations are in on that from the beginning right now. Some might require you to start with something smaller from the low hanging fruit bucket and let's start experimentation there. So it really is about your business, your business's appetite for risk for innovation and investment in something as a capability.

Speaker 3 (00:39:17):

That was a really great way you chunked up across the three buckets and loved the way you thought about those. And actually you made a great point about the metrics now in while building products, be it zero to one or to scale, you have your regular metrics B, it B two B or B two C with respect to either growth or customer acquisition or revenues and whatnot from that perspective. But are there any additional metrics now that you would have to think about when you're using AI to develop some of these products and capabilities?

Speaker 1 (00:39:55):

I don't think so. Not at that top level. And that's the key piece is people keep trying to, let's put accuracy in there or things like that. Scientists are talking about F one score. It's like those are not business or product metrics. And I think that's really where as PMs, we need to be rock solid on keeping that north star around. The thing that hasn't changed is if we're not delivering business or user value, what are we doing? So I think a lot of the metrics that are getting introduced, a lot of the time they belong at a lower level. These are technical metrics, latency accuracy, hallucination rates, things like this. We need to develop a model where you can translate those into who's hallucination rates go down and we assume users will have better goal completion rates or users will come back more often or how is that going to affect the user? How is that going to affect our bottom line?

Speaker 3 (00:40:57):

That's a great call out to say the product and the business metrics are still so key and the rest are still the second level or the technical kind of metrics, but you can't forego the product and the business metrics that are in there. And one of the things you talked about was, hey, is this good enough to launch? Is this what is good enough? And how are you kind of making those decisions? If you have any examples to share, that would be great to listen to

Speaker 1 (00:41:26):

For sure. Yeah. So for example, in the use case with the news, that was a great, because it was something that we had to get a lot of different stakeholders aligned on the internal, just the process for deciding this. We did things like look at human benchmarks. So we looked at the technical benchmarks on what's the best that's been done out there and we believe we can do better than that. Now. We had a really world-class data science team, and the whole point was like, yes, we can do better than the current state-of-the-art as far as the papers that have been published. Let's do that. That's the problem, the challenge we want to take on. So then we look at human benchmarks. So for example, we had found research that showed generally something like 2020 5% of news articles written by humans contain factual errors.

Speaker 1 (00:42:16):

So the same as we see with benchmarks around detecting breast cancer and images, et cetera, that combo of humans are fallible too. And so if you could do better than a human system, that says a lot. Now we went through a complex categorization. Of course, just getting someone's name wrong or sports score wrong is very different than announcing At the time we were working through covid, we didn't want to announce that a political leader had covid who didn't or it's something like that, right? Right. I remember there was one in testing, we had some articles coming out that were stating that there was an Ebola outbreak in Minnesota. We're like, that's the last thing we need right now. Right. Ebola too as well.

Speaker 1 ([00:43:07](00:43:07)):

So we definitely had different categories of errors and we had done a lot of modeling where we're simulating how many of these would happen in a day, a week, a month. And it was really all about discussions and getting on the same page of leadership of how many of these are acceptable if given that the benefits versus the costs, the costs of this. And definitely having all the mitigation processes in place as well, that to, we needed that to earn a lot of trust. So we did things like we have processes for if the minute we are actually monitoring for any incorrect news, we have humans in the loop doing vetting and checking of a certain percentage of our articles. We are monitoring Reddit posts about Alexa, we are monitoring social media and all the folks who are monitoring know the process to take down an article that's incorrect so we can get that taken down within hours as soon as it's reported. So all of those kind of risk mitigation practices help make people more comfortable with the degree of risk we're taking as well as things like reactive or proactive PR and knowing exactly what would happen if something went.

Speaker 3 ([00:44:32](00:44:32)):

Yeah, that's an awesome example. And taking this a step further into really the product development lifecycle itself, what do you think is different or what new layers have been added when you're, you're developing products using AI as a technology? I know one of the things you mentioned was about the ML side and the engineering ops side of the house. Anything you want to elaborate with respect to that or add more from just the AI product development lifecycle perspective?

Speaker 1 ([00:45:05](00:45:05)):

In general, the AI product development lifecycle, it's kind of like agile on steroids. It's not just a simple, there's the traditional agile figure eight and you're like, okay, let's develop a feature and put it out and get feedback and put out the next one. With more teams, there's just more complexity along with managing that uncertainty. So from the beginning, you're investigating the data, you're assessing the risks of the data. You might not know right away whether the data has high enough quality or is enough data to accomplish your metric goals. And so there's definitely an iteration of going back to, okay, do we need more data? Could we identify other sources of data that the data science team is iterating on? At the same time, the engineering team is trying to build the data pipelines you'll need in production and set up the engineering systems to support whatever latency you need for your models and production, et cetera. And they need to know early what it is that you're launching, which you often don't know right at the beginning. It requires some experimentation, it might require several different approaches, very different models that have very different engineering requirements. So there's a lot of uncertainty in the beginning to manage and understanding what do we know, what can we start building right now so that we don't end up with a beautiful model ready to deploy, and then the engineering team saying, okay, now in six months we can deploy it.

Speaker 1 (00:46:36):

So I think that's where it's more stakeholders and it depends a lot on how your company structure is set up, whether you have a team and you have engineering and data science working together and attending the same scrums, or whether you have an infrastructure team or an engineering team that's separate from the data science teams. The flexibility to work in all of those different systems is something that I thinks can really lean on and show like, Hey, throw any org structure on me and I could be the person who ties together these groups, whether they are working arm in arm right next to each other at the desk across from each other, or they're on a different floor in a different building. I can really make sure that be the glue that pulls it together.

Speaker 3 (00:47:27):

Got it, got it. No, that's very insightful. From that perspective, I know we touched upon data scientists in a lot of these places. What do you think or what constitutes a successful collaboration between the product managers, the product leaders, and the data scientists?

Speaker 1 (00:47:49):

Yeah, I think a lot of PMs, the first time they work with data scientists, I remember I made this mistake too, tend to treat them a lot like engineers. And it's very different working with data science teams, whereas engineers always wanted more detail on the tickets and exactly when am I done. Often I would find if you try to manage data scientists like that, they kind of bristle and they're not wrong because a lot of what they're doing, they're like, well, I have to go and sit and figure out what I'm doing and try say things before I try the fourth one. And you're asking me to spell that out beforehand and tell you how much time that'll take. Often the tickets we would write data scientists would be experiment with approach A or A, B and C, and their output would be the results.

Speaker 1 (00:48:42):

So the success of something like that is making sure that you have the trust, that you get the visibility into the data science team's progress and results without having to demand when's the deployable thing going to be done. So their output is no longer features, its results that are like, Hey, is approach A or approach B getting me closer or on a higher velocity to get me closer to the metrics we've decided we need to iterate to. We were having monthly data science reviews where the data scientists would share broadly with the whole team, like, Hey, here's the different approaches. Here's what we can include from them. Here's what we can't conclude. And having that collaborative decision making about the next steps, they would definitely come with recommended next steps. We want to try this completely new approach that just emerged. It was also a chance for the whole team to get a little more educated and if they came in and said, oh, I heard from a TikTok that there's a new approach to X, Y, Z, it was a chance for them to say, that's not really what this is used for, or Yes, we are considering that we want to throw that in the mix for next month's results.

Speaker 1 (00:50:02):

So that kind of collaboration is always a big green flag for me. If you see them explaining as they go and willing to keep other stakeholders, stakeholders in the loop. A lot of the time folks involved do want to know more about the data science side and get more familiar with it. So finding a team that where you can build that collaborative relationship is key.

Speaker 3 (00:50:25):

That's a great insight. I think one of the key things for them also to help them is working with them for them to understand what's the use case, what's the problem that we are really trying to solve and what they're doing kind of fits with that and why are you asking for certain approaches? Exactly sense, right,

Speaker 1 (00:50:46):

Exactly. And that's absolutely key is that the learning is both ways that you don't have the egos involved where they're just like, oh, we're here to explain how the world works to you. And it's like, no, as the product manager, especially if you have a system line out in production and you're looking at the business metrics, you're talking to users, you're getting really qualitative feedback that helps contextualize that. They should be just as curious about what you have to say as you are about what they have to say. So really strong teams, it's two-way learning all the time.

Speaker 3 (00:51:16):

That's a really great call up there, Polly. And maybe one last question. What opportunities do you see or foresee, especially for the product management function in, I wouldn't say in the future, future because I think the

Speaker 1 (00:51:31):

Future is here.

Speaker 3 (00:51:34):

So how do you see them evolving? What's kind of the next in your perspective from a product management function standpoint?

Speaker 1 (00:51:43):

I am so excited about product management and its evolution right now with the sort of rapid democratization and access to large language models that's happening. And I really see it as the time for product managers to rise up and seize the reins in a way that I think a lot wanted to. I spoke to so many hundreds, not thousands of PMs across lots of different industries. And it's so common that people are really frustrated that they know exactly what they want to do in terms of, they start with user research and understanding user value that they've been forced into kind of top down situations where they're either it's a tech led initiative or a sales led initiative, and they're really just the order takers who get things done. A glorified project manager. I'm so excited to see that in this case, there's an opportunity to really, the place where companies can deliver value finally is going to be recognized as being customer first focusing on the value, the technological innovation isn't going to be the biggest differentiator anymore, and that's getting commoditized so quickly.

Speaker 1 (00:52:57):

So now human factors and things that PMs can deliver, whether it's relationships with other companies, data partnerships, right? Ideas on how the business can build an advantage outside of just that technical side that build the both together. I think that's one of the most exciting things that comes out of this. The power is now in the hands of PMs to develop prototypes, to get them out there to show them to people. And so I think it really opens the door to a lot broader set of product managers with more diverse backgrounds than just traditionally computer science, data science backgrounds to be a part of this PMs who think of human factors and responsible AI factors first and not as an afterthought. So I'm just really, really excited to see that become more common in industry and really leaning into the human side of AI products.

Speaker 3 (00:53:48):

That's awesome. And on that note, this has been a really, really insightful conversation. Anything else you want to add Polly to this?

Speaker 1 (00:53:59):

No, I am just super excited to work with more folks. My company, AI Career Boost. We've done a bunch of courses that we've grown over the past year that have really helped accelerate people's careers in AI, product management and AI leadership. It has been an amazing ride and I've met some of the best folks, most inspiring folks in my career. So I hope to continue to do that in the future and hope I'll get to meet some of you along the way.

Speaker 3 (00:54:26):

Yeah, amazing. Awesome. Thank you so much for your time.

Speaker 1 (00:54:29):
Thank you. Thanks so much for having me.