Andrew Obwocha: I saw that your Master's research focused on explainable AI, and recently I saw you also did a certification from Anthropic in AI fluency. This suggests you're big on interpreting AI systems and validating how trustworthy they are. While the headlines right now are focused on extreme AI topics, is there a specific trend in AI safety, in particular, that you feel most people should be paying attention to—maybe even specifically undergraduate students?

Justin Stevens: AI safety is definitely a very important area to be working in these days. A lot of people have pivoted their attention, even their entire labs, such as Yoshua Bengio and Geoffrey Hinton, towards AI safety. The angles that I care about a lot are really how people are being educated about AI. This is especially important when we look at the era of deepfakes, where it is so easy to produce realistic-looking videos and images.

I did an exercise the other day where we were just looking at images and trying to detect if they were real or not. As a group, we got about 60% correct, so only slightly above chance. It's really tough to tell sometimes, even as someone who's been in the AI field for a while. So, I think being educated about AI is going to be one of the most important things we can have going forward, because everyone needs to know that this technology exists and how not to be tricked by it.

There have been disturbing cases where people have gotten supposed phone calls from a family member that was in need, and the voice of the family member was totally fabricated. It was not a family member in need; it was a scam saying, "Hey, I need $10,000 right now." Those are really unfortunate cases. So, I am not so big on the existential risk of AI yet. I'm thinking a lot more about how to make sure that people know what this technology is doing and how we can ensure people do not get manipulated by AI systems.

Andrew Obwocha: One angle on that, in particular, is that the people who end up as victims have very little recourse. I don't know how many legal rules are in place for them to take substantial action against somebody using their likeness. It often ends up in a case where your likeness is used in a deepfake, somebody leverages that against someone else, and both of you end up in a very bad situation. Do you see any laws being passed by governments in the future to try and hinder or limit that in some capacity?

Justin Stevens: Yeah, definitely. The European Union probably has the most restrictive AI laws right now with the AI Act they've enacted. I know the Canadian and U.S. governments have thought about that too. I think a lot of folks are thinking about how to avoid being manipulated by AI systems. I'm definitely not a legal expert by any means, so I can't comment on exactly how the legislation will look, but I think it's good that governments are beginning to take this seriously.

Andrew Obwocha: That's one example of people using AI in a malicious manner, but there's also its use in industry, which leads perfectly into the next question. You're in a unique position because you've conducted research, but you also work with Art of Problem Solving, so you have an industry perspective as well. From this vantage point, are there any tools in AI that are being pushed in research as the next revolutionary thing, but their practicality is lacking? Or conversely, something that is being widely adopted in practice, but the research community doesn't view it as particularly unique?

Justin Stevens: That's a great question. So just to make sure I'm understanding right, the question is: are there tools on the research side that industry isn't adopting enough, or tools on the industry side that researchers aren't adopting enough?

Andrew Obwocha: Yes, exactly that.

Justin Stevens: Cool. I think there are definitely tools on the industry side that people in academia are not as privy to. I know from talking with friends and colleagues at some of the big AI companies that their opinion is that a lot of research is being done behind closed doors. You can see this with OpenAI and how they're releasing the latest GPT models. They were built on the idea that their work would be open, however, they don't release any details besides benchmarks. They do that because they don't want their competitors to know how they're training their models.

I think there are also trends on the research side that aren't being adopted in industry yet. For instance, I've been pivoting towards a lot of work in user studies—how people actually interact with and learn from AI systems. I know some companies like Art of Problem Solving and Anthropic value that type of research, but many other companies will just release AI products without thinking about how they're actually impacting their users. That is a concern for me. ChatGPT was released almost three years ago now, and it has significantly impacted the world. I don't think any research was done when it was released, aside from the goal of making a bunch of money, to ensure that it would be used safely and fairly.

Andrew Obwocha: Just on that, I don't know if you read the article, but about two weeks ago, a memo was leaked from Meta about the specific rules they have for their AI modeling. It had some pretty grotesque stuff in there. It was very shocking to see. If we keep allowing these AI companies to black-box their research and release these models, especially considering they're trained on the entire internet, there's going to be bad or manipulative data in there. How can we combat this if there's no legal penalty for them keeping their competitive edge secret, while behind closed doors they're engaging in unethical practices?

Justin Stevens: I think it ultimately has to do with public perception of these AI companies and whether or not people are willing to boycott the products of those they view as unethical. You mentioned the thing from Meta; I won't go into the details either, but we can link to what they were permitting their AI models to do. Personally, I uninstalled Facebook from my phone after that. I still use Instagram, unfortunately, but it was a moment for me where I was like, "Okay, I need to use less of this product."

I don't have a full answer for how we can stop these companies, but with Meta in particular, it has been shown time and time again that they will value profits over fairness and ethics. There were reports a while ago that they knew they were actively promoting harmful content because it would increase engagement. There's a principle, I think it's Goodhart's law, which is: "When a measure becomes a target, it ceases to be a good measure." That's really important to keep in mind here, because if the goal is just profits, companies will continue to do unethical things. But if the goal is actually releasing AI in safe and fair ways, they will gain more trust with their users, their products will be better liked, and ultimately it will be better for them in the long run.

Andrew Obwocha: On that note, Meta has been gaining a reputation for offering huge contracts to catch up in the AI race. For someone looking to get into AI, seeing this kind of money floating around—let's say a first, second, or third-year student—should they focus on being a mathematician first or a computer scientist first?

Justin Stevens: Oh, that's an amazing question. I'll share my personal experience, but of course, everyone's mileage varies. I was a mathematician first, and I still identify as a mathematician. I did a double major in mathematics and computing science in my undergrad at the U of A, and I think that mathematics underpins everything else going on in the field right now. So I would really recommend that anyone who wants to get into AI take at least a few math courses: linear algebra, calculus, multivariable calculus, probability and statistics, and some differential equations. All of these are important for understanding the latest AI models.

That said, I think the barrier to entry for AI is a lot lower today than it was a couple of years ago. In all honesty, when I work with AI models, I might do a little bit of math, but I'm mainly using pre-built things with PyTorch, TensorFlow, or JAX. So now is one of the best times to learn AI because it is so easy, especially in the era of AI coding assistants. A math and computer science double major is one of the best ways to go, but everyone is coming at it from a different perspective. A biologist might do a computer science and biology combination, for example.

Andrew Obwocha: That's true. You mentioned you were a double major, went into professional research, and then worked in industry. These are different stages a typical AI student might go through. There are the foundational mathematics, the hands-on engineering projects using tools like TensorFlow and PyTorch, and the ability to communicate these complex ideas. These three skills feed into each other but are distinct. What would you say is the bedrock between the three: the foundations, the projects, or the communication?

Justin Stevens: My personal opinion is that technical skills can be learned on the job. I feel like communication skills and being a good person to work with are some of the most important things. Obviously, if you're coming in with no knowledge of the technical skills needed, it's going to be a tough time. But I would really recommend to anyone that one of the most important things right now is learning how to work with other people, how to work effectively in teams, and how to be a good individual contributor or leader. If I were in a hiring position, I would look for the soft skills a lot. I'd say all three are important, but I'd almost emphasize the soft skills as one of the most important.

Andrew Obwocha: That's something I've heard very commonly in these interviews, and it's not really what you hear in university. In university, it's almost like you need to get better, do more projects, contribute more to open source, and get your technical skills up. Which I think is true to an extent, but there's also the limitation that you need to be able to communicate and come across as approachable.

Justin Stevens: Yep, it's very important. And another thing that's so important is finding a place that feels like a really good fit. For me, with Art of Problem Solving, it was a natural fit. I've been teaching for them for a long time, I did an internship with them, and I'm continuing some of that work. These were my people from the beginning because I love mathematics and education.

Anyone could try to chase the big money at a company like Meta, but I would say it's more important to find a place where you will wake up every morning and think, "I can't believe I get paid to do this kind of work. This is just so cool." That's the feeling I really want people to have in their careers.

Andrew Obwocha: Just on that, is there a way you could abstract that advice for a wider range of people? I've heard that for a new grad job, you want to target very good mentors rather than very good pay because they tend to set you up on a steep trajectory. But others say if you start at a high pay, you can climb that ladder faster. There's a lot of conflicting advice. Do you have any heuristics for a new graduate to find a community where they feel accepted while also earning a decent salary?

Justin Stevens: There was a talk at a conference a couple of years ago that emphasized three pillars of a good career: fun, learning, and money. Obviously, you want a job that pays you enough to live a good life, but you also want a job that's fun and where you are learning skills that are relevant to your career and personal interests.

So, when you're looking at companies, try to find one you feel very mission-aligned with. Read their mission statements online. If you get the opportunity, talk to someone who works at the company and understand their goals and their five-year plan. If you want to be a part of that mission, try reaching out to someone and say, "Hey, can I just grab 15 minutes with you to chat quickly?" I'd argue that is more important than getting the highest-paying job immediately out of graduation because you want to set yourself up for a long career. I have friends in finance that probably make half a million a year, but at the same time, they get burned out really quickly because it's a grilling industry. So everyone's mileage will vary.

Andrew Obwocha: That is true. So one last question before we wrap up. You're the founder of the Undergraduate AI Society. Looking back, what was an unexpected lesson you learned from starting this community that ended up shaping your career or had a significant impact on you?

Justin Stevens: Oh, yeah. UAIS has had a really big impact on my life. I've been out of it for a while now, but I'm still really impressed by the work that you guys are doing. I love seeing all the posts on LinkedIn. I know you have Dr. Richard Sutton giving a talk at your welcome event soon, so I'm really excited for you all to have him there.

It was such a great lesson in leadership and networking skills. Honestly, founding UAIS was one of the things that connected me the most with a lot of professors and members of the community. Jonathan Schaeffer was one of the early mentors of the club, and he helped set us up with a speaker series that introduced me to a whole network of people. Honestly, without UAIS, I don't know if I would have done a Master's degree or pursued the career I'm doing now. I knew when I came to the U of A that I wanted to do artificial intelligence, and I wanted to make it more accessible for undergrads to get involved early on. It makes me so happy to see, several years after its founding, that the club is still going strong and inspiring people like you. I hope to continue to be a resource for others in their careers.

Andrew Obwocha: It is a wonderful community, and it has grown to be quite large. I don't know if you had foreseen it becoming this big, but I do have one question. When you were starting out, was there a particular challenge you faced? Was it getting people to join, getting speakers, coming up with ideas, or something else logistically difficult?

Justin Stevens: Logistically, the most difficult thing was honestly getting the paperwork approved. I spent so many hours going through the paperwork for starting a new club, getting a bank account set up, and all of that. I had no idea how much work that would be. The interest in the club was there from the beginning. I was lucky to be based out of the Student Innovation Centre, and the person running it helped send some emails out. To my surprise, we had 15 to 20 people at both of the first two sessions.

Andrew Obwocha: Right now, when we run a session, it depends on the season. During winter the numbers dip, but during the fall we expect between 20 and 40.

Justin Stevens: That's awesome to hear. The interest was there, and one of the most important things for me early on was finding an executive group of people that were all really tied to the mission. This goes back to what we were talking about with companies. I found a group of people so passionate about the club's mission. Two of them went on to be later presidents, Giancarlo and Paul, who I think you also interviewed. Some of them are still close friends of mine, so I'm really grateful for that initial group.

Andrew Obwocha: That's fantastic. We'll continue to carry the mantle. I'm looking forward to seeing you at the event in November.

Justin Stevens: Absolutely. Yeah, thank you so much for your time. It's been great talking to the founder of the club.

Andrew Obwocha: Thank you so much, Andrew. It was a really great chat with you too. All the best to you. I really hope your undergrad continues to go nicely.

Justin Stevens: Thank you so much. Take care.