SEI Cyber Talk (Episode 3)

What's It Like to Intern at the SEI? **Ritwik Gupta, Dominick Gurnari and Sandra Sajeev**

Ritwik Gupta: Hey, everyone. Welcome to another SEI Cyber Talk. I'm Ritwik Gupta, your host. I'm a machine learning research scientist here at the Emerging Technology Center, <music fades> and with me today I have two of our great interns, Sandra and Dominic, who will be talking to us little bit about what it's like to be an intern at an FFRDC like the SEI.

For those of you who may not know, an FFRDC is a federally funded research and development center, and so what that means is that the government, in our case the DoD, the Department of Defense, gives us a bunch of money to do a whole bunch of research from them as trusted advisors. We don't bid on contracts, we don't sell anything to them, but our job is to basically take cutting-edge software, cutting-edge research coming out of CMU and other places, and get it working for DoD relevant applications to work on the mission, and so Dominic, want to tell little bit, us, about yourself?

Dominic Gurnari: Of course, yes. I'm currently a junior at the University of Pittsburgh, studying Computer Science, and I'm currently working with Ritwik right now on applied machine learning, implementing inverse reinforcement learning into the ROS 2 system.

Sandra Sajeev: So my name is Sandra Sajeev. I am a senior in ECE right now at Carnegie Mellon, and my focus areas are computer vision, as well as machine learning engineering, and I get to do some of that here at the SEI.

Ritwik Gupta: Cool. Well, I mean I know you guys get to do that because I get to tell you what to do. Ha. So how'd you guys get here? What's your background then?

Dominic Gurnari: Well, I primarily have background in web development and a little bit of mobile app development. However, I wanted to get into a lot more research based, being able to apply mathematics to a lot of my problem-solving skills, and use coding as the means of doing it.

Ritwik Gupta: Got you.

Dominic Gurnari: So based off that I met one of our co-workers last year at Hackathon, and he told me about this place and I thought it'd be great idea to check out in the fall, as I already had a summer internship, and I've really liked that there's a lot more freedom here. So I, you know, sent a application in and here I am today.

Ritwik Gupta: Cool. What about you, Sandra? What's your background?

Sandra Sajeev: Yeah. So I guess I kind of started out with coding and stuff back in high school, but I primarily worked on, like, data visualization stuff.

Ritwik Gupta: Okay.

Sandra Sajeev: So I always knew that I had an interest in learning more about data and how we can kind of tell stories with data, but it wasn't until I got to CMU that I learned more about machine learning and, like, computer vision, and those two areas really stuck out to me and I hadn't really been able to apply it in industry yet, which is why I decided to come here and kind of explore it more during the semester as well, after taking courses on it.

Ritwik Gupta: Got you. And so that's interesting, because you say you came from industry, or you wanted to apply in industry, but we're not really industry.

Sandra Sajeev: Yeah.

Ritwik Gupta: Is there--so actually, let's start with this. What do you do here? Like, what's your project? What are you working on?

Sandra Sajeev: Sure.

Ritwik Gupta: Let's talk a little bit about that.

Sandra Sajeev: Yeah. So right now, I'm working on two different projects. So one of them is like a natural language processing project that I'm working on with another researcher at the ETC. So it's part of this thing called Pitchfest, where we get to kind of explore a new idea, and we're trying to apply NLP to academic research and, like, academic papers, technical papers, to see what type of insights we can derive from that. So I've been working with clustering, textual summarization, as well as hypothesis generation. So we're kind of looking at these new ways to apply NLP for more technical reviews or technical papers, because current research has been on stuff like Amazon reviews or Twitter.

Ritwik Gupta: Got you.

Sandra Sajeev: Twitter, I guess, like, little Twitter posts, right?

Ritwik Gupta: Sure, yeah.

Sandra Sajeev: Yeah.

Ritwik Gupta: Like tweets.

Sandra Sajeev: Tweets, yes, so--

Ritwik Gupta: So let me--so you're taking basically like paper reviews, right? Like, so if I submit a paper to a conference--

Sandra Sajeev: Yes.

Ritwik Gupta: --and I get a review back, you're taking those reviews and you're doing, like--

Sandra Sajeev: So we're looking at abstracts of papers.

Ritwik Gupta: Okay.

Sandra Sajeev: And then seeing what we can derive from that.

Ritwik Gupta: What do you mean what can we derive? Like, what's the task?

Sandra Sajeev: Like, basically if we can determine different similarities between lots of papers.

Ritwik Gupta: I see. Okay.

Sandra Sajeev: So starting to see trends in the field--

Ritwik Gupta: Got you. Okay.

Sandra Sajeev: --based on NLP.

Ritwik Gupta: So the idea is by kind of using NLP on the abstract of the paper, you want to figure out what papers are similar to each other, which ones are dissimilar to other.

Sandra Sajeev: Which ones are contradicting each other, and which ones are--

Ritwik Gupta: I see.

Sandra Sajeev: --supporting each other. So then we--

Ritwik Gupta: Oh, that's interesting.

Sandra Sajeev: You can kind of see how your stance kind of looks in the overall field.

Ritwik Gupta: Got you.

Sandra Sajeev: So we're looking at quantum papers specifically for this project.

Ritwik Gupta: Okay.

Sandra Sajeev: Yeah.

Ritwik Gupta: And then when you said hypothesis generation, what do you mean?

Sandra Sajeev: Yes. So hypothesis generation basically is if you can create one sentence about a given topic. So for quantum it is--it's like quantum circuits perform better than, like, regular circuits.

Ritwik Gupta: Okay.

Sandra Sajeev: Perform better. So that could be a hypothesis. You can see which--I guess you can try generating these hypotheses from an abstract.

Ritwik Gupta: Oh, awesome.

Sandra Sajeev: So--yeah.

Ritwik Gupta: Okay. So basically, given an abstract, try to figure out what--the research statement they're trying to say.

Sandra Sajeev: Yes, the main idea.

Ritwik Gupta: Oh. That's really cool. So made any progress?

Sandra Sajeev: Some progress. It is still, like, a very hard task, because NLP in general is a very hard task, as well as--

Ritwik Gupta: Right.

Sandra Sajeev: --trying to--because what we're trying to do is going into, like, machine translation and, like, machine understanding and that's, like, a very difficult problem.

Ritwik Gupta: Sure.

Sandra Sajeev: But we have been able to do, like, some basic things, like, with clustering and textual summarization. That gives us, like, a good step forward.

Ritwik Gupta: Yeah. No. We have a whole Language Technologies Institute here at Carnegie Mellon who can probably attest to the fact that this is a really difficult task, so--

Sandra Sajeev: Yeah.

Ritwik Gupta: --I mean, since NLP isn't really your background, is this something that you've been enjoying or...?

Sandra Sajeev: It's definitely been something that I have enjoyed. I've, like, taken a class on it before, but I definitely think computer vision is where I really excel and what I really like, and that's why I'm really enjoying starting to work on the project with you.

Ritwik Gupta: Okay.

Sandra Sajeev: The DIU project, where we get to kind of generate a data set of different satellite images, multi-band satellite imagery, and trying to assess buildings damaged from that. So that kind of goes directly into the field that I'd like to be working in, so, like, machine learning engineering. Kind of working with very large data sets, being able to really, I guess, look at scalability and efficiency within machine learning, which I feel like isn't really stressed as much in an academic setting, in a research setting, so I've been enjoying starting to work on that and starting to create the pipeline for that.

Ritwik Gupta: Yeah. I've been enjoying giving you some work to do, so--

Sandra Sajeev: Yeah.

Ritwik Gupta: Takes it off of me, right?

Sandra Sajeev: Yeah.

Ritwik Gupta: Dominic, why don't tell about what you're working on here?

Dominic Gurnari: Currently I'm working with Ritwik on inverse reinforcement learning and applying it to robotics. So the whole goal of our project is to detect anomalies in a system. So primarily we've been using something, a robot with wheels, but, you know, it could be applied to something as like a self-driving car, and it's using this inverse reinforcement learning to detect what can be anomalous activity in a given state, for example.

Ritwik Gupta: Oh, that's really interesting. So basically, you know, you get to take a really advanced research, which is inverse reinforcement learning, and apply it, like, actually get it applied in a real physical manifestation, right?

Dominic Gurnari: Right. So I think that's one of the most interesting parts about working here is being able to take a lot of this cutting-edge research developed here at Carnegie Mellon University, and apply it to something physical that can be used in industry work. The whole point is trying to be able to prove that we can apply this in real-time, so it's really interesting to see that, you know, academics and academia and just research in general can actually be applied in real-time.

Ritwik Gupta: So that's interesting. So, I mean, obviously there have be lot of challenges with all this, right? Inverse reinforcement learning is a really kind of cutting-edge field of research and so is ROS 2, this really new kind of still untested robotics middleware, so what are some of the challenges that you're facing when you're implementing all this?

Dominic Gurnari: Well, some of the challenges that I've been facing is the lack of documentation and lack of research outside of the university on these two topics.

Ritwik Gupta: Okay.

Dominic Gurnari: However, I think one of the helpful things working at the ETC is being able to work with a group of, you know, very intelligent individuals who have all had experience working with new technologies before. So it's not very stressful when it comes to learning a new technology and being able to use that in real-time.

Ritwik Gupta: I will note, for the audience, my mom would dispute that very intelligent claim for myself.

So, you know, again, it's really not your background, right? You know, you've--because you'd never done inverse reinforcement learning. Most people probably haven't. You probably haven't done robotics. So how does that, you know, how's the transition of going from working in kind of web development, like you said, to this cutting-edge research, pretty much?

Dominic Gurnari: Well, I think it's very interesting because when I was doing a lot of web development, a lot of it's very--it's very task-based, and since I had to be able to complete a task by the end of the day and make sure it worked, make sure there was no bugs in it. However, working with such new technologies and being able to, like, apply machine learning to real-time, I've been able to do a lot more of my own thing. I've been able to put my own spin onto it and really develop my interest in the field by working on it. It wasn't very tasky, but it was a lot more free-form, being able to develop my interest and put my own spin on it.

Ritwik Gupta: And actually, speaking of your background, you both had internships at industry previously, and I'll let you guys talk about that little bit, but what's the difference between

interning at a company, you know, somewhere in industry, and interning at an FFRDC, like the SEI? Sandra?

Sandra Sajeev: Okay. Yeah. So I guess, like, the past two summers I've interned at Tesla, so that's definitely been a very different experience from working here.

Ritwik Gupta: They're a lot bigger.

Sandra Sajeev: They're a lot bigger. They're very fast-paced. You don't know a lot of the people in your overall team. So I have, like, enjoyed the intimacy of, like, working here. It feels more kind of like a startup, but at the same time we're not always--it's not as, like, cutthroat as, like, a startup.

Ritwik Gupta: True.

Sandra Sajeev: Because we're not always looking at deliverables. We have time to kind of explore different technologies and do basically research, which I've really enjoyed. I think when you're doing an internship during the school year, it's nice to have that flexibility to kind of go out and learn more, because this is kind of like an extra class, in a sense.

Ritwik Gupta: Okay.

Sandra Sajeev: Except you do a project too.

Ritwik Gupta: Yeah, it's kind of an interesting perspective, so--

Sandra Sajeev: Yeah.

Ritwik Gupta: --I mean, again, like, for the audience who may not know, the ETC, we're about 30 people, and the--most of the stuff that we do is basing applied research in the areas of artificial intelligence and human machine interaction, and I've never heard anyone describe it like an extra class where get do a project. So I assume--

Sandra Sajeev: It's--

Ritwik Gupta: --that means that you're learning a lot, right? It's not--

Sandra Sajeev: Yeah. You're learning a lot.

Ritwik Gupta: Okay.

Sandra Sajeev: So I think, like, learning is, like, the biggest part of it. Because yes, you learn in an internship.

Ritwik Gupta: Yep.

Sandra Sajeev: But at the same time, you're always doing something. You always have, like, something to show, something to show at your next agile meeting.

Ritwik Gupta: Sure.

Sandra Sajeev: Whereas this is different from that. We don't--

Ritwik Gupta: Yeah, there's no--

Sandra Sajeev: --strictly do (inaudible) or like--

Ritwik Gupta: There's no clear goals all the time, right?

Sandra Sajeev: Yeah, yeah.

Ritwik Gupta: Yeah. It is tough, I would say then, right? Like, it's because, like, you know, normally it's really nice to have this list of things to do.

Sandra Sajeev: Yeah.

Ritwik Gupta: Would you say you prefer that kind of environment where you kind of do like task A, task B, task C, you know, it's on your Jira ticket, you finish it and move on, or kind of prefer the openness?

Sandra Sajeev: Well, Jira annoys me.

Ritwik Gupta: Okay.

Sandra Sajeev: But I do see the, I guess, the positive aspects of it too, when you look at-- because then you have, like, a measurable scale of how much you actually did.

Ritwik Gupta: Sure.

Sandra Sajeev: But at the same time, I think working here, it gives you kind of like a different perspective. I would ideally like something that has both, you know.

Ritwik Gupta: Okay.

Sandra Sajeev: Time to explore but still, like, "Let's do these tasks," and, like, Jira tickets and, like--

Ritwik Gupta: Yeah.

Sandra Sajeev: --see your measurable progress, but--

Ritwik Gupta: And, I mean, with your--that kind of fits in with your interests, right?

Sandra Sajeev: Yeah.

Ritwik Gupta: You want to be ML engineering, which is the engineering side of things.

Sandra Sajeev: Yeah.

Ritwik Gupta: And then the ML is--so that's pretty cool.

Dominic, like, same thing. Same question for you, man. You know, working industry, working here, you know, what's cool, what's not?

Ritwik Gupta: Well, in particular, I think one of the most challenging things about working in industry as an intern was the amount of impact I was actually making as an intern. You know, I didn't really feel like I was doing a lot. I felt like I was on a side project. I didn't really get--I got to learn a lot, but in terms of experience and pushing something out, I really didn't think I was making a huge impact.

However, one of the things I really love about working here is being able to make sure that I am making an impact, which is with this cutting-edge research and actually being able to apply it, I'm not just, you know, 1 out of 100,000 people anymore. I'm 1 out of 30 people, who's, you know, applying this in real-time for the first time. So a lot of the work I am doing, it's a lot more impactful and a lot more meaningful to just my career in general.

Ritwik Gupta: Sure. What would you say--both of you, what would you say are some of the down-sides of working FFRDC, compared to industry?

Sandra Sajeev: Mm--

Dominic Gurnari: I think it could be difficult because it is a little more open and free-form. So you're not always 100 percent sure of what your goals are. You could be researching something

and trying to understand what you're trying to accomplish, but it's not always very clear. I know it takes me, you know, asking questions of three or four people at times just to find out, "Okay. What is the task I need to complete today?" because I have no idea what I can do with this research.

Ritwik Gupta: Sure.

Dominic Gurnari: So I have all this great data, all this great material to work with. I have no idea what I need to accomplish though.

Ritwik Gupta: Sure. Yeah, and I think some of that falls on us too is to be able to set clear things for you guys. But again, the way research goes, what we may do today may not hold up tomorrow, so I can relate.

Sandra Sajeev: Yeah.

Ritwik Gupta: Anything for you, Sandra?

Sandra Sajeev: Well, yeah. I guess going off that. It can be really frustrating when you've worked on something for so long and then realize that it isn't really working or it's not possible.

Ritwik Gupta: Yeah.

Sandra Sajeev: But I think that's just the difference between doing research versus being in industry, and if you are doing something that is, has some sort of research value in industry, you're going to be faced with the same issue. So it's not something that I feel like is specific to the ETC or SEI as a whole. It's more just specific to the research field.

Ritwik Gupta: Right.

Sandra Sajeev: And the fact that we're working on these cutting-edge technologies.

Ritwik Gupta: Yeah. I'm sure if you go work at Google Brain or something they'll have the same exact issues, right?

Sandra Sajeev: Yeah.

Ritwik Gupta: They're also working on some super-cool stuff, but, again, as it happens with super-cool stuff, kind of really don't know where it ends up, so, you know, we're almost out of time here, and so, you know, I'd like to just close with, you know, for the audience, right? So we'll look at the camera over here.

Sandra Sajeev: Mm-hm.

Ritwik Gupta: Any recommendations or any suggestions for interns who are kind of considering SEI or FFRDCs, you know, or research areas in general?

Dominic Gurnari: I think it's really important that you give it a try, regardless if you're interested in it or not, because it's different perspective from a career standpoint, rather than just your general old industry completing a task every day. I think working here has definitely helped me realize lot of my career goals and, you know, put them into effect a little bit.

Sandra Sajeev: Yeah. I guess, like, for me personally, I think it's been a really cool experience to actually apply some of the stuff that I've been learning in the classroom. It's not always easy as an underclassman to get into internships that use machine learning, especially since you haven't had as much exposure to it. So I think internships like these are really good if you're planning on going into any field that involves computer vision, machine learning, or artificial intelligence, because it really gives you the opportunity to test these technologies out in a real setting, while at the same time looking at things like scalability and efficiency, which you might not get at a research lab here.

Ritwik Gupta: Yeah. Well, I mean, both you guys are terrific. I know both you guys are going to go on and really achieve your career goals, so best luck to both you both. Both of you both.

Dominic Gurnari: Thank you.

Ritwik Gupta: But one thing I'd like to say. We're always looking for terrific interns here at the SEI. You know, we do a lot of cyber security research, applied artificial intelligence, software engineering, hardware engineering, what have you, and so if you're interested in kind of doing some research for the summer or the fall or the spring, and, you know, you have little bit of passion of working in a environment where things are kind of open-ended and there's no real set task all the time, please do shoot us an email at info@sei.cmu.edu or shoot me an email at rgupta@sei.cmu.edu <musical outro> and we'd be happy to talk to you and see what we can do.

So again, thank you, guys. Thank you, Dominic. Thank you, Sandra, and we hope to hear from you soon.

Related Resources

Internship Opportunities

AI at the SEI