EPISODE 62

[INTRODUCTION]

[0:00:10.6] SC: Hello and welcome to another episode of TWiML Talk, the podcast where I interview interesting people doing interesting things in machine learning and artificial intelligence. I'm your host, Sam Charrington.

What you're about to hear is the third of a series of shows recorded at the Georgian Partners Portfolio Conference last week in Toronto. My guest this time is Graham Taylor, Professor of Engineering at the University of Guelph who keynoted day two of the conference.

Graham leads the machine learning research group at Guelph and is affiliated with Toronto's recently formed Vector Institute for Artificial Intelligence. Graham and I discussed a number of the most important trends and challenges in artificial intelligence, including the move from predictive to creative systems, the rise of human in the loop AI and how modern AI is accelerating with our ability to teach computers how to learn to learn.

Georgian Partners is a venture capital firm whose investment thesis is that certain tech trends change every aspect of a software business over time, including business goals, product plans, people and skills, technology platforms, pricing and packaging.

Georgian invest in those companies' best position to take advantage of these trends and then works closely with those companies to develop and execute the strategies necessary to make it happen. Applied AI is one of the trends they're investing in as our conversational business and security first.

Georgian sponsored this series and we thank them for their support. To learn more about Georgian, visit twimlai.com/georgian, where you'll also be able to download white papers on their principles of applied AI in conversational business.

Before we jump in, if you're in New York City on October 30th and 31st, we hope you'll join us at the NYU Future Labs AI Summit and happy hour. As you may remember, we attended the inaugural summit back in April. The fall event features more great speakers including Corinna Cortes, Head of Research at Google New York; Davide Venturelli, Science Operations Manager at NASA Ames Quantum AI Lab; and Dennis Mortensen, CEO and Founder of startup X.ai. For the event homepage, visit aisummit2017.futurelabs.nyc. For 25% off tickets, use code TWiML25. For details on the happy hour, visit our events page at twimlail.com/events.

Now, on to the show.

[INTERVIEW]

[0:02:44.4] SC: All right, everyone. I am here at the Georgian Partners Portfolio Conference and I've got the pleasure of being seated with Graham Taylor. Graham is a Professor at the University of Guelph here in Canada. He did a really interesting talk today on some of the challenges and opportunities associated with machine learning and AI, and particularly around his research area in deep learning. We're here to spend a little bit of time chatting about that.

Graham, welcome to the podcast.

[0:03:13.3] GT: Thanks for having me on the program. I told you this is my first time doing a podcast, so I'm really excited to be a consumer of podcast actually, get back to the podcast community. So thanks for having me.

[0:03:21.4] SC: Nice. Absolutely. Absolutely. Why don't we get started by having you tell us a little bit about your background and how you got involved in machine learning and AI and what you're up to nowadays?

[0:03:32.5] GT: Sure. Currently, I'm working as a professor at the University of Guelph. I'm also a member of the new Vector Institute for Artificial Intelligence, which has started up and getting ready to move in in November here in Toronto. I'm the academic director of a program called NextAI, which is a founder development program for startups, specifically working on AI technologies.

I wear a number of different hats, but they're all focused on artificial intelligence and machine learning, so let me tell you a little bit about how I entered that space. I often get asked this question of how I got into AI and fortunately, I can point at one specific point in my life, which really convinced me. That was an inspiring professor when I was an undergrad student at the University of Waterloo.

I had a course. I believe the course was called machine intelligence. The way that course was setup was to actually encourage us, all the students to write AI programs, to play each other's AI programs in this game called Abalone.

It's amazing looking at this effectively. I would say it's been at least 15 years now since we did this, but with all the news last week in terms of this new AlphaGo system by Google DeepMind, playing the game of Go and how it was trained entirely by self-play, this is exactly what we were trying to do on this assignment.

It was an easier game, but it was really inspirational to build these agents, play against each other. Our team ended up winning the competition and made us really proud and excited and eager to do more work. Yeah, so that's what started it all off.

[0:05:09.3] SC: Okay. What's your path then so far?

[0:05:12.5] GT: From that point as an undergrad, I got so excited about the potential of AI. I wouldn't say it all, I was sort of going to predict what would happen to this time and how huge it's growing. But I just observed being a technical person. I was really excited about building those tools and wanting to learn more.

I went to University of Toronto and knew they had a good machine learning program. They had a number of faculty there who I was aware of their work, lots of graduate students. It seemed like a great place to be. It was not too far away from Waterloo and London, Ontario where I grew up.

It seemed like a natural choice. Now I had no idea how big that group would become, and the influence that Toronto machine learning group would have on deep learning today. Some people ask me about this and I say, "Well, I don't know. I stumbled upon this group and it seemed the right place to be for all the right reasons." But it ended up being an amazing time to be there. This is for me 2004 to 2009, really the start of the deep learning movement.

A lot of the individuals that are leaving the major industrial research labs or even the non-profit efforts like OpenAI, for example. They were students and a group at the time. A lot of the key papers and key ideas that were published and disseminated at that time have gone on to be the foundations of it. A lot of it came out of that group.

The group said we're having close collaboration with, including the Montreal – now they're called as Montreal Institute for Learning Algorithms run by Yoshua Bengio. Then the group at NYU, which at that time was led by Yann LeCun. Mentioning NYU, that's where I went immediately after PhD.

We had a good working relationship between these three labs; Montreal, Toronto and NYU. I considered both as post-doc options, but ultimately decided to go to New York for a couple years and work there as a post doc, but felt the pull to come home. After that it was really exciting working in New York with Yann, with Rob Fergus, another professor named Chris Pregler.

That really actually got me working in computer vision more. Then I came back in 2011, and my heart was really pulling me not just towards come back to Canada, but also towards an academic position. I had the opportunity to join the faculty at Guelph in 2012, and that's when I started. It's been five years.

[0:07:37.2] SC: Nice, nice. You mentioned a bunch of names there, but you didn't mention that your advisor for your PhD was Geoff Hinton.

[0:07:42.7] GT: Yeah. I didn't mention him about the names. I didn't credit my PhD advisor. I'm sorry, Geoff. I'm sorry.

[0:07:48.9] SC: I've talked to so many people who he's impacted via advising and other ways.

[0:07:55.5] SC: That's right. I was co-advised by Geoff and another individual named Sam Royce who is really influential in machine learning. He passed away actually in 2009, right when it started at NYU. It was tragic to lose him, but I certainly wanted to note his influence on my PhD as well.

It was really amazing having one very senior advisor, Geoff; really experienced having worked in the field with also someone quite junior. Sam had started as a faculty I think around 2005 or so. He was full of energy and also helped me along the way.

[0:08:34.7] SC: Awesome, awesome. So you did a talk here this morning. Tell us a little bit about what your talk was about.

[0:08:39.6] GT: Sure. I broke my talk up into three parts. The first part was just introducing myself and telling people a little bit about the work that we do in Guelph and the types of machine learning problems we're interested in.

The second part of the talk was focused on the challenges and also the opportunities in AI. That was more of a technical discussion, what was coming around the corner. Then the third part of the talk was about some of the barriers, some startups might be facing. We're here at the Georgian Partners Portfolio conference, so there is many startups in the audience. Done a lot of work with startups. I tried to focus on, again some of the barriers they might face building companies.

[0:09:20.3] SC: Why don't we start with kind of a rundown of the challenges and opportunities as you see them?

[0:09:25.7] GT: Sure. I started by talking about the technological changes coming towards us, and I think the first one that I started with was the move from what I would call largely predictive systems to creative systems.

When I say predictive systems, I mean these systems that we're used to interacting with on a daily basis, the systems that might give us a temperature forecast tomorrow, or we might get up our mapping application and it would tell us an estimated time to get from A to B, or the people on the financial side might be interested in forecasting the price of a particular financial instrument the next day.

Those types of inputs are either category or they're a number. They're pretty low-dimensional and there's usually a single right answer. When I talk about the movement towards creative systems, I'm talking about systems that produce high-dimensional output and where there is no single right answer.

Examples of this is more on the creative side would be art creation, or poetry, or music. While these are some of the more culturally-flavored activities which get some attention, there is also some real commercial applications such as automatic e-mail reply, or conversational dialogue systems, or showed a proposed design for a robot that's creating meals and serving them to you every day. I talked about some of the challenges in building those kinds of systems.

[0:10:49.9] SC: I thought that one was pretty interesting. Folks that listen to the podcast will be pretty familiar with the idea of generative networks and style transfer and creating all the efforts we've seen to create movie scripts and poetry and all these kind of stuff.

The kind of art example is really – has been really front and center for me for a while. But then when you described the recipe creation, that's a totally different domain and one that we hear about all the time. Maybe because there are a bunch of different disciplines that need to come together for us to really explore or fulfill that, the Jetson's vision.

[0:11:31.1] GT: Totally. Yes.

[0:11:31.6] SC: But we're quickly moving towards an area where a lot of opportunities and value exist around generative using known or specifically an AI in general for generative purposes. What are the key challenges there in your mind in that transition?

[0:11:49.0] GT: I'm particularly interested in this as an engineer. I work in an engineering school. I build things. I think we're seeing design migrate over from purely human design to at least the next few years being machines and humans working together on design and building objects, whether they'd be recipes or they'd be parts for vehicles or aircrafts.

I think the major challenges in working towards a more algorithmic design would be what I pointed out this morning, namely the fact that there is no single right answer for design. You have potentially infinite number of solutions to a problem or designs that would be acceptable, and this makes it very hard to come up with reward signals, or what we would call objectives for machine learning systems depending on how they're trained.

For us, when we think about a simple task like image classification; image goes in, category comes out, we compare it with the ground-truth category. There's usually a single right answer. For a system going back to this recipe example, how do you measure the output of the system that cooks your meal?

I mean, you can I guess get some subjective judgment of the person eating the meal, but it's not really calibrated with other people. It's also just that single reward – these are the types of rewards maybe they're being used in reinforcement learning systems. They're very weak signals as well. It would be nice to maybe find some sort of medium between this weak subjective reward, or the explicit guidance that works for supervised learning systems.

[0:13:27.4] SC: How far along are we in that – against that challenge, like that particular approach?

[0:13:34.1] GT: Yeah. I think we've seen – as you've mentioned, the listeners on the podcast are going to be familiar with examples particularly on the visual side of generative systems. That's kind of where we're stuck right now is evaluating generative systems, again coming up with quantitative metrics; one, to evaluate them, but also maybe as a way of feeding this kind of quantitative metric back into the system to make them better. How do we improve the objectives to train them?

Then we've also seen a lot of progress really recently on the reinforcement learning reward side. But we aren't really anywhere I think on this sort of merger of those systems. I think there's a lot more to be done.

Particularly, we've seen a lot of nice examples in the generative space of language. So machine translation systems, conversational dialogue systems. But I think we're still stuck in coming up with the right kinds of metrics. So you haven't say an English sentence going in or a French translation coming out.

Again, there's so many possible valid translations. We're still in most cases stuck at measuring a single, maybe I would say a canonical example giving some data set with the output of the system, rather than really considering the space of the potential answers it could give.

[0:14:53.1] SC: Okay. One of the examples that you used was Inbox by Google, which I also use. You went a little further that you have a percentage in your mind of the time that you use that for responses; quite there yet. But I do use the responses every once in a while. But you also talked about – you talked about a bunch of concepts, you know transfer learning, meta learning and keyshot learning.

One of the questions that I had as you were going through this was what are the mechanisms and approaches for using in a large-scale system the feedback that you provide by selecting one of these responses in conjunction with the broader model that's trained for everybody.

What's that problem called, what are the approaches how far are we along in the developing a body of thinking around that

[0:15:50.7] GT: Right. I think I was referring to this Google Inbox client. I'm a big fan of it, a big user of it. Like you said, some percentage of the time, the auto e-mail reply feature. It's what I would call a human in a loop system. What I was saying earlier this morning is essentially – I wouldn't want to necessarily hand over all my e-mail to an automatic reply system.

AI is not at that stage where I could stop, write e-mails and people would just interface with me through this agent. But it's working at the level where it can propose several candidate replies, and I can still execute judgment over there. I can decide not to send the e-mail at all, I can decide to not accept the proposals and write an e-mail myself. I'm still completely in control.

It's making me more efficient when that once in a while I propose that something that I can just click on and it will send. I would say this is a system, it's a human in the loop system, it's where I maintain the judgment over what goes out and I see this is a effective paradigm of humans and machines working together over the near term.

But I also really like this idea of the transition from full human control over a particular task, all the way to fully automated performance, but this gray area in between. A company that I co-founded in Toronto named Kindred, they're actually exploring this for robotics, where essentially the company is teaching robots to perform tasks that are very difficult to automate, by allowing a human operator to control one or more robots.

The robot will be autonomous, but when it gets into trouble, it can be taken over by a remote operator and sort of gets it out of that whatever it stuck doing. The hope is again to, if this happens enough times, the robot learns about the way that the human assisted it and getting out of that particularly difficult situation, so that it becomes more and more autonomous. Again, it's not 0% or a 100% automation. We're exploring that gray area in between. I really like this paradigm.

[0:17:53.5] SC: As you're using Inbox, it's presenting you these possible e-mails that you might want to respond with. It gives you three. You choose one. That's potentially augmenting the set of training, label training data that the system has.

One way for Google in particular or someone building a system like this in general is to throw that all in and continuously update the model and produce better models that are trained on more data. It strikes me that another way for a system like this to operate is that there is – there's a general component of the model, but then there is a sub-component of the model that's personalized to me and the way I respond.

The question is really, is anyone doing that? Does that have a name? Are there architectures for that? Have you ever come across that?

[0:18:51.9] GT: I would say this fits into the idea of personalization. I think it's important for a product like Inbox to have some element of personalization. A colleague actually told me that he doesn't use Inbox, because it makes him sound like a California dude. He said and puts exclamation marks on everything he says and uses terms like "Awesome!" Which he wouldn't say himself.

[0:19:14.4] SC: Okay. Interesting.

[0:19:16.2] GT: He also claims that when I send him e-mails, he can tell when it's coming from the Google Inbox auto-reply system. Again, what would fix something like this and maybe make him an adapter is a system that would adapt to his own style of writing e-mails.

[0:19:32.2] SC: This also ties back to the bias element of the conversation in a more subtle way than we sometimes think about it.

[0:19:39.8] GT: Exactly. Why is it making him sound like a California dude? Well maybe it was fit on a bunch of e-mails from people in California, right? It certainly ties back to that. I think in terms of how to talk about this, and you even raised the idea of having a model that's been built from a whole lot of data, sort of a master model and then personalized model for each of the people and sort of adapting.

I think we do see it. It's an instance of transfer. I was talking about how do you deal with these problems where you have a very limited amount of labeled data? I said, "Well, it's very popular right now to train a model in a big generic data set, and then cut off the top of it and then replace that with something more specific and then train on a very much smaller set of data." So you can take something like generic object recognition and big image net style data set, and then tackle a task like bird species classification, which is fine-grain, but you have much less data.

That works, because in the big system there are birds in it right? The data set or image net has birds. So you can learn about feathers and wings and colors of birds and beaks and those sorts of features. That works when there is good match between a two different domains.

This is what you're saying in terms of personalization it can be viewed that way as well, like you have a large data set of a whole bunch of different speakers and you can learn a model on that, but then you want to transfer, adapt the system to a particular individual where you have a smaller subset of data.

It make sense. You wouldn't want to necessarily have a model for each individual person at work in isolation, because there's probably not enough data there to generalize well. In that case, you want to capitalize from all of the e-mail that Google is holding in its servers, from people using Gmail.

[0:21:26.4] SC: Okay. Okay. Additional challenges that you were describing.

[0:21:30.0] GT: Yeah. I think we had gotten. We really only gotten across the sort of the two opportunities coming across. I can move in to challenges, or I could tell you about another couple things that are coming across and sort of trends. So maybe I'll try to finish those off.

The two trends that I hadn't mentioned yet, one was this idea of moving from careful human construction to learning to learn. Right now, like the assistants were the output of the hard work of graduate students and the faculty members advising them and researchers and practitioners.

I mentioned the migration from feature engineering to architecture engineering, the way that people describe deep learning is that, "Oh, it's the end of feature engineering. We no longer have domain experts who craft very specific features." We can learn all the features of deep learning.

[0:22:20.3] SC: I saw you had a picture of Stephen Merity's article on your slide, which I've talked about on the podcast a while ago.

[0:22:27.1] GT: Okay. Fantastic. I love that blog post and it really – I think it's totally accurate. We've moved into the world of architecture engineering. One way of getting out of this is essentially having these meta-learning style algorithms. I mentioned a specific example in our lab where we're dealing with multimodal data.

In this case we might have video and audio and we had motion capture coming in. We're figuring out how to actually merge those different modalities. I mean, the nice thing about deep learning models is with multimodal learning you have so many opportunities of how to extract representations from the different modalities and how many levels of that representations you should go for each of those and when they should be merged together and which modality should be merged.

But there is all these decisions to be made. You can either have a grad student like we had who was really skilled at figuring this all out and spends a year working towards a competition, but ultimately we like to hand that over to an algorithm that figures that out, and that's what we've done. That's one instance of learning in architecture. I know Google Brain has been working on this with a reinforcement learning. They worked on learning optimizers. They're now working on – other people are also working on learning activation functions.

Really like it's, yeah, handed over to the algorithm. This meta learning is really exciting. There is some great work that was done at Twitter before you go – show and move it over to Google Brain on. Learning an algorithm that's just good at few shot or one shot learning, also an instance of meta learning. It's an exciting area.

[0:24:00.0] SC: I think after Steven's article came out, I spent a long time trying to – through my interviews, trying to understand the process of architecting deep neural networks. I guess it took me longer to – retrospectively, it took me longer than it should have to figure out the radiant descent by graduate students. There are a couple of different versions of this and you just –

[0:24:25.2] GT: Graduate student descent. Yeah.

[0:24:26.3] SC: Graduate student descent. You know to, at least, it seemed like a year ago or so, like that was the state of the art. But since then, you described a number of methods for automating architecture. One of the ones that you mentioned was a Bayesian-based approach and some others. Can you go into a little bit more detail on the various ones that you mentioned?

[0:24:51.9] GT: Sure. So we explored a couple different approaches in my lab for the multimodal learning problem. One approach is Bayesian optimization, which a lot of people in the deep learning field are familiar with from the point of view as doing model search or hyper-parameter optimization.

These are the decisions that we all need to make about how many layers and how many units per layer and what kind of activation function. Then on the learning algorithm side, how long do we train for – should we use Adam optimizer, or should we use RMSprop, or what should our regularization coefficients be? There's all these decisions. With deep learning, there is more of these decisions in classical machine learning models.

People have proposed Bayesian optimization as a suitable tool, and it's actually been very successful in automating some of the hybrid parameter search. In our first example, we just viewed architecture as another hybrid parameter and we propose essentially a search base of potential architectures in which this modality fusion could happen, and then we had a Bayesian optimization algorithm with a –

The main technical achievement was a kernel or a way of assessing similarity between different architectures, and that was thee building block for the Bayesian optimizer to basically search over that space of potential fusion architectures. It came up with one that would beat the graduate descent method in about 30 or so proposals, different architectures.

The downside to that system is when I'm seeing 30 different architectures, each of those had to be trained and evaluated and then that result given to the Bayesian optimizer, such that it could propose a next one. It's this iterative method in which you're training full architectures to convergence, or evaluating them. You're choosing another one going back evaluating.

It gets quite slow. We explored a second approach, in which we do – we view architecture search as the caustic regularization. It's a meaty thing to say, but it's what you see in methods like dropout where people just knock out activities randomly. In neural networks, there's also drop connect, where people knockout weights.

This is done on an example, white example basis. So every time you present a new example to the model, you knockout a different subset of hidden activities or you knockout a different subset of weights. So they call this the caustic regularization, and it's been shown to make networks generalize better, and it was very popular until some things like BatchNorm came along and people started working with that.

Still, it's this general principle. For us, we did this kind of block-wise knocking out certain weights inspired by an approach by a graduate student at CMU called block out. What this student found with this block out is that if you knockout blocks of weights, this can give you very different architectural patterns made through this weight structure.

You can have mergers of groups of hidden units or splits, or you can just completely ignore certain features that are being discovered in a network, and basically propose a modality where a version of this.

As its training explored many different multimodal fusion architectures, then eventually converged to one network pretty well. That ended up being more efficient than the Bayesian optimization approach. That's pretty technical.

[0:28:10.8] SC: No, that's great.

[0:28:11.7] GT: That's good. Okay.

[0:28:13.6] SC: Then there was another challenge.

[0:28:16.4] GT: Yeah. So I talked about the idea, which is both an opportunity and a challenge of explainability in AI. I don't know if you've talk much about explainability yet.

[0:28:26.0] SC: I've talked a little bit about it. I don't think you mentioned it in your talk, but I did an interview with Carlos Guestrin, who has a paper called LIME, which seeks to do explainability. I appreciated you were quoting someone else, I believe and you took issue with – we often talk about neural networks as block boxes and I think you – well, you can tell the distinction on that.

[0:28:51.4] GT: Yeah, sure. I can talk about that. It was actually a quote by KyungHyun Cho, a researcher at NYU. He came to Toronto last summer very graciously to be part of this NextAI program for startups. A part of that program, we bring in world-class individuals like Cho and they talk about various things with doing a course on NLP.

But he criticized people calling neural network black boxes and said they're actually white boxes. That was kind of neat, because this – or after my talk, Nikola Paprano talked about block box versus white box subtext. It's the same concept here. Neural net is in some sense you can open them up. You can look at their parameters. You just happen to have hundreds of millions of parameters most of the time and they're uninterpretable.

If you're accessing them through an online service, like in Paprano's work they were trying to attack a method that have been deployed. One was on MetaMind Services, one was on Amazon, one was on Google. If you're interacting through the predictions, yes it's black box. But if you're the person evaluating machine learning system, or maybe you're –

[0:29:57.0] SC: If it's your model.

[0:29:57.6] GT: It's your model, it's black box. General –

[0:30:00.6] SC: Still uninterpretable. That's why we're keen to move to more interpretable systems or explainable systems in certain setups. We've looked at it in the medical space. We've looked at it in the financial prediction space forecasting, and then we've looked at the classical vision problems on the benchmark data sets that everybody else benchmarks on.

Yeah, I guess it's like when you teach about software and you're talking about requirements gathering and how much money you're going to spend on each stage of the software development life cycle, the same thing, it's as the rest. Like some problems require more careful consideration of risk, others don't.

I'd say that the same thing about interpretability and explainability. Let's see what the application is, who are the users, what are the risks involved. In many cases, like we want to make the system more explainable.

[0:30:57.5] SC: Okay. Then opportunities. We have arrived to opportunities.

[0:31:02.5] GT: Sorry, we've gone through a lot of opportunities and I'll move more into sort of barriers. Let's say that the last part was some of the barriers. When I talked about barriers, it was – I'll just go quickly through them. Data, the next one was talent, then the third one I talked about was building trust.

We can maybe go in and dissect each of these. First of all for data, I think in deep learning, we've seen a tremendous number of really cool examples of deep learning working in practice, but I would argue that it's been done in fairly limited set of domains.

Once I mentioned where the big three vision, speech and audio processing in a natural language. These are generally unstructured domains where there's a lot of data, labeled data in particular. These are the sorts of applications being pursued by the commercial internet giants, right?

Actually, it's something I've had a struggle with in my lab just motivating some students to tackle other kinds of problems where benchmark data sets are not available. I actually mentioned today, for example some agricultural applications that we've worked on.

But again, you're rewarded more as a researcher to conduct your experiments reasonably quickly, get your papers out and compare to other people in the literature. You know, so you download them internet, you propose your new architecture, you publish paper on it. At the end of the day if you want to solve a problem that actually, really important problem like growing food in an environmentally-friendly way, in a sustainable way and that gives decent yields for the farmers.

You want to explore say deep learning for remote sensing in agricultural fields, this involves a crazy amount of data collection. It takes a lot of work to get on the field, do those flights. You may be in flights, do the ground-truthing, which involves actually collecting samples here. We are looking at soil properties or nitrogen properties of plants.

This might take a summer, might take multiple growing seasons. You actually don't see the effects of any interactions until the end of the growing season when you actually can measure any yields. This is not the same timeframe of a lot of the experimentation that happens in machine learning.

Again, going back to the AlphaGo example, the system is able to play two and a half million games against itself, because it can carry out a game and get the reward in less than a second. In an agricultural situation, you can't get a reward in less than a second. It's six months, or longer.

Anyway, this is a bit of a ramble just saying that we're fairly limited in the ways that we're applying deep learning right now. It's a lot about the data, like how do you collect the data, where do you get the data, how hard is it to gather that and process it.

Anyways, there is a quote by Chris Dixon that I gave at the end, which was data is really the key ingredient to AI, because it's the missing ingredient. We publish our algorithms, like there are just great out algorithms out there, they're available to people.

Compute power has really grown and it's become cheaper, so we have access to great compute. It's really the data that we don't have. That's the missing ingredient for these sorts of problems I'm talking about. It's also for companies, it's the proprietary ingredient. People aren't publishing data sets as quickly as they're publishing papers on algorithms. But it's one of the challenges that I see is data.

[0:34:37.3] SC: Well, if you can maybe quickly summarize the other two and leave us with any final thoughts as we wrap up.

[0:34:44.0] GT: Sure. In terms of the other two, one is talent. I think this is the idea of companies faced with, "Well, how are we going to fill these positions where we need really skilled people in machine learning?" Whether your question is around, "Do we need PhDs or masters graduate is good enough? Can we take some of the training in a different area and move them into machine learning field?"

I think there's a lot of amazing stuff going on here, particularly in not just Canada, but Toronto. There is an announcement last week by the provincial government to fund back their institute with 30 million dollars to work with Ontario Universities to develop professional graduate programs and AI machine learning.

In five years, we're going to be looking at graduating a thousand students per year. That would be the goal for steady state in this province. I think we're addressing that issue with talent right now. But as I mentioned today, we also have a lot of professors leaving academics, going to working industry part-time or fulltime and we need to work on retaining those individuals.

I think decisions like starting – the government supporting AI initiatives, like Vector and Mile and Amy in Alberta, those are all working toward making academics attractive in this field versus industry. But I think we need to do more to encourage the people staying in academics or moving in to academics that continue to train in the next generation.

Then the final topic was on trust, right? Building trust. Actually, I've already touched on a couple of those issues. Explainability was one, bias and fairness. I tend to like the idea of using technology as – actually it's Nikola Paprano who said in his talk following mine, these ideas are on differential for privacy preserving algorithms to increase people's trust in machine learning systems.

Ruling out technology like fair representations for removing bias, from algorithms that make predictions. I feel pretty good about the future of AI, and I guess I'll summarize it down. It's a great field to be working in. This Toronto area is a great place to be working on these technologies. There's a lot more to come and I think in terms of the problems, I do see a diversification in the future of the types of tasks we're solving.

I think they're at least working with the startups in the next AI program. I also see a lot of interest both from the companies building these technologies, but also from the investment side, and companies that are doing social good as well. So building both profitable companies, but also solving real important issues. That's what I look forward to.

[0:37:25.9] SC: Awesome. Well Graham, thanks so much for taking the time to sit with me and share all that you shared about your vision for this and how you see it.

[0:37:34.6] GT: Thanks a lot. My pleasure. Great to meet you.

[0:37:35.7] SC: Great. Thank you.

[END OF INTERVIEW]

[0:37:40.3] SC: All right everyone, that's our show for today. Thanks so much for listening and for your continued feedback and support. For more information on Graham or any of the topics covered in this episode, head on over to twimlai.com/talk/62. To follow along with the Georgian Partner series, visit twimlai.com/gppc2017.

Of course, you can send along feedback or questions via Twitter @twimlai, or @samcharrington, or leave a comment on the show notes page.

Thanks once again to Georgian Partners for their sponsorship over the show. Be sure to check out their white papers, which you can find by visiting twimlai.com/Georgian.

Thanks again for listening and catch you next time.

[END]