Build 2017
Harry Shum, Executive Vice President, Microsoft AI and Research Group
May 10, 2017

HARRY SHUM: Good morning. It's great to be here. Thank you for listening to Satya, Scott and now me.

Let me begin my talk by taking you back to the start of the day. Satya talked about this new app pattern, this new world view of the intelligent edge and the intelligent cloud. Central to this is artificial intelligence. And this is what I'm going to talk to you about now.

I have been in the field of AI for a long time. It's amazing to see AI now becoming more a reality. I hope many of you by now know Microsoft Cognitive Services, but I'm not sure if any of you were in that small breakout room two years ago when we launched Cognitive Services with only four APIs. Maybe some of you were in that big breakout room last year when we launched it with 22 APIs. And this year apparently I have been promoted by Satya to the main stage.

Seriously, the time for AI is now. At Microsoft, we believe AI is about amplifying human ingenuity. This is our vision for AI at Microsoft. You may ask why now? Well, there are three big forces converging that are making AI possible. The first, big compute, especially how you can take advantage of huge computing power in the cloud. Second, powerful algorithms with amazing progress in deep learning and deep reinforcement learning. Number three, massive amount of data that infuse powerful AI applications.

Microsoft is in a very, very unique position to help you take advantage of these forces, because of three important and unique assets we provide for you. First, the Microsoft Cloud, you have heard all about the power of the cloud at a global scale from Scott and others. That's Microsoft Azure.

Second, we provide breakthrough AI algorithms coming from Microsoft Research. Microsoft Cloud plus AI brings all this innovation to you as a developer. And the third, what about the data? To have great AI you need great data. In particular, Microsoft Graph gives you access to the most important data for your business, that's your work data, all in a trusted platform where you own your data, and you have full control of it.

Let me tell you about some of the innovations that we have been driving. Let me start with my own area of computer vision. You have all heard about the deep neural nets making great progress for image recognition, for object recommendation. Two years ago at Microsoft Research we created a new approach to recognize images with a very deep neural network architecture called Resnet. With 152 layers we won the first place in all the three categories we participated in the Image Net Competition. I'm sure many of you in this audience now use Res Net to train your deep neural nets for your applications.

We are also making great progress in speech. Six months ago we became the first to reach human parity in speech recognition with the famous Switch Bot data set that is a collection of phone call data. And we beat the human error rate of 5.9 percent and we continue to improve.

While the researcher is inspiring, what's more exciting to me is what you can do with this exciting technology. Microsoft is a platform company. Developers will always be at the heart of everything we do. Let me show you how we are bringing AI to every developer.

A great example is Cognitive Services that I have been working on. Since we launched two years ago more than half a million developers have been using our services. Thank you. Cognitive Services provides the broadest and the deepest set of developer services for AI in the industry. Today we are releasing additional services, taking us to a total of 29 APIs across vision, speech, language, search and knowledge.

This includes the new Leader Index Service. It actually combines a lot of features from speech to face to vision to semantic understanding. And we're opening our Cognitive Services Lab for you, as well, so you can try a lot of new preview services, we'll start with some fun hand gesture recognition.

Since we launched the Cognitive Services we have been talking to many of you developers understanding what you like and what you want. Our big learning is that you really like customization, for example, Louis, our custom language understanding tool. And today I'm very excited to announce deeper customization across our services. So this set of services can custom train for your scenarios, while still using REST APIs, without the need of designing very complex models.

So this will make it so much easier for you to develop your own AI applications. Let's take a look at how an indie game developer, Human Interact, is applying these custom services, let's roll the video.

(Video segment.)

Wow, isn't that fantastic? (Applause.)

I couldn't even understand some of those names and terms in the video game. But Cognitive Services can. So our advancement with speech and language are enabling an important concept, I call it conversational AI. To me that's the most exciting thing about AI now.

It builds on what Satya talked about last year at Build, conversation as a platform. It represents the next user interface paradigm shift. Just like many years ago when we moved from command line to graphic user interface, we made it easier for people to interact with computers. But we still had to learn a lot about the computer, even with GUI. With conversational AI we move from a world where we have had to learn about the computer to a new world where computers understand humans and AI speaks human.

Imagine how you can build a whole new generation of conversational AI applications. Let me invite Cornelia up on stage to show you a demo bringing all of this together.

Cornelia.

CORNELIA CARAPCEA: Thank you, Harry.

And hello, everybody. (Applause.)

So Harry spoke earlier about customization and he showed you the LIS custom speech service and he also mentioned custom vision service. Together we see services make it really easy for you to build your own custom AI models across speech understanding, language and text, and now image understanding. And they're really easy to use. You just need three steps.

One, bring your own training data, usually a few dozen samples for training. Two, click a button to actually train your model within minutes. And three, you get your model automatically hosted behind the REST API that you can bake within your app.

Today I will spend some more time introducing custom vision service, which is our new cognitive service now available on Azure. What it does is it allows you to build your own image recognition systems and it's pretty flexible. It can work across identifying retail objects to food, to even landmarks. And to show you how it works I actually have this little app today that I'm going to show you and what it does is it takes a picture of a plant, or of a leaf, and it tries to identify what it is. And yes, we went outside and plucked a rhododendron.

All right, so it identified that it is a rhododendron from that one picture that I gave it. And the cool thing about this is that I didn't have to use any deep learning frameworks, I didn't have to find thousands of training images, and I didn't have to wait for hours for my model to train. It was actually pretty easy to do because of custom vision service.

And all I had to do is really upload a few dozen images to a photo album, which is what I have here, and you can see all of the plants that I've trained it over.

Then I click on this big green button, and within minutes I got my own endpoint hosted that I use in the app that I showed you earlier. You can follow these same steps to build your own model.

Now, cool, once you've built it, how do you actually keep it learning, how do you keep it improving over time as your app sees more and more images from your users?

Well, we've got you covered with this thing we call active learning. And what active learning does is it looks at all of the images that may flood your app, and it smartly selects only those that will add the most gains to your classifier.

So what it tells you is say we picked this image here. It's actually an image of a tea chamomile, but it only has about 52 percent confidence that you got it right, which is not very high. So it surfaces it to you, you can go and reinforce that, yes, it is a tea chamomile, save, retrain, and then you get your model smarter.

This in a nutshell is custom vision. It's available today. So go and try it out at customvision.ai. Take it for a spin. We can't wait to see what you guys build with it.

All right, customization is one of the ways in which we are improving Cognitive Services. Another way is by allowing you to use our state of the art machine learned models in a highly parallelized fashion.

So what that means is if you have, say, petabytes and petabytes of images, you can actually use our text understanding or image understanding model to get those images all analyzed.

And if you're familiar with SQL, then you can understand these 20 lines or so of U-SQL code that I'm showing you here. And what this code does, it uses our landmark understanding model across a load of images that I uploaded in Azure Data Lake and it's able to find all the ones that contain Mount Fuji in it.

Because of Azure Data Lake, it runs within minutes, finishes executing, and then you get your whole images analyzed.

Now, we ship these tools and services because we want to ensure that you can bake AI wherever you need it, whether it's a little consumer app like what I showed you earlier with my leaf classifier, or whether it's a massive data pipeline in Azure Data Lake.

But that's not where we stop. We want to ensure that you can reach your users wherever they may be, whatever channels they may use. That's why we talk a lot and we spend a lot of time focusing on conversations. And that's why we shipped Bot Framework last year with nine channels.

This year, in Bot Framework, we're introducing three additional channels where we can reach your users -- Cortana Skills, Skype for Business, and Bing. And with Bing, we want to ensure that bot integration and bot discovery are as seamless as possible.

And what I'm showing you here is an experience that is live today, you can go try it out. For those from out of town, I highly recommend this restaurant.

So if you want to know if this restaurant has veggie dishes, so let's ask it. You just ask that question inline. You don't have to dig through apps or dig through Web pages to find out if there are any vegetarian dishes. It's as easy as that.

We have now 12 total channels in Bot Framework and your users may use a bunch of those, or they may use a lot more other apps, platforms, or operating systems. So it's a lot of UX code for you to think of. And we don't want you to have to rewrite it over and over again to have it handle natively in your apps.

Now, that's why today we're introducing Adaptive Cards, which is a new open framework already available on GitHub, so you can go try it out.

What Adaptive Cards does is it allows you to build your JSON once, write code once and then have it displayed over these multiple apps, platforms, or operating systems.

And let me show you actually how it works. This is the sample JSON that I have here. And it uses basic building blocks like text blocks or images or buttons.

And then the framework is the one that allows me to display it natively in things like Microsoft Teams. This is how the same JSON would look like on Skype. And this is that same JSON on Slack with the Bot Framework.

This is Adaptive Cards, also available for you to use today and take it for a spin.

Now, these are all the demos that I've been able to fit today. And there's a lot more that I didn't get to show you, which is why I highly encourage you to go to our sessions later today or to check more information out on Microsoft.com/AI. We can't wait to see how you guys infuse AI into your apps. And with that, I'll give it back to Harry. Thank you. (Applause.)

HARRY SHUM: Thank you, Cornelia.

Adaptive Cards is just the beginning of how we are bringing cross-platform experiences together. Tomorrow, you will hear from Terry Myerson more about that.

All of this AI capability is now available today on Azure.com. In addition to new Cognitive Services, we are also releasing updates to the Bot Framework such as Adaptive Cards you have seen, and new controls and channels for the analytics and payments.

We're also releasing a new service in Azure to train your deep neural nets efficiently. And this supports all your favorite deep learning framework. Whether you like Cognitive Toolkit, TensorFlow, Caffe or others.

So I hope you have seen how Azure plus AI provides a comprehensive platform with everything you need to infuse AI into your applications.

They give you a powerful data platform with integrated AI like we have seen in the demo using Azure Data Lake.

They give you greater computing infrastructure with Azure CPUs and GPUs to train and deploy your algorithms and with FPGAs in the future.

With Cognitive Services, Cloud Services and the great tools like Azure Machine Learning, you can create amazing AI applications.

Let's see a real example from a partner, AirDoc. They have been using AI and Azure to improve healthcare. Let's roll the video.

(Video: AirDoc)

HARRY SHUM: (Applause.) It's really great to see how developers around the world are using AI to improve our world.

At home, we are also using AI to redefine Microsoft. From Windows to Xbox, from Bing to Office, every product in Microsoft is being transformed with AI.

Let me use an example of PowerPoint. And let me invite Yina on stage to show you how PowerPoint is using Cognitive Services to create a new experience. Yina?

YINA ARENAS: Thank you, Harry. (Applause.) Hello, Build.

I'm a proud engineer on the Office team, and artificial intelligence is pervasive in Microsoft Office. You can see it with Researcher and Editor in Word, or maybe Clutter and focused inbox in Outlook.

In Excel, we've got dynamic mapping. And how about Designer and Quick Starter in PowerPoint?

These are all great examples of infusing AI into the experience. And today, I'm going to show you how PowerPoint and AI can help us reach new audiences.

I am pleased to announce and show for the first time the Presentation Translator add-in for PowerPoint. It uses the same custom speech and translation services that Harry just mentioned and that you can use as a developer. Let me show you.

When I launch it, a dialog will allow me to select my language, Spanish, and the language I want to display the subtitles, English. Know that there are more than 60 languages supported.

Before the presentation starts, a few slides will be added to help attendees get transcriptions on their own device and language.

Harry, please join as an attendee on this iPhone Translator app.

HARRY SHUM: Great. All right.

YINA ARENAS: You will be asked to provide your name and the language you want to use.

HARRY SHUM: I look at the bar code, okay, great.

YINA ARENAS: Attendees can also use this in the same language to help meet accessibility needs. Okay, Harry, are you in?

HARRY SHUM: I think so.

YINA ARENAS: Click here, good. Now, I'm going to unmute my side to demonstrate the real-time translation capabilities between the presentation and the attendee devices.

I want you to take a look at the translations at the bottom of the slide.

(Spanish.)

HARRY SHUM: Wow. (Applause.) I can even read here in Chinese.

YINA ARENAS: The PowerPoint Translator add-in custom trains the model based on my voice and my slides using the power of the Custom Speech Service.

And we're also making audience participation easier. The PowerPoint Translator add-in is bi-directional, and I can use it to unmute the audience and let them submit questions and comments into the presentation.

Harry, do you want to say something on your end in Mandarin?

HARRY SHUM: Well, I have to try my rusty Chinese. Okay. (Chinese.)

YINA ARENAS: Let's try again.

HARRY SHUM: (Chinese.) Let me try again. (Chinese.)

YINA ARENAS: There you go. How's that for an impressive demonstration of bringing Office to life with AI? There it is. (Applause.)

HARRY SHUM: Great. Thank you.

YINA ARENAS: To learn more, visit aka.ms/powerpointtranslator. Thank you, Harry.

HARRY SHUM: Thank you, Yina. This is great. You should really the Translator demo and the app, it's really fascinating. We support nine spoken languages and 60 languages in text.

Another key product being infused with AI is Cortana. Satya talked already earlier about what we are trying to achieve with Cortana.

We want Cortana to be your personal intelligent assistant that goes wherever you go -- at work, at home, at the gym, and on the go.

Cortana continues to expand its reach across platforms and devices. In addition to our partnership with Harmon Kardon, today we are excited to announce that two new partners, Intel and HP, have signed on to help develop Cortana-enabled devices.

As developers, you will now have the power to build Cortana skills using the skill kit which utilizes Microsoft Bot Framework. I cannot wait to see the skills you will build for more than 100 million Cortana users.

Finally, let me talk about data. Data is the magic that enables AI. Microsoft offers you an unmatched data graph, including org knowledge from Office 365, Dynamics, and even your devices and the world knowledge in Bing and LinkedIn.

When you combine that Microsoft Data Graph with your own business data and your understanding about your users, you are finally ready to build incredible AI applications.

Let's take a look at how this comes together with one of our partners, Tact.

Tact is a company on a mission to make enterprise software more human friendly by transforming the daily sales experience.

Let me welcome Yina to come onto the stage again to show you what they do, Yina?

YANI ARENAS: Thank you, Harry.

Artificial intelligence is more powerful when it is applied to valuable data through the Microsoft Graph, which I have been working on.

Let's take a look at how Tact is elevating their solution with Microsoft and how easy it is for developers.

Here, I'm going to show you how Tact, who authors a smart assistant for CRM, that helps users interact with line-of-business applications in a more natural way.

In the video, you saw how Tact is adding Cortana support into their solution. They're also working to build in conversational AI on top of the Bot Framework to help connect with Dynamics, Bing, LinkedIn, and Office 365. Let me show you.

Here, Tact has used the Bing Knowledge Graph to identify an account alert and send a proactive message to my account team. Because Tact is connected to Office 365, it can

leverage the Microsoft Graph to find information about any previous interactions with this new executive.

The Microsoft Graph provides a single API end point for accessing the data and intelligence in the Microsoft Cloud.

And it looks like we've done business with her before because Tact, through the Microsoft Graph, found ten references to her across conversations, styles, notes, and meetings. Tact even performs sentiment analysis on the sources using the power of the text analytics service with the data that it gets from the graph and giving me back a nice indication of the previous interactions with her. Just imagine the possibilities.

Now, let's try to find her on LinkedIn. Behind the scenes, Tact will query LinkedIn for people that match my request, and give me the ability to see their profiles and connect with them.

I'm going to ask Tact to send a connection with the right person. Tact is also deeply integrating with Dynamics 365 and has automatically created a lead for Disney executives. The confirmation includes a deep link that I can use without leaving Microsoft Teams. Tact does this by creating a custom tab that aggregates all the valuable data about this account.

Tact can also use the Microsoft Graph to add a reminder to my calendar. I select top collaborators, and the Microsoft Graph will use intelligence to find the people that I collaborate with the most in this particular deal.

If you take a look at my calendar, you can see I'm pretty busy, but Tact, through the Microsoft Graph, will find meeting availability and schedule the reminder.

It's going to take a second. There it goes. Tact and the Microsoft Graph updated my calendar in real time, you can see the appointment right there.

Now, let's take a look at some code. Do you recall that I showed how Tact uses the Microsoft Graph to find conversations where Julia was mentioned, and then perform sentiment analysis on the response?

You can do this with two simple HTTP requests while maintaining the secure and consistent access to your data.

First, we're going to call the Microsoft Graph and the Groups end point to search for conversations where Julia was mentioned.

You can do this in any language. Here I'm using Node.js and TypeScript. Also, you will notice a token, this is how Azure Active Directory secures your data in the Microsoft Cloud.

The response is simple JSON of the matching conversations that I can pass into the text analytics service.

That call is also very simple. All I do is pass in the JSON data and get back the sentiment score.

This is how easy it is to integrate AI with the vast amount of data in Office 365. To learn more about integrating with Office 365 through the Microsoft Graph, join me this afternoon in the Microsoft Graph breakout session and visit graph.microsoft.com, thank you. (Applause.)

HARRY SHUM: Thank you, Yina. I'm very excited about what Tact is doing. They really demonstrate everything about the data that I have been talking about and shows how AI can truly transform your business.

AI is going to disrupt every single business app, whether an industry vertical like banking, retail, or healthcare; or a horizontal business process like sales, marketing or customer support.

Our goal as a company is to help you embrace AI and apply it to your business.

I love this quote from Alan Kay, as I truly believe that the best way to invent the future -- the best way to predict the future is to invent it.

Go to Microsoft.com/AI and get started.

I'd like to close my presentation with a video that I think demonstrates the transformative power of Microsoft AI to amplify human ingenuity.

We are very privileged to have this opportunity to work with all of you. We invite you to get on the AI journey together with us.

Thank you, and enjoy the conference and the Seattle sunshine. (Applause.)

END