[Applause]

[SWANSON:] Thank you guys all for coming. This is great to be here. So, if we want to protect the world around us, we need to know how it works. But as you guys have been learning about today, natural ecosystems are incredibly complex, right? Many different species interact with many other different species in this really complex web of interactions. So you have to study many different species across many different spatial scales for long periods of time. And this is really hard to do.

Camera traps are revolutionizing ecological research by allowing researchers to basically let the animals take pictures of themselves. So cameras are remote, automatic cameras. They're triggered usually by a combination of heat and motion. So when an animal walks by, the camera snaps a picture, and you can deploy these cameras in really remote areas, studying many different species for long periods of time, in places from the East African savanna to the bottom of the globe, to your backyard.

So the thing is, to turn images like these into usable scientific data, we need to know not just when and where they were taken, but what's inside of them. Right? So we need to know what species are there, how many of them there are, what they're doing, and any number of other questions. And when you're talking about hundreds of thousands or even millions of photographs for some of these large-scale camera trapping surveys, that's a lot for a small research team to go through. And despite advances in computer vision research, computers just really aren't good at this type of pattern recognition. Take this image, for example.

Now, I bet if you send this to your 10-year-old cousin, he or she could probably tell you that there are zebras in there. They might even be able to tell you how many zebras there are. If we run this through some of the best freely available computer vision algorithms on the web, they tell us, there, 66% sure that this is an image of a landmark.

[Laughter]

[SWANSON:] And then it's a nature scene-- and there might be a bird or a wildcat or something else in there. Hang gliding is my personal favorite.

[Laughter]

So my group spends a lot of time running pictures like-- Yeah. So that's not that helpful, right? So at Zooniverse, we get people to help do the types of pattern recognition that computers are just bad at. Zooniverse is an online citizen science platform based primarily out of the University of Oxford, and we connect researchers who have too much data with volunteers to help turn that data into usable information. And we have projects that range-- asking people to identify everything from images to videos to sounds, from things as diverse as plankton to penguins, from archaeology to astronomy. So we have a lot of projects. And we have 1.4 million registered volunteers participating in these projects around the world. We have folks in almost every single country. And they're not experts. They're just regular everyday people. Some of you guys--hopefully all of you guys at this point-- have participated on projects like WildCam Gorongosa.

So you might wonder how we can turn volunteer information into expert-quality scientific data. There's a couple of things. First, the interfaces for our projects are designed so that you can make a valuable contribution even if you don't start out knowing what you're looking at. If you don't know what you're seeing on WildCam Gorongosa, for example, you can filter the species options by things like color or whether it's spotted or the sort of general gestalt of the animal. So the interface, the project is designed to guide you to the best possible answer.

On top of that, because we send every image to many different people, it's okay if you make a mistake. You might have noticed that we don't let you say "I don't know" in these photographs. And that might be a little frustrating. But we actually get a lot of information about the image when we don't let you say "I don't know." And so because we send every image to many different people, when you make a mistake, it doesn't mess up our research. In fact, it actually tells us some really important information about how hard that image is.

Take this image, for example. I know it's getting late in the day, but does anybody want to hazard a guess as to what they're looking at? OK, this is an elephant. It's very obviously an elephant. Fortunately, when we look at the answers that we get back-- some of these might be from you-- everybody who saw this image says it's an elephant. There's one of it, and it's standing. That's good. I can look just at these answers and not even at the picture and tell you I am certain there is an elephant in that image.

Not all images are that easy. Take this image. So there is a critter there. It's right there. You can see its eyes shine. It's moving. I have no idea what that animal is, and I spent six years in East Africa studying wildlife. No idea. On the upside, when we look at the answers that we get--

[Laughter]

[SWANSON:] --nobody else has any idea either. They're all over the map. What I can tell you from these answers, though, without ever seeing the picture, is that there is something in that image. It's small. There's one of it. And it's moving. So I'm actually getting quite a bit of information out, even if I don't know what this animal is. So most images are somewhere in between.

Take this image. This is a little trickier. It's got its head down. There are a handful of stripey critters in Gorongosa. When I look at the answers, when we look at the answers, we can see that there is some confusion about what this is. But for the most part, everybody says it's a--well, most people say it's a nyala. Now, here's the thing. I actually don't work in Gorongosa National Park, and so I'm really hoping that this is a nyala. Okay, phew.

So this image is a little trickier. But to turn this into usable information, we need to come up with one single final answer about what we're seeing. So how do we do this? Well, we run what we call an aggregation algorithm. We basically, at a very simple level, we take the species that got the most votes across all the different volunteers, the nyala. We take the median number of individuals reported in the image. Everybody said there was one individual. We take the percentage of classifications that reported every behavior that was there. So two-thirds of people said it was moving because it's blurry. A third said standing. And then there's this thing, the disagreement score, which is especially exciting. This measures the disagreement among all the different answers. It's a metric that ranges from 0 to 1. So images like the elephant have a score of 0. That's 0 disagreement; everybody says exactly the same thing. Images like that mystery critter are 1. Everybody says something different. And as you can see, this nyala falls somewhere in between.

So hold that number. Remember this, because this comes into play later. The thing is, we have these answers. We have this consensus data set. And now, we need to know how good it is. When we compare it to experts-- which we haven't done for WildCam Gorongosa because it's a brand-new project, which is why it still needs your classifications, but we have done for older projects, like Snapshot Serengeti-- we find that the volunteer consensus data set is right 97% of the time. That's incredible. That's actually almost as much-- as often as experts agree amongst themselves for what they're looking at. And there's a lot in this 3%, that 3% of error.

First, not all species are equally easy or difficult. That 3% of error is distributed differently among different species-- rare species are harder than more common ones. But the coolest thing about that 3% error is that we can predict which images fall into that 3% using that disagreement score. We can predict it almost perfectly. We can just look at this disagreement score for images and, based on our comparisons to expert data sets, we find that there is a threshold where images are likely to be right or wrong. So the closer they are to 1, the more likely they are to be wrong--meaning to be incorrectly classified or impossible to resolve altogether; meaning that we can target the expert effort, the effort of the researchers, the small research team, to look just at those images that need an extra set of eyes.

So the point of all this is that we can use these general approaches not just for WildCam Gorongosa but for all types of projects that need volunteers to process image data. And so you've got camera traps that are letting us collect information like never before and citizen science allowing us to turn these images into usable scientific data to make the types of discoveries about ecological processes that Roland is going to talk to you about next.

[Applause]

[KAYS:] Thank you very much. Great. Well, it's really exciting to be up here in front of you and have a chance to talk to you about patterns from camera traps. And this is a great time for this lecture because you've already been hearing a lot about patterns in nature, how we study patterns in nature, why they're important. You've got a great introduction to camera traps. So now, I'm going to put those two together. And of course, camera traps produce all sorts of great images of patterns in nature, like this zebra. And other animals that have beautiful patterns in their coat, like this leopard from China, clouded leopards that you can-- so you can identify different species based on the patterns. Some of them are really crazy and weird-looking and sort of mysterious-looking. Others of them--they're not all just cats like this aardwolf. Or how about this one-- anyone know this animal?

[Crosstalk]

[KAYS:] I heard a lot of mumbling, and you were all wrong. That's a ring-tailed vontsira from Madagascar, really beautiful creatures. Wait, did anyone say that? Raise your hand. No, I didn't think so.

So cameras catch all these beautiful animals with amazing patterns. Usually, that's not the pattern we're talking about, that we actually use for science. Now, the one exception is actually identifying individual animals based on their patterns, which is sometimes possible. Tigers are the most famous example. So if you look up here, you should be able to figure out that in these four pictures, there's two animals that are exactly the same individual, that we can recognize based on the pattern. So that can be useful because when you can recognize individuals, you can really count and get really good density estimates.

But most animals out there don't have these unique patterns in their fur that we can use to identify. And so usually, there's lots of other things that we can do, lots of questions that we can ask about these animals out there by looking for patterns not in their pelage, but in the data that we collect. Because there are a lot of important unanswered questions out there in nature that we have to go after and try to understand. Anybody get this? Over some of your head-- does a bear crap in the woods.

[Laughter]

[KAYS:] Crapping in the woods, all right. So what we're talking about here is not looking necessarily-- the pictures are the data. And the pictures are the fun part, and they're what's exciting, and you get to see what animals are out there. But when it comes time to look for patterns in nature, we're looking at the data. We're turning these pictures into data.

For example, you might have three sites that you sample, and each of these sites detects a certain species on certain days and not other days. And you can look for patterns in these data, and you can ask questions, especially about changes over time and about changes over space and looking to determine what patterns you find in the data and what you think caused those patterns. And these can be done at different scales as well, which you learned earlier.

We talked about earlier, ecologists are often thinking about scale. So for time, we might be talking about what time of day is an animal active. Or we might be talking about how have things changed in Gorongosa over 20 years of ecological change. So these are the types of things we're looking for. And you can see patterns in data sometimes by just making a simple plot. Or you can see the absence of patterns in data, a simple biplot comparing two variables. It looks like there's no pattern. You might have a positive relationship between the two. You might have a negative relationship between the two. And you can do the same thing with camera trap data.

Here's some data from this region from deer. And you can see, if you look at how many houses were within 250 meters of a camera trap and how many pictures of deer we got per day, you can see that there's generally a positive relationship, especially driven by a really low amount down there. And then I didn't plan this, but it looks like, when you look at it with bobcats, you get a different relationship. Is it a threshold? It might be a threshold. If you see right here, any of these sites that have houses nearby, there's no bobcats at any of those sites. This is within parks.

So these are some of the examples of how you can look at patterns in the data of-- so this is where animals live. And look for patterns based on the data that we get from camera traps and relate that to different things in the landscape. Sometimes, the patterns can be very simple. Sometimes, it's easiest to see the patterns by just moving them into two groups, group A and group B. In this case, it's yes and no. Was there hunting allowed at a park or not, and did that impact-- was there a then similar relationship with deer. And you would expect that-- predict probably that, where people are shooting deer, there's going to be less deer. And that's what we found. But this is a way to show that. So it's not all complicated models or even all correlations. You can see that in simple--see these patterns in simple comparisons. Sometimes, this is what pulls out the patterns the best.

Here are some other interesting patterns we've gotten from some of the data that you guys looked at from camera traps. In the eMammal project, I actually don't run very many of the cameras ourselves. We loan them out to volunteers, to school groups, to people who just want to see some pictures of animals, and they run them in their backyard, in the woods, in the state parks. And here's a pattern that we found in Raleigh, North Carolina, where the opossums were a lot more common out in the rural area than in the suburban area. And interestingly, when you start to expand and look at other areas, you see some different patterns. So here now, we have a comparison with D.C. and North Carolina, and you can see it looks like there's a lot more opossum activity in Raleigh than there is in D.C., and not quite the same trend. So, interesting. Why is this? We don't know. This is--we've found a pattern. One hypothesis might be opossums aren't very good at crossing the road, and there's a lot more roads in the suburban areas. So this isn't conclusive information, but it gives us a pattern to start coming up with some hypotheses about what's driving it.

Interestingly, we find exactly the opposite pattern with foxes. Look at red foxes in Washington, D.C. They are much more common in the suburban area than they are out in the rural area. So if you guys live in D.C. or anywhere-- you all live somewhere in this area-- you probably have a red fox that runs through your backyard. And the other interesting thing is we got very few gray foxes. Here's one--a few gray foxes out in rural D.C. When we look at the Raleigh data, it shows the same pattern, but with gray foxes and very few red foxes. Why are there more gray foxes in Raleigh and red foxes up here? I don't know. But it's interesting that both of them are actually finding the suburban habitat to be preferable habitat over the rural habitat. And we don't know exactly why. We could come up with some hypotheses related to a lot of the mechanisms that you heard about earlier.

So you can also look at species abundance. Here is another way to look at patterns across the whole community of animals. This is showing how many pictures we got in eastern North America compared with Gorongosa. And you can see--what's interesting is there's actually a pretty similar pattern up to here. There's one species that's really common, another species that's pretty common, and then it drops off. And so here, it's deer, raccoons, blah, blah, blah, blah, blah. And in Gorongosa, it's baboons, warthogs. So, interesting, look: waterbuck, that we've heard so much about today, is actually the third most commonly photographed species by camera traps. And I think a lot of the surveys were done from the air for the-- so it's hard to count baboons when you're flying over in an airplane, but the cameras get them much better. But what's interesting, I think, is how the difference is in the tail. Look at the long tail of diversity of species that are rare in Gorongosa, but they just have many, many, many more species, which is what we find a lot of times when we look at the difference between tropical systems and temperate systems.

So here's another interesting pattern that we found when we looked at our data from across the region, with cats and coyotes. We broke this up here in this graph into protected areas, which are basically state parks, national parks, and there were some coyotes there and almost no cats. We found almost no cats in any of these protected areas, which is probably a good thing, that the feral cats aren't running around in these protected areas. Backyards, meanwhile, lots of cats--not surprising; that's cats' habitat. But it lets us directly compare that there's 300 times more cat activity in backyards, very few coyotes in backyards. And then in small urban forests, there's a mix of both of these. So, one, this suggests that coyotes-- and you probably know; coyotes are known to eat cats. And cats are known to avoid predators.

So this certainly leads to a hypothesis. There is a correlation. Cats are rare where coyotes are common. Does this mean that coyotes are keeping-- have we shown this definitively, causation that coyotes are keeping cats out? We haven't. This isn't an experiment. We didn't manipulate the experiment. You can maybe start to think of some, if we took a whole bunch of cats and dumped them into a park full of coyotes. Should we try that? Does anyone want to write that protocol and try to get it approved? Or maybe we could take a bunch of coyotes and dump them into people's backyards. Some of these experiments are hard to do.

You heard about some earlier today that are really elegant, great experiments, and some other ones, like the Gorongosa, where we take advantage of natural experiments, natural variation in which animals live where and how that changes over time, to try to understand the mechanisms. And we're actually doing something like this with this question, because coyotes are now everywhere in North America. They never used to be in D.C., by the way. Twenty years ago, there were no coyotes here. They moved here. There's one decent-sized chunk of real estate in the United States that doesn't yet have coyotes, and that's Long Island, New York. And so we're studying the cats there now to see if they're moving in different ways than the cats in other places that do have coyotes.

But there's other types of experiments that you can do. And this is, like, the classic design, the before-after control design. And you saw some of this with the treatment and the control. This is just an example from a student I'm working with in North Carolina, looking at: does trail construction in a state park affect the wildlife? So she's got a great before, before there's a trail. She runs cameras on the trail and off the trail. She ran cameras during the construction. And then now that the trail is in use, she's running cameras as well. Here, you can see a nice picture of a bobcat walking down this brand-new, super-nice constructed trail. And then she also ran cameras the whole time at a control site nearby where there was never any construction. And we're still analyzing the data, so I don't know the answer. But this is just to show you that you can do these proper controls with wildlife with some types of things that we have control over.

But when we're talking about looking at patterns in nature, it's very rarely one thing. It's very rarely just the coyotes or just the trail that's affecting these animals. You have to start incorporating multiple factors at once. And those are some of the models that you heard about earlier today. This is one example of how we found bobcats responded to the number of people hiking on the trails differently in parks that were hunted versus parks that were not hunted, which is pretty cool because parks that aren't hunted, people aren't really a threat to the bobcat; whereas parks where there is hunting, those people have guns sometimes. Those bobcats should see that these people are potentially predators. And it looks like we did find this response.

So this shows you that you have to take multiple things into account, and this is just two for bobcats. You can imagine the habitat, the number of bunnies, what the fragmentation is like, the number of houses nearby. These all could come in together. And so this is where some of the statistical models come, to start pulling out all the different parts of the patterns. And you guys saw some great examples of that earlier today about multivariate, multiple variables at once because we know it's very rarely just one thing. It's not just the rain. It's not just the fire, et cetera.

So I want to briefly talk about temporal patterns, because I mentioned you can also look at temporal patterns. This shows camera trap data for time of day for red foxes and coyotes. You can see here's time of day and what proportion of the pictures we got. Both of these animals are mostly nocturnal. One of the things we found in this region, in the temperate region, sunrise and sunset times change a lot during the year, and so they kind of end up blurring the patterns. What we did was we looked-- well, let's just look at the early morning activity. And we find coyotes really like to be active. They have this peak a little bit before sunrise. And red foxes really like to have a peak at the same time, but not when there's a lot of coyotes around. So if we look at red foxes in areas with high coyote abundance, they avoid that period. And so this is an indication of temporal avoidance, like you heard earlier of how do species coexist. Sometimes, temporal avoidance, they can partition the niche temporally, even as fine as a couple hours around the peak activity when coyotes are out, because foxes do not want to run into a coyote. It'll end bad for the fox.

So, just to kind of wrap up, there's lots of randomness out there in the world. And when we start to see patterns, we know that there must be something underlying them that's causing the patterns. And by searching for these patterns in ecology, whether it's in the patterns of the plants on the ground or the patterns of where the animals are in space and time, we can start to understand what might be causing them, what are the processes that are important, and how does that relate to conservation, and start to answer some of the questions about these awesome animals like coyotes that are out there, but we hardly ever get to see with our own eyes. You get the camera traps out there; it catches them; it doesn't bother them; it records this voucher information, even for small guys-- like squirrels that are out there and bobcats. And we can help learn-- ask questions about the animals that are right in our own backyards that we still don't know what's determining their patterns of abundance.

So that is the end, and I guess we're ready for some questions. Thank you.

[applause]