2016 Conference Transcription

Date	Thursday 31 March 2016
Session Title	Fireside Chat
Session Time	11:45 - 12:45
Speakers	Moritz Stefaner, Ben Still
Notes	n/a

Intro	Hello and welcome to FutureEverything's 2016 Festival Podcast series. Over two days, in Manchester's iconic Town Hall, we tasked designers, artists, scientists, and many more, to rethink our resources, from life, earth and intelligence, to community and uncertainty, our speakers asked what we might need less, and more of, in our near future. One of the highlights of this year's festival is the launch of Project Ukko, a new climate service for wind energy. Project Ukko breaks new ground in the effort to improve the resilience of society to climate variability and change. For this session, we invited Ukko's interface designer, Moritz Stefaner, to join physicist Doctor Ben Still into an intimate conversation on the challenges facing scientific data and the best way to make sense of it.
Moritz Stefaner	Hello, thanks for coming. My name is Moritz Stefaner and this is…
Ben Still	Doctor Ben Still.
Moritz Stefaner	Together we are having a Fireside Chat.
Ben Still	Yes. We had a chat a little while ago on Skype and we found ourselves talking for far longer than we thought so we are going to be telling you a little bit about ourselves and then just generally chatting. I hope you'll enjoy it and join in with us. That's the idea.

Moritz Stefaner	It's true. So the broad theme is uncertainty, predictions, sensing the world through technology. Personally I'm really specialised on data visualisation so I've been working for more than ten years now as a freelance independent designer of data visualisations. I call myself Truth and Beauty Operator just to make sure there's no-one else who has the same job title as me but it's also because I really try to combine the aesthetic and central aspect and the analytic ones in my work. The type of things I do ranges from very clean analytic information design over high-end interactive experiences to data art and I try and combine all these things into interesting experiences. Many of my projects do have an art and a science component to them and I think good data visualisation always plays a bit with both worlds and also stimulates both our analytical thinking but also our senses. If it works out, of course it can be hopefully a magical experience, this potent mixture that comes together. And many of my projects do have also a science background and science and visual arts have an interesting relation for many years now of course, decades and centuries. We had great visual thinkers like Leonardo Da Vinci who used visuals very effectively to communicate existing knowledge or the inventions he made. So visuals can be a strong way to store information. [Inaudible 03:05] came up with these great illustrations of nerve and cell patterns that to me could easily go into a gallery, what he produced there and it's also ground breaking science. And of course we have the [bubble chamber? 03:19] experiments, a very direct imaging technology, like a physical way of imaging and making the invisible visible which is, of course, at the heart of information visualisation. The big revolution has come in the Sixties only and so we are now just for fifty years having this opportunity to actually do exploratory data analysis that means we just load a data set into a computer and see how it looks. Before you had to figure the structure of the patterns out yourself in your mind or somehow play with visuals or arrangements maybe to help you think. But now we can become surprised by data or just see what's there. And I think that's very exciting. A very exciting perspective. [Name? 04:01] was very important in founding this field. Here's a short clip from the [inaudible 04:06] system they built in the late Sixties and early Seventies. This is actually the first time there was an interactive exploration of data on a screen. That wasn't possible before. I'm just very excited about that. I'm a very data positive person. I also see the dark sides and we'll talk about that maybe later towards the end but at the same time I'm really excited. I think data visualisation can provide us with all these new ways to see the world. You can almost see a new way of glasses or goggles that we can construct that help us make sense of reality, see alternate realities. What's exciting about it and especially today that's important is it puts the humans in the loop. We talk a lot about automated decision making, the power of algorithms, how to profile, how to optimise. But only, I think, data visualisation

	puts us in a position to actually make sense of all these complex patterns that we can spot in data and make actual human decisions. I think it's one of the key culture techniques we need to make sense of this complex world. What I'm personally also really interested in is how these views of the world we create with technology, how that changes our world view, how that changes our mind-sets. One classic example is here, for instance, this Hobo-Dyer equal area projection which looks like a wrong way to map the world but physically it's as well motivated as the Mercator Projection or any other projection. This one is actually more accurate. The countries here and the continents, they have the right size. So it's an equal area projection. You can actually see how big Africa is. And if for once you don't put Europe in the centre, you realise it's just a small collection of exotic countries there on the bottom right. This image makes us almost a bit nervous because it shakes our world views a bit and makes us question about how maybe the way we have presented the world and what we think is normal is so much shaped by arbitrary decisions often. These things have huge impacts and I really like this map and think about how the world would have changed if we had used this one maybe. So that's just a brief intro. I will hand it over to Ben to also introduce him and his work and then we'll look a bit more into Project Ukko art science calibration here at FutureEverything [inaudible, 06:42] and I'll show some of the design background. But first, over to Ben.
Ben Still	So my background is I'm a particle physicist. So I'm interested in breaking down nature around us to see what the fundamental things that everything that we see is made from. These are what we call elementary particles. You smash something apart until you can no longer smash it up anymore and essentially you find that our universe is actually quite simple when you break it down. There's actually only twelve building blocks that everything around us can be made from and most of it, ninety nine point nine nine nine percent of the universe, is actually made from just three of those. The beauty of it is we can simplify our world but these things that we break it down to aren't immediately visible in the traditional sense and so we rely on data visualisation to actually be able to image these things and learn more about the fundamentals of our universe through looking through new eyes. And just in case, there's my credentials. Cheesy poster. Sorry. So, how do we see? Traditionally we see from photons bouncing off of stuff and scattering in different directions. These photons, these particles of light, have different energies and they enter our eye and they interact with the cells at the back of our retina and we can build up an image through that. But as I mentioned, these particles are so small that we can't actually see in that traditional sense because these photons are far too small to actually be able to resolve the particles themselves. We can extend a little further by using the particles themselves to actually image other particles but again it breaks down.

So the way we see our world at the very smallest scales is through the interactions of these different building blocks with the world around them and not directly the building blocks themselves.

A nice example of this is something called a cloud chamber. Moritz mentioned the bubble chamber which is just a more sophisticated type. We can see here ionised particles are actually just affecting the atoms around them, ionising those atoms in the air, it's actually alcohol vapour, and turning that alcohol vapour into droplets. Much like you get contrails behind an aeroplane, you get to see these particles and the path that they make through. But we're not seeing the particle directly. We're only seeing the effect of the particle on the world around it.

So data visualisation when you move forward to modern particle physics experiments is extremely important and how we do this is quite interesting.

I've got a little bit here about my research but I'll skip through that and focus on just the data visualisation side of things. Right now you're being bathed by particles all the time. They're raining down upon us constantly. The particles that I'm most interested in are things called neutrinos. If you look at your thumbnail, right now there's two hundred million neutrinos going through your thumbnail every single second. So there's billions going through you every single second, trillions, quadrillions going through you every single day. The interesting thing about this is that they interact so rarely with the stuff around us that it'd be a very rare chance that one would interact in your body in your entire lifetime.

Yet we can image these particles. We can see them on the rare chance they interact and an experiment which is part of an experiment I'm involved with, Super-Kamiokande, was actually able to image the Sun in these neutrino particles. So by sitting patiently and over many, many years taking and adding together the interactions they saw from these neutrino particles, they were actually able to image the sun in a very unique way.
This original picture is actually the Sun in x-rays but for the first time we're able to go beyond that spectrum of light, whether it be visible, x-rays, gamma rays, and actually see the world through new eyes. So it's not just the fact that we're looking at these particles as small constituents of an atom but we're looking at it as actually seeing how our universe is constructed from these different particles.

I've just got a few examples from some of my CERN colleagues. Data visualisation is something that is really important when we're trying to get a grip on the displays of these particles and how they're colliding together, transferring their energy to produce new matter, new particles. So these are just a few from CERN because I think they're the most impressive.

This is the ATLAS experiment. This is actually a candidate Higgs boson that has been created and decayed in the middle and it's spat out four particles called muons. You can see them quite clearly here. It's quite obvious to the trained eye if you look at this data visualisation, that the Higgs boson, this is a candidate event. This is one of those 'God Particles', I don't like that term, but one of those 'God Particles' that gives mass to the rest of the particles around them.

Moritz Stefaner	Hold on. Is that an actual visualisation of data or is that more of a composition, like an illustration? Because it looks like it's from a movie almost, Tron style.
Ben Still	It does but no, that is actually… the detector itself is constructed in CAD and then the position of every single piece of electronics that is fired is detailed in the actual…
Moritz Stefaner	[Inaudible 12:30]. Wow.
Ben Still	Yeah. And so then it's reconstructed and you get to see – these muons are very, very clear because they fly through the entire detector, the onion skins of the detector, right to the outside. They don't really lose that much energy as they pass through. Again, there's lots of different ways in which these different particle interactions are visualised and each of them give a different perspective on the individual interactions. Again, this is another Higgs boson decaying and you can see quite clearly again those four muons spraying out, right to the outside of the detector. What I think is quite impressive is this is from the run where they were smashing heavy ions together, ions of lead, into each other. So rather than having two protons which are fairly small bags of groceries if you like, in terms of particles, these heavy ions are full of groceries. So this is like smashing together two boxes of oranges and seeing the mess that ensues. But still, being able to make sense of this, it helps to be able to actually visualise it because otherwise you haven't got that tangible link to the data sometimes. This enables you to get another perspective on what you're actually chunking through. That's about it really for now.
Moritz Stefaner	Cool. Shall we switch back?
Ben Still	Go for it.
Moritz Stefaner	So that's rocket science almost. It seems like. Here's a project, Project Ukko. I can tell you a bit about the design rationale and how it came about. So like a little case study, how to make complex data sets more accessible. It's a collaboration here between FutureEverything, the Barcelona Supercomputing Centre, the Met Office and a few other partners in the context of an EU funded project, Euporias. The whole project investigates many ways of how climate science can help us in dealing with the changing world and how we

	can build a resilience there. What we were looking at is how to make use of seasonal wind forecasts. Typically wind is very hard to predict and typically you would only have weather forecasts in the range of a couple of days to know what the wind will be like. But there's also seasonal wind forecasts and now our new computing powers allow us to have better methods there, how the wind will be in the next few weeks and months; the upcoming seasons. We took this data set. It's a fairly large data set already, what I was confronted with although that's already a [inaudible 15:17]. So we have around a hundred thousand points around the world, seventy by seventy kilometre cells basically, and we have fifty one different predictions, different versions of a possible wind future that we try to distil into a meaningful forecast that is still digestible. The idea was to bring that to outside the scientific community towards, for instance, energy traders, wind farm managers or anybody who has to rely on having some clue of how the wind will change in the coming seasons. It was a really interesting project to work on. Overall we've been busy with it over a year now and you now see the results of this longer process. When we first started, of course there was first coming up with the idea and the brief which is already a challenge in itself but once that was clear, we started to do user workshops, trying to figure out who would actually be interested in that data and what type of perspective will people need on that data. So it's not just enough to say 'here, you can download the spreadsheet' but we need to find out what's the type of information you would want to draw from information like this and what are you most interested in? What came out of that is that the main interested, for instance for energy traders, is to know which regions will have unexpected changes. So if there's business as usual, you can just rely on what has been the case in the past but what we're of course interested in are anomalies or strong changes in certain regions because there's something you need to react to or something you could even have a benefit from. We brainstormed a few ideas, had some interviews with potential stakeholders and tried to sharpen all these vague ideas of how the data could be used into a useable application. The other side – I always try to do in parallel – is just explore the data. So just take the data as it is and see what we can do with it, what it affords, what the interesting patterns are, what the texture is. For instance, here in the beginning I made these very blurry world maps based on here the skill level, how well the model has performed in the past. It's a very blurry image. It's a bit like something you might encounter in scientific presentations. But then I worked a bit on how can we distil a few patterns that are more interesting or that highlight interesting aspects of the data? So in this case I decided to just take away all the data that has a skill below zero because there the prediction quality is not great and suddenly we gain more space to work with. Not everything is so busy and crammed. We highlight certain spots in the data. You can also make the

decision to highlight the very high and the very low points of the data, like here the red and blue ones, and deemphasise the middle range.

A lot of my work is finding out what the important stuff is and bringing the important stuff to the front, building up these visual hierarchies so that you see something straightaway like a big pattern and then you can still go into all the details. But you're not just bombarded with all the information at once.

I like to say it's not about simplification but clarification. The end product can be still very complex and very nuanced and have a lot of details but it needs to be clearly shaped. It can't just be a random mess of stuff.

The other big challenges, what I also enjoy most doing, is finding the best visual device for a certain data set, for a certain challenge. I'm always convinced there's one best way to do it. I'm sort of searching for that. To me it's often about finding something unique. Originality alone is not enough but it's really good if you have something unique so people can recognise it and people will be intrigued by it and say 'oh what's that? I haven't seen that before.' That can be a good thing. It needs to be expressive, it needs to be able to transfer all of the information we want. It should be intriguing, as I said, and I like elegant solutions as well.

I try to experiment a bit with different visual forms based always on realistic data because if you just take random data, you will lie to yourself or make it look nice with random data and then the real data is totally different [inaudible, quiet, 19:36]. But I prototype quite a bit and try and understand what certain design solutions in this very special context help or whether they're not so helpful.

The unique visual device we came up with in this project is this probability cone. It's a nice, very simple visualisation but it encodes a lot of really important aspects of the data set and it also, as we'll see later, translates quite well to other views. How it works is here on the left hand side we have all the historic wind speeds in one region in the world, that's 1981 and 2014, and you can see they fluctuate quite a bit here. Generally wind speeds tends to fluctuate quite a bit. They're divided into three different sections. Here, the middle, are the white dots and this is the middle third of observations; the typical wind speeds. The yellow ones are the higher wind speeds, the blue ones, the lower ones. We can say a typical wind speed in that region is eight point seven metres per second over the whole season.

This is the past. This is what we know. Now we present what we think the future might hold. Here we have fifty one different dots and we can see here this full spectrum of wind speeds could rise, wind speeds could stay the same or they could go lower. So in this case, the models don't really agree. There's a wide spread of potential future outcomes. In other cases it's clearer.

The nice thing here is this presents a whole visual model of how to think about the data. This idea that the past is set and then there's a certain point here and then there's this cone of future possibilities.

This works well for showing the data as we have seen here in detail but it also works as a map symbol. This situation here on the large maps we've produced is summarised by a single line but it uses the same visual metaphor. The yellow points here, these are the ones that indicate higher wind speeds. There's a lot of them here so they result in a yellow line that points upwards basically. In a way you could say this is a summary visually of this situation here. As you zoom out, because colour and tilt show you the trend, you can still spot patterns even if you zoom out very far.

So this simple design idea for the detail [inaudible 22:03] has informed the design of the whole project and the whole visual language. I always try to find one coherent way of formulating the data in a coherent visual language. The way it worked out here in the end, it's quite a complex multi-dimensional map that encodes a lot of things at once but the basic principle is the thickness of the lines indicate the wind speeds. So here we have high wind speeds and here we have lower ones, thinner lines. The change in wind speed is encoded twice. It's both in the tilt and the colour. So here we have decreases pointing down, blue lines, and increases pointing up. Again, because we used both colour and tilt, we can have these different zoom levels which is quite nice.

Finally, the skill is encoded in the opacity. The skill is roughly how well our predictive model has performed in that region in the past. Different regions have different difficulty levels in terms of prediction and the skill is a measure of how well in the past, if we look back, the same model has predicted the right wind category. Like if it's higher than usual, normal or lower than usual. So this gives us some idea of how much we can actually rely on that prediction and if the skill is below zero, we don't draw any symbols. So basically the blank spots on the map are the ones where we say predictions are difficult. And so also this idea of uncertainty or unknowingness is already part of that display as well. And I think that's interesting also because we have these two levels of uncertainty. So this is the skill, how well has the model performed in the past, and then we have all these different predictions that could also be distributed and half of them might point upwards and downwards. So you see already it's quite a complex affair.

In addition, you can also see how much wind power is installed in the regions and how much wind power is produced. So with this map you can basically look for a region where there's a lot of wind power and predicted more wind speeds and maybe there will be more energy produced in that region, for instance.

So on display here you can try it out. It's on a big touchscreen and you can hop around the whole world and see how the wind will change and you can also try it on the web. The full map is available and you can use it in your browser and play with it.

The artistic counterpoint that's also on display here is a video loop that was produced specifically for the festival, to have a nice immerse ambient and a bit more artistic experience. This uses exactly the same data and the same encoding, the same basic principles but now we're in this strange type of crystal ball type Earth. It's an exotic geographic projection. It's a video loop. You zoom in to the view, it rotates a bit, it's very immersive. Over there it's on display. It

	also has a soundtrack. You can just meditate on the data for a while. Find some rest in this busy world. For me, first of all I wanted to see the data all at once for once because the map, you can show quite a lot but you never see it all and I wanted to see it all. Just a personal urge. I also wanted to show a different way, because there was this ongoing discussion also in the project, is it okay to summarise on the map in this strong way? And so we just show the average values basically in the big overview map when there's fifty one different scenarios and all of them are a bit different. This one actually cycles through all of the individual scenarios. So instead of summarising all of them together, it cycles through each single one and I thought that was really interesting to look at. Here's a close up and a sped up version. You can get a sense of how different these are. In some regions of the world, the lines always stay the same so the predictions are fairly consistent. In others they just go up and down. There's a couple of potential outcomes and you don't really know what's going on. I think this gives you a different, much more anecdotal and intuitive access to the data but it's much harder then to understand the exact distribution. So even if you watch that for five minutes or so, you still don't know. It's hard to put a number then on the probabilities. So I think we need to have both things. This project is also about not just providing one way into this data set and into this problem but many different readings on many different levels. I think these multitudes are a big part of it. So this leads us to the wider questions of how to deal with uncertainty and predictions and how to make sense of the world, right?
Ben Still	Yeah. It's very difficult, to put it lightly. Just quickly, comments on the project. One of the projects I've been involved with which also combined my field of science with art, these are just some plots from my thesis. I was working with multivariate techniques such as neural networks and boosted decision trees to try and extract physics parameters from this multitude of data. And this was a very different way of condensing down and visualising the data. It leads directly to talking about how confident we are about using certain statistical techniques. A lot of these techniques often fall prey to something called overtraining where we train these multivariate analysis to be able to distil down some data and give us an answer either whether it's black or white or whether it's one thing or another. Sometimes we fall prey to becoming biased and biasing our own statistical techniques. So figuring out how to make sure that we're not focusing on one region and open to all of the data in a very different way is important. I showed a lot of these plots to an artist friend of mine and the Institute of Physics asked if we would use these as inspiration for an installation, which was actually three months in the Canal Museum in London. This installation wrapped up many of the conversations we had not only about the shape of the detectors and the technologies but also… and I know you're not a massive fan of the rainbow colour scheme but it seems to be a mainstay in particle physics at least, which is a shame and I think we could move away and find new ways of

	visualising our data. There does seem to be a lot of traditional ways of doing things within the community. But the idea was behind this, it also included an aspect of the way data analysis was done in the 1950s at CERN where armies of secretaries, they were actually called computers, were scanning the bubble chamber photos that you saw earlier, marking the tracks and then passing on the distilled data to scientists who actually analysed. And the artists actually mimicked this by painstakingly handcrafting each one of these discs. There's well over a hundred thousand of those plastic beads and they're all at different levels on those discs to represent the histograms and the data that we are trying to distil down from these experiments. And with the colours changing throughout the piece as well, just showing how we used that extra dimension to try and display data. But I still think within particle physics we could certainly benefit from a project like Project Ukko to be able to visualise some of these things in new ways and certainly look at the more macroscopic picture of what's going on. That's all I have really.
Moritz Stefaner	So we have a couple of loose blocks of content we could present but we're also ready to take questions. Maybe we should do that right now?
Ben Still	Yeah.
Moritz Stefaner	So if we any concrete questions we're also hearing from you. There's a microphone so it will be recorded. It's better if you use the mic.
Audience	You said you aim for clarity and uniqueness in your work. Is it mainly an aesthetic aim you are following or is there a political and advocative target to your work too?
Moritz Stefaner	That's a great question. I think the first [inaudible 31:10] is really that even if you provide a tool, that the tool itself is not neutral or objective in anyways. Often people tend to think because it looks sciencey and it has a rainbow colour pallet, it's an objective measurement of the world. But what's important to me is that there is an author behind that. And I don't think it's a problem, I think it's great. I think it's equality. But you need to acknowledge that you are an author and you're putting out something under your name and with your certain perspective with your own limited view on the world. Generally I think it's good to look a bit behind the curtain and maybe work with data analysis yourself or understand a bit how it can go wrong, to also understand that maybe the design decisions that went into a certain representation of the world. And again I think it's easier to critique when you do clear work. Like if you make a clear statement and it's very well shaped and it's clear what it is, you're

	opening yourself to critique. If you make everything very muddled and blurry and seemingly you pull yourself back and say it's just the data, it's all there. It's harder to critique that but I think it's good to invite critique and make a statement. Like 'this is what I found characteristic for this phenomenon. This is what I found important to talk about. This is what I found worth visualising.' Because then you can start a debate if that's the case or what the shortcomings are or what's missing maybe. But if you pretend you're not there and the data's all there and there's no more problem then the conversation stops at this point.
Audience	But that requires a fair bit of knowledge about what data can do. So you're not aiming at laypeople I take it?
Moritz Stefaner	That has a lot to do with data literacy and understanding how the sausage is made, I think. But maybe that's the first step. But I mean, that's the whole crux. Also this project, Project Ukko, is ultimately about bringing something very complex or very sciencey to wider audiences and having them understand a bit better what's going on. You don't need to understand everything in detail but just get some sense of how is that made, how does that come about, what is the specific nature of that process that generates this data? Things like this.
Ben Still	I think the key thing is the fact that we're honest about uncertainty in data. And that's I think what we've been discussing about a lot as well, the fact that the visualisation actually brings you in and is honest with you about uncertainty in data. That says something a lot about the scientific process behind it and even if someone does go and have a look at the visualisation and just takes that away, that is very important because uncertainty in data is there and it needs to be communicated a lot clearer. I think it does it in a beautiful way with the changing transparency of those bars.
Moritz Stefaner	Here's a great example from the New York Times. I think simulation can be a great way to make a bit clearer what the wriggle room in the data set is or in a prediction. What they do here is they say the draft report has quite a big error margin basically. It's an [untrialed? 34:41] prediction. They say even if the actual [job? 34:43] growth were steady over the next twelve months, the prediction at a given point in time could have all these different shapes just because it's such a noisy prediction. To understand that all these different shapes, there will be big press articles about how the economy is recovering or reclining and all these things. And you take it for granted that this is what the experts say. But if you understand how much variation can be [inaudible 35:10], you will interpret the results differently. If you think wind is hard to predict, try economy because economic forecasts are all over the place. I love this chart because it shows so well how at any given point in time… these are IMF, GBP, [inaudible 35:28] growth forecasts, this is the actual growth, the black line – at any point in time, IMF says 'it's going to continue a while like this but then we're going to rebound.' 'Actually it's going to

	continue a bit further but then things will be good.' And again, and again. So the forecast is always much more optimistic than reality. I think that also shows that not all forecasts are actually neutral but I think these are political instruments as well or you try to effect a certain change in the world by making a certain statement about the future, hoping for a self-fulfilling prophecy or something.
Audience	Just a reflection on what you were saying before. When you were mentioning about the tension there was in the project between presenting all of the information and then summarised information. I was linking this to what you are talking about now about the authorship of the visualisation, the fact that you have [inaudible 36:23] and there are some decisions there. So my question is related to the limitation that the scientists or the sense of ownership that the scientists have of the data they've generated and the control they want to maintain on the data. So going forward, what do you think is that [inaudible 36:40]? There is a tension there between the sense of 'I generated this data. I am the legal or moral owner of this data. I am the only one who really understands this data. I want to keep control of it.' On the other side there is someone who comes with a completely different take on that data which might be beneficial, it can be useful but can be disruptive in a sense. How do you see that sort of tension going forward?
Moritz Stefaner	It is a big tension and probably the central challenge in science communications. I have one slide on there. First of all, one does not simply communicate science. That's an important fact. So there's always this fear of misrepresentation that I encounter as well. That made the rounds two years ago or something. I think every scientist has this fear of this headline that is just horribly wrong and their name is, until the end of the world, [inaudible 37:38] or something else perhaps. I think that's at the very heart of this problem. At one end, how do you stay true to what your research actually states about the world and what you know we can say and scientists have a very intricate vocabulary also to formulate how well they think something works. So for instance there's this whole problem, like when we talk about predictions, what does 'likely' actually mean? Or 'probable'? Or 'there's a very good chance that …'? There are actually scientific studies that have people rate them on a scale from zero percent to a hundred percent 'what do you think does "probable", "likely" or "we believe" mean?' And it's wildly divergent. It really depends a lot on where everybody's coming from, what they think to be a plausible future or not. I think the way out of this could be just to accept that there's not one single right map of the world and there's not just one single right reading of a data set but it's a debate. Everything's a debate. You can put different poles in the ground and then see how they relate to each other. Then maybe you take out one of these poles again and put some strings around them and suddenly something appears. Again I think in Project Ukko we tried to have these parallel tracks and different levels of explanations. I think the hope is that each combination of artefacts people encounter make sense together and they draw something from that. But

	maybe we need to give up this, as you say, exclusive ownership and exclusive rights to interpret the data to some degree and hopefully gain something from that.
Ben Still	I entirely agree. I think certain particle physics communities and most communities I think within science are going more towards open source data. In particle physics, one of the main difficulties of open sourcing data in a lot of scientific communities is that it's petabytes of data and how do you make that accessible to people? So at the moment a lot of it is being boiled down and it is being filtered a little. That isn't simply because we don't feel that the public will understand that data, it's more that we want to give a clean data set that is something that people can work with. And it's actually the data set that PhD researchers do work with and that ownership is important to us because of the amount of work and effort that people have put in to creating these detectors and creating all of the underlying framework which actually extracts the data from the machines. I think it's only right that they have exclusivity to that data to begin with but then getting it out into the wider community, into everyone's hands can only be a good thing. It's the whole thing about widening discussing and participation in these subject areas. If you wanted to build a football team, you wouldn't literally go to one village and build your football team from whoever was the best players in that village. You'd open your net out and you'd try and find the best players from all of the different regions and then build your football team with that. So opening it up to people who otherwise wouldn't have thought about playing with data from the Met Office or playing with data from particle physics might actually just think 'I'll give this a go.' And that's really important. I think the whole idea of that and citizen science around it will help us find new talent and new ways and perspectives of looking at these things. And that can only be a good thing.
Audience	The statement that you made that the forecast is so much more interesting than the reality, I've just come from a workshop which is caring and sharing and building community. One of the things that we've been talking about is government and restriction and making that different from governance. Basically what we were saying is the 'sod it principle' – sod it and do it anyhow. In terms of my training, I'm not a statistician but I think I can call myself a management scientist and I look at some stats and I go 'that is b***ocks.' And I don't know how it's b***ocks but what I'm getting at is what I'm picking up that you're talking about, the importance of intuitive presence as a means of disrupting constructively to make something that wouldn't happen if we remained attached to what we believe is the status quo.
Moritz Stefaner	Yeah. This question of intuitive judgements is also one we discuss quite a bit because it's at the very heart of both the problem and probably the solution. Because on the one hand we are really notoriously bad in making probabilistic, quick judgements. In many areas. In other areas we are great. We can play tennis which requires advanced probabilistic, really high speed and really precisely but just making quick risk statements or condition of probabilities, we're really bad at that. Sometimes when you read studies there, you wonder how we've been able to survive so long.

Ben Still	This is one of the things. It's a throwback from evolution. I remember reading it's the whole thing of the caveman in a cave hears a rustling outside. Obviously the chances are it's probably the wind but paranoia tells you it could be a bear. So the paranoid individual that goes to check will be safer and will live longer and reproduce more than the person who is, I guess you would say, more rational, a ninety nine percent chance it's just the wind blowing through rustling the bush. But on the odd occasion that person's wrong, they'll get eaten. So evolution has given us this quick way of thinking which has been short, sharp, reactionary decisions based on data. But I think it's quite interesting the way that our brain puts things into patterns to help us but when you're reading data in articles and you're reading data in the press, I think we need to be a lot more honest with the uncertainty all the time rather than it just being quoted as, for instance, again I apologise for going back to particle physics, but the discovery of the Higgs boson. 'Oh it's been discovered. That's it.' Whereas the discovery was announced when the certainty was pretty high but still wasn't, for some people, a hundred percent certain of what type of new particle this exactly was. But yet it was 'discovery'. That was it. It was just black or white. It was either it was there or it's not. Whereas the discussion… I trust data much more if people put error bars on it. If there's no error bars on the chart, I'm sceptical straightaway and I will not trust that data. But if we're open and honest, again, it's about this discussion of uncertainty and making sure everyone understands and accepts that anything comes with caveats. Life is not one thing or another. Actually, can we play the Count Beep?
Moritz Stefaner	Yeah.
Ben Still	We've got a great demonstration of how your mind places things into things. I don't know if any of you have seen this before but it's just how our mind will jump to conclusions and fill in gaps and come up with conclusions from information that is incorrect. This is perfectly clean for all ages.

[Video Plays 46:03 to 47:00].

Ben Still	It carries on. And so I think there's a little bit of that. Because of evolution, we have this fast thinking which immediately when we're presented with data we sometimes extrapolate from it. And then there's the slow thinking which we need to get involved with a lot more and that involves the discussion of uncertainty. It think that's where confidence should be.
Moritz Stefaner	Yeah, that's automatic pattern matching. It's a curse and a blessing and we always need to have systems to check on these, both the scientific world and the quick judgements.

	If we switch back, so speculation is an interesting topic anyway and it's also one I've been investigating in a project called False Positive last year. If you're interested in that, you might want to check that out. The basic idea there is to investigate what happens when classification systems start to speculate. So we rely a lot on automatic pattern matching and analysis and the theory we were investigating, Mark Shepherd and I with a few others, what actually happens when that goes beyond what's reasonable. So last year, I don't know if you had seen that, there were a lot of these artworks created basically forcing artificial neural networks to over-interpret images and get into these feedback loops of seeing things that were not present. It's basically like an artificial network hallucinating. I think this is something we will encounter quite a bit in the future, that they make up stuff based on incomplete information. This is what we do. This is what they will do.
Ben Still	And again, my boring plots earlier were my attempt during my thesis to make sure that this wasn't happening, that my neural networks weren't dreaming of physics parameters that I was trying to measure. Because if it's overtraining and it's thinking it's going to see a dog then it's going to dream that dog because it's been trained to find them. And there's some really great ones on Google.
Moritz Stefaner	We had all these really bad situations where young kids would bring clocks to school and because the policeman has been trained or the teacher, by being exposed to a lot of TV I guess, they match the wrong patterns. So they're like 'there's this ethnic looking person with some wires. That must be a bomb.' And that's again a totally overshooting classification. Google has had a really bad moment there with Google Photos where African Americans were tagged as gorillas. These all can have really real consequences of course, very hurtful. And in this project, this was the basic phenomenon we were looking into and to investigate that and talk about that, we invented a company. It's a new cell phone provider, a hip new company for your phone data plans and so on. It's called Candygram. It had a presence then, a different… it could have been the future. Everything was at Ars Electronica and [inaudible 50:06] in Brussels. So we had booths there and people could come up and see what's going on. We had very motivated assistants handing out candy as well to draw people in. Inside the booths there was a little fake cell tower so we could actually pretend to be an actual telecommunications provider. We'd send people cell phone messages like 'welcome' and 'here's a free data plan. Just register.' They were already drawn in this strange conversation with the machine, the system. What then happened in the background is once they'd provided their email address which we asked for, we would search the web for data related to them and create these automatic profiles. Which is what a lot of companies actually do. They try to find out how much money do these people have? Are they good to have as a customer? How long should they be in the hotline queue or not? So we tried to simulate that process and tried to pull together a bit of information about the people. It all happens in the cloud of course.

	Then later, when they came to the [tent? 51:16], the idea was to have a personal consultation so we could recommend a good plan for them. They would be confronted with these profiles we'd created from them. They begin very simple. 'What's your employment history? Where do you live? What are your interests?' And then towards the presentation it starts to become more speculative. Then we'd start to combine different words that seemed to be characteristic for them. In this case it actually was quite plausible, he's actually an information designer. In other cases, it's more obscure what's coming out. But for all of them there was one or two combinations of terms that was really striking and like 'how do we know that?' That's again this bias that we look for, these interesting things. They were ignoring that four of them were wrong or totally bogus and one was interesting. Then it gets more and more speculative and then you see this movie of how the network dreams basically about the person, of what the images are that come up, the terms and so on. It ends with a personality profile. So we have five different personality traits and we just give people fairly random scores actually but they always react very strongly to that. Like 'can you do that? Is that legal? Is that okay?' This of course the baseline. The question we want to raise, 'what was actually being speculated about you out there by companies?' So we made a little tour of that project and then we retired and announced that we were acquired by a joint venture between Google and GCHQ and we're now ceasing operations and it's been such a wonderful ride together. There's documentation also online if you're interested in that. I think it's also an interesting for of data visualisation, this sort of performance and [inaudible, 53:14] just not making a chart but actually creating a whole situation of where people get confronted with data and the effects there.
Audience	Did you let them know that it was fake [inaudible 53:23]?
Moritz Stefaner	Yeah. In the end they realised at some point. It became obvious. But also at the end we gave them a bit of material of how to protect your privacy online. There were also workshops and we ran crypto parties so they could sign up for a crypto party afterwards. We did not try to just depress people and not offer a solution but be a bit positive towards the end. That was sort of a diversion but I think it's important. Other questions or general comments on the situation?
Audience	I was just thinking from a history of science perspective, what you're saying about bringing objectivity back is quite interesting considering that with the advent of photography, when scientific drawings and things went from illustrations, that was their fear. So 'if we use photography then it removes the objectivity.' And then much later on people start going 'well actually, you've edited anyway so …'

	But I was thinking what you were saying about people trusting data, I don't know how much of that is that hang up of that was the moment where people were like 'this is true because it's documented in this specific way.' But then when you present people with their personal data, their immediate thing is 'well that's not true. That's not me.' So I'm wondering, with that kind of objectivity question in mind, information that's presented as metaphor in that [inaudible 55:11] piece and then the information as data and how people interact with this interpretation of science as metaphor and interpretation of science as something solid that they're almost conditioned to believe has to be true.
Ben Still	I also teach at a school. I teach GCSE physics and I encourage students to always question what I'm saying, never take anything as writ, always go and double fact check because I think as we've said before, scepticism and not taking things and believing in things is very important with science. So when it comes to actually talking about new scientific discoveries or new scientific studies, I think there is an onus on the science community to be able to explain and make sure that the media understands the importance of actually communicating the uncertainties on each piece of information that is talked about. A lot of people that I know in the academic world will say 'the media have taken this out of all proportion' but at the same time, have you been able to talk it through with them in detail enough that you've got across how important the uncertainty is. Or have you been so passionate about the discovery itself that you've also made the selling line of the actual piece? So I think that's an important part of how we deal with it.
Moritz Stefaner	It's a big challenge. Nobody wants to write a story that has a lot of asterisks and this long caption below each figure but I agree with Ben that probably also from the science side you need to offer something, some model or some way of thinking about the results that stays true to reality but is also compact enough that somebody will want to report on it. But it's a big challenge. The other part is really the DIY aspect. I think you get the best sense of what data analysis can or cannot do if you try it out yourself. And it can be a really lo-fi experiment. Like, maybe tomorrow you will hear about your data potentially. So it's a very hand-made approach to collecting data and representing data. I think then you only realise that it's not even clear what to measure or how to count or what are the main categories I'm looking at? All these little decisions that go into data representation even before the visuals. I think you only get a sense of that when you actually do it. So I think that's the best.
Ben Still	Just one last thing, in terms of the metaphor, metaphors are by their nature imperfect and so they break down. But the key thing I think to get across to everyone and all audiences is that science is a metaphor for nature. Science breaks down. Science has a horizon. Science has a point at which we're uncertain about it. I'm not sure where it's come from but there's inherent belief that science has the answer when we don't and the exciting thing about science and the most exciting thing for the next generation will be the fact that there are

	unanswered questions and there are things that they need to find out for themselves. So what we have now is still an imperfect metaphor for nature but you could find a more perfect metaphor by doing science and you can find out another way. It's written in mathematical language most of the time but it's still getting that across. So I think introducing metaphors with being again honest about where that metaphor breaks down compared to the other scientific models we have is again quite important in communicating science.
Moritz Stefaner	That's a beautiful ending. It's a quarter to one so thanks all for coming. Check out the project. It's down the hall, check out the project and see you around.
Outro	We hope you enjoyed this Fireside Chat and thanks for listening. You can hear the rest of the talks from 2016 at futureverything.org/2016podcasts.

[Transcription ends]