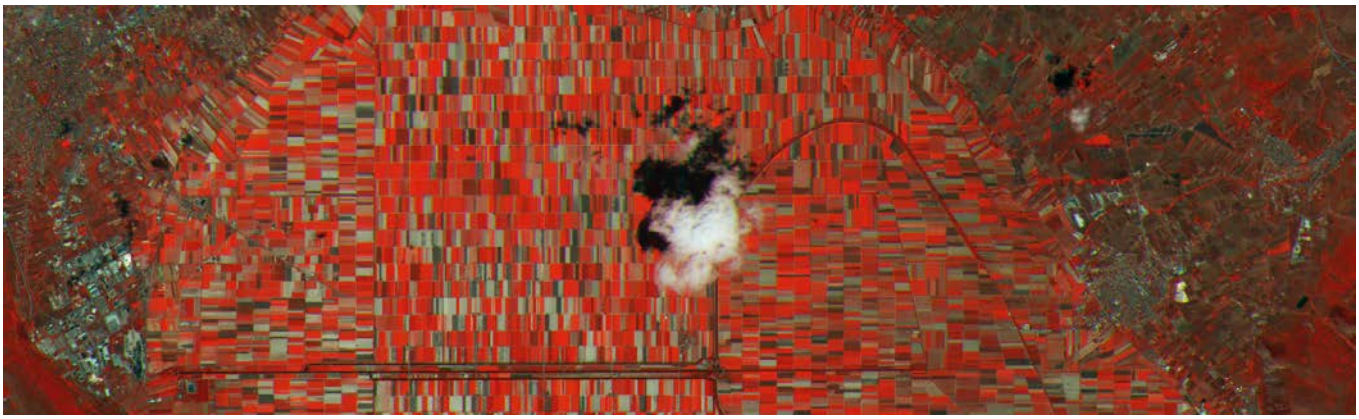
The Cambridge Forum for Sustainability and the Environment

Meeting 2: 19th November 2015 in Downing College



Aims

Our topic this year is 'Risk, resilience and response'. As in previous years, a panel of three expert 'witnesses' from across and beyond Cambridge will join us each month. They will be a rich mixture of policy- and decision-makers from governments and businesses, and technical experts and each of them will provide their perspective on the gaps in our knowledge and burning questions for future research.

This is the second in a series of three meetings, and the aim of it is to use examples provided by the witnesses to explore ways in which overlaying big data sets and remote sensing can assess and communicate risk and resilience in food supplies and changes in biodiversity.

We are jointly hosting these with the Joint Research Centre (JRC) of the European Commission and we are hoping to co-produce a brief white paper for EC policy makers identifying future research gaps and opportunities in this area.

Agenda

All the witnesses will give a 10 minute introduction and their perspective on the two core questions followed a general discussion:

5:00pm	Welcome by the Chair
	Each witness gives a short introduction and thoughts about the questions (10 mins)
	Questions and beginning the open discussion
6:00pm	Coffee break
	Continue the discussion
7:15pm	Reception and dinner, which will include a working session

Witnesses

This month, the three witnesses are:

Dr Francois Kayitakire	Food Security Assessment Team in the Monitoring Agricultural Resources Unit (MARS), Institute for the Environment and Sustainability (IES), Joint Research Centre (JRC) of the European Commission (based in Ispra, Italy)
Dr Matthew Smith	Scientist in the Computational Science Lab at Microsoft Research
Craig Mills	CEO of Vizzuality

Questions

The witnesses have all been asked about what they perceive as being the main gaps in our knowledge and what they would include in the 'next generation' of research questions. This month, we are asking everyone to choose a particular example from a project or a challenge that sparks their interest and then to give a bird's eye view of:

1) The problem
2) What we know already
3) What we need to know

These ideas then act as a springboard for a round table discussion until 7pm which continues over dinner.

UNIVERSITY OF CAMBRIDGE

The Cambridge Forum for Sustainability and the Environment

Meeting 2: 19th November 2015 in Downing College

Witness profiles

Dr Francois Kayitakire

A senior scientist in the Food Security Assessment Team in the Monitoring Agricultural Resources Unit (MARS), Institute for the Environment and Sustainability, Joint Research Centre (JRC) of the European Commission (based in Ispra, Italy)

Francois leads a team working on resilience and on food and nutrition security assessment within the Food Security (FOODSEC) Group. His current activities focus on resilience for food and nutrition security, in particular the resilience measurement issues, food security assessment and classification methods and on agricultural risk management in developing countries. His team provides early warning on food security crisis using various data types and in particular satellite imagery and meteorological data, and they conduct research on modeling food security indicators. His area of interest is mainly Africa but also other developing countries.

Francois' first assignment at the European Commission was within the Unit for Global Security and Crisis Management at the JRC. His work focused on building pieces of an armed conflict early-warning system and understanding their root causes. He worked also on monitoring natural resources that are susceptible to fuel armed conflicts, and on the use of satellite imagery to support post-disaster needs assessments.

Francois holds a Ph.D. degree in Agricultural Sciences received in 2006 from the Université catholique de Louvain (UCL), Belgium. While working as researcher at the UCL, from 1998 to 2004, he focused on forest mapping and urban green area management using satellite imagery.

e-mail: francois.kayitakire@jrc.ec.europa.eu

Dr Matthew Smith

Scientist in the Computational Science Lab at Microsoft Research

Matthew works in the Computational Science Lab at Microsoft Research, and is committed to improving societies (people, businesses, governments) abilities to predict geotemporal phenomena (properties and processes that can be associated with geographical space and time). He has worked in both theoretical and applied ecological science since he left high-school and has come to realise the enormous untapped value in predictive models of ecological and environmental systems and aims to unleash that potential on the world. In recent years he has also discovered so many other geotemporal phenomena that we can predict, anticipate and make decisions about much better than we have done to date, especially in the domains of agriculture, utilities and energy, to name some major business sectors.

He is currently working on some research projects with UK companies to investigate the value of predictive models of geotemporal phenomena to their businesses. While doing that, he maintains research interests in predicting crop dynamics, carbon and vegetation, human responses to climate change, and ecosystem structure and function.

e-mail: mattsmi@microsoft.com

Craig Mills

CEO of Vizzuality

Vizzuality is a science and technology company focused on data visualization, web-GIS and tool development and committed to working on projects related to conservation, the environment and sustainable development. As CEO, Craig is responsible for figuring out what problems they should be trying to solve, guiding the company towards important world improving projects and working with NGOs to help them tell their stories.

Before joining Vizzuality, he spent 7 years working with the UNEP World Conservation Monitoring Centre. As Head of Infomatics, he led a team of software engineers, geospatial developers, designers and product managers to help scientists and government/non-government organisations share their knowledge and data on the web and reduce the time it takes to get from raw environmental data to useful knowledge for policy decisions. These projects ranged from applications mapping every national park on the planet, ipad apps for scientists to monitor mangrove forests and tools for governments to monitor the trade of endangered species. Previously, he spent 7 years with the UK government at the Centre for Environment, Fisheries and Aquaculture Science (Cefas) undertaking marine geospatial research.

e-mail: craig.mills@vizzuality.com

UNIVERSITY OF CAMBRIDGE

The Cambridge Forum for Sustainability and the Environment

Meeting 2: 19th November 2015 in Downing College

Expert Witnesses

Dr Francois Kayitakire (FK)
Dr Matthew Smith (MS)
Craig Mills (CM)

Forum Members

Alan O'Neill (AO)
Alison Smith (AS)
David Coomes (DC)
Emily Shuckburgh (ES)
Graeme Buchanan (GB)
Gregoire Dubois (GD)
Helen Curry (HC)
Hildegard Diemberger (HD)
Ian Leslie (IL)
Keith Richards (KR)
Koen Steemers (KS)
Kristen MacAskill (KM)
Marla Fuchs (MF)
Martin Rees (MR)
Martin Roberts (MR2)
Neil Burgess (NB)
Pamela Kennedy (PK)
Paul Linden (PL)
Rosamunde Almond (RA)
Rosemary Ostfeld (RO)
Shailaja Fennell (SF)
Stephen Peedell (SP)
Susan Owens (SO)
Therese Rudebeck (TR)

Transcript

Duration: **2:02:00**

MR:	Good afternoon everyone and welcome to this meeting of the forum. We have some sort of core participants but quite a few people who are new and so I think it might be a good idea if we go around the table and everyone just say a sentence about what they do and where they're from. But before that I'd like to welcome our three witnesses.
	So I'll start off, I'm Martin Rees, I'm a space scientist and I'm the Chairman of this group.
PL:	I'm Paul Linden and I'm the Director of the forum and my day job is an applied mathematician.
RA:	Hi, name is Roz, I think you all know who I am because I've e-mailed you a lot, but my background is in conservation biology and I'm based out in the Department of Applied Mathematics with Paul helping to bring all of this together.
NB:	Neil Burgess, Head of Science at UNEP-WCMC and part-time professor of conservation biology in the University of Copenhagen and I guess it's my first time and I'm interested in the practical application of science to conservation.
DC:	David Coomes from Plant Sciences here in Cambridge and I work on how forests around the world are changing.
IL:	I'm Ian Leslie, I'm a computer scientist here in Cambridge and have had fingers in various pies including conservation and sensing of the environment.
KM:	I'm Kristen MacAskill, I'm a civil engineer based in the Engineering Department in the Centre for Sustainable Development.
GD:	I'm Gregoire Dubois working at the Joint Research Centre of the European Commission based in Ispra, Italy and managing a small team of researchers and developers focusing on conservation.
HC:	I'm Helen Curry and I'm a lecturer in the Department of History and Philosophy of Science here. I look at the history of the life and environmental sciences in the 20th and 21st century.
AO:	I'm Alan O'Neill, I'm a visiting professor at the Cavendish Laboratory here in Cambridge,

	I'm a climate scientist who likes using satellite data.
MF:	I'm Marla Fuchs, I am the ESRC impact coordinator and confusingly I'm a mechanical engineer.
MS:	I'm Matthew Smith, I'm a computational ecologist based at Microsoft Research just up the road.
FK:	My name is Francois Kayitakire, I work on food security in the Joint Research Centre of the European Commission.
CM:	My name is Craig Mills, I'm the CEO of Vizzuality which is a data design company based in Cambridge and Madrid and we spend most of our time trying to figure out how to communicate large volumes of satellite data into interesting ways.
KR:	I'm Keith Richards, I'm from the Geography Department, my sort of main areas of interest are hydrology and water resources.
TR:	My name is Therese Rudebeck and I'm a PhD student looking at private sector involvement in water governance.
KS:	I'm Koen Steemers, I work in sustainable design in the Department of Architecture and I lead a group on behaviour and building performance.
SO:	I'm Susan Owens from the Department of Geography, I work on science and politics and the role of evidence in policy-making.
RO:	Hi, I'm Rosemary Ostfeld, I'm a PhD student in the Land Economy Department and I'm researching the impact of the Roundtable on Sustainable Palm Oil.
PK:	Hello, good evening, I'm Pam Kennedy. I'm also visiting from the Joint Research Centre of the European Commission. I have a background in use of Earth observation data and agriculture and forestry and my organisation and I am personally interested in how we can inject science into policy-making.
MR2:	Martin Roberts, I'm the Director of the Institute for Sustainability Leadership and I work with global businesses on natural capital.
AS:	I'm Alison Smith from the Department of Plant Sciences, I'm a biochemist but I'm interested in using algae as sustainable feedstocks for chemicals and bioenergy.
HD:	I'm Hildegard Diemberger, I am a social anthropologist working with [inaudible 0:10:53] and I'm also the Director of the Mongolia and Inner Asia Studies Unit and I have general interest in changing climate in especially vulnerable environments.
GB:	I'm Graeme Buchanan from the Royal Society for the Protection of Birds, the conservation NGO, I work with satellite technology around spatial data, I'm interested in applied science for conservation for decision-making and policy.
SP:	Hello I'm Steve Peedell from the Joint Research Centre. I'm a specialist in geospatial technology and I'm looking particularly at biodiversity in protected area management in Africa, Caribbean and Pacific countries.
MR:	Thank you very much. Well for those of you who are new I'll remind you the format which is that we have our three expert guests and witnesses and they'll each speak for about 10 minutes and then we'll have some discussion and we have a short tea break at about six o'clock and then for those who are staying for dinner we'll perhaps have a bit of informal discussion maybe over the dessert. So there will be plenty of chance for informal discussions.
	I'd like to especially welcome our friends from JRC who have been here yesterday and today and I hope it will lead to continued interaction with the forum and of course to welcome the three people who have agreed to lead our discussion today. I didn't check the order in which they would speak but if you are happy to speak with the order on the

	paper with Francois first. Is that all right or is there a different order you prefer?
FK:	No, it's okay.
MR:	So Francois, there's more details about the speakers on the back. So please go ahead.
FK:	Thank you Chair. It's a great pleasure and honour for me to talk to you tonight. In the last summer I visited the other Cambridge, the Harvard University in the United States with my kids trying to assimilate them for the future. Because I started…I did all my studies in Rwanda and then I did my university studies in Belgium, but I think Cambridge [inaudible 0:06:19] big university have been always a dream for us, but now I think it's a kind of realisation of that dream to be in Cambridge talking also to you and I hope that we will have now these nice discussions tonight.
	For the start of the discussion I think I will start now telling you what we do in the JRC in the food security area. In my team we mainly work on food security assessment and this is done to advise the European Commission policies for food aid and also development cooperations. The commitment of the European Commission is to help developing countries to eradicate hunger, that was the first Millennium Development Goal which is now becoming the second Sustainable Development Goal. The European Commission is one of the largest donors, I think it is the first donor in that area, in agriculture and food security. When they have to decide through interventions they want to have a scientific-based evidence on the interventions. When we do food security assessment we use different tools, mainly we use satellite imagery to assess the crop conditions and sometimes also to estimate production. That's how we can have an idea of food availability.
	But having food is not enough. This food must also be accessible for people, the household must have the means to access the food, either if they are a household that are getting food from the markets they must have sufficient income but also compared to the prices. So we look at also the evolution of prices to see if the household are in the position to get access to food.
	There are also the physical access. Sometimes they might have conflicts that broke the food flow and all the trade flows and put some areas at risk of food crisis. As an example now is countries like Yemen or South Sudan where you have some areas that are completely isolated and at risk of famine in South Sudan in some small areas and in a country like Yemen that was used [inaudible 0:08:43] enough food to be relatively sufficient, but now because of the conflict they have a big food crisis.
	Then when you have food it is important that also this food can be utilised correctly, which means that people also need to have clean water to prepare food, but also clean water to avoid waterborne diseases like diarrhoea, that will actually impact the food utilisation by the body. Because when you have food but the sanitation and hygiene conditions are not good, we have realised that then the nutritional status of the population is really bad. In some areas there is enough food but still the prevalence of [inaudible 0:09:31] which is…I don't know if you…maybe you are not aware of or familiar with that, the prevalence of malnutrition is quite high because of hygienic conditions.
	But that aspect means that we have to look at the food security with different angles and with an interdisciplinary approach including socio-economic access, the agriculture which is the main provider of source of food, [inaudible 0:10:02] and hygienic food consumption and dietary habits.
	We have a team that try to analyse all those conditions in different countries and every year we also look at globally to all countries that are susceptible to have food crisis and we come up with an estimate of the number of people in food need and then the Commission can take the decision on how to allocate the money to support those people.
	In terms of research also the interventions there are immediate interventions, immediate

	interventions that are decided and then in that case it's about saving lives. But now we are trying also to move [inaudible 0:10:53] to building resilience to food crisis which means that we try to make sure there is a continuum from humanitarian interventions up to long-term interventions for development cooperations and these long-term interventions must try to look at all the angles of resilience. Which means we have to take into account risk management, we have to take into account the response mechanism and those response mechanism for the short-term can be either in forms of food aid or in terms of income support, which then take us into the area of social protection. Those are new, emerging areas of development cooperation that we are not used to…it was not used to be the case in the 90s, these are new areas because social protection was mainly thought in terms of pensions in developed countries but now we have realised that it can be more productive, more interesting to have this safety net mechanism instead of providing food aid because the money injected in the system for this social protection system can be also used to invest in productive means like in agriculture.
	That's in the area of resilience. That we have for the moment two main areas of intervention, we have the West Africa, I mean Sahel, so the Sahel is the area which is the transition between the Sahara Desert and the coastal areas which are rain and that have a tropical equatorial climate. So the Sahel is very vulnerable to drought risk and the other area which is very vulnerable is the Horn of Africa, all the areas including Ethiopia, Sudan, Kenya, Somalia, Djibouti.
	It means that in those areas the main threat is the drought and with the climate change now there is a risk of more drought even, it should be actually increased which means that we have now to come up with policies that will help people to resist and to adapt to those risk by climate change.
	There are other areas that have other risk based on natural hazard like cyclones. Usually we approach these areas with colleagues that are working on disaster risk management and mainly we will be assessing crop losses after a natural event like a cyclone. Those areas are mainly in Central America, countries like Haiti which is very vulnerable because it's poor, but sometimes exposed to these kind of natural hazards.
	In terms of when we have to advise about policies we have also taken into account the conflict about resources. Because we have limited resources in terms of land, water, but also limited resources in terms of inputs and agricultural fertiliser because it can be limited either by [inaudible 0:14:34] but also by the capacity of the farmers to access, to buy the fertilisers. In terms of land and water what we do for the moment is mainly about mapping the cropland areas because what we have discovered is that statistics coming from FAO, the United Nations Food and Agriculture Organisation, but those statistics are not really reliable and…those are statistics. And on the other side also we have to know where land is, where cropland is. There are many land cover maps but usually land cover maps that are dealing with all the kind of land cover types and not accurate enough for the cropland which is interesting for us, either for calculating production but also seeing the areas where there is a possibility to increase and invest for food production. So updating the cropland is also an area where we think that it's important and now with all the satellite imagery available and the capacity of processing large dataset we think that can be possible to have regular update of cropland maps and if possible mapping also the crop types at a regular basis.
	That is going to be a challenge for [inaudible 0:16:11] sensing people, I think my two colleagues maybe knows this challenge about getting the crop information from satellite imagery.
	Then this will bring us into the big data discussion that maybe we will discuss later on.
	That's what we do. But now in terms of what I would [inaudible 0:16:36] as a debate for you, I was talking about resilience but how to build resilience to food crisis. There are

	several areas of interventions and I think of the people here around the table are from different disciplines and I think it can be good to have some ideas on how can build resilient food systems. Think in terms of food systems in developed countries but also in developing countries.
	I think I will stop here and then the final discussion will come later.
MR:	Thank you very much. We will have the main discussion after we've heard all three but is there any particular question for Francois or comment from other JRC people?
KR:	I was particularly interested in your remark about the land cover aspect and I think that kind of information gathering. Now we ought to be with the technology and the data handling capabilities just about able to do this in a very systematic way, I would have imagined. I'm wondering what initiatives you are aware of to kind of bring this about in a systematic way, so we can get for example climate change resilient resilience through such information gathering. Are you aware of significant initiatives that are trying to set about this challenging task?
FK:	Yeah. I think the most known now in Europe is the Copernicus programme that is developing land services and those land services we have some specific variable for monitoring around ecosystems, vegetation, water and oceans, but yeah, land also, water and vegetation. But I think ESA also has a programme on land cover mapping. There are a number of initiatives on land cover mapping but which are one shot exercise. I think the Chinese are developing…produce a land cover map with some problems. The University of Maryland produced a forest cover map with the change and this forest cover change map and also the water mapping exercise that JRC is doing actually is shifting the paradigm of land cover mapping from the traditional general land cover mapping, but to now mapping one specific class, one specific type of land cover and then at the end, depending on the use and the interest you can bring together the different layers.
	I think this is going to be the new paradigm in terms of land cover mapping.
KR:	But I think it needs to be highly targeted. What concerns me about some of the Copernicus services with respect to them is they can be quite generic and I think in some sense there's got to be some real focus on those critical areas that you were touching on and I'm not clear yet whether that degree of cohesion between these different initiatives has come about.
FK:	I don't know if maybe my colleagues who are more in the land resources management can cover…yeah, can tell them. But the cohesion and accordance between the different services will depend…now the new approach is to go global, whatever you produce you produce to global, because now the capacity for computation is there. Before we used to work on small regions because we did not have the computer capacity, now it's there, we go global. Now we will have probably several products that may be conflicting. There are some people who try actually to combine different products and getting the best information that you can have from each product and probably with probabilistics approach you can come up with the most accurate product among all the existing products. It's kind of blending the different land cover maps. This was done by IES and we collaborated on this exercise. Now we try to do the same for the cropland mapping, specific cropland mapping.
	But I think today we should also take the opportunity of availability of high spatial resolution data and the computer capacity to start the process to recalculate all of this. Because all the land cover maps that we are using are based on low spatial resolution data and we should be a little bit a pity to continue to work at 1 km spatial resolution whereas we have data at 30 metres spatial resolution, and probably in the next future 10 metres [inaudible 0:21:45].
MR:	No other comments? Let's go on and here from Matthew.

MS:	Thank you. Thanks for inviting me. Thanks Roz for giving me the opportunity to speak. I like Francois will tell you a little bit about what we do at Microsoft.
	But what we're doing right now is trying to establish a home in Microsoft, a big software organisation, for geo-temporal analytics. So that's our phrase for things that vary in geographical space and time and of course other companies also do geo-temporal analytics and we're trying to find out where it resides within Microsoft. The company is a cloud and devices company so it wants to help people get the information they want on whatever device they want and that information to be taken from…well for the cloud to help enable you get that information.
	In terms of problems to do with helping the other problems to do with sustainability or how things are influenced, how things are expected to vary over time etc, it sure would be great if any farmer could see the predictions of how their crops are likely to change under the next weather forecast or a country see how the resilience of their crops is likely to evolve over a decade or two and so on.
	From a computation perspective we see the solutions, the question is where is the market for it and how to support it as a business, in what way would it be valuable for companies as large as Microsoft to do it and provide that capability. For example would it make sense for Microsoft to become a brand, to get very big into agricultural predictions, they call that a vertical…its own agricultural prediction, it's very unlikely. But what they would like is something like…we would like to enable the agricultural prediction verticals using our platform, make it easy for those things to happen.
	So we're working hard on seeing what sticks within the company. So how do we get there? We set up about eight years ago by my boss Stephen Emmott convincing Bill Gates to fund…set up our group and have a team of scientists comprised of biologists and ecologists rather than the traditional computer scientist in the industrial research lab to work on societally important problems. So in ecology and environmental science it was food security, water security and climate change and I got particularly involved in the climate side change of things when I came in. And at the time it seemed like a big problem. So our expertise is in predictive modelling, so we would like to enable people to make more intelligent decisions in the light of information from models they can rely on. Whether that's a crop prediction or whether there is likely to be a forest fire or something like that.
	What we would like to enable that is for you to have for anywhere in space and time and obtain the geo-temporal information that you want, it might be your crop predictions, it might be your satellite images, it might be whatever. Do predictive analytics on top of that, so combine the data…say you want to build a predictive model of forest growth, say you're trying to estimate how much carbon is going to be fixed over a 20 year period under a forest restoration programme, you want to claim your carbon credits, something like that.
	Then how do I evaluate the reliability of the predictions of such a model? And we would ideally like to enable that to be spun up by say any Non-Annex I country or something like any REDD qualifying country to be able to do that for anywhere in their country without having to build up the infrastructure of something like the Met Office in order to do such a thing, because that would eat up all the financial benefits you would gain from doing it. So we want to enable people to combine the predictive models with the evidence, including uncertainty, use the analytics to predict something new, produce new information like crop yields or the risk of a disease or something like that, and see that probabilistically as a function of all the inputs and the science in between. Indeed we'd like the people doing the expression of the models to be domain experts. So having the expert knowledge, so we would like the people developing the forest model to be forest ecologists and people doing the disease spread model to be epidemiologists or whatever. Rather than often what you get from the software industry as well, so neural networks

	and machine learning techniques and stuff will happen that we don't understand but we've got training data and 'Look it's got a great predictive accuracy' and they show great success at doing that and it's hard to compete against them. They can predict what you're likely to click on when you type in your Internet search within a split second and then you get this kind of ranked list of things that are most relevant. So there you go, there's prediction, we don't care about causality which is quite infuriating as a domain scientist.
	So in terms of a concrete example, from the climate change perspective then we undertook a project at a time when one of the big questions is how the vegetation around the world is going to respond to climate change. I think most of you will know about this problem that if some…so we have models of vegetation on land that we put into the climate models that predict how much carbon is going to be absorbed by that land surface. And in a modelling to comparison exercise about 15 years ago now one of the models predicted that by 2100 under climate change the vegetation would absorb as much carbon dioxide that we emit into the atmosphere per year. The vegetation would absorb as much carbon as we currently emit into the atmosphere every year, so they would do us a giant favour in being a buffer against climate change. The other extreme one of the models predicted that the land vegetation would emit as much carbon every year as we currently emit, so you would get this horrible feedback. So that spanned the range of predictions.
	But what had not been done at the time was a rigorous assessment of in one of these models how consistent our ability to predict aspects of terrestrial vegetation is with the empirical data. And this is a world that we would like to enable, that when you have on demand predictive analytics, say it's an ecological, a forest model, I can see how well my model explained carbon in the past, present and into the future, but more than that of course carbon is broken down into photosynthesis and that carbon in leaves and the proportion of evergreen deciduous and different plant functional types and stet wood, wood carbon and soil carbon and all the complexities of that. In an ideal world you would be able to see how all those things are varying over time alongside the forest forecast. It seems like a fantastical situation yet that's the very situation that the Met Office has with their weather predictions, that you can see how well…if you had run the model in 2050 you could see how well the forecast would have done. They'd have a very good characterisation of the space and timescales over which the predictions do shadow reality. We're like the Met Office for biology or things like that.
	So anyway, what we did was join all these different datasets including data from the JRC. Global datasets are mostly global in coverage but never at the same site, there was never one site that had multiple different datatypes, so we had photosynthesis rates, fire frequency, that fraction of plants that are leaves…that are evergreen versus deciduous, soil carbon, so over 10 different datasets. Sorting out the licences for all of that was…the permission for that was a nightmare as you can imagine.
	Then we do a machine learning technique where for things we declare to be uncertain in our model we infer uncertain components. So you do…Bayesian parameter inference was the technique we used. But actually the technique is less important as the key insight is the…or rather the key thing we learned that is really useful for the future was the methodology one uses to characterise…rivers that characterise the scales over which the model has predictive skill. So it's been verified to a predictive skill. So for our model it predicts the global distribution of land carbon under specific conditions with about 70% accuracy, for example.
	That started the process for us of doing this with more and more models and realising, hold on, we can get remarkable degrees of predictive accuracy with these models if we train them against data and setups, if we do this right and thinking well this could be really useful. We've been doing it recently for crops and things like that.

	It would be great if we could make it easier for people to do this or maybe we're fooling ourselves, maybe we're just making this seem as though it's really hard but actually everyone is like 'Who cares about it?' and actually finds this sort of thing really easy. This is we're fooling ourselves. That's definitely not the case.
	So after…about two years ago we said so what do we really want to achieve with our research? Do we want to continue to build up our CVs and show how clever we are or do we want to actually have societal impact? We said the latter, so stop doing research and start working with companies because we thought well let's prove that we can make value for companies by helping with their predictive analytics. So far we've been working with an agricultural company, the UK's largest salad producer, Sainsbury's is another one which is a supermarket retailer and a water company. And we've been involved in a lot of conversations with different businesses who are wanting to make decisions about…more from an agricultural perspective. A common problem in agriculture that we see is that people have uncertain supply and they want to be able to forecast supply, they want to build a predictive model of that and see how well they can predict it and how well that model does and over what scales does it fail to track reality and if they can use that they can maybe…well, hold that thought, so they can see how well they can predict supply. They also want to be able to predict demand but represent demand uncertainly.
	Then what they want to do is do supply and demand matching to minimise wastage. So a lot of the growers overgrow by 30% because they…one of the things in the UK you can't do is fail to supply the supermarkets and so if they ever…for example, if one of…an example company, it's not necessarily one I'm working with, told me that if they are likely to fail to meet the demand from the supermarket this week, if they know three weeks in advance they'll ship the crop over from the United States on a container ship. If they know a week in advance they will fly a plane, charter a jet to get the crop there, it will make them huge financial loss but not as much financial loss as failing to meet the supermarket demand.
	But very easily by being able to match these models of supply and demand we can see ways in which we can reduce…adjust the planting schedule in order to minimise wastage. That's just a very simple little kind of practical example but we see potential revolution in agriculture and in general in environmental services by being able to do predictive modelling, but being able to do it more rigorously than we've done it to date. The computational capabilities don't seem to be the limiting factor, we can do it, the machine learning, getting the data, getting the devices. Where we seem to be lacking is the support for doing the middle layers, the analytics part. I am biased because I'm an analyst so I'm always going to say I'm sorely lacking in my particular discipline, but it seems to be true and is backed up by more independent sources. But that's where we see there is a potential market and there is going to be great benefits by providing this predictive capability. And the form in which it would take is unclear, whether it's going to be individual start-ups specialising in crop modelling and modelling particular crops or whether it's going to be a large European global centre in doing predictive analytics and a large problem is a lack of trained people in being able to do that. So a problem that I end up talking about most frequently recently is how do we build the generation of people who can combine the domain understanding with the predictive skills. Because normally if you go down a quantitative route you don't end up in biology and environmental science and vice versa. These companies that we've been working with who are trying to predict supply and demand problems they were frequently getting their new employees in who know the new technology in the UK from Harper Adams College and things like that where you get an agriculture type degree, but they don't have any of the predictive analytics skills. So now they're starting to get them from Oxford and Cambridge in the UK and teaching them the domain knowledge there, but they're now employing environmental scientists and ecologists rather than quant II mathematicians and so on who either they can't attract, it's most likely they can't attract but also a bit harder working terms of getting their head around. But we need to find a way of tooling up the people to

	provide this new valuable information, this thing that is going to drive the agricultural revolution for us is agricultural intelligence. It's not saying there wasn't intelligence in the past, it's saying that we're going to get new insights, we're going to learn new things as a result of being able to assess things systemically with uncertainty, being able to join things up like never before.
MR:	Thank you very much. Any questions on this?
	Could I ask is it going to be competitive with Google and other big companies, are they doing the same thing?
MS:	Yeah, yeah they are. Yeah it is competitive so my closest colleague went to Google a month ago, yeah, there's a lot of competition now for doing this sort of stuff. There's not enough people to fight over, that's why we have a fight.
MR:	If there are no specific questions for Matthew let's go on to Craig's talk.
CM:	Thank you very much and thank you for letting me talk.
	I was just thinking yesterday I was talking with a friend of mine who runs a company called CartoDB and he was telling me all about how they're pivoting their business model to focus on business analytics and using pretty much the same sorts of things that you were talking about and he was reeling off 20 Silicon Valley companies that are all doing the same thing right now, you know there's a big push to commoditise that space which is kind of interesting.
	So normally when I talk, because we are a data visualisation company I have a big screen with me showing lots of fantastic data visualisations, this is a bit of a new challenge for me to do this just purely with…
MR:	Yes, just wave your arms.
CM:	Exactly. I can't draw graphs with my hands and things.
	But Vizzuality we've probably spent the last eight years trying to figure out ways to help scientists and non-profits tell their stories with data. So we don't…I guess I mean in some sense we do do research but really we arrive after a lot of the research has been done. We always say to people can you let us know before the research has been done and maybe we can influence that a little bit, but quite often we are faced with a position where we have large volumes of often satellite data, but not necessarily. We recently were working on a project with BBA, the bank, looking at the distribution of people and how they spend money in Spain and when the newspapers picked it up what they discovered was the British spend a huge amount of money on beer in southern Spain, so it wasn't great insight to be honest, I think everyone kind of knew that. But it was also…what they were trying to do was to create a function within the bank for their customers to be able to figure out where to put businesses because they can figure out where consumers are spending. I don't know, there was about 10 million data points across a summer looking at spending and what we were able to do was create a story around those 10 million data points in the browser on a website which people could very quickly start to get insight and understand things. Actually something around banks is not really our main area, we typically work in conservation and human development, so I guess…I was just trying to think of a good project which is a good example and I guess we work on a project called Global Forest Watch, so we built an application online which allows you to view deforestation events across the world. So the University of Maryland, I want to say Matt Hansen…
MS:	Hansen, yeah.
CM:	Yeah, Matt Hansen created models to predict where to look at land use change and where forests are disappearing and where they are appearing across the planet. And this is not something that's I guess particularly new, JRC and others have been doing this for

	quite a long time, but what was novel was we were able to create something online which you could interact with, so you could very quickly zoom in and out and look at areas of the world where deforestation is happening and it became very popular very quickly, and I'll try and explain why I think it became popular. In the course of a year there were literally millions of people have been looking at this information across a spectrum of society, so from students to just people kind of half interested in environment to people who have stumbled on it because we created maps which could be embedded inside things like the BBC and the Guardian and the Washington Post and Al Jazeera and well lots of websites across the world.
	So when we took up the challenge what we were trying to do is figure out whether we could create stories around deforestation which would appeal across the board. Typically when things are made, when outputs are created around…scientific outputs are created they are very technical and they are very hard to understand from a layman's perspective when they first enter it. So what we were trying to do is design an application which will allow you to go 'Okay, I don't know a great deal about forests, I don't know a great deal about deforestation, I certainly don't know where in the world it's happening', and so very quickly you can understand where in the world it's happening just from a map. We're trying to create kind of an emotional feeling that 'Okay, this is happening in the world and this is a problem', and to create the connection which allows you then to delve into it a bit more. So you delve in and you can start layering up other information to figure out why it's happening and where it's happening. Then if you're really interested you can get to the data itself and that was kind of important, often people go 'Right, we'll make data open and available', but then you're actually only making it open and available to a very small section of people, so we wanted to kind of create something which would go broader than that.
	In that particular case we…it's much more impressive by the way if I show this on a screen rather than talk it in person. So one of the things we were really trying to focus on doing which is a big friction point when we're talking with scientists around this particular…when we're trying to visualise their data is we're trying to remove stuff from their science, we're trying to remove data, we're trying to remove some of the methodology, we're trying to remove quite a lot of things just to drill down to a very simple message which is clear. That friction point is always there in any project we work on and this kind of removal of…I don't know if it's removal of facts, it's removal of noise. So with Global Forest Watch we did want people to delve in and look at the data in a much more detailed way. But because it's the Internet you can get people to do that later, so you can progressively disclose this information to people which is what we're trying to do. So you start incredibly simple. So the example I always use is the remote control, the evolution of a remote control for your TV is a great example of progressive disclosure. So when the TVs first came out there was two buttons…oh no, probably four buttons, the channel and the volume and that was quite nice and it had a little cable running to the TV. Then the TV makers decided to differentiate by how many buttons they could fit on the remote control, so in the 80s you'd get these humongous great remote controls with 300 buttons on it and that was the thing they were trying to sell to people. Then obviously consumers went 'Well I've got absolutely no idea how to function my TV anymore, I just want to change the channel', so they put that little slidey thing on top of all the buttons, so you were just left with the channel and the volume and all the other stuff was hidden behind the slider. So that's kind of the classic example of progressive disclosure, we try and do that with our websites, we try and just show the channels and the volume and then if you're really interested, if you're the scientist or the data scientist or the analyst or whoever you can reveal all the data and more detailed visualisations and you can take it away and you can do things with it yourself. So across our applications we try and replicate that.
	Right now as a CEO it's quite nice to say that we're quite popular at the moment, there's a huge demand for figuring out how to interpret large volumes of data and making it

	available on the Internet. That space if you're ever thinking of setting up a business I'd recommend going into that one because there's more and more…it's not like we're going to have less data any time soon and it's not like we're going to have to try and interpret it in more and more ways. That's been commoditised quite heavily but there's a big role for that. One of the things that when we created Global Forest Watch one of the things that we were kind of uncomfortable with was how does this relate to the person near to the forest with a phone in their hand and how are we going to use that information for them to actively manage or to change the way in which they go about their day, to change their behaviour in some way. So this is the sort of challenge we're trying to figure out now and we're not actually doing it with forests, we were contacted by a foundation called The Zeitz Foundation who are based in Kenya and their goal is to improve the conservation farming practices of small-scale farmers in northern Kenya. So they came to us asking us initially could we create kind of dashboards with information based on remotely sensed data that could feed into the planning of a lot of these small-scale farms in the Laikipia area of Kenya. So we were sort of talking with them and we went and we visited and realised that actually most people there have got regular feature phones, so you can just get text messages and stuff. So then we started to think can we figure out a way of going from this humongous volume of data all the way down to a small text message that can influence the way a farmer may go about their farms. Like can we create something, I'm not sure what the answer to that is, whether we can or not yet, we've just started it but what we can do is we can connect all those pieces of technology, that's all there and that's fine. When we did Global Forest Watch every year we…Google they actually run the analysis on their ad servers which is where they have most of their capacity, because it was such a big volume of data and every year you go kind of kerchunk and you get this new dataset of deforestation or forest gain as well. But that's kind of solved, that stuff is sort of solved, it's hard but it's sort of solved. Texting people is easy now, you can do it within about that much code you can create automatic text messages. So you can do like big analysis very quickly and you can feed that to text messages very quickly, so that stuff is quite easy. The stuff that we're now struggling with is what do you put in that text message and is the person who has the phone in their hand going to pay any attention at all to something that is coming up on their phones. I'm pretty sceptical but we're going to try it anyway, but you know, we're going to see what happens with that. That's I think where it's a different form of visualisation but that's kind of one of the aspects of the company where we're moving to because obviously there's the obvious patterns and increase in mobile phone usage in Africa. There's the obvious trends in big data, more coverage of satellite data, better analysis, better way of predicting objects on images, all those things are kind of all going along and being commoditised. So we're kind of moving a little bit into that space. We still do beautiful things on the web and we're still doing that sort of thing but we are now focusing a little bit more in that space.
	So I guess one of the questions that I think might be worth asking is how can we test whether humans are going to be willing to change their behaviour using this methodology or a similar type of methodology as people convert to smartphones and are using the Internet more in places where they just don't use it now. How are we going to be able to use that scale effectively to change the behaviour in terms of productivity and sustainability and the resilience in terms of farming. For me that would be really interesting for the research community to be looking into because then when we go about building these applications we can use that information, you know we can build them in a way which has more impact. We're not going to do that, that's not our thing, that's not where we sit.
	So yeah, that would be interesting in itself. I think there's a broader challenge which I'm not sure it's a research challenge but is still a problem in the way academic institutions and non-profits go about communicating their data. There's way too many data portals out there which no one looks at, a huge amount of money is being spent on this and time and time again if you have a look at the analytics of how people are using this stuff it's typically only looked at by a handful of people. Normally they are male between the ages

	of 25 and 35, they are based in the West and they account for a very small proportion of the world. I think, yeah, is it a research community's job to figure out how to create these applications better, maybe yeah. I mean it's good for Vizzuality because we try and do it too, but I think there's definitely space there so maybe that's something worth thinking about.
	I know with the Copernicus programme something that I often hear grumbles about is a huge amount of money is being spent on things like satellites and storage of data but not a great deal of money is being spent on how you make that data accessible and easy for people to use, whether you are a data scientist or not, and I think more needs to be done there as well.
	Is there anything else worth saying? Oh yeah, so there's one other thing I think which I think is a good example, so it's not related to food but it's kind of interesting that we're working with the University of San Francisco where they've created models to predict malarial risk in Swaziland and in Zimbabwe, based on a whole bunch of environmental parameters and trained with field data. So we're creating an application now to allow the people, the health services that are planning where to do the spraying and to put the mosquito nets we're creating an application for them to hopefully be more coordinated in the way in which they…where they coordinate their efforts to be more effective there. So that's one angle to it and the next one is making a mobile phone application which allows the people who are doing the spraying and doing the nets and doing the detections of malaria and feeding that back into the models that are going into [inaudible 0:52:23] risk.
	Very rarely you see…so what I think tends to happen, it might be a sweeping generalisation but what tends to happen is scientists will run a model and they will create an output of things like malarial risk and it will be a one-off thing and then maybe it never gets done again, maybe it gets done in a different place with a slightly different methodology so you can't compare the two things, in a different place in the world. Operationalising those things, because they are becoming easier and cheaper I can see that happening more and more and so with malaria is a perfect example of that. So we are using Google Earth engine, we're using just standard web technology and technology which you can make applications on phones to allow people to do real things on the ground and then feed that cycle back round in terms of malarial risk analysis. So yeah, I'm sure there's plenty of space in which we could do similar things in agriculture, in conservation, maybe that's an area that we could look at too.
MR:	Thank you very much. Any questions on that?
GD:	I have millions of questions. I think it's great, it's very complementary that obviously we are able to monitor and assess faster high resolution on environment, we are still pretty bad at understanding the causality and the drivers of these changes. Very often what is happening here is physically decided thousands of kilometres away from that, especially there is a big gap I think with social sciences, for natural sciences we have better models but there is a big difference between being able to predict the growth and speed of a tree with what is going to happen to land because someone else bought it as well.
	But regarding the job of [inaudible 0:54:20] and bring all this information down, I mean we are focusing on making all this information available and we start to see good services to provide tools for people but I think the lack of involvement of the society in all this, they all see this information coming but they have very little means to get feedback. We had very similar discussions in the Global Biodiversity Informatics Conference, you may be aware organised by the GBIC in Copenhagen. At that time I had suggested to develop something called adopt a pixel, you know people should have a pixel, a virtual pixel on one single platform where someone can spot those, someone can spot the environment, someone can say "Well this tree is growing." Getting feedback on that in the way we have the Galaxy Zoo that was allowing people to classify planets and gave feedback and they could see the results of the interaction with that information system. Here we still

	have information systems that bring us a lot of information, we have no question that anywhere interesting is decided as well, like the climate change, I mean we are approaching [inaudible 0:55:24] that will have a number of options regarding our response to that but I don't think the citizens still feel involved in that because they don't understand. It's like the Global Forest Watch is very successful because people start to report of illegal logging, they start to understand that they can make a difference, but the [inaudible 0:55:43] a lot of people to give the feedback is almost non-existent. I mean there are very few examples. I think one way we can make a difference is through this adopt a pixel approach really.
KS:	I have a question for all three of you. I wonder to what degree you think it's important to start from the other direction and look at what the problem is and whose problem it is and how you solve it, rather than having the datasets and everything and then trying to wrestle them into a useful format.
MS:	I think it's really important to go in that direction. All models are wrong, every abstraction is wrong so it's about working in the light of that wrongness. And so the state-of-the-art approach is robust decision-making methodologies where you define your problem well and you look at risks of…you look at what factors will alter you making a decision and there's some good examples of that. There's some companies like the RAND organisation in the US that's been doing that for many years, so it's hard. And there's starting to be public examples of it, so one of the good examples from the climate perspective is how the Netherlands decided on their flood defence infrastructure on the basis of the climate model projections. They took an excellent approach in terms of what types of outcome would most affect with the chances of flood and the impacts of flood on the Netherlands and they split all the model projections down into four scenarios which are just then split down to two scenarios for the decision-makers, either this or this. But that was then simplifying it down to I think it was small or large global impact, small or large local impact and different outcomes for the climate that people have…air temperature and things like that versus sea level and made some decisions on the basis of that. So they did use the climate model projections to make decisions and many high-value decisions as well, very expensive decisions.
KS:	So it's a sort of practical conservation kind of question. I mean I've spent lots of time living in Tanzania and I quite like the kind of remote sense tools and applications and so when I take these nice maps down to people they do work up to a certain point, I mean and then people look at them, they go "Oh yes, forest has changed or this and that has kind of worked out" and then when they say "Okay, let's go and actually do the work", they're going to leave them in the office, they get in the car, they go and they go and talk to the villagers and ask them what happens and they do household surveys, they do health surveys, there's a whole other world of kind of…you must see in West Africa all the time, there's just a mass of this other world of data gathering and the gap between kind of that world and the remote sensed world kind of happens when you try to actually go to a place and use your remote sensing stuff. And they like it up to a point and then they go "Let's just go and talk to people" and they've all got simple mobile phones and all of that is completely correct. So to me that's the kind of the challenge is bringing the remote sensing down to the human scale at the village or the agriculturalist in Tanzanian mountain village, or even worse the pastoralist moving their cows around which even more tricky to deal with them, taking it down to that level and then informing upwards from that level. Solving that is a vast challenge but also would make the models much better and much more useful and more real to the kind of people who want to use them. Answer that!
CM:	So I think the malaria project I mentioned is trying that. Like they are trying to combine the fieldwork…well the work that is going on anyway, making that work more efficient, so creating services, I think the way in which it's described is they use that technology because it's the easiest and best way for them to do their jobs and if it's not they probably won't use it, so they will have failed as a project and I think starting from that metric is a

	good place to start. Does it work? I mean they've only just started so it's hard to say whether it works or not in that case but certainly there is no reason why you can't go into…if there's fieldwork going on on the ground there is no reason why you can't go into those villages, do your standard surveys, do all your stuff, so we're creating forms for people to fill in in a very quick simple way which then synchronises with remote sensed data and improves the quality of the modelling, the analysis that is done on that data. Yeah, I mean that's always the challenge and this is something that we found with why we wanted to move more into that space of can you text farmers with information, can you get the health workers with information on their phones which is going to help them in their jobs. There is a huge opportunity in terms of scale in that sense but then it's massively difficult to kind of market and advertise the things that you're trying to create for people to use. That's often missing a lot in these initiatives is if you're a company you have a whole wing of people who are dedicated to marketing your thing and getting people to use it, whereas that doesn't really happen very much in a lot of these initiatives.
FK:	I think also in terms of bringing those technologies to people on the ground it depends also on the scale and also the responsibility with different people you are talking to. If I'm talking to people in offices who are [inaudible 1:02:05] who are responsible for this for example satellite imagery would be useful and they would be really keen on using it, but if you talk to individual farmer the satellite imagery that we usually use which has 1 km spatial resolution does not mean nothing for individual farmer. But when you start to go to 1 m spatial resolution then it becomes very interesting. There is for instance now a recent case in Rwanda, they mapped all the land parcels using aerial photographs, so each individual household has a map of its land and the people were excited to have this picture, saying "Okay, that is my land" and they have the photograph of this land. So people realised how important it is to have these pictures. But if you are…now the key issue as Craig said, how do we communicate easily this information. We have Google Earth which is great but how we manage to have Google Earth on very slow Internet connection, how do we get it on mobile phones. Because now in Africa you can forget about landlines, everything now is about mobile phones and there are several mobile phone services which are being developed from mobile money to all kind of service like [inaudible 1:03:41] data collection or household service. So there are a number of things that have been developing based on mobile phones, then now the [inaudible 1:03:47] will be to be able to deliver this information on mobile phones with the problem of connectivity that also we know.
MR:	Matthew, any comments on that or otherwise we'll have the tea break?
MS:	Well just that I'd like to see more of…you know in the UK people get some benefit from - it's difficult to put a number on it - from flying drones over their farms and seeing their farm like they've never seen it before, they get some benefit at least in terms of satisfaction, of course we'd like to see more demonstrable cases of commercial benefit from that. And I think that's what we're sorely lacking, whether it's in Africa or anywhere in the world and there is a problem enabling that insight, so what I would like to see is connection between the information that is delivered and the decision-makers' bottom line. But the trouble is that that's a big risk project and the people taking on that risky project often are very technically capable, say it's the academic community, they're going to have a lot of research that is going to not end in publication, it's going to end in failure before you land something and when you land something it might lead to an all right publication but will potentially lead to you running off and starting up a business. And how do we support that? Ultimately you need to have lots of failures to find things that are going to work. So who is going to do that joining of the dots? Who pays them?
MR:	I think we should take the tea break now and come back in about 10 minutes to continue the discussion.

	After Tea Break
MR:	Okay, we've got a bit over half an hour before dinner of discussions. Thank you again to the three witnesses. Keith Richards you had a question?
KR:	Well I was toying with a question and here it goes. I think I'm going to ask the question that Sue would normally ask and Sue is fed up with asking it so it's my turn to do it. So my question is it's about the impression that one has from a lot of the discussion that data is somehow value free and I'd like to talk a little bit or us to have a little discussion in a sense about the politics of data and what the implications of then those politics might be for the way in which the kind of data that we've been talking about are used and it goes back a little bit to your question I think.
	Here are two illustrations of the problem: I supervise a PhD student at the moment who is working on nexus problems in China and so she needs to have quite high resolution land cover data and there are reasonably high resolution land cover data available quite readily around the globe. But of course the classifications of the entities mapped are entirely contingent on who collected and classified the satellite imagery and are completely different between the sort of commonly available global mapping and Chinese land cover maps and it's extremely difficult to put them together and it's actually quite difficult politically to access the Chinese data and so we find ourselves in this slightly odd circumstance of doing a study in China using globally available data and then gradually trying to access data that are not usually available to non-Chinese scientists. But then there's another sort of illustration of this problem which perhaps is even more interesting which takes me back to a paper which I read some time ago which I know that the research group I work with always found quite interesting, it's by someone called Claire Winterton at Lancaster and I don't know whether any of…I think you probably know this. She did this really interesting study about mapping vegetation and she went into the field with a group of people and mapped the ecology somewhere in the UK and showed that in the field you can't actually find any things that map onto the entities that exist in the UK vegetation classification. So when you drill down if you like to the real world on the ground what an ecologist can identify at a local scale is nothing like the pixel that that location represents in the UK land cover map. And then if you compare this with the European CORINE classification for example all of the categories, all of the entities are different, do not map onto the classes in the UK land cover map and in a sense the reason for that is that that's a European project that has a politics to it and the representation of the information about land cover is based on a Europewide classification process that doesn't produce entities that really are representative of things in individual countries in Europe. But the politics of this is to pretend in a sense that they do and everybody knows it's a pretence and those who know it's a pretence, and this is part of the Winterton argument, use it for their own purposes and manipulate the mismatches that exist in order to help their own particular objectives come about. So I think this is an interesting area perhaps to explore because so far I think we've had the impression that somehow data tell us things that we will all believe when in fact of course we won't.
FK:	Okay…
CM:	You absolutely have a position on that.
FK:	Let me comment on that. I think to start with the data policy and the value of data and value and cost of data it's true that usually there has been a policy of retaining data, not sharing data. Now I think with the Internet there is more data available, people are now keen on sharing data because they are proud of demonstrating what they are capable of, so people are producing data and making them available. But it's still true that most of the data that they are producing with public money are still hold and not made available publicly. So the case of China you mentioned we see also the same situation in a number of countries where the data is produced but is not available to scientists, to other

	people. This is a problem maybe that we should try to raise also in terms of general information communication and also data availability policies at high-level. Usually it comes from the environment deciding not to disclose data because for instance it's sensitive or they want to retain data because they want to have advantage of having this information, because having information is an advantage. But there is a tendency now to be more open on data.
	Then I think you mentioned another problem which is about the actual scale for your applications or if you have a land cover map and then for your application it will not work. There is always this problem of the scale. The usual I think example which is given is if you ask somebody what is the length of Britain's coast, what is the length of the coast, there is no solution, there is no answer, it depends on the scale. It's the same with land cover. When people were working with satellite imagery at 1 km spatial resolution they were happy to say vegetation, no vegetation, water and perhaps urban areas, built-up areas. When Ikonos and QuickBird satellite imagery arriving we start to not talk in terms of urban areas but in terms of building, gardens, trees, so then you have to downscale. That's [inaudible 1:35:03] the question then, you start to go into the [inaudible 1:35:10] of sociology, I don't know if it's called that in English because I studied in French.
KR:	It's the European school of sociology.
FK:	Then satellite imagery will not be really very useful but I think somebody was talking about sensors, there are now many other sensors that are coming in and all this big data, how to put all this data. You can imagine that you will probably go in the field you will use your cell phone and while you are taking notes about different species, at the same time taking several pictures with your cell phone and putting all those cell phones in one place and people can share those and then you can somehow map large pieces of land with ground information.
KR:	Crowd source classification?
FK:	Yeah. I think this is one of the things that will probably happen when people will be keen on actually publishing what they have, sharing information, what they do now with Facebook and the Twitter, yeah. I think that's something that we can think about for the future.
MR:	Hildegard, have you had problems with Tibet and China with any of this data?
HD:	Well I was just thinking, I was listening because I'm thinking especially about certain very, very remote communities who have very recently been connected with mobile phones and they are exposed say to glacial floods for example and they recently have been discovering that there is a possibility of having satellite images of their glacial lakes. And the question is to what an extent can for example access to that sort of information become something that could be drawn in in taking some strategic decisions, should we move the village, should we build particularly safety measures in relationship to that particular risk? And it's very difficult because it's a situation in which decision-making happens on multiple scales. So on one hand it's about empirical knowledge, so you have people going out checking the glacial, trying to understand what's happening. On the other hand there may be information comes from previous cases or there is past experience but there is also a point at which there is total uncertainty and then we call in the diviner. So the question is at that point to what extent having access to this kind of information could really make a difference, but what is the place of this kind of information within a much more complex mechanism of decision-making and this is where I think there are challenges at multiple levels. One of which is how we present the data in a way that is accessible to people who are not professionally engaged in decoding this kind of very complex information and I think this is very interesting. But on the other hand is how do we also work out the place of this kind of information within a wider setting and certainly a setting that is highly politically sensitive. So sometimes the very high resolution maps are possible but not available just because they are considered to be a

	military secret. So I think the complexity of the decision-making process is really challenging, while I see very much the potential usefulness of making things more accessible and do you see any way in which we could simplify that to some of these very, very remote communities which often are actually the most vulnerable to extreme events that might be climate change induced?
CM:	So we were just talking at tea with Neil about…that it's already happening, certainly with farmers getting things like commodity prices of whatever they have, cabbages or something. Was that the example you gave? Yeah, cabbages. So the private sector has already cottoned onto this and they are already pushing into this space to provide information in all sorts of ways, if they can make some kind of profit on it obviously. What we've found with our examples is just using very simple and basic design principles in a way in which you display information in a very simple very digestible form makes a massive difference and quite often it's just not done that way, so that would just be a start, just a start. The technology is there but it's limited if you've got a feature phone rather than a smartphone so it's going to be hard, I mean I don't think it's an easy problem. I haven't got any experience in how you put that into one of these complex decision-making processes, I've got no idea, but yeah, yeah, I don't know about that.
	In terms of the free and open data and the perception that data are free now this is a big problem. It's okay if you're working in the public sector and your government or your region has decided it's important to spend lots of money on maintaining the infrastructure and the support around data, whether it's satellite data or whatever it is. But what I've found in the conservation world is that they don't have access to those same resources so what tends to happen is this kind of project grant-based approach to raising money to support these big conservation datasets where everyone else is expecting you to hand them out for nothing. And so there's this weird situation now which certainly in Cambridge where you've got WCMC where I used to work and IECN and all these other guys that are trying to maintain biodiversity datasets and protected areas datasets off the back of project funding but they're being told that they have to do it for free. So the model there is they restrict the commercial access to the data so companies will pay to use the data, which seems kind of reasonable in one sense but actually it's stopping all these other companies using this great conservation data to make the types of decisions you're talking about on… So I think there's a challenge in that.
	Now what the satellite companies are doing now, the likes of Planet Labs and UrtheCast and Skybox who have just been bought by Google, they've realised that their model, they're going to get all this investment in creating satellite-based data, they're really struggling to find a market to sell it because no one wants to pay for it. The military might pay for it but some of these companies don't really want to sell it to the military, you know, they want to do other things with it, sort of starting in a good way. So what they're starting to do is build a layer, a valuable layer on top of the data. So the data itself is just a given, it has to be free, and then they create applications on top which will do things like tell you where to put your hairdressers or whatever, whatever it happens to be, and that business intelligence is the thing that is going to drive a lot of the innovation around data in the next few years. Yeah, I don't know…but the problem with that is then that all those applications that they're making are not designed for the cases that you're talking about yet, because there's not great business there.
HD:	Can I return to that? Because I think that is one of the big challenges in the sense that actually the people who might be in the most vulnerable position are not the ones who are in the position to pay for these kind of resources and I think that that is a real issue and they will be able to perhaps use this material only if it's available for free.
PL:	It seems to be one of the issues with this kind of data and it's the sort of thing that Hildegard is referring to is the ability or not to predict what might happen as a result of this information and therefore to make a decision based on that and I think it was Matthew who mentioned about having a Met Office associated with agricultural biology or

	something. It strikes me that there is a real difference between meteorology for example and conservation in that we understand, we know, if we measure the pressure distribution of the atmosphere and we measure the temperature we know the relationship of how strong the winds will be. I mean the relationship is complicated, it requires big computers to work it out but we actually know cause and effect and we all agree about that and the uncertainty is really in a result of the inaccuracy of the measurements and the inability to compute with sufficient accuracy and so forth. So what we've seen if you look at weather forecasts they sort of improve one day extra per decade basically, so now I think we have about a five day forecast that we can rely on, so wait 10 years and we'll have six days or something like that. But that's then useful, so we now can look up the BBC weather and it will say 'It will rain in Cambridge 10 o'clock' and so you hop on your bike at nine o'clock and get there before 10 o'clock and that sort of information is useful and we can all act on that. But if it just said as it often used to say 'Well showers' that's kind of useless information in some sense and I think that seems to me to be one of the issues that we are facing in this sort of conservation and agricultural world, how do you make these connections, so I just wonder whether you think about that and do you see a way of getting over this?
MS:	I think what we are most excited about there is the two UI, user interface concepts and user experience concepts that we talk about most is enabling people to take a systemic perspective, so you define the system. Imagine a way in which you can set up your problem as a system and the second component being able to handle uncertainty and uncertainty comes in umpteen different forms, is as natural as you would handle normal numbers and where that comes together from the decision-making perspective is it allows you for example to understand in what context for what phenomena is the uncertainty important. Would you have made a different decision given the span of uncertainty? Maybe the uncertainty is huge but it doesn't matter a jot to the final decision you make. What we've got from doing this systemic modelling is being able to trace the importance of different sources of uncertainty to the end decision, and that…so that does a few things. I mean if we can take systemic perspective of how uncertainty influences decisions you can pinpoint the most important sources of uncertainty and invest resources in that. So when we did the carbon model we had umpteen choices for data on plant photosynthesis rates and we trained the model to them all individually together but it didn't matter a bit to our predictions because the carbon is all very similar and we had all the satellite campaigns coming up with better estimates of things relating to the photosynthesis, but that was the least important part in terms of our ability to make predictions for certain things. Whereas if you had that capability 20 years ago you might have gone maybe one less satellite and invest a tenth of it in a completely different part of the carbon cycle if that was what you were putting up the satellite in the first place. So that enables you to make those sorts of intelligent decisions.
	I think in doing that, taking a systemic perspective, understanding how uncertainty influences your decisions it sets up a market for providing estimates of uncertainty on top of the information layers so that people will get some benefit from producing an estimate of how accurate are these land cover maps or how inconsistent are they. By using more sophisticated techniques on top to infer what the truth is, knowing that all of the data has error in it and there could be a market for that in particularly valuable datasets. One of the final exciting things in terms of developments in cloud computation is the ability to keep everything anonymous in the cloud such that if a Chinese government wanted to keep its military photographs secret but wanted people to integrate over it to come up with an estimate of land cover or something like that they could do that analysis but never actually see the maps, so they could put their algorithms into the cloud, they could operate on the dataset but the person never has to see where the military base is.
	You're absolutely right that we can't forward predict a lot of these conservation problems but then I think being able to define it systemically and then conduct analyses on top of that. We call them constraint reasoning problems, how do these different aspects…these

	different levers we can pull influence the end outcome, therefore what should be the order in which we pull the levers to maximise getting the best outcome, what are the trade-offs?
IL:	I think I just want to come back to Keith's point which I think ties back into this, that it's important in all these things to have some idea of providence or trust in datasets and if you've based decisions on datasets that you believe people have treated in a particular way you should be able to demonstrate that that has had some effect, greater or lesser on the final result and then if you find that your result is quite sensitive to a particular dataset that you think has been tweaked by someone trying to make a point then that causes you to place less trust on the final result. But I think in all these things if we're going to be relying on large datasets that come from various places we have to know where they came from in case later on we decide that someone actually has done something they shouldn't have done or just have a different perspective on what the truth is.
MS:	Yeah, for our effect climate we had this design where you go for the information, so you go for air temperature and it runs algorithms…the default user just cares about air temperature and gets air temperature but doesn't know the provenance, but if they want to they can go in and see well the air temperature was recorded from [inaudible 1:50:07] place on the earth, but CRU over here.
SF:	I just wanted to make two points. I'm sorry for being late, I got delayed on the road from Oxford to here. I'm involved in a project, but I just wanted to talk about the issue about data so one of the things in the developing country context is that data even if it is collected for considerable periods of time then stops being collected for some periods of time. There is an inconsistent data collection. So related to your point but it's not because it's on the back of research projects it's government budgets collapse. So for example I've been tracing data collection in India on irrigation and soil quality, soil quality is collected quite regularly and particularly leading up to the Green Revolution in technologies until the 60s, 70s and 80s but it's disappeared now. It's almost like the Green Revolution in technology has happened, you've got the results, so I'm working in the state of Tamil Nadu where we're trying to use a mobile app application which I intend to take to Africa, but we're working in the public domain because we're working with the University of Pisa and the Portland [sounds like 1:51:13] download because we would want to bring all the players in to see the value of trying to work with data. So the Portland project is actually very simple, the SIM pings up to the satellite, it comes down, the GIS maps where you are and it measures the strength of each SIM from 100 to zero in terms of actual strength compared to that. That's very useful for computer scientists in the University of Pisa because they want to look at topography of Internet density. But for us on social science side and why do we get engaged with the computer scientist because our understanding on the ground in agricultural communities is this rupture in datasets that you get. So I guess related to the point that you're making that it might not be a bias, it may simply be a financial crunch, but what has happened and obviously you're going to think about this is the next generation of farmers because they were not trained to collect soil data do not do it in a manner that is effective because if soil is removed using a cylinder method it has to go into lab within 24 hours. So we're working with IIT-Madras because the engineering solution possibility, how do you get…? So in the past because there was a system linked to the sensors the soil engineer would go out every month, collect the data and bring it back. There is no budget, it's disappeared, the soil engineer no longer does it, the government departments don't see it as their responsibility but we'd like to work with IIT-Madras as a platform where they send information through mobile text to farmers saying 'What's your problem?' Well the biggest problem that emerged in the last two years is nobody collects our soil and we can't do it so we're thinking of a private solution, let's get up to 500 grams because that can go as a package through private courier and maybe that's a method that we think. So it's not the lack of or opposition to private, in a sense it's trying to understand how do

	we support government systems that had good data and now have good intentions but not good data, because you want to get that kind of long-term [inaudible 1:53:16] meteorological data so you actually get meteorological data similarly sorted so we can predict. So one of the biggest problems we're getting is increasing collapses in agriculture and a lot of that has to do with the greater risk that comes with existing datasets.
FK:	I fully agree with you, that is something that we observe now in Africa. The collapse of the data collection systems including the meteorological data, there are very few now meteorological station data that are still reporting globally. Sometimes you have just the meteorological station that are at the airport and then for the country this is not useful but nobody is reporting. The problem is that also in many countries the government have priorities and sometimes they feel that collecting data is not really a priority. And this is something, another thing, since there are some people here from the remote sensing communities which is quite strange is that at some time remote sensing gave the impression that it can replace ground data. And then…there is a case, a colleague who was used [inaudible 1:54:36] who told us how somebody was supposed to report on meteorological station observations, especially on rainfall. Instead of reporting the observation he was just going and collecting data and publishing on the Internet from…there was a rainforest, image from satellite imagery and reporting that until somebody told him that…
	There are a number of meetings where I've heard people saying “Oh yeah, yeah, we have satellite imagery, they are more accurate than ground data” and then people “Yeah, we believe you”, they say “Oh fine, so we don't need to collect it”, those are things that are happening now. But then [inaudible 1:55:24] this is also a question that we have sometimes with colleagues who are on [inaudible 1:55:31], when they support information systems in Africa they say “Should the European Commission for instance fund data collection systems or it's the responsibility of the government?” Because for sustainability, because if it's a project for two or five years and then after five years if the government doesn't buy in it will stop. This is an open debate, we don't know what to do, usually the idea is to say okay, it's up to the government to decide and put money there. Maybe in the future it will be maybe the crowd sourcing, to engage citizens in collecting data, then there are questions about data analysis because we are used to analysing data that is collected in experimental systems and we know this distribution, we know the statistical method to use, but when it is collected by crowd we have to figure out how to analyse this data.
GD:	This is where we don't have a proper model at the moment. I mean the business model of getting money to collect the data is changing and this is why the Conservation Committee is struggling so much, is that they have all this data, they have been paid by the public and we expect them to provide a service to deliver the data, but many do not have the proper means to maintain the data, it has a cost and this cost is never recognised and neither is the long-term management of this information. So I guess what is happening nowadays there is no strategy to even define…something again I put on the table, is this critical biodiversity informatics infrastructure. We recognise a few key datasets about fundamental…for most decisions, [inaudible 1:57:18] systems, we have the [inaudible 1:57:19], there is a lot of information posted by a few institutions but even these institutions do not have real resources to ensure and guarantee a long-term availability for this information. It's quite tricky to explain this to the policymakers who have a relatively short-term vision but if you want to sustain and maintain this information over the next 20 years we need to have a strategy and the strategy will not please most of the actors because we can't maintain and support every single individual operator hosting information, who has collected information. He will have to put it in the cloud or in a reference information system because we will never have the means to collect all this information and distribute all this information from all sources. So recognising that only a few actors and something naturally economically we have Google, we have Amazon, we

	have ideas on putting…Microsoft, sorry about you! That are putting big infrastructures and even within the Commission fighting…who holds this critical information, for the long-term it's a serious issue. The cloud seems to be a magic solution but politically it's almost impossible to accept that you don't have control of something that you also have responsibility for so I think it's very important to identify what are the critical data, how we are going to maintain, who will have the responsibility to maintain this and this is never really addressed anywhere.
MS:	There is not enough studies looking at the information content of the data that's been collected. I mean that's a real great study for the boffins is to see how much value in relation to the problem you are getting from the data. So should you collecting every week, every month, every year, is it okay to take five years' break depending on the problem. What's the information content? From a conservation perspective, Lucie Bland, I think she's an Australian actually, she did her PhD with ZSL, she did this fantastic study looking at the amount of information related to conservation decisions you get from collecting information of where the species are in the field of versus doing predictive modelling and looked at the ratio…so like came up with financial estimates of how much effort, how much it would cost to do the modelling versus the data collection and showed that for…I can't remember the application but in a lot of circumstances she would recommend reducing the data collection effort by an order of magnitude and you'd get the same information content when you combine the modelling with the data collection.
GD:	[inaudible 2:00:02] if you put more money in preserving the world history of the Kardashian family than the biodiversity data then you will get quality information. There is no strategy regarding what is important and what is not.
MR:	Is there any way of involving lots of sorts of amateurs, I mean like Galaxy Zoo and Astronomy and things like that who would actually help with some of these tasks and get into the science themselves?
CM:	Yeah, I mean that is happening actually already certainly in biodiversity and the infrastructure that's built around that, things like Naturalis are doing those kinds of things. I was thinking about how you could fill the data rupture, how you could fill that gap, is it even possible and who do you kind of prod to do that and I wonder… So satellite data was supposed to create this constant that would cover all those things, sure there would be project spikes in the middle of that and there would be gaps in that and that would be okay. So you would have a scale of data which was appropriate for a bunch of things and then you'd need different more detailed information and that would happen in spikes because it happens in trends of funding. But I don't think that's actually happening in itself, we don't have kind of the standards around remote sensing and the data coming out of that to even create that constant at the very basic level. I wonder if that's the first step, if we could create that because then there's…it's not reliant, I mean it's partly reliant but it's not entirely reliant on the stability of a government or the funding flow of research into a particular area, it's like a broader bigger thing, so I wonder if that would be a start, if you could just get that very basic level… I guess the Europeans are trying to do that with the Copernicus and NASA and others are doing it and a company called Planet Labs have got a mission to map I think it's the entire planet every 24 hours, those kinds of things.
	So if you can spin up value-added data products out of that which are consistent and scientifically rigorous and all that stuff then that would be great. In terms of what's different I think now to maybe 10, 20, 30 years ago when all these remote sensors were telling us this was going to solve everything, I think Gregoire alluded to this a little bit that the infrastructure in place for people to put things on the Internet and to publish data is virtually free. I mean you can publish data into Twitter and Facebook and you can put it on your e-mails in Google and now obviously Google read every single letter of that and they do things with the data you properly don't want them to do. But the point is there are kind of infrastructures on the Internet now which will allow you to store data even if not in

	kind of the structured way which you'd expect governments to do it which are going to sustain for a while. There's what, a billion active users on Facebook everyday now, that's not going to go away any time soon. So I wonder if one of the ways in which we could do it in terms of supporting infrastructure which is it used to be the problem that people wouldn't pay for the servers and the administry anymore, you don't need that anymore, you just start putting data into these places which most people are anyway - well not most - but a lot of people are anyway. Maybe that's a solution to that as well, I don't know.
GD:	Risky business.
CM:	Yeah, yeah, of course, of course. There's no real European alternative right now, so that's a problem.
GD:	No.
PL:	I was struck by how fragile this whole system is when I was in San Diego and the Keeling Curve which we all know and love is now…that data is now collected by Ralph Keeling who is Dave Keeling's son and every three years he has to apply for more funding from the US government to keep that going. I assumed naturally that was just a standard station and we were just watching this graph, and every three years he has to go through peer review to get this thing restarted for the next three years. And you think that's unbelievable, I mean it's the curve that we all know and love and talk about all the time and it's not just a standard measurement made by the US government, which I think shows just how fragile this whole system is.
MS:	After I did my… I'm interested in different ways to fund it and after doing the carbon study I got involved in discussions within Microsoft about our offsets, our carbon offsets which are then going to projects to conserve carbon and things like that and of course every big company does that. What I was thinking was what's the relative benefit to society? Say you've got £10 million or £100 million or something like that, would society benefit more by allocating that money to current offset projects, however allocated, or putting up more satellites to understand the system better so we can make even more intelligent decisions? I haven't seen that case well-articulated to the private sector that an alternative to just throwing money at various green schemes might be to invest in the research that would make the planet greener overall to support the satellites, so support Keeling and so on. But there's certainly always going to be a lot of money sloshing about for those green initiatives, it will go up and down but it won't drop below a certain level for a long time and that could be a good source to tap.
MR:	We've got time just for one more question before dinner.
GB:	That would be great, thanks. I'm just going to combine this and just make a couple of points and then ask a question. But about citizen science data collection we all know perfectly well the distribution and survey effort is not matched through the provision of biodiversity. So we know exactly what happens in the UK especially in terms of rare birds turning up but we've got vast tracts of Africa where we know very little about what's going on, so you're going to have to work on that one an awful lot more about developing the capacity to get people to collect those data if we're going to start using that approach.
	It was interesting what you were saying about the basic level of data provision and the meteorological example as well. In many cases we don't even know if it's raining outside let alone being able to predict at what time in two weeks' time if a shower is going to come across and the Conservation Committee needs to know first of all is it raining, yes/no, they don't need to know how heavy the shower is, it would be nice but we don't need to know what the weather is going to be in two weeks' time.
	I've just been doing a bit of work, a colleague and myself in NASA, we're just trying to put together examples of where we have operationalised systems using satellite imagery and remote sensing to inform regular daily and really quick yearly conservation decision-making processes and there's precious few examples about that. Okay, I know you can

	always think of a few, but if you think of the number of scientific papers that have been published on this topic and say 'We can do this, this and this', okay, how many have been turned into an operational system that can either be used on a regular basis, very, very few of those.
	But to get to my question, you mentioned there Facebook and content on Facebook and Google reading e-mails and suchlike, don't you think that bringing together the content of social media and trends in what topics have been mentioned from social media with satellite images in order to tighten up accuracy of incident detection and land cover change.
CM:	We tried to do something like this with detected areas a while ago, I think the conclusion was it didn't work very well, just that it was incredibly patchy and it was very hard to detect a signal in any sort of meaningful way. Yeah, and especially with things like Twitter now they are so heavily kind of polluted by advertising and marketing that it's quite difficult to find signals in that. So I was just trying to think if there's anyone who has done a study on that, I couldn't think of any. The closest we came was something kind of different which was we worked on a project with Oxford University looking at…it was looking at light patterns, so trying to detect planets. So from the telescope when the planets went in front of the sun then it changed the light signature that was getting detected from the telescopes and computers couldn't pick out the patterns but humans could from these graphs of dots of light. So we created an interface which allowed citizens to then just to look at the graphs and see where the spikes were and they detected loads of planets or potential planets and I think one or two of them actually turned into be real ones.
GB:	[inaudible 2:09:02] I didn't explain it very well, people mentioning on Facebook, people mentioning on Twitter land cover changes happening just on this particular site.
CM:	Well what would they say? Like what would be the thing they said? They would say "There's a fire here" maybe, maybe that one you could do.
GB:	Yes, [inaudible 2:09:19], but fire is it deliberate.
CM:	Yeah, exactly, is fire... I don't know. Certainly you could try, like it is possible to try but I don't know of anyone who's done it. It would be kind of fun to do it.
MS:	There's a lot of work on social media but it tends to be about inside cities, finding where communities work and [inaudible 2:09:37].
CM:	Yeah.
AS:	But there has been species spotting hasn't there? Like the sort of 'I have seen this species here' and that kind of thing.
GB:	And naturalists.
CM:	Yeah, so there's systems set up to do that.
AS:	But again that's really patchy, right?
CM:	Yeah, it's really patchy and the challenge isn't so much like people…well the challenge is getting people to structure their information in a way which can be easily detected by computers…
GB:	Enough signal to run it.
CM:	Yeah, exactly, exactly and I don't…just my impression is I don't think there is now but I don't know really, I mean we could try… When we tried it it was about four or five years ago and there was less people on these social networks so maybe we could try it again.
PK:	I know the Royal Bank has done a project recently with taking the whole of the Twitter Fire Hose and looking at…they were trying to verify the government statistics in the

	Philippines or Indonesia or somewhere on food prices and using the Twitter Fire Hose feed and particularly searching for mentions of in whichever Indonesian language is something to do with rice. But from what I read there was a huge number of false positives and it was really, really, really difficult to actually get…
CM:	One thing we have been able to do is detect when the sun rises from Instagram, so when people take pictures of sunsets and the sunrise you can actually…I wish I could show you the map, but we've got a map where you can see…like we sort of mapped the sunrise point to going across the world and if you look just along that edge it's like a burning edge of people tweeting saying 'Wow, look at this sunrise, isn't this sunrise amazing?' You don't even have to look in the sky anymore to know when the sun is going up. So you can detect some stuff, like that you can detect because lots of people have taken pictures of sunsets.
MR:	Good way to finish. We've run over time but before we…Roz do you want to make any announcement?
RA:	No I don't think so.
MR:	No, okay. So can I first thank our three panellists for this interdisciplinary discussion, it's been fascinating and thank you very much and especially our JRC friends for being here. There's a dinner for those of us who are staying which I think is most of us, officially in five minutes so there's time for a quick drink downstairs and then the dinner is in the next room downstairs.
	END OF AUDIO