Topic 1c - The role of citizen science in Big Data Analytics

What do we see a lot when we do land use modelling, we find that there is a big data gap, especially understanding what type of management practices are present in different places. Farmers, what type of management, of crop management, they do. Do they apply fertiliser or not? What type of crops are grown there? Still, we have very little information about the precise distribution of global crop land and crop type in particular.

So we are working on collecting data which supports this modelling exercise, but also improving the underlying global databases, which come partly also from remote sensing. And for that, we need to increase the number of training data that we improve the global mapping of land cover and land use, for example. A very important parameter which enters the global land use change models.

I think one important component which citizen science and the crowd can bring in is collecting massive amounts of data. So an individual, an expert, could never collect that much data. And as we know, in 2050, we will have 9 billion people living on this planet. So these are potentially 9 billion data collectors which have all devices, such as mobile phones, which actually can be used as data collection devices. And there is massive potential to use this crowd for collecting all sorts of type of data, also addressing the SDGs (Sustainable Development Goals), for example.

All the different SDGs can be supported with citizen science. In particular, the monitoring, but also the implementation side. So on the monitoring side, if you look at the different indicators, 269 indicators, which are used to monitor track towards addressing the SDGs, you will find that they are so-called tier 3 indicators. These tier 3 indicators, they don't have yet an established accepted methodology on how this data should be collected. And here, citizen science can play a really important role in closing those data gaps, but also in combination with remote sensing, citizen science has a massive potential in terms of collecting calibration and validation data for improved remote sensing, algorithms, and processing, for example, for the production of land cover and land use maps.

There has been a lot of criticism in terms of how good is citizen science data-- is it really good enough? And what we find in our research is that there is a big diversity in terms of quality and what we get. The quality can actually be really good in the crowd land project where we use the Fotoquest Go app and where we have a near real time quality check. We can make sure that citizens learn to improve their quality of contributions, and some citizens are much willing to improve, and we've seen very good results.

In particular with what we have seen in the development of new technologies, having all these sensors present everywhere, the potential has increased drastically. Also, with the connection to the internet, we can do near real time quality checks of the data. And the potential is enormous and endless, but also, there can be online citizen science projects. Galazy Zoo, for example, has shown it. And we do a lot of work with the interpretation of very high resolution Google Earth type of, Bing type of imagery. And we find that the quality is actually, in most cases, very good, but we also do checking of the quality internally, where we have ground control points set and where we give citizen immediate feedback on how good they have been performing, though facilitating the learning process as well.

There is a very wide range of applications where citizen science can really play a very important role, so it starts with humanitarian applications. For example, humanitarian OpenStreetMap is doing a lot of work. But also there are a lot of projects in the environment. So for example, GBIF, the Global Biodiversity Information Facility, has a lot of data. They have more than one billion records, and more than half of those records are collected by citizens. So in the field of biodiversity, there has been big progress on using citizen science for observation data or current status of species.

That is very much used also for reporting. For example, the Aichi Target 19 is important, which is on biodiversity loss. And there, citizen science data is already feeding into this reporting process of the Aichi Target 19, which is important. But there are other applications, such as beach LitterWatch and so forth, and this would be another element where contributions to the SDG could be made because that's also an indicator. Beach litter, the amount of beach litter, which is going to be monitored. That's a typical tier 3 indicator, for example, where a methodology yet doesn't exist and where citizen science could come in quite strongly.

Another, I think, very interesting implication is what we are currently carrying out in the duration of the LandScence project. There we are focusing on the perception of green space in cities. In particularly, we are looking at the parks, and what is the perception in parks where these so-called oasis of well being in cities, especially when it gets really hot, is a park sufficiently shaded and so forth. And we try to capture the perceptional component of humans where humans, of course, are particularly good at. And this also relates a lot to the SDG monitoring, access to green space, well being, the feeling people have in cities. And as we know, more and more people are moving towards urban areas, and the well being in the urban space is increasingly needed, and there needs to be improvement made in the future in a lot of places. In Europe, but also in the developing world.

One important element of doing citizen science is that you have a two-way communication platform. So basically, you are, on the one hand, collecting data, but on the other hand, you're also feeding data back to the citizen. And I think this is really critical because citizens, they want to get something back. They want to see what's in it for me, what's my personal benefit. And here, there is also a very interesting link to remote sensing because remote sensing can feed information back to citizen.

One example would be work we do currently in Mexico where we provide farmers with remotely sensed based data. For example, Vegetation Indices has infield variability. You can see, based on the pixel how your fields differ, how potentially the productivity and the yield of your field is changing. And individually gives you infield variability as one of the indicators where you might want to apply more fertiliser.

So this is an example where we use apps on the one hand to disseminate data, but on the other hand, also to collect data on crop management, which helps then to also improve our modelling. It's a so-called win-win, we can see there.

I think another application of this two-way communication platform where remote sensing is used is in an app we are currently deploying in Spain where we work with an international community BirdLife International. They have been very active in collecting data on birds, so far occurrence data, but now they are moving into threat mapping. And for that, we are currently developing this app, which

additionally to the bird observation, also allows people to tag areas which are threatened in terms of biodiversity loss, habitat loss. And this is another place where remote sensing can come in, providing hot spots of change, for example, which citizens then can document and go to and they find areas where they are requested to provide information about.

There is big potential to use artificial intelligence and citizen science data together and to blend it, because citizen science data collected by people can provide very important training data, but also, validation data of products coming out from AI algorithms. And I think this is something where humans are still needed and will be needed in the future to do the initial classification of objects, for example.

We have all these sensors now embedded, for example, in mobile phones, but there are also sensors in cars. So sensors are basically everywhere, plus we have our five human sensors, the natural sensors. And one is the seeing, and that's what we also use still in our apps where we collect land use and land cover data, where we ask people to observe the landscape and then report back on what type of land cover is present at a very specific location.

The brain is, indeed, still the best computer we have. Technology is moving extremely fast, artificial intelligence is moving extremely well at the moment, and we are all getting a little bit scared in terms of what's going to happen in the future. Are we going to be replaced, to some degree, by machines? Machines can learn, we know that by now. But I still think that the human component will always play a very important role, such as feelings, such as emotions, such as perceptions, and I think those elements will never be replaced by machines.