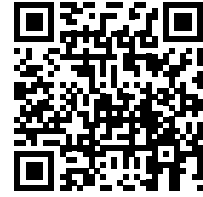


# Radical Futures: Data and Data Coalitions



## Audrey Tang

Lecture presented on July 8th, 2021, at the event Radical Futures, of ENAP.



**Lecture presenter:**  
Bruna Santos

**Abstract:** With the theme Data and Data Coalitions, Audrey Tang, digital minister of Taiwan, reported on the country's experiences in tackling Covid-19. In her definition, coalition is the use of data in a participatory way. Audrey presents that the success in tackling the pandemic without confinement was achieved by also fighting infodemic, that is, false information that spreads, especially in digital environments. In addition, the first non-binary minister posits that for the data coalition to be effective, it needs to be fast, fair, and fun.

**Keywords:** data coalition, infodemic, pandemic



**BRUNA:** Hi, I'm Bruna Santos, Director of Innovation at ENAP. Data has generated innovation in the last two decades at an accelerating rate. Economics has brought us structures that concentrate wealth and power. So to unlock the resources and potential, how can we find new ways for individuals to participate in decisions? Data is the new resource, and we need tools to calculate its value and find ways for it to give people more say in decisions. These are questions that will guide our conversation today. And to inspire us we will have Taiwan's digital minister, Audrey Tang, who will share the worldwide work done on shared control.

Audrey Tang is from Taiwan and the 1st Digital Minister. She is a hacker activist and the first non-binary minister from Taiwan. She is an influencer for many people like me, working on governance and innovation. Audrey, thank you so much for being here with us. You have the floor.



**AUDREY:** Hi, I'm really happy to be here virtually to talk a little bit about data coalitions and I look forward to your questions and interactions with the panelists. Now let's share my screen and see if the data coalition mascot appears. You can probably see a Shiba Inu, a dog.



So in Taiwan, since the beginning of Covid-19, the pandemic, on January 1st of last year, we have gone almost 500 days without a local outbreak. Recently we had an outbreak, but only for a few months, and now we are back to where we were, less than 20 local cases a day. The key to enabling this coronavirus fight without lockdown is fighting the infodemic, the conspiracy theories. This is because of the use of data in a participatory way or data coalition.

The data coalition mascot is a Shiba Inu named a dog that lives with the Ministry of Health and Welfare's participation officer. So when there is a new data coalition, the Participation Officer, who is a public servant, career public servant, and every minister, to engage the public, he goes home and takes new pictures of the dog and establishes rules of social distancing. Or that you need to cover your mouth to sneeze. And the participation of the data coalition, in which the participant must protect his or her face, so that you don't do something that the dog shows here. I brought some data coalition cases. It's fast, fair, and fun. The quick part, I mean the collective intelligence, we need to build safe spaces for the digital public and the infrastructure. This is saying instead of using the antisocial corners of social media where it's easy to spread anger, discrimination, and revenge, we need to design the interaction in a way that doesn't serve the interest of advertisers or shareholders.



That is, the civic infrastructure has to be covered by the social sector, this has been happening for 25 years. Here, what we see is a form that we call PDT that has the sharing of the ideas of people my age, because it has been running for 25 years as collaboratively managed open source. It is a student project at the university, there is no commercial interest, it is subsidized by the national budget for academia.

Because the national university has a liberal code of ethics, its very active participation in public health and other public issues are brought up for discussion.

So, as you can see, on the last day of 2019, on PDT there is a post by a young doctor, she posted something that said that in Yuhan market there were seven cases of confirmed Sars and maybe other things like that have been posted on the internet, but only that one got a number of votes and people screened the legitimacy of the message and came to the conclusion that Sars 2.0 had been released. So this resulted in health inspections, in people leaving Yuhan and coming to Taiwan.

Which shows that a civic, social sector driven space can easily gather collective intelligence without being distracted by the not so important health issues. This stems from expectations that people without digital access or digital capability are a responsibility for the state to bring connectivity, under the banner of broadband as a social right.

Infrastructure requires committed investment for anyone not connected to become connected. But even in Taiwan where we have mountains over 4,000 meters, even up there you are guaranteed 10 megabits per second for 16 euros a month for a limited data connection. This ensures that people can participate, look at daily streaming throughout the epidemic. After the message was posted on PDT, people watching live streaming interacted as a hub on things like the 1992, which is related to Covid. Over 2,000 calls were made to 1992, anyone can participate by sharing what they see, on the spot, or bring constructive criticism of the policies brought by PCC. So it's two ways and even the very young non-voting people, for example, last April there was a young man who criticized the pink masks.

“I'm a boy, I don't want to wear pink in my school, maybe other boys in my school have blue masks to wear, do something about it.” The Participating Officer took this to the Minister, immediately, and the following day, everyone at the conference wore pink masks. The Minister even said that the Pink Panther was his childhood idol, so the mask was adopted. This quick response instead of waiting 60 days, as in literally minutes and hours, is a direct response, and has more reliability. It's the basis on which fairness guaranteed by data coalitions can be done, because if people don't have a way to quickly correct data bias, people are not going to trust the data coalition and bring it into interest.

Now the part about being fair, I talked a little bit about the masking and everything, it looks like this map was built by the government, but it wasn't. Last February, some hackers, civilians, people who are independent in Taiwan, built these maps, in a few days, showing the availability of personal protective equipment, specifically medical grade equipment in the stores and pharmacies nearby. In the beginning it relied a lot on crowdsourcing, meaning people would use these maps and they could report if something was out of stock.

Bringing people to where they had stock, obviously, so they could get those PPEs, but it was all done with crowdsourcing, with a lot of risk, because if nobody or not many people participated, it wouldn't work. In a few days, and because I am part of these movements here, the GOV movement, a project that prototypes outsourcing without government participation, in every government service is something that has already been read here, and then hackers can do the same kind of service, in the same way as GOV, just changing the O to a 0, which is very similar in the address bar of your browser. People started to see a swarm of hackers doing reimagining of state digital services, so for example the mask map was an example of that, because it is always open source, which means that the creators will set aside property rights.

So we don't have to bid or tender to use that, instead we just say "ok let's do a reverse bidding, you can't get rid of this interface that people already like, but you can provide the real time data in the API's." And then you can see that when people buy these masks they use the national health card, this one that is maintained by the national health insurance, not only for natives, but for all residents and when people use this card, actually, we know every 30 seconds how many people bought, how many masks in each pharmacy.

And then we decided that instead of publishing a daily summary, we would publish a collection and then every 30 seconds it is published as a distributor with over 100 tools that put their own copy of this here every 30 seconds. And that allowed people to participate in auditing the fairness of that system, to see if it was fair or not. The people in line could swipe their card to check where and when their turn was going to be after 30 seconds and see the transactions that are being done in real time.

And that encouraged people to participate more and also to trust each other more, instead of accusing the pharmacies of lack of supplies and so on. They can see that people do a good job. It also helps people to build their own dashboards to track distribution efficiency and also allows people to see if they have biases, because initially we distributed on the map the pharmacies based on population centers. The pharmacies almost align perfectly with the population center, however, according to analysis of the dashboards built by the social sector, people's time and the opportunity cost in people's time of going to the pharmacy is not the same.

Not everybody has helicopter, obviously, and because of this we can't say that just because it's within a 5 km radius on the map it means that the person has equity of access and in fact, normally, people may even have to wait for public transportation and once they get to the pharmacy, it is already closed, out of stock, and there is a lot of that in our presentation. But the beauty of publishing this data as soon as it is collected is that no public service takes the blame for how they present the data.

It's much more about fulfilling the requirements of presenting the information by having the data before it's published. And once the pharmacies have the data they already publish it immediately, so nobody takes the blame for anything. That is why we work together and suggest to the legislators the modified distribution to make it more fair.



Then, I believe that the immediate correction of these information biases is only possible when people have equal access in real time to access the data in the public domain because everybody understands that if you don't have every district, every area, with more than 75% equally distributed masks, that would not work as a physical vaccine and everybody would suffer from that. And that's why this common purpose unites everybody. Finally, to bring in this graph, to contradict the information that we already thought we had, that is to say, a collection of data now human to human. As you can see, for example, this here makes sure that we give explanations in enough time for people to be able to access the information.

How do we make sure that we focus on exactly that information that is really there as a trend that has a higher volume of basic reproduction? Well, we can rely on measures here, so the major antivirus companies, for example, as well as Russcall – which is a company that blocks phone spam and also has a text box – that even if you have closed tools like LINE (which would be Taiwan's whatsapp), you can see that information; and if you are not too sure, you can follow it to the virus detector and see if it has that already in the group text box.

And just like a normal antivirus, it will check that against a database of clarifications and immediately it will put that clarification of service to people who are there who might want to do the same search. Based on the number of that information, the way it is shared you can quickly see. Even before it goes to Facebook or Twitter, we can already see which of that information has the most value and which is toxic and therefore would need further cladding.

We call it a public notification and a private notification, so the point is not to take everything away, but to inoculate each other to be able to share more human messages; because when people pay attention and the conspiracy theories are no longer viral, they become ridiculous and then you can talk to people in the same way, with one foot in reality. So I will now share another collection of data that was developed a few months ago. It is called SMS 1922. It is a checking system that introduces how you put a place. You can scan this QR code here and check the banner. I know many countries have done this kind of system, but the Taiwanese system is unique in the sense that it doesn't require downloading any applications and it doesn't introduce new data drivers to the mix.

It is compatible with phones that don't even have a rear camera, it is maximally affordable even. It is also fun. The trick is of course instead of solving all this on a website, you can scan a code and you will see that it will take you to your SMS application and on every phone has a camera and you can swipe it, point to the code and immediately send an SMS without going through a block. That takes about two seconds to complete before you move on to what you want and then if you have an old Android, all the messages like on LINE, they all have their own QR code scanner.

Nowadays you can use LINE or Bluetooth to use a notification display and they offer the scanning capability. So of course, it's going to cost about 5 seconds, but if you're not on the phone, of course, people who are with you can scan and if you're alone without a phone we still have the paper to help you, so it's more inclusive. Another important thing is that the text of the message besides showing a random 15-digit code is also just for epidemic control.

That's why people use it in bulk, because they know that the national health service would never use it for commercial purposes. It is forbidden by law. And the same thing for SMS check-in, that would never be used for commercial purposes, because again it is forbidden by law. People understand that they are not going to get advertising calls etc. if they participate in this kind of data collection, meaning that it is in line with the public interest and in line with public health values. But as I said before, if you don't have a camera you can type in the digits manually and then the QR code is much more transparent because it does the same thing, only here it is done with the text instead of the code.

Just by entering 15 numbers into your phone, you also complete this registration. In Taiwan, we have five major telecom companies, the operators do a data check every four weeks and they only allow legal access for example from trackers and for exposure notifications and all access is auditable and kept as a record.

We will offer it so that people can do a check themselves to find out who got the data from whom, when and also with the exposures from the exposure notification areas before they are deleted. We have a whole record of audits that allow people to have the exercise of their dignity and sovereignty of public health.

And again, if you don't think it's necessary, if we don't do this everybody's going to have to go back to pen and paper, which of course is much more risky of infecting people, but people are going to say "isn't this an overconcentration of data?" The 1922 is a short code that when you send an SMS it stores that on your carrier and nowhere else. In a way, it is not even a message, it is just a data storage offered by the agency that already knows where you are anyway and then when you send that message to a store, it decentralizes and makes a structure out of that because you have that 15 number code that is only known to the seller and the trackers. The companies have no information about that to be able to compare it to these digits that you have. It's safe, they can't do a triangulation to know where you are.

The people at the technology companies, the users and the vendors now have a way to complete this in this system, only when these four happen together can we actually have the tracking and ultimately the exposure of the notification. This also significantly decreases the risk of selling this data because it is not commercially useful. I believe the idea here is that both vendors, suppliers who need to do this for themselves, and trackers who can ask for this when there is an outbreak, can use this data because it is fast, it is very convenient for everybody, it is very fair, it is useful and used just for that, with equity, and it is fun.

If you interact with this code for a long time, for a few days, you quickly get used to it, without having to use your phone much. You just swipe to the left, enter a code, and that's it. People have gotten into this habit of using this as data collection, and they enjoy the benefits also of fighting the pandemic with lockdown, and they've gone down to only 20 local cases after a month or two. In the first moment of the pandemic, this ensured that more than a million SMS's were sent by the trackers and this would send a message to everybody in Taiwan that if we like to have universal broadband and digital coverage in basic education, if we support mechanisms like this for open innovation, this will ensure equity of access to data and then we can fight against communicable diseases, whether it is covid or any other.



**ENAP**  
•••••

**BIBLIOTECA**  
•••••