

Transcript

Title: Webinar - Social Data in Action: Emerging Social Data Governance in Europe

Creator: Prof Jane Farmer, A/Prof Anthony McCosker, Marina Micheli

Year: 2021

Audio/video for this transcript available from: <http://commons.swinburne.edu.au>



JANE: So thanks to everyone who's here. It's very exciting for us to have you here in this second webinar in our series around social data in action. And today we're super excited that Marina Micheli has joined us from Farese-- I think you've said, Marina-- in Italy, where it's approached-- well, it's just probably just gone 8:00 AM in the morning.

And Marina is going to talk to us about social data governance models in Europe. So I might just do the acknowledgment and then introduce Marina. So I'd like to acknowledge that I'm hosting webinar from the lands of the Wurundjeri People of the Kulin Nation. I acknowledge the traditional custodians of all the various lands on which you all work today and the Aboriginal and Torres Strait Islander people participating in this webinar. And I pay my respects to elders past, present, and emerging and celebrate the diversity of Aboriginal peoples and their ongoing cultures and connections to the lands and waters.

And I'm going to quickly just give some housekeeping here. So as Marina is speaking, you might come up with a lot of questions, and we do encourage you to come up with questions. If you post these questions into the chat, then once Marina has finished her talk, Anthony will be fielding the questions. And he may hand it to you to ask the question, if we can figure out who you are.

And obviously, you can use the Raise Hand symbol at that point and Q&A as well. And we are recording this section. And if you don't want to be recorded or if you have any questions about the recording, then contact Paul, paullavey@swinburne.edu.au.

OK, so-- and I want to take the opportunity to introduce Marina. So Marina, as I've said, is joining us from Italy. And Marina is a scientific project officer at the European Commission's Joint Research Centre. Her current work explores data governance for the public interest, examining, in particular, the perspectives of social actors and the emerging governance models.

She is interested in the social value of data, digital inequalities, and data power. And her background is at the intersection of media studies and sociology. Prior to joining the commission, she was a senior researcher and teaching associate at the University of Zurich in Switzerland and a postdoctoral fellow at the Department of Sociology and social research at the University Milano-Bicocca, perhaps, Italy. OK, I'm now going to hand over to you, Marina. It's super great that you're here. And thanks very much.

MARINA MICHELI: Thank you. Thank you. Thank you, Jane. I will very quickly share the slides. OK, I guess it's fine. So thank you. Good afternoon, and thanks a lot for having me here, really. It's an

honour to me to have the opportunity to present, and discuss the results of this exploratory research project called Digitranscope that ran at the Joint Research Centre of the European Commission from 2018 to 2020.

First of all, a parenthesis about what is the Joint Research Centre, the JRC, it is the science and policy interface of the European Commission, which is in charge of providing independent scientific evidence and support for European policymaking. So in the context of this project, this exploratory research project, we, among other things, we investigated on the emerging approaches for personal data sharing, control, and use that are put forward by different social actors, beyond big tech corporation.

So we were-- doing so, we identified four alternative data governance models. I say alternative because they differ from the, let's say, take it or leave it approach for an empirical data that is typical of these tech platforms. And they differ in the sense that they grant more decision-making power to other actors, beyond the problems themselves. And they foster the production on different kind of values.

So the work was moved by two main goals that you see-- first, to develop an heuristic tool to examine and better understand prevalent data practises, and to do so through a social science perspective that I will explain in a bit. And so overall, of course, the aim was to contribute informing European data policy for fostering a fair data economy. So in the slide, you see some of the outcomes of the project that are those that I will mainly present today during the talk.

OK, well, the work starts from the premises that there is a wide variety of imbalances when it comes to not only accessing data, but also in the ability and possibility to derive meaningful insights and data, the so-called big data divide. So it is widely recognised that from extracting value from big data has become a significant source of power for large players in the internet platform.

So we moved from celebrating the emancipatory opportunity of social media in 2005 to now addressing, critically, the power of big tech platforms. But in particular, what we see here and what they examine is how to go beyond the economic model of data governance that is embodied by big tech platforms.

So the fact that the rules for running data and controlling data are basically completely set by the platforms. So it's not only a matter of how platform access data, but also they control their governance. And in particular, it has been said that data subject have lost control of their personal data.

So for instance, Tim Berners-Lee, the inventor of the World Wide Web, claimed among others. But this issue is not only a matter of individual rights or social technical development of the World Wide Web-- it's also, of course, a political and economic issue. So for instance, as this is shown by the quote from Chancellor Angela Merkel in 2018, from a talk at the World Economic Forum, of the-- well, I will not go in detail about the geopolitical context.

But of course, this work stems from the premises that we see now, some sort of delineation of three different macro approaches to the digital transformation to data and that Europe is trying to take a different strategy to the digital transformation, both from the data commercialisation model that

distinguishes platform capitalism, but also from government control of citizens personal data that can all characterise China nowadays.

Two main things to be said about the European way to the digital transformation-- one is that the digital transformation is expected to be in line with European Union fundamental values of human dignity, human rights, freedom, democracy, equality, and the rule of law. And on the other hand, that Europe, let's say, does not want to be a digital colony of US internet giants. So it's working to protect its digital sovereignty by enhancing data flows across member states, across sectors, and across actors in Europe, besides, of course, also regulating the power-- the market power of these corporations.

But in the terms of data, there is a sort of shift from an open data paradigm to a shared data paradigm, which is based on really using data to create more economic and social benefits. So most of this is seen, for example, in the current European strategy for data, which is a set of action and new regulation that are being proposed, basically, this year.

It is in this context, in this economic, political, and regulatory setting that we place our work. So we ask, to what extent such data asymmetries are tackled by the practises of the social actors with the governmental data. So we moved our attention to, as I said, the social practises and also to such emerging concept in terms of, sometimes, buzzwords that, at the time of the study, circulated for defining various approaches to data science.

So to address this question, we propose a social science informed definition of data governance. With these definitions, we want to emphasise the power relation between the actors that are involved in the way the data is handled, the kind of value that is produced through data and the reuse of data and how it is distributed, and also doing so, accounting for the social technical mechanism that sustain these approaches.

Such mechanism are both technological infrastructure, but they are also principle and system of thoughts that underpin and sustain these emerging models for data governance. So basically, we understand data governance not as data management at the micro level, neither strictly only as data regulation at the macro level. We see data governance as the result of a process that develops from the practises of value actors and stakeholders from the interaction, cooperation, and negotiations between these stakeholders for controlling data and then the model that consolidate from it.

So basically, data governance is considered who can make a decision about data and who can influence the way that data is accessed and benefit from. Of course, the model that we analysed emerge from the economic, political, and regulatory setting that I've briefly described before. But the idea of this work is also that from the social practises that could be influenced to future regulation and future development of the digital economy.

So the definition of data governance that I just-- that I just briefly sketched is drawn from [INAUDIBLE] technology studies and critical data studies-- in particular, from two concepts which allow to investigate the power dimension within social complex, social technical complex arrangement.

So there is this ideal data infrastructure. So the data in other models are social technical systems in which a plurality of actors with multiple agendas and interacting with an array of tools mechanism, values, and systems, and also the idea of data politics. So the assumption that new power relation are generated in the process for accessing, controlling, and using data, and that is especially worthy, of course, to look at the perspective of some data subjects that are governed often by platform and institution, but they can also have an important role in intervening in data regimes.

So in 2018 and at the beginning of 2019, we reviewed scientific and green literature, looking for these emerging models. And so we basically looked at emerging or even implemented alternative practises for the governance of data. The models that we came up with that you can see in this slide that is taken from our publication on this are-- such models, basically, are ideal types.

On the surface, they are abstract conceptualization. So they want to emphasise certain traits in order to synthesise a phenomenon, but of course, in reality, differ for the degree of affiliation to those traits, because reality is always messier than abstract concepts. But such concepts and models, we think they can provide a basis to understand how such emerging data governance models are addressing the power asymmetry between big data platforms and data [INAUDIBLE].

These models are not the exhaustive and final pictures of all possible models out there. But let's see together a bit more of this. So I'll give an overview of these four model. Then I will move to the conclusion and the main takeaway from the study. So data sharing tools are understood as horizontal joint initiative among two or more data orders to organisation that could be both private and public entities when established partnerships for sharing, aggregating, and analysing together each other data.

Well-- sorry. They set up specific infrastructure, such as APIs or platforms. And they are built on legal contractual framework in which parties jointly define the modalities for data sharing. So basically, our data can be angled and for which purposes. But the goal of such model is mainly to produce data-driven innovation, new services, and economic benefits among all the parties involved.

Data sharing pools are the most widespread model among those that we described. Still, of course, it's important to highlight the main challenges and limitation of this model. First of all, of course, there are several obstacles that are identified in the literature to enter and establish data-sharing rules, such as the fear of breaching commercial confidentiality or competition rules of the GPR, fear of security and [INAUDIBLE] risk, and the technical resources that are needed to prepare the data set.

Then once data sharing rules are established, their [INAUDIBLE] that I said before is easily challenged by the entrance in the pool of one data order that possess very more valuable data set. So such data will [INAUDIBLE] have greater power to set the terms on our data's access and use and so recreate [INAUDIBLE] and maybe a monopolistic position with.

So another point of this model is that, as you see, data subjects are not included. They are mainly described as the recipient of the innovation and the outcomes that are generated from it. Well, an example that I place here, even if, example, as I said before, are a bit messier than an analytical

construct is-- oh, sorry, is the ways connected citizen Centre, which is a collaboration between the platform wave and local administration and municipal governments.

So as a part of the pool, basically municipal governments share their real-time construction over road closure data, and in exchange, wave share the community-collected data about traffic in real time. But one of the underlying principle of this model is that data as a market quality in the sense that more data means creation of more value.

So the main goal here is filling knowledge gaps to create more profit. So to a certain extent, the assumption are not completely different from that of the Germanic model. Instead, data cooperatives-- you are probably already familiar with this concept-- and understood as, of course, the grassroot-driven and decentralised organisation in which members of certain communities voluntarily pool their data together, and such data can only be used to rules that members approved.

So that cooperative is basically an entity that, on the one hand, facilitates the collaborative pooling and management of data, and on the other hand, it is co-owned and democratically controlled by the members to which there is a relationship of trust. So they are self-organised governance systems that allow democratic controls.

All members share equitable benefits that are produced with data. So data cooperatives are ideally generating benefits for the members of the community-- for instance, empowering data subject with the use of their health data. And at the same time, data cooperatives foster public interest or social change, such as ad research or even workers' rights and such other issues.

So basically, data cooperative is the model that support greater participation and transparency. But at the moment, of course, there are limitations. What are the limitations that there are, of course, at the moment, niche initiative with limited uptake, especially because they mainly rely on data altruism, and they do not have public incentive. And also, there are difficulties in the implementation and sustainability.

Also difficult to scale up because of the low uptake, and also, in terms of keeping democratic control and engagement of all members once they scale up. Well, the principle extends from the co-operative movement established in the UK in 19th century, but also from more recent platform cooperativism. And basically, the principle of the cooperative movement is to promote fair condition of value production in a monopolistic and democratic setting.

So what does it mean? The data cooperative are explicitly aimed to provide an alternative to an extractive model of platform capitalism. And the third model we'll call public data trust in which a public actor, such as a local administration, is the lead actor and establish a relationship of trust with citizens and manages their data or their data rights on their behalf.

So the public actor assume the role of a trustee, is guaranteed that citizen data is an ethical and securely and, most importantly, that data is used for the public interest or the benefit of society. Also, access to public data trust is open to other actors, such as research institution or assignees and so on. But it's the public body to have the responsibility of managing it.

But the main, let's say, limitation of this model is that it is a mostly similar conceptual model. There are very few complete implemented project in full, in reality. And there are-- [INAUDIBLE] there are a lack of incentives for public bodies to actually establish these. And also, they are, of course, challenged by the bureaucratic structure. And somehow, in certain cases, for the lack of resources, that they are funded by public taxes.

There is, of course, the issue of data culture within public institution, which is not always present. And a final point to raise is that many could see a danger of state surveillance and loss of privacy. So public data traffic require public putting in place trust building governance mechanism, such as citizen consultation and living [INAUDIBLE] or independent data stewardship from non-profit organisation that can act as trusted intermediary or other public oversight mechanism.

The principle, a planned principle of public data trust is that all data with a public interest component-- for instance, data collected as a byproduct of utilities consumption or a byproduct of moving through a local public transportation. All of this kind of data should be part of the public infrastructure.

So the information that it affords should produce value for citizens and society. So also, when [INAUDIBLE] of this approach are, for instance, [INAUDIBLE] and [INAUDIBLE] and also the city of Barcelona with the project [INAUDIBLE]. But the first model is called personal data sovereignty in the paper.

It's a model in which data subject individually regain control of their personal data through adult digital services. So the main actor in this conceptual model are the individuals. A key mechanism is the existence of an ecosystem of personal data spaces, which are, as I said, intermediary digital services that citizen can choose to store their personal data or collect-- use to collect and aggregate their personal data that we've dispersed it across platforms and they can use, also, to control data sharing with the third parties and for other purposes.

So data sharing will empower individual citizen because they improve their ability to make choice about their data and leverage the GDPR data possibility right. And they foster, at the same time, some economic growth in the sense that new commercial digital services are launched by SMEs and other entities. The main limitation of this model is that somehow, since personal data databases are business entities, of course they might have interest in creating a more favourable position in the market and nudging users to be attached to their platforms. So there might still be a network effect if a safeguard measure are not put in place. And also, for now, the majority of citizens, because this system already exists, they're also called PIMS, personal information management systems.

The majority of citizen do not have skills or time to maybe take advantage and really dig into this ecosystem of personal data spaces. Well, the underlying principle here is that the technological sovereignty intended as reducing, limiting the influence both of big tech platforms and states or governments in the control of technology.

But this is applied specifically at the level of the individual users. So it's mainly understood in this model as individual empowerment. So [INAUDIBLE] you see that I've briefly summarised, identified

four emerging models that, in different ways, foster more democratic forms of data governance that are different from the "take it or leave it," as I said, approach of big tech [INAUDIBLE].

Well, we think that thanks to the social science informed definition of data governance that is developed in the paper, the typology brings to the surface the key differences that are among the models, beyond maybe [INAUDIBLE] similarity. So of course, three out of this model involve data subject, all besides data sharing pools that you see in orange. And of course, [INAUDIBLE] is important step for a more democratic regime. But we can see from the description that involving data subject might have very different implications because all the models have a trusted data intermediary that gave more power to data subject.

But all have very different dynamics because these intermediary, these data intermediary could be a data cooperative that is a democratic entity that allow control and produce benefits for the members of a community. Or it can be a public body that is entrusted by citizen, that has to be entrusted by citizens to use their data ethically and for the public interest.

Or finally, for state intermediary in the last example could be a commercial service that is chosen from an ecosystem of new services that compete in the market. So basically, he would point that the same notion or buzzword that we see at the time when we began to study is indeed used to refer to data practises that are very different at the core.

So it's very different in terms of power relation that are actually supported. And moving to the conclusion, well, with our work, we pointed out that value social actors are currently thinking and developing new approaches that are alternative to the data-- economic data governance model that we have been used to in the last year.

But most of these models are in their early stages, their initiative, their pilot project. And they have challenges and limitations. Still, the models provide conceptual ideas, but not much information, probably, about the practical implementation. That has to be examined. One conclusion, of course, is that there is no one-size-fits-all data governance model to be recommended from [INAUDIBLE] for an equitable data landscape, but instead [INAUDIBLE] models-- so a combination of modest models should be proposed.

And eventually we should discuss how to mix centralised models, such as the public data trust idea, with decentralised models such as the data cooperative approach. Of course, the results show that civil society and public bodies could be key actors for promoting models that are truly alternative in terms of power emission and value production. So their role in data governance should be supported to avoid that market forces determine what say over data and how data is used.

And finally, a social science perspective on data governance is useful to unveil data power relation and differences in alternative data governance models, and in particular to promote, in the future, more inclusive data governance that accounts and balance the interest of all stakeholders, including those that, at the moment, have less negotiation power over data.

So we hope that the typology presented in the paper can work as an heuristic device for scholars and policymakers that want to design empirical research or develop theoretical commentaries that

address issues of data power for fostering the development of more desirable future while our data [INAUDIBLE].

And, well, finally, a disclaimer, of course-- in my talk, I'm not representing the commission, the European Commission as a whole, but I report the work and the thinking that were developed across the DG [INAUDIBLE] project by a team of researchers and policy officer. Well, if you're interested, I suggest maybe to have a look at the collaborative final report of the project and the other paper that was published to have a better understanding of our work. That will be good. Thank you.

- Thank you very much, Marina. We have plenty of time here and plenty of amazing people working in and across data projects, data governance, data sharing. And so I'm hoping to be able to open this up to discussion. I'll kick off with a question from Jane. Jane, do you want to ask it yourself?

JANE: Yeah, why not? Thanks, Marina. That was really great and really clear and really structured and easy to understand. I guess, I'm interested in whether the EU has, like, visages an active role in this, from your perspective, in encouraging or kind of is likely or showing some interest in supportive mechanisms around some of these models that are particularly in your citizen empowerment or included, oriented. Do you see anything like that happening?

MARINA MICHELI: Yes, of course, when one action of the European Data Strategy that I showed before is the Data Governance Act, which was-- I mean, the proposal of the Data Governance Act was released, I think, in December or late November last year. And in that proposal, there are many ideas that overlap with this work.

And it's interesting because, also, when we started the project, we were not really, at the very beginning, familiar with the Data Governance Act regulation. But across the months that we were working on it, there was more exchanges also with the DG CONNECT that was developing with the [INAUDIBLE]. And finally, when we saw-- with the final proposal of the regulation, we saw that data cooperatives and teams or personal information, management system, or personal data spaces are indeed present into the regulation because the intention of such act is to support different actors in enabling access to data and especially to support the creation of data intermediaries that can help people to share their data for the public interest.

It's really one of the goal of that act, as you maybe already know. But it's important because in that act, it emerged this idea of creating a data economy that also foster-- that is mixed in the sense that it also foster social value, and it also generates space for not only SMEs, but also for cooperatives, public interest organisations, or NGOs. And it's also interesting because mainly, the first take, one of the first paragraphs of these proposal really mentioned that this is done to limit the power of big tech platforms in terms of the monopolistic power on data.

And I think it's interesting because it is in parallel with other measures, such as the Digital Service Act, the Digital Market Act that really work on the competition issue, the more traditional economic issues is that in this one, basically, the same data asymmetry is tackled from the perspective of, what kind of other measure can be implemented to reduce data for the public interest?

Of course, it's still a proposal, and there are still question regarding how it will work, especially because the big question is in terms of the sustainability of data cooperative. These are personal

data spaces, especially when they are not economic actors. I mean, if they are economic actors, there is the question of how to add safeguard to guarantee that they do not, at the end, reach a monopolistic position or too big a position.

And if they are social actors, such as cooperatives, well, then is the question of how to help these to grow in terms of organisation and in terms of sustainability. So there are still many question on this regard. But indeed, in the Data Governance Act, data cooperative is and [INAUDIBLE] are really seen as the enabler of our sharing of data, and explicitly so as not necessary economic actors, or at least not, I mean, pursuing that kind of objective in the end. More so, well, maybe there is another question I was seeing popping up [INAUDIBLE].

JANE: Thanks, thank you.

MODERATOR: Great. Yeah, there's a question from Anne. Anne, did you want to ask a question yourself or frame it in your own--

ANNE: OK, thanks, Andrew. Yeah, I was kind of-- this is really very interesting and, I think, very important, high level, kind of overarching sorts of arrangements. But I was sort of wondering what your work is-- I'm sorry, I've forgotten your name. Michelle, is it?

JANE: Marina.

ANNE: Marina, sorry. [INAUDIBLE]. Marina, my apologies. What is your work beginning to tell you about how to engage both institutions and organisations. And I'm particularly interested in social or community organisations who are not heavily endowed with resources, if I can put it mildly, in these matters, but also recognise that it's really important to engage. And it just seems so big, so huge, so out there. So I'm sort of wondering, are you starting to get some sense of what's required to engage those more on-the-ground voices in this kind of conversation?

MARINA MICHELI: Right, thanks. Of course, thanks for the important point, for sure. Well, basically, what this word was-- as I said at the beginning, it was an exploratory research project. So the idea was to include new themes that at the Joint Research Centre, at the time when we started, were not even yet into the mainstream research agenda. So that's probably why we took very high level.

We wanted to make straight what is the field and what can be done. So indeed, this is a preliminary work for more empirical research that indeed has to engage with community and social services or other kind of actors. So the idea was to continue the work [INAUDIBLE] empirical research. But indeed, the road that we take so far was to engage with the public bodies and local administration first.

So this is-- what you are asking is indeed an area that we are thinking of including more in our research but is not, I have to say, yet maybe so prevalent in terms of data governance. There is other kind of research, yes.

ANNE: Just one thing. I don't want to monopolise anymore time. But one of the things I've learned was I was previously an academic. I work in the practise space, and I go backwards and forwards within the spaces in between in many ways. So I'm a bit-- now, I'm not muddled, but it's just more

complex. And what I'm learning is that the sooner you get people on the ground, engaged, not just kind of working through it, but the actual way of framing it, it can actually also really inform your theoretical and heuristic devices.

And so the way you think about it. So my encouragement would be to get that into the thinking, like right at the beginning, in a way. I know it's easier said than done. Yeah.

MARINA MICHELI: Thank you, I appreciate it, yeah.

MODERATOR: Thanks, Anne. I've got a follow-up question. Maybe it's related, maybe it's just an additional question. But I'm interested in how social value gets defined in these models, in and around these models, whether it's through the proposed Data Governance Act, for example, or through the models themselves or through the actors and organisations that are kind of championing the different models. And how did that feature in your work?

Is that something that has been defined, that is defined, that is being defined? Or maybe there are mechanisms for thinking about what social value is and maybe even economic value, not like does-- who gets to extract additional economic value out of reuse of data, for example, is clearly a different question to the social value that might be extracted out of reuse of data. Yeah, exactly.

Of course, since we were mostly analysing what it really tried to document and scientific paper of, let's say, conceptual or projects that were not even sometimes implemented in practise. So we were looking mainly at discourses or what kind of-- what normative system were at the ground of this different idea of controlling and sharing data.

So we didn't have a system to and assess, really, the value that was produced by such models because we didn't look at their implementation. So what we mainly look, what all kind of orientation involved for its value production that were different to what extent from the extractist model, as you are familiar with [INAUDIBLE]. So we really wanted to see, in general, what kind of understanding and what kind of goals were at the core of this model.

It was mainly looking at a discourse [INAUDIBLE] which is also this other question, how to measure the social value produced by the implementation now. But I think it's also interesting to think in the abstract and because you still point to the fact that the governance data is not necessarily aimed to increase economic profit for certain platforms.

I mean, this is also what we wanted to see. But there is besides extractivism with data because, of course, it's a simple question, but this is what we ask, what there is around beyond, because if you want to measure social value sometimes, you still, I think-- I mean, you might risk to be in the game of measuring economic value [INAUDIBLE].

I don't know what to say-- what to explain it well. But if you want to see the alternatives, also the, sometimes, [INAUDIBLE] even, the idea was to not stick to the measurement of the value produced by looking at normative understanding of value in [INAUDIBLE]. But I would be happy to know more about your take, also, on these issues because I know you have done a lot of work on these issues. Feel free to chip in with your experience.

Yeah, and I think it comes up-- like this question comes up sometimes when you're speaking with, or speaking across different kinds of organisations with different sorts of missions and when there's a bank involved, for example, that might have some sort of social good agenda that they want or community-based agenda. Sometimes it appears to be couched in those terms of social value, but essentially still about extracting additional resources for their commercial enterprise.

And of course, it becomes even more complex and muddy when it's about social enterprise where there is a kind of profit as well as social aspect to, you know, a data project or some, yeah, some aspect of data reuse or data use. So I don't have any solutions, that's for sure. But it's clearly, clearly complicated.

MARINA MICHELI: No, no, no, clearly-- OK, yeah, I understand now better, your question. Of course, we cannot take the aim to produce social value at face-- as a reality. It's just an intention. But once again, there will be the limitation of the [INAUDIBLE]. So the idea is to look at what they are planning to address.

But at the same time, we introduce this element of power relation also to have another tool to actually check what this model has supported, are supporting. And in terms of going beyond, let's say, the facade of doing social good that sometimes is actually more complex. In the other study that we conducted later on, local administration, access to private sector data for the public interest, for instance, which is related to this one.

And it was a more empirical study. We pointed to the fact that data-- for instance, data donorship or data altruism, I mean, you can give it different names. But from data corporation to local administration. So giving data for good, nonprofit, without-- no monetary exchange, was indeed not the best solution and the best knowledge in terms of power relation between local administration and [INAUDIBLE].

So because-- for many reasons, but one of the reasons, for instance, is that only certain local administration are more likely to be beneficiaries of data altruism initiative, at least in Europe, because, of course, the most famous large cities or the biggest capital are more likely to be conducted by [INAUDIBLE] to give data for free and this reinforced position. This is just an example to how, looking at the power issue, indeed, it's helpful to go beyond these claims of produce-- of doing good with data, which is more complex for sure.

MODERATOR: Are there any other questions? I have more myself, but [INAUDIBLE].

JANE: Amir's got one.

MODERATOR: Amir.

AMIR: So right now, I have a question for you related to the data equality and heterogeneous data co-op. So data co-op projects, in the way they explain, a lot of them are for kind of like individual citizens that they contribute data to a bigger pool. It not always work like that.

Sometimes data cooperative forms as a result of different organisations coming together. And each one will contribute data into that pool. Usually when that happens, then the data co-op is not

homogeneous. it's heterogeneous. They have different-- different value data added to the pool and also that often are from different-- representing different aspects of the cooperatives.

If those data cooperatives drive certain decision-making in the community, then the whole-- the role of the players of contributing data into that cooperative make a big difference on what the data insights come out of the cooperatives. Now, in your experience in Europe, do you have experience of problems like this, that the data cooperative drives certain public policies or certain community engagements?

And the data sets that we put into that cooperative drives the conversation, but there's always conflicts. And there is kind of battle of ideas at what data actually needs to go into the cooperative and how they have to be kind of connected in a way that project different perspective. This really goes back to the principle of the data-driven policymaking or policy-driven data-making. In the cooperatives, it's very, very easy to game the system. So just, I want to hear your experience with this in Europe. How does it go with this challenge?

MARINA MICHELI: K, well-- one example on data cooperative-- I mean, first of all, I would like to say that, as I said at the beginning, in reality, these models are different from what I've described because they do these ideal types. So in reality, the affiliation of these traits can be partial.

So in reality, indeed, it's more likely to find [INAUDIBLE] model, as you also, you mentioned in your example, [INAUDIBLE] in the sense that there are data cooperatives that are more towards individual empowerment, so they are maybe similar more to floor model, and data cooperatives that are more truly embedded in communities and develop from initiatives of civic groups.

So there are different-- say, there's some cooperatives can be state cooperatives in the sense that, indeed, they are other digital services that are used without much engagement and knowledge in terms of what a data subject can do and how it can democratically participate.

So that is actually the idea of this model, but to see the degree to which, in reality, the initiatives and data practises do actually conform to the model that they say that they are going toward. For instance, of course, I know that in terms of data cooperatives, there is a nice, recent article from [INAUDIBLE] setting on data co-op whitewashing and also with another colleague, ex-colleague of mine, [INAUDIBLE].

So it's an interesting article because they pointed subtly I think to what you mentioned is-- so the fact that, if there is not the involvement, the actual involvement of bottom-up community into the establishment of the data cooperative organisation, it's not really a data co-op. So it's white cooperative washing, in the sense.

And also what I wanted to say in terms of an example that comes to my mind, one example could be that the project [INAUDIBLE]. That is very much-- maybe you are also familiar in Australia. I'm not sure how you follow a European project. They call this project a decentralised system, own data ecosystem. But it was a project that ran in several cities in Europe, from Barcelona and Amsterdam, and was put forward by Francesca Bria.

And it's interesting because it was with the local administration, as I said, but it also based on the establishment of civic group at the city level that use data-sharing initiatives to make their voice heard on different [INAUDIBLE]. So I think it is a combination between public data trusts and data cooperatives, which is sustained by more clear social control, also thanks to involvement of the municipalities in this case. So this is just to say that, in reality, this much, much more complex than the models, indeed.

MODERATOR: Thanks. And thank you for-- I have a question from Kimberly as well. Did you want to jump in, Kimberly?

KIMBERLY: In some ways, you've just addressed the sorts of questions that I was raising in the chat, which is that I've read the big data and society paper. It's terrific work, thank you for that. But one of the interesting things is you're sort of painting a picture of some models, but without really-- maybe this is the lawyer in me-- without kind of stating any kind of minimum conditions for things to actually count as something like a public data trust or a data cooperative.

You started to hint at some of that in that answer just then, that there are things that make something more truly cooperative. Is there a next stage to this work that actually starts to think about, what are the kind of minimum conditions for something to actually count as that? Maybe there are, maybe there aren't, maybe that's just something that we need to keep working on.

And maybe the real answer is, nothing will conform to a perfect model anyway and we're going to end up with a bunch of hybrids. The other question that I was going to ask that, again, gets at some of the stuff you started to talk about then is what's the link between all these moves towards data sharing and the existence of intermediaries? Because what we're observing with some of the data sharing legislation in Australia is there's this idea that government will share with companies or with researchers.

But a lot of those relationships are being set up as kind of contractual relationships rather than the creation of any kind of intermediary that could manage that process. And I'm wondering, is it similar in Europe, or is data sharing being talked of in conjunction with some of these ideas of creating in trusted intermediaries?

MARINA MICHELI: Well, thank you. Thank you, Kimberly. It's a great question. Indeed, probably--

KIMBERLY: It's nice to see you, by the way.

MARINA MICHELI: Yeah, OK. Nice to see you, indeed. Well, I mean, it's a great question. And it would, indeed, make a great project. The thing is that the project was-- well, we wanted to start from the practises, you know? I know that it's maybe-- we wanted to really see and challenge what is out there.

So we wanted to look at how the practises can actually fulfil the model that they claim to perform. So we want to compare the social practises that are currently being developed in unit to thinking and implementing different ways for accessing data. What do they say about the power relations?

So we didn't want to be normative and say how they should be act. But we wanted to say, what can we learn from what is going on? Of course, that is, as you said, completely-- that is the first step from such work. And what we did was these other studies on local administration, for instance, that we then went to see what it meant exactly to access [INAUDIBLE] data, basically, just a little piece-- just an angle to look at a part of a public data trust and what many challenges and power inequality already emerged from that one.

So yes, the idea is not to put normative, I don't know, parameters, but I want to say that there is a new project that is started that wants to more precisely identify what are the enablers and what are the obstacles for creating more inclusive data sharing. So in that sense, there is something planned, but it is not something that we have already done [INAUDIBLE]. But I think, well, I think your point is very important, indeed. So yes, I see you also wrote something about [INAUDIBLE], but you didn't mention in your question.

KIMBERLY: No, thank you. That's helpful. That helps me understand where you're coming from on that. Thank you. Appreciate it.

MARINA MICHELI: Yeah.

MODERATOR: We're at the hour, everyone. I'd like to wrap up. And sorry if there were more questions. I had more questions, but maybe we can continue the conversations offline. And it's really good to connect up over these kinds of issues and topics. I'd just like to quickly shout out a reminder for the third in our Social Data in Action webinar series, which will be with Sarah Williams next Tuesday at 9:30.

So get on to the registration for that one. And I'd really like to thank Marina for your time and insights. The work is really important to the work that we're doing in the Centre of Excellence, Centre of Excellence for Automated Decision Making in Society, but also in the Social Innovation Research Institute at Swinburne. So it's really good to have your input and to everyone who came along and to the questions that I think were really pointed and shows that we're all working in this space, in this emerging space. So hopefully we can keep these conversations going. Thank you, everyone.

JANE: Thanks so much, Marina. Have a good day.

MARINA MICHELI: Thank you. Thank you Thank you. Thank you. That's great meeting you, and thanks for the great question [INAUDIBLE]. Thanks a lot.

JANE: Cool.

MARINA MICHELI: Have a nice evening.

JANE: See you.

KIMBERLY: Thank you.

MARINA MICHELI: Thank you. Thank you, really. Thanks.

[END OF TRANSCRIPT]