

Transcript of IIC/BEREC Telecommunications and Media Forum 2021

Day One - Wednesday 26th May 2021

Panel 3 – Manipulative practices: Why they’re hard to regulate, but why we should

JEREMY GODFREY: I think you’re off stage now. Can I say just how nice it is to get to talk do you all again, meet so many old friends I hope that sometime this year we will all get to meet in person rather than virtually, can I congratulate Alejandra, and Jack, and all the panellists on the earlier panels, fascinating debate and importantly for finishing on time. It is a great pleasure to introduce Helen Nissenbaum for the closing keynote today. Helen has is background in mathematics and philosophy and she’s a Professor of Information Science at Cornell University, and a leading researcher and thinker in the field of privacy and security. 15 years she was one of the creators of Track Me Not, which was one of the first browser extensions to help protect user privacy when using the web. Helen is co-author of a paper online manipulation and abusive practices that was published last year in the Georgetown Technology Review. The paper contains a masterful survey of different ways in which people can be influenced and it shows us why some of the influencing techniques are legitimate and why some should be regarded as unethical and harmful. There is a lot to learn from the paper and I believe it should be required reading for policy makers who want to protect the world, particularly in light of the way that AI can be used to manipulate users.

I was privileged to hear Helen speak about these issues at a session at Harvard earlier this year, and I am really pleased she has agreed to join us today. After hearing her talk, I hope you will agree with me, the issue of online manipulative practices is something that should

be addressed as the EU institutions consider further, not only the Digital Markets Act and the Digital Services Act which have been discussed but also particularly the Artificial Intelligence Act. After hearing Helen it became something of a mission for me to get this on the agendas in Europe during those discussions. So if I can hand over to Helen and invite her to give the talk, Helen the floor is yours.

DR HELEN NISSENBAUM: Hi everybody, this is not a group I am normally familiar with or spoken with, so I am doubly honoured to be here and to present this work. Thank you so much Jeremy for that very super generous introduction. This work on digital manipulation, I am just wondering, there is yes, okay great. I want to grab more of the screen so I can see what I'm talking about. I understand that if I want I will just ask for the next slide to follow but, right now, as Jeremy mentioned, this is work that I have done and it is work together with two other collaborators and you will see their names in a while. But Beata Russler who is a Professor of Philosophy of the University of Amsterdam and Daniel Sasso who is a Professor of Ethics and Technology at Penn State University in Pennsylvania. I think I am mentioned his, it is the first time I worked with three philosophers. Now, we are all inspired by what we are seeing, all of us have been doing work on privacy, ethics and technology, but all three of us realised we very much affected by some of what is going on, on digital platforms and generally on the web and online.

So, if I could have the next slide?

Okay, alright. So, some of you may have seen this particular instance where it came to light that Facebook told advertisers that just based on how teams were operating online, or on Facebook platform, they could see moments when teams were questioning themselves and felt vulnerable and low confidence and they thought ah this could be a good moment to advertise to them. Generally speaking, if I could have the next slide, we were all of us, ourselves, the three of us sorry, next slide. Is there a way I can operate these slides from here? Yes.

Thanks. So, we were inspired by real world cases of course there was the Facebook targeting ads to vulnerable teens. There was of course many of us are familiar with the Cambridge Analytica where people were shunted to fake news websites and manipulated

in political views. There was Uber has ways of manipulating their drivers by using means that are commonly now called 'dark patterns' where they would send certain kinds of teasers or they would show incorrectly that certain areas of a city, there was high demand and so surge areas where they could charge surge prices but when the drivers got there, it wasn't the case. Because in fact Uber, it serves Uber's interests to have a lot of drivers in certain places. So we considered all these to be manipulative.

Now we also motivated by prior research and scholarship, and we can move to the next slide, where there was a lot of interesting work done in, by people likes Alessandro Acquisti and Laura Brandimarte, Ryan Calo and so who were already looking at manipulation via digital media platform, services, workplaces. Uber as an example of manipulation of gig workers. There is quite a long trajectory of philosophical work on manipulation that's conducted by ethicists not necessarily in the context of digital media and we have our own, the three of us had been working separately on issues of profiling and privacy and so forth.

So we decided that actually, we needed a more rigorous understanding of manipulation. We felt that our own work on privacy and data practices and so forth was important but it didn't address this notion of manipulation and we didn't think that any of the other work, as important as it was, either fit the application area very well or was rigorous enough for, to justify ethical analysis or even policy analysis.

So, we, we forged ahead and we wanted to create a definition of manipulation. If we can move to the next slide. Okay. Sorry, just a little bit of, doesn't matter. But, we wanted to try and present a concept of manipulation that was rigorous, that was true to the meaning of the term. At the same time, distinguish practices that looked potentially manipulative but didn't fit that definition, and then offer the conceptual account that, it could be fuzzy, often this happens with important concepts but that it could nevertheless inform ethical judgment, policy and system design because we were all interested in that.

The little notice, this little caption, is because we realised that the term 'manipulation' is a success concept. Which means in order to point and say this manipulates, you also have to show not only what the practice is, but the outcome and we didn't feel that we had enough empirical work on the other side to make that link but we thought it was nevertheless

important to develop this concept of manipulative practice and as you see in the next slide, we have a definition of manipulative practice which is and this, I am going to, there is a lot of this worked out in the article but I am going to be very quick to say that what we found to be key to our definition of manipulative practice and I am not going to read the slide, I will assume you will read some of this material while I am talking, is that whatever the action is, it is hidden, the influence that is being imposed on people is covert. It is hidden. And it's an attempt to affect a person's decision-making processes in ways that is are, that are subliminal to human consciousness.

Now that is fine and this is, this could be a philosophical analysis but what we were really keen to explain is why digital technology and - next slide - I call it the digital technology turbo charge, because the question is, if you accept this definition of manipulative practice as hidden or covert influence on another person's decision making, and all of this by the way is a little bit contentious. But I won't go into it here, is what is it about digital technology that makes it so critical that we think specifically about manipulative practices and in this case, the definition becomes the use of digital technology to covertly influence another person's decision making and what we argue about digital technology is that it makes concealment, the concealment aspect is easily achieved. That digital technology can exploit general human vulnerabilities and a lot of the work here in dark patterns is revealing this. That digital technology exploits and discovers individual vulnerabilities and I love this Ryan Calo quote which is, "Firms will increasingly be in the position to create suckers, rather than waiting for one to be born".

So now a bit of elaboration on the next slide. This is really at the heart of the presentation about digital technology and you know, what is so lethal about digital technology? It is that on the one hand digital technology is outstanding at discovering these vulnerabilities because of the 360 degrees surveillance, multiple party aggregation, seamless and profiling. We have platforms and there's this dynamic interactivity with platforms which allows lightning A/B testing to be taking place. Are people more likely to press a red button or green button? The creation of suckers. I think we have seen how difficult and just harkening to the previous panel, it's almost impossible to convey to people what your data practices are and pretty easy to get them to consent. This is critical in the arena.

A third aspect of discovery, digital technology likes to disappear, it is a mediating technology. So this is from my colleague Daniel Sasso, it's more like eyeglasses, it's not like a magnifier. The importance is that you forget that you are doing this via technology and you fail to recognise that it is actually having an impact in the way you see the world.

This affects the way the digital technology can deliver manipulation which is, it can deliver a view of the world. This is what we saw in the US in the 2016 elections which is people were given, they were manipulated into a reality, that had them responding in a certain way. We can undermine reliable assumptions and these things can be done at scale.

So now when we take this definition and we return, - next slide - and when I presented before, I would take actually a poll to say, well what do you think about, are these manipulative practices but I won't do it today. I will tell you how our definition plays out in these different cases, so you can flip to the next slide and see how we come out on these different cases.

Now, I am happy to discuss but, I know that my time is relatively short and what, I wanted to get to what I think might be of interest to this community which is one question is digital manipulation unethical? Our answer is yes, it is. Oh sorry go to the next slide.

Then, you can flip to the next slide. What we argued and this really, I want to acknowledge Beata Russler who is really the expert, a leading expert I should say in this area of philosophical autonomy, which is the casualty of manipulative practice is autonomy. Human autonomy, and the problem with this concept of autonomy, it is considered to be some kind of grandiose concept but it really is what we are after in a free society. We need to create space for individuals but they can see that their life can be shaped by themselves, that they can understand what the factors are. They can choose values, they can plan what they are going to do, and if they are manipulated, the exercise of autonomy becomes increasingly difficult. Even the ability to vote in a democratic election and this is why for us, manipulative practices are antithetical to free societies because your brain is manipulated to see reality in a certain way.

The next slide is where we got to in the article that I am mentioned which is this Online Manipulation: Hidden Influences in A Digital World, and a lot of it is well worked out. We

realised that we were kind of sorry that we never developed the next section of the article and this is where I have been doing some thinking which is should digital manipulation be regulated? In my view, what I am going to, just in the few minutes I have remaining today, say the answer to that is yes, it should be regulated. I don't know how many of you are familiar with the Norwegian Consumer Council, they have several really incredibly useful reports, some of them for example on mobile location tracking and so forth. But they also have done some really important work on dark patterns, the way if you have an Amazon Prime account. It is really hard to shut down your Amazon Prime account. Sorry, could you flip, so we can pause for 10 seconds and then to the next slide.

So there are, the Council, Norwegian Consumer Council is strongly critical of the practices, focusing specifically on dark patterns, but I want to talk more generally about regulating manipulative practices which I think, I don't know how it is in your Europe, but in the US it is going to be an uphill battle so could we go to the next slide. I will explain why in the US, it's going to be difficult because, because it goes against some of the truisms that guide regulation. First of all, the harms which I talked about in terms of autonomy, they are unclear. They are not like, "I lost a thousand dollars" or "look I have a physical injury", it's this a notion of being decision maker, choosing your own values, sort of freedom to define your own path and do things for reasons that you determine to be the right reasons to take one action or another.

Second of all, there's a very strong sense in the US and again I am curious about Europe, that anything that constrains other people's behaviours, so, don't stop the advertiser from saying whatever they want to say or other individual liberties. So any attempt to regulate manipulative practice might be cut down because of argument like this. Finally, there is a tendency to say oh, this is paternalistic, you know, people you don't have to protect people against manipulative practices, they can do it for themselves, they can see through it.

Next slide. Now, of course I mean there is some kind of Pollyana-ish optimistic notion and we know that sometimes we have to protect people against themselves like, us against ourselves. So when you are, there is a lot of evidence showing the addictive nature of some of the technologies - next slide - we may, here it is back to this issue. So these are

you know, these are challenges and then I was, I was really fascinated. You can go to the next slide. Laura Brandimarte who I mentioned over here, had given a talk in my class and she was talking about how in fact one can manipulate, and here the marketing literature has plenty of examples showing that a familiar face much more easily can sell you stuff. What can AI do for you? Here this is her example. We are able to merge faces. So this isn't Alessandro Acquisti selling you whatever it is that this guy is trying to sell you. But because we merged and you have a model on the one side, we have merged Alessandro Acquisti's face with this paid model who has agreed to have his face used. Now, you have someone who kind of resembled Alessandro, maybe has that sense of familiarity and you're going to say, oh this guy looks trustworthy to me, yes, I will buy those whatever those things are.

So there are ways in which manipulation can get to the root of how we think and decide, that I believe are problematic. So I am going to - my final slide is the next slide in which I am going to argue, that the reasons nevertheless to regulate against manipulative practices because of the loss of autonomy that is not only harmful to individuals but also challenges the foundations of free societies. That when we, that in consumer regulation, we do regulate against the deceptive practices but that deceptive can mean a lot more than outright falsehoods and in the US, there tends to be a really narrow interpretation of what deceptive practices are. Also, regulation does happen for severely asymmetric situations, in this case, of both knowledge and power. So we do have regulation against exploitative labour practices or contracts, individuals are exposed, machines can be very opaque and also these digital practices create novel sources of vulnerability that I think we need to study. I don't believe that the GDPR and traditional consumer practice is going to be enough to meet these challenges.

Thank you. That is it for me and I am delighted, happy to take comments and questions.

JEREMY GODFREY: Thanks very much Helen, another tour de force. I'm going to maybe kick off the questions, but can I just say to people in the audience, do type in questions and they will come up on my screen and I will put them to Helen. Helen I was very struck by your last example, which I don't think - I may have forgotten – but I don't think you

mentioned the idea that I might be sold something by some kind of composite face that looks a bit like someone I trust? But not so much like someone I trust that I would absolutely notice the manipulation of the image. That's I think a great example of the covert manipulation that you are talking about.

I was going to ask you a little bit about the AI regulation that has come, the proposal that has been put forward in Europe, and I was quite excited when I saw the leaked draft from our friends Politico and rather less excited when I saw the final version. They do talk about subliminal techniques in the explanatory note but they really limit that in the regulation to techniques that can cause psychological or physical harm to people. So nothing like the loss of autonomy that you were talking about.

The Commission's view is that other manipulative practices can be dealt with by data protection, by consumer protection, by digital services legislation. Their argument is that legislation ensures that people are properly informed and have a free choice not to be subject – I'm just reading here from the proposal, have a free choice not to be subject to profiling or other practices that might affect their behaviour.

I wonder if you might say a little bit more about why you think that data protection and consumer protection you know, is not enough and why you think something more specific ought to be used?

DR HELEN NISSENBAUM: Well, I don't - I am not sure, you know it could be that when they use this concept of subliminal, I mean it could be that we meet somewhere in the middle. If we are saying subliminal, often it is like the old-fashioned ideas if you are watching a movie and they flash below your consciousness an image of Coca-Cola, and at the intermission a lot of people go and buy Coca-Cola, this was the advanced hack hidden persuaders. So you're willing to -if they tie that to a narrow definition of harm, then we're in trouble and when we say, harm we say people acting in ways that are against their benefit or interest. So if we have a really restrictive, but if we can expand subliminal to, into 'covert' you know, in a way, we could achieve some of that, some of what I am hoping for when we, when we look at working at against the kind of covert influence that gets underneath our ability to decide. So that's on the one hand, and if we can enlarge what we

mean by harm, that also can be useful for me, in other words we could go that direction and it could be useful.

How do I feel about the data protection? I mean here, how much time do we have? There is no way you can inform people. I do not agree with informed consent. I think it is a bad approach to data protection. You cannot inform people sufficiently to assume that they are going to make a good decision and even what we are seeing now, when we see these little pop-up dialogue boxes, oh we are using cookies and what is that? So, I am not optimistic about that at all. I think that we, it's time for people who are expert to step in and give a lot more guidance on this. So it is partly yes, maybe we can achieve something with subliminal - I don't think that that aspect of GDPR will you know, is getting us far.

JEREMY GODFREY: Okay. Thanks Helen, I think we have got probably time for one more question. I remember you talking about how manipulation, even if it's potentially in somebody's interests, is also unethical. Could you say a bit more about that and maybe give an example of the sort of thing you had in mind there?

DR HELEN NISSENBAUM: Well, I mean it's about, this is the kind of bad paternalism where we maybe push people to do what is good for them but really if you want to develop your full humanity, each of us needs to develop our decision-making capacities and when we're undermining the decision-making capacities, I think we are damaging what is essential about humanity. I mean it's much longer, but I wanted to quick maybe we could get in another question?

JEREMY GODFREY: I don't actually have any more questions from the floor. But I think, I think yes, you are kind of saying that even if it was, we were trying to promote say the take up of COVID 19 vaccines, you would say that it would be just as bad to pop-up in people's timelines you know, to kind of suggest that they take the vaccine at a time when they are particularly worried about COVID rather than to make use of that. So to time of the intervention at that time. That is as bad as suggesting they vote for political party or buy a product?

DR HELEN NISSENBAUM: I think that's an interesting point because sometimes what we, there are layers of manipulative practice and 'nudge' is, I don't know if you are familiar with

this work on nudge, which is to say, when we, if you want to organise a cafeteria to encourage people to take healthy options you make sure that you put the salads and so forth at a certain level and you then make them bend right down for that cream cake.

That is manipulative in a sense but as a society, and in our moments of clarity, we have acknowledged that this is good for society. So that when we impose on people nudges, that have generally speaking been approved by thinking autonomous human beings because we know we have weakness as well, then I think that is a different kind of manipulation. It's not a manipulative practice that I would have difficulty with and there are many sorts of safety issues, like wearing a helmet and so forth. So it has to do with making the decision at a higher order that yes, we know we are weak of will and scared and reasonably so. We need to, we need to insert a certain kind of encouragements, we have already agreed on that. In the case of some of the one-sided manipulative practices that are on the web, they are beneficial for some parties and at the expense of other parties, there hasn't been a societal discussion, so I think this needs to be had.

JEREMY GODFREY: Yes, the equivalent of the cafeteria rearranging the shelves around my weaknesses in order to sell me something expensive as opposed having the salads at the same level for everybody.

DR HELEN NISSENBAUM: Yes, yes.

JEFFERY GODFREY: Helen that's been a great, a great exposition, a quick canter through it, but I really would encourage people to, to read the paper, read your paper because I think when I look at the legislative proposals and I look at the current legislation, this is a big area in which people's autonomy and interests are affected and I think as you said Helen, we can't rely on the existing legislation and at a time when Europe is just considering three major pieces of legislation in the digital area, the Digital Markets Act, the Digital Services Act and the AI Act, it would be a huge shame if these issues were not thoroughly debated and the right place was found for some regulation on this area. So, anyway, thank you very much Helen.

DR HELEN NISSENBAUM: Can I just, just wanted to add one thing. First of all I would love everybody to read the article but be aware that the work of the policy recommendations are not there and therefore we really need everyone on board here,

but at the same time there's a student of Beata Russler at the University of Amsterdam in his PhD dissertation has done a really fantastic job. One of the arguments he has made, because he is very familiar with the EU regulatory framework, is that it really, the manipulative practices cannot be addressed, and so I am going to provide a reference to that and make the connections, but that work really is really awesome.

JEREMY GODFREY: Thanks very much Helen. I have been asked, this is the final session of the day, so the IIC asked me to wrap-up the session as well and again thank everyone who spoke, everyone who participated and everyone who asked a question, and to welcome you all back tomorrow morning for the next session of the day. So I can wish you everyone a very good lunchtime if they are in America, and a good evening and sweet dreams if you're still awake in Australia. Thank very much everybody and goodbye.