

Open data: a shared responsibility

Interview with **Dr. Btihaj Ajana**
by Hein Schellekens



Individuals are increasingly sharing their personal information - such as medical health records – on online platforms and communities. In addition to seeking support and exchanging knowledge, the provision of personal information can support third parties in the development of new services or therapies for the benefit of all. Some have therefore claimed that sharing data can contribute to a sense of solidarity. In this interview, Dr Ajana argues why the current application of open data promotes a narrow conception of solidarity. She stresses the importance of informed consent, distributed ownership and shared responsibility as pre-conditions for the promotion of a more structural and inclusive notion of solidarity.

Open data evolves around the idea that some data should be freely available to everyone, without restrictions or mechanisms of control. What are your thoughts on this concept?

I am very sceptical about the notion of open data. The hype about what can be derived from open data can overshadow the possible downsides. My main problem with it is that people are often not consulted about how their data is shared and used. To take the case of health data: the UK government is trying to work with industries to make health data available for other purposes such as research. But as a patient, you don't consent to repurpose the data that is collected when you consult a doctor.

Why do you think such consent mechanisms are currently lacking?

They are not easy to implement. We do have the General Data Protection Regulation at the moment, which is a really welcome addition to the regulatory framework around data use. But it is not enough. The problem with the law is that it is always lagging behind technological developments. Open data will only work if we have mechanisms in place that give people the choice to either opt in or opt out. You should not be opted in by default. That is problematic from both a legal and moral perspective. In my opinion, it is only fair that if I produce a piece of data, it should belong to me. And I should have a say in terms of how it is being used. There is even a debate at the moment about putting a price on personal data and paying people for producing data in the first place. The other issue with open data is that, even when anonymized, it can have unintended consequences and be used in ways for which I did not give permission. For instance, my data can be used to develop differentiated insurance premiums by categorizing between different health risks, and perhaps thereby reinforce income inequality. It is not just about what can harm me as an individual but also how my "small" 'personal data can feed into "big data" to create differential categories, risk levels, etc. that can have material effects of many groups in society.

Some believe that sharing one's data for the benefit of the common good, such as medical research, can enhance a sense of solidarity between people. Assuming for a moment that people do have the opportunity to give consent, would you say that open data can foster solidarity in such cases?

Of course, there are some positive elements of sharing data between entities. Platforms such as "PatientsLikeMe", where people share their medical information and treatment experiences with fellow patients, enable people to support one another. So there certainly can be some positive elements in terms of support and solidarity. At the same time, these practices exclusively foster a narrow and, at times, simplistic understanding of solidarity. Sharing a post or commenting on a link is seen in itself as an act of solidarity. But that is a very reductive understanding of solidarity. Moreover, we have to recognize that not everybody has access to such platforms or has access to the means by which they can actually make sense of their data.

How is this “narrow” understanding of solidarity different from how we typically define solidarity, in terms of unity or agreement among individuals with a common interest or within a group?

The key terms to pick up here are “common” and “group”. Solidarity, in essence, brings with it the question of community itself which is also about the sharing of something in common and belonging to a group. But this is precisely the issue here: the fact that one has to have something in common with others in order for solidarity to manifest. This leads to substantialist understandings of solidarity that may risk creating forms of exclusion for those who don’t necessarily share something in common with a group. Instead, a broader, deeper and more inclusive approach to solidarity should be developed, one that goes beyond commonality. As such, we should not restrict solidarity to the mere act of sharing medical information with one another, so that people can come together regardless of whether they share the same health issues or not. And provide people with the means to understand what happens to the data they produce and share, in other words, “the social life of data”. In order for solidarity and a community to work, we need to have all these elements in place; informed consent, education around the social life of our data, inclusive notions of belonging. I am not undermining or underplaying the importance of the solidarity that is being promoted by healthcare platforms such as “PatientsLikeMe”. At the same time, I am weary that they are only promoting a narrow sense of solidarity. And sometimes, unconsciously, they even risk putting people up against each other. If you share your data, you are seen as a good person. But for someone who is still holding onto privacy, that person can be perceived as selfish. So-called data philanthropists are increasingly being pitted against those who still hold on to the value of privacy.

At the same time, one could say how this example illustrates how solidarity is established in the first place, namely by defining the borders of one’s community.

Exactly, that is the heart of the issue: the definition of solidarity itself. It can also be an act of solidarity to argue in favor of privacy, by making people aware of the possibly harmful consequences of sharing your data, especially at a time when, for instance, health services are being increasingly privatized resulting in power being transferred to corporate companies, and when there is a lack of legal systems in place that can provide effective remedies for people who have been harmed by data use and sharing. Unfortunately, the issue is often not perceived as such because the overall debate is overtaken by this philanthropic zeal towards open data. Solidarity, in my opinion, should not negate personal privacy nor individual choice as to whether to have one’s data shared with others or not. An “inclusive solidarity”, so to say, would in fact respect the diversity of opinions and positions rather than dissolve this diversity to totalistic and pre-determined understandings and formulations of “public good”.

One can observe a trend within the scientific community towards open data, which is generally perceived as positive development. Do you think we should distinguish between different sorts of data and value their public nature accordingly?

Absolutely, ethics is all about the context. Health data is one of the most private sorts of data because it directly relates to the body. But if data does not concern individuals per se, research data for example, it is considered somewhat less problematic. So, I think the notion of open data should be looked at in light of the context. Who has produced the data, who will benefit and who will be harmed by open data. Those are the questions that should ultimately drive our arguments in favor or against open data. And they are context-based questions.

Do you think health data holds any particular features compared to other types of data?

It does. Some philosophers therefore differentiate between indexical data and bio-data. Indexical data is data that is not directly derived from your body. An example would be a picture. It is a representation, as opposed to something that is derived directly from the body, such as your DNA. However, this distinction is not always so clear-cut. For instance, when I am using my Oyster card to travel on the London underground, I produce information about where I have been. This information is just as personal as the physical data. It still concerns my body, moving through space. I think we need to broaden our concept of what should be considered as personal data.

Would you say digitization and open data are changing the normative concept of solidarity and if so, how?

I think they definitely are. They are redefining solidarity in terms of very specific values. Let us take the example of social media. Sharing or liking a post can be seen as solidarity. We may seem connected, but the underlying social fabric is increasingly fragmented and consequently prevents solidarity. As Sherry Turkle has put it, we are “alone together”. In addition to the online space, we need actions in the actual world too in order to make that solidarity manifest in a more concrete way. I am not necessarily positing a sharp distinction between the online world and the offline one, as the boundaries between the two are increasingly blurred. Nevertheless I do believe we should not restrict our acts of support and reciprocity to online platforms only.

Reciprocity and solidarity are closely related concepts. Nowadays, data from wearable technologies is used to offer personalized health insurance products, for example by rewarding healthy behavior with discounts. Do you think this particular employment of data, whereby risk is transferred from the collective to the individual level, has an effect on legitimation of reciprocity and insurance solidarity?

The problem I perceive with all these developments is that the underlying problems are not dealt with. Technologies of personalization and prediction for instance will always advantage the people who are at the center of society. If you have a secure position in a society, you will benefit from having personalized offers. You will benefit from services that are effective, efficient and convenient for your own use, according to your own needs. But people at the margins of society do not benefit from these developments. These technologies create a world whereby some people are benefiting from these developments, because they get personalized and convenient services, whereas others have to endure the exclusion that comes along with this process of categorization.

So open data could actually increase existing inequalities?

Absolutely, because I think the very aim of data is categorization, especially in particular contexts such as insurances or border management. Data intensifies the process of categorization. In order to create a fair and just open data society, we need to address the assumptions and biases that underpin society itself. Notions of solidarity, openness and sharing will remain empty concepts unless we address the underlying social problems. Often technology is seen as a panacea, as this magical silver bullet that will solve all societal problems. But it does not. It aggravates certain problems. It might help in certain situations, depending on the context.

It might be wishful thinking, but I think people need to actively refuse to be part of this categorical design of society. That goes back to the problem of solidarity. If we restrict solidarity to online spaces, to sharing things on Facebook, we will never have any real change where it really matters, in the actual spaces that define so much of our everyday life.

So there should be an element of effort in the definition of solidarity?

Absolutely. And the mobilization of various actors with different backgrounds. It is also the responsibility of citizens themselves. It is a shared responsibility. Solidarity should be seen as something that is shared and distributed responsibly. People need to get informed about how they can use data and how they can benefit from it. Of course we can have open data, and we can anonymize it. But we will still have the same entities that are benefiting from it. The companies that have the skills to use that data. Who owns data, who benefits from it and who is harmed by it? Ultimately, these are the three questions that need to drive any debate surrounding open data and solidarity.



Dr. Btihaj Ajana is an international scholar in the fields of digital culture and social analysis at King's College London. Her academic research is interdisciplinary in nature and spans many areas of expertise including the critical study of new media technologies and identity systems, digital health and self-tracking technologies, museum developments and curatorial processes, immigration and citizenship governance, and the socio-political and ethical dynamics of surveillance culture.

Hein Schellekens works as a project manager for W.I.R.E., where he analyzes the normative, political and economic consequences of long-term trends in society.