Paul-Olivier Dehaye
(PersonalData.IO, MyData & HestiaLabs)

In order to address data power, it is necessary to understand how market power is derived. First, data is pooled to derive insights. The quality and exclusivity of these insights define their commercial value. Volume of data collected tends to affect quality, while variety tends to affect exclusivity. Due to the asymmetric nature of the relationship created, both of those factors will actually negatively impact the trust an individual might have in the entity processing their data. From there, in a second stage, a feedback loop is created whereby small competitors operating under the same premises are unlikely to gain sufficient foothold (due to suspicion and direct competition). This, in turn, makes the entire space a winner-take-all market to those trying to confront it with the same logics. However with this size and dominance comes fragility: these giants' business models, or rather their scale, tend to produce externalities, which current operators are not well equipped to address.

This opens a gap, wherein much smaller projects focused on directly addressing these externalities do have an opportunity to find a niche and effect change. These projects can take advantage of a few distinct levers. Most likely they can survive by expanding the use cases they address, which suggests a first focus on variety rather than volume. In order to achieve this, it is essential that they maintain trust throughout their operations, through technical and governance means. While these projects might be thought as a unit, we have found that such projects regroup different functions that are best located in the hands of different entities.

Governance and technical development, for instance, need not happen within the same actors. In fact, for reasons of cost and trust, it might be highly desirable to split the two functions. On the technical side, approaches through open-source should be favored, with copyleft licenses as well as data standards being almost essential tools to reduce costs. We personally believe it is necessary that some commercial actors also be involved on this technical side in order to provide continuity. The democratization of this technology should in turn favor the emergence of autonomously governed entities using those tools. Should these take the form of trusts, cooperatives, foundations, etc.? It is still unclear at the moment, and it could very well be that several layers of this governance intertwine, particularly from the perspective of an ecosystem. A necessary condition seems to be that through a first layer individuals should be able to granularly direct both how their data is used, and for what purpose.

We leave with a practical question. Over the past two years or so, Uber drivers throughout Europe have had some good success at challenging the company's handling of their data. Leveraging rights of access and portability, they have made efforts to better understand how the platform takes decision about their work. They have recently obtained some success in a Dutch court. Their explicit goal is to pool this data and e.g., address issues of discrimination. I am a firm believer – but definitely in the minority – that

such a reverse engineering effort is possible from a technical standpoint, if the data protection law is properly leveraged and applied. In any case these Labor Union 2.0 efforts also raise many questions of governance: should such a structure's audience be "all the Uber drivers of the world"? There are many alternatives: all the for-hire drivers of the world, of a given country, of a given city. Or all the gig workers, or all the Uber/UberEats workers? All of them would have overlapping purposes related to workers' rights, but for different subsets of workers. But the data of concern could also have other uses, for instance, addressing issues of pollution in a city. This vast set of possibilities is thus dizzying in three directions: what data and whose data gets pooled into these collectives, and for what purpose. We, thus, don't think there is one answer for the governance question: many of them should be tested, in some form or another, but it will be inescapable that both the data and the purposes have a natural structure that will be reflected in the governance hierarchies that get created for these collectives.