"Privacy gives us the freedom to be ourselves"

Those words were uttered not by a philosopher or a politician; they came from the mouth of Mark Zuckerberg as he stood on stage at the annual Facebook conference. Behind him was an unambiguous statement of Facebook’s new self-proclaimed values: “The future is private.”

Facebook is not alone among tech giants with this newfound privacy focus. Sundar Pichai, the CEO of Alphabet, stated in 2019 that “we think privacy is for everyone – not just for the few” and the company has begun rolling out new features to allow users to manage and delete their data across Alphabet’s myriad products. Has Silicon Valley caught the privacy bug?

**Why privacy, why now?**

In the aftermath of events such as the Cambridge Analytica scandal of 2016, in which a political consulting firm exploited user data to micro-target political ads for Donald Trump and Brexit, there has been a popular awakening to the reality that our every move – online and, increasingly, offline – is tracked, catalogued, and used to sell ads or train sophisticated machine learning models. As popularized by *The Economist*, data has become ‘the new oil’, with massive amounts of value being extracted by companies that have access to the biggest reserves and the most talented minds. Of the ten most valuable companies in the world in 2019, seven were primarily technology- or data-based.

Meanwhile, the subjects of the data – everyday people – have grown increasingly unsettled with the state of affairs. In a poll from the Open Data Institute, only 5% of U.K. respondents trusted social media companies to handle their data ethically. One click of the ‘accept’ button at a time, we have slowly ensconced ourselves in a world of tracking and measurement where we ourselves are the product. This new regime threatens to severely undermine the fundamental civil, political, and economic freedoms that form the bedrock of liberal democracy.

In this essay, I will explain how freedom is threatened by the current lack of a coherent data governance regime. I will then consider government regulation as a possible solution before introducing a more comprehensive solution: data trusts.
How freedom is compromised

In 2015, riots broke out Baltimore, U.S.A., in response to the death of Freddie Gray – a black man who died in police custody. Unbeknownst to protestors, a company called Geofeedia was helping Baltimore police to target and even arrest activists using social media data. As the Internet of Things (IoT) becomes more adopted, devices like smart speakers and home cameras, as well as fully-connected ‘Smart Cities’ will collect mountains of data that could be used by advertisers and law enforcement alike.

Since the Cambridge Analytica scandal, Facebook, Google, and other large online advertisers have restricted what political advertisers are able to do with user data. However, the fact remains that 43% of American adults read at least a portion of their news on Facebook. This news is heavily curated based on user preferences, contributing to the formation of ideological echo chambers, exposed only to the ideas, opinions, and facts that confirm their pre-existing beliefs. This has greatly aggravated political polarisation and made the polity vulnerable to populism and false news.

Data concentration also threatens market freedom by reducing business’ ability to compete with the tech giants. Incumbent businesses are able to leverage huge stores of data to out-compete potential challengers. As stated by Kai-Fu Lee, former researcher at Apple, Microsoft, and Alphabet,

“the more data you have, the better your product; the better your product, the more data you can collect”.

Taken together, the hegemony of tech giants over our personal data will continue to infringe on the civil, political, and economic freedoms on which liberalism relies. To restore these freedoms, we need to comprehensively rethink data governance.

Government regulation as a possible solution

Recent regulations such as the EU’s General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA) have jointly imposed new responsibilities on users of data. These include the delineation of clear data protection principles and the requirement that consent be “freely given, specific, informed, and unambiguous.”

Although a step in the right direction, these regulations face several key limitations. First, the globalised nature of the internet means that regulations from any particular jurisdiction may be unenforceable where a breach occurred. Second, even if, as the GDPR insists, regulations apply regardless of jurisdiction, this is hardly a desirable feature, since the government with the most restrictive regulations would set internet norms, leading to a world where the ‘netizens’ of one country are de facto governed online by a government of another.

Another, more extreme course of action might be for regulators to break up tech giants under antitrust law. If data is the new oil, then perhaps Alphabet, Facebook, and Amazon are similar enough to Standard Oil in 1911, which was broken up into dozens of smaller companies. However, the antitrust approach neglects the fact that the quality of the algorithms that power Google search and Uber ride pairing depend on network effects, meaning that breaking up a company like Alphabet (Google) might significantly decrease the quality of the service provided. Furthermore, breaking up the largest firms does not ensure that the new smaller firms will implement necessary data governance. At best, a dramatic breakup of tech giants would only be a partial, temporary solution to a much
deeper problem. At worst, it could deplete political will, decrease the quality of online services, and stifle AI research.

**Solution: data trusts**

I propose data trusts as a solution that avoids these pitfalls. A data trust is defined by the Open Data Institute as "a legal structure that provides independent stewardship of data." It accomplishes this by mimicking the traditional legal structure of a trust in order to safeguard vast amounts of personal data in the interests of individual users, rather than technology companies.

There are five elements that comprise a data trust:

1. **The asset:** the data and/or code that generated it.
2. **The grantor:** the party who initially has the data and gives it to the trust.
3. **The trustee:** the individuals in charge of managing the data. Trustees have a fiduciary duty towards the beneficiary, meaning that they are legally required to act in their best interests.
4. **The beneficiary:** the users in whose interest the trustees act.
5. **The purpose:** why the trust exists. If the purpose is subverted or ignored by the trustees, then the beneficiaries are entitled to legal recourse, such as suing in civil court.

In short, "a grantor gives an asset to a trustee in order to ensure that it fulfills a purpose that is in the interest of the beneficiary." In the case of a data trust, this would mean that all companies would need to give personal data over to one of several independent trusts to store on its own servers. Any subsequent use of the data would need to be verified by the trust as in line with its purpose and the interests of its beneficiaries.

**Benefits of a marketplace of trusts**

**Flexibility and negotiation**

In comparison with governmental legislation, data trusts have several key advantages: (1) they are accountable to and can affect only those individuals whose data they control; (2) they can adapt to changing technology much more rapidly, since they are hyper-focused on issues of data collection and use; and (3) they allow for a plurality of values systems with respect to privacy and openness.

An optimal scenario would be one in which there was a marketplace of data trusts, each allowing different levels of data sharing. When choosing where to put their data, users could choose trusts whose values were most closely aligned with their own. These values could be described broadly, such as ranking the relative value of “privacy” and “personalization”, leaving the fine-grained interpretation to the more data-literate trustees. This is much more effective than the current regime of ‘all or nothing’ opt-out systems for most online platforms, where the only choice an average user has is to consent to a huge list of possible data uses upfront or to abstain from a service entirely. Instead, a data trust with control of a significant database could effectively negotiate with individual companies to change their terms and conditions in exchange for access to the trust’s data. For instance, based on the widespread backlash towards use of personal data for political targeting and ideological sequestration, it is highly likely that most data trusts would partially or fully restrict companies from using data for those ends.

**Data sharing for solving problems**

A data trust could link together many diverse datasets, which could then be employed to solve problems in the public interest. For instance, genetic data could be combined with health app data to aid in discovering new risk factors for rare conditions. A 2018 poll found that 73% of people would share their personal data in an effort to improve public services if there was a simple and secure way of doing it. A key reason for the failure to maximise public value in the status quo is the lack of institutions that are sufficiently trusted to make judgements in the public interest. Data trusts, in contrast,
guarantee that trustees are accountable to users through the fiduciary relationship.

**Levelling the data playing field**

Instead of data being concentrated in the data-centers of few tech giants, a system of data trusts would guarantee a fair level of access for any company that could demonstrate it would use the data responsibly. This would break the cycle of more data leading to better models and yet more data, since the data collected by one company could be used by its competitors. However, companies would still be incentivized to contribute data to the trust because if they were to withhold personal data, they could have access to the entire database revoked. As well, since an individual data trust could pool data from many different sources, the temptation to gain access to such information might make companies willing to hand over their data to such a trust.22

**Funding model**

How will data trusts secure the capital they need to maintain huge servers, process galaxies of data, and handle countless requests from prospective data users? Although governments and non-profits should help with covering start-up costs, more sustainable would likely be a ‘pay-for-access’ funding model, where data users (companies like Amazon and Alphabet) would pay for the use of the data. Companies may be willing to do this because (1) all data would be held through trusts, so they would have no other option; and (2) the trust would be providing a certification service, demonstrating to users that the company was, at all times, using their data responsibly.17

At the moment, nearly all the (monetary) value of the digital economy is concentrated in the hands of companies. With a ‘pay-for-access’ model, data trusts could return a portion of revenues to beneficiaries as a dividend, thus allowing individuals to share in the value created by the digital economy.

**Steps to implementation**

Already, we have seen pilot projects of data trusts around the world. The City of London piloted three data trusts for urban data, wildlife management, and the food & drink industry,6 while in California, the Silicon Valley Regional Data Trust has successfully governed data sharing among several school boards.6 In Toronto, Canada, a data trust is currently being discussed for implementation in Alphabet’s Sidewalk Labs initiative.23

The regulatory framework is beginning to take shape. Already, articles 77-79 of the GDPR gives individuals the rights to portability, erasure, and access with respect to their data, all of which are necessary for data trusts to become a reality.79 The next steps could include laws that required all companies to store their data with a trust, as well as the creation of ‘regulatory sandboxes’ to allow trusts to experiment with a variety of structures.24

**Conclusion**

In the coming decade and beyond, we will undoubtedly see the acceleration of the trend of increased data collection in every aspect of our lives. If we are to defend the civil, political, and economic freedoms that are essential to liberal democracy, then a new, comprehensive data governance regime is sorely needed; one built on an understanding that data is a common good that merits collective ownership and control. Only then will we ensure that the future really will be private.
References


16. Hardinges, J.W., Peter; Blandford, Alex; Tennison, Jeni; Scott, Anna, Data trusts: lessons


23. Barter, A.Y., Andrew; Ryan, Alex; Coleman, Beth; Finley, Charles; Makris, Chris; Greenwood, Joe; Stolarick, Kevin; Verner, Kristina; Austin, Lisa; Girard, Michel; Zolli, Mikayla; Sud, Sasha; De Lara, Sergio; Bauer, Sharon; Banks, Timothy; Matos, Vanessa, A Primer on Civic Digital Trusts. MaRS: Toronto. Available from: https://marsdd.gitbook.io/datatrust/.