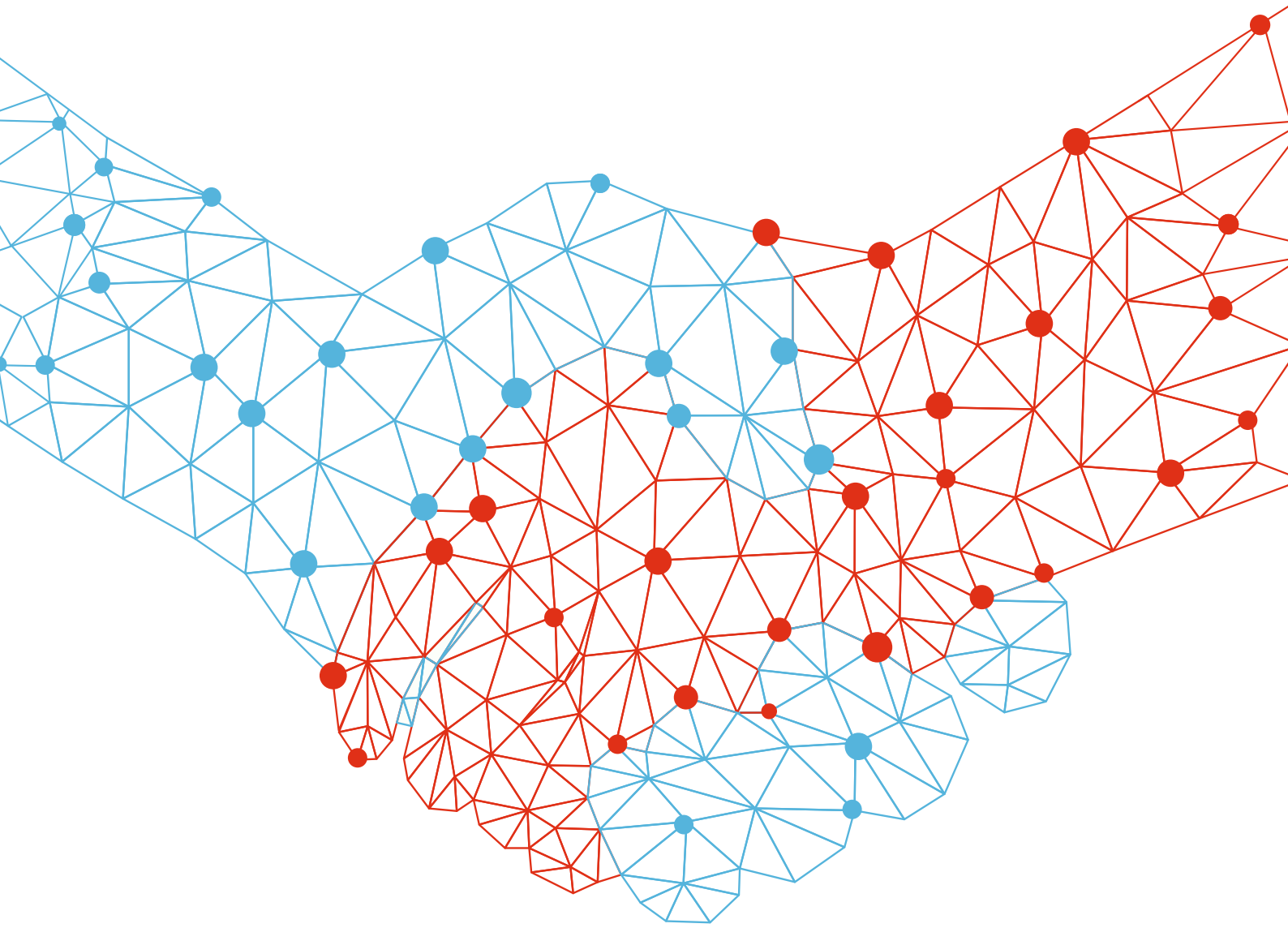
Data trusts in Germany and under the GDPR

Anouk Ruhaak



Published by

ALGORITHM WATCH

With support from

mozilla

ZEIT-Stiftung Ebelin und Gerd Bucerius

Contents

Executive summary

The ongoing collection of personal data and the use of this data in automated decision-making systems (ADMS) raises questions about the effectiveness of current approaches to data governance.

Background:

Personal data is collected and used to automatically predict what content we are most likely to engage with. There is much to be said about the accuracy of such predictions, but while their benefits may be unclear, their harms are certainly not.

The vast majority of use cases of data-driven, automated decision-making systems recorded in the Automating Society 2020 report tend to endanger, rather than help, people. A systemic lack of transparency makes it difficult to research and, consequently, provide an evidence-based judgment concerning the overall contribution of such systems to society. However, there is ample indication that most of the time, this opacity is exploited precisely to prevent scrutiny.

Governments around the world are taking steps to protect their citizens and their right to privacy. Depending on where you live, you may now have the right to allow or prevent the collection of data about you. These rights are vital, but on their own, they are insufficient to protect individuals and society against the worst harms of mass data collection and use. In addition, they do little to promote the collection of data for uses that benefit us.

We need data governance models that emphasize both the individual and collective risks of data sharing and help us decide when and how we want to make data about ourselves available.

In this report, we consider three major points:

— The shortcomings of our current approach to data governance that mainly focuses on individual data rights.

— How the reduction of the collective harms of data sharing and the simultaneous activation of collective benefits of our data require approaches to data governance that rely on greater democratic control over our data.

— The specific role of data trusts as independent intermediaries with a fiduciary duty to act on behalf of data subjects.

Key recommendations:

— We recommend greater clarity on the various legal uncertainties that currently undermine the creation of data trusts and similar data intermediaries.

— We argue for the creation of a new legal role: an intermediary that can represent the data rights of data subjects that would have to adhere to a strict set of safeguards and duties.

— We recommend a series of trials within the safe confines of a regulatory sandbox.

The need for collective data governance models

Whom do we want to decide how data about us and where we live is collected and used? In this section, we explore why it is not enough to give individuals rights over their data. Instead, we argue for the need for new models of democratic data-governance that bring to the foreground individual and collective *agency*.

/ Big data does not care about you

In 2017, ProPublica, an investigative journalism platform in the US, learned that people in minority neighborhoods in the US pay higher car insurance premiums[1], even though insurance companies reportedly do not collect data on ethnicity. But they don't need to. The only information they need to discriminate on the basis of ethnicity is to know the ethnic profile of someone's neighborhood. Aggregate statistics thus become a proxy for individual identifiers. Facebook's advertising algorithms follow a similar logic. Facebook does not need to know your age, or whether you go to school, it can infer that information from your likes and dislikes. For instance, if most people who like Miley Cyrus have revealed themselves to be 15-year-old high schoolers in the US, then it's easy for the platform to profile anyone else who likes Miley Cyrus as also belonging to the same demographic. This profile then determines what advertisements they're most likely to click on or what posts will keep them on the platform.

That is the problem with decision algorithms based on large datasets. They don't care about you individually, but, instead, they use data about you to make inferences about people who are, in some way, like you. As a result, you are as much - if not more - affected by other people's decision to share data, as you are by your own decision to do so. This is one reason why protecting individual privacy by anonymizing data sets - that is removing identifiers that allow data to be traced back to an individual - does not remove all risk.

/ Does personal data exist?

Most so-called personal data is in fact relational, interpersonal. We are social creatures and, as such, data about us often, if not always, also describes our relationship to other people or things. The data about your COVID status does not impact you alone; it impacts those around you as well. Your communications are relational by definition. The Instagram picture of your lunch is also about where you choose to eat or your local food supply. Your genetic data not only describes you; it also describes your family members and even as of yet unborn family members. Who gets to make decisions about the collection and use of these categories of data? By reducing these types of data to 'personal', we fail to see the larger social context from which they arrived and we forget to acknowledge the impact on that social context that sharing (or not sharing) that data may have.

1 Julia Angwin Mattu Lauren Kirchner,Surya, 'Minority Neighborhoods Pay Higher Car Insurance Premiums Than White Areas With the Same Risk', *ProPublica* <https://www.propublica.org/article/minority-neighborhoods-higher-car-insurance-premiums-white-areas-same-risk?token=CpfrqXaMuR8UnJ5-FryRkVwuKl3C98Ae> [accessed 1 December 2020].

/ Consent needs to be meaningful

The other problem with our reliance on individual consent is that it's hard for each of us to understand the consequences of making data about ourselves available to others. In addition, the sheer amount of decisions this forces us to make can be overwhelming. In the words of privacy philosopher Helen Nissenbaum: *"Proposals to improve and fortify notice-and-consent, such as clearer privacy policies and fairer information practices, will not overcome a fundamental flaw in the model, namely, its assumption that individuals can understand all facts relevant to true choice at the moment of pair-wise contracting between individuals and data gatherers."2* And so, instead of making real decisions about what terms and conditions we do and do not want to agree to, we end up mindlessly clicking 'OK' on the various windows that pop up as we open websites.

And even if we did have perfect knowledge and ample time to make these decisions, consent is meaningless if we cannot meaningfully opt-out. The choice between opting into having data collected about us, or losing access to our social graphs is often not a real choice. The power imbalance between the individual presented with lengthy terms and conditions and the platforms they rely on for their daily functioning is simply too great. And as we all know, without the option to say 'No', our yesses become meaningless.

/ Collective agency

In light of the externalities of an individual's decision to share data, as well as the power imbalances and information asymmetries that complicate such individual decision-making and consent, we need to look for alternative data governance models that emphasize individual *and* collective agency. At the same time, we need to ensure that those most affected by the current mass data collection practices are protected.

To this end, we will explore alternative data governance models that allow us to bundle our rights and collectively bargain for better terms and conditions; To decide together when we want our data to be collected and how we want it to be used.

/ The societal value of data

In the above, we have described the problems of our current approaches to data governance that are based on protecting individual rights. We have not yet touched on the question of whether mass data collection is a valuable endeavor in the first place. In a talk for RightsCon 2020, Shoshana Zuboff remarked, "*Right now most discussions already begin with 'data' — data ownership, data portability, data accessibility, and so forth — and my view is that once the sentence begins with 'data' then we've already lost," "What should become data in the first place, that is where the line has to be drawn."3*

There are many uses of data that should arguably just be outlawed, or otherwise severely limited. Face recognition, employee surveillance, and micro-targeting are all good candidates for abolition. Simply put, if the ends are unjust, harmful, or wasteful, improving consent mechanisms or implementing collective data governance models will only work to legitimize the harmful practice.

That said, there are legitimate cases where large-scale data collection and storage can help us coordinate responses to crises, monitor progress towards a common goal (e.g., climate change), provide insight into the evolution of a virus, or inform our opinions and policies.

In June 2020, the New York Times published a feature4 detailing how data had been instrumental in identifying the source of a cholera outbreak in the 19th century. It described how a statistician named

2 Helen Nissenbaum, *A Contextual Approach to Privacy Online* (Rochester, NY: Social Science Research Network, 2011) <https://papers.ssrn.com/abstract=2567042> [accessed 1 December 2020].

3 Soshana Zuboff (2020), RightsCon Online 2020: https://www.youtube.com/playlist?list=PLprTandRM961s376IH8TmvbThcxjfPNwL

4 Steven Johnson, 'How Data Became One of the Most Powerful Tools to Fight an Epidemic', *The New York Times*, 11 June 2020, section Magazine <https://www.nytimes.com/interactive/2020/06/10/magazine/covid-data.html> [accessed 1 December 2020].

William Farr was able to trace the origins of the outbreak to a contaminated water basin, using data on new incubations and where they occurred. The article, no doubt inspired by the ongoing COVID pandemic, makes a strong case for the need for data to help us curtail the spread of infectious diseases. This message is echoed by governments all over the world as they seek to understand and contain our current pandemic.

However, in 2020, citizens have grown weary of the mass collection of data about them by governments, fearing that the data collected will be used for other, more nefarious uses both today and in years to come. One fear is that data collected by one arm of the government will later be used by another, for an altogether different purpose. This fear materialized recently in the UK when news broke that police departments had used data from the NHS COVID-19 app[5] to locate suspects. But even if governments were to legally mandate the compliance of citizens with data collection efforts, a lack of trust in the proper use of such data would likely undermine both compliance and the quality of the collected data.

The same can be said for the categories of health data needed to cure diseases and develop new medications. There is a clear social argument for the collection of such data, but the risks are equally sizable: in the right hands, your DNA data may help find a cure for skin cancer. But the same data could also be weaponized by insurance companies looking to weed out high-risk individuals.

Reversing our climate crisis requires data as well. In order to understand the magnitude of the problem we're faced with, as well as the progress we're making towards tackling it, we need data on environmental processes, pollution, CO2 emissions, etc. And, it's not enough for such data to merely exist; it needs to be accessible to all relevant parties: internal reports show that as early as the 1980s, oil companies Shell and Exxon were well aware of the adverse impact of our reliance on fossil fuels on the planet[6]. It failed to make these documents public, thereby limiting the ability of the general public to make up its own mind and for activists to make their case.

/ Access and trust matter

These examples show the social value of data, but they also emphasize the importance of the context in which data is collected and used, who has access to it, and, most importantly, who gets to make these decisions.

In the case of COVID, we might have fewer hesitations about sharing data about ourselves if we could trust that the institutions in charge would only use it for the stated purpose. What guarantees would need to be in place to provide proper assurances? These questions become even more salient in cases where society's need for certain types of data outweighs the individual's need for privacy. Importantly, the question is not just about whom we do and do not trust to make such decisions, but who is trustworthy, and how can we guarantee their enduring trustworthiness.

Equally urgent is the question of access. The same piece of data can be useful in the hands of some and dangerous in the hands of others. And then there are the many cases where data (that is used mostly for exploitative means and, arguably, should not have been collected in the first place) can still prove valuable as a way to get insight into those in power. For example, many people dislike having sensitive data about their financial situation collected by credit scoring companies. However, when AlgorithmWatch asked people to donate data to scrutinize the practices and data analysis of Schufa, the pre-eminent private credit company in Germany, more than 4,000 people submitted their scores that they had obtained

5 NHS Covid app (2020): https://apps.apple.com/us/app/id1520427663

6 Benjamin Franta, 'Shell and Exxon's Secret 1980s Climate Change Warnings | Benjamin Franta', *The Guardian*, 2018 <http://www.theguardian.com/environment/climate-consensus-97-per-cent/2018/sep/19/shell-and-exxons-secret-1980s-climate-change-warnings> [accessed 1 December 2020].

via data subject access requests.[7] The collected data revealed data quality problems and elusive scoring results that may affect more than 60 million people.

And, of course, journalists generally have long relied on Freedom of Information laws to obtain data about government processes and elected officials that allows them to shine a light on corrupt practices.

7 OpenSCHUFA: https://openschufa.de/english/

Alternative data governance models

The problems outlined above, as well as the need for reliable data to help solve society's problems, have given rise to a plethora of alternative data governance models. Most of these are modeled after existing governance arrangements and collective bargaining institutions, in the form of commons, unions, cooperatives, and trusts. What they share is a focus on collective governance, and an attempt to create structures that allow for greater trust in the way data is collected and shared.

In data commons, such as Open Street Map[8], people pool their data and collectively decide how and when to make it available to each other and third parties. This model resembles the natural resource commons studied by Elinor Ostrom[9]. DriverSeat, a cooperative in California, used the cooperative model to allow drivers to collect and aggregate data on rideshares[10]. This data is then used to help unions and policymakers to make better decisions on behalf of gig workers in the rideshare economy. Each driver gets a vote and, if data is licensed to a third party, they share in the profits. Trade unions and consumer watchdogs have a long-standing history of bundling the rights of many to advocate on their behalf. Their experience in organizing collective action provides a useful model for how we can start to tackle power imbalances when it comes to the governance of our data.

Finally, data trusts allow data subjects to hand over the rights to their data to a data steward (the trustee), who will make decisions about data collection use on their behalf. This model protects both individual and collective rights while giving the steward the power to bargain on behalf of the pool of data subjects.

These data governance models are not merely examples for us to draw from as we design future versions. It is often the case that issues around data collection and use intersect with other economic and social activities previously organized through commons, cooperatives, and unions. In those instances, these institutions and organizations can themselves start to act as data stewards. For example, German banking cooperatives hold not only the capital of their members but also vast amounts of data on their financial transactions. Therefore, it is conceivable that they would allow their members to vote on the collection and use of that data. Indeed, a research group at MIT is proposing that that is what they do. Another example is that of trade unions that are increasingly fighting for the data rights of the workers they represent.

/ Key principles

What happens when a group of diabetes patients decides to make data about themselves available to health researchers? And not one single researcher, but various researchers looking into different aspects of diabetes in an attempt to come up with a cure, better treatment, or measures that could prevent diabetes from developing. We could imagine each patient storing data about their personal health and making

8 Open Street Map: https://www.openstreetmap.org/#map=6/51.330/10.453

9 Elinor Ostrom, *Governing the Commons: The Evolution of Institutions for Collective Action*, 1st edition (Dallas, TX: Cambridge University Press, 1990).

10 DriverSeat: https://www.driversseat.co/

that available to whomever they please. But doing so would require them to make numerous decisions. In addition, data that they share about themselves may also contain information about their family members or environments. Finally, who this data is shared with could potentially impact all diabetes patients, not just the patients who decided to share. For instance, if they allow a pharmaceutical company to collect data about them, which contributes to the discovery of a cure, that data would likely be patented and sold on with high margins.

Instead of everyone making these decisions on their own, they could decide to pool their data and elect a board to help them make decisions about what data should be collected and who it can and cannot be shared with. They may even use their collective bargaining power to set conditions on data collection and sharing agreements. For example, they could negotiate a fixed price on drugs that are created with help of the data.

How does this collective, be it a data commons, cooperative, or union, ensure that whoever they place on the board continuously acts in their best interest? What can the board decide on, and what happens when they go against the interests of the collective? And how would any decision-making allow for the diverse sets of needs and interests within the collective? After all, we cannot assume that every diabetes patient has similar values and interests. To start answering these questions, we define some of the core principles to guide our data governance design.

/ Collective agency

Simply put, when a system governs us, we should be able to govern it. We should be able to co-determine the rules we are subject to. This also implies that any governance system needs to be able to negotiate the, at times, conflicting interests and needs of those it serves.

/ Shared benefit

Decisions about our data should weigh our collective benefits as well as minimize the risk to both individuals and groups.

/ Accountability & transparency

We should be able to hold those who execute our decisions and/or decide for us accountable and keep them to their words. In order to hold our representatives and agents accountable we need to know what decisions they make and receive updates in a timely manner.

/ Dispute resolution

A well-working system of dispute resolution should be in place that allows each of us recourse when we have evidence to suggest our collectively established rules have been violated.

Together these principles help us identify key legal, governance, and technological safeguards that should be present in any data governance design.

/ Data trusts

Over the last two years, data trusts have gained attention as a way to allow individuals to collectively enforce their data rights against data controllers and have a greater say in what data is collected and used. Especially the legal foundations of a data trust provide crucial safeguards against the abuses of power outlined above.

Take the diabetes patients described above. Let's say they opted to create a data commons in which they pool their data rights. Every year they elect a board to make decisions about who has access to the pooled data. They may further set some ground rules. For instance, they could decide that the board can never grant access to commercial entities. Now, as we have asked before, how do we ensure that the board members act in their best interest?

A data trust would add a legal relationship to this arrangement. Data trusts are special aspects of a legal trust, a popular legal relationship within common law that allows one party to hand over assets or power to another, to have these assets or powers managed on behalf of a third party. For instance, parents could place a house in a trust, to be managed by a trustee on behalf of their children.

With a data trust in place, instead of handing their data rights to board members of a commons, they would give them to a board of trustees[11]. The board would then have a fiduciary duty to make decisions on behalf of the beneficiaries of the data trust. These could be the members of the data commons themselves, but could also be expanded to include diabetes patients that did not participate in the data pooling. Such an extension would hold one key benefit: it would encourage the data trustees to ensure that data collections that are made available to researchers are representative of the entire population of diabetes patients, not just the set that decided to pool their data. Moreover, the data trust would serve a specific purpose that would bind possible actions of the trustees. In this case that purpose could be 'to cure diabetes'.

Thus, with a data trust in place, we can rest assured that, over time, the objectives set out by a data commons are guaranteed. Members of the commons could still vote (or otherwise negotiate) on key decisions, as well as decide on who should be put in charge. Moreover, the data trustees, imbued with the power to decide on behalf of the beneficiaries would be well-placed to negotiate with data users (e.g., social media platforms) on their behalf, thereby acting as intermediaries between data subjects and data users.

/ The role of fiduciaries

At the heart of a trust lies a fiduciary duty of undivided loyalty, as well as a duty of care. The former means that a fiduciary can only act in the *sole* interest of the beneficiary of a trust, regarding a specific purpose. That means, for instance, that those who have a vested interest in using our data cannot also be imbued with the fiduciary duty to look after it on our behalf. Only an independent third-party, without any interest in the data, can do so. Much in the same way that a doctor, who has a fiduciary duty to look after his or her patients, cannot take money from a pharmaceutical company to prescribe certain drugs and not others. Or, a lawyer cannot represent parties on opposing sides of a lawsuit. A conflict of interest would prevent them from doing so.

The duty of care is a slightly weaker duty to deter negligence. In the world of data, that could mean that a data trustee ensures that the data under its control is properly secured and that adequate measures have been taken to make sure that whoever gains access to this data will use it in accordance with the agreements set out by the trustee.

Why are these fiduciary duties so important? When we entrust a third party with power over our data, it is often hard to evaluate how they will use this power. Monitoring is costly and requires a significant amount of expertise. In fact, we specifically entrust a fiduciary with our data because they have both the time and expertise that we lack. In the words of legal scholar Tamar Frankel: "*controlling the fiduciaries in the performance of their services and the use of entrusted assets may undermine the very usefulness of fiduciary services.*"[12] For instance, the very reason we visit a doctor when we feel ill is that we lack the appropriate expertise to diagnose ourselves and act on that diagnosis. But of course, in the absence of full monitoring and in light of the asymmetry in expertise, we need other ways to protect those who are subjected to the decisions made by the fiduciary. A duty of undivided loy-

11 Sylvie Delacroix and Neil D. Lawrence, 'Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance', *International Data Privacy Law*, 9.4 (2019), 236–52 <https://academic.oup.com/idpl/article/9/4/236/5579842>.

12 Tamar Frankel, 'Unifying Fiduciary Duties in the Common Law and the Civil Law Systems', in *Kapitalismus - Eine Religion in Der Krise II* (Nomos Verlagsgesellschaft mbH & Co. KG, 2015), pp. 102–21 <https://doi.org/10.5771/9783845249094-102>.

alty helps ensure that the trustee continuously acts in our interest and faces significant costs should they fail to do so. If a trustee violates their fiduciary duties, a beneficiary of the trust can hold them liable in a court of law. Of course, this presupposes that the beneficiary is able to determine that their interests were violated.

/ Additional safeguards

While this report focuses on legal mechanisms to ensure those in charge of our data have our best interests at heart, it would be a mistake to rely on fiduciary duties alone. Instead, we might think of fiduciary duties as providing us with a last resort: if other safeguards fail, they allow us to challenge the decisions made by trustees in court. That means a robust data governance model requires additional safeguards to ensure that our collective and individual interests are continuously upheld and we are able to have our voices heard before any laws are broken. Here we discuss some available options.

/ Certification

Certifications provided by existing trusted institutions and based on thorough evaluations of a data trust could help us, the general public, make better decisions about where to put our data, or whom to entrust with our data rights. That said, we are a long way away from understanding what good models look like and, therefore, what to base a certificate on. Thus, before we create certification bodies, we need to create spaces for safe experimentation with various models and approaches.

/ Transparent decision-making

We may not have the resources to monitor a trustees' every decision, but we should have enough transparency to be able to step in and hold a trustee accountable if they start to act against our interest. As a minimum, we should mandate each data trust to share a log of daily decisions about what data can be collected and accessed, by whom, and for what purpose.

/ Election/representation

Who do we trust to make decisions about data, and how do we decide who to trust? One way to ensure underperforming trustees are removed from their positions - without involving a court - is through regular elections by beneficiaries of the trust. In addition, beneficiaries of a trust could be asked to vote on key decisions, or general rules. The level of granularity of the decisions that beneficiaries should vote on will likely depend on both the size and nature of the trust. For instance, diabetes patients could be expected to be relatively involved with their diabetes data trust, as their health might well depend on that. Other data trusts that govern less critical data, may not motivate similar participation levels. They might rely less on direct voting and more on long-term representation (as is often the case in credit unions) or participatory governance models such as citizen juries13.

/ Interoperability

Instead of one data trust to rule them all, we need many data trusts for many different purposes. That also means we should be able to move data between data trusts. Or at the very least, we should be able to take our data out of a data trust and start a new one ourselves. Such an interoperability requirement also means that we need data and consent protocols, such that data and consent statements created in one data trust can be understood by another.[14]

/ (Financial) independence

In line with the fiduciary duties of the trustees and to avoid conflicts of interests, the trust should never rely on the income from selling or licensing data for its continued existence.

---

13 Patel, R. (2020) "The foundations of fairness for NHS health data sharing" https://www.adalovelaceinstitute.org/blog/the-foundations-of-fairness-for-nhs-health-data-sharing/

14 McKenty, J. and Ruhaak, A (2019): "Data Portability, Federation And Portable Consent" https://www.digitalcommoners.org/privacy/consent/2019/07/03/data-portability.html

/ Recourse/dispute resolution

What happens when we disagree with the decisions taken by the trustees? How are we able to update data about ourselves held by a trust? When groups of people make decisions together, disputes are bound to come up. In some cases, such disputes may warrant legal action, while in others a procedure for filing complaints may suffice. In either case, the beneficiaries of a data trust should be able to avail themselves of a number of dispute resolution mechanisms, ranging from informal mediation to legal interventions.

Data trusts and German law

So far, data trusts have mostly been considered within the context of common law, where trusts are a familiar concept. However, as the interest in data trusts is growing, civil law countries are exploring these instruments as well. In Germany, the Data Ethics Commission and the Commission for Competition Law 4.0 recommended that „the feasibility of setting up data trustees be examined"[15]. Furthermore, in its key points, the Federal Government announced a data strategy to analyze „what contribution trusted data spaces and structures of data trustees can make to increasing the voluntary sharing of data."[16]

The question is to what extent the concept of a data trust as a legal fiduciary relationship can be sufficiently translated to civil law jurisdictions, specifically Germany. In addition, it is still an open question whether the rights obtained under Europe's General Data Protection Regulation (GDPR) can be transferred to a data trust.

This chapter considers both questions. First, we will explore to what extent GDPR rights can be represented by a third-party and/or transferred to a trust entity. In addition, we will ask under what condition such representation or transfer of powers can come about and what legal duties and liabilities this places on the data trustee. Secondly, we will discuss whether the creation of a data trust itself is possible under German law and, if not, what other institutional forms we may consider. We will focus specifically on the creation of fiduciary duties under German law. A longer version of our legal analysis may be found here[17].

/ Representation and transfer of GDPR rights

In order for a data trustee to make and execute decisions about our personal data, we need to be able to transfer our data rights obtained under the GDPR to a data trust. This section explores whether this is legally possible. In addition, we will also explore which GDPR rights could be represented by a third party. This scenario differs from the first in that the third party would not act *as* the data subject, but performs specific actions on their behalf. Finally, we will explore whether it is possible to give a third party a relatively open mandate to make decisions about an individual's data on their behalf through the creation of a contract.

/ Full transfer of GDPR rights

A full transfer of GDPR rights would mean that the data trust itself would become the de facto data subject, and all associated powers would be transferred to the data trustee. According to our legal analysis, a full or partial transfer of GDPR rights to a third party is not possible under the GDPR.

15 Bundesministerium für Wirtschaft und Energie, 'Ein neuer Wettbewerbsrahmen für die Digitalwirtschaft' <https://www.bmwi.de/Redaktion/DE/Publikationen/Wirtschaft/bericht-der-kommission-wettbewerbsrecht-4-0.html> [accessed 1 December 2020].

16 Federal Government Germany, Data Strategy 2019: https://www.bundesregierung.de/resource/blob/997532/1693626/e617eb58f3464ed13b8ded65c7d3%20d5a1/2019-11-18-pdf-datenstrategie-data.pdf

17 AlgorithmWatch (2020): Legal feasibility of Data Trusts in Germany, https://algorithmwatch.org/publication/gutachten-data-trusts-dsgvo/

First of all, the GDPR clearly states that any rights can only be held by a data subject, which is defined as an identifiable or identified natural person to whom the information contained in the personal data relates (Art. 4, No. 1, GDPR). A data trust could not, therefore, act as a data subject. Secondly, a blanket transfer of data rights would mean that the data subject removes any protection of their personal data and, in the process, waives fundamental rights, as guaranteed under Art. 8 of the Charter of Fundamental Rights of the European Union (CFR). The GDPR guides the waiver of fundamental rights to a degree, in a controlled way, in order to balance the interests of the data subject with those of the general public. A blanket transfer would undermine this balanced system that the legislature deliberately specified.

Furthermore, a partial transfer of GDPR rights is also problematic. Data subjects can give permission to process data through contract or consent (Art. 6 Para. 1 lit. a and b). No one else is allowed to give permission to process a data subject's data by means of consent. There are good reasons why this is the case: with a partial transfer of rights, the data subject would no longer be able to determine what data is processed and by whom. For instance, if someone transferred the rights of their health data to a data trustee, they would no longer be able to decide how this data is used themselves - all decisions would have to be taken by the data trustee. This goes very much against the intention of the European legislator.

/ Representation of GDPR rights

The creation of most common law trusts requires at least an initial transfer of rights, which as we have observed above would violate the GDPR. However, we may consider the case of representation, in which an agent would agree to represent a data subject, but not obtain a right of its own (so to speak: become the data subject themselves). In this case, the data subject would have to conclude a contract with the third-party that instructs it to safeguard their interests and to exercise their data subject or defense rights and, within the framework of this contract, extend the agent's power of representation according to § 164 BGB. This option is less problematic, but comes with a significant degree of legal uncertainty.

The question of whether representation of a data subject's data processing rights is a legal possibility depends on whether the GDPR allows for consent by proxy. This question is a source of fierce debate. On the one hand, one could argue that a data subject has the right to decide for themselves whom they want to speak for them. In addition, European Law generally acknowledges the legal possibility of representation. However, opponents maintain that representation of data processing rights is not provided for within the GDPR. They further point out that there is no provision for member states to determine this on their own. Finally, one could argue that consent by proxy would undermine the GDPR's consent mechanisms.

This picture is further complicated by Art. 8, which regulates special representation for children, as well as Art. 80, which looks at the right of representation related to defensive rights (e.g., the right to be represented in court when GDPR rights have been violated). Given that both these cases were specified within the GDPR, one could come to the conclusion that other forms of representation are, therefore, not allowed and that individual member states do not have the authority to regulate legal representation. Then again, one could just as easily argue that the GDPR does allow for broader representation and merely emphasizes these specific cases here.

All in all, while the representation of data rights may be legally justified, it is also the subject of significant legal uncertainty. The recent proposal for a Data Governance Act by the European Commission suggests that the Commission itself holds that representation of GDPR rights by third parties is not possible, as it states that: "It is important to acknowledge that the rights under Regulation (EU) 2016/679 can only be exercised by each individual and cannot be conferred or delegated to a data cooperative"[18].

18 European Commission (2020). Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act), page 18.

/ Representation through contract

Another way to design representation would be through a contract between the data provider and the agent, in which the data provider would contract the agent to process their data. In such a case, the agent would be entitled to process data on behalf of the data subject *and* allow others to do so as well, as long as it falls within the scope of the contract.

This construction is also not without its problems. For one, there is a concern that, 'Art. 6 para. 1 lit. b) GDPR[19] should only cover such processing in which the data processing is ancillary to the contractual services.' That means that data processing itself should never be the subject of the contract, as this would be all too easy a way to circumvent the specific legal requirements for consent. Another open question is whether personal data itself could be the subject of a contract.

On the other hand, one could argue that the GDPR does not raise such restrictions. While it is true that consent places special requirements to protect the data subject, it is also true that as an essential part of signing the contract a person needs to be able to make an informed decision. Moreover, German contract law arguably offers sufficient protective measures to protect persons concluding a contract. These include the condition of good faith (§ 242 BGB).

It is, therefore, legally justifiable to design the representation model in this form. However, given the untested nature of this option and the legal uncertainty surrounding it, it is equally possible that it will be struck down in court.

/ Data trusts in Germany and under the GDPR

Given the absence of trust law in Germany, a data trust as previously envisioned would not be possible in this jurisdiction[20]. However, we might be able to create an institution that mirrors some of the most important elements of a data trust. This section specifically discusses how fiduciary duties are created under German law and then goes on to explore what other legal entities could be created to hold these duties.

/ Fiduciary duties

Where do fiduciary duties come from? In common law jurisdictions, prevalent in most Anglo-Saxon countries, fiduciary duties are distinct from obligations that can be created through contract. Common law assumes that the signatories of a contract stand on equal footing and there are no notable power asymmetries between them. That is, they enjoy equal bargaining power (ref Frankel). In contrast, when it comes to fiduciary relationships common law assumes power asymmetries. As a result, there are special provisions that govern this relationship that are not captured in the written agreement between them, such as a duty of loyalty and a duty of care. Additionally, when fiduciary duties are breached the remedies are far more severe than would be the case for a breach of contract (ref Frankel). This crucial difference explains why, in common law jurisdictions, we favor data governance models that are based on fiduciary relationships rather than contracts.

However, the situation is different in civil law jurisdictions, such as Germany. Here, the line between fiduciary duties and contracts is not as clear-cut, and, in fact, fiduciary duties either result from contract or specific laws.

In German contract law, when one person can legally force another person to do something, a 'relationship of obligations' (Schuldverhältnis) is created between them. Such a relationship can come about in two ways: through contract or by law. The latter occurs when someone harms someone else, possesses something that belongs to someone else, obtains something without a legal reason, or does something

19 European General Data Protection Regulation (2016): https://gdpr-info.eu/

20 In light of the analysis above, the creation of data trust in Common Law countries subject to the GDPR may be equally problematic.

in the interest of someone else without that person's knowledge. For all other relationships of obligations, contract law applies. As a result, contractual relationships are both very important in German law, but are also subject to special rules.

Unlike common law, Germany's Civil Code does not always assume both parties to a contract are equals and takes - to a certain extent - special effort to protect parties from power asymmetries. Also, the German Civil Code defines certain types of contracts that are subject to different statutory rules, depending on what category they fall into (e.g., loan contracts, rental agreements, etc.). One of the categories is a mandate, which allows one party to contract another to do something on their behalf. This is the contract that governs most lawyer-client relationships. It comes with special obligations, such as a duty of loyalty, which is analogous to the duty of loyalty in fiduciary law. However, one could write this duty of loyalty into virtually any German contract. In summary, within German law, it is no problem to create fiduciary duties through contract.

/ Third-party rights

Another core element of a data trust is the ability to commit a data trustee to act in the best interests of a third party, the beneficiary. The beneficiary never has to sign any contract or even be aware of the creation of a data trust. However, if the data trustee were to violate its duties, the beneficiary would be able to challenge their decisions in court. Without a legal data trust, how might we create such a construction in German Law?

The German Civil Code allows for one party to contract another party, the data trust, to perform a service on behalf of a third-party. This is called a contract for the benefit of third parties ('Vertrag zugunsten Dritter'). For instance, let's say Anna wants to make her DNA data available for research. Of course, as her DNA data also reveals information about her family members, she wants to ensure that their interests are protected as well. She could contract a data trust to look after her data rights and protect both her and her family's interests. If the data trust fails to act in her family's best interest, her family could sue the data trust, even though they never signed a contract themselves.

/ Entities acting as a data trust

Finally, without a legal trust, what entities should hold the fiduciary duties and act as a *de facto* data trust?

Given that in German law fiduciary duties can be created through a contract, there is no real limit on the type of entity that could undertake such tasks. One option would be a company structure (GmbH or equivalent). Of course, given our additional safeguards outlined in the previous chapter, we may want to avoid an entity that can turn a profit. Instead, we could opt for a non-profit entity, like an association or a non-profit to act as the fiduciary.

Finally, there may be one restriction on who can act as a data trustee. Under German law, only lawyers are allowed to render legal services or legal assessment on behalf of another person. Therefore, if the data trustee is expected to engage in such work (which seems likely) they would have to be a lawyer or have to engage a lawyer.

Policy recommendations

How can policymakers help bring reality closer to the goals and principles sketched out above? We recognize three main interventions that, together, could provide the foundations for new data governance models. The first is to clarify existing legal confusion. The second is to create a separate entity to act as a data intermediary, with fiduciary duties. And finally, we encourage policymakers to create legal sandboxes that allow for safe experimentation with different models. Below we discuss each option in detail.

/ Clear up legal uncertainty

As discussed above, there remains a fair amount of confusion around whether it is legally possible to have our GDPR rights represented by a third party. In order to move forward with the creation of data governance models, we need to know with certainty what GDPR rights can be represented, under what conditions, and by who. Removing the current legal uncertainty is vital.

Similarly, guidance around the possibility of mandating our GDPR rights through contract would be helpful. Here, we should also point out that while such a possibility would enable new forms of data governance, it could also open the door to contracting our rights to data controllers, or other third parties with an interest in our data.

Finally, as the EU Commission is considering a new Data Governance Act, it needs to clarify what this act is supposed to accomplish in terms of the use of personal data. In its current form, the act mainly addresses the institutional arrangement of what it calls ‚data sharing service providers'. This is to be welcomed. But the Act does not clarify how these intermediaries could actually make use of personal data that they would be entrusted with. At the moment, the EU plans to address these topics in a separate „Data Act". It remains unclear whether the adjustments that are needed concerning personal data can and would be made in this Data Act, or whether the GDPR needs to be changed to achieve this. In either case, we call on the EU Commission to soon lay out how it intends to balance the possibility of sharing personal data via data sharing service providers with appropriate safeguards for users, e.g. by creating a new role in addition to that of data controller and data processor.

/ A new data intermediary

We recommend the creation of a new legal role: *data intermediary or steward*. Such an entity would, at the very least, have a fiduciary duty to only act in the *sole* interest of those whose data it represents and would have to adhere to a strict set of safeguards and duties.

As argued above, the possibility of representation of GDPR rights would allow for the creation of data intermediaries, such as data trusts. However, such a reality would give rise to the question of how a data subject is to evaluate these data intermediaries. As discussed, we doubt fiduciary duties alone will be enough to protect data subjects from risks related to this principle-agent problem. Indeed, we recommend additional safeguards are put in place. The creation of this new legal entity would ensure such safeguards are implemented. And, if an intermediary fails to uphold its duties, it would lose its special status. Much like doctors or lawyers risk losing their license if they breach their duty of loyalty.

/ Regulatory sandboxes

Many have researched and discussed the possibility of creating data trusts and similar legal data intermediaries. However, partly due to the legal uncertainties mentioned above, a lot of our theories are as of yet untested. To allow us to gain a better understanding of what works and what does not, and to be able to identify the benefits and risks inherent in each model, we recommend the creation of regulatory sandboxes[21]. These kinds of sandboxes usually exist for a limited period, during which they are given special leeway so that the experimenter can understand the real-world implications of new technology or a new institutional arrangement. Lessons learned from these experiments would help guide both future regulations and the design of these new data governance models themselves.

21 Federal Ministry for Economic Affairs and Energy (2019): 'Making Space for Innovation - The Handbook for Regulatory Sandboxes'

Data trusts in Germany and under the GDPR

Anouk Ruhaak
December 2020

Available online at https://algorithmwatch.org/en/data-trusts/

Publisher:
AW AlgorithmWatch gGmbH
Linienstr. 13
10178 Berlin
Germany

Contact: info@algorithmwatch.org

Copy editing:
Graham Holliday

Layout:
Beate Autering
www.beworx.de

With support from

moz://a

ZEIT-Stiftung Ebelin und Gerd Bucerius