Harvard Journal of Law & Technology
Volume 33, Digest Spring 2020

MY DATA, MY TERMS: A PROPOSAL FOR PERSONAL DATA USE LICENSES

Paul Jurcys, Chris Donewald, Jure Globocnik, Markus Lampinen[*]

TABLE OF CONTENTS

I. INTRODUCTION

In this paper, we propose the creation of a system of personal data licenses that will help individuals determine conditions for granting access to their personal data. The taxonomy of personal data licenses is based on three major premises. First, we believe that in the near future it will no longer be true that only three companies (i.e., Google, Facebook, and Amazon) hold most of the world's data. Second, we submit that it is the individual who is the ultimate source of truth. And, third, we maintain that existing technologies and social sentiment are now mature enough for the emergence of effective user-centric data models.[1]

---

* Paul Jurcys is a cofounder of Prifina and a CopyrightX Teaching Fellow at Harvard University's Berkman Klein Center for Internet and Society. Christopher Donewald is Senior Corporate Counsel dealing with privacy, protection, and trust at Affirm and a Professor of Law at Golden Gate University School of Law. Jure Globocnik is Junior Research Fellow at the Max Planck Institute for Innovation and Competition in Munich. Markus Lampinen is Co-Founder and CEO of Prifina.

1. J. M. Chua, *Direct-to-Consumer's Lasting Impact on Fashion*, VOGUE BUSINESS (Feb. 3, 2020),
https://www.voguebusiness.com/consumers/direct-to-consumer-lasting-impact-on-fashion-levis-nike-samsonite.

Our proposed system of personal data licenses is the result of the ongoing effort to build a user-centric, user-held data ecosystem where individuals maintain digital copies of their raw personal data in one place (a personal data cloud) and have the ability to control who can access that data. From a macro-perspective, two major problems need to be solved in order to make personal data portable. The first roadblock is related to infrastructure: usable data formats need to be standardized or, at a minimum, made functionally interoperable. The second problem relates to user experience: individuals must understand that their personal data is valuable and should be armed with tools that enable them to manage data relationships with third parties.

The underlying paper is structured as follows: First, the overarching trends pertaining to personal data are presented, followed by a discussion on data "ownership" and access to data. In the next section, which comprises the bulk of the paper, personal data licenses are explained and their practical and technical viability is assessed. The paper concludes with a brief evaluation of possible next steps.

II. REVOLUTION IN THE DATA MARKET: THREE CHANGES

There are three noticeable and fundamental developments taking place in the field of personal data. First, the adoption of new laws on different continents reflects an increasing interest in, and concern for, data privacy, by lawmakers, politicians, and society as a whole.[2] Countless data breaches[3] occurring over the past few years have illuminated the vulnerability of data stored in centralized data servers. Every new instance of a major data breach serves as a repeated reminder that individuals have very little control over their personal data and ignites fierce debates among technology journalists, public interest groups, and other stakeholders.

The European General Data Protection Regulation ("GDPR")[4] and the California Consumer Privacy Act ("CCPA")[5] are the most

2. For example, Brazil has passed a data protection law that is inspired by the EU's GDPR, while the Parliament of India is currently discussing such a bill. *See* Saritha Rai, *India's About to Hand People Data Americans Can Only Dream Of*, BLOOMBERG (Jan. 12, 2020), https://www.bloomberg.com/news/articles/2020-01-13/india-s-about-to-hand-people-data-americans-can-only-dream-of (last visited Feb. 4, 2020).

3. In 2005, 157 breaches at U.S. businesses, government agencies, and other organizations were reported, the number of breaches reached 1,251 in 2018. *See ITRC Multi-Year Data Breach Chart, Jan. 1, 2005 - Dec. 31, 2018,* IDENTITY THEFT RESOURCE CENTER (Jan. 31, 2019), https://www.idtheftcenter.org/wp-content/uploads/2019/02/Multi-Year-Chart.pdf (last visited Jan. 23, 2020).

4. Regulation (EU) 2016/679 of Apr. 27, 2016, on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and Repealing Directive 95/46/EC, 2016 O.J. (L 119) 1 [hereinafter GDPR].

5. California Consumer Privacy Act of 2018, Cal. Civ. Code § 1798.100 (2018) [hereinafter CCPA].

aggressive legislative acts aiming to improve how companies manage their customers' data. Both the GDPR and the CCPA set forth a number of new or expanded rights for individuals. For instance, the CCPA (which came into force on January 1, 2020), establishes the right of individuals to request information from businesses about consumer data the business collects about them (§ 1798.100) and to require businesses to delete any personal data they have collected about the consumer (§ 1798.105). One of the cornerstones of the CCPA relates to the sale of consumers' personal data: the CCPA provides that consumers have the right to require companies to disclose what information about a particular consumer they are collecting for sales and business purposes (§ 1798.115) and can even opt-out of sales of their personal data (§ 1798.120).

The second significant change relates to the way companies approach the collection and use of their customers' personal data. New legal requirements result in increasing costs for regulatory compliance.[6] In addition, companies are being forced to look for alternative ways to obtain information about their customers. This is due not only to more stringent data protection regimes around the globe, but also because the companies holding the data (especially Google, Facebook, and Amazon) are less inclined to disclose consumer data they have in their possession.[7] Over the past few years, companies holding massive amounts of data, like Google and Microsoft, have lessened the data types and quantity they are willing provide to the marketplace,[8] and this trend is likely to continue, posing an increasing problem for various companies relying on this information for their business practices.

Third, technological advancement has accelerated at an immense pace in the past few years. Data processing technologies have reached a level of maturity, and decentralized data management models have become feasible. One of the main assumptions of such decentralized

---

6. An economic impact assessment prepared for the California's state attorney general's office estimated the total cost of initial compliance with the CCPA to be approximately $55 billion, which is equivalent to 1.8% of California's Gross State Product in 2018. CALIFORNIA DEPARTMENT OF JUSTICE, STANDARDIZED REGULATORY IMPACT ASSESSMENT: CALIFORNIA CONSUMER PRIVACY ACT OF 2018 REGULATIONS 11 (2019), http://www.dof.ca.gov/Forecasting/Economics/Major_Regulations/Major_Regulations_Table/documents/CCPA_Regulations-SRIA-DOF.pdf.

7. Twitter and LinkedIn have in the past restricted access to data pertaining to their users by companies using the data to provide data analytic services. *See* Thomas Tombal, *Economic Dependence and Data Access*, 51 INT'L REV. OF INTELL. PROP. AND COMPETITION L. 70, 79–80 (2020).

8. Gerrit De Vynck and Naomi Nix, *Google Follows Apple in Ending Third-Party 'Cookies' in Ad-Tracking*, BLOOMBERG (Jan. 14, 2020),

https://www.bloomberg.com/news/articles/2020-01-14/google-plans-to-move-forward-with-changes-to-ad-tracking-tools; *see also* Gerrit De Vynck and Mark Bergen, *Google Stuck Between Privacy, Antitrust With Ad Data Limits*, BLOOMBERG (Jan. 14, 2020), https://www.bloomberg.com/news/articles/2020-02-03/google-gets-stuck-between-privacy-antitrust-with-ad-data-limits.

data processing models is that the most authentic source of information is the individual ("only you know what was on your breakfast table"). Namely, while the data held by a company about an individual might get outdated quickly,[9] obtaining the data directly from the "source" (i.e., the individual) is best to guarantee up-to-date information. From a purely technological point of view, data processing is becoming feasible not only in centralized databases but also "closer" to individuals (e.g., in each person's own device or personal data accounts in the cloud).[10]

Such decentralized data processing models create a solid technological foundation for the exchange of information: for service providers, accessing data directly from their customers means a significant reduction of costs and risk. User-centric, user-held data models liberate service providers from collecting data from third parties (data brokers) and give service providers tools to get the most accurate data directly from their customers (with customer consent). Such new decentralized data models would also help companies create more personalized experiences for their customers and increase competition among companies trying to offer more customer value. Moreover, individuals will benefit from having better control over the usage of their personal information, as well as from receiving better, individualized products and services.[11]

This decoupling of personal data from the big data platforms of today also has broad implications in the data science field. Currently, due to the fact that the most advanced data applications require the most comprehensive depth and breadth of data, the most advanced and sought-after data science positions are with parties that possess such data, often either with data platforms themselves or government-related agencies (e.g., the National Security Agency, military agencies, etc.). This means that a data scientist's pursuit of interesting data science professions is naturally coupled to the interests of these large organizations.

In the future, however, it is likely that user-centric, user-held data models will be built using open-source tools[12] which will be available to any software developers who can integrate such decentralized data approaches in improving the existing business models by building more

---

9. See Vikas Kathuria & Jure Globocnik, *Exclusionary Conduct in Data-driven Markets: Limitations of Data Sharing Remedy*, JOURNAL OF ANTITRUST ENFORCEMENT 1, 12 (2020).

10. The development of edge computing, coupled with the wider availability of 5G wireless technologies, will likely further accelerate this trend; see Weisong Shi, George Pallis & Zhiwei Xu, *Edge Computing*, 107 Proceedings of the IEEE 1474, 1474-1478 (2019).

11. Douglas Elliott & Lisa Quest, *It's Time to Redefine How Data Is Governed, Controlled and Shared. Here's How*, WORLD ECON. F. (Jan. 14, 2020), https://www.weforum.org/agenda/2020/01/future-of-data-protect-and-regulation.

12. *See, e.g.*, Prifina, *Liberty, Equality, Data Model*, GITHUB, https://github.com/libertyequalitydata (last visited Feb. 17, 2020).

personalized applications. Just as technological shifts with App Stores' growth allowed software developers to market services directly to consumers for the first time in a significant way, decoupling data from the contemporary data platforms may ignite a more direct-to-consumer era of services that provide value directly to the individuals themselves.

III. IS THERE A NEED FOR "OWNERSHIP" OF PERSONAL DATA?

The issue of "ownership" in data has been heavily discussed in recent years. While the discussion was centered on non-personal data, ownership in personal data was also given considerable attention. In the broadest sense, the term "personal data" encompasses any information relating to an identified or identifiable individual (data subject).[13] In other words, personal data is linked, by reason of its content, purpose, or effect, to a particular individual.[14]

Nevertheless, the mere fact that a piece of data pertains to a specific individual does not imply that the individual also "owns" her personal data in a legal sense.[15] In fact, data protection laws currently do not allocate ownership of personal data to any subject.[16] There is also no other legal principle or theory that would *per se* justify the allocation of exclusive property rights over data.[17] Therefore, any recognition of a new (intellectual) property right, such as an ownership right in

13. OECD, OECD GUIDELINES ON THE PROTECTION OF PRIVACY AND TRANSBORDER FLOWS OF PERSONAL DATA 13 (2013) (see Art. 1(b)).

14. Opinion of the Working Party on the Protection of Individuals With Regard to the Processing of Personal Data on the Concept of Personal Data, 01248/07/EN (2007) 10–11, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf.

15. Josef Drexl, *Designing Competitive Markets for Industrial Data - Between Propertisation and Access*, 8 J. OF INTELL. PROP., INFO. TECH. AND E-COM. L. 257, 267 (2017).

16. OECD, ENHANCING ACCESS TO AND SHARING OF DATA: RECONCILING RISKS AND BENEFITS FOR DATA RE-USE ACROSS SOCIETIES 100–01 (2019). On the other hand, a Working Paper by the Joint Research Centre of the the European Commission argues that the GDPR defines "[p]artial and limited ownership rights to data ." Nestor Duch-Brown, Bertin Martens, & Frank Mueller-Langer, *The Economics of Ownership, Access and Trade in Digital Data*, 12 (Joint Research Center, Working Paper 2017-01, 2017), https://ec.europa.eu/jrc/sites/jrcsh/files/jrc104756.pdf.

17. Josef Drexl et al., *Data Ownership and Access to Data*, MAX PLANCK INSTITUTE FOR INNOVATION AND COMPETITION 1, 2 (2016) http://www.ip.mpg.de/en/link/positionpaper-data-2016-08-16.html.

(personal or non-personal) data, would be in need of a sound justification.[18] Such a justification currently does not exist.[19]

From a purely economic point of view, the conferral of exclusive rights over data upon certain market players could lead to market distortions. Namely, even if data ownership rights would initially be vested in the individual, due to unequal bargaining power, big service providers might easily be able to request data be licensed to them as a precondition to use a desired service — possibly even on a royalty-free basis[20] — hence consolidating their market power.[21] Exclusive rights over data could create barriers to entry into the markets,[22] and impede the creation of new data-based products and services in neighboring markets.

While data protection laws do not govern the ownership of data, they also do not allocate the economic value generated by way of data processing to any of the involved subjects.[23] This is therefore a factual rather than a legal question. The crucial question in this context is who holds the data, because *de facto* control over data enables the data holder to generate revenue from it. Currently, the majority of personal data is collected and held by companies — either by companies providing various services to individuals or by data brokers — rather than data subjects themselves.[24] Thus, while not owning data in a legal

---

18. *Cf.* Josef Drexl et al., *Position Statement of 26 April 2017 on the European Commission's "Public consultation on Building the European Data Economy,"* MAX PLANCK INSTITUTE FOR INNOVATION AND COMPETITION 1, 5 (2017), https://www.ip.mpg.de/fileadmin/ipmpg/content/stellungnahmen/MPI_Statement_Public_consultation_on_Building_the_EU_Data_Eco_28042017.pdf. For personal data ownership rationales and their flaws, *see* Vaclav Janecek, *Ownership of Personal Data in the Internet of Things*, 34 COMPUTER L. & SECURITY REV. 1039, 1044–45 (2018).

19. In his analysis of the possibility of the application of the concept of data ownership in the IoT context, Vaclav Janecek comes to the conclusion that currently, the introduction of ownership rights in personal data is justified neither from a top-down nor from a bottom-up approach. The top-down approach fails to convincingly explain why ownership-like control is best suited to achieve economic and factual goals as opposed to other models of data control. On the other hand, for the bottom-up approach to function: (1) better factual control of the potential rightholder over data would be needed, and (2) regardless of the approach taken, it is implausible to expect that the law could offer stable protection over personal data as the existing IoT architectures are not transparent enough. Janecek, *supra* note 18, at 1044–46. On the other end of the spectrum, Nadezhda Purtova argues that the benefit of the introduction of ownership rights in personal data would introduce ultimate clarity as to the allocation of data protection obligations; NADEZHDA N. PURTOVA, PROPERTY RIGHTS IN PERSONAL DATA: A EUROPEAN PERSPECTIVE 270 (2011).

20. Drexl et al., *supra* note 18, at 8.

21. If the current regulatory status quo continues, big tech firms might become even more influential. *See* Elliott & Quest, *supra* note 11. This holds all the more true if they were to obtain ownership rights over the data they process, or have the data licensed to them.

22. *See* Drexl et al., *supra* note 17, at 2.

23. Drexl, *supra* note 15, at 267. Data protection laws contain provisions on *how* personal data can be processed, but remain silent on the matter of revenue allocation.

24. In order to justify the ownership of customer personal data, technology companies submit they have made large investments in building their business models and creating complex tools to harness customer data. For one example, see the testimony given to the US

sense, these companies can be considered *de facto* owners of data (owners in an economic sense).[25]

There is a widespread consensus that natural persons should be better able to participate in the wealth generated through the usage of personal data pertaining to them.[26] However, conferring (exclusive) data ownership rights to certain categories of stakeholders may not be the best way forward. Besides bearing significant risks for competition, data ownership would likely be very difficult to regulate. For example, the question of rightholdership is not a straightforward one, as often multiple stakeholders directly or indirectly contribute to data collection and processing. Further, potential co-ownership of data could result in blocking situations,[27] and exacerbate inefficiencies due to the underuse of data. Indeed, a complex system of exceptions and limitations would have to be introduced, taking account of the interests of other subjects.[28] Other ways forward should therefore be explored.

IV. THE IMPORTANCE OF ACCESS TO PERSONAL DATA

As shown above, legal and technological complexities speak against the introduction of data "ownership." Taking this into account, the discussion has shifted to the issue of *access* to personal data.[29] In the existing legal environment, access to personal data is conditional upon, first and foremost, the legislation granting the individuals certain rights vis-a-vis companies related to the data that those companies have collected about them. Access to data could be requested based on the right to get information about what data is collected and processed[30] or

House of Representatives by Equifax CEO Mark Begor, who explains that Equifax always tries to have the most up-to-date and relevant data about their customers. *See generally Who's Keeping Score? Holding Credit Bureaus Accountable and Repairing a Broken System: Hearing Before the H. Comm. on Fin. Servs.*, 116th Cong. (2019) (statement by Mark Begor, CEO, Equifax, Inc.), https://financialservices.house.gov/uploadedfiles/hhrg-116-ba00-wstate-begorm-20190226.pdf

25. *See* Nestor Duch-Brown, et al., *supra* note 16, at 23–24.

26. *See, e.g.*, *Opinion of the European Data Protection Supervisor 9/2016 on Personal Information Management Systems*, at 5 (Oct. 20, 2016), https://edps.europa.eu/sites/edp/files/publication/16-10-20_pims_opinion_en.pdf.

27. Drexl et al., supra note 18, at 7–8.

28. *Id.* at 9–10.

29. *See* Drexl et al., *supra* note 18, at 12–13. In the EU, access to personal data has been regulated in Arts. 15 and 20 of the GDPR and Art. 16 of the Digital Content Directive ("DCD"). Similarly, numerous jurisdictions such as Brazil and India introduced or are discussing the introduction of data portability rights inspired by the GDPR.

30. In the EU, such a right is named the right of access, and gives the individual the right not only to obtain a copy of the personal data undergoing processing, but also to be informed about, i.e., the purposes of processing, the recipients to whom the personal data have been or will be disclosed, and, where possible, the envisaged period for which the personal data will be stored. *See* GDPR, *supra* note 4, at Art. 15.

the right to get digital copies of personal data.[31] These rights are often complemented with other rights, such as the right to request deletion of that data or to opt-out from the selling of personal data.

The rapid advances of personal data management technologies based on the notion of user-held, user-centric data has prompted the development of new consumer-oriented personal data management tools.[32] These tools aim to empower individuals to have their own copies of "their" personal raw data and to get value by controlling who can access that data. Significant quantities of raw data about individuals can be collected by scraping the internet, from the individual's personal accounts with various service providers,[33] or by individuals themselves (e.g., individuals can volunteer their time in order to set their preferences which could be used in various interactions with third parties).

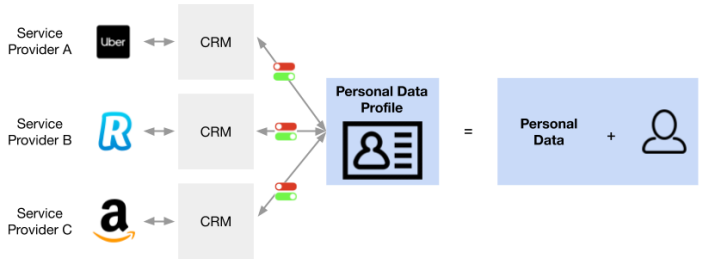

Image 1: The Framework for Accessing Personal Data Profiles

Data portability is mired by various, often proprietary, data formats used today as a result of proprietary data platforms and models that remain disconnected and lack interoperability. While it is unrealistic to

31. In the EU, the right to data portability, enshrined in Art. 20 of the GDPR, gives the individual the right to receive certain categories of personal data concerning her in a structured, commonly used, and machine-readable format, or, where technically feasible, to have these data transmitted directly from one company to another one.

32. For example, one report drafted on behalf of the Swiss federal government (Bundesrat) came to the conclusion that personal data management systems are a promising concept for the governance and usage of own personal data. *See* ROLF H. WEBER & FLORENT H. THOUVENIN, GUTACHTEN ZUR MÖGLICHKEIT DER EINFÜHRUNG EINES DATENPORTABILITÄTSRECHTS IM SCHWEIZERISCHEN RECHT UND ZUR RECHTSLAGE BEI PERSONAL INFORMATION MANAGEMENT SYSTEMS (PIMS) 43 (2017), https://www.itsl.uzh.ch/dam/jcr:26f84429-2aef-47b1-9ba1-6f6e8910c60d/180321%20BJ-Gutachten_final.pdf.

33. Numerous providers already provide tools to download data pertaining to a certain user. Furthermore, in 2018, big technological companies like Apple, Google, Facebook, Microsoft, and Twitter launched the Data Transfer Project, the aim of which is to create an open-source, service-to-service data portability platform enabling users to easily move their data between online service providers. *See* DATA TRANSFER PROJECT, https://datatransferproject.dev (last visited Feb. 19, 2020).

advocate changing existing models today, there are avenues where portability can be fostered. It is reasonable, then, to focus on providing compatibility of data models from one platform to another, where the individual wishes to bring their data from one service to another and establishing open standards as tools for new services and products. Industry bodies have created such standards in various verticals in the past, such as the IAB in advertising and the CINT standard in surveys; however, few standards exist beyond the industry they have been created in. By providing for compatibility with existing platforms and convergence with existing industry standards, as reasonable, new models can bridge the gap by having enough familiarity to be utilized by existing organizations and enough novelty to bring in new aspects such as user set data use licenses to incentivize new use and value.

V. PROPOSAL: PERSONAL DATA USE LICENSES

User-held data means that an individual has her personal data in her personal cloud account which can be accessed only by the individual herself.[34] Hence, in order to get value from personal data in interacting with third parties, the individual has to have tools that enable her to "activate" that personal data. In other words, the individual should be able to decide upon the conditions for the use of her personal data profile by third parties.

Assuming that individuals have copies of their personal data and are able to provide access to it, the following question relates to the scope of access the individual would be willing to grant and what permissions a third party accessing that personal data profile would be given. In one of the most recent empirical studies on the value of personal data,[35] Harvard scholars Cass Sunstein and Angela Winegar showed that individuals value most their health, personal identity, and finance data. However, we anticipate that the actual value of personal data varies depending on personal preferences as well as case-specific circumstances in which individuals interact with service providers.

In the light of the existing data usage practices by consumer-facing companies, it may be envisaged that, in their dealings with such third-party service providers, individuals should be able to set the following conditions for access to a personal data profile:

(1) **Full/limited anonymity**: For example, a reader of an online news portal may choose to remain anonymous and not to disclose any personal information about herself to the site.

34. Such systems are sometimes referred to as Personal Information Management Systems (PIMS). *See European Data Protection Supervisor*, *supra* note 26 at 5–6.

35. Angela G. Winegar & Cass R. Sunstein, *How Much Is Data Privacy Worth? A Preliminary Investigation*, 42 J. OF CONSUMER POL'Y 425 (July 1, 2019).

This means that, without being able to identify any attributes of the reader, the news portal administrator will provide only generic website content to that reader. However, if the reader were to reveal her age range and hobbies (e.g., an 18-25 year-old female interested in fashion trends), the content of the website could be tailored to that reader's interests.

(2) **Permission to track**: By granting access to her personal data profile, an individual can impose an obligation on the service provider not to follow that particular user (i.e., not to track individual's activities during or after that particular session). Such a restriction on tracking can have great practical significance, as many webmasters are now collecting data on user behavior (e.g., how many microseconds users spend watching certain content and how content users behave in the digital space).

(3) **Permission to store data**: This means that even if the service provider is given access to the personal data profile of a particular individual, the service provider is not entitled to retain the personal data profile in its system.

(4) **Permission to bundle data**: Individuals should also have the right to prevent third-party service providers from aggregating that particular individual's personal data profile with personal data profiles of other individuals. This could have important implications for ethical data use practices for technology companies who are deriving the greatest value from aggregated customer personal data.

(5) **Permission to share data**: Individuals should be able to impose a requirement that the service provider does not share that individual's personal data with third parties. This restriction should be especially significant in cases where the individual grants access to her personal data profile and also if companies are allowed to store that individual's personal data profile.

(6) **Permission to sell data**: As mentioned above, one of the most controversial issues currently relates to the fact that personal customer data is sold among companies without customer consent. Accordingly, individuals who grant access to their personal data profiles should be able to prohibit the selling of their personal data to third parties.

The hierarchical flow from the most "private" regime to the most liberal personal data sharing framework can be explained in the chart below:

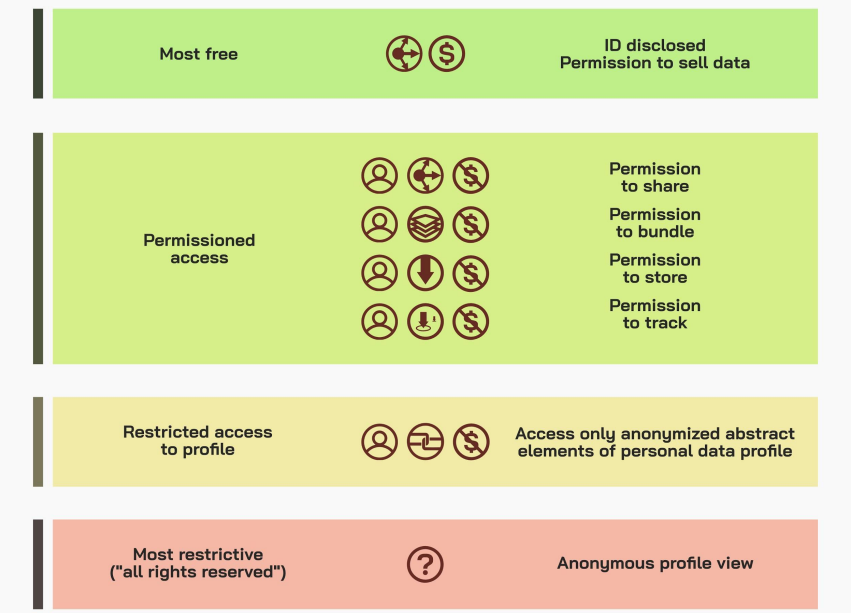

Image 2: Permissions to Access Personal Data Profiles

In practice, it is possible that the above mentioned permissions will evolve into default licenses comprising several combinations of permissions. For instance, an individual may choose to remain anonymous, but give permission to the service provider to share some of the data (demographics, and certain data related to the behavior while using the product). Such licenses are likely to evolve based on certain practices in different various verticals.

In addition to those core types of permissions, and unless the individual is interacting under a completely anonymous setting, each of the above-mentioned licenses may be accompanied by additional sets of permissions:

(1) **Duration of access**: The individual should be able to determine the period for which the access to the profile is granted (e.g., one-time access, for one hour, for 12 hours, for a week - depending on a given case).

(2) **Identification of the Accessing Person**: the name of the person given access to personal data profile is shared.

(3) **Personalized value**: each license will be accompanied with a value consideration. Given the nature of the relationship between the individual and third party service provider, an individual will receive certain personalized value offerings

based on the personal data profile. Such a personalized benefit could take various forms: a discount, better user experience, better personalization, etc.

This proliferation of user-centric, user-held data models will make a significant change in the fundamental approach to data ecosystem: currently prevailing opt-out privacy regimes will be replaced by opt-in frameworks. The ability to have copies of personal data in personal data clouds will enable individuals to better understand the amount and value of their personal data. More specifically, based on raw data in personal data clouds will be translated into easily understandable data dashboards. Such dashboards will also be accompanied by additional tools that help the individual keep track of the consents she has given.[36] Besides being informed about the companies that are processing her data, the individual would also be able to alter the permissions given without even contacting the company processing the data.[37]

The shift towards more user-centric, user-held data models does not mean an apocalyptic revolution but, rather, a transition towards a more nuanced and more ethical data ecosystem. This is so because in certain sectors, legal obligations requesting certain subjects to possess certain data will continue to exist. For example, banking sector regulations would impose such obligations with the aim to combat money laundering and terrorist financing, and tax laws may request the same in order to combat tax fraud and tax evasion. Under the existing legal regimes, the presence of such a legal obligation renders data processing lawful.[38] Similarly, such legal obligations should also take precedence over the will of the individual, expressed in the data licenses she has granted.

VI. FUTURE IMPLICATIONS TO THE DATA ECOSYSTEM

As we enter the next generation of the Internet, which will be based on the use of such commodities as personal data, labor, resources, services, and interaction between individuals based on trust,[39]

36. For example, unless any other legal basis for data processing is fulfilled, EU law requires consent be given for the data processing to be lawful. However, in the "traditional" opt-out consent system the individual does not have any possibility to get an overview of the consents given. To find out which companies she has given consent to, she has to exercise her right of access vis-a-vis every such company. *Cf. European Data Protection Supervisor*, *supra* note 26, at 7. Additionally, the German Data Ethics Commission has endorsed research on such user-centric approaches. *See Gutachten der Datenethikkommission*, DATENETHIKKOMMISSION (2019) 133–34.

37. *Id.*

38. See, for example, Art. 6(1)(c) GDPR.

39. If individuals feel more secure about how data pertaining to them is being processed, they may be more inclined to take advantage of digital services; Elliott & Quest, supra note 11.

automated agreements and personal AI tools will help maximize individual utility. Personal data will not be a commodity *per se* (why would you sell your genetic data for $10.00 in a finite transaction?), but will actually serve as a medium that facilitates interaction between different parties and curtails information asymmetries.

Accordingly, personal data use licenses will drive data interactions based on explicit prior consent — that is, individuals will be able to communicate the terms of access and use of personal data before or at the time when access to their personal data profile is given. This will provide more legal certainty and clarity. In addition, individuals will be able to change data access conditions at any time. User-centric, user-held data models are likely to play an important role in facilitating competition between service providers in order to provide valuable services that are as aligned as possible to the interests and expectations of service users.

This system of personal data licensing terms is based on a bottom-up logic where data is held by the individuals themselves. Allowing individuals to determine the terms for accessing their personal data profiles could have great potential as a flexible, market-driven approach in solving problems related to the use of personal data. Creative Commons could be seen as a similar project which has been implemented in order to allow creators of copyrighted content to set forth the terms and conditions for the use of the content.[40]

The Montreal Data License is one of the recent efforts to address the problem of how data is used by businesses for AI and machine learning purposes.[41] Montreal Data Licenses are focused on databases and big data; the main purpose is to help database creators easily generate licenses to facilitate the use of those databases by other entities for research or commercial purposes.[42] Similarly, Microsoft has presented drafts of data sharing agreements which are also designed for data-sharing scenarios between companies rather than individuals.[43]

Recently, there have been some efforts to develop so-called "privacy-icons" which should help individuals easily understand how their data is used by service providers.[44] They aim to make data use

40. For an overview of Creative Commons licenses, *see Six Licenses for Sharing Your Work*, CREATIVE COMMONS, https://wiki.creativecommons.org/images/6/6d/6licenses-flat.pdf (last visited on Jan. 25, 2020).

41. Misha Benjamin et al., *Towards Standardization of Data Licenses: The Montreal Data License*, ARXIV (Mar. 21, 2019), https://arxiv.org/pdf/1903.12262.pdf

42. *Id.* at 3.

43. *Removing Barriers to Data Innovation*, MICROSOFT, https://news.microsoft.com/datainnovation (last visited on Jan. 23, 2020).

44. *See* Privacy Icons, I4BIWIKI (2012), https://cyber.harvard.edu/i4bi/Privacy_Icons; *see generally* Zohar Efroni, Jakob Metzger, Lena Mischau & Marie Schirmbeck, *Privacy Icons: A Risk-Based Approach to Visualisation of Data Processing*, 5 EUR. DATA PROTECTION L. REV. 352 (2019).

terms more user-friendly and reduce transaction costs.[45] Such privacy icons were advocated by Mozilla, which tried to simplify existing privacy policies.[46] The GDPR and CCPA contain some fragments of this effort. Art. 12(1) of the GDPR requires that the information about the processing of personal data be provided in concise, transparent, intelligible and easily accessible form, and Art. 12(7) of the GDPR sets forth that this information may be provided in combination with standardized icons. If presented electronically, the icons should be machine-readable. Under the implementing regulations for the CCPA, there is a requirement for the California Attorney General's Office to develop a "do not sell my data" icon.[47]

User-centric, user-held data models present a great opportunity for service providers and merchants alike. Several companies have already taken the stance of not selling data (e.g., Microsoft) and respecting their users privacy (e.g., Apple) and similarly, by empowering individuals to choose and decide for themselves, we believe business will see ways to create more customer value, by being able to better serve their customer in a timely manner and more efficiently than possible before. The most challenging task is revolving around user experience: someone will have to come up with a solution allowing individuals to take-back their data in two or three clicks without compromising security of data.

In conclusion, the coming years will see interesting developments in the field of personal data technology and the legal regulation of this data. Looking ahead, it is worth discussing more broadly the appropriateness of existing legal concepts to address legal, social and economic challenges in this fast-moving field.[48] While user-centric data models might not replace the need for baseline data protection entirely, they could streamline the complex web of data management responsibilities, and allow the monetization of data.[49] Licenses for the use of personal data should be one of the most effective tools for creating a fair, individual, and market-driven ecosystem.

---

45. Benjamin et al., supra note 41, at 8.

46. *Privacy Icons*, MOZILLAWIKI (2011), https://wiki.mozilla.org/Privacy_Icons and https://wiki.mozilla.org/Privacy_Icons_v0.2.

47. *See* CCPA § 1798.185(4)(c) and Proposed Text of CCPA Regulations, § 996.306(e)(1).

48. The discussions mainly center around the potential need to readjust data protection legal regimes due to the advent of AI, and the suitability of consent-based systems to give individuals control over the data pertaining to them. *See, e.g.*, Daniel J. Solove, *Introduction: Privacy Self-Management and the Consent Dilemma*, 126 HARV. L. REV. 1880 (2013); Sandra Wachter & Brent Mittelstadt, *A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI*, 2 COLUM. BUS. L. REV. 494 (2019).

49. Elliott & Quest, supra note 11.