UC Irvine

UC Irvine Journal of International, Transnational, and Comparative Law

Title

Data as Public Goods or Private Properties?: A Way Out of Conflict Between Data Protection and Free Speech

Permalink

https://escholarship.org/uc/item/6ch8f8q9

Journal

UC Irvine Journal of International, Transnational, and Comparative Law , 6(1)

Author

Park, Kyung Sin

Publication Date

2021-05-27

eScholarship.org Powered by the California Digital Library
University of California

Data as Public Goods or Private Properties?: A Way Out of Conflict Between Data Protection and Free Speech

Kyung Sin Park*

In this Article, I will review the origins of data protection laws and reestablish the concept of "data surveillance" as the primary evil that data protection laws should try to abate. From this review, I discover a transnational principle that strong data protection laws are must-haves for all jurisdictions wishing to protect privacy for their people, but that data protection laws should not be applied to data that have been made publicly available through legitimate process. I then find legislative examples embodying such principle. Next, I will look at "scientific research" exemptions from data subjects' control on pseudonymized data, and using GDPR's exemption as an example, will demonstrate that ownership-like control by data subjects is not absolute. Finally, I will examine the possibility and morality of data socialism whereby data (including personal data) are regulated as public goods or infrastructure like scenery, sunlight, air, etc., and whereby data silos are replaced by a data commons for the benefit of all. "Data socialism" is proposed despite its negative connotation among contemporaries intentionally in order to highlight the libertarian pitfalls of the mechanistic application of data protection law.

* Professor, Korea University Law School, Director, Open Net Association (Korea), kyungsinpark@korea.ac.kr.

INTRODUCTION

What many overlook is the relationship between data and speech. Transfer of data to another is speech. Collection of data is cognition, involving the right to know—the alter ego of freedom of speech. Using data for a new purpose is research, as research involves finding new meanings behind the pre-existing data. Data protection laws give data subjects (living persons that the data refer to) ownership-like control over transfer and collection of, and research into, personal data (the data to which data subjects exist).

If you recognize this, various conflicts between data protection laws and freedom of speech are not surprising. For instance, data protection law has been invoked under the banner of the "right to be forgotten" to restrict freedom of access to data that has been made publicly available through legitimate process.

In this Article, I review the origins of data protection laws and reestablish the concept of 'data surveillance' as the primary evil that data protection laws should try to abate. From this, I discover a transnational principle that data protection laws should not be applied to data that have been made publicly available through legitimate process and find legislative examples embodying such principle. Next, I look at "scientific research" exemptions from data subjects' control on pseudonymized data and, using GDPR's exemption as an example, demonstrate that ownership-like control by data subjects is not absolute. Finally, I examine the possibility and morality of data socialism whereby data (including personal data) are regulated as public goods or infrastructure, such as scenery, sunlight, and air, and whereby data silos are replaced by a data commons for the benefit of all.

I. ORIGINS OF DATA PROTECTION LAW

The concept of the personality right (*Grundgesetz*) has dominated the European scene of privacy[1] with a narrative that individuals should have the right to lead their own lives without being paralyzed by mental distress over what others think of or know about them.

> If individuals cannot, with sufficient certainty, determine what kind of personal information is known to their environment, and if it is difficult to ascertain what kind of information potential communication partners are privy to, this may seriously impair the freedom to exercise self-determination. In the context of modern data processing, the free development of one's personality therefore requires that the individual is

1. *See* Maryann McMahon, *Defamation Claims in Europe: A Survey of the Legal Armory*, COMMC'NS L., 24, 31 (2002); 2 BASIL S. MARKESINIS, *Developing an English Law of Privacy*, *in* ALWAYS ON THE SAME PATH: ESSAYS ON FOREIGN LAW AND COMPARATIVE METHODOLOGY 321, 401–15 (Raymond Youngs trans., 2001).

> protected against the unlimited collection, storage, use and sharing of personal data.[2]

The personality right, based upon a provision in the German Constitution,[3] is a unitary concept under which the concerns for privacy, reputation, portrait right, and other personal information are subsumed and are adjudicated through balancing all of those legal interests against the legitimacy of and the need for expressions and communications describing that person.[4]

Balancing is often facilitated by differentiating the areas of human life into domains in which personality rights are given different degrees of protection. These domains are characterized as intimate zones, private zones, individual zones, social zones, and public zones. In intimate and private zones, even truthful information cannot be collected or revealed, while in social and public zones, photographing individuals is allowed, but the permitted use of the photographs is zone dependent.[5]

French law also recognizes personality rights,[6] the facets of which correspond to privacy, portrait right, reputation, and so on.[7] France does not differentiate areas of life into zones, but the method of balancing the level of publicness of the concerned activity, the public or private purpose of the activity, and the countervailing public interest is similar to that used in Germany.

The significance of the personality right jurisprudence stems from the lack of distinction between reputational interest and privacy interest. Importantly, a true statement may still be considered defaming if it concerns a person's *private* life.[8] It is not explained how disclosure of positive facts about a person—e.g., that an artist is Banksy—can be defaming to that person. Similarly, in Sweden, if a truthful statement (e.g., person A has been convicted of rape) is spread with the intent of causing that person harm, or harming his or her standing in society, it is still

---

2. BVerfG, 1 BvR 209/83, Dec. 15, 1983, http://www.bverfg.de/e/rs19831215_1bvr020983.html.

3. GRUNDGESETZ [GG] [BASIC LAW] art. 2(1), translation at http://www.gesetze-im-internet.de/englisch_gg/englisch_gg.html (Ger.) ("Every person shall have the right to free development of his personality insofar as he does not violate the rights of others or offend against the constitutional order or the moral law.").

4. Paul M. Schwartz & Karl-Nikolaus Peifer, *Prosser's Privacy and the German Right of Personality: Are Four Privacy Torts Better than One Unitary Concept?*, 98 CALIF. L. REV. 1925, 1943 (2010).

5. McMahon, *supra* note 1, at 24. Basil Markesinis, Colm O'Cinneide, Jörg Fedtke & Myriam Hunter-Henin, *Concerns and Ideas about the Developing English Law of Privacy (And How Knowledge of Foreign Law Might Be of Help)*, 52 AM. J. COMPAR. L. 133, 188–91 (2004).

6. Jeanne M. Hauch, *Protecting Private Facts in France: The Warren & Brandeis Tort is Alive and Well and Flourishing in Paris*, 68 TUL. L. REV. 1219, 1223 (1993).

7. F. Jay Dougherty, *Foreword: The Right of Publicity—Towards a Comparative and International Perspective*, 18 LOY. L.A. ENT. L.J. 421, 435 (1998). *See also* Code civil [C. civ.] [Civil Code] arts. 9, 1382 (Fr.).

8. Patrick Wachsmann, *La liberté d'expression* [Freedom of Expression], *in* LIBERTÉS ET DROITS FONDAMENTAUX [Freedoms and Fundamental Rights] 498–99 (Rémy Cabrillac ed., 2013).

considered *defamation*,[9] even though the operational interest is ostensibly more of privacy than the rapist's reputation.

Covering privacy under the general rubric of personality right unavoidably makes European privacy *incremental* in nature. In other words, there is no categorical separation between privacy and publicness. Even information voluntarily disclosed to a large portion of the public (therefore taking place in the most public zone) still remains redressable since it may negatively affect a person's right to "free development of personality" even if it is truthful and nondefamatory. Recognizing someone participating in a gay pride march and sharing that information with the public through a magazine article will be redressable in Europe; it will not be redressable in the United States.[10] What is disclosed to a circle of march-goers remains private, compared to its disclosure to an outer concentric circle of magazine readers.

Compounding these differences, Europe has also gone beyond incrementalism in an effort to protect privacy with heroically ample margins: all data about identifiable persons are by default to be controlled by those persons ("data subjects") even if the data do not implicate privacy or reputation. We call this "data protection law."

Data protection law began—ironically—with the American concept of "fair information practice," which Alan Westin developed in a chapter of his influential volume *Privacy and Freedom*.[11] Other chapters of Westin's volume were influential in generating the legislative momentum behind the world's first wiretapping law. In the equally influential third chapter, Westin develops the concept of "data surveillance"—the idea that surveillance can take place even through voluntary disclosure of personal data. According to Westin, accumulation of voluntarily disclosed data can amount to surveillance if the receiver of the data uses it in a manner contrary to the data subject's original wishes at the time of the original disclosure, or if he shares such data with others previously unknown to and unapproved by the data subject. To prevent "data surveillance," Westin advocated for recognition of the right of the individual to "decide for himself who shall know . . . facts [about himself] at what time under what conditions."[12]

Importantly, Westin envisioned the right to be similar to a *property right* where "you own data about you." He anticipated that this property right would address a *market failure*—namely, where powerless individuals submit personal data to governments and companies to obtain necessary services, yet lack the acumen or meaningful opportunity to impose, not to mention enforce, the conditions under

9 . LIBRARY OF CONGRESS, LIMITS ON FREEDOM OF EXPRESSION 1, 73 (2019), https://www.loc.gov/law/help/freedom-expression/limits-expression.pdf (citing Brottsbalken [BRB] [Penal Code] 5:3) (Swed.).

10. Hauch, *supra* note 6, at 1220–21.

11. ALAN F. WESTIN, PRIVACY AND FREEDOM 365–99 (1970).

12. *Id.* at 368.

which their personal data are to be used after being collected. The distinction between a property right and a contractual right makes a huge difference on the default rule. Under a contract regime, the person collecting your personal data will be bound only by the promises that he or she made to you. No promises, no restrictions. Often you do not have bargaining power or acumen to demand promises. Under a property regime, you "own" the data about you, which means that anyone wanting to collect, use, or transfer your data will have the burden of obtaining your consent before using it for any purpose or sharing it with any third party. There will be restrictions even when no promises are made.

Under the data ownership scheme, data collectors have an affirmative duty to obtain express consent—the data subject's silence is insufficient. Such a property-based norm was established in the United States in the form of a *Code of Fair Information Practices* (the name being inspired by a code of fair labor practice) in 1973 as a voluntary guideline for government agencies handling data about citizens.[13] It subsequently became popular among many other agencies in the United States and overseas. The Code was enshrined in the 1980 Organisation for Economic Co-operation and Development (OECD) Privacy Guidelines[14] and ultimately as a binding legal document, the 1995 European Union (EU) Data Protection Directive.[15]

The escalation of fair information practices into a binding law was anticipated by a hugely influential 1983 decision on census data in which the Federal Constitution Court of West Germany decided that:

> [I]n the context of modern data processing, the protection of the individual against unlimited collection, storage, use and disclosure of his/her personal data is encompassed by the general personal rights of the German constitution. This basic right warrants in this respect the capacity of the individual to determine in principle the disclosure and use of his/her personal data. Limitations to this informational self-determination are allowed only in case of overriding public interest.[16]

In this judgment, the Court struck down a federal law authorizing the sharing of census data with other administrative and local agencies in the absence of the census respondents' consent to such sharing. Since then, the concept of data protection—

---

13. ROBERT GELLMAN, FAIR INFORMATION PRACTICES: A BASIC HISTORY 2–5 (2019), https://bobgellman.com/rgdocs/rg-FIPshistory.pdf.

14. ORG. FOR ECON. COOP. & DEV., OECD GUIDELINES ON THE PROTECTION OF PRIVACY AND TRANSBORDER FLOWS OF PERSONAL DATA (1980), https://www.oecd.org/sti/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm [hereinafter OECD GUIDELINES].

15. Directive 95/46/EC, of the European Parliament and of the Council of 24 October 1995 on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of Such Data, 1995 O.J. (L 281) 31.

16. BVerfG,1 BvR 209/83, Dec. 15, 1983, https://www.bundesverfassungsgericht.de/SharedDocs/Entscheidungen/EN/1983/12/rs19831215_1bvr020983en.html.

whereby all data, whether public or private, commendatory or defamatory, or true or false, are presumed to be under the data subjects' control—has been codified in the 2016 EU General Data Protection Regulation.[17]

Under the structures of data protection law, a data controller using data about another person is subject to the same restrictions as someone borrowing a car over a weekend for grocery shopping. To collect it (car or data), you will need the owner (data subject)'s consent. You cannot transfer it to a third party or use it for a new purpose (e.g., a long trip out of town) without the owner's updated consent. You will have to return it upon the owner's demand. You will have to let the owner inspect it if he or she so demands. Such an approach provides water-tight protection for anyone who does not want to be stressed about what others know about him or her against his or her wishes, whether it is true, false, commendatory, defamatory, confidential, or public. The project of nurturing "free development of personality" is thus consummated.

II. LIMITS OF OWNERSHIP-BASED DATA GOVERNANCE

The "all data are mine and therefore under my control" approach obviously has its drawbacks, lying as it does at the opposite end of the spectrum from the U.S. position, which rather too quickly deems information to be available for public consumption without data subjects' consent. No matter how publicly the information has been made available (either by the data subject's choice, by operation of law, or by natural social discourse) the person referred to in that information can claim control over it, exercising censorship on the social discourse around that information, even when the information is true and lawful in other respects. For instance, where tax information is already public as a matter of law, sharing that tax information with other people via SMS text message rapidly becomes controversial under data protection law.[18]

The framers of GDPR noticed this problem and tried to abate it by adding exemptions for "processing carried out for journalistic purposes or the purposes of academic, artistic or literary expression"[19] and by also adding that in cases where the data subject is a public figure or there is a public interest in the processing, such personal control by the data subjects is nullified.[20] However, the scope of "journalistic purposes," "academic[,] artistic or literary expressions," and "public interest" is not only hard to define (as evinced by the two conflicting decisions from

17. Regulation (EU) 2016/679, of the European Parliament and of the Council of 27 April 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 (General Data Protection Regulation), 2016 O.J. (L119) 1, 59 [hereinafter GDPR].

18. Case C-73-07, Tietosuojavaltuutettu v. Satakunnan Markkinapörssi Oy and Satamedia Oy, 2008 E.C.R. I-09831.

19. GDPR, *supra* note 17, art. 85.

20. GDPR, *supra* note 17, art. 6, § 1(e), (f).

the two highest courts in Europe[21]), but is also too narrow to protect people's right to expression and information. This is because requiring exchanges of information to fit certain collectivistically determined molds (e.g., public interest) contradicts the pluralistic ideal of that human right. A maxim "thou shalt speak only for public interest" is applicable only to broadcasters enjoying the state grant of monopoly. Furthermore, whether information has journalistic, artistic, or academic value, or is of public interest, can be determined only by people who have not yet accessed the information and yet whose access will be threatened by the data protection law. Without such access, the balancing between one's personality right and the public nature of information adds even more uncertainty that will have a chilling effect on those who wish to share the information.

The European proposal of giving all persons ownership of all personal information about themselves was heroically (or antiheroically, depending on your perspective) implemented in *Google Spain v. AEPD and Costeja*,[22] where the Court of Justice of the European Union (CJEU) ordered a search engine to drop legally required public notices of judicial auction of a data subject's house, which had taken place sixteen years prior, from its search results. For our purposes it is important to note that the CJEU itself specifically excluded any proof of "prejudice" in relation to the data subject as a requirement for issuing the "delisting" order and was unconcerned about the fact that the information was truthful (nondefamatory) and publicly available by operation of law (nonconfidential). For the CJEU, the decision was a simple extension of the 1995 EU Data Protection Directive's correction/deletion right for information that "became no longer relevant . . . by passage of time."

The decision met objections in Europe and around the world.[23] At this point it is worth recalling that the data protection law was an evolution of the personality right which protects one's right to personality development free from mental distress over what others know about or how they perceive one. The actual application of this right was supposed to depend on balancing privacy, the impact on data subjects, the public nature of the information (that is, other people's freedom of expression and information), and so on. The CJEU seemed to view the gravamen of the case as the fact that one of the top results seen by people Google-searching Mario Costeja's name was the sixteen-year-old house auction, which put

21. Satakunnan Markkinapörssi Oy and Satamedia Oy v. Finland, App. No. 931/13, Eur. Ct. H.R. (2017); Case C-73-07, Tietosuojavaltuutettu v. Satakunnan.

22. Case C-131/12, Google Spain SL v. Agencia Española de Protección de Datos, ECLI:EU:C:2014:317 (May 13, 2014).

23. LA QUADRATURE DU NET & REPORTERS WITHOUT BORDERS, RECOMMEND-ATIONS ON THE RIGHT TO BE FORGOTTEN 2 (2014), https://rsf.org/en/news/recommend ations-right-be-forgotten-la-quadrature-du-net-and-reporters-without-borders; *IFLA Statement on the Right to be Forgotten (2016)*, https://www.ifla.org/publications/node/10320; *IGF 2017 - Day 4 - Room IX - DC on Publicness*, https://www.intgovforum.org/multilingual/content/igf -2017-day-4-room-ix-dc-on-publicness (last visited Aug. 28, 2019).

him in worse light than his real financial condition. However, the CJEU did not require proof of any reputation-based "prejudice" against the data subject (not even false light!) when it enacted the delisting order, other than the fact that Costeja did not like it. Hence, reputational injury did not carry the day. Also, there was no privacy injury, since the information was as much publicly available as it could have been (i.e., the judicial auction was published in the newspaper advertisement).

The only remaining justification, lack of relevance, is merely an attempt to restate the conclusion: "Costeja is unfairly connected to irrelevant information." However, it is a poor justification, because relevance is not measured in a vacuum but only relative to a purpose. What is the purpose of a Google search using Costeja's name as the search word? The house auction sixteen years ago may not be relevant to someone interested in Costeja's current financial status but may be very much relevant to others interested in Costeja's previous financial status (for example, a researcher studying the socioeconomic conditions of Spanish judicial professions using anecdotal methodologies). Personality right does not mean any person's absolute right to be free of others' scrutiny but instead is subject to balancing. It is not clear what interest tipped the balance in Costeja's favour in this case.

Personality right cannot exist without its components and yet CJEU, too pre-occupied with incrementalism, did not even attempt to articulate which component of personality right was at work.

III. OWNERSHIP OF PUBLICLY AVAILABLE INFORMATION SCRUTINIZED

A tenet that "one owns data about him or her (and therefore should have control over that data)" sounds good but is not always sustainable and compatible with respect for others' freedom of thoughts and expressions. The sentence "K.S. Park is a professor" is data about me that is known to many already. That I, K.S. Park, can control circulation of such data will be impossible in a free society, especially after I have introduced myself as a professor to a countless number of people. When and under what grounds can I control perfectly lawful data that resides in another's head, that is non-defamatory and non-privacy-infringing?

Recall, the slogan that "one owns data about him or her" originates from the concept of "data surveillance," a term coined by Alan Westin in his 1967 book, *Privacy and Freedom*. The idea is that when a person discloses personal data to governments and companies, the processing of such data for purposes not contemplated by the data subject or the disclosure to agencies not thus contemplated can constitute surveillance in a sense that the data processor(s) will learn more information about the data subject than the data subject intended to give. For instance, the data processors can make predictions about the data subject's future behavior. Of course, the term "surveillance" usually means acquisition of data about another against his or her will, such as wiretapping or search and seizure. But even voluntary disclosures of data can, if the conditions of the

disclosures are not adhered to, lead to revealing something about oneself against his or her will, and hence the term "data surveillance."

Westin, in an effort to protect people from data surveillance, proposed giving all data subjects some sort of property right to the data about them because enforcing promises about how the data will be used is not sufficient: these promises would be difficult to enforce and, more importantly, powerless individuals would have a hard time demanding those promises from governments and companies. A property right, as opposed to a contractual right, compels the data processor to uphold affirmative duties to obtain consent from the data subjects whenever it takes the data even if the data subject fails to demand such duty. This includes consent on the purpose and scope of data use and disclosure, just as someone borrowing a car from another has affirmative duties to obtain consent from the car owner regarding its use and whether others may drive it. Since then, the property metaphor has hardened into refrains such as "owning data about oneself." Indeed, data protection law is a very effective tool for protecting the right of a powerless individual who, in disclosing data to a mega data processor, does not have the acumen to bargain or enforce the conditions of that disclosure.

In other words, the concept of data ownership was concocted to compensate for this market failure—unequal bargaining power at the point of disclosure. The goal in fixing this market failure is to prevent unwanted subsequent use and disclosure of data about individuals who have little to no say in the matter in making the *initial* disclosure. Given that this is the goal, it is very important that the concept of data ownership is not applied mechanistically to all data. *Publicly available* data has no point of disclosure where the concept of data ownership is needed to intervene to strengthen the data subjects' bargaining power. The paradigmatic situation for such intervention works like this: When a data subject has kept certain personal data within a zone of privacy and later transfers that data out of such a zone to a government or company, the data crosses the boundary between two zones. The concept of data ownership kicks in at that point of crossing to ensure that the data's subsequent use or disclosure does not depart from the data subject's original will. This protection is provided with a stronger force than that of contractual law.

This means that the ownership concept should not be applied to personal data that has already been published to the public on a voluntary basis without any condition. That "K.S. Park is a professor" is a perfect example of such data. In the same vein, the data lawfully compelled into disclosure by operation of law also should be free from data protection law's stricture (for instance, the publicly noticed data of a company, when such data is personal data of a shareholder) will be included. Such restriction on data protection is consistent with common sense: It is not surveillance to acquire data that everyone knows.

A closer look at the world's data protection laws already reveals a thread of such philosophy in Australia,[24] Canada,[25] Singapore,[26] Taiwan,[27] Belgium,[28] and pre-GDPR Germany,[29] which explicitly leave "publicly available data" out of the purview of data protection laws. The 2005 APEC Privacy Framework also states that a data subject's right can be limited with respect to "publicly available data."[30] In 2000, the EU and the United States entered into a safe harbor treaty on application of the 1994 EU Data Protection Directive on U.S. data processors, which also left out publicly available data.[31] In Germany, in a highly publicized *Spickmich* case, publicly available information on school teachers such as the courses and schools in which they taught were left out of protection pursuant to the then-effective data protection law.[32]

If we, as data subjects, can control even publicly available data about us —as is currently the case in the European data protection regime led by *Costeja v. Google*— then the data protection law aimed at preventing surveillance will instead cause us to not only censor but also conduct surveillance on one another. We need to watch what data others are acquiring and are able to intervene when we want. The goal was to protect privacy through data protection law, but the current data protection regime does not achieve that goal because of its focus on data ownership. Thus, we should stop talking about data ownership and start talking more about privacy, which was the original goal of concocting the metaphor to ownership.

My proposed framing would modify the result of decisions such as *Costeja* since that case involved information that was made publicly available by operation of law. One can even liken the facts in that case to voluntary disclosure because the forced public nature of the judicial auction was due to the policy objective of maximizing the number of respondents to the auction and therefore the price of the sale, which would contribute to the amount of Costeja's debt satisfied by the

24. *State and Territory Regulation of Privacy*, AUSTL. L. REFORM COMM'N (Aug. 16, 2020), http://www.alrc.gov.au/publications/2.%20Privacy%20Regulation%20in%20Australia/state-and-territory-regulation-privacy.

25. Personal Information Protection and Electronic Documents Act, S.C. 2000, c. 5, art 7 (Can.); Regulations Specifying Publicly Available Information, SOR/2001-7 (Can.).

26. Personal Data Protection Act 2012, No. 26 (Sing.).

27. Personal Data Protection Act, art 19, sec.7 (Taiwan).

28. Loi relative à la protection de la vie privée à l'égard des traitements de données à caractère personnel [Law on the Protection of Privacy in Relation to the Processing of Personal Data] of Dec. 8, 1992, MONITEUR BELGE [M.B.] [Official Gazette of Belgium], Mar. 18, 1993.

29. BUNDESDATENSCHUTZGESETZ [BDSG] [Federal Data Protection Act], Jan. 14, 2003, BGBL at 66, last amended by art. I, Aug. 14, 2009, BGBL at 2814 (Ger.), https://www.legislationline.org/download/id/5438/file/Germany_DataProtection_act_am2010_en.pdf.

30. ASIA-PAC. ECON. COOP., APEC PRIVACY FRAMEWORK 7 (2005), http://www.apec.org/Groups/Committee-on-Trade-and-Investment/~/media/Files/Groups/ ECSG/05_ecsg_ privacyframewk.ashx.

31. Commission Decision 2000/520, 2000 O.J. (L 215) 7 (EC).

32. *See* Claudia Kodde, *Germany's 'Right to be Forgotten' – Between the Freedom of Expression and the Right to Informational Self-Determination*, 30 INT'L REV. L., COMPS. & TECH. 24, 26–27 (2016).

auction. Costeja himself, at the time of judicial sale, would have voluntarily disseminated the fact of the auction as widely as possible for his own benefit even if the law did not require such disclosure.

IV. WHAT IS "PUBLICLY AVAILABLE"?

I proposed above that we should not blindly follow the impossible mantra that one owns data about oneself but should instead maintain our focus on the original legislative purpose of data protection law: privacy.[33] According to that proposal, I argued that publicly available information should not be subject to data protection law: Westin's proposal of ownership-based protection of privacy was concocted to intervene at the point of boundary-crossing from the data subject's control into the data collector's. Publicly available information does not have such an intervention point.

However, what is publicly available can be disputed. For example, a publication about an openly gay but otherwise private individual is treated differently depending on the jurisdiction.[34] There is serious dispute as to what constitutes a privacy violation, even between Europe and the United States. Is obtaining publicly available information about another person (a practice called profiling) a privacy violation? What about facilitating others' efforts to obtain such information (as a search engine does)? How about aggregating online, otherwise scattered, publicly available information on another person (as a reputation platform does)? With no transnationally reconcilable delineation of what is private and what is public, disputes over issues like these can implode the proposal to remove publicly available information from the structures of data protection law.

Especially concerning is the United States' position on privacy, which has been defined along the lines of "reasonable expectation of privacy"[35] in criminal procedure. The idea is that the law will protect the privacy of individuals by warrant protections only with respect to the things, persons, and places that people reasonably expect to be afforded privacy (i.e., when such expectation is reasonable given the efforts limiting other people's access thereto). However, a corollary of the objective reasonableness requirement is that one loses the right to privacy and is therefore deemed to have permitted a warrantless search by law enforcement of one's things, persons, and places if he or she leaves them in "plain view" of others.[36] For instance, even if one keeps the phone numbers one has dialed confidential only between one and the phone company, the Supreme Court has ruled that leaving the phone numbers available for observation by a "third party" (the phone company in

33. Kyung Sin ("K.S.") Park, *'One Owns Data About Oneself', Myth? Metaphor? Rule?: Implications for RTBF*, OPEN NET (KOREA) (Sept. 16, 2015), http://opennetkorea.org/en/wp/ 1421.

34. *See, e.g.*, Hauch, *supra* note 6, at 1220–21, 1254–55, 1288.

35. *See* Katz v. United States, 389 U.S. 347, 360–61 (1967) (Harlan, J., concurring).

36. Arizona v. Hicks, 480 U.S. 321, 326 (1987).

this case) renders one's expectation of privacy in this regard unreasonable, making them available for warrantless search by law enforcement.[37]

This so-called "third-party" doctrine has been seriously challenged by the U.S. Congress, which, through the Electronic Communication Privacy Act of 1986 (ECPA), added pen register/trap and trace provisions. These provisions permit law enforcement agencies to trace telephone numbers which called or were called from the suspect's phone only upon the issuance of court orders.[38] However, the third-party doctrine remains good law to the extent that, under other parts of ECPA, all information stored on a third-party server for more than 180 days, as well as subscriber information submitted to communication service providers, can be accessed by law enforcement without any judicial approval.[39] Fortunately, at least one intermediate appellate court found otherwise. The Sixth Circuit found that the contents of emails are private and thus require a warrant for law enforcement access.[40]

The idea of privacy as a protection *against the eyes of fellow citizens*, as opposed to law enforcement, had an entirely different root in the United States: Warren and Brandeis's seminal article of 1890.[41] The authors proposed creating, for the first time, a legal theory whereby one would be held liable to another for describing (or disclosing facts about) another person against his or her will, even if such a description or disclosure does not invade the other party's reputation by libel or interference with the other party's property right. The authors, while drawing a clear line at the ever influential maxim that "[t]he right to privacy does not prohibit any publication of matter which is of public or general interest," also argued that "to publish of a modest and retiring individual that he suffers from an impediment in his speech or that he cannot spell correctly, is an unwarranted, if not an unexampled, infringement of his rights."[42] Here, the connotation is that publishing facts not held confidential (i.e., speech or literary impediment) may still violate privacy. So, according to Warren and Brandeis, the third-party doctrine does not hold for private individuals' publication about others.

Decades later, Prosser conducted a comprehensive analysis of almost all cases mentioning the term "privacy" after the article by Warren and Brandeis had been published. Prosser concluded that the seventy-year development of relevant court decisions had crystallized into four theories, yet Prosser recognized only one of these theories as a legitimate avenue for further enquiry and suggested "calling [the rest] to a halt."[43] The theory he approved was the one based on the idea of "public disclosure of *private* facts," which requires publication of facts previously shrouded

37. Smith v. Maryland, 442 U.S. 735, 744–46 (1979).
38. Electronic Communication Privacy Act of 1986, 18 U.S.C. § 3121 (2020).
39. Stored Communications Act, 18 U.S.C. § 2703 (2020).
40. *See, e.g.*, United States v. Warshak, 631 F.3d 266, 282, 285–86, 288 (6th Cir. 2010).
41. Samuel D. Warren & Louis Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193 (1890).
42. *Id.* at 214–15.
43. William L. Prosser, *Privacy*, 48 CAL. L. REV. 383, 389, 398, 401 (1960).

from public observation. The three other judicial evolutions (that is, *intrusion*, *misappropriation*, and *false light*) concern physical invasion, commercial interest, and reputation—interests which have been developed more readily and prominently in other areas of law such as trespass, publicity rights, and defamation, respectively. Prosser did not delve into what is considered private and not private, leaving unclear whether the proposal by Warren and Brandeis is consistent or conflicting with the third-party doctrine.

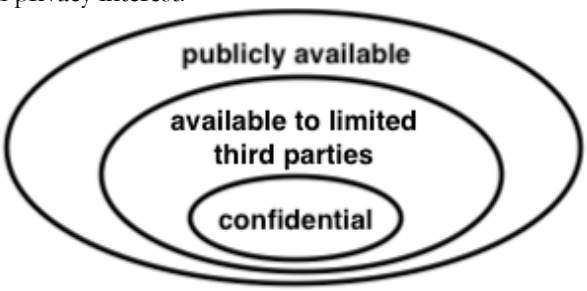
Since then, the U.S. jurisprudence on privacy has remained relatively static. On the one hand, it has consisted of the draconian third-party doctrine (which disregarded privacy whenever voluntary or unavoidable disclosure was made to any third party and was tamed partially by the ECPA) and, on the other hand, the vague Warren-Brandeis proposal that private publication of non-confidential facts may still violate privacy. One may make distinctions between the *publicly-interested* state's mere *access* to the information and a private gratuitous intermeddler's invasive *publication* of the same, and apply the third-party doctrine strictly to the former and leniently to the latter.

Depending on how the discrepancy in the transnational definition of "publicly available information" is resolved, there is a risk that the proposed exception of publicly available information may swallow the rule or make the entire data protection law moot. The goal of data protection law was to broaden privacy beyond confidential information to include the information voluntarily made available to third parties. Strict application of the third-party doctrine in defining what is publicly available will leave all voluntarily transferred information outside the purview of a data subject's privacy. A culture of privacy built on the third-party doctrine is a barren soil for any conception of data protection law aiming to give data subjects control over data about them after the data leave their physical dominion.

Fortunately, at least three times recently, the U.S. Supreme Court has adjusted its answers to the question of whether aggregating information available to numerous third parties may constitute a privacy violation: first in *United States v. Jones*,[44] where the police left a GPS device attached to the suspect's car for a whopping twenty-eight days (avoiding the need to follow him around on public streets) and then in *United States v. Riley*,[45] where the police, in conducting a lawful search incident to arrest, searched the contents of an arrestee's cell phone. In each case, a majority of the Justices approved, though only informally in *Jones*, the idea that modern communication technology increases the collection of and processing of information about people to such an extraordinary scale that some restraint is needed in order to protect individual privacy. The Court found this to be true even in connection with information otherwise already accessible by "third parties" (as

44. United States v. Jones, 565 U.S. 400, 402–03 (2012).
45. Riley v. California, 573 U.S. 373, 378–79 (2014).

other people who have witnessed the suspect's car in *Jones* or as law enforcement in *Riley*). Most recently, in *Carpenter v. United States*,[46] the Supreme Court finally chipped away at the third-party doctrine by imposing the warrant requirement on the geolocation-revealing component of metadata, that is, the cell site location information tied to particular users, which is already accessible by "third parties." These are positive steps toward European incrementalism which recognizes that all personal information, unless it is completely publicly available, has some, even if minimal, residual privacy interest.



Jettisoning the third-party doctrine in favor of incrementalism does not mean that transnational differences will disappear. I believe these differences will persist due to cultural reasons. For instance, the opposite results on the forced coming-out case may still arise. However, the difference will be that of line drawing, not categorical. American judges may still find public street marches devoid of privacy interest, not because of the mere presence of others' plain view but as a result of the incremental analysis of how many people were there, how open the data subjects' actions were, etc. Yet this transnational ordering, which leaves out publicly available data, makes it possible to reestablish privacy as the gravamen of data protection law. It does so without demolishing the great legal invention of the twentieth century: data protection laws.

V. USE OF NON-PUBLIC PERSONAL DATA FOR COMMUNITY PURPOSES

Publicly available information is not the only information on which the data ownership metaphor ought to and does loosen its grip. GDPR loosens ownership-like control on a subject's personal data when it exempts the consent requirement for the use of data when the data is used for the *communal* purposes of the society.

For a starter, data controllers may subject personal data to "further processing" (i.e., processing for a new purpose—"off-purpose processing" or "repurposing," as I call it) as long as it is processed "in a manner that is not

46. Carpenter v. United States, 138 S. Ct. 2206, 2219–21 (2018).

incompatible with" the original purposes for which the data was collected.[47] This phrase connotes a much broader scope of further processing beyond the original purpose compared to analogous provisions in other data privacy laws, such as "reasonably expected from [the original purpose]" in Singapore, "duly related to [the original purpose]" in Japan, "within the scope of [the original purpose]" in South Korea, "for the purpose collected" in India,[48] and "for the purpose collected" in pre-GDPR Germany.[49] The Article 29 Working Party interpreting the Data Protection Directive made clear that there will be a balancing of social interests and individual interests in setting the scope of "not incompatible use" as follows:

> Rather than imposing a requirement of compatibility, the legislator chose a double negation: it prohibited incompatibility. By providing that any further processing is authorized as long as it is not incompatible (and if the requirements of lawfulness are simultaneously also fulfilled), it would appear that the legislators intended to give some flexibility with regard to further use. Such further use may fit closely with the initial purpose or be different. The fact that the further processing is for a different purpose does not necessarily mean that it is automatically incompatible: this needs to be assessed on a case-by-case basis, as will be shown below. In some situations, this additional flexibility may be needed to allow for a change of scope or focus in situations where *the expectations of society*—or of the data subjects themselves—have changed about what additional use the data may be put to. It is also possible that when initially specifying the purpose, neither the controller nor the data subject thought additional purposes would be necessary, although it subsequently transpired that the data could indeed be very useful for other things. In some of these (and similar) situations, a change of purpose may be permissible, and further processing may be considered not incompatible, provided that the compatibility test is satisfied.[50]

The reasoning based on *changed expectations of society* gave birth to a new affordance of non-consensual use, and under other provisions of GDPR, further processing for "public interested archiving, scientific and historical research, and statistics" are

47. GDPR, *supra* note 17, art. 5 ("Article 5 Principles relating to processing of personal data. 1. Personal data shall be (a) processed lawfully, fairly and in a transparent manner in relation to the data subject ('lawfulness, fairness and transparency'); (b) collected for specified, explicit and legitimate purposes and not further processed *in a manner that is incompatible with those purposes*; . . . ('purpose limitation') . . . .") (emphasis added).

48. GRAHAM GREENLEAF, ASIAN DATA PRIVACY LAWS: TRADE AND HUMAN RIGHTS PERSPECTIVES 295, 242, 141, 418 (1st ed. 2014).

49. *See* BDSG, Jan. 14, 2003, BGBL I at 66, last amended by Gesetz [G], Aug.14, 2009, BGBL I at 2814 (Ger.) (repealed 2016).

50. *Opinion of the Article 29 Data Protection Working Party on "Purpose Limitation"*, 2013 O.J. (WP 203) 1, 21 (emphasis added).

across the board considered not "incompatible with the initial purpose."[51] Such a generous exemption for non-consensual use is moderated by the additional requirement of data minimization which may be satisfied by pseudonymization.[52] Furthermore, as for scientific research, consent does not have to be as specific as in other situations. This means that further processing for the purpose of scientific research is more freely permitted in reliance on the breadth of the original scientific research. [53] In order to protect the communal value that such off-purpose processing generates, ancillary data protection rights, such as rights to access, rectification, restriction of processing, and objection to processing, are left subject to national derogations.[54]

The social nature of these exemptions is apparent: the definition of "scientific research" refers to the Treaty on the Functioning of the European Union as a guidepost and the referenced article promotes "strengthening its scientific and technological bases by achieving *a European research area in which researchers, scientific knowledge and technology circulate freely*, and encouraging it to become more competitive, including in its industry."[55] The European Data Protection Board (EDPB, formerly

51. GDPR, *supra* note 17, art. 5 ("Article 5 Principles relating to processing of personal data. 1. . . . (b) . . . further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall, in accordance with Article 89(1), not be considered to be incompatible with the initial purposes. . . .").

52. GDPR, *supra* note 17, art. 89 ("Article 89 Safeguards and derogations relating to processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes. . . . 1. Processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes, shall be subject to appropriate safeguards, in accordance with this Regulation, for the rights and freedoms of the data subject. Those safeguards shall ensure that technical and organisational measures are in place in particular in order to ensure respect for the principle of data minimisation. Those measures may include pseudonymisation provided that those purposes can be fulfilled in that manner. Where those purposes can be fulfilled by further processing which does not permit or no longer permits the identification of data subjects, those purposes shall be fulfilled in that manner.").

53. *See* GDPR, *supra* note 17, ¶ 33 ("It is often not possible to fully identify the purpose of personal data processing for scientific research purposes at the time of data collection. Therefore, data subjects should be allowed to give their consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research. Data subjects should have the opportunity to give their consent only to certain areas of research or parts of research projects to the extent allowed by the intended purpose.").

54. *Id.* art. 89(2).

55. *Id.* ¶ 159 ("Where personal data are processed for scientific research purposes, this Regulation should also apply to that processing. For the purposes of this Regulation, the processing of personal data for scientific research purposes should be interpreted in a broad manner including for example technological development and demonstration, fundamental research, applied research and privately funded research. *In addition, it should take into account the Union's objective under Article 179(1) TFEU of achieving a European Research Area*. Scientific research purposes should also include studies conducted in the public interest in the area of public health. To meet the specificities of processing personal data for scientific research purposes, specific conditions should apply in particular as regards the publication or otherwise disclosure of personal data in the context of scientific research purposes. If the result of scientific research in particular in the health context gives reason for further measures in the interest of

known as the Article 29 Working Party) defines the term "scientific research" to mean "a research project set up in accordance with relevant sector-related methodological and ethical standards, in conformity with good practice," [56] implying a requirement of minimum disclosure to and oversight by the sectoral ethical bodies. Also, European Data Protection Supervisor said: "the special data protection regime for scientific research is understood to apply where . . . the research is carried out *with the aim of growing society's collective knowledge and wellbeing*, as opposed to serving primarily one or several private interests.[57] Similarly, in defining "research" for the purpose of the GDPR exception, the Swedish Ethical Review Authority has taken into account "whether the research will be conducted by researcher(s) and/or *if the results of the research will be made available to the public, e.g., by publishing the results in scientific journals*."[58]

To reiterate, data ownership was concocted to create a default rule that one cannot collect or control data about another without explicit consent to such collection or use. The purpose of this default rule is to protect people without sufficient acumen to impose or negotiate conditions on the disclosure of data about themselves, such as the purpose of use or the number of people with access to the data. I have argued that such legal effects may not be necessary for already publicly available data. Also, I just established that GDPR allows certain forms of non-consensual off-purpose processing under the justification of serving certain *communal* purposes such as science and history, even when the data are not publicly available. Note that these consent exemptions for research, statistics, and other communal purposes apply to information that has not been made publicly available as well as publicly available information.

Caution: This should not be seen as evidence that GDPR digresses from its original purposes to protect privacy and prevent surveillance. The global assumption throughout this discussion is that data protection laws were created to provide enhanced protection for privacy, i.e., protection for the data that has already left the data subject's physical dominion out of his or her own volition. The parallel assumption is that if there is any involuntary extraction of otherwise confidential data, whether for purposes of scientific research or not, other laws such as wiretapping law or civil tort of invasion of privacy will intervene to address such harms even if GDPR itself does not punish such action.

---

the data subject, the general rules of this Regulation should apply in view of those measures.") (emphasis added).

56. Eur. Data Prot. Bd. [EDPB], Guidelines 05/2020 on Consent Under Regulation 2016/679, ¶ 153 (May 4, 2020).

57. EUR. DATA PROT. SUPERVISOR, A PRELIMINARY OPINION ON DATA PROTECTION AND SCIENTIFIC RESEARCH 12 (2020) (emphasis added), https://edps.europa.eu/sites/default/files/publication/20-01-06_opinion_research_en.pdf.

58. Fredrik Roos, *Processing of Personal Data for Research Purposes*, SETTERWALLS (May 16, 2019) (emphasis added), https://setterwalls.se/aktuellt/artikel/processing-personal-data-research -purposes.

"Scientific research" exemption inspires us to think about ways to maximize publicly beneficial use of personal data by freeing them from the strictures of private ownership and control.

VI. PRIVATE OWNERSHIP OF DATA

One may wonder why we should adhere to what may be viewed as a parsimonious interpretation of data ownership. Why can't we just declare that we own data about ourselves or each of us owns data about him or herself? Or, at least we could start from that position and allow erosions and derogations on a case-by-case basis in order to accommodate external interests, such as right to know and freedom of expression?

In this Section, I delve deeper into the nature of data and data ownership to provide a set of endogenous justifications for limiting the applicability of data ownership as I proposed above.

There is no doubt that data, though non-rivalrous, can be "owned" just as copyright, patents, and other non-rivalrous intellectual properties can be owned. My observations here are directed more toward the other natures of data that resist ownership in favor of data subjects or in favor of data controllers. So far in this article we have talked about the former sense of ownership (i.e., in favor of data subjects), but, on the other end of the spectrum, there have been attempts to strengthen data ownership in the latter sense.[59] My observations should add to the preemptive warnings against those attempts as well.[60]

Firstly, data are the result of perception of a sensed object by a sensing agent. The earth itself is not data. Its presence translates into data only after there is that interaction between the object—the earth—and sentient beings sensing that object. Data is created to be shared among sentient beings as surrogates for perception of the entities or their features that the data describe. For example, blood types precede the data used to describe them, "A," "B," "AB," and "O." Then, if data are the result of that interaction, how much control should the sensed being or the sensing being have on the result of that interaction? This should be an open question, not a deontological absolute.[61]

---

59. *See* Osborne Clarke LLP, Final Report on the Legal Study on Ownership and Access to Data, at 13–14, 27 (2016), http://publications.europa.eu/resource/cellar/d0bec895-b603-11e6 -9e3c-01aa75ed71a1.0001.01/DOC_1.

60. *See* JOSEF DREXL, RETO M. HILTY, LUC DESAUNETTES, FRANZISKA GREINER, DARIA KIM, HEIKO RICHTER, GINTARE SURBLYTE & KLAUS WIEDEMANN, DATA OWNERSHIP AND ACCESS TO DATA: POSITION STATEMENT OF THE MAX PLANCK INSTITUTE FOR INNOVATION AND COMPETITION OF 16 AUGUST 2016 ON THE CURRENT EUROPEAN DEBATE 4 (Max Planck Inst. for Innovation & Competition Rsch. Paper No. 16-10 ed., 2016).

61. *See, e.g.*, Cameron F. Kerry & John B. Morris, Jr., *Why Data Ownership is the Wrong Approach to Protecting Privacy*, BROOKINGS (June 26, 2019), https://www.brookings.edu/blog/techtank/2019/06/26/why-data-ownership-is-the-wrong-approach-to-protecting-privacy/.

Much of human civilization is data transfer, and often, transfer of personal data. What is poetry? An attempt to describe one's perception (data) of the things around him or her in succinct and beautiful words. If you write a poem about a mountain, it is data about that mountain. If you write a poem about another person, it is personal data about that person. Should that person own (or control as if to own) that poem or should its author? It is in this same vein that a pro-privacy scholar Jerry Kang has acknowledged the tension between copyright and privacy.[62]

As long as data is the result of interaction between you and another person who perceived you or your features, there is no deontological reason why you should own (or control as if you own) data about yourself.

Secondly, data is produced for the purpose of diffusion. As explained above, existences precede data. We make deliberate efforts to create the data in order to share perceptions of real existences with others, e.g., blood types created to inform medical professionals treating us. Now, the fact that data is created for transferability or shareability does not compel us to withdraw a norm against such transfer or sharing when the subject matter of the data is personal and its sharing or transfer violates the data subject's privacy. However, my point is that data, including personal data, is created with the assumption that other people (i.e., other than the data subject) will receive and share it.

The prime examples are the daily uses of our names and faces with respect to people around us. We use names in order to efficiently refer to individuals. Instead of calling me "that curly-haired immigrant from South Korea in 1986 who later studied physics in Cambridge and law in Los Angeles who now teaches back in Seoul," you can call me "Kyung Sin Park." When I introduce myself, the first personal data I share is my name. We all don ourselves with the dandiest outfits every morning but always reveal our faces whether days are hot or cold, all in order to be efficiently recognized by others as our individual selves. Of course, when we do not want to be recognized, we wear masks, and when we do not want to be introduced by names, we refuse to give our names, but that is another story.

When the data created for diffusion by data subjects are used for the very purpose of creation, i.e., diffusion, we should not become too overzealous in trying to gain control over those personal data, all under the slogan of data ownership. To what extent can we say that we own our faces or names when we threw them out there into the multitude of people around us so they can call us, refer to us, and visually recognize us through those faces and names? Who are the people who collect the facial images and names for identification purposes? Certainly not us.

Thirdly, some personal data are inherently produced *relationally*. Let's take an example of the data that I am a professor. I can be a professor only because there are students willing to listen to my lecture. You cannot be an attorney in a vacuum

62. *See* Jerry Kang & Benedikt Buchner, *Privacy in Atlantis*, 18 HARV. J.L. & TECH. 229, 241–43, 255 (2004).

but only in relation to clients, hypothetical or real. Doctors' identities equally rely on a relationship to patients. In this sense, much personal data are thought to be created by presupposing relations among people. Much data thought to be "about us" are actually about our relationship with others.

Unless one side can claim exclusive dominion over that relationship, it seems *unfair* to grant one person dominion over the data corresponding to that relationship (i.e., I am a professor, or you are an attorney). After I give lectures to numerous students, I can technically invoke my data protection right to stop my former students from commenting on me in teacher reputation platforms such as on *Spickmich*, as we discussed earlier. Such invocation may be unfair in the sense that my identity as a professor came into being due to contributions from the students who suffered through my lectures. The same goes for the personal data that you are a lawyer.

I am not attempting to make a Marxist argument that contribution to creation of a thing should decide who controls it. I am saying that there is no deontological reason why one should have control over relationally-originated personal data to the exclusion of control of others in that relationship. Such relationally-originated data, even if not publicly available, should be carefully reevaluated: there may not be any meaningful market failure that calls for intervention by the concept of data ownership.

To further emphasize the problem of unfairness, think of the personal data that X is a domestic violence offender. His existence presupposes a relationship or a transaction with his victim, and the data that he is an offender was created *relationally* or *transactionally*. We do not let that offender censor the victim into silence, even if the victim's public accusations of the offender will infringe the offender's right under data protection laws to control data about himself; we think that the victim should also be granted *some dominion* over that data. We usually rely on a "public interest" defense in exempting such crime victims' voices from the data ownership of criminals and the core of that "public interest" is the freedom of speech of the current victim and "the right to know" of potential victims for the future. However, before we even apply this "public interest" exception, we should re-examine the rule itself: Just as the data that I am a professor originates from a relationship with my students, the data that X is a domestic violence offender cannot be wholly owned by him because it originates from X's relationship or transaction with his victim. A statement that "X hit Y" is as much victim Y's personal data as it is X's.

The problem of unfairness even affects statements that are seemingly non-personal. For instance, a statement that "A had lunch with B at restaurant Z" is not just A's personal data or B's personal data but also potentially C's personal data when C is the owner of the Z restaurant." These two statements taken together are semantically indistinguishable from another statement "C owns a restaurant Z where A and B had lunch." Given this, it seems inappropriate to give only A or B

exclusive dominion over this third statement (data), especially when C would like to disclose that data to others.

This insight is important because all data, even seemingly non-personal, are potentially personal data, and calling something someone's personal data immediately puts the data under the strictures of that data subjects' control. Even a mountain's location is not just that mountain's location. There must be people who own or live near that mountain. A mountain's location is the location of something that those people own or live near and therefore becomes personal data of the owners and the habitants. For instance, suppose that a murdered body was found on a mountain owned by a data subject. Technically, the data subject can argue that the fact that the body was found on his mountain is his personal data and therefore he can regulate how that data is circulated. The data subject may well have an incentive to do so to keep land prices from falling.

The concept of *joint ownership* may be invoked to answer my challenge. In real property law, transfer of joint tenancy requires the approval of all joint tenants and therefore, when more than two people have jointly originated personal data, one will need consent from all joint owners for collection and further processing. However, will this be true when I have taught thousands of students? Do I really deserve ownership-like control over the fact that I am a professor to the point of silencing each of my students who would be able to disclose this fact to non-students only by obtaining consent from me?

Fourthly and finally, what does it mean to "own" something? Exclusive possession and control do not suffice because you can obtain that by contract. In the 1980s, Hegelian scholars met in the United States to figure this out. Ironically—and somewhat tautologically—they concluded that when we say someone owns something, what we mean is that he or she can disown it. This means that the "owner" has the legal right to transfer ownership to another person, however restricted that ownership may be.[63] Both real properties and intellectual properties satisfy this requirement. If you are the author of a book, and therefore owner of the intellectual properties embodied in that book, your ownership means that you can transfer whatever rights and privileges you have over the book to another person such that the other person will have the same rights and privileges, even to the exclusion of the transferor. This is a legal feat that no other person can achieve other than the so-called owner of the book.[64] This means that ownership may have been created to facilitate the acts of disowning, in other words, transfer to another. Also, the Rule Against Perpetuities and other property rules show that ownership is a concept that actually facilitates free use of the owned thing for any purpose that

63. Peter Benson, *The Priority of Abstract Right and Constructivism in Hegel's Legal Philosophy*, *in* HEGEL AND LEGAL THEORY 188–91 (Drucilla Cornell, Michel Rosenfeld & David Gray Carlson eds., 1991).

64. Some argue that data cannot be owned because of its non-exclusive and non-rivalrous nature. As shown in this hypothetical, data can be owned just as copyright can be owned.

the owner desires. Maybe, ownership exists not just to prevent use and transfer but to promote use and transfer, an objective traditionally seen as antithetical to privacy, which is often equated with "the right to be left alone." Maybe, privacy is not the only axiomatic value that we should try to protect; maybe privacy is not the primary indicator by which we should define market failure.

Indeed, data portability is both a value and a norm enshrined in GDPR.[65] It relies on the autonomy of the data subject. However, if a data subject's failure to impose or negotiate future use of his personal data stems from a market failure, it is equally possible that the data subject's decision to port the data from one service to another will be based on flawed autonomy (e.g., lack of sufficient information) and therefore constitute a market failure as well. That GDPR stipulates the right to data portability despite this risk demonstrates that ownership, if we deploy it at all in data governance, should not be understood as a presumption against portability. A strong recognition of one's ownership of data about him or herself may require implementing the data subject's reasonably inferred desire to offer their data for communal purposes, obviating "ownership"-based restrictions.[66]

In a surprising survey of 700 outpatients of a northern German university hospital, data subjects showed not only strong willingness to give broad consent for secondary data use of their personal data (468 of 503 responders, 93.0%) but even strong approval of abolishing patient consent (n = 381, 75.7%) for such usage.[67] GDPR's "scientific research" exemptions may have been attempts to codify this apparent reverse presumption for portability and repurposing with respect to a subset of data, and my proposal for data socialism is an extension of such attempts.

CONCLUSION: WHICH IS A MORE PROGRESSIVE VISION OF DATA GOVERNANCE?

The idea of private land ownership originated from our experience with the tragedy of the commons.[68] In response to similar inefficiencies, we created institutions such as intellectual properties. Like these two cases (land ownership and copyright/patent ownership), data ownership was concocted to respond to a type of market failure. In order to call something a failure, we need to have axiomatic values by which it is evaluated. In the case of the tragedy of the commons, the axiomatic value was production of livestock. In the case of copyrights and patents, the axiomatic value was advances in science and the arts. With data ownership, the axiomatic value is to protect privacy.

---

65. GDPR, *supra* note 17, art. 20.

66. NATCEN SOC. RSCH., PUBLIC ATTITUDES TO DATA LINKAGES 3 (2018) (showing social services recipients' strong desire to make available their data for socially beneficial research purposes).

67. Gesine Richter, Christoph Borzikowsky, Wolfgang Lieb, Stefan Schreiber, Michael Krawczak & Alena Buyx *Patient Views on Research Use of Clinical Data Without Consent: Legal, But Also Acceptable?*, 27 EUR. J. HUM. GENETICS 841, 841 (2019).

68. Garrett Hardin, *The Tragedy of the Commons*, 162 SCIENCE 1243 (1968).

Data ownership addresses a specific *market failure*—the inadequate protection of privacy that occurs when individuals lack the power to propose or negotiate the permitted scope of using and transferring data about them at the point of making a disclosure. However, as I have shown, the data ownership rule is unnecessarily over-protective with respect to publicly available data. I have demonstrated affirmative endogenous reasons why we should limit the influence of the data ownership metaphor in protecting privacy. As discussed, ownership may not be a good fit for protecting privacy because (1) data may have an author whose claims compete with data subjects; (2) data may have been created for diffusion after which it will be awkward for data subjects to maintain control; (3) data may have relational/transactional origins that make any one individual's exclusive control unfair to other participants in those relationship and transactions; and (4) ownership itself may include within it an inherent goal toward diffusion and repurposing, which may explain why, for a subset of data used for "scientific research," GDPR, while being ownership-based in the overall architecture, exempts transfer and repurposing across the board."

Data ownership, introduced to address a market failure, may itself cause another market failure: the "tragedy of the data commons."[69] The traditional tragedy of the commons describes self-interested actors incentivized to overuse the common resources, leading to depletion for all. The tragedy of the *data commons* describes self-interested actors incentivized *not* to make data available about themselves to the point of rendering research of common benefit to all impossible.

> Each individual has an incentive to remove her data from the commons to avoid remote risks of reidentification. This way she gets the best of both worlds: her data is safe, and she also receives the indirect benefits of helpful health and policy research performed on the rest of the data left in the commons. However, the collective benefits derived from the data commons will rapidly degenerate if data subjects opt out to protect themselves. . . . The bulk of privacy scholarship has had the deleterious effect of exacerbating public distrust in research data. Rather than encouraging the public to fervently guard their self-interest, scholars should build a sense of civic responsibility to pay their "information taxes" and participate in research datasets.[70]

According to Fred Cate, state consumer statutes requiring consent for collection of data may be the cause of this sort of market failure.[71] Although Cate refers to anonymized and therefore privacy-free data, the argument might as well apply equally to personal data at lower levels of de-identification (i.e., pseudonymized data). As a matter of fact, Cate does not even need to make the above argument for anonymized data because opting-out of anonymized data sets is either impossible

69. Jane Yakowitz, *Tragedy of the Data Commons*, 25 HARV. J.L. & TECH. 1 (2011).
70. *Id.* at 4–5.
71. FRED H. CATE, DATA AND DEMOCRACY 15 (Ind. Univ. ed., 2001).

(since the data no longer has an owner) or unnecessary (since there is no risk posed to one's privacy). Cate's argument made originally for anonymized data applies well to processing pre-anonymized data as long as the privacy risk is properly abated. Indeed, Germany's post-GDPR data protection law does not even require pseudonymization or anonymization of personal data for non-consensual further processing for research purposes.[72]

In this sense, data may be not only inconducive to ownership in certain contexts, but it may even be politically *libertarian* (in the sense of over-sacrificing the societal good in favor of individual liberties) to recognize data ownership. On the opposite end of the political spectrum from libertarianism sits socialism, an idea that some essential things, such as housing and minimum subsistence, should be left to public control for the maximum benefit of all. Data socialism, then, would be the idea that data is one of such essential commodities or infrastructures that need to be left to public control, namely public properties.

While this may sound like a radical idea, data socialism has been already proposed in different contexts though from a different motive. India recently proposed to compel data controllers to share non-personal data with other data controllers.[73] Although the proposal is limited to non-personal data, much of the non-personal data will be harvested by anonymizing personal data. It is a legislative response to the potential harms of "data monopolies" by *data controllers*.[74] Also, the International Committee of Medical Journal Editors (ICMJE) proposed that all submissions for publication be accompanied by the underlying raw data, properly deidentified. Upon request, this data will be made available to other researchers no later than six months after publication.[75] ICMJE believes there is an ethical duty owed to research participants to share the benefit of clinical trials with the greater society.[76] Both initiatives address the perils of data ownership being monopolized by data controllers. In parallel, we might consider what a legislative response would

72. *See* BDSG, June 30, 2017, BGBL I at 2097, last amended by Gesetz [G], Nov. 20, 2019, BGBL I at 1626, art. 12 (Ger.).

73. MINISTRY OF ELEC. & INFO. TECH. GOV'T OF INDIA, REPORT BY THE COMMITTEE OF EXPERTS ON NON-PERSONAL DATA GOVERNANCE FRAMEWORK (2020).

74. *Id.*

75. Darren B. Taichman, Joyce Backus, Christopher Baethge, Howard Bauchner, Peter W. de Leeuw, Jeffrey M. Drazen, John Fletcher, Frank A. Frizelle, Trish Groves, Abraham Haileamlak, Astrid James, Christine Laine, Larry Peiperl, Anja Pinborg, Peush Sahni & Sinan Wu, *Sharing Clinical Trial Data: A Proposal From the International Committee of Medical Journal Editors*, ANNALS OF INTERNAL MED. (2016), http://www.icmje.org/news-and-editorials/M15-2928-PAP.pdf; INT'L COMM. OF MED. JOURNAL EDS. (ICMJE), RECOMMENDATIONS FOR THE CONDUCT, REPORTING, EDITING, AND PUBLICATION OF SCHOLARLY WORK IN MEDICAL JOURNALS (2019), http://www.icmje.org/recommendations/.

76. Thomas Sullivan, *ICMEJ Proposes Data Socialism – Data Utopianism Has Its Cracks – Comments Due April 18*, POL'Y & MED., https://www.policymed.com/2016/01/icmej-proposes-data- socialism-data- utopianism- has- its -cracks- comments -due- april- 18-1 .html#respond (last updated May 5, 2018).

look like to prevent the tragedy of the data commons caused by a widely shared entrenchment of data ownership by *data subjects*.

Currently, government agencies and companies justify building closed silos of personal data and not sharing them with people, citing none other than the concerns of data protection laws.[77] Some of these concerns are justified but other concerns must be moderated with or balanced against the people's need for that data to achieve better democracy and a better society.

Some governments have already made positive improvements on the existing ownership-based data protection law by carving out a family of personal data that can be freely used for social discourse such as "publicly available data" as recognized by Singapore, Taiwan, Canada, and Australia or a family of exempted use such as "scientific research" as recognized by GDPR. We need more of this. Granting only case-by-case exceptions, as the current data protection laws do, would maintain chilling effects on people wishing to use the data for other socially beneficial uses. Further categorical exceptions (e.g., court judgment databases as in Canada and Australia) need to be carved out so that within these exceptions, people can engage in research and discourse benefitting society without worrying whether their research and discourse satisfies any then-current collectivistic (or majoritarian) notions of "public interest." These collectivistic notions of "public interest" can often crush pluralistic visions of a society. In order for such categorical changes to take place, we need to squarely face the libertarian perils of data ownership and further examine the progressive vision of data commons.

77. The Korean judiciary does not disclose court judgments to the public unless the costly anonymization process is completed for each of the judgments requested to be disclosed. *See* Minsasosongbeob [Civil Procedure Act], Act No. 547, Apr. 4, 1960, *completely amended by* Act No. 6626, Jan. 26, 2002, *amended by* Act No. 10859, July 18, 2011, art. 163-2 para. 2 (S. Kor.), *translated in* Korea Legislation Research Institute's online database, https://law.go.kr/LSW/eng/engMain.d o (search required); Hyeongsasosongbeob [Criminal Procedure Act], Act No. 341, Sept. 23, 1954, *amended by* Act No. 10864, July 18, 2011, art. 59-3 para. 2 (S. Kor.), *translated in* https://law.go.kr/LSW/eng/engMain.do (search required).