Buffalo Law Review

Volume 68 | Number 2 Article 4

4-1-2020

Data Management Law for the 2020s: The Lost Origins and the New Needs

Przemysław Pałka
Yale Law School

Buffalo Law Review

VOLUME 68 APRIL 2020 NUMBER 2

Data Management Law for the 2020s: The Lost Origins and the New Needs

PRZEMYSŁAW PAŁKA†

In the data analytics society, each individual's disclosure of personal information imposes costs on others. This disclosure enables companies, deploying novel forms of data analytics, to infer new knowledge about other people and to use this knowledge to engage in potentially harmful activities. These harms go beyond privacy and include difficult to detect price discrimination, preference manipulation, and even social exclusion. Currently existing, individual-focused, data protection regimes leave law unable to account for these social costs or to manage them.

This Article suggests a way out, by proposing to re-conceptualize the problem of social costs of data analytics through the new frame of "data management law." It offers a critical comparison of the two existing models of data governance: the American "notice and choice" approach and the European "personal data protection" regime (currently expressed in the General Data Protection Regulation). Tracing their origin to a single report issued in 1973, the Article demonstrates how they developed differently under the influence of different ideologies (market-centered liberalism, and human rights, respectively). It also shows how both ultimately failed at addressing the challenges outlined already forty-five years ago.

To tackle these challenges, this Article argues for three

† Research Scholar at Yale Law School, Fellow in Private Law at Yale Law School Center for Private Law and a Resident Fellow at the Information Society Project at Yale Law School. For invaluable comments and conversations, I thank Daniel Markovits, Jack Balkin, Filipe Brito Bastos, Julie Cohen, Nikolas Guggenberger, Johan Fredrikzon, Claudia Haupt, Agnieszka Jabłonowska, Thomas Kadri, Bonnie Kaplan, Kate Klonick, Hans Micklitz, Giovanni Sartor, Rory van Loo and participants in the YLS ISP Writers' Workshops in January and July 2019, and the 8th Annual Conference of the Younger Comparativists Committee (YCC) at McGill University Faculty of Law in May 2019.

normative shifts. First, it proposes to go beyond "privacy" and towards "social costs of data management" as the framework for conceptualizing and mitigating negative effects of corporations' data usage. Second, it argues to go beyond the individual interests, to account for collective ones, and to replace contracts with regulation as the means of creating norms governing data management. Third, it argues that the nature of the decisions about these norms is political, and so political means, in place of technocratic solutions, need to be employed.

TABLE OF CONTENTS

INTRODUCTION

In the data analytics society,[1] a person who allows a company to collect data about herself enables that company to indirectly learn things about other people.[2] Some of this inferred knowledge might pertain to private matters, which the other prefers to keep unknown. However, even knowledge inferred about non-private matters, which people might freely disclose, can be costly when captured in the vast databases of online companies. Non-private information can be used to engage in fine-tuned price discrimination,[3] to harness behavioral effects like preference manipulation[4] or addiction,[5] or to sustain social segmentation, subordination, and even exclusion.[6] These costs, conceptually distinct from the problem of "privacy,"[7] are unaccounted for by contemporary privacy law, in particular the "notice and choice" model governing consumer data collection and usage.[8] When I allow others to collect data about me, I

---

1. By "data analytics society," I mean the socio-technological reality where a significant number of individuals' daily activities is mediated by technology, recorded in tech companies' databases, and analyzed in order to refine the environment, including via personalized communications. *See infra* Section II.A.

2. *See infra* Section II.A.

3. *See* Oren Bar-Gill, *Algorithmic Price Discrimination: When Demand Is a Function of Both Preferences and (Mis)perceptions*, 86 U. CHI. L. REV. 217, 254 (2019).

4. *See* Eliza Mik, *The Erosion of Autonomy in Online Consumer Transactions*, 8 L., INNOVATION & TECH. 1, 1 (2016); *see also infra* Section II.B.

5. *See* Jack Balkin, *Fixing Social Media's Grand Bargain* 4 (Hoover Working Grp. on Nat'l Sec., Tech., & Law, Aegis Series Paper No. 1814, 2018) ("The more digital companies know about people's emotional vulnerabilities and predispositions, the more easily they can structure individual end-user experience to addict end users to the site"); *see also infra* Section II.B.

6. *See* Latanya Sweeney, *Discrimination in Online Ad Delivery*, QUEUE, Mar. 2013, at 10-16, https://queue.acm.org/detail.cfm?id=2460278; *see also* Muhammad Ali et al., *Discrimination through Optimization: How Facebook's Ad Delivery Can Lead to Skewed Outcomes*, PROG. ACM HUM.-COMPUTER INTERACTION, Nov. 2019, at 2–4, https://dl.acm.org/doi/abs/10.1145/3359301.

7. *See infra* Section II.C.

8. *See* DANIEL J. SOLOVE & PAUL M. SCHWARTZ, INFORMATION PRIVACY LAW

impose various costs on you and our fellow humans. My consent, even well-informed, cannot justify the costs imposed on others.

This state of affairs is caused both by the changes in socio-technological practice and by the law. Widespread usage of internet-connected smart devices allows service providers to continuously collect data about millions of people.[9] Data analytics techniques, like machine learning, enable companies to detect patterns in the enormous databases and to infer new knowledge from them.[10] People's reliance on online communications makes it possible for companies to act upon that knowledge through automated, personalized communications.[11] As a result, the music you listen to on Spotify can reveal your race,[12] and influence what job advertisements you will receive through Google Ads.[13] The type of food you order on Uber Eats can reveal your political convictions[14] and be used to encourage or discourage you from voting through the content you see on Facebook.[15] Whose posts you like on Twitter can help reveal how rich you are[16] and enable online sellers to price-discriminate against

785–918 (2018); Daniel J. Solove & Woodrow Hartzog, *The FTC and the New Common Law of Privacy*, 114 COLUM. L. REV. 583 (2014); *see also infra* Section III.A.

9. *See* ETHEM ALPAYDIN, MACHINE LEARNING: THE NEW AI 9–10 (2016).

10. *Id*. at 10–19.

11. *Id*. at 125–39.

12. Shantal R. Marshall & Laura P. Naumann, *What's Your Favorite Music? Music Preferences Cue Racial Identity*, 76 J. RES. PERSONALITY 74, 74 (2018).

13. For the empirical evidence of racial discrimination in online ad-delivery, see Sweeney, *supra* note 6.

14. *See* Daniel DellaPosta et al., *Why Do Liberals Drink Lattes?*, 120 AM. J. SOC. 1473, 1474–75 (2015).

15. *See* Jonathan Zittrain, *Facebook Could Decide an Election Without Anyone Ever Finding Out*, NEW REPUBLIC (June 1, 2014), https://newrepublic.com/article/117878/information-fiduciary-solution-facebook-digital-gerrymandering.

16. *See* Yannick Leo et al., *Socioeconomic Correlations and Stratification in Social-Communication Networks*, 13 J. ROYAL SOC'Y INTERFACE 1, 1 (2016).

you.[17] All these things are possible not because you directly disclose information about your race, political convictions, and wealth. Rather, the bits and pieces you disclose, analyzed in the context of data about millions of others, allow the data holders to infer these facts about you.[18]

Law plays a significant, even if background, role in the creation of this world. Technology, and the way people use it, renders the practices described above factually possible; but it is the law that deems them normatively permissible. As of today, companies derive their rights to collect, analyze, and use the personal data from the individual users' "choices" or "consents."[19] In their "privacy policies," online service providers inform us what they plan to do with the data they collect, and we grant them to a right to do so. However, it is by no means obvious that it should be up to the individual to make these decisions or grant these permissions. Not only the consequences of accepting a policy are often unknown to the users,[20] but also the effects of the data-driven actions might affect people who are not even parties to the agreement that accepting the privacy policy creates.

For these reasons, an individual consumer is neither cognitively capable nor normatively competent to "agree" to data collection and usage that will affect not only her but also other individuals and indeed society as a whole. And yet, the currently dominant "notice and choice" paradigm of American consumer privacy law is based precisely on the paradigm of individual consent.[21] With a few exceptions, no

17. Ramsi A. Woodcock, *Big Data, Price Discrimination, and Antitrust*, 68 HASTINGS L.J. 1371, 1386–87 (2017).

18. To be clear, this is not "knowledge" in the philosophical sense, but rather what Ethem Alpaydin calls "a good and useful approximation." ALPAYDIN, *supra* note 9, at 13.

19. *See* Solove & Hartzog, *supra* note 8, at 592.

20. Jonathan A. Obar & Anne Oeldorf-Hirsch, *The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services*, 21 INFO., COMM. & SOC'Y 128, 143 (2018).

21. *See infra* Section III.A.

federal rules setting absolute substantive limits on the types of usage of personal data exist.[22] By refraining from stepping in, the law creates certain markets and incentives to construct business models around obtaining consent to effectively unrestricted data collection, analysis, and usage.[23]

This Article proposes to address the problem of social costs of data gathering, analysis and usage within a new legal frame: "data management law." This proposed reconceptualization aims to create a new legal field and includes not just a call to enact new rules but also to reconsider existing rules within a different paradigm. "Data management law" would encompass both the existing and the future norms stipulating who is entitled to do what with which personal information and under what conditions. For this paradigm shift to occur, the law must reform along three dimensions.

First, the legal community needs to go beyond "privacy" when pondering harms of data analytics, and move towards a more inclusive category of "social costs of data management."[24] Various types of data-driven social costs exist; many of them have little to do with disclosure of personal data. For example, data-driven price discrimination, preference manipulation, or social segmentation can hurt individuals that remain anonymous from the point of view of the company conducting such actions.[25] As a result, the privacy protections applying to

22. *See infra* Section I.C.

23. *See* Julie E. Cohen, *Between Truth and Power*, *in* INFORMATION, FREEDOM AND PROPERTY: THE PHILOSOPHY OF LAW MEETS THE PHILOSOPHY OF TECHNOLOGY (Mireille Hildebrandt & Bibi van den Berg eds., 2016); *see also* JULIE. E. COHEN, BETWEEN TRUTH AND POWER: THE LEGAL CONSTRUCTIONS OF INFORMATIONAL CAPITALISM (2019).

24. For the explanation of the concept of a social cost, see Ronald H. Coase, *The Problem of Social Cost,* 3 J. L. & ECON. 837 (1960).

25. *See infra* Section II.B.1.

"identified or identifiable"[26] persons might not be triggered in situations most harmful to society and individuals. Not everything done with personal data should be governed solely by privacy laws.

Second, we need to go beyond the individual interests and account also for the collective ones. Matters like non-discrimination, the integrity of the political process, or mental health, should not be the domain of individual preferences. They are of interest to the society as a whole. In the data analytics society, they are endangered. Instead of relying on individual consumers to consult the privacy policies, and to decide whether they accept the rules or not, we need public-oriented disclosures ("societal notice") and collectively taken decisions ("societal choices"). This will mean separating the fact-conferring function currently played by corporations' privacy policies from the norm-creating one and supplementing contracts with regulation.[27]

Third, the policy-makers need to abandon the technocratic normative standards and replace them with politics. The rules governing data management should not come from the market,[28] as they currently do in the US, or be derived from general principles like human rights[29] or "fiduciary duties."[30] Rather, they should be decided through public deliberations, based on reliable knowledge (coming from societal notices) and considering various interests at

26. *See* Paul M. Schwartz & Daniel J. Solove, *The PII Problem: Privacy and a New Concept of Personally Identifiable Information*, 86 N.Y.U. L. REV. 1814, 1818–19 (2011).

27. *See infra* Part IV.

28. For the proposal of managing external effects of data analytics using economic solutions, like the "Pigouvian taxes," see Omri Ben-Shahar, *Data Pollution* (Coase-Sandor Working Paper Series in Law & Econ., Working Paper No. 854, 2018).

29. *See infra* Section III.B.2.

30. Jack M. Balkin, *Information Fiduciaries and the First Amendment*, 49 U.C. DAVIS L. REV. 1183, 1233 (2016). *But see* Lina Khan & David Pozen, *A Skeptical View of Information Fiduciaries*, 133 HARV. L. REV. 497, 509–10 (2019).

stake. These decisions, given the enormous variety of uses of personal data, and the heterogeneity of risks imposed by them, should not come in the form of "one-size-fits-all" regulation, as they currently do both in the US and the EU. Rather, they should occur on a sectoral, issue-by-issue basis.[31]

My argument is based on the lessons learned from the critical analysis of the development of the "data management law" in the United States, and its comparison with the alternative approach adopted by the European Union, currently expressed in the EU General Data Protection Regulation.[32] I provide a new history of both legal regimes seen together, tracing the origins of both models to a single document, namely the 1973 Report of the U.S. Health, Education and Welfare Secretary's Advisory Committee on Automated Personal Data Systems titled *Records, Computers, and the Rights of Citizens*.[33] Its influence on the American law is acknowledged,[34] though as I show, often misunderstood. Regarding the European system, this article nuances the self-understanding of European data protection lawyers by demonstrating how the GDPR's regulatory framework can be traced back to a document issued by Americans for the American administration.[35]

---

31. *See infra* Section IV.C.

32. Regulation (EU) 2016/679, of the European Parliament and 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. (L 119) [hereinafter GDPR].

33. U.S. DEP'T OF HEALTH, EDUC. & WELFARE, RECORDS COMPUTERS AND THE RIGHTS OF CITIZENS: REPORT OF THE SECRETARY'S ADVISORY COMMITTEE ON AUTOMATED PERSONAL DATA SYSTEMS (1973) [hereinafter 1973 REPORT].

34. *See* JONATHAN ZITTRAIN, THE FUTURE OF THE INTERNET AND HOW TO STOP IT 201 (2008); SOLOVE & SCHWARTZ, *supra* note 8, at 36; Marc Rotenberg, *Fair Information Practices and the Architecture of Privacy (What Larry Doesn't Get)*, 2001 STAN. TECH. L. REV. 1, 13–17, 15 n.74.

35. *See infra* Section I.C.

This document turns out to have been not only influential but also prescient. I show that, already in the early 1970s, certain problems we tend to treat as new were foreseen. These include the ease of combining data from different sources and using it for unintended purposes, organizations' ability to use this data to influence the behavior of individuals, the possibility of masking policy-decisions as technical ones, and the risk of imposing the costs on the most vulnerable members of the society.[36] The report's authors foresaw the need to closely monitor technological developments, and to take regulatory decisions regarding the limits of automated data analytics.[37] The recommendations they put forward called for legislative intervention, enabling public oversight over the organizations' data practices. The proposed legislation, based on three pillars: general principles, public-oriented disclosure, and data subjects' rights; was supposed to facilitate political deliberations about the acceptable frames for data management.[38] However, neither the American nor the European data management frameworks managed to fulfill the tasks the report set forth. Today, we suffer the consequences of these failures.

I trace how the existing American approach has been developed within the 1990s mindset favoring the markets, self-regulation and individual choice, thereby ignoring the questions of externalities, and opposing regulation. The European model, on the other hand, was forged largely by human rights lawyers, first and foremost concentrating on keeping the state power in check, thereby ignoring the reality of market operations. On the surface, the two regimes seem distinct.[39] The American law is sectorial, market-

36. *See infra* Section I.A.

37. 1973 REPORT, *supra* note 33, at 12–30.

38. *Id.* at 48–71.

39. For the comparisons of both systems, see Lindsey Barrett, *Confiding in Con Men: U.S. Privacy Law, the GDPR, and Information Fiduciaries*, 42 SEATTLE U. L. REV. 1057, 1057 (2019); Paul M. Schwartz & Karl-Nikolaus

oriented and based on self-regulation, while the European is general,[40] rooted in the human rights logic, and enshrined in legislation. However, they both commit the same mistakes: frame the costs of data management solely as the problems of privacy (or personal data "protection"), concentrate only on the individual interests, and employ technocratic means of dealing with these costs (market solutions, and human rights, respectively). Both the American and European law forgot the insights provided already forty-five years ago. This Article serves not only as a reminder, but also updates the recommendations, to make them directly useful for the socio-technological reality of the 2020s.

The Article consists of four parts. Part I analyzes the 1973 Report in detail, paying attention to the types of risks that its authors noticed 45 years ago, as well as the exact recommendations they made. Part II fast-forwards to 2020, and focuses on the features of today's data analytics society, explaining how three phenomena: (i) seamless data collection; (ii) inferred knowledge; and (iii) automated decision making further facilitate practices in which companies impose costs on individuals and the society as a whole. I study three examples of such data-driven, costly behavior: price discrimination, behavior manipulation, and societal segmentation. Part III analyzes how these practices are sanctioned by law. It describes the American "notice and choice" and European "data protection" models of data management law, tracing their historical and ideological origins, and pointing out their shortcomings regarding the management of social costs of data analytics. Part IV draws lessons from these reconstructions and proposes a way of thinking about "data management law" for the 2020s. It argues for: i) moving beyond "privacy" law and towards a more inclusive category of "social costs of data management"; ii) accounting for collective interests in addition to individual

---

Peifer, *Transatlantic Data Privacy Law*, 106 GEO. L.J. 115, 180 (2017).

40. Sometimes called "omnibus." *See* SOLOVE & SCHWARTZ, *supra* note 8, at 1141.

ones; and iii) replacing technocratic means of decision-making with politics. Conclusions follow.

I. THE ORIGINS. OR, HOW THE AMERICANS "INVENTED" THE GDPR IN 1973

On June 25, 1973, Willis H. Ware sent a letter to a Caspar W. Weinberger, U.S. Secretary of Health, Education and Welfare in the Nixon administration. Ware, a pioneer in the fields of computing and computer security,[41] served as a chairman of the Secretary's Advisory Committee on Automated Personal Data Systems. The Committee had been formed to analyze the "harmful consequences that may result from using automated personal data systems."[42] In the early 70s public and private organizations were rapidly adopting computers as a technology allowing for more efficient storage and processing of records about people. In the midst of the Cold War and the Space Race, the citizens voiced numerous fears connected with this phenomenon, and the Government decided to investigate.[43]

The Committee, comprised of twenty-five members, including lawyers and engineers, as well as social workers, managers, state legislators, private citizens, and a labor union official met nine times between April 1972 and March 1973.[44] It heard testimony from more than 100 witnesses, consulted similar groups from Canada, Great Britain, and Sweden, and contacted about 250 trade and professional associations and public interest groups to gather information. The result of its work, which Ware attached to his letter to Weinberger, was a 300-page long report, titled *Records, Computers, and the Rights of Citizens.*

This report would prove immensely consequential in the decades to come. It has influenced, though in different,

---

41. Michael Rich, *Eulogy for Willis Ware*, RAND (Mar. 28, 2014), https://www.rand.org/content/dam/rand/pubs/corporate_pubs/CP700/CP775/RAND_CP775.pdf.

42. 1973 REPORT, *supra* note 33, at ix.

43. *See* ALAN F. WESTIN & MICHAEL A. BAKER, DATABANKS IN A FREE SOCIETY 465-86 (1972); *see also* Solove & Schwartz, *supra* note 8, at 36.

44. 1973 REPORT, *supra* note 33, at 147–59.

unobvious, and today sometimes forgotten ways, both American data management law[45] and its European counterpart.[46] The vocabulary it used, concepts it coined, and axiology it proposed continue to shape the laws of today. However, as I argue in this Article, both the American and the European data management frameworks did not meet the report's challenges. Instead, each regime has been distorted in various ways and failed to follow-through on certain foundational tasks. To understand how and why, let us first take a closer look at this magnificent document.

A. *The Rise of the Machine and the Mindset Forty-Five Years Ago*

The Report's authors stated in the Foreword:

> It is important to be aware, as we embrace this new technology, that the computer, like the automobile, the skyscraper, and the jet airplane, may have some consequences for American society that we would prefer not to have thrust upon us without warning. Not the least of these is the danger that some recordkeeping applications of computers will appear in retrospect to have been oversimplified solutions to complex problems, and that their victims will be some of our most disadvantaged citizens.[47]

Today's debates about fairness of algorithmic practices regarding policing,[48] creditworthiness,[49] teacher's assessments,[50] or hiring[51] render these predictions, made in

---

45. *See infra* Section III.A.1.

46. *See infra* Section III.B.

47. 1973 REPORT, *supra* note 33, at v–vi.

48. *See generally* Elizabeth E. Joh, *The New Surveillance Discretion: Automated Suspicion, Big Data, and Policing*, 10 HARV. L. & POL'Y REV. 15 (2016); Andrew D. Selbst, *Disparate Impact in Big Data Policing*, 52 GA. L. REV. 109 (2017).

49. *See* Mikella Hurley & Julius Adebayo, *Credit Scoring in the Era of Big Data*, 18 YALE J.L. & TECH. 148, 157 (2016).

50. CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY 3–6 (2016).

51. *See* McKenzie Raub, *Bots, Bias and Big Data: Artificial Intelligence, Algorithmic Bias and Disparate Impact Liability in Hiring Practices*, 71 ARK. L.

the early 1970s, sadly correct. Various "algorithmic tools" put to work with the aim of optimizing assessment processes indeed oversimplify the world, and the victims of these oversimplifications are disproportionately women and racial minorities.[52] Moreover, ironically, in retrospect, we see that the legal frameworks developed as responses to the emergence of computing were in themselves simplified solutions to the complex social problems. The authors begin by situating the development of computers in a wider social and historical context, then explain what new capacities result from their introduction, highlight the risks associated with these changes, and gradually move to their own policy proposals. Let us briefly look at each of these.

Three consequences of computers' introduction and widespread adoptions attracted the authors' attention: (i) substantial enlargement of organizations' data processing capacities; (ii) easy access to data within the organization and across firm boundaries; and (iii) emergence of a "new class of record keepers."[53] Each of these changes, inherent in the process of computerization, can be beneficial for an organization and its clients, but also comes with its own "latent effects."[54]

First, increased data processing capacity might lead to "too much data" to process. Data overload could force organizations to try to simplify the world, by creating various categories *ex ante* and forcing individuals to "fit [themselves] into" these categories while filling in the forms and providing information.[55] Second, "easy access" to information within and across the organizations might lead to combining of information from various sources, gathered for different

---

REV. 529 (2018).

52. *See* Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671 (2016); Mark MacCarthy, *Standards of Fairness for Disparate Impact Assessment of Big Data Algorithms*, 48 CUMB. L. REV. 67 (2017).

53. 1973 REPORT, *supra* note 33, at 12.

54. *Id.* at 12–30.

55. *Id.* at 14.

purposes, into single individual dossiers. As a result, "neither the data subject nor the new holder knows what purpose the data may someday serve."[56] Third, entrusting more and more data processing into hands of engineers and system developers might lead to a situation where "questions of record-keeping practice which involve issues of social policy are sometimes treated as if they were nothing more than questions of efficient technique."[57]

These three developments are particularly problematic, in the view of the Report's authors, when the organizations' purpose is not only to predict, but also to "coerce" (or, as we would say nowadays, when such practices became normalized, "influence" or "nudge")[58] some types of behavior. They write:

> [W]ords like "control" and "coercion" may have an objectionable ring, but the coercive potential of the surveillance component, especially in some other area of application, is evident.[59]

The authors of the 1973 report were aware of the fact that, as a society, we face policy trade-offs. They enlist examples of political decisions to be made:

> Should a national credit-card service be prohibited from using a sophisticated personal data system to prevent its card holders from going on irresponsible spending sprees? Should school districts be forbidden to use personal data systems to help prevent children from becoming delinquents? These are difficult questions to answer.[60]

---

56. *Id.* at 21.

57. *Id.* at 23.

58. *See* RICHARD THALER & CASS SUNSTEIN, NUDGE: IMPROVING DECISIONS ABOUT HEALTH, WEALTH, AND HAPPINESS 53–55 (2008). *But see* SHOSHANA ZUBOFF, THE AGE OF SURVEILLANCE CAPITALISM: THE FIGHT FOR A HUMAN FUTURE AT THE NEW FRONTIER OF POWER 293–328 (2019).

59. 1973 REPORT, *supra* note 33, at 26.

60. *Id.* at 27.

They immediately indicate why these decisions will be difficult not only policy-wise (agreeing on values and goals), but also in regard to the knowledge necessary to make them:

> Often the immediate costs of not using systems to take preemptive action against individuals can be estimated (in both dollars and predictable social disruption), while the long-term costs of increasing the capacity of organizations to anticipate, and thus to control, the behavior of individuals can be discussed only speculatively.[61]

The Report's authors knew that automated personal data processing, even if conducted for noble purposes, increases the potential for coercing individual behavior by private and public organizations. They acknowledged that this "coercion" might sometimes be socially desirable, but they claimed that a political decision about whether to allow it or not, might be difficult to make if the processes are secret,[62] policy decisions are framed as technical ones, and data gathered for the originally accepted purpose is combined with other data, and used for some other purpose later on. In other words, they acknowledged that in order to take some policy decisions on a societal level, what is necessary is to establish mechanisms for bringing these questions to the fore, both concerning those tradeoffs we face already, as well as questions concerning the future "long-term costs of increasing the capacity of organizations to anticipate, and thus to control, the behavior of individuals."[63]

Interestingly, the authors directed their attention to the question of "privacy" only after they discussed all these other concerns. Following the claims made in the literature of their times,[64] they argued that the concept of "privacy" must be

---

61. *Id.* at 27–28.

62. *Id.* at 28–29 ("Today it is much easier for computer-based record keeping to affect people than for people to affect computer-based record keeping.").

63. *Id.* at 28.

64. ALAN F. WESTIN, PRIVACY AND FREEDOM (1967); OFFICE OF SCI. & TECH. OF THE EXEC. OFFICE OF THE PRESIDENT, PRIVACY AND BEHAVIORAL RESEARCH (1967); Charles Fried, *Privacy*, 77 YALE L.J. 475, 493 (1968).

refined and expanded. They stated that in the era of automated personal data processing "an individual's personal privacy is directly affected by the kind of disclosure *and use* made of identifiable information about him in a record."[65] The "and use" part is absolutely crucial here. It suggests that the meaning assigned to the term "privacy" was expanded in the Report to other individual interests of persons, qualitatively different from the "disclosure" actions which were traditionally associated with "privacy" in the canonical Warren and Brandeis article,[66] as well as the typical formulations of the privacy torts.[67]

One should notice that this conceptual move was not inevitable. Arguably, an individual's view on her personal information being disclosed to others, and the use to which this information is put, need not be strongly correlated. For example, I might strongly prefer that no one learns about the "guilty pleasure" music I listen to on Spotify but simultaneously welcome algorithmic suggestions of similar songs. Or, conversely, I might be fine with everybody knowing what my religious convictions are, but would strongly prefer that this information was not used in the process of assessing my creditworthiness, during the hiring process, or for directing political communications at me. These two "dimensions" of privacy: disclosure and usage of information about a person, can be treated jointly (as the authors did) or separately, as I argue would be beneficial.[68] Nevertheless, it is important to remember that when making

65. 1973 REPORT, *supra* note 33, at 40–41 (emphasis added).

66. *See* Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193, 204 (1890).

67. *See* William L. Prosser, *Privacy*, 48 CALIF. L. REV. 383, 389 (1960) (enumerating four types of privacy torts: "1. Intrusion upon the plaintiff's seclusion or solitude, or into his private affairs; 2. Public disclosure of embarrassing private facts about the plaintiff; 3. Publicity which places the plaintiff in a false light in the public eye; 4. Appropriation, for the defendant's advantage, of the plaintiff's name or likeness").

68. *See infra* Section IV.A.

recommendations about "privacy," the Report's authors treated these dimensions jointly.

Within this frame, the authors of the 1973 Report expressed their normative view that: "personal privacy, as it relates to personal-data record keeping, must be understood in terms of a concept of mutuality."[69] Since individuals and organizations often have different though legitimate interests, the role of law is not to endow any of them with the unilateral choice of the ways in which data will be used, nor is it to make the choice regardless of the parties' preferences.[70] Rather, the law should create conditions for *a process* through which all the interested parties will be able to voice their concerns.[71]

What were these conditions supposed to be?

B. *Recommendations Made: Legislate to Create Public Deliberations*

The Report divided its recommendations into several parts, but from this Article's perspective two are significant: "Safeguards for Privacy" (Chapter III) and "Recommended Safeguards for Administrative Personal Data Systems" (Chapter IV).[72] The former contained what came to be known as "Fair Information Practice Principles," later picked up and adopted (in a morphed manner) in the U.S. "notice and choice" model.[73] The latter laid the ground for the axiological and conceptual framework which, through the 1980 OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data,[74] the 1981 Council of Europe

69. 1973 REPORT, *supra* note 33, at 40.

70. *Id*. at 43–44.

71. *Id.*

72. These were followed by specific questions concerning statistical and research data analysis, usage of Social Security Numbers as universal identifiers, and other more technical questions, in Chapters V–IX.

73. *See infra* Section III.A.

74. Ministerial Council of the Org. for Econ. Cooperation & Dev.,

Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data,[75] and the 1995 EU Directive on the Protection of Individuals with Regard to the Processing of Personal Data and on the Free Movement of such Data[76] made their way to the GDPR.[77] Importantly, however, the authors of the Report saw both parts as elements of the same response. Since "privacy," in their understanding, encompassed entitlements to participate in deciding about how personal data will be used, both the general principles and the specific frame turning them into concrete practice were designed as parts of the same system.

The general frame of the proposal was based on five "Fair Information Practices Principles":

> 1. There must be no personal data record-keeping systems whose very existence is secret.
>
> 2. There must be a way for an individual to find out what information about him is in a record and how it is used.
>
> 3. There must be a way for an individual to prevent information about him obtained for one purpose from being used or made available for other purposes without his consent.
>
> 4. There must be a way for an individual to correct or amend a record of identifiable information about him.
>
> 5. Any organization creating, maintaining, using, or disseminating records of identifiable personal data must assure the reliability of

---

Recommendation of the Council Concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data, OECD Doc. C(80)58/FINAL (Sept. 23, 1980) [hereinafter 1980 OECD Guidelines], https://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm.

75. Council of Europe, Convention for the Protection of Individuals with Regard to Automatic Processing of Personal Data, Jan. 28, 1981, E.T.S. No. 108 [hereinafter Convention 108], https://www.coe.int/en/web/conventions/full-list/-/conventions/rms/0900001680078b37.

76. 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, O.J. (L 281) 31 [hereinafter 1995 Directive].

77. *See infra* Section I.C.

> the data for their intended use and must take reasonable precautions to prevent misuse of the data.[78]

These principles expressed the general axiology of the Report's authors' argument: transparency, mutuality and the procedural nature of the regulatory frame. To turn these principles into social reality, the authors suggested:

> The enactment of legislation establishing a Code of Fair Information Practice for all Automated personal data systems . . . [providing for] both civil and criminal penalties . . . injunctions . . . right to bring suits for unfair information practices to recover actual, liquidated, and punitive damages, in individual or class actions.[79]

They expressly argued against self-regulatory approaches, pointing out that the companies might not have sufficient incentives to self-regulate.[80] The foreseen legislation, applicable to both private and public organizations, would divide safeguards applicable to every data management system into three parts: (i) "General Requirements"; (ii) "Public Notice Requirement"; and (iii) "Rights of Individual Data Subjects."[81]

"General Requirements," among others, stipulated: prohibition of data transfers to other organizations without the explicit consent of the data subject, and mandatory requirements for the systems' security and personnel's training.[82] The former is extremely important, as this is one of the major divergence points between the US and the EU. In the American model, an acceptance of the privacy policy containing a blank right to transfer to "business partners" renders such transfers lawful, while in the EU each transfer should be separately "legalized," for example through an

78. 1973 REPORT, *supra* note 33, at 41.

79. *Id.* at 50.

80. *Id.* at 52.

81. *Id.* at 53–64.

82. *Id.* at 53–55.

informed consent; and notified every time it occurs.[83] The latter remained present in the data management regimes in the US and the EU, playing a significant role in the GDPR's framework.[84]

The "Public Notice Requirement" stipulated that every organization processing data in automated manner shall make publicly available the information on, among others: "the nature and purpose(s) of the system"; "the categories and number of persons on whom data are (to be) maintained"; "the categories of data (to be) maintained"; "the organization's policies and practices regarding data storage, duration of retention of data, and disposal thereof"; "the categories of data sources; a description of all types of use (to be) made of data."[85] This recommendation would later lead to the idea of "privacy policies," i.e. documents serving as vehicles of "notice" for consumers in the American model; and the transparency requirements in the European model.[86] However, in the view of the authors, the role played by the public notice was supposed to be fact-conferring, not norm-creating. Note that within this view, the "public notice" was not supposed to be a contract. Rather, it was meant to serve a purpose equivalent to the "nutrition facts" table on food products. Companies would be liable for misleading the public, but a person's "agreement" was not necessarily to be equated with granting the company a right to engage in these practices. The norms stipulating what data can be used for were to be enshrined elsewhere, in sectorial legislation.

Finally, every person about whom data is automatically processed was to be endowed, by legislation, with a set of mandatory rights. To realize these rights, every organization processing personal data was supposed to, among others: enable an individual to know whether he or she is the subject

83. *See* GDPR, *supra* note 32, arts. 6, 12, 14.

84. *See id.* arts. 32–34.

85. 1973 REPORT, *supra* note 33, at 57–58.

86. *See* GDPR, *supra* note 32, arts. 12–14.

of data in the system; assure that no use of individually identifiable data is made that is not within the stated purposes of the system; inform an individual, upon his or her request, about the uses made of data about them, and about the identity of all persons and organizations involved; maintain procedures that allow an individual who is the subject of data in the system to contest their accuracy completeness, pertinence, and the necessity for retaining them; and permit data to be corrected or amended when the individual to whom they pertain so requests.[87] "Data subjects rights," though absent in the "notice and choice" model of American data management law, became one of the cornerstones of the European model. The (in)famous "right to be forgotten"[88] can be directly traced back to the 1973 idea that a person should have a right to "contest necessity for retaining a piece of data." Arguably, this part became the source of the "individualist bias" in the further development of the American and European regimes.[89]

Importantly, all three elements: general rules, public notice, and individual rights were intended to give substance to the Fair Information Practice Principles. In the Report authors' view, such a legal frame was most likely to create the conditions for public deliberations on the limits of personal data processing. The intended effects of the enactment of the Code of Fair Information Practice were described as follows:

> The proposed safeguards are intended to assure that decisions about collecting, recording, storing, disseminating, and using identifiable personal data will be made with full consciousness and

87. 1973 REPORT, *supra* note 33, at 59–63.

88. For a legal-theoretical discussion of the right to be forgotten, see Giovanni Sartor, *The Right to Be Forgotten: Balancing Interests in the Flux of Time*, 24 INT'L J.L. & INFO. TECH. 72 (2016). For a policy analysis from the domestic perspective, see Michael J. Kelly & David Satola, *The Right to Be Forgotten*, 2017 U. ILL. L. REV. 1 (2017); Daniel Lyons, *Assessing the Right to Be Forgotten*, 59 BOS. BAR J. 26, 28 (2015); Jeffrey Rosen, *The Right to Be Forgotten*, 64 STAN. L. REV. ONLINE 88, 92 (2012).

89. *See infra* Section IV.B.

> consideration of issues of personal privacy—issues that arise from inherent conflicts and contradictions in values and interests. Our recommended safeguards cannot assure resolution of those conflicts to the satisfaction of all individuals and groups involved. However, they can assure that those conflicts will be fully recognized and that the decision-making processes in both the private and public sectors, which lead to assigning higher priority to one interest than to another, will be open, informed, and fair.[90]

In other words, the authors saw the role of the proposed legislation not as the ultimate answer to the question of what practices should be allowed and what not. They understood that such legislation is necessary to create a forum where, in the societal and political processes, further decisions about particular uses will be made. They were also aware that no one-size-fits-all solution is feasible.[91] Even though they do not use the word, the authors were aware of the possibility of externalities, and the need to regulate in order to mitigate them:

> The past two decades have given America intensive lessons in the difficulty of trying to check or compensate for undesirable side-effects stemming from headlong application and exploitation of complex technologies. Water pollution, air pollution, the annual highway death toll, suburban sprawl, and urban decay are all unanticipated consequences of the too narrowly conceived and largely unconstrained applications of technology.[92]

All this taken together: the plea for public disclosure, for deliberation and for balancing interests—uttered by people who, in principle, saw development of computers as a positive phenomenon—signals the authors' hope that the enactment of the Code of Fair Information Practice by the U.S. Congress would have been just the first step. They did not explicitly outline how they hoped the further deliberations would look like. Nevertheless, we know that the problems they foresaw included the rise of uncontrolled power to coerce behavior, the emergence of engineers being

90. 1973 REPORT, *supra* note 33, at 43–44.

91. *Id.* at 43.

92. *Id*. at 45.

de facto policymakers, and negative effects affecting the most vulnerable people. To remedy them, they wanted to install procedural guarantees providing for societal control and, potentially, further regulation of particular sectors.

What followed was a much more limited intervention.

C. *What Followed: The Transatlantic Split and the Technocratic Turn*

The most direct and immediate effect of the Report's publication was the enactment of the 1974 Privacy Act.[93] This legislation closely mirrored the recommendations of the Report. Several substantive rules were put in place, including the purpose limitation principle, and the citizens were granted various rights, among others to access and correct their personal data.[94] Nevertheless, the Privacy Act applied (and applies) only to the Federal Agencies.[95] General in its material scope (as a matter of a rule, it governs all the personal data gathered by the federal agencies about citizens for all the purposes), Privacy Act remained limited in its subjective scope of application, governing just one part of only the public sector.

Several other statutes followed. Importantly, the approach taken by the legislature has from the beginning been sectoral. The legislative acts adopted by Congress included: Family Educational Rights and Privacy Act of 1974,[96] Cable Communications Policy Act of 1984,[97] Video

93. Privacy Act of 1974, Pub. L. No. 93-579 (codified at 5 U.S.C. § 552a (2012)).

94. *See* ARTHUR A. BUSHKIN & SAMUEL I. SCHAEN, THE PRIVACY ACT OF 1974: A REFERENCE MANUAL FOR COMPLIANCE 7 (1976).

95. *Id.* at 11–13.

96. Family Educational Rights and Privacy Act of 1974, Pub. L. No. 93-380 (codified at 20 U.S.C. § 1232g (2012)).

97. Cable Communications Policy Act of 1984, Pub. L. No. 98-549 (codified at 47 U.S.C. § 551 (2012)).

Privacy Protection Act of 1988,[98] Telephone Consumer Protection Act of 1991[99] and many others.[100]

The sectoral, "issue by issue" approach remains the defining feature of the American data management law. In itself, this approach has been faithful to the 1973 Report, in which the authors argued against one, top-down solution to all possible negative effects of data processing.[101] However, the condition for making sure that the "issue by issue" approach would be, in the authors' view, "open, informed, and fair"[102] has never been fully realized. Within this legal framework, defined by the sectoral approach and lack of general data transparency, the "notice and choice" paradigm would be born in the 1990s, influenced however by different ideologies and tacit normativities. The Report would be cited and referred to, though as we will see, the ultimate shape of the end product would look little like the original idea. The addressee of the disclosure would change from the public to the individual, requirements about its concreteness would be watered down, and the free market would be treated as choice making mechanism superior to politics and regulation. As a result, since the late 1990s, essentially no new federal legislation governing data collection and usage has been enacted. New types of costs created by the business models of Google, Facebook and Amazon were not followed by new sectoral legislative interventions. The Report called for collective decision-making mechanisms, in essence political. The law of today leaves the decisions to the individuals and is in essence market-based.

---

98. Video Privacy Protection Act of 1988, Pub. L. No. 100-618 (codified at 18 U.S.C. §§ 2710–2711 (2012)).

99. Telephone Consumer Protection Act of 1991, Pub. L. No. 102-243 (codified at 47 U.S.C. § 227 (2012)).

100. For the list of the acts adopted as a response to the emergence of internet, see *infra* p. 53. For the list of all privacy acts adopted in the United States, see SOLOVE & SCHWARTZ, *supra* note 8, at 36–39.

101. 1973 REPORT, *supra* note 33, at 42–43.

102. *Id.* at 44.

In Europe, the story took a different path. In 1980 the OECD Guidelines[103] were published, and in 1981 the Council of Europe's Convention 108 would follow.[104] This Convention would closely mirror both the vocabulary of the Report ("personal data," "data subject," "processing") and its general axiology (public notice, purpose limitation principle, need for regulation). However, it would do so as a part of its mandate to protect the human right to privacy.[105] It would oblige the Member States to enact legislation on personal data "protection"; a general law applicable to both private and public sectors. National laws enacted to fulfill these obligations would be further harmonized in 1995 via the European Union Directive[106] and unified and modernized in 2016 with the GDPR.[107] In the meantime, the European Union would declare "personal data protection" a fundamental right, essentially equivalent to the position enjoyed in the US by the Bill of Rights.[108]

The European regimes would closely follow the recommendation of the 1973 report, by operating within the frame of (i) general principles; (ii) public notice; (iii) subjects rights.[109] However, since the body through which the legislation would be adopted would be the Council of Europe, an international organization created to protect human rights, the reception of the idea would be conducted within the "privacy is a human right" ideology. This, taken together with the "direct effect doctrine"[110] would largely influence the

103. 1980 OECD Guidelines, *supra* note 74.

104. Convention 108, *supra* note 75.

105. *See infra* Section III.B.1.

106. 1995 Directive, *supra* note 76.

107. GDPR, *supra* note 32.

108. *See generally* GLORIA GONZALEZ FUSTER, THE EMERGENCE OF PERSONAL DATA PROTECTION AS A FUNDAMENTAL RIGHT OF THE EU (2014).

109. *See infra* Section III.B.1.

110. Broadly speaking, in the European Union law, unlike in the United States, constitutional rights are directly applicable in the private relations between the market players. *See* ELENI FRANTIZIOU, THE HORIZONTAL EFFECT OF

shape of the final product. The GDPR, direct descendant of the 95/46 directive and the 108 Convention, would impose much more strict general rules than the American system. However, it would also focus primarily on the individual, rather than collective interests, apply only to "identifiable" persons, and employ technocratic solutions rather than political ones.

To be clear, proving that the 1973 Report directly influenced the 1981 Convention 108 is a difficult task, and I would not defend a strong claim that without the Report, the European system would necessarily look differently. Other European countries, like Great Britain or Sweden, have been adopting their own data protection laws at the time, and we know that the authors of the 1973 Report were in contact with them.[111] One can only assume that the exchange of ideas happened in both directions. However, two things can be demonstrated beyond doubt. First, the final product of the Council of Europe, and ultimately the GDPR, follows the structure and terminology of the 1973 Report's recommendations. Second, when Council of Europe commissioned its own recommendations in 1973[112] and 1974,[113] they used different terminology and ultimately adopted that of the 1973 Report (they spoke of "personal information" instead of "personal data," "individuals" instead of "data subjects" etc.). However, from the policy perspective, it does not matter whether the drafters of the European regime read the 1973 Report and decided to implement it, or whether these ideas the Report enshrined made their way to

---

FUNDAMENTAL RIGHTS IN THE EUROPEAN UNION: A CONSTITUTIONAL ANALYSIS (2019).

111. 1973 REPORT, *supra* note 33, at x.

112. Council of Eur., Comm. of Ministers, Res. (74) 29 On the Protection of the Privacy of Individuals vis-à-vis Electronic Data Banks in the Private Sector (Sept. 26, 1973), https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680502830.

113. Council of Eur., Comm. of Ministers, Res. (74) 29 On the Protection of the Privacy of Individuals vis-à-vis Electronic Data Banks in the Public Sector (Sept. 20, 1974), https://rm.coe.int/16804d1c51.

Europe through the conversations among experts indirectly through the 1980 OECD guidelines. What matters is that already in 1973 the American administration was in possession of the report suggesting something very similar to today's GDPR and that both the American law and the GDPR failed to properly respond to challenges that this Report set forth.

Before discussing the current shape of both regimes, the ways in which they lead to the creation of the data analytics society, and the ways in which they fail at mitigating the social costs, let us first see what the last forty-five years of socio-technological changes brought about. To fully appreciate the mistakes which were committed while forging the "notice and choice" and "personal data protection" models, one needs to better understand the socio-technological reality at the dawn of the 2020s. What goes without saying, a lot has changed since 1973.

II. SOCIAL COSTS OF DATA MANAGEMENT

When identifying the threats of using computers to process personal data, the authors of the 1973 Report were engaging in speculation, but it turns out that they were prescient. Forty-five years later, we have empirical studies documenting what the negative effects of technological data processing have become, and are able to see which of them resulted from not having lived up to the recommendations put forward. In this Part, I analyze the developments in the techno-sociological practice of collection and usage of personal data and study the social costs they generate. These costs encompass both the knowledge that companies are able to infer about people and the various negative effects of using this knowledge in their interactions. I look at three examples: price discrimination, behavior manipulation, and social segmentation; and show how the law we have today sanctions these costly behaviors, as well as why it is unable to prevent them. After this analysis, I will move to the closer study of both models of data governance in Part III.

A. *Technological Foundations of the Data Analytics Society*

Three significant technological breakthroughs have occurred since the publication of the 1973 Report: the miniaturization and widespread distribution of (mobile) computers,[114] the emergence of the public internet,[115] and the rise of machine learning techniques regarding data analytics and problem-solving.[116] As of today, in 2020, it is

---

114. *See* MARTIN CAMPBELL-KELLY & WILLIAN ASPRAY, COMPUTER: A HISTORY OF INFORMATION MACHINE 233–300 (1996); ADAM GREENFIELD, EVERYWARE: THE DAWNING AGE OF UBIQUITOUS COMPUTING (2006); GEORGES IFRAH, THE UNIVERSAL HISTORY OF COMPUTING: FROM THE ABACUS TO THE QUANTUM COMPUTER 302–47 (2001).

115. *See* JANET ABBATE, INVENTING THE INTERNET 195–200 (1999); ZITTRAIN, *supra* note 34, at 26–30.

116. *See* ALPAYDIN, *supra* note 9; Nello Cristianini, *The Road to Artificial Intelligence: A Case of Data Over Theory*, NEW SCIENTIST (Oct. 26, 2016), https://www.newscientist.com/article/mg23230971-200-the-irresistible-rise-of-

common for individuals to own personal computers and mobile devices,[117] constantly connected to the internet, serving as access points to various online services. Some previously analog activities, like shopping, news consumption or banking have gradually moved online; some new forms of social behavior, like social media, "sharing economy"[118] or multimedia streaming have emerged. At the same time, these devices and services function as data collectors for tech companies.

The "data analytics society" we currently live in, I argue, is characterized by three latent effects: (i) *seamless data collection*, facilitated by people's usage of smartphones and other connected devices; (ii) *inferred knowledge*, that is corporations' ability to find patterns in data sets and establish probabilities using machine learning techniques; and (iii) *automated decision making*, where "decision" should be understood broadly, as a computer's choice to generate some output without a direct command by a human, for example, to show a particular ad to a particular person in a given moment.[119] All three phenomena are interrelated, but for analytical purposes let us consider each separately.

Seamless data collection is possible due to individuals' continuous reliance on technology. When a person carries an internet-connected smartphone on them all the time, it is difficult to draw a sharp distinction between being "online" and "offline."[120] All the activities undertaken via a

artificial-intelligence/.

117. According to PEW Research, as of June 2019, 81% of Americans own a smartphone, three quarters own a laptop or a personal computer, and nearly half owns a tablet device. *See Mobile Fact Sheet*, PEW RES. CTR. (June 12, 2019), https://www.pewinternet.org/fact-sheet/mobile/.

118. For the definition of the concept, see ARUN SUNDARARAJAN, THE SHARING ECONOMY: THE END OF EMPLOYMENT AND THE RISE OF CROWD-BASED CAPITALISM 26–27 (2016). For the overview of the legal challenges posed by the sharing economy, see THE SHARING ECONOMY: LEGAL PROBLEMS OF A PERMUTATIONS AND COMBINATIONS SOCIETY (Maria Regina Redinha et al. eds., 2019).

119. *See* Mik *supra* note 4, at 1.

120. Philosophers of technology call this phenomenon the "onlife world." *See*

smartphone leave a digital footprint.[121] From dating to searching for information to shopping, to transport and to communications, more and more data is generated as a side product of our daily activities. If one halts a taxi on a street and pays cash, there is no record. Uber makes a record. When one eats in a restaurant, there is no record. Grubhub makes a record. When one consults a physical encyclopedia, there is no record. Google knows what we are searching for. All this data collection occurs in the background, sometimes without individuals' knowledge, and usually without any human's additional effort to store information on top of conducting the service. The "easy access" that the authors of the 1973 Report worried about has become a reality, both due to the technological developments, and the law's sanctioning of these practices. An individual's "choice" in the American law, and a "consent" in the European model are sufficient legal basis to collect personal data and share it with other organizations.[122]

As a side product, large amounts of data are generated. Ethem Alpaydin calls this the "dataquake",[123] while the term best known to the general public is "big data."[124] Big data is often characterized by the "3 Vs": volume, variety, and

---

Mireille Hildebrandt, *Dualism is Dead. Long Live Plurality (Instead of Duality)*, *in* THE ONLIFE MANIFESTO: BEING HUMAN IN A HYPERCONNECTED ERA 27 (Luciano Floridi ed., 2015).

121. *See* Carrie Leonetti, *Bigfoot: Data Mining, the Digital Footprint, and the Constitutionalization of Inconvenience*, 15 J. HIGH TECH. L. 260, 300 (2014); Agnieszka A. McPeak, *The Facebook Digital Footprint: Paving Fair and Consistent Pathways to Civil Discovery of Social Media Data*, 48 WAKE FOREST L. REV. 887, 893–94 (2013); Sandi S. Varnado, *Your Digital Footprint Left behind at Death: An Illustration of Technology Leaving the Law Behind*, 74 LA. L. REV. 719, 721 (2014).

122. *See infra* Part III.

123. *See* ALPAYDIN, *supra* note 9, at 10–16.

124. For the definition of the concept, see Thomas Hoeren, *Big Data and Data Quality*, in BIG DATA IN CONTEXT: LEGAL, SOCIAL AND TECHNOLOGICAL INSIGHTS (Thomas Hoeren & Barbara Kolany Raiser eds., 2018); *see also* Talia B. Gillis & Jann L. Spiess, *Big Data and Discrimination*, 86 U. CHI. L. REV. 459 (2019).

velocity.[125] From the perspective of a human mind, there is too much data to process, the sets are too heterogeneous to make sense of, and the speed of analysis necessary to process them surpasses our abilities. Recall that "too much data to process" has also been identified as one of the latent effects of computerization by the authors of the 1973 Report. What they feared would happen was the organizations "simplifying the world," distorting the picture of reality by forcing people to fit themselves into the pre-defined categories. In some spheres, this did indeed occur.[126] However, there was another possible strategy for organizations to adopt. Instead of simplifying the data, one could try to increase the performance of data analyzing technologies. And this is what happened. Data analytics, and in particular machine learning, underwent significant progress in the last decade, turning the "too much data to process" burden into a data blessing.[127]

Inferred knowledge results from organizations' ability to apply sophisticated data analytics techniques to detect patterns in the big data collections.[128] Data now reveals more information about individuals than what one could directly observe. Studied in isolation, the fact that I am of a given gender and age, live in a specific town, like chicken waffles and Marvel movies, might say not much more than simply that. However, this data, seen against data about millions of other people who in some regard are "(not) like me," and additional data about them, allows companies to *infer* what are my political views,[129] what products I might be willing to buy, and what social cause to donate to. This is not "knowledge" in the sense in which philosophers would define

125. Andrea De Mauro et al., A Formal Definition of Big Data Based on its Essential Features, 65 LIBR. REV. 122, 130–31 (2016).

126. *See* O'NEIL, *supra* note 50, at 3–9, 23–27, 71–76.

127. *See* ALPAYDIN, *supra* note 9, at 15.

128. *Id.* at 10–20.

129. *See* DellaPosta, *supra* note 14, at 1473.

it, but is what Ethem Alpaydin calls "a good and useful approximation."[130] If a company knows something about me with an 87% probability, this "knowledge" presents significant value.

This knowledge can be further refined, through automatic communication and feedback loops. Have you ever seen an irrelevant ad and thought "how would they ever think that I can be into that?!" One of the reasons why the algorithm displayed it to you might have been to confirm the probability that people within your demographics are not interested in such things. Various technological advancements, in particular machine learning, are currently being employed to not only find patterns in existing data sets but also to refine these data sets, test hypotheses and fill in the loopholes.[131] This "knowledge" can later be used in an automated way. To stay with the example of advertisements, a task that a human gives to a machine could be "display this ad to 1000 people with the highest probability of clicking." The machine would estimate, based on the existing data, who has the highest chance of clicking (or even start at random) and then through numerous feedback loops improve the estimation over time.[132]

Apart from advertising, this technology can power "decisions" on what flight price to display,[133] whose profile to show on dating apps, or what music to suggest. Automated "decision" making pertains not only to "decisions" in the strict sense (grant or refuse a credit card, approve a refund or invite for a job interview); but essentially all activities of computer-systems that are not 100% pre-programmed. Given the machine-learning basis of these systems, humans

130. *See* ALPAYDIN, *supra* note 9, at 14.

131. *Id.* at 113–68.

132. *Id.*

133. *See* Jeffrey K. MacKie-Mason & Michael P. Wellman, *Automated Markets and Trading Agents*, *in* 2 HANDBOOK OF COMPUTATIONAL ECONOMICS (Leigh Tesfatsion & Kenneth L. Judd eds., 2006).

developing and deploying them will have a general idea on what the software will do (the goals are specified manually); but might not be able to fully predict all possible outcomes (resulting from the available data).[134] Importantly, these "automated decisions" are, at the same time, a tool to provide services *and* to collect new/feedback data. In this sense, when a "smart" algorithm performs a task, it simultaneously tries to achieve the best possible outcome and generates feedback on its own performance, so that the next time, it will be more efficient. It is these characteristics taken together that led to a widespread popular referral to machine-learning-based technologies as "artificial intelligence."[135]

Note two things. First, for all this to function it is by no means necessary that you are "personally identified" or "identifiable."[136] In many cases, companies engaging in commercial activities based on data collection, analysis and automated communication really do not need to know what is your name, Social Security Number, address or any other

134. The humans' ability to understand the logic of the big-data fueled and machine-learning powered systems is being discussed under the label of the "black box" problem. *See* FRANK PASQUALE, THE BLACK BOX SOCIETY: THE SECRET ALGORITHMS THAT CONTROL MONEY AND INFORMATION 6–7 (2015). *But see* Agnieszka Jablonowska & Przemyslaw Palka, *EU Consumer Law and Artificial Intelligence*, *in* THE TRANSFORMATION OF ECONOMIC LAW: ESSAYS IN HONOUR OF HANS-W. MICKLITZ (Lucila de Almeida et al. eds., 2019).

135. Machine learning is a subset of artificial intelligence. *See* ALPAYDIN, *supra* note 9; Cristianini, *supra* note 116. Several scholarly and policy initiatives aimed at identifying legal challenges associated with the emergence of artificial intelligence have been launched in recent years, including Harvard's and MIT's "Ethics and Governance of Artificial Intelligence," overview available at: https://www.media.mit.edu/groups/ethics-and-governance/overview/ (last visited Mar. 17, 2020), and Yale's "AI, Ethics and Society," overview available at https://aiethicsyale.wordpress.com/. Nevertheless, one should remember that as long as discussions about normative implications of artificial intelligence are concerned with the real-world developments, and not thought experiments, they pertain to the development and usage of machine learning. *See* MACIEJ KUZIEMSKI & PRZEMYSLAW PAŁKA, EUR. U. INST. SCH. OF TRANSNAT'L GOVERNANCE, AI GOVERNANCE POST-GDPR: LESSONS LEARNED AND THE ROAD AHEAD 1 (2019) ("[G]overnance of AI will in practice mean policies regarding both the design and access to algorithms, as well as collection and usage of information.").

136. *See* Schwartz & Solove, *supra* note 26, at 1818.

personal characteristic. What they need to know is what features you have and how you behave in response to what stimuli. There needs to be an ability to track you somehow (for example, by assigning a number to your browser or your device), but from a corporation's point of view, it is often perfectly fine if you are anonymous. Legally speaking, many companies might *prefer* not to personally identify you—the less they know *who* you are, the more they are legally allowed to learn about *how* you are.[137] That is one of the reasons why "privacy" approaches based on the concept of "personally identifiable information" do not fully match the potential risks in the data analytics society. It might be the case that in situations where individual and societal interests are most at stake, privacy protections simply are not triggered.

Second, notice the externalities. By allowing a company to collect and study information about myself, even if we assume I am fully aware of what they plan to do with my data, I impose a cost on (or at least make a decision about) you and other people. If I make it clear what my political views or religious convictions are, essentially everything I do online can later be used to infer knowledge about other people's religion and politics. "Everything you say can and will be used against other people" would be a fair statement to include in the privacy policies written within the "notice and choice" paradigm. The more things I buy on Amazon, the more refined suggestions to other customers will be. If I am convinced, or not convinced, by a political ad, I help ensure that its next reiteration will be even more successful in manipulating people's preferences and behavior. Even if I fully agree to the collection of my data, I impose costs on you and our fellow humans. I should not be allowed to do this so easily.

These two aspects, the possibility to harm people while they are anonymous, and the risk of being harmed as a result

137. *Id.*

of other people's actions, are significant weaknesses of the current approaches to data management. In the following section, I study three examples of these harms, i.e. social costs of data management, in detail. Having done so, in Part III, I will demonstrate how this state of affairs is sanctioned by the existing data management laws, and pinpoint where I believe they should be modified.

B. *Direct and Indirect Social Costs: Knowledge In-Use and Knowledge In-Itself*

Two general types of social costs exist in the data analytics society: new knowledge about persons, and the potentially negative consequences of acting upon that knowledge. On the one hand, there are external effects of individual behavior concerned with knowledge: when disclosing some personal data, I indirectly reveal information about other people. This does not yet bring about actual harms to others, but increases the potential for harm. On the other hand, there are costs of the uses to which companies can put this knowledge. Data-driven price discrimination, behavior manipulation or social segmentation are actions costly to some people, and the society as a whole, resulting from the technological capabilities of tech companies and the law's sanctioning of these practices.

The distinction between "knowledge in-itself" and "in-use," or between potential and actual harms, matters not only conceptually, but also policy-wise. Depending on one's normative theory, one could argue that the mere fact that Amazon, Google or Facebook know so much about their users is problematic and should be prevented; or emphasize that we should prevent only certain activities based on this data. Further, as I will show in Part IV, both collection and usage can be objects of regulation, and the choice will depend on one's regulatory objectives. With these distinctions in mind, let us consider three examples of costly external effects: price discrimination, behavior manipulation, and social segmentation.

1. Examples: Price Discrimination, Behavior Manipulation and Social Segmentation

First, consider price discrimination.[138] Put simply, price discrimination means charging different customers different prices for exactly the same good or service. Imagine two people, Amy and Barbara, none of whom directly reveals what their income is. Amy is friends on Facebook with colleagues boasting about their high-paid jobs, watches videos about golf, and searches with Google for piano classes for children.[139] Assume that this suggests, with a high probability, that her income is above $100,000. Barbara, on the other hand, engages in online activities associated with a low-income social group. Now consider a company, let us call it Umbrella Corporation, selling widgets. Imagine that a total cost of producing and marketing a widget is $10, so any price above $10 generates profit. Assume that in a brick-and-mortar retail store, where the same price must be used regardless of who is the buyer, Umbrella would sell widgets at the supply and demand equilibrium price of $15. Online, however, Umbrella has access to information about Amy and Barbara and can try to display different prices to different customers. Based on the knowledge inferred about Amy's and Barbara's income, it charges the former $20, while the latter $11. This is costly for Amy, who ends up paying $5 more than she normally would, but beneficial for Barbara and for the Umbrella Corporation.

Second, think about behavior manipulation.[140] Imagine that Umbrella Corporation engages in online advertising,

---

138. *See* Bar-Gill, *supra* note 3, at 271; Woodcock, *supra* note 17, at 1372–74; *see also* Julie Cohen, *Irrational Privacy?*, 10 J. ON TELECOMM. & HIGH TECH. L 241, 245 (2012).

139. For empirical evidence that such behavior is a good predictor of higher income, see ELIZABETH CURRID-HALKETT, THE SUM OF SMALL THINGS: A THEORY OF THE ASPIRATIONAL CLASS (2017).

140. *See* ZUBOFF, *supra* note 58; Mik, *supra* note 4; Rory Van Loo, *Helping Buyers Beware: The Need For Supervision of Big Retail*, 163 U. PA. L. REV. 1311, 1331–43 (2015).

and wants to increase the chance of consumers purchasing widgets while browsing the net on their smartphones. What makes e-commerce different from brick and mortar shops is the option of immediate purchases, clicking on an ad and having concluded a contract three minutes later. Umbrella collects data about past engagement, and its algorithm establishes that people with characteristics similar to Amy are most prone to buy the widget when seeing the ad early in the morning, especially if they went to sleep later than usual. Amy-types also like the widget being advertised by fit, young people. At the same time, people with characteristics similar to Barbara, most often buy widgets when shown the ad late at night, and especially when coming back from their friends'. They are most convinced by ads that have widgets advertised by people who look like a happy family. With this knowledge, Umbrella fine-tunes the timing, content, and form of the widget ads, and manages to increase the sales by 20%. This is profitable to Umbrella but can be costly to Amy and Barbara. If for most of the time, they would not be willing to buy a widget, but do so when shown an ad at the moment when they are most prone to engage in unnecessary spending, the actual cost they pay is higher than the overall utility they derive from owning a widget. This, again, is possible not because they have revealed anything about their preferences or behavior patterns, but because other people, with similar characteristics, have done so.

Third, there is a risk of social segmentation, discrimination, or even social exclusion.[141] Imagine that Umbrella Corporation wants its products to be associated only with certain social classes. It will attempt to display the widget ads only to those who seem to belong to these classes. In the world of billboards, TV ads and brick and mortar stores, most members of the society can see ads of the same products, and most of them have the ability to buy those

141. *See* Barocas & Selbst, *supra* note 52; Gillis & Spiees, *supra* note 124; Sweeney, *supra* note 6.

products in physical stores. However, in a data analytics society, when everyone sees mostly their own set of ads, displayed on their own smartphones, social segmentation can occur. Only people like Amy are offered to buy widgets, while people like Barbara are not. Or imagine that Umbrella Corporation wants to hire new workers, and for some reason ends up displaying the job-ads only to people like Amy. Or, in this case, probably people like Andy.[142] This is costly for Barbara-types, who not only are not offered to buy a widget or work for Umbrella, but might even not be aware of these products' and jobs' existence. This is also costly for society as well. If different members of a community are exposed to different types of content, the social division, resulting from exposure to different communications, might increase.

2. The Politics of Data-Driven Social Costs

One could argue that all the costs described above are neither new nor necessarily bad. In the end, one could say, sellers always wanted to maximize profits by charging the highest possible price and engaging in the most convincing advertising. Consumers and workers have always bought different products, and worked different kinds of jobs, depending on their material situation. This is how capitalism works, and even though the system is not perfect, and some inequalities will necessarily persist, the overall outcome is the best possible one.

These two objections must be addressed together.

I argue that the question of whether data-driven social costs are new or not is a wrong question to ask. No one can deny that the above-mentioned practices existed, in some form, long before the emergence of computers, not to mention the internet and machine learning. At the same time, it is difficult to refute a claim that the amount of knowledge that

142. Umbrella Corporation's employment ad might target a man, given the widespread gender bias in online job ads delivery. *See* Barocas & Selbst, *supra* note 52, at 681–87.

corporations have accumulated about individuals, paired with their ability to directly act upon that knowledge, is unprecedented. The source of this knowledge has shifted as well, from individuals freely disclosing information about themselves, to having knowledge inferred from the behavior of others who, themselves, might not realize what data is being collected. Whether these shifts make up a qualitative or only quantitative change is, ultimately, of secondary importance. What matters is that they have occurred, and normative questions can be asked about them.

This paper makes a plea to bring these questions to the fore. Whether the practices discussed above should be assessed negatively and mitigated through regulation, or not, is ultimately a political decision to make. Let us get back to the examples for a while.

Imagine that Umbrella Corporation shows higher prices to people whom it believes are wealthier and lower to those whom it believes possess fewer funds. One could argue that, given the inequalities in society, it is fair if richer people subsidize purchases by those less fortunate. Or, one could argue that these practices should be discouraged, as they diminish the overall efficiency and lead to a quasi-monopolistic behavior. However, imagine that Umbrella does the opposite. Since it wishes to have widgets associated only with some social classes, it ends up showing a higher price to Barbara, and lower to Amy. Does this change our normative assessment? Does the type of product they offer matter? Would it make a difference if Umbrella was selling pharmaceutical products, or energy, or books?

Further, consider a world in which Umbrella does not sell widgets, but leads an election campaign of a certain political party. Instead of encouraging people to buy its products, it uses all the data it manages to find to discourage people from voting for the other party.[143] Or, it manages to

143. As was the case with the Cambridge Analytics scandal. *See* Nicholas Confessore, *Cambridge Analytica and Facebook: The Scandal and the Fallout So*

modify people's news feeds in a way that leads to a social segmentation of the content they ultimately receive.[144] Would our assessment of its activities change if, instead of marketing products, Umbrella Corporation engaged in data-driven behavior manipulation and social segmentation aimed at influencing an outcome of a political process?

The answer is yes, such details matter. Possibly, certain outcomes of data-driven behavior invite more regulatory scrutiny than others. However, to know that, we need to understand the data-driven practices of tech companies, as the authors of the 1973 Report argued. And we need to engage in deliberations about these practices, in political, issue-by-issue manner. Unfortunately, as of today, the answer regarding the acceptable uses of data about persons is given by individualistic, technocratic, "one-size-fits-all" solutions offered by "notice and choice" and "personal data protection" regimes.[145]

In the following Part, I show how the currently binding legal regimes were created, to help legal reformers understand why the law ended up sanctioning these types of costly practices, and in what aspects these regimes are ill-suited to manage the costs of data analytics.

Far, N.Y. TIMES (Apr. 4, 2018), https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html.

144. The Wall Street Journal's "Blue Feed, Red Feed" Project provides insights on this front. *See Blue Feed, Red Feed*, WALL ST. J., http://graphics.wsj.com/blue-feed-red-feed (last updated Aug. 19, 2019).

145. *See infra* Part III.

III. THE EXISTING PARADIGMS: "NOTICE AND CHOICE" VS. "PERSONAL DATA PROTECTION"

As of 2020, two competing visions of data management law exist in the transatlantic sphere: the American "notice and choice" model and the European "personal data protection" model.[146] Both can be traced back to the 1973 Report and the recommendations it made. In many aspects, they look very much unlike one another. However, even though the American system is grounded in market logic, while the European system is in the logic of human rights, both fall prey to the same three mistakes. Both look at the problem of social costs of data management through the lens of "privacy," both concentrate on individual interests instead of collective ones, and both employ technocratic means to address these costs.

A. *American Model: Individual "Notice and Choice"*

Within the "notice and choice" paradigm, companies that want to collect and use personal information about the users of their websites and apps should inform the customers about their data practices ("notice"). In turn, consumers decide whether they are willing to use the services and disclose personal information to companies or not ("choice").[147] "Notice" usually occurs through a publication of a "privacy policy" (sometimes called a "privacy notice"),[148] while the choice might be either expressed explicitly when ticking a box next to a statement like "I have read and accept the privacy policy," or implicitly through accessing the service or using the app.[149] In this sense, "notice and choice" posits the

146. *See supra* note 39 and accompanying text.

147. *See* Solove & Hartzog, *supra* note 8, at 585–90.

148. *Id.*

149. Various privacy policies contain a "consent by use" clause. *See, e.g.*, *Under Armour Privacy Policy*, UNDER ARMOUR (May 20, 2018), https://account.underarmour.com/en-us/privacy ("By using the services, you consent to the collection, use and transfer of your personal data for processing in the united

decision about acceptable data practices as a market transaction between a user and the company.

This market logic has profound consequences for the shape of the normative framework governing data practices, and ultimately for the nature of the data analytics society. The answer to a legal question a company might ponder: "what data am I allowed to collect, and what can I do with it?" stems from a contract. This contract is written by the company itself, and offered to consumers in a boilerplate form, "take it or leave it," with no space for negotiation.[150] As long as the company clearly states its plans in the "privacy policy," and the consumers accept it, the contents of the contract essentially make up the norms the company must abide by.

Whether a privacy policy should be considered a "real" contract or not is disputed in the legal scholarship, with arguments being raised both for the affirmative and the negative answer.[151] However, this disagreement is about the most common form of enforcement of privacy policies, not their norm-creating nature. What suffices to state here is that the Federal Trade Commission (FTC) is ready to police companies that do not abide by their own policies.[152] It does so based on Section 5, giving the FTC mandate to police "unfair and deceptive trade practices."[153] FTC will both initiate proceedings against companies who do act contrary

states as described in this privacy policy.").

150. For a discussion of legal, sociological and philosophical implications of the mass-usage of boilerplate contracts, see Daniel Markovits, *Good Faith Bargaining in the Shadow of a Form*, *in* DEFENCES IN CONTRACT (Andrew Dyson et al. eds., 2017).

151. For the arguments supporting the affirmative answer, see Oren Bar-Gill, Omri Ben-Shahar & Florencia Marotta-Wurgler, *Searching for the Common Law: The Quantitative Approach of the Restatement of Consumer Contracts*, 84 U. CHI. L. REV. 7 (2017). For the opposite view, see Solove & Hartzog, *supra* note 8, at 595–97.

152. Solove & Hartzog, *supra* note 8, at 598–605.

153. *Id.* at 598–99.

to their own policies,[154] as well as analyze the contents of these policies to determine whether the stipulated practices could be deemed "unfair."[155]

In other words, the freedom of these "contracts" is not absolute, but most of the constraints do not come from any statutes. Daniel Solove and Woodrow Hertzog argue that, even though the majority of the FTC's enforcement activities end up being settled, corporations attach such high importance to the contents of these settlements, that one could even treat them as a new type of "common law."[156] Nevertheless, in their own words:

> There is no federal law that directly protects the privacy of data collected and used by merchants such as Macy's and Amazon.com. Nor is there a federal law focused on many of the forms of data collection in use by companies such as Facebook and Google.[157]

The internal logic of the "notice and choice" model is individual-focused and market-based. Ultimately, users' acceptance of the privacy policies is being equated with their agreement to the practices foreseen, and this agreement is treated as a sufficient legal basis for norm-creation. This model ignores the potential of externalities, disregards collective interests, and posits the market as a superior tool to politics for social ordering. At the same time, when forging this regime, its creators referred to the 1973 Report and, allegedly, were inspired by it.[158] The context of this model's origin is crucial for understanding its logic, and the source of its shortcomings.

---

154. *Id.* at 628–30.

155. *Id.* at 638–43.

156. *Id.* at 610–19.

157. *Id.* at 587.

158. FED. TRADE COMM'N, PRIVACY ONLINE: A REPORT TO CONGRESS (1998) [hereinafter 1998 FTC REPORT], https://www.ftc.gov/sites/default/files/documents/reports/privacy-online-report-congress/priv-23a.pdf.

1. The Emergence of "Notice and Choice": the 1990s and the (Neo-)Liberal Fever

A sympathetic reading of the history of "notice and choice" would try to explain its emergence through a genuine belief of its forgers that the economic freedom leads to other freedoms, competition generally works well, and so markets are the most favorable means for deciding the data management rules. A cynical reading, on the other hand, would be that the "notice and choice" model was essentially invented by the lobbyists of the marketing companies and tolerated by the Government, which saw an opportunity for creating a privately run surveillance system, useful especially in the aftermath of 2001. The correct interpretation does not necessarily lie in the middle, but almost certainly lies between these two positions.

"Notice and choice" was born in the 1990s, continuing the sectoral trend in American privacy regulation. When the Internet went public in 1995, both its potential for commerce and for abuse was quickly recognized. In June 1996 the FTC Staff held a "Public Workshop on Consumer Privacy on the Global Information Infrastructure," and in December published a report summarizing the proceedings.[159] The document opens with a characterization of "limitless opportunities" offered by the new, online marketplace.[160] According to the 1996 report's authors, the benefits stemming from anyone's ability to easily gather personal information online "are apparent, both for consumers and for industry".[161] Overall efficiency will rise because marketers will spend less money on reaching the potential clients, while

159. FED. TRADE COMM'N, STAFF REPORT: PUBLIC WORKSHOP ON CONSUMER PRIVACY ON THE GLOBAL INFORMATION INFRASTRUCTURE (1996), https://www.ftc.gov/reports/staff-report-public-workshop-consumer-privacy-global-information-infrastructure [hereinafter FTC STAFF REPORT].

160. *Id.* ch. 1.

161. *Id.*

consumers will spend less time looking for the information that interests them.[162] However, the authors continue:

> The proliferation of readily available personal information . . . could jeopardize personal privacy and facilitate fraud and deception. These risks may make consumers reluctant to use the Internet or participate in online transactions and therefore could prevent consumers from obtaining the benefits promised by online commerce.[163]

Note the subtle move from privacy intrusions and deception being legitimate concerns on their own terms, to being problematic as potential barriers/deterrents in access to the market. Consumers' primary source of benefits is participation in the marketplace, and so circumstances preventing them from participation should be taken care of. Within this setting, the Workshop participants sought to find the appropriate response. Acknowledging that some cases, particularly health data, financial data,[164] and data about children,[165] might require specialized responses, the overall tone of the report seemed much more favorable to "soft solutions" like education and business self-regulation than legislative intervention.[166] This document, publicly available but still internal, shaped what would follow in the year to come.

In 1998 the FTC submitted its official report, titled "Privacy Online,"[167] to Congress. Reiterating the belief that internet is "an exciting new marketplace for consumers,"[168] the FTC observed that "there are also indications that consumers are wary of participating in it because of concerns

---

162. *Id.*

163. *Id.*

164. *Id.* ch. 2.

165. *Id.* ch. 4.

166. *Id.* ch. 3.

167. 1998 FTC REPORT, *supra* note 158.

168. *Id.* at i.

about how their personal information is used."[169] Having considered both, the FTC stated that, in its view, the proper way to proceed is "to encourage and facilitate effective self-regulation as the preferred approach to protecting consumer privacy online."[170] The 1998 Report contained its own, morphed set of "Fair Information Practice Principles." Referring to the 1973 Report (which the 1998 Report labels "seminal") and several international documents, including the 1980 OECD guidelines and the 1995 EU Directive, the FTC listed five, "widely accepted" principles:

(1) Notice/Awareness;
(2) Choice/Consent;
(3) Access/Participation;
(4) Integrity/Security;
(5) Enforcement/Redress.[171]

Although they are clear and relatable, these principles look little like those expressed in their alleged source, the 1973 Report. On the one hand, this re-formulation must be applauded for its concise and succinct communication of the 1973 Report's plea for public notice, for facilitating data subjects' rights to access and correct the data, and for ensuring security and proper enforcement of the normative frameworks governing processing. On the other, several key aspects are missing.

The 1998 FTC Report does not mention the social aspects of data management at all. Among the threats, it sees privacy and deception as potential barriers to the market; but does not speak about the "coercive potential" so strongly underlined by the authors of the 1973 Report. Apart from direct negative effects stemming from unwanted disclosure of personal information, or failures of the systems to be secure, the Report sees no policy-decisions to be taken

169. *Id.*

170. *Id.* at i–ii.

171. *Id.* at 7–11.

regarding the tradeoffs between more efficient markets and the rising power of private organizations to predict and influence the behavior of individuals. As a result, the recommendations it makes differ from those expressed in the 1973 Report not only regarding the perceived superiority of self-regulation over legislation (i.e. the form of intervention), but also the substance. The 1998 Report does not see a place for the "purpose limitation principle," i.e. the rule according to which data gathered for one purpose should not be used for other purposes. With the exception of data about children, to which the 1998 Report attaches special attention,[172] the view of the potential risks and necessary responses has been much more relaxed than in 1973.

The legal response on the side of the Government closely mirrored these views. Several statutes, regarding the areas identified as sensitive, were passed, including the Health Insurance Portability and Accountability Act of 1996,[173] the Children's Online Privacy Protection Act of 1998[174] and the Gramm-Leach-Bliley Act of 1999.[175] For the rest of the internet-driven data collection practices, self-regulation, to be enforced by the FTC, has been chosen as the preferred approach. As a result, a one-size-fits-all, market-grounded and individual-centered approach has become the paradigm of the American data management law.

One thing is perplexing. How was it possible that in the 1970s, the emergence of huge, bulky and expensive computers used in private and public organization gave rise to worries about individual freedom and necessary political choices; while in the 1990s the emergence of a global network connecting personal computers in everyone's homes to these

172. *Id.* at 4, 12–13.

173. Health Insurance Portability and Accountability Act of 1996, Pub. L. No. 104-191, 110 Stat. 1936 (1996).

174. Children's Online Privacy Protection Act of 1998, Pub. L. No. 106-170 (codified at 15 U.S.C. §§ 6501–6506 (2012)).

175. Gramm-Leach-Bliley Act of 1999, Pub. L. No. 106-102 (codified at 15 U.S.C. §§ 6801–6809 (2012)).

organizations' databases did not? Why, when the potential for abuse got higher, did the political response become more relaxed? Four possible reasons, not necessarily mutually exclusive, come to one's mind.

First, the geopolitical context has changed dramatically. 1973 was the middle of the Cold War, with nuclear proliferation, Space Race, and proxy wars around the globe giving people a general sense of worry. In the 1990s, the overall climate in the West was triumphant. With the defeat of communism, a capitalist liberal democracy in a globalizing world seemed not only the inevitable, but also the best political-economic choice.[176] This triumphant liberalism has been particularly strong regarding the internet, with legal articles expressing doubts about nation states' ability to govern the "cyberspace"[177] and activists expressing opposition to the desirability of doing that.[178] Lawrence Lessig describes these societal feelings well in his *Code 2.0*.[179] Second, one could argue that in 1998 computers were

176. *See* Francis Fukuyama, *The End of History?,* 16 THE NAT'L INT. 3, 3–4 (1989).

177. *See* David R. Johnson & David Post, *Law and Borders—The Rise of Law in Cyberspace*, 48 STAN. L. REV. 1367 (1996).

178. John Perry Barlow, *Declaration of the Independence of Cyberspace*, ELECTRONIC FRONTIER FOUND. (Feb. 6, 1996), https://www.eff.org/cyberspace-independence. Barlow famously expressed the sentiment:

> Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather. We have no elected government, nor are we likely to have one, so I address you with no greater authority than that with which liberty itself always speaks. I declare the global social space we are building to be naturally independent of the tyrannies you seek to impose on us. You have no moral right to rule us nor do you possess any methods of enforcement we have true reason to fear.

179. *See* LAWRENCE LESSIG, CODE 2.0 3 (2006). Professor Lessig writes:

> [I]n the spring of 1995, while teaching the law of cyberspace, I saw in my students these very same postcommunist thoughts about freedom and government. Even at Yale—not known for libertarian passions—the students seemed drunk with what James Boyle would later call the

no longer "new." People got used to the idea, and nothing terrible had happened since the 1970s, so naturally, the worries were lower. Third, one could explain the change by corporate capture. Lobbyists sold the idea to the lawmakers. The 1998 Report is clear about their participation in the FTC Workshops.[180] Marc Rotenberg, writing in 2001, argued:

> The traditional complement to "notice" had long been "consent," and the problem that attracted privacy scholars and policymakers was to determine what would constitute adequate or meaningful consent. . . . The "notice and choice" formulation put forward by the Direct Marketing Association in 1996 provided an opportunity for the marketing industry to avoid resolving the difficult problem of what would constitute meaningful consent.[181]

Seeing benefit in being allowed to collect and use personal data freely, without burdensome regulatory requirements, corporations, in particular marketing professionals, promoted not only the self-regulation as a means to realize the policy goals, but also the very philosophical foundation of market-centered individual freedom. As a result:

> Subtly, but powerfully and profoundly, the substitution of "notice and choice" for "notice and consent" transferred the protection of privacy from the legal realm, and from an emphasis on the articulation of rights and responsibilities, to the marketplace, where consumers would now be forced to pay for what the law could otherwise provide.[182]
>
> . . . .

> "libertarian gotcha": no government could survive without the Internet's riches, yet no government could control the life that went on there. Real-space governments would become as pathetic as the last Communist regimes: It was the withering of the state that Marx had promised, jolted out of existence by trillions of gigabytes flashing across the ether of cyberspace.

180. *See* 1998 FTC REPORT, *supra* note 158, at i.

181. *See* Rotenberg, *supra* note 34, at 10.

182. *Id.* at 11.

> One cannot escape the conclusion that privacy policy in the United States today reflects what industry is prepared to do rather than what the public wants done.[183]

To be clear, not least because the lobbying has not been perfectly documented, establishing a causal link between particular events and the ultimate outcome is a very difficult task. On top of these considerations, one should also note the fourth possible reason, and that is a silent alliance of business and the US Government, where the latter saw an opportunity in creating a privately-owned surveillance technology. I do not want to argue for or against that, though I would mention that the NSA scandal showed that even if this was not already the Government's intention back in the 1990s, it clearly grasped that potential in the years to follow.[184]

The overall ideology in which the "notice and choice" model has been forged was a market- and individual-centered economic liberalism. As a result, the problem of data management by tech companies has been reduced to "consumer privacy," and the solution to this problem was by definition going to concentrate on individual interests, best protected (in the view of the model's creators) by the market. Furthermore, the "privacy" framing meant that law governs only the situations when an individual can be "personally identified," and not those where she is anonymous, even if she can still be affected by data practices. These beliefs make up the ideological and legal foundation of the data analytics society in 2020.

2. The Original Sin: A Market-Individual as the Sole Subject of Disclosure and Decisions

There are three problems with constructing a data-management law around an individual who is expected to

---

183. *Id.* at 34.

184. *See* Karina Rider, *The Privacy Paradox: How Market Privacy Facilitates Government Surveillance*, 21 INFO., COMM. & SOC'Y, no. 10, 2018, at 1369, 1369 (2018).

make decisions as a market actor. First, individuals' understanding of "privacy policies" is necessarily limited. Second, even if individuals understood these documents, given the existence of social costs, they are not normatively competent to decide to impose these costs on other people. Third, individuals' say about these matters should not be voiced as a part of a market transaction, where they might be in haste and have immediate interests at stake.

The idea of a "privacy policy" in the "notice and choice" model is very different, even if seemingly connected, to the "public notice" requirement advocated by the 1973 Report. The addressee of the document is the consumer, not the public. And the function of the document became a norm-creating one, not merely information-conferring. The authors of the 1973 Report imagined a world where public notices allowed the public to engage in political deliberations and, ultimately, legislation setting the limits on data collection and usage. However, the paradigm adopted in the late 1990s turned these "privacy policies" into contracts, ceding the norm-making to the individual market actors.

Series of empirical studies now demonstrate that people do not read privacy policies,[185] and even if they do, they do not understand them.[186] Findings like these led Omri Ben-Shahar and Carl Schneider to call mandated disclosure "the most common and least successful regulatory technique in American law."[187] This has to do with the human condition, or what Daniel Solove calls structural and cognitive reasons,[188] but also with the condition of the privacy policies

185. *See* Yannis Bakos et al., *Does Anyone Read the Fine Print? Consumer Attention to Standard-Form Contracts*, 43 J. LEGAL STUD. 1, 32 (2014); Obar & Oeldorf-Hirsch, *supra* note 20, at 129.

186. Joel R. Reidenberg et al., *Disagreeable Privacy Policies: Mismatches between Meaning and Users' Understanding*, 30 BERKELEY TECH.L.J. 39, 69 (2015).

187. OMRI BEN-SHAHAR & CARL E. SCHNEIDER, MORE THAN YOU WANTED TO KNOW: THE FAILURE OF MANDATED DISCLOSURE 3 (2014).

188. *See* Daniel J. Solove, *Introduction: Privacy Self-Management and the*

themselves. Not only is their form often difficult to comprehend, but also do not contain enough information to make an informed decision.[189] However, even if one assumes that these problems could somehow be mitigated, the other two remain.

Let us come back to the example of data-driven price discrimination. Imagine that I do fully understand that disclosing what my job is on Facebook will make it easier to infer what the wealth of other people is. Imagine I do understand that liking certain pages, or listening to particular music, makes it even easier to draw such links. Imagine that I even understand that as a result of my acceptance of Facebook's privacy policy, some people will be shown higher prices of goods and services. Whether I consider this a good deal or bad deal is beside the point here; what matters is that I should not be given this choice in the first place. The same holds with regard to other social costs of data management: discrimination or behavioral effects. I should not be the one to make this decision, and especially not while looking for something online, and hastily "accepting" privacy policies.

The limits for data collection and usage should be established in a political process and enshrined in regulation. The genetic code of the "notice and choice" model is contrary to these requirements. The "notice and choice" model essentially allows the companies to write the rules on these limits, put them in the consumer "contracts," and derive their rights from individuals' "choices."

Consent Dilemma, 126 HARV. L. REV. 1880, 1881 (2013).

189. *See* GIUSEPPE CONTISSA ET AL., BEUC, CLAUDETTE MEETS GDPR: AUTOMATING THE EVALUATION OF PRIVACY POLICIES USING ARTIFICIAL INTELLIGENCE 3 (2018), https://www.beuc.eu/publications/beuc-x-2018-066_claudette_meets_gdpr_report.pdf.

B. *European Model: "Personal Data Protection" Approach*

In the meantime, our European colleagues have developed a very different approach to dealing with data management. The "personal data protection" model, currently expressed in the GDPR, is a general law applicable to all private and public organizations "processing"[190] data about residents of the European Union.[191] Unlike the "notice and choice model," the GDPR does not see the relationship regarding data as a market transaction. Rather, it creates a range of administrative rules, governing any operation undertaken on personal data; rules identical for businesses and public authorities. These rules cannot be changed by contract, and their violation results in administrative fines imposed by specialized supervisory authorities.[192]

---

190. GDPR, *supra* note 32, art. 4, para. 2. The category is defined broadly:

> [A]ny operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organization, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

191. GDPR, *supra* note 32, art. 4, para. 1. "Personal data" is defined as follows:

> [A]ny information relating to an identified or identifiable natural person ("data subject)"; an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

Hence, any operation regarding any data about a person who can be identified falls within the material scope of the Regulation. It governs processing activities of both private and public entities, *id.* art. 2, by firms incorporated in the EU or "not established in the Union, where the processing activities are related to: (a) the offering of goods or services, irrespective of whether a payment of the data subject is required, to such data subjects in the Union," or "(b) the monitoring of their behavior as far as their behavior takes place within the Union." *Id.* art. 3, para. 2.

192. *See* GDPR, *supra* note 32, art. 83.

GDPR divides its regulatory framework into three substantive groups: general principles,[193], data subjects' rights[194] and data controllers' obligations.[195] However, within the chapter on the subject's rights, three substantial articles are devoted to the issue of "transparency," being basically an obligation to engage in the public notice.[196] If one, for analytical purposes, separates these three from other rights, and at the same time treats the data controller's obligations as correlated with subjects' rights,[197] one will see the familiar triad of (i) general principles; (ii) public notice; and (iii) subjects' rights, exactly as advocated by the 1973 Report. However, one significant difference must be brought to the fore: the GDPR treats "data protection" and "privacy" as human rights. Among its three stated objectives, the Regulation lists: "[protection of] fundamental rights and freedoms of natural persons, in particular their right to the protection of personal data."[198]

This has significant consequences for the shape of the "personal data protection" model. The legal requirements for consent to be valid are much more rigid than those necessary for a "choice" in American law.[199] Further, any individual interest violated through an operation undertaken on their

193. *Id.* ch. II.

194. *Id.* ch. III.

195. *Id.* ch. IV.

196. *See* Article 29 Working Party, *Guidelines on Transparency under Regulation 2016/679* (Apr. 11, 2018), https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=622227.

197. Rights are correlated with duties. *See* Wesley N. Hohfeld, *Fundamental Legal Conceptions as Applied in Judicial Reasoning*, 26 YALE L.J. 710, 717–19 (1917).

198. GDPR, *supra* note 32, art. 1, para. 2. The other two stated objectives include "[laying] down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data" and facilitating "free movement of personal data within the Union." *Id.* art. 1, para. 1. The last objective stems from the EU's goal, establishing a common market.

199. *See* GDPR, *supra* note 32, art. 7; *cf. supra* Section III.A.

personal data might be treated as a human rights violation. And human rights, unlike economic interests, cannot be a subject of "cost-benefit analysis,"[200] either in regulation or in adjudication. They might be balanced against other human rights,[201] but not against interests like economic efficiency or innovation. This, taken together with the EU's "direct effect doctrine,"[202] results in an extremely rigid governance framework, difficult to comply with by corporations collecting data as a part of their business model, as many of the tech companies do.[203]

At the same time, this law is still focused on the individual, and frames her interests as a problem of "privacy," just like the "notice and choice" model. This renders the GDPR blind to collective interests and completely toothless when data cannot be directly linked to an "identified or identifiable" person. Finally, GDPR also applies a one-size-fits-all, technocratic approach to decide what data practices are acceptable or not. The difference being that the technocracy here is the domain of human rights, not microeconomics. Either way, no space for political choices has been foreseen. To understand the internal logic of this system, let us take a look at historical and ideological conditions of its emergence.

1. GDPR's Genealogy: European Convention of Human Rights

The first issue regarding the European approach to data management is that the laws under consideration are

200. For a contemporary restatement of the cost-benefit analysis theory, and its place in the functioning of the regulatory state, see CASS SUNSTEIN, THE COST-BENEFIT REVOLUTION (2018).

201. For the theoretical treatment of the rights-balancing in law, see Robert Alexy, *On Balancing and Subsumption: A Structural Comparison*, 16 RATIO JURIS 433, 436, 439 (2003).

202. *See* FRANTZIOU, *supra* note 110.

203. Scholars have gone as far as to call the GDPR "incompatible" with the very idea of big data analytics. *See* Tal Z. Zarsky, *Incompatible: The GDPR in the Age of Big Data*, 47 SETON HALL L. REV. 995, 1017 (2017).

explicitly not a part of consumer law. Even though the European Union has a robust system of consumer law in place, in some ways mirroring the American approach (rules against unfair and misleading commercial practices)[204] and in some ways going further (rules against unfair contractual terms);[205] these tools have not yet been put to work regarding consumer data. Consumer agencies (FTC's counterparts) are separate entities from the Data Protection Authorities.[206] In 2017, two legal articles suggesting using consumer law tools to advance data protection goals have been welcomed as a new and refreshing view.[207] How does one explain this surprising difference between the two systems?

American "notice and choice" model came in to fill a normative void. In the 1990s there were no rules governing consumer data collection online, so something was needed. In Europe, however, there was no such void. In 1995, the year the Internet went public, the EU adopted the Data Protection Directive, GDPR's predecessor, harmonizing many of the already existing national data protection laws.[208] These national laws, in turn, were either enacted or sustained to meet Europe's nation states' obligations under international law, stemming from the Convention 108 of the Council of Europe. So, when the Americans were discussing what to do about internet-facilitated data collection, their

---

204. *See* Directive 2005/29/EC, of the European Parliament and of the Council of 11 May 2005 Concerning Unfair Business-to-Consumer Commercial Practices in the Internal Market, 2005 O.J. (L149) 22.

205. One could, roughly, summarize them as a "legislatively enshrined unconscionability doctrine." *See* Council Directive 93/13/EEC, 1993 O.J. (L 95) 29.

206. These are called "Independent Supervisory Authorities" in the GDPR. GDPR, *supra* note 32, ch. VI.

207. *See* Natali Helberger et al., *The Perfect Match? A Closer Look at the Relationship Between EU Consumer Law and Data Protection Law*, 54 COMMON MKT. L. REV. 1427, 1429–31 (2017); Nico van Eijk et al., *Unfair Commercial Practices: A Complementary Approach to Privacy Protection*, 3 EUR. DATA PROTECTION L. REV. 325, 325 (2017).

208. *See* FUSTER, *supra* note 108, at 55–65.

European counterparts already had their own national data management laws, by having just adopted new, shiny, state-of-the-art legislation to tackle the problem. These laws contained the same catalogs of principles, rights, and obligations as the GDPR does nowadays, and were applicable both to public and private actors. However, they were enacted within a human rights mindset.

Council of Europe, the body which adopted the Convention 108, was established back in 1949, in the period directly following the atrocities of the Second World War and Nazism, and when the Cold War between democratic West and communist East was materializing.[209] Among its various duties, it was tasked with the protection of human rights.[210] Most notably, in the early 1950s, it adopted the European Convention on Human Rights (ECHR),[211] followed by a series of issue-specific international human rights treaties. Among its many provisions, the ECHR contains Article 8: "Right to respect for private and family life." The article reads:

> 1. Everyone has the right to respect for his private and family life, his home and his correspondence.
>
> 2. There shall be no interference by a public authority with the exercise of this right . . . .[212]

For an American, this is a familiar language. The character of this provision could be, roughly, compared to the American Bill of Rights. Not in content, but in form. The story goes like this: we had some really bad experience with the government in the past, a big part of that bad experience was the government not respecting our privacy, and so now we put a constitutional constraint on the government to prevent it from doing so.

---

209. *See* BIRTE WASSENBERG, HISTORY OF THE COUNCIL OF EUROPE 9–12 (2013).

210. *Id.*

211. Convention for the Protection of Human Rights and Fundamental Freedoms, Nov. 4, 1950, 213 U.N.T.S. 221.

212. *Id.* art. 8.

In 1981, when (automated) information processing was becoming widespread, human rights lawyers realized that a qualitatively new possibility for a state to intrude in the life of an individual emerged. Hence, the Council of Europe adopted the Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data.[213] It did so as a part of the Article 8 mandate. It read:

> Article 1: The purpose of this Convention is to secure . . . for every individual . . . respect for his rights and fundamental freedoms, and in particular his right to privacy, with regard to automatic processing of personal data relating to him ("data protection").[214]

What follows is a text extremely similar to the 1980 OECD guidelines and quite faithful to the 1973 Report. This framing will make its way to the national laws of European countries. In 1995 national laws enacted to fulfill the obligations imposed by the Convention 108 were harmonized,[215] and in 2000 the European Union enacted its own Charter of Fundamental Rights.[216] In 2009, this Charter became a part of the EU's primary law (functionally the European Union's "Constitution").[217] Within it, one finds a new fundamental right to "protection of personal data," expressed in addition to the right to privacy:

---

213. *See* Convention 108, *supra* note 75, art. 1.

214. *Id.*

215. "Harmonization of laws" is one regulatory strategy of the European Union, occurring through an adoption of a directive. Directives are not directly binding on individuals, but oblige the Member States to enact state-legislation achieving the desired effect. *See* KAREN DAVIES, UNDERSTANDING EUROPEAN UNION LAW 56–62 (2019).

216. Charter of Fundamental Rights of the European Union, Dec. 7, 2000, 2012 O.J. (C 326) 391 [hereinafter the EU Charter].

217. This statement is not entirely accurate in technical jargon, but best manages to explain the role played by the primary law in the European Union. For the explanation of the technical and political issues at stake when invoking the term "constitution," see JOSEPH H.H. WEILER, THE CONSTITUTION FOR EUROPE 102–07 (2001).

> 1. Everyone has the right to the protection of personal data concerning him or her.
>
> 2. Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified.
>
> 3. Compliance with these rules shall be subject to control by an independent authority.[218]

In 2016, with the right to protection of personal data enshrined in its "Constitution," the EU unified its data protection law in the GPDR.[219] This new law would be noticed internationally for several reasons. First, it applies extraterritorially, also to companies with a seat outside of the EU, if only they direct their goods and services at the EU residents, or monitor their behavior.[220] Second, unlike its predecessors, the GDPR gives the Data Protection Authorities competence to fine violators up to 4% of their yearly revenue.[221] Third, the GDPR introduces several innovations, potentially interesting for other reformers. Those include the (in)famous right to be forgotten,[222] right to explanation,[223] legislatively-mandated obligations to

---

218. The EU Charter, *supra* note 215, art. 8.

219. "Unification of laws" is another regulatory strategy of the European Union. It occurs through an adoption of a regulation which, unlike a directive, is directly binding on the individuals and does not require implementation by the Member States. *Regulations, Directives, and other acts*, EUROPEAN UNION, https://europa.eu/european-union/eu-law/legal-acts_en (last visited Mar. 17, 2020).

220. *See* GDPR *supra* note 32, art. 3, para. 2.

221. *See id.* art. 83.

222. *See* Rosen, *supra* note 88, at 88.

223. *See* Margot E. Kaminski, *The Right to Explanation, Explained*, 34 BERKELEY TECH. L.J. 189, 193 (2019).

introduce "privacy by design,"[224] or the institution of a "data protection impact assessment."[225]

Substantively, however, the GDPR is not a revolution, but a subtle evolution of the system existing as binding, in one way or the other, since 1981. The framework of general principles, public notice mechanisms, subjects' rights, and controllers' obligations existed already in the 1980 OECD Guidelines and the Convention 108; and their contents, even though updated and made more rigid, very much resemble these documents. Documents which, in turn, can be traced back to the 1973 Report. However, even though faithful to some of the tasks it set forth, the European "personal data protection" model also forgot some of the most important insights of its intellectual predecessor. Just like the "notice and choice" regime, it ended up constructing a system focused on privacy (and data "protection") and on individual interests. And just like its American counterpart, it employs technocratic means of decision-making in place of political ones.

2. The Original Sin: Individualistic, Technocratic, Human-Rights Mindset

Unlike the "notice and choice" model, the "personal data protection" approach did not give in to the corporate capture and retained several useful data management tools proposed by the 1973 Report. First, it established a set of general principles and processors' obligations through legislation, taking certain questions away from the market. Second, it retained the separation of "fact-conferring" and "norm-creating" functions of "privacy policies." In the European legal frame, these documents are not contracts, but rather transparency mechanisms,[226] playing a role similar to

224. *See* AURELIA TAMÒ-LARRIEUX, DESIGNING FOR PRIVACY AND ITS LEGAL FRAMEWORK: DATA PROTECTION BY DESIGN AND DEFAULT FOR THE INTERNET OF THINGS 84–87 (2018).

225. GDPR *supra* note 32, art. 35.

226. *See* FREDERIK J. ZUIDERVEEN BORGESIUS, IMPROVING PRIVACY PROTECTION

"nutrition facts tables" on food packaging.[227] In this sense, the GDPR does account for collective interests to a certain degree. Just as one cannot contract-out from traffic laws when entering a taxi, one cannot contract out from the prohibition to profile people's political beliefs or religious convictions enshrined in the GDPR.[228] Third, the internal structure of the system, visible in its very name ("general") leaves space for further legislative interventions, tackling certain problems issue-by-issue. However, no such specialized laws have been enacted.

Despite some good intuitions, the GDPR commits the same mistakes as the "notice and choice" model. It focuses on "privacy" and "data protection," and does not apply when data is not-personal. Further, it concentrates on the individual interests and individuals' "control over their personal data." The notice that companies must engage in needs to be concrete and comprehensive,[229] but ultimately should be phrased in "plain and intelligible language, easy to understand."[230] The disclosure rules belong to the chapter on subjects' rights, further indicating that the individual, not the society, is the intended recipient of these documents. Finally, the GDPR also does not leave much space for political decisions, replacing the market approach with human rights, like the technocratic means of establishing rules.

Probably the greatest omission of the European system is a lack of a political mechanism to further specify data

---

IN THE AREA OF BEHAVIOURAL TARGETING 106–09 (2015).

227. If companies lie about their data practices, they will be fined. GDPR, *supra* note 32, art. 83. However, the mere fact that an individual "consents" to what the privacy policy stipulates does not yet mean that the company is allowed to engage in these practices. They must comply with all the other norms of the GDPR, including purpose limitation, data minimization, etc.

228. One can consent to processing these types of data about oneself, but this consent does not allow the company to process this type of data about other people. *See* GDPR, *supra* note 32, art. 9.

229. *See* Article 29 Working Party, *supra* note 196, at 8–9.

230. *Id.*

management rules for particular sectors. The "purpose limitation principle" makes it unlawful to process data for reasons other than those stipulated, but the GDPR does not say anything about which purposes should be treated as lawful, and which not. The questions visible in the examples discussed in Part II: "can a company use data to price discriminate? to engage in behavior manipulation? to segment the society?" are not, and cannot, be answered by the GDPR's frame. Other legal instruments should step in. However, no such rules have been enacted, and currently, no such rules are pondered in the EU. Pervasively, despite everything written about the differences of the American and the European approach, it is the "consent" of the data subject, given in the market conditions, that shapes the decisions on the acceptable limits to data management. This consent is much more difficult to obtain; the limits of what one can consent to are much stricter.[231] However, ultimately, it will be the individual who decides to impose costs on others. And if a data practice is questioned in a court, the decision regarding its permissibility will be based either on the interpretation of technocratic procedures of the GDPR itself or on the general human rights provisions.

To be clear, I do not want this criticism to sound as if I oppose human rights protections and human rights movements. On the contrary, I consider them to be some of the most important achievements of the 20th century legal systems. However, I oppose the idea that when it comes to Facebook's or Google's business models, human rights will be the ultimate answer to what should be allowed. Obviously, they might offer us a view on what are the boundaries that should not be crossed; but should not be used as the source

231. One should mention, among others, the prohibition of processing of the "special categories of data," defined as "personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation." *See* GDPR, *supra* note 32, art. 9.

of all normative arguments. What exactly to do with data collection enabling behavioral manipulation or societal segmentation are political decisions, and should be taken in a political manner.

***

Having shown where the current legal regimes are lacking, what socio-technological aspects escape their reach, and how to historically explain the reasons for their insufficiency, I would now like to sketch a proposal on how to go about amending these legal frames.

IV. DATA MANAGEMENT LAW FOR THE 2020S

In every jurisdiction, one can describe the contents of its data management law. If we ask the question: what are the rules governing the collection, analysis, and usage of data about people; we will always get an answer. The answer might be (and currently is): there are very few rules; most of the decision-making is left to the market. However, there is always some answer. In this sense, the novelty of the "data management law" approach lies not only in the new rules it calls for but also in the change of the mindset allowing one to grasp this feature of the data analytics society. Some rules always emerge, there is always someone to draft them. There is no escape. The question is: who will draft those rules, what form will they take, and whose interest will they protect?

I argue that in the 2020s, we should focus our efforts on developing a data management law able to account for, and mitigate, the social costs of data analytics. Data-driven technologies, making our lives more efficient and more convenient, will continue to be developed. Entities developing them will continue to be profit-driven companies. Someone will have to pay the costs. The authors of the 1973 report feared that those will be "some of our most disadvantaged citizens."[232] Empirical research, for example on data-driven discrimination, now shows this prediction was correct.[233] Moreover, we are paying costs as a society as a whole. Some of these costs, we already see and begin to understand, as behavior manipulation in the political process.[234] Some others, like social segmentation in the commercial communications we receive, can be discussed only speculatively, just as in 1973. These costs, unless we want to ban the technology as a whole,[235] cannot be entirely

232. 1973 REPORT, *supra* note 33, at vi.

233. *See* Sweeney, *supra* note 6, at 33–34.

234. *See* Confessore, *supra* note 143.

235. Which, in itself, would generate enormous social costs.

wiped out. However, they can be channeled in a much more conscious manner than the current regimes, the American "notice and choice" and the European "personal data protection," allow us to do.

The policy recommendations of this Article could be summarized in three claims. First, we must go beyond "privacy" (or data "protection") as the frame for conceptualizing social costs of data analytics, and replace it with a more inclusive "social costs of data management." This is necessary both to account for interests other than individual privacy and to govern situations in which individuals remain anonymous from the point of view of the company targeting them. Second, we must account for collective interests on top of the individual ones. The society, not the individual, should be the addressee of disclosure about data practices; likewise, the society, not the individual, should have a say on the limits of these practices. The integrity of the political process, non-discrimination or public health, are not matters to be decided by the sum of individual preferences. Third, these decisions are political in nature, and so should be taken via a political process. Societal views might differ when people are asked about the desirable limits of behavior manipulation, or social segmentation, in the market situations and during elections. They might differ when the product offered on the market is clothes and when it is pharmaceuticals. And many interests will have to be balanced against one another. That is why we need to tackle the social costs of data management one by one, issue by issue; and tackle them in a political manner, on the public forum, and not only in the offices of the technocrats.

Just like the authors of the 1973 Report, I would like to make clear that the policy recommendations outlined below are not meant to give us a final answer on any given data-management issue. Nor can they guarantee that in the end, everyone's interest will be accounted for and no one will have to bear any costs. However, the envisioned paradigm will guarantee that the decisions about these costs will be taken

by the polity, not the by the corporations alone, and that they will be made in a process that is as "open, informed, and fair"[236] as possible. Let us now take a look at each of the three pillars of this new framework.

A. *Beyond "Privacy": Social Costs of Data Management*

Privacy obviously matters, but achieving privacy will not shield us from other social costs of data management. Concentrating on privacy might make us overlook other interests put at stake by corporate data collection and usage. There are at least two reasons to replace the conceptual frame of "privacy"—currently employed by the "notice and choice" and "personal data management" regimes—with "social costs of data management."

First, when speaking about "privacy," lawyers tend to conflate two distinct social phenomena: disclosure of data about individuals, and use of data about individuals. I argue that the term "privacy" should be used to speak about the former,[237] while the latter should become the domain of "data management." The authors of the 1973 report treated these issues jointly, following the writings of their time.[238] However, both for the sake of conceptual clarity and of effective policy-making, we should look at them separately. Consider two examples. Imagine that I enjoy listening to Dutch marching music on Spotify. For social reasons, I would prefer that my colleagues, who think this is silly, do not learn about it. At the same time, I welcome Spotify suggesting me similar music, for example, Belgian marches. Or, imagine that I share a photo of my face on Facebook, and accept that everyone can see it. I am fine with my friends "sharing" it, or Facebook "disclosing" it to all its users. At the same time, I do not want Facebook, or anybody else, to be able to use this

236. 1973 Report, *supra* note 33, at 43–44.

237. This is consistent with the traditional understanding of privacy torts. *See supra* note 67 and accompanying text.

238. *See supra* Section I.B.

photo to develop facial-recognition technologies, and certainly not to learn how to recognize my face. The mere fact that I do not want my data to be disclosed does not yet mean that I oppose using this data by a company, and the other way round. Referring to both issues with a single term "privacy" leads to conceptual confusion, and renders policy discussions less nuanced.

Second, mechanisms of "privacy" and "personal data protection" laws are triggered only when dealing with "identified or identifiable" individuals.[239] If a company knows (or could know) that it is collecting data about my friend Bob Smith, its actions fall within the obligations imposed by the "notice and choice" or "personal data protection" regimes. However, if the company takes proactive steps *not* to identify him, and only collects data about his behavior and demographics, they are not bound to respect privacy laws.[240] From the point of view of someone who wants to price-discriminate, or show some ads only to people who belong to a certain social class, or to manipulate someone's behavior in the economic or the political sphere, it really does not matter what is your name, or Social Security Number. All they need is to assign you some (random) identification, and analyze data about "someone" with that number, not *you* personally. From the "privacy" point of view, such a world would be wonderful. If all the personally identifiable information "stayed" at users' devices, and only anonymized data was transferred, our wish for privacy would be fulfilled. However, from the point of view of the social costs of data management, this changes nothing. There will still exist data-driven discrimination, exclusion, addiction, and behavioral manipulation.

---

239. *See* Schwartz & Solove, *supra* note 26, at 1817–18.

240. Scholars dispute whether perfect anonymization is practically feasible. I do not offer an argument for or against that claim. Rather, I show that even if it was, and privacy considerations were taken care of, the problem of social costs of data management would persist.

In the data analytics society, data-costs can be imposed on individuals and social groups, even when the data collected and analyzed is not private (individuals freely disclose it) and even if the individuals are anonymous from the point of view of the corporation using it. This is why, to counter the social costs of data management, we need to move beyond "privacy" and towards "data management law." Within it, we must identify other individual and societal interests: decisional autonomy, paired with the integrity of economic and political processes; freedom from addiction, paired with public health considerations; non-discrimination, paired with social justice and equality; and many more. Axiologically, those are not new problems, and the values at stake here belong to the foundations of a liberal, democratic society. However, the data-driven means of intrusion upon these interests, based on seamless data collection, inferred knowledge and automated decision-making call for new legal responses. These responses must tackle the costs of data management but must occur outside of the frame of "privacy."

Two caveats are due. First, I am by no means claiming that "privacy is dead," or "not important," or that one should not protect it. On the contrary, I admire and support the work of privacy advocates, and believe that we should stand our ground regarding the privacy matters. However, as I demonstrated, the majority of interests at stake in the data analytics society are conceptually distinct from privacy and can be better tackled through other channels. Second, I am not claiming that theoretically speaking, these problems cannot be explained and accounted for using privacy language. Privacy is an intellectually rich field, full of fine-tuned arguments, distinctions, and theories.[241] If one really

241. *See, e.g.*, Lisa M. Austin, *Privacy and Private Law: The Dilemma of Justification*, 55 MCGILL L.J. 165, 210 (2010); Ignacio N. Cofone & Adriana Z. Robertson, *Privacy Harms*, 69 HASTINGS L.J. 1039 (2018); Julie E. Cohen, *What Privacy Is For*, 126 HARV. L. REV. 1904 (2013); Daniel J. Solove, *I've Got Nothing to Hide and Other Misunderstandings of Privacy*, 44 SAN DIEGO L. REV. 745

wanted to, one probably could employ the mix of "information privacy" and "decisional privacy" theories to tell the story of the social costs of data management under that one label. However, as I have shown, the stakes are too high to sacrifice the possibility of managing all the costs at the altar of conceptual austerity.

Once we are able to go beyond privacy when framing the problems associated with data management, we should concentrate on the second shortcoming of the "notice and choice" and "personal data protection" models. And this is focusing solely on the individual interests while remaining blind to collective ones, and to the overall systemic health.

B. *Beyond Individual Interests: "Societal Notice" and "Societal Choices"*

Current data management laws value individual interests higher than collective ones. In a way, they are created solely to protect individual interests. The American "notice and choice" model concentrates on the individual consumer expressing her market preferences, while the European "personal data protection" model exists to protect the human rights of an individual data subject. However, given the potential for externalities, and the fact that one person's disclosure imposes data-driven costs on other people, individuals should not be the only decision-makers. Certain collective interests, or interests of certain minorities, can only be protected as a result of collective action.

In both systems, an individual is the intended recipient of the disclosure (occurring through the "privacy policies") and the ultimate decision-makers (acting through "choice" or "consent"). I argue that these two elements of the data management laws should be modified.

Let us begin with disclosure. The 1973 Report called for legislatively mandated "public notice," aimed at facilitating political deliberations. Forty-five years later, in a world in

(2007).

which we failed to live up to the tasks set forth in the report, Shoshana Zuboff observed: "Surveillance capitalists know everything about us whereas their operations are designed to be unknowable to us."[242] Currently existing "privacy policies" not only are written in a language difficult to understand[243] but often simply do not contain the information necessary to fully comprehend corporations' operations.[244] Usage of open-ended phrases like "including," "among others"; conditional forms like "we might," or vague terms like "business partners," "research purposes" have been documented by researchers[245] and criticized by experts.[246] Hence, the reasons why so few people fully understand what companies do with personal data pertain not only to the limitations of humans [247] but also the fact that the information often simply is not there. To be fair to the companies, there are good legal reasons for vagueness. If a policy must be understandable to the individual, one cannot expect a company to produce a 100-page long document containing a comprehensive list of all the business partners, types of uses and categories of data. Arguably, the requirements of "comprehensiveness" and "comprehensibility" cannot be fully reconciled.

For these reasons, data management law for the 2020s should change the recipient of disclosure from an individual consumer (or data subject) to the community as a whole.[248] We should require that companies collecting, analyzing, and using data disclose the identity of all the corporations that

242. *See* ZUBOFF, *supra* note 59, at 11.

243. *See* Reidenberg, *supra* note 186, at 39, 51.

244. *See* Contissa, *supra* note 189.

245. *Id.*

246. *See, e.g.*, Article 29 Working Party, *supra* note 196, at 6–7.

247. *See* Solove, *supra* note 188, at 1880.

248. This would also include the active regulators, should they be established by law. For the arguments supporting this type of disclosure, see Rory Van Loo, *The Missing Regulatory State: Monitoring Businesses in an Age of Surveillance*, 72 VAND. L. REV. 1563 (2019).

they share data with (or receive data from), all types of data they analyze, and all uses to which they put it. For the disclosure to be meaningful, these "notices" should be absolutely concrete and comprehensive. We should forbid the usage of open-ended ("including . . .") and conditional ("may") phrases and require concrete examples and descriptions of all the general terms, like "personalized advertising" or "improving services." What types of goals are pursued? What types of improvements envisioned? This should be concretely specified. Other transparency mechanisms could be introduced as well. For example, one could ponder the creation of a national registry of data brokers (which already exists on the state level in Vermont)[249] and the national repository of data sales and licensing contracts (similar to the repository managed by the SEC).[250]

One could argue against these types of disclosure obligations, claiming that they would generate "too much data to process." If every company was supposed to concretely describe *all* their data practices, would these documents not end up being hundreds of pages long each? This is a legitimate, but an unfounded, worry. Yes, these documents could be extremely long. However, we need to remember that when the corporations encountered this problem, they developed new data analytics technologies. There is no reason for the public not to use similar tools. Researchers have shown that techniques like machine learning, including natural language processing, can be employed to empower civil society by increasing their data processing capacities.[251] Just like Google or Facebook equip

249. *See* Douglas MacMillan, *Data Brokers Are Selling Your Secrets. How States Are Trying to Stop Them*, WASH. POST (June 24, 2019), https://www.washingtonpost.com/business/2019/06/24/data-brokers-are-getting-rich-by-selling-your-secrets-how-states-are-trying-stop-them/.

250. *See* James A. Overdahl, *A Researcher's Guide to the Contracts of Firms Filing with the SEC*, 34 J.L. & ECON. 695, 695–99 (1991).

251. *See* Marco Lippi et al., *CLAUDETTE: An Automated Detector of Potentially Unfair Clauses in Online Terms of Service*, 27 ARTIFICIAL INTELLIGENCE & L. 117, 117–18, 135–36 (2019).

their human employees with powerful data analytics technologies; activists, journalists, and public authorities could increase the factual capabilities of lawyers working for them with the usage of the same techniques.[252]

This would also separate the fact-conferring function of "privacy notices" from the norm-creating one. Another objection one raises against increasing the requirements for concreteness and comprehensiveness of these documents has to do with corporations' mixed incentives. If companies can only do what they state in the privacy policy, and the privacy policy must list all the business partners and all purposes of using data, etc., would that not stifle innovation, or incentivize companies who want to innovate not to disclose? This worry holds only within the current paradigm of "notice and choice," where the "privacy policy" is, at the same time, a disclosure mechanism and the contract. However, under the data management law, these two would be separate. The norms governing what data uses are acceptable would be enshrined in legislation, and if the companies would like to conclude additional boilerplate contracts, these contracts would be different documents that the "public notices."

The purpose of these legislatively mandated disclosures would be to increase to social scrutiny of data practices and to enable political decisions about the necessary rules. Further, they would enable proper oversight and enforcement of these rules. These rules, in turn, should stem not from technocratic means like the operation of the markets, or human rights adjudication, but from political choices made by society. What would that look like?

C. *Beyond Technocracy: On the Necessity of Politics*

Data management aimed at minimizing and allocating social costs of personal data processing necessarily entails choices. The question is: who should take them, how should

252. Harry Surden, *Machine Learning and Law*, 89 WASH. L. REV. 87, 87, 102 (2014).

they be taken, and based on what normative considerations. Having argued for transferring these choices from the individual to the collective sphere, I would now like to outline how this should occur in a sectorial, issue-by-issue basis; how the normative considerations should be political and not technocratic; and present some deliberative and regulatory strategies for the legal reformer.

1. Against One-Size-Fits-All, and for Sectorial Solutions

The authors of the 1973 Report argued against one, overarching system of managing all social costs of data analytics, writing:

> The number and variety of institutions using automated personal data systems is enormous. Systems themselves vary greatly in purpose, complexity, scope of application, and administrative context.[253]

However, both the "notice and choice" and the "personal data protection" regimes ended up instituting one-size-fits-all solutions. Even if American privacy law can be called "sectoral," the "notice and choice model" is, in itself, a one-size-fits-all regime. Every instance of data collection, analysis, and usage occurring in the online commercial context is governed by corporate self-regulation and market transactions. It does not matter whether the question concerns using my shopping history to show me recommendations of other products, using information about me to infer knowledge about others, or using that inferred knowledge to price discriminate or manipulate their behavior. The decision always occurs in the form of a boilerplate contract. Similarly, the European "personal data protection" regime applies exactly the same set of principles, rights, and obligations to all situations. A pizzeria keeping my phone number to text me promo codes, and online advertisers trying to convince people to vote "yes" in Brexit referendum based on data I have disclosed, are governed by

253. 1973 REPORT, *supra* note 33, at 43.

the same law. As if the stakes were the same, and the societal choices identical.

To address the social costs of data management, we should return to the approach suggested by the 1973 report, and practiced in the US until the late 1990s, i.e. sectoral legislation. In the data management law for the 2020s, we should remain faithful to the practice of addressing various types of data-driven costs one by one. Legislative deliberations regarding behavioral, targeted advertising; addictive design employed by social media, or price discrimination are long overdue. And the outcomes of these deliberations, both regarding the questions of whether to regulate, and how to regulate, need not, and should not, be always the same. A statute governing price discrimination in consumer products like clothes needs not to employ the same rules as a statute governing price discrimination in pharmaceuticals. The principles used in these laws need not be the same as the principles used to counter behavior manipulation in consumer markets. Which, in turn, need not be identical to those aimed at preventing behavior manipulation in the political sphere. Prevention of data-driven social-media addiction can be achieved using different channels than non-discrimination in job advertisements.

Note that in some spheres we might even accept the normative frames enshrined in the existing approaches. For example, when data analytics in criminal law are concerned, human rights is probably the correct standard to apply. Similarly, in some market conditions, where the potential for social costs is low, we might be fine with leaving some decisions to the market. However, none of these two should be the single standard aimed at solving all the problems; and the question what standard to employ where is, in essence, a political one.

2. Bring the Questions to the Public Debate

As a society we allow individuals to engage in activities that might impose costs on others. Sometimes, we are willing

to accept very high costs. In his seminal *The Decision for Accidents*, Guido Calabresi pointed out that we allow car traffic, despite the fact that, statistically speaking, we know how many Americans will lose life on the road this year.[254] If we *really* believed that human life is priceless, we would simply outlaw cars, and accept the world with lower mobility and higher prices of goods. We are not willing to pay such a price.

If this is the case with matters of life and death, one can easily assume that also in data analytics society, where autonomy, equal treatment and mental well-being of some individuals are at stake, we will choose to pay that price in exchange for the benefits of increased efficiency, new possibilities, and convenience. As of today, these "choices" are made by the markets in the US, and by human rights experts (at least in theory) in the EU. However, they should be up to debate.

Consider price-discrimination. As explained in Part II, ubiquitous data-driven price discrimination can lead to increased efficiency, but also to social segmentation. The choice we face is not whether to allow it or prohibit it, but rather what social costs of allowing it are we willing to tolerate. Imagine, for example, that personalized price discrimination leads to a world where the consumers earning three-digit salaries have to, on average, pay 50% more for the plane tickets than those earning less than 60,000. Whether this should be acceptable or not is a policy decision that we should take as a society; not a random effect of allowing, or applying a general prohibition on, analyzing shopping-history data. We might be willing to accept price discrimination if it traces wealth, or when it traces preferences,[255] but oppose it when leading to social segmentation, even if the price is lower efficiency.

---

254. Guido Calabresi, *The Decision for Accidents: An Approach to Nonfault Allocation of Costs*, 78 HARV. L. REV. 713, 716 (1965).

255. *See* Bar-Gill, *supra* note 3.

Or think about targeted advertising. As a society, we might be willing to accept some level of data-driven behavior manipulation in some consumer markets, but oppose them in the political sphere. Or, we might be willing or accept it in the political sphere, provided that all sides have access to the same technology. Further, when issues like discrimination in ad-delivery are concerned, we might want to prohibit it altogether or try to come up with some mitigating mechanisms. Instead, what we see today, is economists discussing what is more efficient, or human rights lawyers debating what respects human dignity, without giving much say to the people regarding the type of world they would like to live in.

One caveat: I am not, by any means, arguing against the involvement of experts in the law-making process. Participation of people understanding the intricacies of technology, of the human mind, and of the functioning of the market, is necessary both at the problem-positing and solution-implementing stage. However, what I do argue against is leaving the normative choices, the political choices, to the experts only. It should not be the case that what is acceptable in Facebook's or Google's business model will be decided by an economist applying some theory of efficiency, or a human rights lawyers subscribing to some theory of privacy. Those are choices to be left to the people. This is the last "latent effect" identified by the authors of the 1973 report: "questions of record-keeping practice which involve issues of social policy are sometimes treated as if they were nothing more than questions of efficient technique."[256] It is time to re-gain them for politics.

How exactly should we do it? Again, the answer will differ depending on the issue. Certain matters, like the integrity of the political process or non-discrimination, could and should be fought for, and discussed, by various civil society groups representing the interests of the people. Other

256. 1973 REPORT, *supra* note 33, at 23.

matters, like whether to allow business models where the service is "free," but comes with a cost of data collection and usage, should be expressly posed to the people. Become a part of the political process. Part of the political candidates' platforms. Let us see what is it that people want. As of today, the market gives us just the illusion of choice.

3. Direct and Indirect Data Management Law

The last observation I would like to share pertains to the object of potential regulatory intervention. Data management laws can govern either data-practices themselves, like collection, analysis, sharing and usage of personal data; or the data-driven practices themselves, like behavioral advertising, price discrimination or online service provisions. There is a dialectical relationship between the two, where putting limits on data practices can tame the social costs of other practices; while regulating these practices can minimize the collection of data in the first place. Consider two examples.

First, one could imagine legislation largely limiting the possibility of price-discrimination online. For example, it could contain a straightforward prohibition of charging different consumers different prices. Such a law would not address the practices of data collection and usage explicitly but minimize the incentive to collect and analyze data by some companies. If the effects cannot be used to price discriminate, why bother to collect and analyze data in the first place.

On the other hand, it is possible to address the problem of data usage and collection to tackle social costs directly. For example, there could be norm prohibiting using data on consumer behavior to target political communications. As a result, even without passing a law on limits of political speech, an intervention aimed at data practices would bring about changes in the creation and distribution of social costs.

Whether to employ one way of intervening or the other, again, is a matter for political choice.

CONCLUSION

This Article argued for a new way of thinking about the negative effects of data analytics: data management law. Instead of concentrating solely on privacy (or data "protection") I suggested we should adopt a more inclusive category of "social costs of data management." Instead of focusing on individual interest only, we should account for the collective ones as well. And instead of employing one-size-fits-all technocratic solutions, like markets or human rights, we should adopt a more nuanced, issue-by-issue, political approach to setting the limits on data practices. Human rights might delineate the boundaries of what we can allow, but a huge space for political deliberation remains within those boundaries. The question: what society do we want to live in? can be answered in numerous, equally legitimate, ways. Markets might be a useful tool for achieving some of our goals, but they should function within politically agreed upon legal frames.

The history of the "notice and choice" and "personal data protection" approaches teaches us several valuable lessons. The problems we face today are not new. Already in 1973, the authors of the HEW Report foresaw the world characterized by the easy access to data about people; corporate ability to use this data for many purposes, including to influence human behavior; and the risk of imposing the costs of data-driven practices on the most vulnerable members of the society. The development of the legal response to these threats in the United States has been halted by the liberal euphoria of the 1990s, while its European counterpart has been overshadowed by the human rights mindset of "data protection." However, at the dawn of the 2020s, we see how the markets neither accounted for, nor properly distributed, the social costs of data management. We see how human rights cannot be the only normative criterion in debating the limits of data management. We must see that it is time to move on.

Whatever the path we ultimately take, three observations will remain true. First, in the data analytics society, a person disclosing personal information imposes costs on other people, by enabling companies to infer new knowledge about others. Second, these costs can be imposed on people without personally identifying them, thus without violating their "privacy." This is how technology functions and can function. The question is whether this should be permissible. Which leads me to the final observation: there is always a political decision taken. Regardless of whether it occurs explicitly through legislation, or implicitly, by leaving the decision to the market, and allowing corporations to write their own rules, some decision occurs. We know what the rules written by the corporations are. Hence, it is up to us to decide. Will the substance of the data management law result from our political action, or our failure to act?