

Data and Manure: Are Data Subjects Investors?

Tae Wan Kim, * Joohe Lee, ** Joseph Xu *** and Bryan Routledge ****

ABSTRACT

This article explores the status of data subjects in the era of data capitalism. Data-driven companies rely on data subjects offering personal information to train the algorithms they use to provide their services. We maintain that it is time to seriously investigate whether data subjects should be considered as investors. First, we preview our thesis and provide a functional definition of an investor. Then, we develop our argument that data subjects are better understood as investors rather than consumers or labor providers. Finally, we examine the balance sheet impact of a data contribution to the firm and the existing legal regime requiring subjects to retain an ownership interest in their data even after it has been transferred to the firm.

Abstract	65
Introduction	66
I. Are Data Subjects Consumers, Workers, Or Investors?.....	66
II. Defining Investors	70
III. The Balance Sheet Impact	72
IV. The Parallel in the Balance Sheet.....	76
A. Example 1: Amazon.....	78
B. Example 2: Facebook.....	79
C. Example 3: Google	79
V. Is Data Manure?	80
VI. What Differentiates Data Subjects from Consumers?	85
VII. Is This Whole Idea Feasible?.....	86
Conclusion.....	89

DOI: <https://doi.org/10.15779/Z38DR2P915>

*. Associate Professor of Business Ethics, Carnegie Mellon University

** . Assistant Professor of Business Ethics and Law, Pepperdine University

***. Assistant Professor of Operations Management, Carnegie Mellon University

****. Associate Professor of Finance, Carnegie Mellon University

This paper has benefited from seminar participants at Carnegie Mellon’s Privacy Seminar, Finance Brown Meeting, and Georgetown Institute for the Study of Markets and Ethics. We are grateful to Norman Sadeh, Thomas Rucht, Duane Seppi, Brent Glover, Lars-Alexander Kuehn, Pietro Bonaldi, and John Hasnas.

INTRODUCTION

The status of data subjects has evolved in the era of data capitalism.¹ A data subject is a person whose personal data is being collected, held, or processed by data collecting or processing companies (e.g., Amazon, Google, or Facebook).² Such companies train their algorithms using data, including the personal information provided by data subjects.³ These better-trained algorithms lead to general data analytics insights, more efficient sales, and streamlined operations processes which, in turn, increase the company's profitability. It is time to consider such subjects as investors in the firms that use their data. We justify this proposal by providing a functional definition of an investor in the context of our thesis. Then, we develop our main argument that data subjects are better understood as investors, as opposed to simple consumers or labor providers.

I. ARE DATA SUBJECTS CONSUMERS, WORKERS, OR INVESTORS?

How should we characterize the relationship between data subjects and firms that collect and process their data? Our expectations of what is "normal" significantly influence formal institutional arrangements—and thus our lives. Take, for instance, the idea that the shareholders of a corporation own the corporation—and that the managers of the corporation are their agents.⁴ This idea is not the only possible interpretation of the relationship between the corporation, its managers, and its shareholders.⁵ The proliferation of this view has led to

1. See *Fuel of the Future: Data Is Giving Rise to a New Economy*, THE ECONOMIST (May 6, 2017), <https://www.economist.com/briefing/2017/05/06/data-is-giving-rise-to-a-new-economy>; Anita Pratap, *The Era of Data Capitalism*, THE WEEK, Mar. 5, 2017.

2. We follow the definitions in the European Union's General Data Protection Regulation (GDPR), Article 4: "Personal data means any 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." 2016 J.O. (L 119) 33.

3. See Stacy-Ann Elvy, *Paying for Privacy and the Personal Data Economy*, 117 COLUM. L. REV. 1369 (2017); Mara Paun, Book Note, 10 L., INNOVATION & TECH. 153 (2018) (reviewing BRUCE SCHNEIER, *DATA AND GOLIATH: THE HIDDEN BATTLES TO COLLECT YOUR DATA AND CONTROL YOUR WORLD* (2015)).

4. Michael C. Jensen & William H. Meckling, *Theory of the firm: Managerial Behavior, Agency Costs and Ownership Structure*, 3 J. FIN. ECON. 305 (1976); Henry Hansmann & Ugo Mattei, *The Functions of Trust Law: A Comparative Legal and Economic Analysis*, 73 N.Y.U. L. REV. 434, 443 (1998).

5. One prominent example of an alternate interpretation of the relationship between the corporation, its managers, and its shareholders is that managers of the corporation are required to balance the various interests of the corporation's stakeholders. See, e.g., Margaret M. Blair & Lynn A. Stout, *A Team Production Theory of Corporate Law*, 24 J. CORP. L. 751, 752-53 (1998); R. Edward Freedman & David L. Reed, *Stakeholders: A New Perspective on Corporate Governance*, 25 CALIF MANAGEMENT REV. 88, 106 (1983); Oliver Hart & Luigi Zingales, *Serving Shareholders Doesn't Mean Putting Profit Above All Else*, HARV. BUS. REV., Oct. 12, 2017 at 2; Eric W. Orts, *Beyond Shareholders: Interpreting Corporate Constituency Statutes*, 61 GEO. WASH. L. REV. 14 (1992). One reason why shareholders should not be given priority over other stakeholders is that shareholders only own shares in corporations rather than the corporation or its assets. See, e.g., Alan Strudler, *What to Do with Corporate Wealth*, 25 J. POL. PHIL. 108, 108-09 (2017).

DATA AND MANURE

significant institutional changes, including the passage of new regulations that attempt to better align the incentives of a corporation’s managers with those of its shareholders. In the dawning era of data capitalism, what is characterized as the “normal” relationship between data subjects and data firms may have a similar impact.

According to the conventional view, data subjects are consumers who do not have any meaningful rights to their data once it is transferred to the firm. They are seen as having a consumer relationship with the data firm, one in which they allow data firms to collect information about them in exchange for a service. The transaction thus entails a service in exchange for access to information. The information itself is useless and only has value due to the efforts of the data firm. Such treatment of data is reminiscent of an analogous approach to abandoned manure in a well-known nineteenth century case. In *Haslem v. Lockwood*,⁶ the Connecticut Appellate Court held that an entrepreneur has a property claim to the manure abandoned by horses pulling wagons because he enhanced the value of it through the application of his labor.⁷ Under such an approach, data processors such as Google and Facebook are the rightful owners of data because they apply their labor to convert this raw material—some form of unprocessed information about the data subject’s online activity that, like abandoned manure, is useless to data subjects—and thereby greatly enhance its value.

But many are not satisfied with the view that data is like manure. For instance, Eric A. Posner and E. Glen Weyl—a Chicago law professor and an economist at Microsoft—offer a thought-provoking perspective in their book, *Radical Markets*:

The powerhouses of the digital economy, firms like Facebook, Google and Microsoft, exploit the lack of public understanding of AI [artificial intelligence] and ML [machine learning] to collect for free the data we all leave behind in our online interactions. This is the source of the record profits that make them the most valuable companies in the world. Facebook, for example, pays out only 1% of its value each year to workers (programmers) because it gets the rest of its work for free from us! In contrast, Walmart pays out of 40% of its value in wages. People’s role as data producers is not fairly used or properly compensated.⁸

This message attempts to alter the way corporate America understands and treats data subjects by challenging the conventional view. It is a message so powerful that a social movement, known as “Data as Labor” (DaL), has come into being, with over one hundred chapters having been formed around the world within a year.

6. 37 Conn. 500 (1871).

7. CMU’s Block Center for Technology and Society, *Data Subjects and Manure Entrepreneurs: When It Comes to How Your Data Is Being Used to Drive the Technologies of the Future, You Have a Seat at the Table; It’s Just Been Empty*, MEDIUM (Nov. 2019), <https://medium.com/consequential-podcast/data-subjects-and-manure-entrepreneurs-e3bc61cef7eb>.

8. ERIC A. POSNER & E. GLEN WEYL, RADICAL MARKETS: UPROOTING CAPITALISM AND DEMOCRACY FOR A JUST SOCIETY, 209 (2018) (bracketed parentheticals added); see also Imanol Arrieta Ibarra et al., *Should We Treat Data as Labor? Moving Beyond ‘Free,’* AM. ECON. ASS’N. PAPERS & PROC. 1, 38-39 (2017).

DaL gives a real operational option to younger people enthusiastically inquiring about developing an ethical governance structure for the digital economy. We agree with the position that “[m]ost people do not realize the extent to which their labor—as data producers—powers the digital economy”⁹ and that “ROD [return on data]—the relationship between the data price consumers pay and the benefits they receive—is unknown.”¹⁰

Table 1: Comparison of the Three Views

<u>Data subjects'</u>	<u>Conventional view</u>	<u>Data as Labor</u>	<u>Data as Investment</u>
Status	Consumer	Workers	Investors
Activity	Disclose data for services	Sell data as labor	Provide data as capital
Due claim	Good service	Fair compensation	Financial security

However, it is our view that DaL fails to capture a critical dimension of the data subject-processor relationship regarding the equitable nature of the data subject's contribution. It is from this perspective that we develop an approach that will better legitimate claims that subjects can make on firms that collect and process their data.

In this article, we offer a distinct view—which we call “Data as Investment” (DaI)—that argues that most people who offer personal information to firms are better understood as investors. Data subjects are better understood as investors, we will argue, because they meet the *objective* condition necessary to be considered investors, regardless of their *subjective* mental condition. Someone meets the objective condition of being an investor if they are justified in receiving a financial security from a firm in exchange for a resource contribution. A financial investor, for instance, meets the objective conditions necessary to be considered an investor because they explicitly negotiate for a financial security (and are thus justified in receiving it) in exchange for a financial capital contribution to the firm.

Data subjects, we will argue, make a data contribution to the firm and are similarly justified in receiving a financial security in exchange. However, unlike financial investors, who are fully aware of the financial security received in the transaction (with varying levels of due diligence depending on the regulatory framework for the transaction), data subjects typically do not recognize that they meet the objective condition for being an investor. Nevertheless, the lack of the subjective condition is not necessary nor particular to the issue of classifying someone as an investor. The classification of other groups of a firm's stakeholders does not depend on a subjective recognition of such classification.¹¹

9. POSNER & WEYL, *supra* note 8, at 208.

10. Noam Kolt, *Return on Data*, 38 YALE L. & POL'Y. REV. 77, 103 (2020).

11. As a general matter, the existence of contractual and agency relationships between firms and their stakeholders are determined according to objective criteria rather than the subjective intent of the parties.

DATA AND MANURE

For instance, classifying a labor provider as an employee—rather than an independent contractor—is not based on the labor provider’s subjective recognition of their classification, either. Instead, the law clearly classifies labor providers based on tests that look at objective factors.¹² In other words, we—via labor and tax laws—determine whether or not someone is an employee of a firm based on the extent to which an individual is judged to be an employee, not the extent to which the person in question actually believes herself to be an employee. Similarly, we will argue that current accounting standards and data privacy laws provide sufficient justification for the data subject receiving a financial security in exchange for a data contribution to the firm. Therefore, data subjects as investors may deserve several rights, such as control and information rights.¹³ The defining of specific rights for different kinds of investors is a separate question we will not discuss in full here.

Lastly, by arguing that data subjects are better understood as investors, we do not deny that they can be simultaneously consumers, workers, or members of some other stakeholder group. Data subjects as consumers (or workers) might deserve proper consumer (or labor) rights and relevant protections.¹⁴ DaI attempts to move beyond consumerism and labor theory to address problems specific to the data economy by positing that data subjects should have *additional* rights as investors. Neither do we deny that other stakeholder groups might also be considered investors in a broad sense. They may have a special claim to a firm’s assets in ways that parallel the claims of data subjects. Nevertheless, we focus on data subjects because there is no denying that their initial and continuous support (e.g., allowing firms to collect and process personal information) significantly affects firms’ business success. Realistically speaking, those who provide capital have power within a capitalist system. Offering a new paradigm—that data subjects provide important capital to a central mode of production within the data economy—can radically change how society should treat such subjects.

See, e.g., Wayne Barnes, *The Objective Theory of Contracts*, 76 U. CIN. L. REV. 1119, 1119-1120 (2008) (“mutual assent to a contract is determined by reference to external acts and manifestations, not by evidence of subjective, internal intention.”); *A. Gay Jenson Farms Co. v. Cargill, Inc.*, 309 NW 2d 285, 290 (Minn. 1981) (“An agreement may result in the creation of an agency relationship although the parties did not call it an agency and did not intend the legal consequences of the relation to follow.”).

12. See, e.g., *Robicheaux v. Radcliff Material, Inc.*, 697 F.2d 662, 667 (5th Cir. 1983) (“An employee is not permitted to waive employee status. . . . A person’s subjective opinion that he is a businessman rather than an employee does not change his status.”). For more on the application of objective factors to classify workers, see, e.g., Richard R. Carlson, *Why the Law Still Can’t Tell an Employee When It Sees One and How It Ought to Stop Trying*, 22 BERKELEY J. EMP. & LAB. L. 295 (2001), Orly Lobel, *The Four Pillars of Work Law*, 104 MICH. L. REV. 1539 (2006), and Jooho Lee, *The Entrepreneurial Responsibilities Test*, 92 TUL. L. REV. 777 (2018).

13. Julian Velasco, *The Fundamental Rights of the Shareholder*, 40 U.C. DAVIS L. REV. 407, 416-21 (2006).

14. Kirsten Martin & Katie Shilton, *Why Experience Matters to Privacy: How Context-Based Experience Moderates Consumer Privacy Expectations for Mobile Applications*, 67 J. ASSOC. INFO. SCI. & TECH. 1871, 1881 (2016).

First, we will develop our functional definition of an investor as *someone who is justified in receiving a financial security from a firm in exchange for a resource contribution*. Second, we propose to understand data subjects as investors by analyzing the impact their data contribution has on the firm's balance sheet and the type of financial security that they are justified in receiving based on the type of data contributed. Third, we argue that data subjects meet the objective condition of being investors because they retain a property interest in their data even after it is transferred.

II. DEFINING INVESTORS

What do we mean by “investor”? In general, and within the view that a firm is a “nexus for contractors/patrons,” it is hard to distinguish between suppliers, customers, and investors.¹⁵ Employees and suppliers being considered as providers of labor and goods—and banks or venture capital firms as suppliers of capital—are similar. All these groups are usually owed money. There are also others who do not have a contractual relationship with the firm but are nevertheless owed money. Take, for example, someone who wins a legal judgment and is thus “owed money.” Which groups should be considered as investors of the firm?

The current commonsense understanding of investing entails a resource contribution with an expectation of profits. For instance, Warren Buffett defines investing as “forgoing consumption now in order to have the ability to consume more at a later date.”¹⁶ Similarly, the legal definition of an investment contract requires a capital contribution to a common enterprise with a reasonable expectation of profits derived from the entrepreneurial or managerial efforts of others.¹⁷

In contrast, we propose a broad definition. An investor is anyone who is justified in receiving a financial security from the firm in exchange for a resource contribution. Like in the term “investor,” there is also ambiguity in the definition of a “financial security.”¹⁸ Nevertheless, it suffices for our purposes to define a financial security as a bundle of rights to the firm's assets that can be

15. HENRY HANSMANN, *THE OWNERSHIP OF ENTERPRISE* 18-19 (1996).

16. Warren Buffet, *Berkshire's Corporate Performance vs. the S&P 500*, BERKSHIRE HATHAWAY, 17 (Feb. 25, 2012), <http://www.berkshirehathaway.com/letters/2011ltr.pdf>.

17. *Reves v. Ernst & Young*, 494 U.S. 56, 64 (1990).

18. The U.S. Securities and Exchange Commission recently confronted the definition of financial security during the 2017–2018 wave of cryptocurrency “Initial Coin Offerings.” The SEC has jurisdiction over anything that is defined by the Securities Exchange Act of 1934 as a security, including “any note, stock, treasury stock, bond ... investment contract ... or, in general, any instrument commonly known as a ‘security.’” See 15 U.S.C. § 7811l. In the course of ruling that certain coin offerings can constitute a security, the SEC distinguished between “utility tokens,” which are akin to prepaid goods and services such as airline miles or a reward t-shirt offered on a crowdfunding site, and “financial tokens,” which have many properties that look like financial assets (such as stocks or bonds). A similar distinction is helpful here as we argue that the data subject should receive a financial security. See *What Is an ICO?*, BITCOIN MAGAZINE (Aug. 10, 2017), <https://www.nasdaq.com/article/what-is-an-ico-cm830484>.

DATA AND MANURE

characterized as falling within four areas: (1) cash flow, (2) control, (3) information, and (4) resale/liquidity rights.

For individual securities, the characteristics of these areas are shaped by a mix of contracts, regulations, and practices. For example, a stock or equity financial security has cash flow rights that are residual. An equity share is a claim on profits after expenses and other capital providers are paid. Equity has broad control rights over the decisions of the firm. In practice, the exercise of the control rights is via proportional voting for a board of directors.¹⁹ Shareholders receive annual financial statements, including for public companies a host of U.S. Securities and Exchange Commission (SEC) filings. Finally, in most public companies, transferring ownership is a simple and frequent transaction.²⁰

In a similar fashion, we can describe the (typical) features of a debt financial security such as a bond or bank loan. Debt contracts are a schedule of contracted cash flows. Similarly, debt contract control rights are fairly narrow and specific, similar to a covenant that bounds inventory. Conditions of bankruptcy highlight that, like equity, cash flow and control rights are contingent.²¹ Debt contracts typically require periodic information. Finally, debt contracts are often assignable so they can be sold.²²

However, the rights associated with a financial security can also be quite limited. Consider employee stock options, a common benefit provided by many start-ups. Cash flow rights are contingent on the option being exercised and therefore contingent on the value of the underlying equity. The options might also be contingent on a vesting schedule tied to employment. The control rights are also nil until after the options are converted to stock. Information rights are often limited (particularly in a private company).²³ Finally, employee stock options are typically not transferable and cannot be sold (i.e., “locked up”).²⁴ Nevertheless, an employee stock option is a financial security because it gives the option holder a bundle of rights to the firm’s assets; the holder of an employee

19. In large public companies, the ability of shareholders to exercise their control rights—corporate governance—is an active area of research and practice. See Andrei Shleifer & Robert W. Vishny, *A Survey of Corporate Governance*, 52 J. FIN. 737, 737–764 (1997).

20. The annual volume of transactions on the NYSE is trillions of dollars—close to the magnitude of the GDP of the U.S. economy. FXCM GROUP, *New York Stock Exchange (NYSE)*, <https://www.fxcm.com/insights/new-york-stock-exchange-nyse/>.

21. For an overview of corporate bankruptcy types in the United States, see Arturo Bris et al., *The Costs of Bankruptcy: Chapter 7 Liquidation Versus Chapter 11 Reorganization*, 61 J. FIN. 1253 (2006).

22. Collateralized Debt Obligations and Asset Backed Securities are examples of debt contract assignments. See e.g., Joshua Coval et al., *The Economics of Structured Finance*, 23 J. ECON. PERSPS. 3 (2009).

23. Rajesh K. Aggarwal et al., *Strategic IPO Underpricing, Information Momentum, and Lockup Expiration Selling*, 66 J. FIN. ECON. 105 (2002); Bernard S. Black & Ronald J. Gilson, *Venture Capital and the Structure of Capital Markets: Banks Versus Stock Markets*, 47 J. FIN. ECON. 243 (1998); Arnoud W. A. Boot et al., *The Entrepreneur’s Choice Between Private and Public Ownership*, 61 J. FIN. 803 (2006).

24. For a discussion of the “lock-up” of stock options, see Laura Casares Field & Gordon Hanka, *The Expiration of IPO Share Lockups*, 56 J. FIN. 471 (2001).

stock option is an investor because her claim is justified in light of her provision of labor to the firm.

Our suggested definition of an investor has an important advantage over the commonsense definition of an investor as someone who forgoes consumption now in order to have the ability to consume more at a later date.²⁵ The commonsense definition places too much emphasis on the subjective condition of being an investor and, thus, does not allow for the consideration of what justifies one's rights and obligations as an investor. For instance, imagine someone inheriting a modest amount of money and using it to purchase a cheap house as her primary residence in a neglected neighborhood. Then suppose she was fully ignorant of the principles of finance and real estate investment; her intention was simply to eliminate paying rent. Finally, suppose that, a few years later, the government enacts legislation granting a specific tax credit to real estate investors in neglected neighborhoods. If the formal definition of an investor emphasized subjective conditions, she would be barred from claiming the tax credit because she did not meet them. But it would be arbitrary to employ such a definition to bar her claim without considering *why* such a definition *should* be used. A justifiable reason for defining an investor for the purposes of the tax credit should consider relevant reasons—such as the purpose for which the tax credit was created, moral issues related to fairness and justice, etc.

On the other hand, our definition of an investor makes room for such considerations. We will argue that data subjects should be considered investors because their data contributions to the firm have an impact on the firm's balance sheet in a way that justifies their receiving a financial security from the firm in exchange for their data. The value of the data must be reflected in the firm's balance sheet somewhere. And since data subjects have claims to their data—and thus perhaps some portion of the value of their data—that are not extinguished simply because they provide data firms access to such data, the firm should recognize such claims on its balance sheet by offering a financial security to those who make the data contribution to the firm.

III. THE BALANCE SHEET IMPACT

To explain our claim that data subjects are justified in receiving a financial security in exchange for their data, we begin by developing an understanding of the relationship between a firm and its capital investors by examining the balance sheet impact of a capital investment. The balance sheet in this context serves two purposes: it provides a standard foundation for how investment is handled practically, and it provides a robust framework for measuring the impact of economic value generated by an asset. Based on this understanding, we build a case for data subjects as investors by replicating the balance sheet impact of a capital investment.

25. See *supra* note 17.

DATA AND MANURE

We use standard accounting definitions such as assets, liabilities, and equity based on Generally Accepted Accounting Principles²⁶:

- **Assets:** A property with probable *future economic benefits* that is presently obtained or controlled by an entity, as a result of past transactions or events.
- **Liabilities:** A property with probable *future sacrifice of economic benefits*, due to present obligations of an entity, as a result of past transactions or events.
- **Equity:** The residual interest in the assets after deducting liabilities.

In addition to these definitions, we also refer to two rules that form the foundation of the double-entry accounting system, a standard methodology for keeping track of the effects of transactions and events on financial statements. The first is that all transactions must be recorded and reflected in the accounting statement (i.e., balance sheet) of all parties.²⁷ The second, known as the balance sheet equation, states that an individual entity's assets always equal its liabilities plus equity.²⁸

The impact of a capital investment on a firm's and its capital investor's balance sheets develops over time. Capital (cash, physical asset, etc.) is an asset simply because it can be used to generate future economic benefits. When a firm acquires capital from an investor, this transaction must be reflected in both the firm's and the investor's balance sheet (the first general rule). For the firm, the transaction would cause an increase in the asset side of the firm's balance sheet because it has received capital. Assuming no changes to the firm's other assets, this transaction must result in a corresponding increase to the liabilities/equity side of the firm's balance sheet (based on the second general rule, the balance sheet equation). For the investor, the changes take place all on the asset side of their balance sheet, as the investor transferred an asset (capital) to the firm and received another asset (a financial security) of equal value from the firm in exchange.

Consider how a firm's balance sheet changes over two periods— $t=1$ and $t=2$ —under a capital investment transaction. At $t=1$, the investor provides capital to the firm and receives either a debt-like financial security and/or an equity-like financial security. We will assume that the firm also owns some existing assets that are 100 percent equity financed. Between $t=1$ and $t=2$, the firm generates cash flow using its original assets and the additional investment. At $t=2$, the capital investor can lay claim to the firm's assets tied to the financial security received in $t=1$.

To make this scenario more concrete, suppose that at $t=1$ the capital investor provided \$1 million in exchange for a bond maturing in $t=2$ (a debt financial security) and another \$1 million in exchange for common stocks (an equity

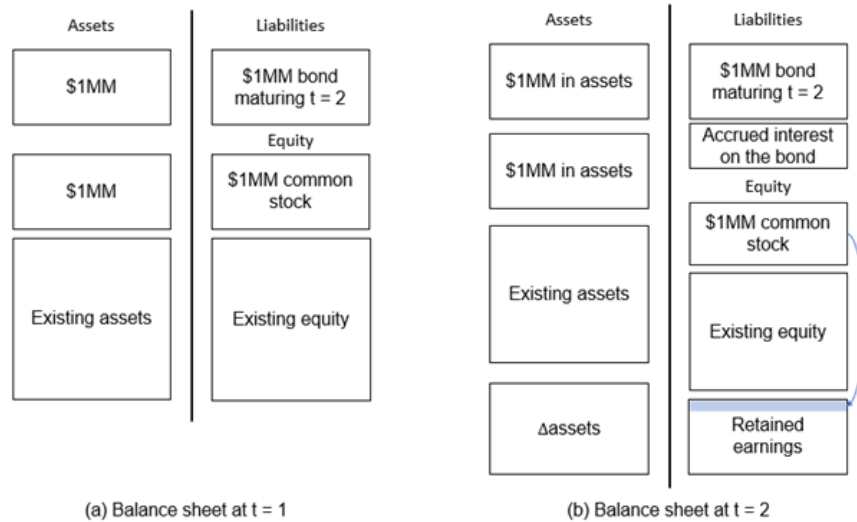
26. FINANCIAL ACCOUNTING STANDARDS BOARD, STATEMENT OF FINANCIAL ACCOUNTING CONCEPTS NO. 6, ELEMENTS OF FINANCIAL STATEMENTS (1985).

27. PETER D. EASTON ET AL., *Module 3 Transactions, Adjustments, and Financial Statements*, in FINANCIAL ACCOUNTING FOR MBAS (7th ed. 2018).

28. *Id.*

financial security) simultaneously. Figure 1(a) illustrates the firm’s balance sheet at t=1. The firm operates between t=1 and t=2, generating some amount of cash flow, which is reflected in t=2 as some change in its total assets (this can be positive, zero, or negative). Over the same time interval, the bond accrues interest. Figure 1(b) illustrates the firm’s balance sheet at t=2.

Figure 1: The Balance Sheet Impact of Capital Investments



We can clearly identify the assets that the capital investor can lay claim to at t=2. The capital investor in this example first claims the assets tied to the \$1 million that corresponds with the bond principal and the assets tied to the accrued interest. In addition, the capital investor in this example can then claim the assets tied to the \$1 million common stock investment at t=1 (and the corresponding share of retained earnings):

$$(\Delta\text{assets} - \text{accrued interest}) \times \frac{\$1\text{MM common stock}}{\$1\text{MM common stock} + \text{existing equity}}$$

The maturity of the bond results in an asset outflow for the firm that was already scheduled at t=1. The common stock investment does not result in a scheduled asset outflow. However, the firm surrenders a portion of the change in the residual asset value (after servicing the debt obligation) to the equity investor.

DATA AND MANURE

Generalizing from this example, we can characterize debt-like and equity-like financial securities that are involved in capital investment transactions. A debt-like financial security is defined primarily by a scheduled asset outflow to the capital investor, including some guarantee on the principal. This represents a probable future sacrifice of economic benefits for the firm and is therefore classified as a liability on the balance sheet. Products such as bank deposits—or utility tokens such as travel mileage—also fall under the category of a debt-like financial security and are treated as liabilities. On the other hand, an equity-like financial security can be defined by its lack of debt-like property (equity = assets - liabilities). An equity-like financial security does not involve a scheduled asset outflow to the financial investor, and there is no guarantee on the principal. A common stock would fall under this category.

We have defined the relationship between a firm and its capital investors with the balance sheet impact as the basis. A firm receives an asset from the capital investor and offers a financial security. This is necessary because the balance sheet requires a liability or equity entry to balance the asset increase.

The relationship between a firm and its data subjects would work in exactly the same fashion. An asset is defined as a resource that can generate future economic benefits; data is such an asset because it can be used to generate such an outcome. Due to the first general rule,²⁹ when a firm acquires data from a subject the transaction must be reflected in both the firm's and the subject's balance sheet. In addition, based on the second general rule (balance sheet equation) and assuming no changes to the firm's other assets, the transaction must result in a corresponding increase on the liabilities/equity side of its balance sheet.

If a firm capable of using data for future economic benefits is required to classify data as an asset upon receiving it from the subject, then the firm needs to balance its financial statement through a corresponding entry in its liability or equity. If the firm classifies this as a liability, then the subject should receive a debt-like financial security in exchange for the data. If the firm classifies this as an equity, then the subject should receive an equity-like financial security in exchange.

In the next sections, we show that as long as we accept the notion that data is an asset, it must follow that the subject is entitled to receive a financial security of some form in exchange for their data. In doing so, we further clarify why we believe that it is more appropriate to treat data as an asset that should appear on a company's balance sheet as opposed to a simple exchange of goods and services.

29. See Financial Accounting Standards Board, *supra* note 26.

IV. THE PARALLEL IN THE BALANCE SHEET

We propose that the nature of the financial security (debt-like or equity-like) that the subject receives depends on whether the data transferred is personally identifiable (identity data) or unidentifiable (non-identity data).³⁰ To clarify, identity data is any data that can be associated with an individual subject. Non-identity data is economically valuable data that does not have to be associated with an individual subject.³¹ For example, suppose that John Doe, who lives in Palo Alto, opened an Amazon account and ordered a copy of *Le Morte d'Arthur* on February 8, 2019, at 7:00 p.m. In this case, his identity and address would be considered identity data, whereas the fact that someone in Palo Alto ordered a copy of *Le Morte d'Arthur* on February 8, 2019, at 7:00 p.m. is non-identity data.

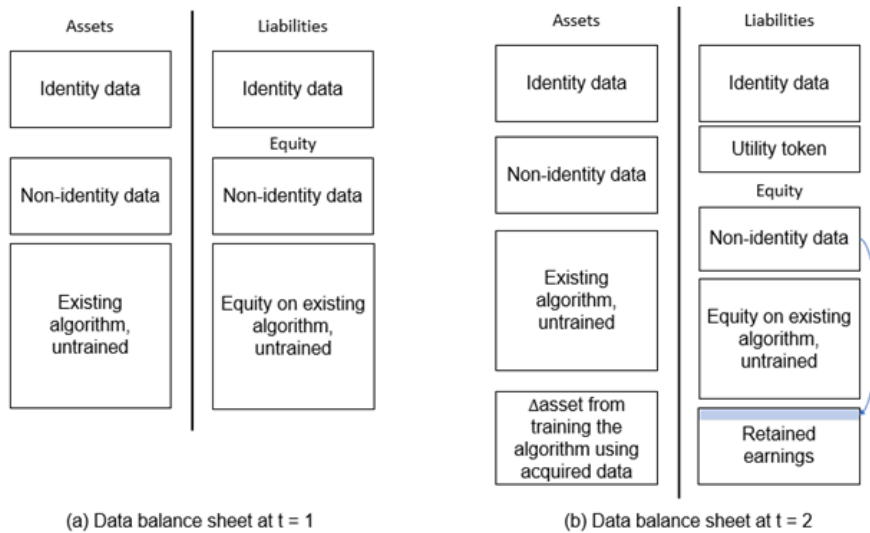
In the world of data capitalism (or the “analytics age”), data subjects provide firms with not only their identity data but also a constant stream of valuable non-identity data.³² We maintain that subjects are entitled to a debt-like financial security in exchange for their identity data and an equity-like financial security in exchange for their non-identity data, given the parallels that can be drawn between the contribution of data subjects to a firm’s balance sheet and that of financial investors.

30. We make a distinction between identity data and non-identity data to resolve the question on the balance sheet entry that is necessary to offset the addition of data as an asset. Based on the rules of double-entry accounting, there needs to be an offsetting entry on the liability, the equity, or a combination of both (assuming no change in other assets). Based on the nature of different kinds of data and how companies use different data, we propose the distinction of identity data and non-identity data and further create a parallel between these types of data and different types of financial securities (debt-like or equity-like).

31. It is a thorny question to specify the boundaries of “personally identifiable data.” To answer it, we would need to join the contested debates about the legitimate scope of the “privacy zone,” which would be beyond the scope of this article. It suffices to say that companies are interested in collecting and processing a variety of identity data that are within the scope of the privacy zone. See Herman T. Tavani, *Informational Privacy: Concepts, Theories, and Controversies*, in *THE HANDBOOK OF INFORMATION AND COMPUTER ETHICS* 131, 141-154 (Kenneth E. Himma & Herman T. Tavani eds., 2008).

32. Examples of these non-identity data include interactions with the technology firm’s websites or clickstream data. Clickstream data summarizes a user’s online activity through sequence of information such as the websites visited, links clicked, and the amount of time spent on each website. See Alan L. Montgomery et al., *Modeling Online Browsing and Path Analysis Using Clickstream Data*, 23 *MARKETING SCI.* 579, 580-581 (2004).

Figure 2: The Data Balance Sheet Impact of Data Investments



In making the parallel between financial investors and data subjects, we consider how a firm’s data balance sheet changes over two periods: $t=1$ and $t=2$. At $t=1$, the data subject can provide personal data to the firm and receive a debt-like financial security and/or provide nonpersonal data to the firm and receive an equity-like financial security in exchange. We will assume that the firm also owns some existing algorithm that is 100 percent equity financed and is not yet trained with actual data. Between $t=1$ and $t=2$, the firm increases the value of its algorithm by training it on the new data acquired from the data subject. At $t=2$, the data subject can lay claim to the firm’s assets tied to the financial security received in $t=1$. Figure 2(a) and 2(b) illustrate the firm’s data balance sheet at $t=1$ and $t=2$, respectively.

At $t=2$, the data subject can first lay claim to the debt principal and accrued interest. The principal is the original identity data, and accrued interest is the utility token (e.g., access to a product or service). We assign a utility token to the role of accrued interest because interest is a liability and so are utility tokens. Other assets can also be used as an interest payment on the identity data, but utility tokens seem to be used most often in practice. At the same time, the non-identity data can be used to lay claim to a portion of the increase in the total asset value if the data is used to generate business analytic insights or to improve the performance of existing algorithms. In Figure 2, we consider the case where the firm improves the economic value of its machine learning algorithm by training

it on the subject's data.³³ This results in a financial increase in the overall asset where the subject can claim the assets tied to the non-identity data and the corresponding share of data retained earnings, which is:

$$(\Delta \text{assets} - \text{utility token}) \times \frac{\text{non identity data}}{\text{non identity data} + \text{existing algorithm}}$$

Note that, similar to the case of capital investment, there is a potential scheduled outflow of data assets in terms of identity data or utility tokens. These items are therefore classified as liabilities, but are best treated like a bank deposit (and accrued interest on those deposits) rather than as a bond. After liabilities are subtracted, any remaining assets correspond to equities by the accounting definition. The non-identity data provided by the subject represents a claim on a portion of the residual asset value of the firm.

Here, the financial security considered in the case of the data investor is both debt-like and equity-like, but not necessarily a full debt or equity. In the case of an equity-like financial security, we can imagine a financial security that is similar to an employee stock option where the cash flow right is contingent on certain conditions—such as exercise or a vesting period. How much cash flow, control, information, and resale rights should belong to the data subject are important questions for the practical implementation of such a system that we wish to tackle in the future. However, we emphasize at this point that unless we have a good reason to accept a double standard, subjects should be entitled to receive some form of financial security in exchange for their data in accordance with data balance sheet accounting.

We conclude this section with three examples based on the ways major technology companies acquire and process data—and how these activities should correspond to an exchange of debt-like versus equity-like financial securities.

A. Example 1: Amazon

Amazon acquires data on a customer's name, address, and product order. In exchange for this mixture of identity and non-identity data, it is obligated to provide fulfillment services, and if requested by the customer at some later point in time, must delete the identity data (the customer's name and address) from its database. Thus, Amazon should issue a debt-like financial security (e.g., the data principal and the utility token for Amazon's fulfillment network) in exchange for the customer's identity data that was acquired when the customer opened an Amazon account and placed an order.

33. Machine learning defines a class of algorithms that exhibit better performance with more experience, such as being trained on a larger dataset. TOM MITCHELL, MACHINE LEARNING (1997). Therefore, the performance of machine learning algorithms improves when given more data by definition.

DATA AND MANURE

At the same time, Amazon acquires non-identity data: the customer's location, time of order, and product ordered. This information is economically valuable to Amazon—even if the data is ultimately anonymized—because Amazon can use its algorithms to generate economic value by streamlining its supply chain in the future (e.g., lowering logistics costs by achieving better demand forecast and inventory placement within its fulfillment network).³⁴ Thus, Amazon should issue an equity-like financial security in exchange for the subject's economically valuable non-identity data.

B. Example 2: Facebook

Facebook acquires personal identity data when a user creates an account. Once Facebook acquires this data, it is obligated to provide the user with access to its social network platform. It is also obligated to maintain the personal identity data securely and remove it upon the user's request, which represents a liability to Facebook. Thus, Facebook should issue a debt-like financial security (e.g., the principal and the utility token for the Facebook platform) in exchange for the personal identity data required to open a Facebook account.

At the same time, Facebook acquires non-identity data through the user's interactions with the platform. This information is valuable to Facebook even if the data is completely anonymized because Facebook can use it to generate economic value by improving its algorithms (e.g., improving its News Feed algorithm) or selling data analytics to advertisers.³⁵ Thus, Facebook should issue an equity-like financial security in exchange for the subject's economically valuable non-identity data.

C. Example 3: Google

Suppose that an individual submits a search query to Google on a public computer. Google acquires data on the search query, but there is no identity data. Google is technically not obligated to respond to this search query by returning links that best match the keywords, nor is it required to maintain confidentiality on the specific keywords submitted. Therefore, Google does not need to issue a debt-like financial security in exchange for this search query data.

However, Google should still issue an equity-like financial security in exchange for this search query data. Even if it is fully anonymized, the information on the keywords used in the search query is economically valuable non-identity data. For example, the company can use the data on the popularity

34. Jason Acimovic & Stephen C. Graves, *Making Better Fulfillment Decisions on the Fly in an Online Retail Environment*, 17 MFG & SERV. OPERATIONS MGMT., Winter 2015, at 34-51. (demonstrating how anonymized past-order data can be used to improve fulfillment decisions in e-commerce to lower operational costs).

35. Dokyun Lee et al., *Advertising Content & Consumer Engagement on Social Media: Evidence from Facebook*, 64 MGMT. SCI. 4967, 5015-31 (2018) (demonstrating how anonymized Facebook usage data can be used to generate insights on how to improve social media engagement).

of specific search queries to optimize pricing on its Ad Exchange platform.³⁶ Therefore, the individual submitting the search query to Google is entitled to receive an equity-like financial security, even if that entitlement is ultimately forfeited. We note that this arrangement is not entirely theoretical because it is feasible (and straightforward given the available software toolkits) for Google to construct a search tool that works in such a way. In practice, Google can and does track the IP address of all search requests that are being submitted to its servers to monetize its search queries.³⁷ Google and its users can manage claims to the equity-like financial security either on a centralized system (as is done with many standard payment applications such as PayPal³⁸) if log-in is enforced, or using the same private-public key pair cryptography (for example, on the Ethereum public blockchain like the Brave web browser³⁹) without having to log in or identify each other for every search.

V. IS DATA MANURE?

A crucial premise in our proposal is that subjects retain some property interest in their data even after they contribute it to a firm. Consider, for example, when someone hands a barista a \$20 bill for a \$2 cup of coffee. If it were the case that the act of receiving a cup of coffee from the barista extinguished the customer's property interest in the \$20 bill, the customer would not be justified in receiving \$18 change. They would only be justified in receiving change if they retained some property interest in the original \$20 bill. In the same way, the subject is justified in receiving a financial security from the data firm only if they retain some property interest in their data—even after it is transferred. Otherwise, the balance sheet argument might simply imply that booking the data as an asset ought to increase the equity stake for the firm's shareholders.

For the sake of argument, we are willing to grant that a subject has intentionally transferred her personal data to the firm to collect, hold, and/or process. Even so, determining whether or not a subject retains some property interest in their data upon transfer cannot be settled by pointing to mere circumstances such as a change in possession, an exchange of goods/services, or subjective intentions of at least one of the parties. As the above example

36. Negin Golrezaei et al., *Boosted Second Price Auctions: Revenue Optimization for Heterogeneous Bidders* (2017) (unpublished manuscript) (on file at SSRN), https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=3016465 (illustrating how Google uses anonymized search keyword auction data to improve its ad auction process, leading to a 3 to 6 percent improvement in its ad exchange revenue).

37. Identification of the IP address of the request origin is a requirement in the Hypertext Transfer Protocol (HTTP) used to communicate over the Internet. Google simply needs to keep a log of these inbound requests, so tracking is technically quite trivial. See Roy T. Fielding et al., *Hypertext Transfer Protocol – HTTP/1.1* (June 1999), <https://www.rfc-editor.org/rfc/pdf/rfc2616.txt.pdf>.

38. Computerized Payment System for Purchasing Goods and Services on the Internet, U.S. Patent No. 5,757,917 (filed Nov. 1, 1995) (issued May 26, 1998).

39. Brave Software, *Basic Attention Token (BAT): Blockchain Based Digital Advertising* (March 13, 2018), <https://basicattentiontoken.org/BasicAttentionTokenWhitePaper-4.pdf>.

DATA AND MANURE

illustrates, the mere fact that the money is now in the possession of the barista does not extinguish the customer's property interest in the money transferred. Neither would it be sufficient to claim that the customer has given up the entirety of her property interest merely because they received a cup of coffee—or because the barista honestly thought that the customer had intended the remainder of the cash to constitute a rather generous tip. Instead, the determinative issue is whether the customer has a property interest in the money. Although social custom and intentions will certainly play a role in determining the scope of one's property interest, the focus must be on the property interest itself. Similarly, the issue that requires attention is whether data subjects actually have some property interest in their personal data even after they transfer it.

There is much legal and philosophical controversy over what the concept of property actually entails. “[E]veryone knows what [property] is, but no one can define it.”⁴⁰ We begin with a standard law school textbook case: *Haslem v. Lockwood* (1871). In the United States of the late nineteenth century, horse-drawn wagons were a primary mode of transportation. As a result, manure regularly accumulated on roadways. Thomas Haslem saw an advantageous opportunity and ordered his servants to gather it into piles. After gathering the piles, the servants did not have a cart with which to move the manure, so they left the piles overnight—planning to come back with a cart the next morning. However, before Haslem and his servants returned, William Lockwood carted the piles away. Litigation ensued and, eventually, the Connecticut Supreme Court had to decide who owned the piles. The Court concluded that Haslem owned the manure, relying on the standard Lockean theory of property: it is created by the *fruits of one's labor* when individuals mix their labor with a portion of the earth itself. The court wrote, “after the plaintiff [Haslem] had changed [the manure's] original condition and greatly enhanced its value by his labor, [Lockwood] seized and appropriated to his own use the fruits of the plaintiff's outlay.” The piles of horse manure in the case are analogous to big data.

Data processors (e.g., Google, Facebook) are tempted to make a parallel argument using the Lockean idea. Data as raw material—some form of unprocessed record of the data subject's online activity—is like manure abandoned on a public road. Data subjects give off data to be collected by the data firm in an analogous way to horses giving off manure to be collected by others. Raw data *per se* is useless to data subjects, so they do not care about it. Useful data is the fruit of the processor's labor because its value is greatly enhanced. Therefore, the processed data is the processors' property.

However, there are significant differences between data and manure. No one claimed to own the manure, so no one objected when Haslem wanted to use it.

40. Abraham Bell & Gideon Parchomovsky, *A Theory of Property*, 90 CORNELL L. REV. 531, 533 (2004).

In contrast, firms should—and do via rarely-read user agreements—get data subjects’ permission via informed consent. Firms should and do inform data subjects of the ways they will use their personal information through online consent forms. Firms obtain subjects’ permission through consent forms before putting significant efforts into processing that data. The outcome of this—the processed data—is more than a mere collection of personal data, and it is often technically difficult to pinpoint which part of the final product is a particular data subject’s contribution. Personal data has been transformed into a corporate asset. Nevertheless, such a transformation is not the sole determinant of whether data subjects continue to have a proprietary interest in their data. The actual content of the rights transferred via user agreements and the mixture of mandatory rules imposed by the government also play a role, as will be discussed further below.

The Lockean theory purports that property refers to “exclusive control over something—to be able to use it as one wishes, to sell it, give it away, leave it idle, or destroy it.”⁴¹ But careful scrutiny quickly unravels this understanding of property. For instance, consider someone who sets up a trust, commits to funding it with his own money according to a specified schedule, and then conveys the trust to a trustee for the benefit of a charitable organization. Who is the owner of the trust, let alone of the money committed to the trust? In such an instance, it can be helpful to think of property not as a unitary concept but rather a bundle of rights that can vary according to context.⁴² The rights within the bundle can be divided and apportioned in a myriad of ways.⁴³ In instances like a trust, it can be more helpful to think about the property interests that each person has within it—the “particular limited rights each of the co-owners has with respect to the thing”⁴⁴—rather than attempting to identify the “true” owner. Given that our argument centers around the conditions under which the data subject is entitled to receive a financial security, we will refer to the idea of “property” and “ownership” interchangeably, referring to a set of cash flow, control, information, and liquidity rights that individuals possess with respect to their data.

Given the definition of property above, we contend that—as a matter of property rights—the relationship between data subjects, data firms, and personal data is more akin to a trust than a typical commercial transaction. A simple deal involving the exchange of coffee for money might involve both parties alienating the entirety of their property claim to the items being exchanged. Once I give my

41. Thomas C. Grey, *The Disintegration of Property*, 22 NOMOS 69, 69 (1980).

42. See A. M. Honore, *Ownership*, in OXFORD ESSAYS IN JURISPRUDENCE (1961).

43. Considerable controversy concerns whether or not there is an identifiable bundle of rights that constitutes the idea of property or ownership. Whereas some argue that some stable bundle of rights underlies the idea of property (see, e.g., Grey, *supra* note 41), others argue that the very idea of property is meaningless (see, e.g., Honore, *supra* note 42). Our argument does not depend on the resolution of such a debate because there is no controversy about the fact that there are multiple rights associated with a common sense understanding of property and that these rights can be divided and apportioned in a myriad of ways.

44. See Grey, *supra* note 41, at 69.

DATA AND MANURE

money to the barista, I might transfer to the corporation for whom he works the entirety of my property claim to the agreed-upon price of the coffee; the corporation might similarly transfer to me the entirety of its property claim to the cup of coffee. On the other hand, we contend that it is more accurate to view data subjects as alienating only certain aspects of their data property rights—while retaining other aspects—rather than to view them as alienating the entirety of their property interests in their personal data.

As a descriptive matter, modern property rights are primarily determined by law. And although legal regimes are still attempting to catch up to the fact that data is becoming increasingly more valuable, a significant body of laws already grants data subjects property rights—even after their consent to transfer their data to processing firms. The right to receive cash flow is largely a matter of contract and, as a result, subjects already have the right to demand financial compensation in exchange for transferring their data to firms (even if they do not exercise such a right). And the subjects' right to information about how the firm is using data—and how it will be used in the future—is already protected by most modern privacy regimes in the world.

Furthermore, newer legal regimes go further in protecting other property rights of data subjects. Take, as an example, the “right to be forgotten” that is part of the European Union’s General Data Protection Regulation (GDPR),⁴⁵ which implies that users have the right to control their data even after the user transfers it to a data firm. Or the California Consumer Privacy Act (CCPA),⁴⁶ which grants data subjects the right to opt-out of the sale of their personal information to third parties. This right implies that subjects have some liquidity right in their data even after they transfer it to a data firm, since it gives subjects the right to allow—or not—a firm to sell their data to third parties. Although these regulations are too recent to reliably represent case law—and therefore predict how exceptions and contractual provisions might limit such rights—there is no doubt that these laws recognize data subjects as having at least *some* property rights to their data *after* they transfer it to data firms.

One might argue that such rights are not property rights. For some, a property right to something refers to the right to dictate the terms of its transfer.⁴⁷

45. Article 17 no. 1 of the 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).

46. CAL. CIV. CODE § 1798.120.

47. There is a lively debate about the appropriateness of categorizing one’s rights to information about one’s self as a property or a privacy right. Those who are in favor of categorizing it as a property right see the ability to sell the rights to one’s information as a major advantage. *See, e.g.*, Richard S. Murphy, *Property Rights in Personal Information: An Economic Defense of Privacy*, 84 GEO. L.J. 2381 (1996); Lawrence Lessig, *The Architecture of Privacy*, 1 VAND. J. ENT. L. & PRAC. 56, 63-65 (1999); Pamela Samuelson, *Privacy as Intellectual Property*, 52 STAN. L. REV. 1125 (2000). Those who are more wary of propertization point to various inadequacies of the market mechanism to adequately protect one’s interests in information about one’s self. *See, e.g.*, Jessica Litman, *Information Privacy/Information Property*, 52 STAN. L. REV. 1283 (2000); Daniel J. Solove, *Privacy and Power: Computer Databases and*

According to such a definition, the data subject would not have any property rights in their data at all because the subject has no ability to fully transfer their right to their data by waiving their right to be forgotten or to opt out of the sale of their information. But such a disagreement would amount to nothing more than a semantic disagreement. What matters for our argument is whether a data subject is justified in receiving a financial security in exchange for their data contribution. And if the law mandates that the data subject retain certain rights to their data, the data firm must recognize those rights on its balance sheet, regardless of what those rights are called. If one does not—or should not—have any property interest in one’s data, data users might be required to grant firms access to their data while retaining all relevant rights that they might have over it. In such a world, data subjects would be operating within a paradigm in which they act as lenders or depositors to the data firm. In other words, resisting the definition of one’s right to one’s data as a property right would require granting data subjects an even stronger claim to a financial security from the data firm, consisting of robust control and information rights over their property along with a total claim to all potential liquidity/resale rights.

Furthermore, regardless of how contracts between subjects and firms might limit the legal rights that are protected by newer data privacy legislation, there is nothing in online consent forms or privacy policies that suggest that data firms are granted *complete* rights to user data on their platforms. If that were so, a company like Facebook could exclude users from accessing their own personal information, which is currently not the case. Instead, online consent forms and data firm privacy policies typically specify the rights that firms have to their users’ data. Even when such rights are quite broad, they do not give *exclusive* rights. Thus, the rights that are not specified within such agreements are retained by users, regardless of how one characterizes them. As a result, it is descriptively *more* accurate to characterize the typical data transaction as one in which data subjects *allow* firms to access and use *their* personal data—as a patient allows a surgeon to touch their body during surgery—than it is to characterize it as a transaction in which the subject *sells* their data to the firm. Just as it would not make any sense to claim that a patient consenting to surgery somehow transfers the full set of rights they have over their own body to the surgeon, it does not make sense to say that a data subject transfers their full set of rights to the data firm. Whether such rights are characterized as property or privacy rights does not change the fact that subjects retain certain rights associated with the control and liquidity of their data—even while it is in the possession of a data firm.

Of course, we are not claiming that subjects retain *all* of their rights to their personal data nor that firms do not gain *any* rights to user data. These firms put significant efforts into data processing; a significant amount of usable data is the

Metaphors for Information Privacy, 53 STAN. L. REV. 1393 (2001); and Carole Pateman, *Self-Ownership and Property in the Person: Democratization and a Tale of Two Concepts*, 10 J. POL. PHIL. 20 (2002).

DATA AND MANURE

result of the firm's observation and collating of the data rather than any simple transfer between the subject and the firm. In most cases, the value of a user's data is the result of an amalgamation of the efforts of the firm and subject (who either provides information to the firm or consents to being observed). As a result, after the data is processed, it is often technically difficult to pinpoint which part of the firm's data is the direct contribution of a particular data subject. However, regardless of what may be the actual extent of the user's rights to their data, it is clear that they have *some* rights to it. Even if the data firms' efforts were significantly larger than the data subjects' contributions, subjects would still be justified in receiving a financial security from the firm for reasons similar to why the holder of a single share of company stock still has a financial security in the firm.

VI. WHAT DIFFERENTIATES DATA SUBJECTS FROM CONSUMERS?

The discussion thus far may seem to suggest that there is no meaningful difference between viewing data subjects as consumers and viewing them as investors. Subjects, one might argue, are not any different from other providers of capital or labor inputs for a firm—including consumers who transfer capital to the firm in exchange for the goods or services. From this perspective, whether the data subject is justified in receiving a financial security from the firm is simply a matter of bargaining within the market. Just as a labor provider or a supplier might be able to negotiate an equity stake as part of their contractual arrangement, data subjects are able to do the same. The fact that they do not currently receive a financial security from the firm may merely be a function of subjects being willing to transfer their data in exchange for access to the various services provided by the firm. From this perspective, our position—that data subjects should be considered investors—would be akin to arguing that a coffee shop customer should be considered an investor because, given sufficient bargaining power, she would be justified in receiving a financial security from the shop.

To a certain extent, we agree with such a perspective. The very idea of a financial security is flexible, given that it encompasses a wide range of potential cash flow, control, information, and liquidity rights. As a result, in addition to the fact that various patrons of the firm might be able to bargain for an equity stake in the firm, even a relatively unimportant patron might also have a financial security in the firm. The simple right of being owed a small amount of money in exchange for services is sufficient to qualify as a financial security—and thus qualify the service provider as an investor.

However, there are important differences between various classes of patrons that arise from the various mixtures of default and mandatory rules associated

with their contractual relationship.⁴⁸ Various providers of inputs to a firm can be distinguished by the differences in the default and mandatory rules that govern their relationship. Ordinary providers of goods, for instance, are governed mostly by default rules that allow for a significant amount of flexibility in their contractual arrangements. A contract that involves the exchange of a good for a financial security, for instance, might impose a default rule in which the provider of the good transfers all of their property interests to the recipient in exchange for compensation. Providers of labor, on the other hand, are governed by a more robust set of mandatory rules governing working conditions, collective bargaining, etc. An employment contract, for instance, cannot waive the protections of a federal law requiring the employer to provide a safe working environment.

Our contention is that the rights of data subjects to their data are already governed by a sufficient set of mandatory rules that distinguish them from consumers. For example, firms with publicly traded securities in the United States are governed under a variety of mandatory rules that require financial disclosure to investors. There is a parallel for data subjects. Various legal and regulatory regimes require firms in possession of user data to, upon request, disclose information on the types of information they possess and how they use it. On the other hand, the typical consumer who transfers some money to the firm in exchange for a good or a service has no mandatory or default right to receive information about how the money is then used by the firm. Once the exchange occurs, the money belongs entirely to the firm. A consumer who has an interest in knowing how a fried chicken sandwich chain spends its proceeds can only attempt to convince the firm to voluntarily disclose relevant information. Given such distinctions, there is sufficient basis to distinguish data subjects from consumers. In fact, the robust set of mandatory rules that govern data subjects makes them more like shareholders and other providers of financial capital than ordinary suppliers and other patrons.

VII. IS THIS WHOLE IDEA FEASIBLE?

A fully developed discussion of the feasibility of our proposal is beyond the capacity of this paper. We recognize that, practically speaking, instituting our proposal will require overcoming several technical and regulatory hurdles. But

48. Default rules govern the rights and obligations of private parties—unless they contract for an alternate arrangement, whereas mandatory rules do not allow for an alternate arrangement. For instance, Delaware imposes a fiduciary duty of care and loyalty on the directors of a corporation. However, courts in the state are also increasingly allowing corporations and their directors to opt-out of liability for the breach of such duties through contractual arrangements. As a result, the rule that governs the fiduciary duty of corporate directors in Delaware is a default rule, with the fiduciary duty imposed by the law limited if the directors and corporations contract for an alternate arrangement. On the other hand, California courts have ruled that corporate directors' fiduciary duties do not permit parties to arrange for a waiver of such duties. As a result, the rule that governs the fiduciary duty of corporate directors in California is a mandatory rule.

DATA AND MANURE

these challenges are not insurmountable. Take, for example, the digital sale of stocks, which has made buying and selling them much easier. Or consider the development and implementation of various sustainability indices (e.g., the Dow Jones Sustainability Index). Such efforts faced similar regulatory and technological challenges; these were eventually overcome by legislators and businesses. There is no reason to think that recognizing the investor status of data subjects would be any different.

Implementing an equity-like claim for investors will be complicated for a data-intensive company like Facebook or Amazon. Even if the aggregate value of the data-equity claim is small, it would still involve millions (or billions) of individual claims. Fortunately, these challenges do not seem insurmountable because the tech challenges are very similar to managing massive amounts of user data. Adding a small data-equity claim to a user's existing Facebook profile may be cumbersome, but it is hardly daunting for a company commanding such resources. In general, managing such a user database—tracking our data-equity claims—is analogous to implementing a program like airline reward miles.

Perhaps a more nettlesome issue is maintaining anonymity and privacy. Without user accounts, how would data firms record and account for data as it was collected? Forcing companies to collect even more data would certainly be an unintended and unwanted consequence. While this problem is also challenging, we are confident that there are clever solutions. Technology like the Do Not Track (DNT) browser setting or Basic Attention Tokens⁴⁹ (used on the Ethereum blockchain platform) hint at solutions that might enable the recording of users' claims to a data equity security in a way that is firewalled from the company's other data collection. We leave that discussion to a future paper exploring this cryptography challenge.

A new financial security will obviously pose challenges to existing financial oversight as well. But regulators are not unfamiliar with adapting to new technologies. For example, the U.S. Securities and Exchange Commission (SEC) has adopted regulations for new “cryptocurrencies.”⁵⁰ The Financial Accounting Standards Board (FASB) maintains an Emerging Issues Task Force to address amending reporting requirements due to changing technologies.⁵¹ Similarly, the Jumpstart Our Business Startups Act of 2012 adapted regulations to facilitate

49. Phillipa Gill et al., *Best paper—Follow the money: understanding economics of online aggregation and advertising*, in PROCEEDINGS OF THE 2013 CONFERENCE ON INTERNET MEASUREMENT 141 (2013); Brave Software, *Basic Attention Token (BAT): Blockchain Based Digital Advertising* [White Paper] (Mar. 13, 2018), <https://basicattentiontoken.org/wp-content/uploads/2017/05/BasicAttentionTokenWhitePaper-4.pdf>.

50. As an example, see Public Statement from Jay Clayton, Chairman, Securities and Exchange Commission, *Statement on Cryptocurrencies and Initial Coin Offerings* (Dec. 11, 2017), <https://www.sec.gov/news/public-statement/statement-clayton-2017-12-11>.

51. See FINANCIAL ACCOUNTING STANDARDS BOARD, <https://www.fasb.org/jsp/FASB/Page/LandingPage&cid=1176164637724>.

“crowdfunding” technology that facilitates micro-investments in new companies (e.g., Kickstarter, Lending Loop).⁵²

Another challenge is how the new data-equity claims might be traded—if they are to be subject to trade at all. But even here, experience suggests that financial market innovation will be up to the task. Note the remarkable decline to near zero in the trading costs (brokerage fees) for trading stock. Alternatively, consider new assets like Bitcoin and the now huge ecosystem of—perhaps far from perfect—exchanges that have grown. A growing body of research is attempting to develop a technically feasible way to monetize and trade data.⁵³

Some investor rights can be granted without trading or monetizing data. Information and control rights can be granted without monetization at all. Through annual meetings shareholders enjoy the information right. A board has a duty to explain how the company has used the shareholders’ investment. Likewise, if our argument is plausible, data investors should have a right to understand how their data investment is used by a company. Shareholders’ right to control is, basically, a voting right in practice; in real life, shareholders’ right to control is more like a right to influence. Likewise, we can imagine a data investors’ annual meeting where subjects could exercise their voting power on how to govern data management and, by doing so, influence the board’s decision about how to use their data investment.

Does this mean that data subjects are corporate owners? Investors may simultaneously be owners, but not always. In one view, ownership—functionally defined as the sum of rights to control and to residual earnings—can be given to any patron group in theory. In practice, it is given to shareholders because doing so minimizes transaction costs in various ways.⁵⁴ In conventional business organizations, shareholders are investors—and owners—in the so-called “nexus-of-contracts” paradigm. But consider venture capital, which perfectly meets the conditions necessary to be an investor. Venture capital’s rights are contingent upon the investee’s performance. They may obtain, for instance, board control if the firm’s performance falls below the agreed-upon threshold. But if the firm’s

52. See Douglas Cumming & Sofia Johan, *Demand-driven Securities Regulation: Evidence from Crowdfunding*, 15 VENTURE CAPITAL 361 (2013). More generally, for an overview see Susan Chaplinsky et al., *The JOBS Act and the Costs of Going Public*, 55 J. ACCT. RSCH. 795 (2017); Mary E. Barth et al., *The JOBS Act and Information Uncertainty in IPO Firms*, 92 ACCT. REV., no. 6, 2017, at 25.

53. See, e.g., Amirata Ghorbani & James Zou, *Data Shapley: Equitable Valuation of Data for Machine Learning* (2019), <https://arxiv.org/abs/1904.02868>; Ruoxi Jia et al., *Efficient Task-Specific Data Valuation for Nearest Neighbor Algorithms*, 12 PROC. OF THE VLDB ENDOWMENT 1610–1623 (2019); Ramesh Raskar et al., *Data Markets to Support AI for All: Pricing, Valuation and Governance* (2019), <http://arxiv.org/abs/1905.06462>; Ali Dorri et al., *Blockchain in Internet of Things: Challenges and Solutions* (2016), <http://arxiv.org/abs/1608.05187>; Guy Zyskind, et al., *Decentralizing Privacy: Using Blockchain to Protect Personal Data*, in 2015 IEEE SECURITY AND PRIVACY WORKSHOPS, 180–184 (2015). See also C. Reimsbach-Kounatze et al., *Exploring the Economics of Personal Data: A Survey of Methodologies for Measuring Monetary Value* (2013) (OECD’s survey on methodologies to measure monetary value of personal data).

54. Henry Hansmann et al., *Ownership and Organizational Form*, in 891 HANDBOOK ORG’AL. ECON. 897-907 (2013).

DATA AND MANURE

performance improves, the entrepreneur may retain board control. In sum, being an investor is distinct from being an owner. Thus, arguing that data subjects are investors is not the same thing as arguing that they are owners. What rights must be granted to data investors is another issue that deserves in-depth analysis.

CONCLUSION

We have argued that it is coherent and not farfetched to think that data subjects, at least under some circumstances, are entitled to certain rights that are afforded to financial investors. Unless there is a strong reason to override the principle of “No double standards (or Fair treatment),” data subjects should be entitled to claim that they are investors in data firms. Most subjects do not meet the subjective condition because there is a lack of transparency about what contribution they are making to companies, and efforts by many firms to make their contribution as opaque as possible. But justice demands that data subjects be made aware of their contributions, be able to claim their entitlements, and be treated like investors.