THE ROLE OF INDIVIDUALS IN THE DATA ECOSYSTEM:

Current debates and considerations for data protection and data rights in the United States

EXECUTIVE SUMMARY

FINTECH EDGE SPECIAL REPORT

FEDERAL RESERVE BANK
OF SAN FRANCISCO

The Role of Individuals in the Data Ecosystem: Current debates and considerations for individual data protection and data rights in the U.S.

Executive Summary

Author
Kaitlin Asrow, Fintech Policy Advisor, Federal Reserve Bank of San Francisco

Publication Date
June 3, 2020

Acknowledgements
The author would like to acknowledge the following individuals for their contributions to this research:

Symposium Partners
- Melissa Koide, FinRegLab
- Kelly Thompson Cochran, FinRegLab
- Aaron Milner, FinRegLab

Paper Reviewers
- Kelly Thompson Cochran, FinRegLab
- Jonah Crane, Klaros Group
- Douglas Elliott, Oliver Wyman
- David Medine, Consultative Group to Assist the Poor
- Andres Wolberg-Stok, Citi
- Chi Chi Wu, National Consumer Law Center

The author would also like to thank the following individuals from the Federal Reserve Bank of San Francisco for their invaluable contributions to the final paper:

Anna Baram, Avery Belka, Brian Walker, Caroline Pao, Cynthia Course, Marshall Eckblad, Mitchell Lee, Mongkha Pavlick, Santa Ram Susarapu, Steve Wertheimer, Walter Yao, Irina Yugay

COVID-19 Impacts and Considerations

The global COVID-19 pandemic has caused unprecedented disruption and change in every aspect of our lives. In spite of these changes, efforts to consider data governance frameworks for the United States (U.S.) appear even more relevant.

Since early 2019, the Federal Reserve Bank of San Francisco has engaged in focused research to understand data as a complex policy area, with a particular focus on the potential roles of individuals in the data ecosystem. This effort has culminated in the white paper, "The Role of Individuals in the Data Ecosystem: Current debates and considerations for individual data protection and data rights in the U.S." The majority of the paper was completed prior to the COVID-19 crisis therefore considerations relative to the pandemic are not explicitly addressed. Despite this, many of the data governance concepts that are presented for consideration and discussion remain directly relevant.

In particular, the current crisis has highlighted the tensions that can occur between individual preferences and societal preferences with regard to data; the importance of broad data protection; and the value of data portability and digital infrastructure.

HEALTH AND DATA IMPACTS

To contain the spread of the virus, public health officials have indicated that it is essential to track infected, and potentially exposed individuals.[1] Technology, and the collection and use of data, can facilitate this, but it can also expose individuals to privacy and security risks both now and in the future. Many countries are using mobile phones, and mobile phone applications, to gather a range of data including GPS locations, phone-to-phone proximity, and individual daily health status.[2] The full impact of these data activities is still unknown, however two essential challenges have already surfaced that underscore issues and concepts raised in the paper.

The first challenge is that for both digital contact tracing and the reporting of health status, it is more effective[3] if a majority of the population participates. Unfortunately, many Americans report that they are unlikely to voluntarily take part in this kind of data collection.[4] This tension highlights the larger issue raised in the paper, of when individual rights around data, such as consent to collection, could come into conflict with other policy goals.

The second challenge that has surfaced thus far is that the information being collected and used in the pandemic response is particularly sensitive, such as health status, and can be challenging to anonymize, or disassociate from specific individuals.[5] This second challenge underscores the importance of strong data protection requirements across entities that incorporate both cybersecurity and internal conduct expectations, a concept raised in the paper.

ECONOMICS, DIGITAL LIVES, AND DATA IMPACTS

To facilitate the economic response to the pandemic there has been a clear and pressing need to make data available quickly and securely for consumer authentication and loan underwriting. Furthermore, COVID-19 has highlighted the importance of effective digital infrastructure to facilitate economic responses such as stimulus payments to individuals, and to enable increasingly digital lives. Unfortunately, these needs have exposed a lack of standards and processes for sharing information, and the limitations of aging technical systems.[6] The value of data portability processes, as well as adequate infrastructure, are also issues raised in the paper.

Currently, there are not standardized systems and policies in the U.S. for moving data quickly and securely between diverse financial service providers in order to facilitate activities such as lending and payments. The paper delves into the potential of data portability as right for individuals, as well as the value of consistent and secure systems to facilitate such a right.

During the COVID-19 crisis there have been reports of individuals, primarily in rural and marginalized communities, that do not have the internet access or computing resources necessary to work or attend school remotely.[7] As lives become increasingly digital it is essential to consider how the U.S. will create, update, and maintain adequate digital infrastructure for everything from baseline needs such as internet access, to more complex systems such as facilitating payments. The importance of digital infrastructure to support both data rights and data protection is emphasized in the paper.

ESSENTIAL COVID-19 DATA CONSIDERATIONS

In addition to the issues and concepts raised in the paper, COVID-19 raises additional questions:

Where does increased data collection have the greatest opportunity to help? There is limited evidence that using data to monitor locations or proximity at the individual level is effective at identifying those who may have been exposed to the virus, compared to traditional contact tracing.[8] More work is needed to determine which types of data are actually necessary and effective.

Which entity/entities will be doing the actual data collection, and for what purpose? Currently, the most important consideration is determining which entities have the capacity to engage in these activities immediately and securely. For the U.S. that capacity rests largely in the private sector. Unfortunately, today there are limited avenues to hold private companies accountable for data security and conduct.

What rights will individuals have relative to data collected and used during the pandemic? This pandemic highlights the tension between individual agency and broader societal needs. An important distinction can be made between limiting choices now, and still providing them in the future. There may be opportunities to provide individuals with rights that can enable them to

review data collected and take actions—such as having data deleted—at a later date, even if those options are not available today.

How do we build data systems for the crisis that can be monitored, and reconsidered after it is over? The data systems that are built to deal with the pandemic today do not necessarily need to define U.S. data governance in the future. Once the pandemic has subsided, the country can hopefully learn from these crisis efforts, as we endeavor to make changes and build more permanent governance frameworks. Some companies[9] have already developed voluntary data conduct guidelines for the pandemic, and these standards, along with consistent security approaches could be used as starting point for future systems.

MOVING FORWARD

While data governance may understandably be a lower priority when lives are at risk, using data for the benefit of citizens and to help the U.S. weather this pandemic can happen in tandem with individual data protection and data rights.[10] The country has an opportunity to examine the role of data under these current challenging circumstances, even as it assembles guardrails and infrastructure that can help respond more quickly and confidently in the future.

Executive Summary

We are living in a time of immense technological change, driven in part by our ability to capture data, convert it into new information, and use that information to inform decision-making, automate activities, and develop new products and services. This use of data, and the insights that it can provide, are impacting every aspect of our lives.

Today, information is fueling massive business growth and helping individuals by enabling more personalized experiences, providing visibility into complex relationships such as financial services and health, and aiding in decision-making.[11]

Unfortunately information is also used to manipulate actions, exploitatively target people, discriminate on improper grounds, and open once intimate spaces up to unforeseen dissection and analysis.[12] These benefits and risks are increasing as more data are collected and used,[13] and they directly affect every person in the U.S., and therefore the country as a collective whole. The global COVID-19 pandemic has further emphasized this tension. As society seeks new ways to track and monitor the virus, and as even more daily activities become digital, concerns around privacy and data security continue to grow.

A range of stakeholders and policymakers are acknowledging the need to more proactively balance the benefits and risks of this data explosion,[14] and there are robust conversations occurring across the country around the potential for improving and expanding U.S. data governance regimes.[15]

An essential question that weaves through this dialogue is the role of individuals themselves in judging these benefits and risks, and managing data directly.

Terms such as “control”, “choice”, and “ownership” over data are increasingly being used to describe the roles that individuals could play in this ecosystem. These words underscore a shared acknowledgment that there is inherent value in individuals playing a role relative to data, but a fundamental question remains around what type, and what amount, of individual management is actually reasonable to expect in such a large and technically complex ecosystem. There are also questions regarding what needs to change within technology systems to make management choices and control truly meaningful.

Currently, information related to each of us is constantly collected and used, and there is often little we can truly do about it.[16] The most meaningful action that individuals can take around data today may be to avoid using digital services at all. However, this potentially cuts people off from the benefits of personalization, improved efficiency, and necessary activities such as school and

work.[17] Individuals themselves are increasingly aware of this reality, and report feeling a lack of both protection and control with regard to data, while still acknowledging the value of information for creating innovative products and services.[18]

Creating a U.S. data governance regime that more deliberately balances benefits and risks and creates an enhanced role for individual choice and autonomy is an exciting proposition and may be more necessary than ever,[19] but it is also uniquely challenging. This kind of undertaking would affect individuals and businesses across the county, and requires careful consideration of issues concerning privacy, cybersecurity, innovation, inclusion, domestic competition, international trade, and more.[20]

The Federal Reserve Bank of San Francisco (SF Fed) has undertaken research, policy analysis, and stakeholder outreach throughout 2019 to help provide a nuanced and neutral perspective on the issues described above, with a particular focus on the potential roles of individuals themselves. **It is the goal of this paper to help examine the complexity of data as a policy area and offer data governance concepts for consideration and further discussion.** Another contribution of this work is to offer consistent terminology that can be used to describe data governance goals moving forward.

This research builds upon a number of public and private efforts that have offered common principles for data governance, including work from the Consumer Financial Protection Bureau (CFPB),[21] the American Law Institute (ALI),[22] the Financial Health Network,[23] the Consultative Group to Assist the Poor (CGAP),[24] the World Economic Forum (WEF),[25] and others.[26] Of particular importance to this work was a symposium hosted by the SF Fed and FinRegLab, a nonprofit innovation center that tests new technologies and data to inform public policy and drive the financial sector toward a responsible and inclusive financial marketplace. This symposium, entitled, "The Role of Consumers in the Data Ecosystem" took place on November 4 - 5, 2019.[27] The event provided a forum to test and refine some of the ideas considered as part of this research. While the Symposium contributed to the concepts in this paper, the detail and analysis stand alone as a larger project.

The analysis and recommendations presented in this paper should *not* be interpreted as opinions of symposium partners, advisors, or participants, reviewers of this paper, the Federal Reserve Bank of San Francisco, or the Board of Governors of the Federal Reserve System.

This paper contends that while control over data has the potential to empower individuals and create new benefits, that concept alone cannot address the full range of challenging, and intersecting policy issues that the data ecosystem gives rise to.

In particular, there is an important distinction between enabling individuals to manage data, and expecting them to take action as a form of protection or to achieve other policy goals.

The paper is presented in two parts. Part 1 is intended to unpack and analyze many of the key debates and challenges that have been raised relative to individuals' role in the data ecosystem. Based on the analysis in each section, **key takeaways are offered that focus on benefiting and protecting individuals**, while highlighting particular legal and technological complexities to keep in mind.

REFRAMING FROM "DATA OWNERSHIP" TO "DATA RIGHTS"

It is easy to say we should "own" data about ourselves, but creating legal and administrative structures for individual "ownership" of data would be exceptionally complex. Furthermore, **an ownership framework that treats data primarily as a commodity, or in monetary terms, could have negative consequences for other policy goals. A shift is proposed away from the framing of "ownership", and instead towards a broader concept of "data rights".**

THE LIMITATIONS OF INDIVIDUAL CONSENT

Notice and consent regimes are broadly used in the United States to give individuals a role relative to data, but given the complexity of technology and the volume of digital interactions today, it places too heavy a burden on individuals to protect themselves. This regime is profoundly broken and there are significant challenges to improving it. One potential approach is to move away from detailed consent and introduce a "legitimate purpose" requirement across all data activities. To meet a legitimate purpose requirement, data activities would need to be necessary for the product or service requested, and not be harmful to the individual.

THE CRUCIAL CHALLENGE OF EQUALITY

The benefits and risks of data collection, processing, and use occur differently across diverse populations and therefore data protection and data rights will impact individuals differently. More research is needed to understand the needs and preferences of diverse populations relative to data, and to explore systems that could customize experiences and reduce management burdens across individuals. A particular focus on populations already under-represented in the design of, and access to, digital technology is especially warranted.

BALANCING INDIVIDUAL RIGHTS AND COLLECTIVE GOALS

There are a multitude of policy goals that interact with data and information. Siloed approaches to data governance that focus only on one policy objective may inadvertently interact with other goals and considerations. A multi-disciplinary, multi-sector approach to the creation and implementation of data governance can help reveal and address these policy intersections.

CURRENT U.S. DATA GOVERNANCE

U.S. laws that include data governance elements can yield important lessons about the effectiveness of particular approaches and where changes may be warranted in policy design and implementation going forward. **There are gaps in protection and confusing overlaps across current U.S. data governance laws. A broad, baseline approach to data governance could help address these issues.** There is also the potential to accomplish shorter-term improvements while larger structures are developed.

Part 1 seeks to demonstrate that while agency is important to respect the dignity and diverse needs of individuals, there are limitations to individual data control as a tool given the complexity of digital interactions today, and in light of other policy considerations. Part 2 of this paper strives to address these tensions through a two-sided data governance framework that includes both individual data protection and individual agency through active data rights. The scope of this idea is broader than any current Federal, State, or sector-based regulation, and is intended to be a conceptual design that can serve as a base for further refinement and adaptation.

THE FOUNDATION: INDIVIDUAL DATA PROTECTION

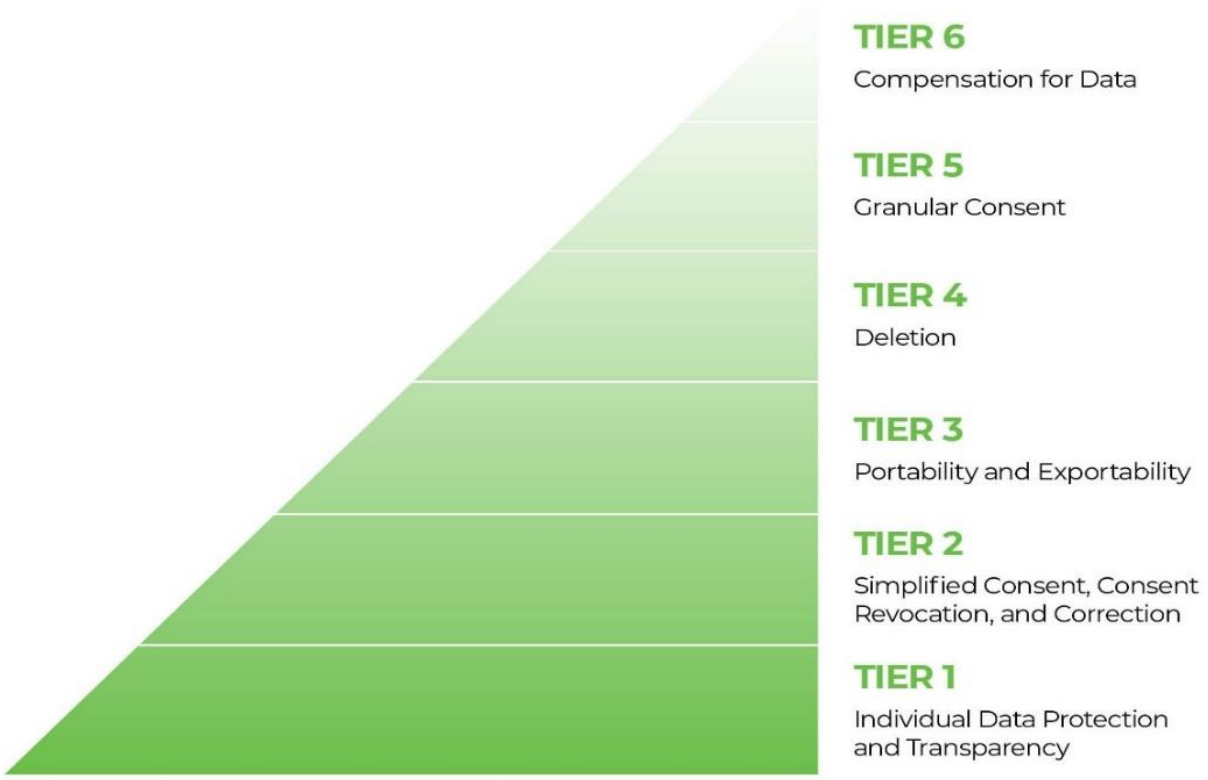
It is challenging, and maybe impossible, to quantify the exact risk that data can pose to individuals. Harm from data can come from external attacks or internal misuse of data by companies. Entities that use data, and how those data are used, blend across traditional sectoral and regulatory boundaries, and techniques used to "de-identify" data in order to protect individuals are largely ineffective. To address these realities, **a comprehensive data protection structure is proposed that would apply to all entities that collect, process, and use data. This structure would specify both security and conduct standards, such as a legitimate purpose requirement, and would extend protections to cover "de-identified" data.** The goal of this proposal is to create a clear, navigable, and baseline ecosystem for all entities, and to enable the subsequent active data rights framework. An essential premise of individual data protection is that **individuals should not need to take any affirmative actions, or exercise any 'rights', in order to experience data protection.**

A PROPOSED SPECTRUM OF ACTIVE DATA RIGHTS

Building upon the foundation of individual data protection, enabling individuals to directly act on information related to them can improve well-being, and may stimulate innovation, inclusion, and competition across markets. However, those considerations must be balanced with other social and policy goals, and an acknowledgment of how active data management on the part of individuals could create new cognitive, time, and resource burdens. To address these considerations **a bundle of active data rights is proposed that is structured into tiers that vary the rights available to individuals across circumstances.** Certain rights are foundational and apply across all situations, and others vary based on the intellectual property that entities contribute, how close data activities are to the original legitimate purpose of collection, and the risk of certain data and data formats.

Figure 1. A Spectrum of Data Rights

The pyramid base applies more broadly, while the availability of rights is reduced in subsequent tiers to account for other policy considerations.



TECHNOLOGY AND BUSINESS MODELS TO SUPPORT INDIVIDUAL DATA RIGHTS

While a broad framework for individual data protection and active data rights can help change the current data landscape for individuals and businesses, there is also **a need for new market-based innovation and business design that can support these data governance structures**. Ideally governance structures can both incentivize these evolutions, and benefit from them.

Across all of the analysis and ideas presented in this paper, there are common barriers to achieving broad changes to data governance: (1) complexity across current and future systems, (2) tensions with existing law and precedent, and (3) limitations to what technology can accomplish for data management today. Because of this, more collaboration and research are needed across all of the elements of individual data protection and active rights described in this paper.

This paper, and the concepts within it, are intended to stimulate conversation and provide a forward-leaning set of ideas to be vetted and refined collectively. The SF Fed looks forward to continuing to explore the potential for broad data governance in the United States.

References

[1] Mayo Clinic Networks. "What is contact tracing, and why is it important in fight against COVID-19?" MSN. April 28, 2019. https://www.msn.com/en-us/health/medical/what-is-contact-tracing-and-why-is-it-important-in-fight-against-covid-19/ar-BB13k54K

[2] "Privacy & Pandemics: The Role of Mobile Apps (Chart)." Future of Privacy Forum. April 2020. https://fpf.org/wp-content/uploads/2020/04/DP3T_The-Role-of-Mobile-Apps-Chart-10.pdf.

[3] "Digital contact tracing can slow or even stop coronavirus transmission and ease us out of lockdown." University of Oxford. April 16, 2020. https://www.research.ox.ac.uk/Article/2020-04-16-digital-contact-tracing-can-slow-or-even-stop-coronavirus-transmission-and-ease-us-out-of-lockdown

[4] Timberg, Craig, Drew Harwell, and Alauna Safarpour. "Most Americans are not willing or able to use an app tracking coronavirus infections. That's a problem for Big Tech's plan to slow the pandemic." The Washington Post. April 29, 2020. https://www.washingtonpost.com/technology/2020/04/29/most-americans-are-not-willing-or-able-use-an-app-tracking-coronavirus-infections-thats-problem-big-techs-plan-slow-pandemic/

[5] Zetter, Kim. "Anonymized Phone Location Data Not So Anonymous, Researchers Find." Wired. March 27, 2013. https://www.wired.com/2013/03/anonymous-phone-location-data/

[6] Collins, Michael. "IRS' antiquated technology could delay delivery of $1,200 coronavirus stimulus checks, experts warn." USA Today. April 4, 2020.https://www.usatoday.com/story/news/politics/2020/04/04/coronavirus-stimulus-outdated-technology-could-delay-checks-experts-say/5112012002/

[7] Goldstein, Dana, Adam Popescu, and Nikole Hannah-Jones. "As School Moves Online, Many Students Stay Logged Out." New York Times. April 6, 2020. https://www.nytimes.com/2020/04/06/us/coronavirus-schools-attendance-absent.html?campaign_id=158&emc=edit_ot_20200505&instance_id=18236&nl=on-tech-with-shira-ovide&regi_id=92128841&segment_id=26646&te=1&user_id=c8211ba7a76400964e6e36c656c9497e

[8] Crocker, Andrew, Kurt Opsahl, and Bennett Cyphers. "The Challenge of Proximity Apps For COVID-19 Contact Tracing." Electronic Frontier Foundation. April 10, 2020. https://www.eff.org/deeplinks/2020/04/challenge-proximity-apps-covid-19-contact-tracing

[9] Brill, Julie and Peter Lee. "Preserving privacy while addressing COVID-19." Microsoft Corporation. April 20, 2020. https://blogs.microsoft.com/on-the-issues/2020/04/20/privacy-covid-19-data-collection/."

[10] Sonmez, Murat. "How personal data could help contribute to a COVID-19 solution." World Economic Forum. March 23, 2020. https://www.weforum.org/agenda/2020/03/covid-19-personal-data-new-commodity-market/

[11] Carlin, Bruce Ian, Arna Olafsson and Michaela Pagel. "FinTech and Consumer Financial Well-Being in the Information Age." January 2019.

[12] Ceglowski, Maciej. "Statement of Maciej Cegłowski, Founder, Pinboard" United States Committee on Banking, Housing and Urban Affairs. United States Senate, May 7, 2019. https://www.banking.senate.gov/imo/media/doc/Ceglowski Testimony 5-7-19.pdf.

[13] Patrizio, Andy, and Andy Patrizio. "IDC: Expect 175 Zettabytes of Data Worldwide by 2025." Network World, December 3, 2018. https://www.networkworld.com/article/3325397/idc-expect-175-zettabytes-of-data-worldwide-by-2025.html.

[14] Various Authors. "The Privacy Project." The New York Times. Accessed March 2020. https://www.nytimes.com/series/new-york-times-privacy-project.

[15] "FTC Hearing 12: April 9 Session 1 Opening Remarks by FTC Chairman Joe Simons Followed by Panels on the Goals of Privacy Protection and the Data Risk Spectrum." Federal Trade Commission, September 26, 2019. https://www.ftc.gov/news-events/audio-video/video/ftc-hearing-12-april-9-session-1-opening-remarks-ftc-chairman-joe.

[16] Herrera, Tim. "You're Tracked Everywhere You Go Online. Use This Guide to Fight Back." The New York Times, November 24, 2019. Link.

[17] Hill, Kashmir. "I Cut Google Out Of My Life. It Screwed Up Everything." Gizmodo, January 29, 2019. https://gizmodo.com/i-cut-google-out-of-my-life-it-screwed-up-everything-1830565500.

[18] Auxier, Brooke, Lee Rainie, Monica Anderson, Andrew Perrin, Madhu Kumar, and Erica Turner. "Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information." Pew Research Center, December 31, 2019. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/.

[19] Uberti, David. "Coronavirus Surveillance Highlights Need for Federal Privacy Law." Wall Street Journal, April 17, 2020. https://www.wsj.com/articles/coronavirus-surveillance-highlights-need-for-federal-privacy-law-11587115801?ns=prod/accounts-wsj.

[20] Bremmer, Ian. "The American International Order Is Over." Time Magazine, November 18, 2019. https://time.com/5730849/end-american-order-what-next/.

[21] "Consumer Protection Principles: Consumer-Authorized Financial Data Sharing and Aggregation." Consumer Financial Protection Bureau, October 18, 2017. https://www.consumerfinance.gov/data-research/research-reports/consumer-protection-principles-consumer-authorized-financial-data-sharing-and-aggregation/.

[22] "Principles of the Law, Data Privacy." The American Law Institute. Accessed March 2020. https://www.ali.org/publications/show/data-privacy/#drafts.

[23] Asrow, Kaitlin, and Beth Brockland. "CFSI's Consumer Data Sharing Principles: A Framework for Industry-Wide Collaboration." Center for Financial Services Innovation, October 1, 2016. https://s3.amazonaws.com/cfsi-innovation-files-2018/wp-content/uploads/2016/10/31152340/2016_Data-Sharing-Principles1.pdf.

[24] Medine, David, and Gayatri Murthy. "Making Data Work for the Poor." CGAP, January 1, 2020. https://www.cgap.org/research/publication/making-data-work-poor.

[25] "The Appropriate Use of Customer Data in Financial Services." World Economic Forum, September 1, 2018. http://www3.weforum.org/docs/WP_Roadmap_Appropriate_Use_Customer_Data.pdf.

[26] The Digital Standard. Accessed March 2020. https://www.thedigitalstandard.org/.

[27] "2019 Data Symposium: The Role of Consumers in the Data Ecosystem." Federal Reserve Bank of San Francisco, July 1, 2019. https://www.frbsf.org/banking/fintech/events/2019/november/role-of-consumers-in-data-ecosystem/.