

UNDERSTANDING CORPORATE DATA SHARING DECISIONS: PRACTICES, CHALLENGES, AND OPPORTUNITIES FOR SHARING CORPORATE DATA WITH RESEARCHERS

November 2017



Acknowledgements

The Future of Privacy Forum and FPF Education and Innovation Foundation would like to thank Leslie Harris (FPF Senior Fellow), the Principal Researcher of this report, and Chinmayi Sharma (University of Virginia School of Law), Research Assistant. FPF gratefully acknowledges the support of the Alfred P. Sloan Foundation for this project.

TABLE OF CONTENTS

I.	Executive Summary.....	1
II.	Factual Background.....	2
III.	Research Methodology.....	3
IV.	Scope.....	3
V.	Researchers' Perspectives.....	4
VI.	Survey Findings.....	5
	a. Do companies make data available for academic research?.....	5
	b. Why do companies share data for academic research?.....	7
	c. How do companies make data available for academic research?	9
	d. What do companies perceive to be the risks of sharing data for academic research?.....	11
	e. How do companies address risks associated with data sharing for academic research?	13
VII.	Recommendations	16
	Appendix A: Corporate Interview Guide	17
	End Notes	18

I. Executive Summary

Large data sets are increasingly crucial tools for researchers, policymakers, NGOs, and commercial entities. Some large data sets are controlled by government agencies, non-governmental organizations or academic institutions, but most are held by the private sector. These data sets include information about key areas of societal concern that are highly valuable for academic research – information regarding human health, social behavior, economic activity, and more. Most corporate data is typically unavailable to academic researchers. While many companies make some data available to researchers through formal, publicly available processes, some sharing is also arranged on a tailored, targeted basis with trusted academics. A select set of elite institutions appear to benefit most from access to a variety of corporate data sources.

More widespread access to corporate data sets would support new scholarship and allow researchers to consider questions that cannot fully be answered from publicly available data alone. The Future of Privacy Forum (FPF) conducted research and interviews with experts in the academic and industry communities to determine:

- The extent to which leading companies make data available to support published research that contributes to public knowledge;
- Why and how companies share data for academic research; and
- The risks companies perceive to be associated with such sharing, as well as their strategies for mitigating those risks.

FPF's research and engagement with academic researchers and industry leaders revealed that:

- A small number of academic researchers have substantial experience with corporate data sharing, but many researchers have no or limited access to company data;
- Academics and companies perceive a pattern of increased sharing in recent years;
- Internet platforms and services have established some of the most robust programs to share data with academics, and select companies in other industries have robust sharing frameworks in place;
- Half of the surveyed corporate/academic data sharing programs began recently – within the past five years;
- Companies typically share data with academics for a variety of reasons. Common goals include gleaning insights to support execution of their corporate mission, demonstrating or unlocking the value of their data, and – to a lesser extent – advancing broad societal goals;
- Close to sixty-five percent of companies interviewed have clear processes or programs in place to share data with academics. Thirty-five percent solicit proposals, while the remainder share through diverse formal and informal models;
- Companies identify concerns about privacy and re-identification, as well as concerns about potentially diminished intellectual property value, as the main risks of corporate/academic data sharing; and
- Companies employ a range of measures – from internal review processes to data use agreements and technical safeguards – to mitigate the risks. These measures typically preclude access, sharing, or linking of individual-level records.

FPF's interviews support several interesting conclusions and opportunities for future inquiry and action:

- There is an opportunity to enhance the positive public profile of company/academic data sharing;
- There is an opportunity to strengthen the dialogue between researchers and academics;
- There is an opportunity to help mitigate perceived risks, particularly privacy and re-identification risks;
- There is an opportunity to develop and share tools for public outreach and community engagement;
- There is an opportunity to encourage peer-to-peer knowledge sharing; and
- There is an opportunity to create a clearinghouse identifying data types desired by academics.

II. Factual Background

Data has become the currency of the modern economy. A recent study projects the global volume of data to grow from about 0.8 zettabytes (ZB) in 2009 to more than 35 ZB in 2020,¹ most of it generated within the last two years and held by the corporate sector.²

As the cost of data collection and storage becomes cheaper and computing power increases, so does the value of data to the corporate bottom line. Powerful data science techniques, including machine learning and deep learning, make it possible to search, extract and analyze enormous sets of data from many sources in order to uncover novel insights and engage in predictive analysis. Breakthrough computational techniques allow complex analysis of encrypted data, making it possible for researchers to protect individual privacy, while extracting valuable insights.³

At the same time, these newfound data sources⁴ hold significant promise for advancing scholarship and shaping more impactful social policies, supporting evidence-based policymaking and more robust government statistics, and shaping more impactful social interventions.⁵ But because most of this data is held by the private sector, it is rarely available for these purposes, posing what many have argued is a serious impediment to scientific progress.⁶

A variety of reasons have been posited for the reluctance of the corporate sector to share data for academic research. Some have suggested that the private sector doesn't realize the value of their data for broader social and scientific advancement.⁷ Others suggest that companies have no "chief mission" or public obligation to share.⁸ But most observers describe the challenge as complex and multifaceted. Companies face a variety of commercial, legal, ethical, and reputational risks that serve as disincentives to sharing data for academic research, with privacy – particularly the risk of reidentification – an intractable concern.⁹ For companies, striking the right balance between the commercial and societal value of their data, the privacy interests of their customers, and the interests of academics presents a formidable dilemma.

To be sure, there is evidence that some companies are beginning to share for academic research. For example, a number of pharmaceutical companies are now sharing clinical trial data with researchers,¹⁰ and a number of individual companies have taken steps to make data available as well.¹¹ What is more, companies are also increasingly providing open or shared data for other important "public good" activities, including international development, humanitarian assistance and better public decision-making. Some are contributing to data collaboratives that pool data from different sources to address societal concerns.¹² Yet, it is still not clear whether and to what extent this "new era of data openness" will accelerate data sharing for academic research.

In this exploratory study, we aim to contribute to the literature by seeking the “ground truth” from the corporate sector about the challenges they encounter when they consider making data available for academic research. We hope that the impressions and insights gained from this first look at the issue will help formulate further research questions, inform the dialogue between key stakeholders, and identify constructive next steps and areas for further action and investment.

III. Research Methodology

FPF conducted a literature review and twelve unstructured interviews with individual academics and leaders of academic data sharing networks (“researchers”).

Most of the researchers interviewed are grantees of the Alfred P. Sloan Foundation. Several others were identified by the principal investigator. Together, the literature review and the interviews informed the development of a semi-structured survey instrument to guide the interviews with companies. For more specific interview information, see Appendix A.

The survey instrument seeks to answer 5 key questions:

1. Do companies make data available for academic research?
2. Why do companies share data for academic research?
3. How do companies make data available for academic research?
4. What do companies perceive to be the risks of sharing data for academic research?
5. How do companies address risks associated with data sharing for academic research?

More than 30 companies were invited to participate; 19 agreed to be interviewed.¹³ The companies represent diverse sectors, including high-tech manufacturing, workforce, education, healthcare, financial services, telecom/ISP, real estate/housing, ecommerce, data-related services, transportation, consumer genetics testing, and online services. Most are large well-known companies that are industry leaders, and in our view are highly representative of their sectors. The final sample included an equal number of long-established companies founded before 1990 (some dating back to the early twentieth century) and those founded since that time. Companies that were invited but did not participate in the survey were, with a few exceptions, older firms established before 1990.¹⁴ The non-participants included companies in the financial, online services, retail and technology sectors.

We conducted confidential interviews with the companies between March 15 and July 26, 2017. In most instances, senior members of the research and privacy teams participated in the interviews; most were conducted by phone. We followed up by email to clarify our understanding and ensure accuracy.

IV. Scope

In this report, we have focused our inquiry on corporate data sharing for research in instances where the shared data is likely to support published research that contributes to public knowledge. We have excluded long-term university partnerships that we understand to be intended exclusively to develop new products and services, new patents, and new business lines. We know anecdotally that some companies have multiple university partnerships, many for sponsored research related to robotics, artificial intelligence, advanced networking, IoT and core data science. We understand that these partnerships may produce enormous benefits to society, but in the absence of additional information about how they are structured, we have excluded them from our analysis. This is an area for further study.

V. Researchers' Perspectives

Of the researchers we interviewed, many had limited experience with corporate data. Only a few said that they or participants in their research networks had accessed significant corporate data sets for social science research, although there were a few notable success stories. Even so, many researchers believed that companies were beginning to share somewhat more data with researchers. Several reasons were suggested for this “partial thaw.” While a few thought that companies were responding to the growing demand for companies to demonstrate that they are “responsible,” others suggested a more practical reason for an uptick in data sharing: the increasing importance of unlocking the value of data for innovation and insights. One researcher also suggested that there was a generational change in companies that supported data sharing: “The new generation knows the value of sharing information they see the value and know it’s a trend and don’t want to fall behind.”

However, many researchers were doubtful that companies were likely to share data in ways that were scalable or useful for their own research. For example, poverty researchers often work with linked sets of administrative microdata in order to track the impact of policy, programs, or other social interventions over time. It seemed unlikely to them that companies would ever make individual-level record data available or permit that data to be linked with external data sets. Researchers who worked primarily with de-identified or aggregate data were doubtful that companies would allow researchers to combine corporate data with other external data sets.

Many researchers expressed concern about unavailability of corporate data in any form and the lack of scalability of the dominant “one to one” sharing model. Researchers said that it took months or more of trust building and negotiation to reach agreement with companies on data sharing.¹⁵ Moreover, each researcher had to begin the trust-building process anew.

Researchers also noted that there was little chance of accessing corporate data if the proposed research didn’t align with the interests of internal researchers. As companies hold more and more data, researchers also expressed concern that companies rather than social scientists will increasingly set the research agenda.¹⁶ A leader of a prominent academic data sharing network explained, “It is not that this data is never available to some researchers, but it is unlikely to be made widely available to all researchers.” Even if companies are sharing more data with an individual university for a specific research project, the barrier to more robust sharing, for example, through academic research networks, remains high “because companies do not want to give up control over the data.” Another suggested that the unwillingness to relinquish control was related to reputational damage that might arise from unfavorable results that cause harm to the company.

Finally, several researchers expressed concern that corporate data sharing was mostly benefiting elite data scientists at elite institutions. While the number of researchers that we interviewed is quite small, we did observe a data science “difference” insofar as data scientists/institutions reported a high degree of success in accessing corporate data in contrast to other researchers. One researcher explained it this way, “Companies have very little control over what someone might say about the company based on the assertion that they have analyzed the data. The risk of reputational harm is great.” By sharing with data scientists fluent in the most advanced techniques, the risks are greatly diminished. At the same time, the benefits of data sharing increase exponentially.

VI. Survey Findings

A. Do companies make data available for academic research?

- 70% of the companies report making at least some data available to academic researchers.
- Of the companies that share data for research, half began doing so within the last five years.
- Companies that provide products/services/user platforms via the Internet make data available for research.

Out of the 19 companies interviewed, 14 (the “sharing” companies)¹⁷ reported making at least some data available for academic research. The sharing companies are almost evenly split between long established firms, and those established in the last twenty years. None are startups or small firms. Half of the sharing companies began making data available to external researchers within the last five years. Of those, two began sharing within the last two years,¹⁸ and one is currently in the process of standing up an academic research program. Of the companies that began sharing in the past five years, five of the seven are Internet companies that provide products, services and user platforms of various types and two are long established companies in the education and data services sectors. Many of the sharing companies reported that they have research teams that engage in outward facing as well as inward facing research. Several noted their internal researchers are often published in academic journals.

We observed differences among industry sectors in their data sharing practices:

- Companies that provide goods/services/user platforms via the Internet all report that they make data available for academic research, and as noted above, most in this sector began sharing within recent years. This observation appears to be consistent with publicly available information about research sharing in this sector.¹⁹ If we were to extrapolate to a larger sample, we would expect to find confirmation of these observations.²⁰
- Companies in the “data sector” – data measurement, aggregation, analytics, and analysis – also report providing data for academic research.
- Companies that hold personal health data, including de-identified personal health and genetic information, share data with academic researchers. This observation is also consistent with available public evidence and we would expect it to hold true in a study with a large sample.²¹
- Companies in the labor/workforce sector and real estate sectors also reported sharing data for research.²²
- Companies in the telecommunications/Internet access sector share data through data challenges, internships, and occasional foundation projects, but do not generally otherwise share data for non-business-related research.
- For companies in the education sector, we observed a sharp division between the K-12 and higher education companies with respect to research sharing. K-12 companies were extremely reluctant to share data for academic research because of the ongoing controversy over collection and use of student data.²³ In contrast, the higher education company interviewed was currently developing a new academic research initiative to further learning science. High-tech manufacturing and semiconductor companies also do not generally share company data for research.

Is there a cultural shift toward data-sharing?

Our observation that half of the companies interviewed began sharing data for research in the last five years led us to wonder whether there was a cultural shift afoot, at least in some sectors and if so, what might be driving it. Both researchers and companies provided interesting insights on this question. Several suggested that as data became more valuable, more companies were establishing internal research groups that engaged in both inward facing and outward facing research. In turn, this provided the infrastructure and the impetus to share with external researchers. “We have seen a big trend over the years. Before most companies did not use it (data) for internal research, it was only operational which meant there were much higher costs and more barriers [to sharing] because nothing was set up inside.”

Others thought the influx of top academic talent into corporate research groups was also an important factor influencing academic sharing. By this account, this new generation of researchers were used to working with shared and open data. They maintained close academic ties and brought a fresh perspective to firms about the value of academic data sharing.

Further, while some researchers saw “data for good” activities as a distraction, others suggested that participation in these activities might provide a “gateway” to academic data sharing, not only because companies became more comfortable with using their data for non-business purposes, but also because data for good sharing required companies to put the internal systems and processes in place that were necessary for academic sharing. One company explained that its recently launched academic research program “could not have happened” without the internal work done to support a research challenge several years before:

We had to invest the time, money and thinking through what data can be made available, under what circumstances, a lot of things that are necessary for working with external researchers.

Another company suggested that data sharing should be viewed as a “competency” and the more data sharing became an everyday part of business activities, the more companies would become comfortable with sharing for research:

Companies are becoming comfortable using data shared from external sources and making their own data available. They are coming around to it, but there are competencies required to share data that few organizations have. By using data shared by other sources, companies will begin to learn how to share their own data.

Finally, several companies noted the importance of leadership support for data sharing. When company leaders valued data sharing with academics, barriers to sharing were easier to overcome.

Query: While a number of our survey participants suggested that there may be a cultural shift toward more academic data sharing and provided some possible reasons why, this question area that would benefit from additional research with a larger sample.

B. Why do companies share data for academic research?

Companies share data with academics for several reasons:

- To gain insights that support the company's mission.
- To unlock and demonstrate the value of company data.
- To support a company's philanthropic mission.

Companies identified several reasons for sharing data with academics. Close to half of the interviewed companies said that the main reason for sharing data for research was to obtain insights that would help the company “better execute” or “better understand” their mission. Our impression was that economic research and large-scale data science research was particularly valuable to many of these companies.

Several companies emphasized that “mission-aligned” research is not limited to research “*about*” the company or its narrow interests. Often, research that provides broader insights, about the workforce, online behavior or the economy, may be equally valuable to some companies and to furthering public knowledge. As one company explained, “the research that we do is almost never about [the company]. It needs to be general interest ... it needs to show core finding in economic science.”

A similar number of companies said that they made data available for academic research in order to “demonstrate” or “unlock” the value of their data for “novel” and “significant” research that advanced the field. Academic papers based on company data testified to the value and trustworthiness of the data. As one company explained, sharing with academics “is one way to distinguish ourselves Our data should withstand empirical rigor and scrutiny. We want to publish things ... it increases the trust in our data.”

Almost all of the companies that shared principally to “unlock value” identify huge databases that are made available to academics for research.²⁴ It was our impression that this approach to sharing enabled a greater diversity of research projects across disciplines and reduced some of the “friction” associated with sharing. For example, QuintilesIMS licenses data from its large proprietary databases and subsidizes the cost for investigators pursuing policy-relevant or clinically significant research that can improve health and the delivery of care. Over 300 academic papers have been published in major academic journals based on company data.²⁵ Zillow recently began sharing all of the company's upstream data used to build products with academic institutions and non-profit organizations for research across many disciplines. Thus far, the company has already signed over 65 data use agreements. Glassdoor makes retrospective survey data from the past 10 years available for original academic research as a way to demonstrate that the data is addressing larger economic questions, outside of its own needs.²⁶

Regardless of the principal reason cited for sharing, companies also said that they found that there were unanticipated benefits to sharing. For example, several companies that shared principally to “unlock” the value of their data found that the research conducted with its data was sometimes also useful to the company, although not the impetus for sharing. One company said that it also learned about its own data every time it shared with researchers. “When we do [share], we've found sometimes that their research has spillover and serendipitous benefits to how we can apply and use it.” Similarly, a company that shared to gain actionable insights also explained that sharing for research also demonstrated that its data was interesting and valuable.

A number of companies also said that sharing data for research helped to build the brand, strengthen relationships with academics and attract talent to the company.

Finally, we observe that a small number of companies share data for academic research in support of companies' philanthropic mission. For example, the Mastercard Center for Financial Inclusion shares aggregated data with academic researchers for research that aligns with its philanthropic mission of financial inclusion and assistance for the unbanked.

What can we learn from corporate data sharing for social good?

We observed that many companies said they were involved in a variety of “data for good” activities, including providing research insights and analysis of company data to governments, non-governmental organizations, and global bodies for better decision-making. For example:

- Facebook researchers provided anonymized insights to UNICEF from Facebook posts about the Zika conversation in Brazil, where more than 90 percent of the population use the platform every month. UNICEF then incorporated those learnings into a data-driven campaign that led to 82 percent of those reached taking action to protect themselves against Zika.
- Nielsen has shared data and insights on food pricing to help Feeding America publish its Map the Meal Gap and address food insecurity through its national network of food banks that provide support for millions of low-income Americans.

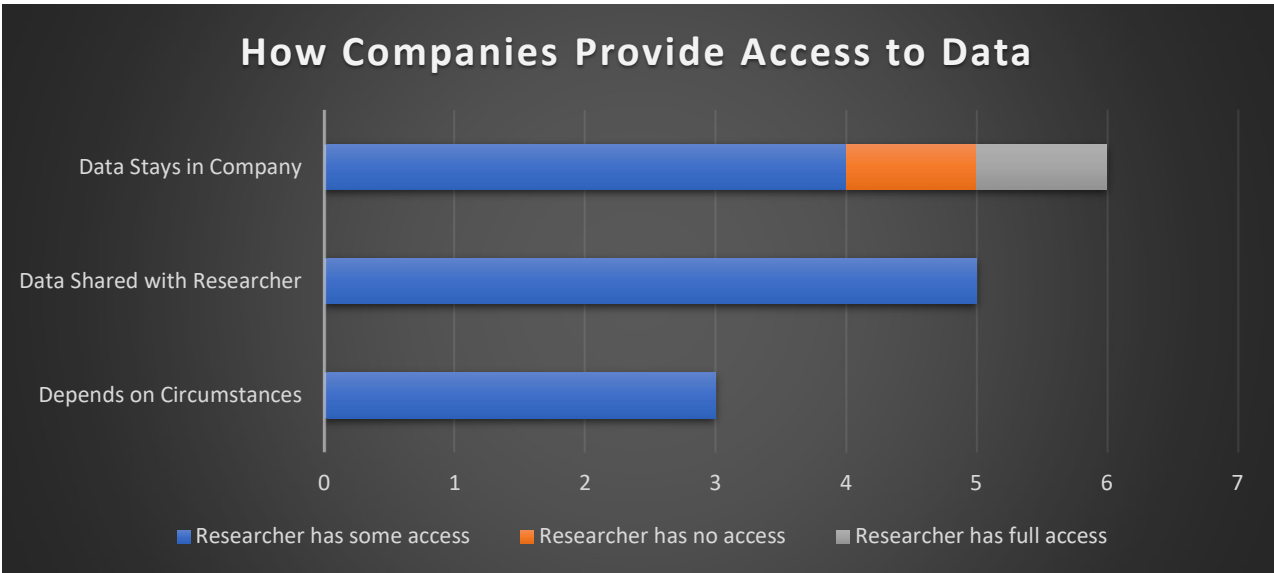
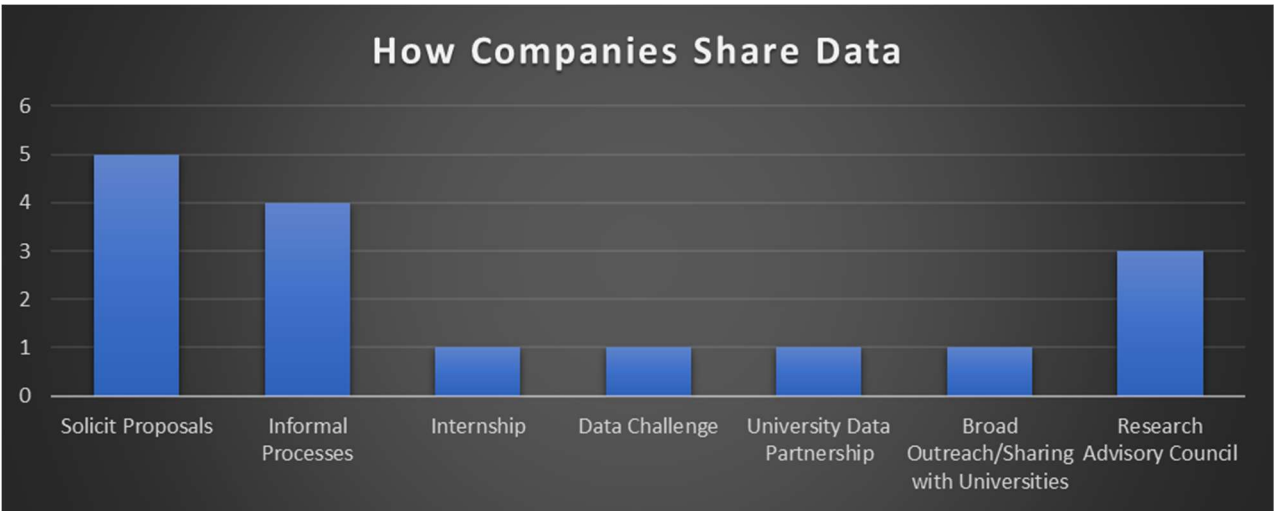
Companies were highly enthusiastic about the important social mission associated with these activities and the value company data brought to solving social problems. “Social good” research activities – often to support better government decision-making – were described by several as “actionable,” “immediate,” and “impactful” ways to deploy and extract value from company data for social good. In contrast, academic research was often described as important but risky, with long time horizons, greater risks and uncertain returns. One company put it this way, “Academic partnerships are great, but actions may go farther if done in conjunction with local and state partnerships. The impact side often occurs through these types of partnerships.”

Query: There is a strong value proposition for social good data sharing including goodwill, brand equity, relationship building, employee/ customer/user satisfaction, demonstration of data value and most importantly “actionable,” “immediate” and “measurable social impact.” How should this insight impact how the research community describes its work to corporate data holders? The individual one-to-one sharing model makes it difficult for companies to understand how any one research project might address a broader societal concern like poverty, homelessness, or the growing skills gap. Do researchers need to be more clear about the social impact of their research or contextualize their research within a broader societal frame. Should individual research be presented as part of a larger initiative with big “actionable” goals like ending poverty or homelessness? Should companies be challenged or enlisted to share their data for research that contributes to finding answers? If so, what are the priorities and who should be making the ask? Are there other ways to involve or recognize companies that share for socially important research that provide or inform some of the value associated with social good data sharing?

C. How do companies make data available for academic research?

- Sixty-five percent of the companies have clear processes in place for data sharing data with academics.
- Thirty-five percent solicit proposals.
- Most seek proposals that are aligned with internal research priorities.
- Others share through diverse formal and informal sharing models.
- Only a few companies share data through an Administrative Data Research Networks (ADRN).

Companies share data with academic researchers through varying formal and informal processes. Sixty-five percent share through clear publicly available processes. Thirty-five percent have established formal processes for the submission and review of research proposals. Companies that share both for insights and to demonstrate and unlock the value of data solicit proposals for research.



Most of the companies that solicit proposals prioritize research that is aligned with current research interests, although those priorities are often described in broad terms. For example, LinkedIn provided researchers with clear high-level guidance about its research interests when it launched its first round of solicitations for the Economic Graph Research program earlier this year.²⁷ QuintilesIMS encourages proposals on most aspects of health care, but also sets out areas of greatest interest.²⁸ A few simply look for proposals that will make good use of company data, are original, rigorous and likely to be published. Some companies engage with researchers informally. Several companies said that they reviewed proposals on an ad hoc basis, but did not have a formal structure for doing so. Others actively engaged with academics. For example, Uber's Public Policy and Economics team regularly engages in outreach to academics to present ongoing work, discuss research ideas and identify new projects and new collaborators.

Other companies make data available to researchers in diverse ways. Nielsen provides retrospective marketing data sets to the University of Chicago, delegating sharing decisions to the Kilts Center, which in turn licenses the company's large retrospective data to other universities for research use by students and faculty. Zillow established a formal program to share the data from its huge ZTRAX real estate databases with academics and nonprofit researchers. To reach interested researchers, the company engages in extensive outreach to colleges and universities across the country. While the company does not approve research topics, sharing is subject to a data use agreement that requires researchers use the data for non-commercial purposes, be properly attributed to ZTRAX and be kept confidential. Further, the code used to clean and structure the data for analysis must also be submitted for publication.

Several companies have made data available to researchers through research challenges. For example, Orange French Telecom has provided anonymized big data sets for an annual big data innovation challenge to support research for development. In the most recent challenge, 60 university groups submitted research papers on topics related to health, transportation, agriculture, and data science.²⁹ LinkedIn's Economic Graph Research Challenge in late 2014, issued a call for research proposals that would further economic opportunity, using the highly-structured data sets from the economic graph. Two hundred researchers submitted proposals and 11 finalists worked with the company to complete their research. The company established a new academic research program as a result of the challenge.³⁰

Internship programs provide another way to share data with academics. Subject to confidentiality arrangements, AT&T makes data available to graduate students through a summer internship program that brings students into the company to work with the company data under the supervision of an internal researcher. The data does not leave the company, though students are allowed to include the results of the research in their published work.

Three companies have established academic research steering committees, advisory boards, or research networks which to varying degrees, advise on research questions, review proposals, identify researchers or engage in collaborative research. McGraw-Hill Education for example, has recently established a Learning Science Education Research Council to guide its internal research and external research collaborations with selected academics.³¹

Only two interviewed companies provide data through ADRNs,³² with a third company expressing interest in sharing through an ADRN because of its commitment to data transparency. Zillow makes its large data sets available through an ADRN because it will "make the data ubiquitous when you are studying housing." Other companies were either unfamiliar with the ADRN model or believed that the model did not meet their rigorous safeguards for providing research access to data.

Finally, several companies also provided a range of open datasets or made some data available through an API, which have been used for academic research. For example, Amazon

recently released an enormous data set of over a half million images of bins and related metadata to data scientists to accelerate machine learning and computer vision research, subject to a creative commons license.

The “Substory” of Trust

Researchers, who had successfully accessed company data, all said that there was always a “substory” of trust behind their success. They explained that regardless of whether there was a formal or informal process available to academic researchers, personal relationships and “trust building over time” were the keys to a successful outcome. As one researcher explained, “it [data sharing] was almost always a network thing.”

A number of companies also emphasized the importance of trusted relationships to academic data sharing. One researcher reported that it was difficult to access company data, but once the researcher had demonstrated that she understood how to work with company data and produce scientifically rigorous research, a level of trust developed that led to more sharing. In some cases, companies established ongoing academic partnerships with trusted researchers. For example, one company noted that a single researcher had published more than 20 journal articles in major journals with company data over the last decade.

A few companies emphasized the high value of company data and the importance of limiting data sharing to academics who would produce “unimpeachable” research. Taking a chance on an unknown researcher was, for some, considered extremely risky because of the “difficulty in gauging credibility and reputation.” Several company researchers emphasized the importance of maintaining close academic ties and sharing research interests with trusted colleagues, particularly those with a successful history with the company. As one company explained there was “a lot of who you know” behind data sharing. Another put it this way, “We know who is interested and they do too.”

Query: There is no question that companies need to trust the academics that access their data for research. But does that mean that only a handful of researchers from elite universities will have access to corporate data? How can trust be built between companies and a broader set of academics, institutions, and sharing networks? What dialogues need to happen, and who should participate? Would credentials from a trusted intermediary make a difference? We believe that this is an important area for additional research and dialogue.

D. What do companies perceive to be the risks of sharing data for academic research?

- Companies are concerned about privacy, particularly the risk of re-identification.
- Companies are equally concerned that sharing data for research might diminish or destroy the intellectual property value of their data.

Privacy and security were cited as the top concern for companies that hold personal data because of the serious risk of re-identification. As one company explained: “[W]e lean in to protect privacy/security over research utility.” Some non-sharing companies cited privacy regulations as a barrier to sharing. A number of sharing companies also held data covered by privacy regulations, but built regulatory compliance into their data sharing processes.

The other oft-cited worry was that data sharing for research would damage or destroy intellectual property rights in the data. Several companies explained that data was a valuable

corporate asset. Sharing big data sets and allowing third parties to run algorithms against the data and share publicly posed a risk that the value would be lost or diminished. Several also noted that negotiating IP rights was the most difficult challenge to successfully reaching agreements with universities on data sharing. Loss of IP was also cited by non-sharing companies as a barrier to sharing.

Less often cited were concerns about consumer/customer/user reaction and contractual or regulatory compliance. Several firms said that carefully scrutinized projects to ensure that they were ethical and would not cause their customers/users discomfort. While public backlash was cited by K-12 companies that did not share data, it did not appear to be a major concern of the sharing companies.

Companies that hold data from corporate or government clients or license data from third-party vendors said that data provenance, customer expectations, contracts, risks to third party IP, as well as regulation limit, prohibit, or complicate data sharing. These concerns were expressed by sharing and non-sharing companies alike.

Company Research Enabling Activities

Several companies noted that they are engaged in “enabling activities” that support secure data storage, exchange and analysis of large complex data sets for academic research. Intel collaborates with academic researchers by providing cutting edge technology and know-how to important research initiatives. For example, the company has partnered with the Oregon Health & Science University to build a Collaborative Cancer Cloud, a precision analytics platform that allows institutions to securely share large amounts of patient genomic, imaging and clinical data to support research and speed diagnosis and treatment. Amazon Web Services (“AWS”) Cloud Credits for Research Program supports academic research by providing credits for free usage of cloud services and analytic tools that researchers can use to run computations and share large data with research partners around the world. Orange has partnered with the MIT Media Lab and the World Economic Forum, among others, to develop and pilot OPAL (Open Algorithms), an open suite of software and algorithms intended to run on OPAL servers in partner companies behind their firewalls, that will provide access to anonymized, formatted and encrypted statistical data to support research and global development. Several other companies also said that they were in the process of developing data sharing platforms.

Query: Secure cloud-based computing/sharing platforms are experiencing exponential growth and are being adopted by small and midsize firms.³³ To what extent will these secure research platforms mitigate privacy and security concerns that limit data sharing? Will sharing platforms reduce the complexity of research sharing for small and midsize firms?

E. How do companies address risks associated with data sharing for academic research?

- Many companies engage in a rigorous internal review of proposed academic research projects.
- Close to half of the companies retain custody and control over the research data at all times.
- Companies employ rigorous data use agreements to limit access to and use of shared data.

Companies employ a range of strategies to mitigate the risks of sharing data for academic research. Many companies report that they employ rigorous risk/benefit processes to mitigate the range of risks associated with making data available for research. Several companies emphasized that their risk management strategies are designed to limit risk through the entire “lifecycle” of a research project.

A number of companies said that external research proposals undergo rigorous internal review, often using the same research review processes required for internal research. Facebook, for example, has a robust set of review processes and systems for both internal and external research projects that look at the risks and benefits of the research that consider scientific and ethical merits of each proposal, the impact of the research on vulnerable populations and platform users, as well as privacy and security.³⁴ Other companies said that internal reviews include rigorous data mapping to understand data provenance and contractual or legal rules, contextual ethical reviews, privacy by design, security reviews as well as a review of research design and methodology. One company noted that it also reviews proposed algorithms and research questions for the potential to create bias. Two companies involve research advisory boards in the risk assessment process, looking to them for advice on research design, ethics, and privacy.

Several companies require that all academic research involving external researchers be treated as a formal collaboration between the company and the researcher.

Technical measures to limit access to the data also play an important role in risk management. A third of the sharing companies either transfer custody of data to the researcher, make it available for download, or make it publicly available. A few said that the custody of the data is highly dependent on circumstances. Close to half maintain custody of the data throughout the research engagement. Some companies provide some access to data remotely through the company’s firewall or an API. Often, access is strictly controlled and shared in a “read only” format. Several companies provide secure “sandboxes” which allow approved researchers to come into the company to work with data under the supervision of internal researchers. Sandboxed researchers can – in some cases – participate in all aspects of the research inside the company, except when highly personal sensitive data is involved. A few companies never or rarely permit external researchers to directly access the study data. In such cases, research algorithms are often developed collaboratively between the researcher and an internal researcher, but the computations are always run by a company researcher. The external researcher only sees “outputs” and “insights” of the research, never the data.

Companies also mitigate risk by limiting the amount and types of data shared. Some companies did not share personal information in any form. Others shared de-identified data or “rolled up” personal data into larger, highly aggregated data sets. Many companies said that they narrowly tailored the actual data shared to the specific needs of each research engagement.

Nearly all sharing companies employ data use agreements. We found common features in almost all of them:

- Strong protections for intellectual property rights and, where appropriate, the rights of third-party entities.
- Purpose limitations and, in some instances, limitations on use to a finite term.
- To the extent that data is allowed to leave the company, strong privacy and security requirements for data storage, including use of encryption.
- Requirements to return or destroy data at the conclusion of the research.
- Contractual prohibitions on re-identification, including, in some cases, provisions to promptly report instances of inadvertent instances of re-identification to the company.
- Prohibitions on data sharing, with the exception of the few instances where the data is provided to a data sharing network.

Companies varied in whether they permitted external researchers to combine company data with other data sets. Three companies allow “all kinds of data” to be combined with company data. One company did not allow anything “sensitive” to be combined with company data. Four companies said that they prohibited combining other data sets. Others did not provide information on this question. In addition to concerns about re-identification, companies are also concerned about the quality, accuracy, and veracity of the additional data sets, and the extent to which other data sets might impact the quality of research findings, damage the company’s reputation, or reveal insights about the company to competitors. Data rights and restrictions of third-party data were also mentioned as a limit on combining data sets.

No clear norms emerged concerning company policies regarding permission for researchers to link records at the individual level. More than half of the companies either did not hold personally identifiable information (PII), did not share PII, or only shared PII in a highly aggregated form. Of those that shared de-identified data, several said that they had permitted linking in “rare instances” and only with individual consent. In those cases, the research related to an examination of the impact of a company service. Another noted that requests to link records required careful scrutiny.

Companies that share highly sensitive personal health and genetic information also included additional provisions in data use agreements to ensure legal compliance with HIPAA and GINA. In the case of genetic information, sharing is always consent based and subject to an Institutional Review Board (IRB).

Most companies required a review before publication, a practice which researchers said is now “standard,” only a few did not. The principal reason stated for review was to ensure that there was no inadvertent re-identification. A few said that they reviewed research to ensure that the terms of the data use agreement were met or that the company’s data was properly described and cited. Others reviewed research in order to make sure that the researcher understood the data and the context for collection. These companies were concerned with the possibility that “unsound” research findings would result from misunderstanding or misuse of the data.

A small number of companies retained the right to control publication, although most said that they had not exercised that right and were unlikely do so because of the high degree of due diligence and care exercised in their sharing decisions. As one company explained, “[W]e work so carefully to craft a meaningful collaboration agreement, that scenario would be pretty surprising.” Several said that their data use agreements give all the parties equal say over the publication decision, or that the decision to publish was always made collaboratively by external and internal researchers.

Finally, companies were evenly split on whether and to what extent researchers were permitted to publish or make data available with publication. About half permit publication of some data in aggregate form. An equal number do not.

Companies were keenly aware of the significant challenge that restrictions on data sharing and publication place on researchers. Several said that the increasing number of academic journals that require that data be made available is becoming a “huge issue” in the company. Several noted that restrictions on data publication were also impacting internal researchers. Companies said that they often worked with researchers to make some data available for publication; for example, developing co-variance statistics to enable comparisons, providing publicly available data from other sources, developing dummy data, and creating limited sets of aggregated data.

The Challenge of Replication

The lack of replication of research is becoming an increasing challenge to the integrity of the scientific method. Researchers expressed concerns that replication and validation was particularly difficult for research conducted with corporate data because of the restrictions companies placed on sharing and publication of data. While acknowledging that challenges to replication were not unique to corporate data – for example, source code is rarely shared – several researchers said that the inability to replicate research using corporate data posed a severe threat to the scientific method “at a time where the legitimacy of science is increasingly being challenged.”

While it does not appear that companies are receiving replication requests, over half said that they either permitted enough data to be published for others to replicate or that they would consider replication requests. One provided open data to researchers. Two said that they did not provide data to researchers to conduct replication studies.

Several companies echoed researchers concerns about the lack of replication studies and the importance of replication research to the scientific method. Another stressed the importance of making both the code and the data available, suggesting that there were “professional returns” for allowing research to be replicated. Companies cautioned that there had to be “boundaries” around replication requests and that they would consider requests from “legitimate researchers” but not allow the data to be used for another study. For one company, business risk would also be a consideration.

Query: Are more companies willing to share data for replication? Our survey suggests that this may well be the case, but more research is needed.

VII. Recommendations

1. **There is an opportunity to enhance the positive public profile of company/academic data sharing.** Currently, companies balance the main perceived benefits of corporate/academic data sharing against the main risks; often, broader public benefits are not well-articulated or understood. There is currently no high-profile governmental or private sector mechanism that provides a social incentive for companies to share data in support of academic projects that benefit society, but may not be closely aligned with a company's research interests or mission. Future work could focus on a range of methods for encouraging, recognizing, and honoring corporate data sharing programs and initiatives. Options include: leveraging senior policymakers, thought leaders, and high profile academics to encourage and enlist companies in greater sharing (call to action, CEO outreach, etc); and/or publicly identifying a set of socially important priorities that can be supported by data sharing (e.g. issues related to employment, poverty or the environment), and recognition of companies that share through awards and other mechanisms.
2. **There is an opportunity to help mitigate perceived risks, particularly privacy and re-identification risks.** Companies noted privacy and re-identifications risks as leading concerns that inhibit broader data sharing with academics. Future work could focus on developing robust, flexible privacy safeguards and norms that can support greater sharing and mitigate re-identification risks and other privacy concerns. Such safeguards could include: high-quality de-identification resources, technical and policy controls that support secure data use, and a central resource for privacy impact assessment tools. Risks could also be mitigated by development of a central repository for resources detailing frameworks for privacy, security, and ethical review of data sharing arrangements.
3. **There is an opportunity to develop and share tools for public outreach and community engagement.** Some companies perceive risk of public backlash as a limiting factor on data sharing with academics. Tools can be developed and shared that support public outreach regarding the benefits of data-driven academic research and data sharing arrangements, as well as describing the privacy safeguards that can protect personal data in such arrangements.
4. **There is an opportunity to encourage peer-to-peer knowledge sharing.** A network of evangelists from companies and academic institutions who participate in data sharing arrangements could discuss challenges, opportunities, best practices, and activities that would support broader, more frequent data sharing. The network could create and share best practices and develop tools to support enhanced data sharing in support of academic priorities.
5. **There is an opportunity to create a clearinghouse identifying data types desired by academics.** The clearinghouse could bring together leading academics to identify core areas of consensus priorities, provide a "matchmaking" service between companies and academics, and support networking opportunities for academics who engage in data-driven research.

Appendix A

Corporate Interview Guide

Semi-structured interviews were conducted covering the following subject areas:

1. Quick level-set on type of company and types of data it collects, (administrative, survey, user-generated, sales, personnel, inventory, etc.) and the general uses of the data that it collected (product/service improvement, new products/services, new IP, analytics, etc.)
2. The various ways that data might be shared with academics or other noncommercial actors (data for good programs, publicly available data sets, formal academic research, etc.)
3. If data is not shared for academic research, then discuss the concerns /reasons/barriers for not sharing (cost, privacy concerns, lack of capacity, no process, legal constraints, no business case, etc.)
4. If data is shared with outside researchers, a discussion about how researchers seeking data interact with the company (how do they find the company, does the company find them, the process for requesting data (formal/informal), the types and formats of data requested and shared/made available, the company decision-making process around sharing data for research, i.e. internal review process on ethics, privacy, who is involved, etc.)
5. How data is “shared” with external academics, for example, data provided to researcher or to university data center, data stays in the company and researcher provides code, researchers paired with internal researcher, researcher comes into the company as a fellow, etc.
6. Research partnerships – ongoing formal relationships with specific universities, university data science or social science institutes, operating under a master data sharing agreement – advantage of partnerships to companies and to researchers.
7. Principal concerns about sharing data for academic research? Example: privacy/security/re-identification, reputation, IP, misunderstanding of the data, customer backlash.
8. How company protects user privacy/security when sharing with an external researcher.
9. Data Use Agreements:
 - a. Terms, for example, limits on sharing, limits on publication of data, IP, privacy/re-identification, right of review, control over publication etc.
 - b. Experiences with negotiating data use agreements with universities and thoughts on barriers to agreements.
10. Knowledge sharing via various models of data sharing within academia and a company’s willingness to share data:
 - a. University-based data commons supported by data scientists, strong governance, access controls, encryption.
 - b. Academic Research Networks – consortia of academic institutions which share research data. Various models – data remains in each institution or is shared, strong data science support, governance, access controls, encryption.
 - c. Ongoing partnership between company and specific university or university entity that allows researcher broader latitude on research topics, greater collaboration with company, etc.

End Notes

¹ Thomson Reuters Industry Forum report, *Unlocking the Value of Research Data*, 3 (Reuters, Jul. 2013).

² Stefaan Verhulst, *Mapping the Next Frontier of Open Data: Corporate Data Sharing*, GovLab Blog (Sep. 16, 2014) <http://thegovlab.org/mapping-the-next-frontier-of-open-data-corporate-data-sharing/>; *The GovLab Index: The Data Universe*, GovLab Blog (Aug. 22, 2013), <http://thegovlab.org/govlab-index-the-digital-universe/>.

³ Stephen Eglash, Stanford Data Science Initiative, *Corporate Data Access and Sharing: Private Data Resources for the Common Good*, Presentation to the Workshop on the Current Practices of Private Companies and Their Use of Big Data and Key Issues and Challenges with Privacy and Confidentiality (Feb. 25 2016) published by The National Academies Committee on National Statistics, http://sites.nationalacademies.org/DBASSE/CNSTAT/DBASSE_170269; See, e.g., Fisher et al., *Quantum computing on encrypted data*, Nature Communications (Jan. 21, 2014), <https://www.nature.com/articles/ncomms4074>.

⁴ “Found data” is distinguished from “made data” in that the former is data collected as a byproduct of a pre-existing function or process, whereas the latter is collected specifically by design for research purposes, usually through surveys and experiments. Examples of found data can include transaction data such as credit card use, administrative data such as hospital records, data from social programs or marriage licenses, user-generated data such as social media content, and operational data such as computer diagnostic logs. See Connelly et al., *The role of administrative data in the big data revolution in social science research*, 59 Social Science Research 1, 3 (Sept. 2016), <http://www.sciencedirect.com/science/article/pii/S0049089X1630206X>.

⁵ For example, the Commission on Evidence-Based Policymaking, which was established by Congress to support better use of data resources among federal agencies to inform and improve decision-making and policy analysis, recently released a final report urging “a future in which rigorous evidence is created efficiently, as a routine part of government operations, and used to construct effective public policy” and noting that “improved access to data under more privacy-protective conditions can lead to an increase in both the quantity and the quality of evidence to inform important program and policy decisions.” See The Commission on Evidence-Based Policymaking, <https://www.cep.gov/cep-final-report.html> (last visited Aug. 30, 2017); See also 5 U.S.C. § 3109 (2016); 5 U.S.C. § 5316 (2016). A similar effort is underway at the National Academies of Sciences, which recently examined how new data sources, including corporate data, could modernize government statistics for policy and research. National Academies of Sciences, Engineering, and Medicine, *Innovations in Federal Statistics: Combining Data Sources While Protecting Privacy*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/24652>.

⁶ Liran Einav & Jonathan Levin, *Economics in the age of big data* 346 Science 715 (2014) (stating that corporate data can “offer researchers a look inside the ‘black box’ of firms and markets by providing meaningful statistics on economic behavior such as search and information gathering, communication, decision-making, and microlevel transactions”); Gideon Mann, *Private Data and the Public Good*, Medium (May 17, 2016), <https://medium.com/@gideonmann/private-data-and-the-public-good-9c94c656ff28> (discussing how corporations hold the key to unlocking research insights otherwise unavailable to the public due to the rising cost of survey research and the limitations of a profit-driven internal corporate research structure, and noting that without concerted changes in corporate data sharing practices, academics will remain unable to tap into these resources).

⁷ Thomas Roca & Emmanuel Letouze, *Open algorithms: A new paradigm for using private data for social good*, Devex (Jul. 18, 2016), <https://www.devex.com/news/open-algorithms-a-new-paradigm-for-using-private-data-for-social-good-88434>.

⁸ Robert Groves, *Improving Government, Academic and Industry Data-Sharing Opportunities*, in Krosnick et al., *The Future of Survey Research: Challenges and Opportunities* 130, a report of the NSF Advisory Committee for the Social, Behavioral and Economic Sciences Subcommittee on Advancing SBE Survey Research (2012), https://www.nsf.gov/sbe/AC_Materials/The_Future_of_Survey_Research.pdf.

⁹ See Stefaan Verhulst & Robyn Caplan, *Open Data: A Twenty-First-Century Asset for Small and Medium-Sized Enterprises*, 42–45 (GovLab 2015), <http://images.thegovlab.org/wordpress/wp-content/uploads/2015/08/OpenData-and-SME-Final-Aug2015.pdf> (discussing the challenges companies face in collecting the right information, processing or synthesizing data, analyzing data to uncover insights, and sharing data while maintaining user privacy); Simson L. Garfinkel, NISTIR 8053: *De-Identification of Personal Information* (NIST 2015), <http://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf> (reviewing the de-identification challenges and risks of re-identification presented by aggregate datasets).

¹⁰ Fletcher et al., *Statistical guidance for responsible data sharing: an overview*, 16 BMC Medical Research Methodology 1 (2016); Sudlow et al., *EFSPi/PSI working group on data sharing: accessing and working with pharmaceutical clinical trial patient level datasets – a primer for academic researchers*, 16 BMC Medical Research Methodology 23, 24–25 (2016); Markus Perkmann & Henri Schildt, *Open data partnerships between firms and universities: The role of boundary organizations*, 44 Research Policy 1133, 1133–1135 (2015) (discussing how the Structural Genomics Consortium (SGC) brought together pharmaceutical firms to determine the three-dimensional shape of proteins and release these insights openly for further research without restriction); See, e.g., *Clinical Trials: Our Data Sharing Commitments*, Sanofi, http://en.sanofi.com/Innovation/clinical-trials-and-results/our_data_sharing_commitments/our_data_sharing_commitments.aspx (explaining one pharmaceutical company’s policies on data sharing and their public support for collaborating with researchers) (last visited Aug. 30, 2017).

¹¹ See e.g., Klint Finley, *Twitter Opens its Enormous Archives to Data-Hungry Academics*, Wired, Feb. 6, 2014, <https://www.wired.com/2014/02/twitter-promises-share-secrets-academia/> (discussing how Twitter shifted from one-off partnerships with researchers to systematically opening up access to its data archives for wide public use of data for social good research); Dr. Andrew Chamberlain, *New Academic Research with Glassdoor Data*, Glassdoor Econ. Research Blog (Aug. 4, 2016), <https://www.glassdoor.com/research/new-academic-research-with-glassdoor-data/> (reviewing the goals and priorities of Glassdoor’s research arm and recent examples of researchers using Glassdoor data to provide important insight into company and employee relations); *Yelp Dataset Challenge*, GovLab, <http://datacollaboratives.org/cases/yelp-dataset-challenge.html> (last visited Aug. 30, 2017) (publicizing a data challenge the company organized in which it provided researchers with access to user data on restaurants across cities to build tools on urban trends and behavior).

¹² Stefaan Verhulst, Andrew Young & Prianka Srinivasan, *An Introduction to Data Collaboratives: Creating Public Value by Exchanging Data*, GovLab (unpublished presentation), <http://datacollaboratives.org/static/files/data-collaboratives-intro.pdf>.

¹³ The Future of Privacy Forum works with Chief Privacy Officers of companies around the world and was able to facilitate identification of appropriate interviewees.

¹⁴ Only a few companies said that they did not want to participate in the study. Rather, most of the non-participants either failed to respond to multiple email requests or were unable to set up interviews within the study time frame.

¹⁵ See, Andrey Fradkin blog, *A Guide to Using Corporate Data for Academic Research* (last visited Oct. 24, 2017),

<http://andreyfradkin.com/posts/2014/02/08/how-to-obtain-proprietary-datasets-for-research-part-1>.

¹⁶ See, Ben Williamson, *The death of the theorist and the emergence of data and algorithms in digital social research*, The London Sch. of Econ. Blog (Feb. 10, 2014), <http://blogs.lse.ac.uk/impactofsocialsciences/2014/02/10/the-death-of-the-theorist-in-digital-social-research/>.

¹⁷ We use the term “sharing,” for convenience. We do not intend to suggest that all companies provide researchers with physical access to the data or permit it to be transferred.

¹⁸ Not all companies provided information on this question.

¹⁹ These companies come from diverse sectors including social networking, workforce, retail, housing and real estate, and personal genetic services.

²⁰ See *supra* note 11 for examples.

²¹ See e.g., *supra* note 10 on increased sharing of healthcare and clinical trial data. For institutions covered by the HIPAA Privacy Rule, there are clear rules that govern the sharing of personal health information for research that reduce the complexity and legal uncertainty of sharing. See, National Institutes of Health, *How Can Covered Entities Use and Disclose Protected Health Information and Comply with the Privacy Rule?* https://privacyruleandresearch.nih.gov/pr_08.asp (last visited, Oct. 23, 2017).

²² President Obama’s Workforce Data Initiative, launched several years ago, has by many accounts encouraged greater sharing of workforce/labor data for research, as evidenced by the creation of the Skills Cooperative Research Database in 2016, a partnership between the Center for Data Science and Public Policy (DSaPP) at the University of Chicago and the Alfred P. Sloan Foundation. The database integrates data from public sources with privately-held data for academic research. DSaPP posting, *Sloan Grant Supports Construction of Powerful New Labor Database* (posted on July 13, 2016). <http://dsapp.uchicago.edu/2016/07/13/sloan-grant-supports-construction-powerful-new-labor-database/>.

²³ Brenda Leong, *Student data privacy: Moving from fear to responsible use*, Brookings Blog: Brown Center Chalkboard (May 23, 2016), <https://www.brookings.edu/blog/brown-center-chalkboard/2016/05/23/student-data-privacy-moving-from-fear-to-responsible-use/>; See Monica Bulger, Patrick McCormick & Mikaela Pitcan, *The Legacy of inBloom*, 25–28 (Data & Society, Working Paper No. 02.02.2017). *But* see *Beyond One Classroom: Parental Support for Technology and Data Use in Schools 6*, (Future of Privacy Forum, Dec. 8, 2016), <https://fpf.org/wp-content/uploads/2016/12/Beyond-One-Classroom.pdf>.

²⁴ While researchers are able to conceptualize research projects based on the availability of data from these huge databases, most of the companies limit the data they actually *share* with researchers to a specific set of extracted data that has been carefully tailored to the research inquiry.

²⁵ QuintilesIMS Research Support, <http://www.imshealth.com/en/thought-leadership/quintilesims-institute/research-support> (last accessed October 29, 2017).

²⁶ Andrew Chamberlin, *New Academic Research with Glassdoor Data*. August 4, 2016. <https://www.glassdoor.com/research/new-academic-research-with-glassdoor-data/> (last accessed October 29, 2017).

²⁷ See, Economic Graph Research Details <https://engineering.linkedin.com/data/economic-graph-research/economic-graph-details> (last visited Oct. 18, 2017).

²⁸ See, QuintilesIMS Research Support <http://www.imshealth.com/en/thought-leadership/quintilesims-institute/research-support> (last visited Oct. 18, 2017).

²⁹ Orange, *The D4D Challenge is a great success!* <http://www.d4d.orange.com/en/presentation/endowment-and-panel/Folder/The-D4D-Challenge-is-a-great-success> (last visited Oct. 24, 2017).

³⁰ Introducing the LinkedIn Economic Graph Research Challenge, <https://specialedition.linkedin.com/> (last visited Oct. 24, 2017).

³¹ McGraw-Hill Education, Learning Research Council, <https://www.mheducation.com/learning-science/research-council.html> (last accessed October 29, 2017).

³² Although only one company said that it was making data available through an ADRN during the interview process, we understand from researcher interviews that one other company in the survey group has contributed data to a research sharing entity.

³³ Louis Columbus, Roundup Of Small & Medium Business Cloud Computing Forecasts And Market Estimates. <https://www.forbes.com/sites/louiscolumbus/2015/05/04/roundup-of-small-medium-business-cloud-computing-forecasts-and-market-estimates-2015/#564992c932b0>. (last accessed November 7, 2017).

³⁴ Molly Jackman and Lauri Kanerva, *Evolving the IRB: Building Robust Review for Industry Research*, Washington and Lee Law Review Online, Volume 72 | Issue 3 (2016).



1400 EYE STREET, NW | SUITE 450 | WASHINGTON, DC 20005 • [FPF.ORG](https://www.fpf.org)

202.768.8950

[INFO@FPF.ORG](mailto:info@fpf.org) | [MEDIA@FPF.ORG](mailto:media@fpf.org)