CENTER
FOR
DATA
INNOVATION

November 23, 2016

Office of Science and Technology Policy
Eisenhower Executive Office Building
1650 Pennsylvania Avenue
Washington, DC 20504

On behalf of the Center for Data Innovation (datainnovation.org), we are pleased to submit these comments in response to the Office of Science and Technology Policy's (OSTP) request for information regarding data portability.[1]

The Center for Data Innovation is the leading think tank studying the intersection of data, technology, and public policy. With staff in Washington, DC, and Brussels, the Center formulates and promotes pragmatic public policies designed to maximize the benefits of data-driven innovation in the public and private sectors. It educates policymakers and the public about the opportunities and challenges associated with data, as well as technology trends such as predictive analytics, open data, cloud computing, and the Internet of Things. The Center is a non-profit, non-partisan research institute affiliated with the Information Technology and Innovation Foundation.

Data is often described as "the new oil," in that it has become an immensely valuable resource that fuels new products and services in the public and private sectors. As such, policies that promote data portability—the ease with which consumers can access and export their data in a machine-readable format from one service and import it into another one—can enable considerable economic benefits for consumers and businesses alike. Most importantly, data portability allows individuals and organizations to maximize the utility of their data by allowing data collected by one service to be reused in another, thereby enabling the development of valuable third-party services. In addition, data portability fosters competition by reducing switching costs and avoiding vendor lock-in. OSTP should be commended for taking steps to better understand the benefits of data portability and identify opportunities to promote data portability.

---

[1] "Request for Information Regarding Data Portability," The White House, Accessed November 22, 2016, https://www.whitehouse.gov/webform/request-information-regarding-data-portability.

Please find our responses to the relevant questions in the attached document.

Sincerely,

Daniel Castro
Director
Center for Data Innovation
dcastro@datainnovation.org

Joshua New
Policy Analyst
Center for Data Innovation
jnew@datainnovation.org

1. THE BENEFITS OF INCREASED DATA PORTABILITY

Data is at the core of many recent innovations in fields as diverse as health care, transportation, and energy. The more easily data can be shared, the more opportunities there are to use it as a platform for innovation. Thus, increases in data portability equate to increases in the social and economic benefits that data-driven innovation can unlock. When consumers can easily access and export machine-readable data about themselves collected by companies or government agencies and then import this data into other services, the opportunities for data-driven innovation increase substantially. For example, by enabling consumers to access data about their utility usage from smart meters, consumers can make more informed decisions about their electricity and water use habits and leverage third-party analytics services to identify opportunities to conserve resources and lower their utility bills.[2] In short, data portability enables the long-tail of data innovation.

Increases in data portability can also promote competition in the marketplace and substantially reduce switching costs for consumers by making it easier for them to export their data and bring it to another service provider. This reduces the opportunity for companies to artificially "lock-in" customers by making it prohibitively expensive to move their data to another company, and instead encourages companies to retain customers by offering the most competitive services.

This type of competition is especially important in data-intensive industries because many high-value innovations result from how data is used, not from how it is collected. It is inefficient to have multiple companies collect the same data. Ideally, data would only be collected once and used hundreds of times. For example, in health care, patients do not want to have their blood drawn every time they visit a new doctor, they want to have one result that they can share with all their health care providers. Most high-value innovations will be with creating better analytics to interpret blood tests, not taking better blood samples. The goal of public policy should be to ensure competition with data use, and that means promoting policies that make data available for sharing and reuse. Moreover, by promoting data portability, policymakers can ensure that companies without the resources to invest heavily in data collection can still access and use data to introduce new products and services.

---

[2] Daniel Castro and Brandon De Bruhl, "How to Promote Smarter Water Use by Giving Consumers Access to Their Consumption Data" (Center for Data Innovation, September 7, 2015), http://www2.datainnovation.org/2015-water-data-green-button.pdf.

Data portability can only work if data protection laws do not interfere. Overbearing privacy rules preventing data sharing or data minimization rules narrowly restricting data collection are directly at odds with the goals of data portability and limit the benefits of data innovation.

2. THE INDUSTRIES OR TYPES OF DATA THAT WOULD MOST BENEFIT OR BE HARMED BY INCREASED DATA PORTABILITY

Overall, the areas in which increased data portability would lead to the most benefits are those where the value of data is not fully aligned with who holds the data, and the greater this discrepancy is, the greater the need for data portability. While there are many sectors of the economy that would benefit from increased data portability, three areas in particular—health care, education, and utility providers—stand out as having among the most to gain.

HEALTHCARE

In health care, a patient's electronic health record (EHR) is a crucial data source. For example, doctors use EHRs to make more informed treatment decisions, hospitals use EHRs to manage patients and prioritize how they allocate resources, and researchers use EHRs to conduct large scale analysis of patient health and the health-care system, leading to new drug discoveries, increases in the efficiency of health-care providers, and more. The more easily health-care stakeholders and patients can access and share EHRs, the more likely these kinds of innovations are to occur. However, health-care providers sometimes engage in information blocking—the practice of knowingly inhibiting the sharing of health information without legitimate justification, such as privacy or security considerations—to anticompetitively protect their market share and prevent customers from bringing their data to competitors.[3] By enacting stronger restrictions to information blocking and ensuring patients can freely access and export their EHRs to bring them to providers of their choosing, the increased portability of this data could help make the health-care sector more competitive, lower health-care costs, and lead to new valuable and potentially life-saving innovations.

---

[3] Joshua New, "Congress Has a New Cure to Stop Companies from Blocking Patient Data," Center for Data Innovation, November 10, 2015, https://www.datainnovation.org/2015/11/congress-has-a-new-cure-to-stop-companies-from-blocking-patient-data/.

EDUCATION

In education, at both the K-12 and higher education levels, increased data portability can improve student achievement and parent engagement, lead to valuable new personalized education services, and help families make more informed decisions about financial planning and higher education. Student data portals for K-12 education can help students and parents easily access detailed data about student performance, ranging from grades to the development of non-cognitive skills such as self-reliance. Students can use this data to be more cognizant of their own performance and parents can use this data to be more involved in their children's education. Additionally, with the ability to export machine-readable data from these portals, the potential for third parties to develop valuable education services increases substantially. For example, parents could export their child's data and share it with a tutoring service that can then offer personalized, supplementary instruction designed around his or her unique strengths and weaknesses.[4] Similarly, college preparation services could analyze student data to provide personalized recommendations for choice of college based on a student's interests and strengths, give families realistic estimates of college costs by estimating the likelihood of a student receiving merit-based scholarships, and provide students with information about their likely future earning potential based on their educational history and choice of field of study and institution—all of which can help students make more informed decisions about pursuing loans or where makes the most sense to apply.

Data portability in education is especially important for students who change schools. Students who switch schools, especially multiple times, face academic challenges, including lower academic achievement, lower school engagement, and a higher risk of dropping out.[5] Students change schools for a number of reasons, but those who are homeless, foster children, or come from military families are more likely than others to switch schools one or more times.[6] Moreover, high-poverty urban schools can have up to half of their students change schools within a single year.[7] Data portability is important for these students so that they do not lose access to valuable data about their learning when they switch schools.

UTILITY PROVIDERS

---

[4] Joshua New, "Building a Data-Driven Education System in the United States" (Center for Data Innovation, November 15, 2016) http://www2.datainnovation.org/2016-data-driven-education.pdf.
[5] Sarah D. Sparks, "Student Mobility: How It Affects Learning," *Education Week*, August 11, 2016, https://www.edweek.org/ew/issues/student-mobility/.
[6] Ibid.
[7] Ibid.

Despite the increasing prevalence of smart meters—electronic devices that collect granular data about a consumer's electricity, gas, and water usage habits—many consumers are unable to access this data in a timely manner or in useable formats, if at all.[8] In 2011, the federal government led a successful call to action for electric utilities to voluntarily join the "Green Button" initiative, a project to provide consumers direct access to their energy usage data in a standardized format by clicking a uniformly branded green button on their utilities' websites.[9] The Green Button initiative created a standard for reporting and exchanging utility usage data among providers, third-party developers, and consumers. Thus, many consumers can easily analyze data about their own electricity consumption and take steps to reduce usage, conserve resources, and lower utility bills. However, there is no equivalent program for water utility data.[10] By making a high-profile call to action to create a similar program for water utility data, consumers could access the same benefits to efficiency and lower costs, and third parties could build similar services for water utility analytics that they have for electricity data.

3. SPECIFIC STEPS THE FEDERAL GOVERNMENT, PRIVATE COMPANIES, ASSOCIATIONS, OR OTHERS MIGHT TAKE TO ENCOURAGE OR REQUIRE GREATER DATA PORTABILITY

When companies do not voluntarily provide customers access to their own data in a reusable, electronic format, policymakers may need to intervene.[11] In most cases, the government should not mandate data portability, but instead foster data access through voluntary agreements and by raising public awareness about the importance of sharing data.[12] The federal government's achievements at promoting electricity data portability with the Green Button Initiative is an example of this type of effort. In other cases, this intervention may include specific requirements

---

[8] Daniel Castro and Brandon De Bruhl, "How to Promote Smarter Water Use by Giving Consumers Access to Their Consumption Data" (Center for Data Innovation, September 7, 2015), http://www2.datainnovation.org/2015-water-data-green-button.pdf.

[9] "Green Button Alliance," 2012, http://greenbuttonalliance.org/history/.

[10] Daniel Castro and Brandon De Bruhl, "How to Promote Smarter Water Use by Giving Consumers Access to Their Consumption Data" (Center for Data Innovation, September 7, 2015), http://www2.datainnovation.org/2015-water-data-green-button.pdf.

[11] Daniel Castro and Travis Korte, "Data Innovation 101: An Introduction to the Technology and Policies Supporting Data-Driven Innovation" (Center for Data Innovation, November 4, 2013), http://www2.datainnovation.org/2013-data-innovation-101.pdf.

[12] Daniel Castro and Travis Korte, "Data Innovation 101: An Introduction to the Technology and Policies Supporting Data-Driven Innovation" (Center for Data Innovation, November 4, 2013), http://www2.datainnovation.org/2013-data-innovation-101.pdf.

for data sharing, such as in highly-regulated industries. This does not mean that companies should be required to give up ownership of their data—and they should certainly not be required to give up proprietary, non-customer data--only that they should provide customers with copies of their own data. In all cases, the costs of enabling data portability should be considered against the benefits. These costs include not only the direct costs of providing the data, but also the indirect costs to a company if it has lost exclusive access to data to which it has added value.

In health-care, policymakers should work to prevent health-care providers from engaging in information blocking. While the Office of the National Coordinator for Health Information Technology (ONC) and the Federal Trade Commission (FTC) have some ability to investigate allegations of information blocking, it is insufficient to solve the problem. In October 2015, bipartisan members of the Senate introduced the Transparent Ratings on Usability and Security to Transform Information Technology (TRUST IT) Act to expand the authority of ONC and the Department of Health and Human Services (HHS) to investigate and crack down on bad actors.[13] Congress should enact the TRUST IT Act, and HHS, ONC, and FTC should aggressively pursue allegations of information blocking that limit data portability and reduce the efficiency of the health-care system.

In education, federal and state departments of education should engage school districts to educate them about the importance of education data portability and encourage them to develop student data portals that provide access to their data. This could include developing software toolkits, sharing best practices, or incentivizing their development, such as through grant funding. At the federal level, OSTP and the Department of Education should continue to pursue the development of the "MyData Button," an initiative they announced in January 2012 to improve the portability of education data that has not made meaningful progress.[14] The MyData Button would allow students to download all of their education data in a common, machine-readable format with the click of a button, and the Department of Education and several major technology companies have committed to supporting its development and implementation.[15] However, the MyData Button still does not exist four years later, due to a variety of technical and

---

[13] Joshua New, "Congress Has a New Cure to Stop Companies from Blocking Patient Data," Center for Data Innovation, November 10, 2015, https://www.datainnovation.org/2015/11/congress-has-a-new-cure-to-stop-companies-from-blocking-patient-data/.
[14] "Fact Sheet: Unlocking the Power of Education Data for All Americans," Office of Science and Technology Policy, January 19, 2012, https://www.whitehouse.gov/sites/default/files/microsites/ostp/ed_data_commitments_1-19-12.pdf.
[15] Ibid.

political obstacles.[16] The federal government should recognize the substantial benefits of increasing the portability of education data and increase its efforts to implement the MyData Button. To achieve this, the Department of Education should make it a policy that its grants can only be used to purchase education technology products and services that provide students access to their own data.

For utility data, just as the federal government spurred the development of the Green Button initiative for electricity, it should do the same for water utility data.[17] Due to the similarities between the business models and technologies involved in both water and electricity metering, the success of the Green Button initiative gives policymakers an easy model to follow for increasing the portability of water utility data.

The federal government should also work to identify other areas were policies to promote data portability would benefit consumers and foster innovation. In general, consumers should expect that if they have lawful access to their own data, then they should be able to empower a third-party to access it on their behalf. For example, banks should not prohibit third-party apps and services that help consumers manage their finances, such as Mint, Yodlee, LearnVest, and Personal Capital, from accessing their customers' data on their behalf. When necessary, regulators such as the Consumer Financial Protection Bureau should intervene to ensure that banks for not engage in information blocking activities.[18] In addition, the federal government, such as FTC or the National Telecommunications and Information Administration (NTIA), should consider developing a model data portability policy, so companies can clearly disclose whether they offer data portability and consumers can easily compare practices across different companies.

Policymakers should also avoid unnecessarily restrictive regulations on the collection and sharing of data. When restrictions on use are necessary they should be implemented with restraint. Legal rules preventing the use of data can lead to a situation known as the "tragedy of the anticommons." This occurs when the existence of too many legal and bureaucratic barriers

---

[16] Glynn Ligon, "MyDataButton: Button, Button, Who's Got the Button?," ESP Solutions Group, May 2016, http://www.arniedocs.info/wp-content/uploads/2016/05/MyData-Button-Disappears-2016-05-03.pdf.
[17] Daniel Castro and Brandon De Bruhl, "How to Promote Smarter Water Use by Giving Consumers Access to Their Consumption Data" (Center for Data Innovation, September 7, 2015), http://www2.datainnovation.org/2015-water-data-green-button.pdf.
[18] Liz Weston, "Why Banks Want You to Drop Mint, Other 'Aggregators'," *Reuters*, November 9, 2015, http://www.reuters.com/article/us-column-weston-banks-idUSKCN0SY2GC20151109.

create high transaction costs that restrict the use and exchange of data. For example, uncertainty over data ownership may prevent a company from creating a useful data-driven application. To avoid undermining beneficial applications of data, policy discussions should focus on resolving how data can be used, rather than on deciding whether it can be collected and exchanged. Uses that result in specific harm should of course be prohibited, but policymakers should craft open-ended policies acknowledging the unpredictable breadth of future data-driven applications, particularly in the health and education sectors.[19]

In summary, the federal government can play an important role in encouraging data portability, but these efforts should largely be voluntary. In some high regulated industries, such as finance or health care, explicit data portability requirements may be appropriate.

4. BEST PRACTICES IN IMPLEMENTING DATA PORTABILITY

For data to be truly portable, it should be available in an open, machine-readable format, and it should be licensed to allow for reuse. Usually, this means making data available under an open license, with no restrictions on how it can be used, however, if the data is a consumer's to begin with, a company should not attempt to apply a new license to that data when it returns it to the user.

5. ADDITIONAL INFORMATION RELATED TO DATA PORTABILITY POLICYMAKING THAT OSTP SHOULD CONSIDER

Policymakers should recognize that data portability is good for competition and keeps switching costs low for consumers, but also that data portability only works when data can be both exported from existing services and imported into new services. Additionally, policymakers should be aware that, in some cases, data portability is not feasible. For example, it would be unrealistic to expect social networking companies to make all their consumer data portable because, by its very nature, one user's social network data is related to other users' data. For example, the "likes" on a post have significantly less relevance without detailed data about who created those likes and their relationships with other users. In these cases, limited data portability, such as the ability to download photos or videos, might make more sense. In addition, the larger policy question is whether third-parties should have access to the data platform, such as to offer additional services

---

[19] Daniel Castro and Travis Korte, "Data Innovation 101: An Introduction to the Technology and Policies Supporting Data-Driven Innovation" (Center for Data Innovation, November 4, 2013), http://www2.datainnovation.org/2013-data-innovation-101.pdf.

or perform research. In such cases, policymakers should focus on ensuring that these companies do not unfairly limit who can access their network in ways that harm consumers and hurt competition. For example, they might encourage these companies to provide a transparent and consistent set of rules for developers to build applications.

CONCLUSION

It is encouraging to see OSTP working to understand the benefits of data portability and identify opportunities for policymakers to increase the ability of companies and consumers to access and share their data. There are several concrete actions that OSTP or the federal government can take to increase data portability, which would increase competition, empower consumers, and promote innovation.