April 26, 2017

Re: Response to the European Commission Questionnaire on Building the European Data Economy

Dear Sir or Madam,

On behalf of the Center for Data Innovation (datainnovation.org), we are pleased to submit these answers in response to a request for comments in response to the European Commission's Questionnaire on Building the European Data Economy.

The Center for Data Innovation is the leading think tank studying the intersection of data, technology, and public policy. With staff in Brussels and Washington, 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 artificial intelligence, predictive analytics, open data, cloud computing, and the Internet of Things. The Center is a nonprofit, nonpartisan research institute affiliated with the Information Technology and Innovation Foundation.

The Center is strongly of the view that there is no reasonable justification for data localization of any kind, for personal or non-personal data, and that the practice eliminates competition between domestic and foreign cloud services providers, raises costs and prices, and stifles innovation. On the matter of access and reuse rights of non-personal machine data, the Center favors the reuse of data by many actors as possible, including both owner-operators and manufacturers. Public policy should support such re-use through proper guidelines, but should not force it with regulations that close down mutually beneficial contractual agreements on the use of machine data. By and large, the Center takes the view that insofar as existing liability rules protect consumers, contractual relationships between producers (such as component manufacturers and software developers) should, in most cases, be sufficient for establishing liability between them in case of damages, and that disputes should be resolved in court on an ad-hoc basis.

Yours faithfully,

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

Nicholas Wallace
Senior Policy Analyst
Center for Data Innovation
nwallace@datainnovation.org


CENTER
FOR
DATA
INNOVATION

1. LOCALIZATION OF DATA FOR STORAGE AND/OR PROCESSING PURPOSES

The main objective of this part of the questionnaire is to get detailed insights into the extent, nature and impacts of data localization restrictions within the EU and what could constitute limited, justified grounds for such restrictions without unduly jeopardizing the free movement of data within the EU (except for restrictions to the free movement of personal data for reasons connected with the protection of natural persons with regard to the processing of personal data. The Treaty on the Functioning of the European Union and the General Data Protection Regulation (GDPR) establish the free flow of personal data within the EU and set out the rules relating to that free movement).

Another important aspect is to find out to what extent businesses store or process data in multiple geographical locations within the EU and what are the reasons for this multiple location and, respectively, local storage or processing. The Commission also seeks respondents' views on the perceived impacts of the removal of data localization restrictions within the EU. The Commission welcomes replies particularly from businesses, including SMEs, and public sector organizations.

Do you know about legislation or administrative rules or guidelines (including those adopted in the context of public procurement) requiring to store or process data in your or other EU countries (please see part 2 of the Staff Working Document linked to on the consultation webpage for the summary of data localization restrictions identified so far)?

-Yes

If yes, please specify:

Legislative requirements and administrative rules

If yes, the legislation, administrative rules or guidelines concern:

-Personal data for reasons other than the protection of natural persons with regard to the processing of personal data,
-Business privately-held data
-Non-personal publicly-held data

Is your business or organization required to comply with any of the measures?

-No

Please describe briefly the requirement

There are multiple regulations in EU member states requiring data to be localized. For example, Germany's data retention law requires telecoms firms to store meta-data in Germany, and Luxembourg's financial transparency laws require banks to store financial data in Luxembourg. French procurement rules make it illegal to store administrative data from local or national government bodies on a "non-sovereign" cloud server (that is, one not located in France).

Is there any impact of such a measure, notably on your business or organization?

We are a public policy think tank, we research the effects of policies on economies and actors other than ourselves. Data localization laws inhibit international competition in cloud services, which in turn diminishes the impetus to lower prices and improve services. For both large and small firms that rely on these services, that raises costs and reduces opportunities, and the costs are further exacerbated when they have to invest in data storage in multiple countries. That in turn inflates their prices, stifles their product innovation, and makes it costlier for them to enter new markets. The overall effect is a less competitive and less innovative economy with inflated prices and diminished choices for consumers.

For your own organization's purposes, do you store or process your data in multiple locations within the EU?

-Yes

If you answered yes, what are the main reasons?

We use reputable, mainstream cloud services. We are unconcerned about where the providers have their data centers.

In your opinion, should data localization restrictions be removed within the EU?

-Yes

In your opinion, what grounds would justify keeping data localization restrictions within the EU?

There are no justifications for data localization.

If you answered yes, how would the removal of the localization restrictions be beneficial to your business or organization?

We participate in this debate as an independent think tank, we do not have a vested interest in the outcome.

What kind of action at EU level do you consider appropriate to address the restrictions?

Legislation against data localization and commitments to oppose data localization in free trade agreements

2. ACCESS TO AND RE-USE OF NON-PERSONAL DATA

This part of the questionnaire aims to understand the data trading practices of businesses, and how all businesses, in particular SMEs, and other stakeholders access and trade non-personal data, and what are the perceived barriers to such trading and re-use of such data. The Commission seeks the views of businesses and other respondents on ways to enhance access to and re-use of data and data trading in Europe today.

2.1. ACCESSING DATA

This section is addressed to businesses and organizations of any size, and especially SMEs and start-ups which are seeking access to non-personal or anonymized data for running their businesses or developing new businesses. For consumer access issues, please see section 4.1 on data portability for non-personal. The aim is to find out whether and to what extent businesses and organizations have access to the data they need to develop or conduct their tasks, and furthermore to find out what role existing legislation plays in today's data markets, and whether there is a need to revise or introduce legislation to support the European data economy.
Do you currently depend to a significant extent on data resources that you acquire from others (for products or services you offer, for your internal business processes)?

-No

Have you had difficulties in acquiring data from other business actors (i.e. limited or no access to the data) or have you been exposed to business practices that you consider unfair with respect to access to such data?

-No

Does current competition law and its enforcement mechanisms sufficiently address potentially anti-competitive behavior of companies holding or using data?

-To a great extent

Have you entered contracts in which certain data was defined as a trade secret?

-No

2.2. HOLDING AND SUPPLYING DATA

This section is addressed mostly to businesses that hold non-personal or anonymized data not subject to significant data processing ("raw" data), in particular data collected by sensors embedded in machines, tools and/or devices and who are in a position to share them. The aim is to get more information about data licensing practices.

Do you believe existing EU legislation sufficiently protects investments made into data collection by sensors embedded in machines, tools and/or devices?

-Yes

Are you including the value of at least some of the data you hold as a business asset in your balance sheets?

-No

Please explain why.

-This is not required by the applicable accounting/financing reporting standards
-I am not sure how to measure the value of the data I have

2.3. POSSIBLE SOLUTIONS

Sections 2.3.1 and 2.3.3 are directed at all respondents, including consumers and businesses. Section 2.3.2 is directed at businesses that deal with data collected by sensors embedded in machines, tools and/or devices. The aim is to receive input on what a possible future EU framework should look like to support a thriving, diverse and innovative European data economy.

2.3.1. GENERAL OBJECTIVES FOR A FUTURE EU FRAMEWORK FOR DATA ACCESS

To what extent do you agree with the following statements (1=not at all,2=to a minor extent, 3=neutral/I don't know, 4=to some extent, 5=to a great extent):

	1	2	3	4	5
Trading of non-personal machine-generated data should be enabled to a greater extent than it is today.					X
The sharing of non-personal machine-generated data should be facilitated and incentivized.					X
Investments made into data collection capabilities and data assets should be protected.				X	
Sensitive business and confidential data should always be safeguarded.				X	
Lock-in effects in the data market should be minimized, especially for SMEs and start-ups.				X	


CENTER
FOR
DATA
INNOVATION

2.3.2. ACCESS FOR PUBLIC SECTOR BODIES AND SCIENTIFIC RESEARCH

Could you agree to an obligation to license the use of (non-personal) data you hold for any of the following purposes (subject to conditions)?

-For the establishment of statistics by public statistical offices
-For government agencies for the prevention of public health or other specified risks
-For government agencies in order to address other societal challenges (e.g. improving urban planning, manage supply of energy)
-For scientific research that is funded from public resources
-For factual information that is publicly available

Could such access be given at no cost or only the cost related to making the data available?

-Yes

Do you consider there should be action at EU level to address access to such data for the entities mentioned in the previous question (the establishment of statistics by public statistical offices, government agencies for the prevention of public health or other specified risks, government agencies in order to address other societal challenges (e.g. improving urban planning, manage supply of energy), scientific research that is funded from public resources)?

-Yes, through legislative measures (for a scope to be defined)

2.3.3. ACCESS FOR OTHER COMMERCIAL ENTITIES

The following questions ask for an assessment of a number of potential measures that might help to make more data held by one commercial entity available for re-use by another commercial entity.

Would you agree with the following statement: More data would become available for re-use if the Commission would issue guidance on how access, use and re-use of data should be addressed in contracts (data usage licenses) – based on existing legislation (in particular the Trade Secrets Protection Directive, copyright legislation and the Database Directive)?

-Sometimes

What impacts (if any, including economic) on competition and innovation would you expect from the solution described in the previous question?

The solution is not described in sufficient detail to support a fair assessment of how it might impact competition and innovation. Clear guidance on who should have access to what data in contracts may help to free up data in some instances, such as public-private partnership investments in the Internet of Things, where confusion between data as a publicly-funded good and as a valuable investment asset can limit the supply of data and stifle investment. But this would depend on precisely what those guidelines are, and in any case, the guidelines would not free-up data where companies believe (justifiably or not) that it is not in their interests to share data.

Would you agree with the following statement: The optimal solution for making data collected by sensors embedded in machines, tools and/or devices available for re-use is to leave it entirely to the parties to decide (by contract) who should have the right to license the usage of these data, how and to whom.

-Sometimes

What impacts (if any, including economic) on competition and innovation would you expect from the solution described in the previous question?

With clear guidelines, some companies will be able to establish mutually beneficial agreements with other companies about the exchange of machine data, without the need for regulation. However, not all data-holders will see an incentive to release the data, which would have negative implications for consumers. For example, rail operators may refrain from opening-up live data from rail sensors, as this could expose poor performance in the network. Such reluctance would reduce opportunities for innovation and competition in the transport sector by reducing the supply of available data.

Would you agree with the following statement: More data would become available for re-use if more data holders used Application Programming Interfaces (APIs) to facilitate access to the data they hold, and these APIs were designed and documented in a way easy to use by third party application developers.

-Yes

What impacts (if any, including economic) on competition and innovation would you expect from the solution described in the previous question?

In general, providing more data through APIs would create more opportunities for re-use of that data, particularly among large datasets that change frequently, and encourage data innovation. However, the extent to which APIs provide value depend not only on their design, but the policies that govern them. For example, some APIs may only be available for non-commercial purposes or

impose unreasonable limitations on the number of times a developer may use the API in a given time period, thereby restricting certain commercial uses of the data.

Would you agree with the following statement: More data would become available for re-use if legislation would define a set of (cross-sector or sector-specific) non-mandatory contract rules for B2B contracts, possibly coupled with an unfairness control in B2B contractual relationships) for allocating rights to access, use and re-use data collected by sensors embedded in machines, tools and/or devices were defined.

-Yes

What impacts (if any, including economic) on competition and innovation would you expect from the solution described in the previous question?

If designed to encourage data sharing, such as by making the default standard for contracts to allow for data reuse, this would "nudge" companies to favor data sharing unless they choose to opt out. This would help establish a baseline norm without creating "one size fits all" rules with unintended consequences for particular actors. Increasing the supply of data helps fuel innovation.

Would you agree with the following statement: More data would become available for re-use if a set of recommended standard contract terms were to be drafted in close collaboration with stakeholders.

-Yes

What impacts (if any, including economic) on competition and innovation would you expect from the solution described in the previous question?

This is also a good idea, provided the contracts are non-mandatory. Contracts could be geared towards data innovation, but sensitive to the concerns of different stakeholders. However, even if the contracts do become industry standards, this still does not mean they are genuinely a single size to fit all, and their terms may not necessarily be equally advantageous to both partners in all cases. As mentioned before, the diversity of interests and power imbalances between firms, though not problems in themselves, are factors to consider in this type of solution.

Would you agree with the following statement: More data would become available for re-use if a company holding data which it protects through technical means against illicit misappropriation had civil law remedies against such misappropriation (e.g. the right to seek injunctions, market exclusion, or to claim damages).

-Sometimes

What impacts (if any, including economic) on competition and innovation would you expect from the solution described in the previous question?

If it is raw machine data, it is hard to see how any reuse could constitute harmful misappropriation. But in any case, whether companies have civil law remedies depends on how they license the data, and the terms in the contract accompanying that license. If the intention of the question is to ask whether or to what extent data access rights should be accompanied with reuse constraints, then the answer is that any limit on the uses of lawfully-accessed data necessarily stifles opportunities for innovation, and opportunities to reuse the data for competitive purposes.

Would you agree with the following statement: More data collected by sensors embedded in machines, tools and/or devices would become available for re-use if both the owner or user of the machine, tool or device and the manufacturer share the right to license the use of such data.

-Yes

What impacts (if any, including economic) on competition and innovation would you expect from the solution described in the previous question?

Owner and manufacturers likely have different incentives for sharing data. For example, it would likely be difficult and costly to obtain aggregate sensor data from each owner of a particular product by manually locating each owner and getting them to complete the necessary technical and legal steps to begin sharing the information with a third-party. In contrast, if the manufacturer had access to data from all of its products, obtaining this aggregate information would be simple. Conversely, a manufacturer may have little incentive to provide to each customer specific data from their product, but this information may be highly valuable to individual users. If both owners and manufacturers had the right to license the use of the data, this could fuel more research, product innovation, and competition.

Would you agree with the following statement: More data would become available for re-use if the companies active in the production and market commercialization of sensor-equipped machines, tools or devices were awarded an exclusive right to license the use of the data collected by the sensors embedded in such machines, tools and/or devices (a sort of sui generis intellectual property right).

-No

What impacts (if any, including economic) on competition and innovation would you expect from the solution described in the previous question?

Policymakers' goal should be to maximize re-use and sharing of data by as many stakeholders as possible, especially competing stakeholders, because this supports socially and economically beneficial innovation and competition. The proposed policy would harm competition and innovation by denying users and owners of sensor-equipped machines the right to reuse and share the data generated by the goods they have purchased.

Would you agree with the following statement: More data would become available for re-use if the persons or entities that operate sensor-equipped machines, tools or devices at their own economic risk ("data producer") were awarded an exclusive right to license the use of the data collected by these machines, tools or devices (a sort of sui generis intellectual property right), as a result of the data producer's operation, to any party it wishes (subject to legitimate data usage exceptions for e.g. manufacturers of the machines, tools or devices).

-No

Please explain.

Policymakers' goal should be to maximize re-use and sharing of data. The proposed policy would needlessly reduce the supply of available data by limiting the exclusive right to share it to one actor. Machine operators would be unlikely to initiate data sharing without some incentive—such as a subscription fee—thereby raising the cost of data as an input unnecessarily, which would stifle downstream innovation with this data.

What impacts (if any, including economic) on competition and innovation would you expect from the solution described in the previous question?

This would harm product innovation by denying manufacturers the ability to analyze information coming from their products in order to improve them. It would also make data-driven research of how certain machinery functions and is used more difficult and expensive. The results could even be dangerous, as it would become harder for third-parties, including government regulators, to identify flaws and bugs.

Providing machine owners the ability to share data with third parties, would support competition by allowing firms to collect data about their competitors' products and learn from the data to make improvements thereby fostering competition.

To what extent would you agree to an obligation to license for the re-use of data generated by machines, tools or devices that you have commercialized under fair, reasonable and non-discriminatory (FRAND) terms?

-Not at all

To what extent would you agree to an obligation to license for the re-use of data generated in the context of your online platform through its users under fair, reasonable and non-discriminatory (FRAND) terms?

-Not at all

3. LIABILITY

This part of the questionnaire aims to understand the level of awareness, as well as the respondents' experiences and issues related to liability for products and services coming out of Internet of Things (IoT) technologies and autonomous systems. The questions are also meant to gather evidence for a proper assessment of the adequacy of the Product Liability Directive (85/374/CEE) to respond to IoT and robotics liability challenges. The Commission seeks the views of producers and users of IoT technologies and autonomous systems in this section.

3.1. EXTRA-CONTRACTUAL LIABILITIES: IOT AND ROBOTICS PRODUCTS AND SERVICES

As a consumer, what (if anything) makes you reluctant to buy IoT/robotics products or services?

-No reluctance at all

Do you think IoT/robotics products and services should be equipped with an event data recorder to track what the device was doing when the damage occurred?

-No

In your opinion, who should bear the liability in case of damages caused by defects or malfunctioning of a smart device which combines tangible goods (a car), digital goods (an app) and services (e.g. data services)?

See answer to question 3.2

3.2. POSSIBLE OPTIONS AND A WAY FORWARD (BOTH FOR CONSUMERS/END USERS AND PRODUCERS OF IOT/ROBOTICS DEVICES)

Do you think a risk management approach in which the party that is best placed to minimize or avoid the realization of the risk (e.g. the manufacturer of the IoT device, or the software designer) could be a way forward?

-I don't have information about what a risk management approach would entail

In your opinion, who should bear the liability in case of damages caused by defects or malfunctioning of a smart device which combines tangible products, digital products and services?

Producers should not be liable for harms that result from changes they do not make or authorize. They should be liable for damages arising from faults in the unmodified product components

they produce. When the source of the fault cannot be determined, the attribution of liability is better handled through contracts on a case-by-case basis, without prejudice to consumer protections in the Product Liability Directive.

What type of liability, contractual or extra-contractual, is, in your opinion, the most consumer-friendly way to deal with damages caused by defects or malfunctioning in smart devices, which combine tangible products, digital products and services?

-Contractual and extra-contractual

Do you think that the liability in relation to smart devices combining products and services require an ad hoc approach at EU level?

There should be basic consumer protections and product safety regulations, which must inevitably include some basic principles about liability for damages, as explained in the previous written answer. But policymakers should focus on providing guidelines that ensure concepts such as damages, negligence, and breach of warranty are clearly-defined, which will allow cases to be settled individually, on their own merits.

Independently of who is considered liable, should there be a liability cap, i.e. an upper bound to the compensation of damages?

-Yes, for all IoT products

What is your opinion on the idea of best practices guidelines and/or expected care and safety standards that, if fulfilled, would automatically exclude/limit liability?

-I agree, for all IoT products

Is there a need for mandatory cyber insurance?

-No

Do you feel protected by the current legal framework (both Business-to-Business and Business-to-Consumer) for algorithms, e.g. in case it can be proven that an accident has been caused by a bug in the algorithm?

-Yes

Please explain.

There is no obvious reason why a bug in an algorithm could not be considered a product fault like any other, meaning the customer should be protected under the Product Liability Directive.

Should some sorts of standard certification or testbedding be envisaged for algorithm based services?

-No

Please explain.

Just because a product or service uses algorithms does not mean it carries some elevated risk that demands special certification. As long as liability is attributed correctly and consistently when do damages occur, then certification would be a pointless bureaucratic hurdle.

Who should be liable for defects or accidents caused by products embedding open algorithms, i.e. algorithms developed through cooperative platforms?

-The producer

4. PORTABILITY OF NON-PERSONAL DATA, INTEROPERABILITY AND STANDARDS

4.1. PORTABILITY OF NON-PERSONAL DATA

This section is directed towards all respondents, including consumers, organizations and businesses. The objective of this section is to explore business situations where portability of non-personal data can unlock opportunities and/or eliminate blockages in the data economy, as well as the effects of such conditions on all the concerned actors.

Are you using or have you used services which allow you to port or retrieve non-personal data that you had previously provided?

-No

If you are a business user of cloud services or online platforms: Have you experienced difficulties in switching providers?

-No

Do you see a specific need for businesses to receive non-personal data in a machine-readable format, as well as the right to license the use of such data to any third party (i.e. the right of data portability under article 20 GDPR extended to any user and to non-personal data)?

-Yes

What types of data should be covered by a portability right, in your view?

-Non-personal data submitted to cloud service providers
-Non-personal data submitted to online platforms
-Data generated by sensor-equipped machines, tools and/or devices

What types of entities should be the beneficiaries of such a portability right?

-All businesses

If you have further comments on portability rights, please insert them below.

GDPR Article 20 is a good basis for establishing data portability rules for non-personal data. It distinguishes between observed and inferred data, thereby balancing the right of the data subject to data portability with the right of the data processor to intellectual property. The same model for non-personal data balances IP protection with the shared value of machine data to the economy. In fact, as personal and non-personal data will often be ported together, that creates a

strong argument to say that the regulations governing data portability should be the same in either case.

What are the possible effects of introducing a portability right for non-personal data regarding cloud services? Please consider positive and possible adverse effects, and consequences for your business and, more generally, for the user of the cloud service as well as the service provider and other concerned actors.

Data portability in cloud services could intensify competition between cloud providers by making it easier for users to switch between them. As cloud computing is an essential utility for data-driven companies, anything that increases competition and lowers prices in this market is likely to support data innovation, business growth, and the creation of new businesses.
When businesses use cloud services simply to store data in traditional file formats, then porting the data is a relatively simple task. But when businesses use cloud platforms to manage business processes and to process data using specialized applications, switching from one platform to another presents a technical challenge, as not all cloud platforms operate in the same way. A data portability law should not penalize providers of complex services or new services where data standards do not exist, as this would create a disincentive for services to move to the cloud. In cases where cloud providers cannot or will not enable data portability, they should disclose this fact to customers to ensure transparency in the market.

What are the possible effects of introducing a portability right regarding non-personal data generated by sensor-equipped machines, tools and/or devices? Please consider positive and possible adverse effects, and consequences for your business and, more generally, for the user of the services as well as manufactures, service providers and other concerned actors.

A data portability right for the users of sensor-equipped machines would allow them to take that data to third parties in order to access new services, which could fuel new economic activity and innovation. It would also help prevent vendor lock-in and allow more competition. This would also support efforts to benchmark and compare competing products, because independent testers could access the raw data.

What are the possible effects of introducing a portability right for non-personal data regarding online platforms? Please consider positive and possible adverse effects, and consequences for your business and, more generally, for the business user of the platform, consumers, intermediary (data) services, the online platform and other concerned actors.

The answer to this question is similar to that for the question about cloud services: online platforms have huge differences between them. While platforms should not prevent users from shifting their data to another platform, and that data should ideally be in an open, machine readable format, it should not be incumbent on the platform to re-structure the data in a way that suits the makeup of a competitor.

4.2. INTEROPERABILITY AND STANDARDS

This section is primarily directed towards businesses and organizations. The objective of this section is to get the stakeholders' opinions on the best approaches to technically support data portability and access to data.
What do you consider as a priority for facilitating access to data and to improve its technical and semantic discoverability and interoperability?

-Commonly-used machine-readable file formats.

What technical instruments should be used for promoting/implementing your priorities suggested in the previous question?

-Recommendations

What legal instruments should be used for promoting/implementing your priorities suggested in the same question?

-Support actions

Do you see the need for the definition of a reference architecture recommending a standardized high-level framework identifying interoperability interfaces and specific technical standards for facilitating seamless exchanges across data platforms?

-No

Please explain.

This is not a job for policymakers. It is up to companies to work out how they incorporate data ported to them by their customers from elsewhere.