Making Europe a Data Economy: A New Framework for Free Movement of Data in the Digital Age

By Paul Hofheinz and David Osimo

New Economy

Single Market

Data

Innovation

Table of Contents

Lisbon Council Policy Brief

Making Europe a Data Economy: A New Framework for Free Movement of Data in the Digital Age
By Paul Hofheinz and David Osimo

Paul Hofheinz is president and co-founder of the Lisbon Council. He holds degrees from the London School of Economics and Yale University, where he received the prestigious John Addison Porter and Dean Henry P. Wright prizes.
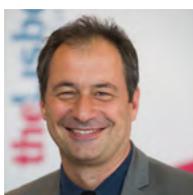
David Osimo is director of research at the Lisbon Council. He holds degrees from University of Milan and University of Cardiff.

REPUBLIC OF ESTONIA
GOVERNMENT OFFICE

European Union
European Social Fund

Investing in your future

The opinions expressed in this policy brief are those of the authors alone and do not necessarily represent the views of the Lisbon Council, the government of Estonia, the European Commission or any of their associates.

I. Why Free Movement of Data is Necessary for Europe and the Digital Single Market

Data, some say, is "the new oil."[1] Others say it's "the new currency."[2] Still more call it "the infrastructure" on which modern commerce will travel.[3] And some even believe that it already counts as one of the three principal factors of production: "labour + capital + data = economic growth."[4]

What, then, is data?

To be sure, in the age of machines, data is the raw material out of which our most potent and powerful economic and social achievements will come. Data is how we come to know ourselves. It is how our businesses communicate. It is how we predict and think about the future. Its dots and blips are what entertain and inform us – and preserve our most precious cultural jewels. In an age where the Internet has united the world around the miracle of instantaneous, zero-marginal-cost global communication, at a time when machines themselves are starting to augment human intelligence in ways that promise to yield some of history's most exciting discoveries, data is the key commodity underlying every transaction and fuelling every new insight.[5] It is the raw material from which our future knowledge will be mined and the likely site of a coming gold rush that will put the 1849 American prospectors to shame.[6]

And its role is set to rise as the economy moves inexorably towards new products and services based on artificial intelligence (AI), machine-learning and other high-performance-computer-driven processes whose vast social utility are hard to calculate today. AI and machine-generated learning will need more and more aggregated data to be effective – data sets that can only be brought together through enhanced European cooperation and a greater emphasis on sharing and exchange.[7]

---

1   The Economist, "Briefing: The Data Economy," *The Economist*, 06 May 2017; Michael Haupt, "'Data is the New Oil' – a Ludicrous Proposition: Natural Resources, The Question of Ownership and the Reality of Big Data," *Medium*, 02 May 2016.

2   William D. Eggers, Rob Hamill and Abed Ali, "Data as the New Currency: Government's Role in Facilitating the Exchange," *Deloitte Review*, No. 13, 2013.

3   Organisation for Economic Co-operation and Development, *Data-Driven Innovation: Big Data for Growth and Well-Being* (Paris: OECD, 2015).

4   The idea is more fully developed in Paul Hofheinz and Michael Mandel, "Uncovering the Hidden Value of Digital Trade: Towards a 21st Century Agenda of Transatlantic Prosperity," *Interactive Policy Brief 19/2015* (Brussels and Washington DC: Lisbon Council and Progressive Policy Institute, 2015).

5   Viktor Mayer-Schönberger and Kenneth Cukier, *Big Data: A Revolution that Will Transform How We Live, Work and Think* (London: John Murray, 2013). See, also, Ray Kurzweil, *The Singularity is Near: When Humans Transcend Biology* (London: Duckworth, 2005).

6   McKinsey Global Institute estimates that rising global data flows have boosted world GDP by more than 10%. Data flows, by MGI's count, have risen 45 fold since 2005, and now account for more than $2.8 trillion (€2.27 trillion at the 2014 exchange rate) of global GDP – a larger impact on global growth than traditional good flows. See James Manyika, Susan Lund, Jacques Bughin, Jonathan Woetzel, Kalin Stamenov and Dhruv Dhingra, *Digital Globalisation: The New Era of Global Flows* (San Francisco: McKinsey Global Institute, 2016).

7   Paul Hofheinz, *Artificial Intelligence and Machine Learning: Opportunity and Challenge* (Brussels: Lisbon Council Discussion Paper, 2016).

'We need a new framework for data access, use, reuse, storing and mining that looks at – and adequately answers – the very real challenges of the future.'

This is why the "free movement of data" has become so important – European people, companies, even countries can't thrive in this environment on their own.[8] They need to come together and create access to the large data sets that will allow analytics to propel and keep Europe at the forefront of the data-driven economy. Put simply, we need a new framework for data access, use, reuse, storing and mining that looks at – and adequately answers – the very real challenges of the future. In this policy brief, we will analyse the leading proposals in this field, and propose a new framework in Section IV.

For Europe – slowed as it is by a decade of disappointing economic growth, facing an ageing population that will strain public resources, challenged by politically revanchist movements, and still confused over why the Internet-era's most innovative companies all seem to originate from the United States – the stakes could not be higher. The countries and regions that embrace and develop the most advanced data analytics – that build up the "data markets" with access to sufficient amounts of good, high-quality data behind them – will have the best, most advantageous services of tomorrow – and the strongest platform for solving the great social challenges that have perhaps defied solution in the past. In an age where individuals, companies,

even countries themselves compete not just on cost but on relevance, insight, speed and creativity, the winners will come from those who most effectively and enthusiastically unlock this coming explosion of data, data analytics-driven services and machine-learning-enhanced outcomes. We truly live in an information age. And it would be to Europe's competitive advantage to unchain and facilitate that.

And yet, the use, reuse, transfer and even the most basic gathering and storing of data are often difficult and subject to controversy, especially in Europe. The famous "single market" intended to weave fractious countries into a seamless whole, capable of giving businesses the economies of scale they need and driving forward innovation, remains a largely incomplete project. Data, for better or worse, is still divided into too many national silos, and even into silos within those silos, where the legal use or potential reuse of this data is often vague and ill-defined. Governments, companies and individuals often refuse to share the data they have despite the potential opportunities, either because they fear the data will be misused, they worry about potential violations in an increasingly complex legal environment or they consider the data proprietary.[9] This poses no immediate problem, *prima facie*.

---

8   European Commission, *Building a European Data Economy* (Brussels: European Commission, 2017).

9   A recent European Commission-funded study showed that in 87% of the 100 cases surveyed companies do not share data among themselves. In most cases, data-holding companies subcontract data-analytic services to third parties. These data-analytic companies can access the data, analyse it and provide results, often enriched with third-party data such as social media or mobility data, but crucially they can't reuse the data to build new products. New data products are typically developed in-house. See Laia Pujol, David Osimo, Jonathan Wareham and Federica Porcu, "Data-Driven Business Models in the Digital Age: The Impact of Data on Traditional Businesses," paper presented to the 3rd World Open Innovation Conference, 2016.

‘Data is the key commodity underlying every transaction and fuelling every new insight. It is the raw material from which our future knowledge will be mined.’

Europe has existed for centuries as a divided continent; why can’t it exist that way for a few centuries more? And European companies could, theoretically, continue innovating in-house using mostly the proprietary data they have amassed or acquired themselves. But in the digital age, this is a distinct disadvantage. Because the truth is, small data sets and individual data points on their own are worth very little.[10] It is only by aggregating them that we can come to the fascinating insights possible through large-scale data analytics. And, typically, the most innovative products these days are built by third parties using and sometimes reusing proprietary and other sets of data for purposes which are not always the ones for which they were gathered.[11] And there, the possibilities for better living are almost endless: we could cut traffic in cities, learn to grow better yields from agriculture, put an end to cancer, keep closer track of extremists, discover new solutions to age-old problems, even make medical disasters that befall us in foreign countries much easier for citizens to solve when they occur.[12]

The sharing and aggregation of data from different sources is a necessary precondition of this progress. Here are five areas where better data analytics could accelerate existing trends and yield concrete results:

1) **Technological innovation.** AI requires data for machine learning: typically, machines need to learn from many similar implementations of the same behaviour in order to be able to perform at a high, value-adding level. One frequently mentioned “rule of thumb” is that to teach a machine how to do a job well, an AI programme will need 10,000 to 100,000 times more data than a human worker will generate during the entire course of her or his professional life.[13] If this is true, industrial applications of AI will only take place when data is shared between manufacturers – because few companies have the required critical mass of raw computing power, engineering talent and access to huge swaths of multifarious data.[14]

---

10 According to a recent OECD study, most individuals estimate the market value of their social security number at $240.00, (€226.84 at the 2015 exchange rate) while its actual market price is about $10.00 (€9.45) – the amount some companies might be willing to pay, hypothetically, for it. See OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being*, op. cit.

11 Joel Gurin, *Open Data Now: The Secret to Hot Startups, Smart Investing, Savvy Marketing, and Fast Innovation* (New York: McGraw Hill Professional, 2014).

12 See, especially, Jon Russell, “Google: Defeating Go Champion Shows AI Can ‘Find Solutions Humans Don’t See.’” *TechCrunch*, 17 March 2016.

13 Kim-Mai Cutler, “The Public Policy Implications of Artificial Intelligence,” *Medium*, 09 December 2016.

14 This is the reason why many current implementations of AI still focus primarily on consumer-oriented activities. Facebook and Google are the top recruiters for AI experts, because they have more data, perhaps, than anyone else except the government. CEO Sundar Pichai has set the goal of making Google a “machine learning first” company, and is encouraging competition in the AI-enabled digital assistant market with Amazon and Apple. Fitbit, a wearable company, is developing machine learning services based on the data gathered through the wearables. Self-driving cars are at the forefront of AI because there are many potential data points (cars) and huge amounts of data. Each connected car generates about 25 gigabyte of data per hour. Quartz, “Connected Cars Will Send 25 Gigabytes of Data to the Cloud Every Hour,” *Hitachi Bulletin*. Steven Levy, “How Google is Remaking Itself as a ‘Machine Learning First’ Company,” *Backchannel*, 22 July 2016.

'In an age where individuals, companies, even countries compete not just on cost but on relevance, insight, speed and creativity, the winners will come from those that most effectively and enthusiastically unlock the coming explosion of data.'

2) **Business model innovation.** Manufacturing itself is evolving, with leading companies moving away from selling widgets and into the provisioning of long-term service contracts based on the products they make.[15] In that context, data plays a key role, with many companies coming to rely on data-driven services to replace their more traditional offerings. Rolls Royce Ltd. is a case in point. It now sells guaranteed "aviation hours," rather than engines, to airplane manufacturers. The "aviation hours" themselves rely on complex data analytics, partly so they can deliver timely "predictive maintenance" to engines based on the data accumulated by sensors. This amounts to a dramatic change in the core business model of the company.[16] And it is happening in many other manufacturing sectors as predictive maintenance and data analytics emerge as the "killer applications" of the "fourth industrial revolution."[17] The increasing prevalence of application programming interfaces (APIs) and the related "API economy" – the process by which businesses are reconstituting themselves as online platforms – has grown largely based on this trend compelling enterprises to use data to develop new, innovative services. Another example is Orange S.A., the France-based telecommunications multinational, which now offers analytical services to shop owners; for a fee, the company will tell shop owners how customers are moving around their stores based on data they collect through smartphone usage.[18]

3) **Market creation.** Through portability of data requirements, consumers can choose new service providers and avoid lock-in. For instance, by accessing personal account data held by banks – as the European Union's payment services directive II now requires – some "fintech" market participants have been able to provide entirely new financial services which increase convenience and lower costs for consumers.[19] But more established players like Banco Bilbao Vizcaya Argentaria (BBVA), the Spain-based financial services multinational, are also innovating based on the new rules; the bank offers carefully curated, anonymised data sets to third parties for analytics, creating opportunities for others to innovate and mobilising additional eyes

15 Christian Reimsbach-Kounatze, "Benefits and Challenges of Digitalising Production" in OECD, *The Next Production Revolution: Implications for Governments and Business* (Paris: OECD, 2017).

16 Veit Dinges, Florian Urmetzer, Veronica Martinez, Mohamed Zaki and Andy Neely, *The Future of Servitisation: Technologies That Will Make a Difference* (Cambridge: Cambridge Service Alliance, 2015).

17 Klaus Schwab, *The Fourth Industrial Revolution* (London: Portfolio, 2017).

18 See Lorenzo Veronesi, Gabriella Cattaneo and Giorgio Micheletti, *Industrial Data Platforms: Key Enablers of Industry Digitisation* (Milan: IDC, 2016).

19 European Commission, *Directive 2015/2366 of the European Parliament and of the Council of 25 November 2015 on Payment Services in the Internal Market, Amending Directives 2002/65/EC, 2009/110/EC and 2013/36/EU and Regulation (EU) No 1093/2010, and repealing Directive 2007/64/EC (Text with EEA Relevance)*, 23 December 2015.

'The most innovative products are built by third parties using and sometimes reusing proprietary and other sets of data for purposes which are not always the ones for which they were gathered.'

to existing trends where further advanced insights might lie.[20]

4) **Social innovation.** Data is a fundamental component of solving complex public problems. Many companies donate data for public interest issues, a concept which the United Nations describes as "data philanthropy."[21] For instance, Orange shared anonymised records of millions of mobile phone users with the research community as part of their Data for Development Challenge, which sought to help find novel solutions to chronic development problems through the study of complex consumer behaviour in developing countries.[22] Among other things, researchers used mobile phone location data to study malaria diffusion related to travelling. Similarly, Yahoo! shared anonymised email messages to allow governments and researchers to map international migration. Google Trends – a website detailing the words being searched globally in real time – has been used to predict price trends, GDP growth and most famously the outbreak of influenza.[23]

5) **Evidence-based policymaking.** Public administrations are increasingly aware of the value of anchoring policy on sound evidence, as well they should be. The world is awash with indications that the best policy comes from – and often originates in – evidence itself. In that sense, governments are not only producers and custodians of data – they are often consumers as well.[24] And the more public administrations can rely on accurate, reliable information about population needs and trends – both prognosis and current assessment – the better.

There are, to date, many alternative frameworks for gathering and sharing data within Europe – systems that are designed to protect individual privacy while making it easier for a successful, large-scale data analytics industry to rise and take shape in Europe. These range from Midata and Open Algorithms to a radical idea to create new "data markets" by allowing citizens to "own" their data, and, by implication to sell it or license its use to others. Many of these ideas have useful elements that contribute to better sharing of data. But we believe they fall short of solving the problem for all time, and they may have missed the essential point: what Europe really needs is a better understanding of why sharing data is good, of how sharing

---

20 Gabriella Cattaneo, Giorgio Micheletti, Alys Woodward and David Osimo, *Data Ownership and Access to Data: Key Emerging Issues* (Milan and Barcelona: IDC and Open Evidence, 2016).

21 UN Global Pulse, *Big Data for Development: Challenges and Opportunities* (Geneva, United Nations, 2012).

22 Visit www.d4d.orange.com for more.

23 Most of the data analytics performed on Google Trends have been developed by third parties.

24 Martina Barbero, Jo Coutuer, Régy Jackers, Karim Moueddene, Els Renders, Wim Stevens, Yves Toninato, Sebastiaan van der Peijl and Dimitry Versteele. *Big Data Analytics for Policymaking* (Brussels: European Union, 2016). The study was prepared by Deloitte for the European Commission.

'What Europe really needs is a better understanding of why sharing data is good, of how sharing that data will lead to better social outcomes and why we should all join forces behind a new idea of "co-ownership" of data.'

that data will lead to better social outcomes and why we should all join forces behind a new idea of "co-ownership" of data. We need a simple framework – one that is short, simple and easy to understand – where the rights and responsibilities of data sharing are clearly defined and the reason why data can and should be shared readily understood. And, given Europe's unique structure – a transnational alliance of 28 (possibly 27) sovereign member states – that awareness needs to extend well beyond national borders, and should take up and comprise a European policy that promotes and facilitates the formation of easy cross-border data sharing and aggregation at the European level, where Europe has the size and scale to make something truly vital happen in this field. We will endeavour to provide such a framework in Section IV.

II. Is Data Ownership Central to the Free Movement of Data?

Some analysts have proposed a novel solution to the data-sharing problem. They say that the challenge of making data flow more easily across borders would be easiest to solve if we acknowledge that individuals "own" the data they create, an idea often referred to as "data ownership."[25] This idea seems logical at first sight. While common, statutory and regulatory law all routinely attribute "ownership rights" to ideas, "knowledge" and intellectual property through practices like copyright and patents, the use and spread of raw data are still mainly governed through contract law.[26] There is today a rich "market" for patented discoveries and other forms of knowledge in the broad economy, embodied by the plethora of registered patents and the emerging "open innovation" framework. But the market for data and raw data inputs, by contrast, is non-existent. Some believe that we could create a thriving market for data inputs by building an ownership framework for data. Under that scenario, enhanced individual control, clearer individual rights and – not-the-least-important – small financial payments would provide the incentives to encourage people to share their data into larger, aggregated and anonymised sets, possibly through a licensing system like the one commonly used today for the sale of software and other "intangible" creations.[27] The effect could be dramatic – much as enclosures kick-started agricultural capitalism and patents allowed knowledge trading to emerge in the Industrial Age.

However, there are several problems with this concept, at least one of them based on what would appear to be a genuine linguistic misunderstanding – something that arises from time to time in international institutions where non-native English speakers are often found using a language other than their own to discuss complex issues and reach joint policy conclusions. The Oxford English Dictionary defines ownership as "the act, state or right of possessing something," which can include the exclusive right to buy, sell or transfer the item in question.[28] This poses grave problems. First and foremost, the General Data Protection Regulation expressly forbids the sale of personal data for any reason. And, even if it were legal, the exclusive sale of personal data could lead to severe, extremely undesirable concentrations of data. As often happens in situations of this type, many individuals might fail to see the value of their data; they could sell the exclusive rights

---

25 See Cattaneo et al, op. cit.

26 European publishers, not surprisingly, have tried to claim an exception to this, arguing that text cannot be mined without a licence based on an intellectual property argument. Reformists have pushed back, arguing that text mining is the same as reading; the value lies in the patterns you identify, not in the database you draw from. See Sergey Filippov and Paul Hofheinz, *Text and Data Mining for Research and Innovation* (Brussels: Lisbon Council, 2016).

27 Officially, European software makers are not granted "patents" for the software they make. But it is their property under a "moral" rights argument, so it is theirs to license. The key point is the licensing is itself "non-rivalrous," meaning a product can be licensed to an indefinite number of people for an infinite number of times.

28 Catherine Soanes and Angus Stevenson (eds), *Oxford Dictionary of English: Second Edition* (Oxford: Oxford University Press, 2003).

'What is needed is a detailed and broadly accepted framework which establishes simple, easy-to-grasp rights and responsibilities for all members of the data ecosystem.'

for a small price. And, ultimately, the policy – if promoted on that basis – would have the opposite goal of the one that was intended. It would lead to serious concentrations of data in the hands of fewer and fewer, and to less and less empowerment for people who "sold" their data on to them for whatever reason. The 1995 Russian "voucher" privatisation is a good example of a large-scale experiment with a similarly sensitive problem – and the distinctly negative outcomes that approaches of this type can reach.[29]

There is, however, another, less formal definition of "ownership," which is closer to what "data ownership" advocates mean when they put forward this concept. It comes from the colloquial English popularised in management-based literature and marketing campaigns in recent decades. To "own" something, or to "take ownership" in this sense, means to feel like an owner, to embrace a personal stake in the outcome or to take personal responsibility for controlling how something is done. This concept is more usable in this debate, but we believe it still provides more confusion than clarity. For one, it hardly addresses the key issues of access, modification, use and reuse where a clear and broadly embraced delineation of rights and responsibilities so clearly needs to

be defined and embraced.[30] In other words, it provides a useful metaphor, but it is still only a metaphor. And at the end of the day what is needed is a detailed and broadly accepted framework which establishes simple, easy-to-grasp rights and responsibilities for all members of the data ecosystem who contribute in one way or another to complex data analytics. That framework must set out the rights of the participants – corporate, individual and state – and make the case for sharing. And it should contribute to an overall framework of trust.

Perhaps the best place to start would be to address the potential bones of contention – the concrete rights and responsibilities where definitions and delineations must be found to facilitate a more open, widely understood and broadly accepted environment for data and data sharing:

- **Access.** Granting "access" to data allows one person or agency to see another person's or agency's data. This happens when a payment service company, for example, is allowed to "access" individual account data held at another financial service company, as is mandated in some instances already today.[31]

---

29 Andrei Schleifer and Daniel Treisman, *The Economics and Politics of Transition to an Open Market Economy: Russia* (Paris: OECD, 1998).

30 See also, the extremely interesting discussion on how these key aspects of data "rights" might be addressed in OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being*, op. cit. The OECD will continue its work in this area with a new initiative on "enhanced access to data: reconciling risks and benefits of data reuse." Visit http://www.oecd.org/sti/ieconomy/data-driven-innovation.htm for more.

31 The OECD provides a useful distinction. Access "in its weakest form" allows a person to see their personal data, giving consumers greater insight into their personal behaviour. Access "in its most extreme form" is when third-parties are given statutory access to people's data. See OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being*, op. cit.

'Complications surrounding the confused legal status of data, and the uncertain value it holds, are discouraging companies from sharing it.'

- **Modification.** "Modification" goes a step further. It allows a person or agent to modify the data. Government agents use modification to make changes in citizens' entries in the population registry. Fintech companies use it to perform payments from accounts held by other companies.

- **Use.** Use is a relatively simple concept. It's what the people and agencies holding the data do with it, i.e., the analytics they run and the questions for which they seek answers. In advanced cases, some individual data producers are already given rights over the use to which their anonymised, aggregated data is put. For instance, farmers who take part in data-gathering processes run by Deere and Company, the U.S.-based agricultural machinery manufacturer, are informed about what the company does with "their" data, with whom it shares it, and for what analytical purpose. Usually, these "consultations" take the form of ex-post communication.

- **Reuse.** This is the Achilles heel of the emerging European system. Under terms of the forthcoming general data protection regulation, companies are forbidden from "reusing" personal data for any purpose other than the purpose for which it was originally gathered. In practice, this puts certain very clear limits on the ability of European companies to combine some data sets with others in search of new insights. Other restrictions – including confusion over the legal status of mining legally held datasets or highly restrictive terms of use for legally shared proprietary commercial data – complicate the task of allowing data gathered for one purpose to be used to train machines or algorithms in other contexts, or to perform exploratory research.

Not covered in this list of key rights is another particularly important aspect of "ownership": namely, the extent to which data producers should have the right to benefit financially from the commercial reuse of data – what is sometimes referred to as a "fair return" on data gathering. The idea is that if data holders could benefit financially from their data, they would be far more willing to sell or to license access to it. However, there are several drawbacks to this approach; first and foremost, the market is highly un-developed at the moment, and, even if it were more developed, it could be a while before anything like price transparency is evident or fair value clearly established enough to provide a fluid market that incentivises data sharing. As is, the value of aggregated databases is easy to see. What is harder to evaluate and/or assign is the market price for individual data points within it. For instance, Monsanto Company, the U.S.-based agro-chemical and agricultural biotechnology corporation, paid $930 million (€678 million at the 2013 exchange rate) for Climate Corporation, mostly to have access to the weather-based data company's 2.5 million sensors installed on cultivated U.S. farm land.[32] But farmers taking part in similar systems have recently sold their individual data points for about $2.00 (€1.47) per

---

32 Bruce Upbin, "Monsanto Buys Climate Corp for $930 million," *Forbes*, 02 October 2013.

'It boils down to two questions: how do we create the necessary trust where people understand the value of their data and are comfortable sharing it? And how do we ensure that once shared that trust is honoured without breach?'

acre.[33] Considering that the average farm in Europe is 16 hectares, or 40 acres, the average European farmer would have received about $80.00 (€58.85) by selling his/her data at this price – not exactly a revenue source destined to incentivise radical changes in behaviour or to finance major shifts in business models.[34] And even then, the notion of introducing a debate over fair value to the discussion we are having here would only provide incentives in the wrong direction – it would lead to more walls, more silos, more debate over who might best and most usefully aggregate what and when. What is needed is more sharing of data, and more pooling of common interest in broad social outcomes. The bottom line is, based on the evidence available today, market mechanisms created by enhanced data ownership rights are not likely to provide the appropriate incentives to unlock broad-based data sharing projects. "Ownership," particularly if implemented in its strictest sense, is rather more likely to lead to even greater concentration of data in hard-to-reach silos and pose additional barriers to data sharing.

Last but not least, the fundamental transaction theory of economics shows that uncertainty typically induces companies to adopt "hierarchy-based" solutions at the expense of "market-based" approaches. In other words, complications surrounding the confused legal status of data, and the uncertain value it holds, are discouraging companies from sharing it – and robbing society of the immense advantages we would enjoy if data could be coaxed out of its silos, aggregated and put to use for broader social good. In the absence of a broader consensus on this – and a framework that facilitates and encourages it – most companies push for in-house solutions or bring the data they need in-house through acquisitions, such as the Monsanto-Climate Corporation transaction described above.[35] In the end, the data remains in silos. Which is precisely the problem we must address.

---

33 Leawood, Kansas-based Farmobile LLC offers farmers $2.00 per acre for the Electronic Field Records (EFRs) – around €1.89 at the 2016 exchange rate. According to its website, one farmer earned $17,952.00 (€16,974.28) for his EFRs in 2016. Visit https://www.farmobile.com/datastore. For consistency, the euro/dollar exchange rates given in the Climate Corp. example are from 2013. Visit https://www.farmobile.com/datastore.

34 The farming data is from Eurostat, available at http://ec.europa.eu/eurostat/statistics-explained/index.php/Farm_structure_statistics.

35 Oliver E. Williamson, "Organisational Forms and Internal Efficiency: Market and Hierarchies: Some Elementary Considerations." *American Economic Review 63*, No. 2, 1973.

III. Potential Ways Ahead: Alternatives and Scenarios

Quite a few experts have analysed the challenge of facilitating greater data exchange. The problem itself could be stated like this: how do we open up data for greater social good without 1) violating personal data privacy, 2) expropriating proprietary data from people, companies or governments that don't want to share it, and 3) breaking existing European Union rules? Most proposals – including the one we will make in this policy brief – all revolve around similar goals. In a nutshell, it boils down to two questions: how do we create the necessary trust where people understand the value of their data and are comfortable sharing it? And how do we ensure that once shared that trust is honoured without breach?

a. **Midata.** This United Kingdom government-led initiative has focussed on providing customers with control over their transaction data history in the banking, energy and mobile phone sectors. It is designed as a voluntary effort by companies to give back the data in machine-readable format in order to foster competition and choice.[36] Similarly, the "Mydata Alliance" was recently formed to develop broader, trans-national interoperability standards around a "human centred approach to personal data management" which allows individuals to put their data in a secure hub, from which they can later authorise who does or does not have access to it.[37] Users use their e-mail address and a password to log on to a secure dedicated website. Once there, they can review what personal data is stored on the secure platform, see who has accessed it, learn the reason for that access, authorise the category of users with whom they will share their data and download some data (if they are authorised) themselves.[38]

b. **Open Algorithms.** Massachusetts Institute of Technology (MIT) Professor Alex Pentland has led this effort, which essentially offers to provide a third-party service to analyse data without actually sharing that data or the algorithm on which it is to be processed.[39] Essentially, Open Algorithms (OPAL) serves as a "black box." A researcher can give a query to the data. OPAL returns an answer to the query, but without ever sharing the data or the way it was analysed – thereby preserving the privacy of all involved.[40]

---

36 See also, HM Government, *The Midata Vision of Consumer Empowerment*, at https://www.gov.uk/government/news/the-midata-vision-of-consumer-empowerment.

37 The quotation is from the website of MyData 2016, an international conference on Midata that is growing in scope and importance. For more, visit http://mydata2016.org/.

38 France has developed a similar system called "MesInfos," led by the *Fondation Internet Nouvelles Generation (FING)*. And both systems are based loosely on ProjectVRM, a U.S.-based initiative, that produces a model for "vendor relationship management."

39 Thomas Hardjono, David Shrier and Alex Pentland, *Trust::Data: A New Framework for Identity and Data Sharing* (London: Visionary Future, 2016).

40 Startups like Estonia's Cybernetica AS offer similar services.

'Data needs to be coaxed out of its silos – not put in more of them.'

As a further level of security, the answers are provided in encrypted format. Prof Pentland has further proposed a rating system for algorithms. Rather than forcing companies to reveal their algorithms so they can be vetted for "fairness," Prof Pentland offers to use the same encryption technology to allow algorithms to be vetted independently without being disclosed publicly.

c. **Data Pods and social linked data (Solid).** Essentially an app – or perhaps a family of apps, all privately developed – "data pods" would host an individuals' data, giving him or her the right to decide with whom it was shared – and on what terms. The system, however, would be dispersed, avoiding the concentration of data in any one place but allowing individuals to aggregate it through a system of permissions. The idea is similar to the de-centralised, free-flowing concept that gave rise to the World Wide Web, which is no surprise, since the principle advocate of "personal online data stores," or "pods," is Tim Berners-Lee, founder of the World Wide Web.[41] The project is being further developed at an MIT lab. The concept behind it is that you would have "social-linked data." The data creator would be the "owner;" he or she could decide how the data was shared, hence the "link" between data creator and data

processor. An interface called "solid" would determine which apps should be given access to your data based on the criteria you define.[42]

d. **Open data.** Some EU member states – and more recently the European Commission itself – have recently proposed widening the definition of open government data by including "public interest" as a criterion for future sharing of commercial data.[43] In particular, France has established that data derived from public procurement, real-estate transactions and energy consumption must be made available to the broader public on the same terms as government data. Finland has acted similarly for data from private transportation providers.

e. **Sector-specific regulation.** Some new regulations – promulgated mostly at the EU level – force companies to share data with downstream service providers to create new markets and ensure a level playing field. This is the rationale behind both the payment services directive II and regulation 715/2007, which force banks and car producers, respectively, to share data with third-party-service providers (payment services and after-sale) in order to ensure competition and avoid lock-in effects.

---

41 Tim Berners-Lee, "I Invented the Web. Here are Three Things We Need to Change to Save It," *The Guardian*, 12 March 2017.

42 Diaspora, an open-source, not-for-profit social network run by the Diaspora Foundation, runs a similar project.

43 European Commission, *Communication on the Mid-Term Review on the Implementation of the Digital Single Market Strategy: A Connected Digital Single Market for All* (Brussels: European Commission, 2017).

'We believe the problem is best solved not with more and more intermediaries, but with better common understandings and greater public awareness.'

f. **Self-regulation.** Oddly, the farming sector – by definition the paragon of the Agriculture Age economy – has emerged as a surprising digital champion in the Age of the Internet, as some of the examples cited above indicate. But the story hardly ends there. Several U.S.-based agriculture sector stakeholders have joined together to design a highly successful "transparency evaluator" for farm data,[44] offering a set of tools to farmers to help them understand 1) how their data is being used, and 2) why they should share it. This certification process for "agricultural technology providers (ATPs)" assesses companies and other stakeholders based on their respect for a set of data principles collectively chosen by the ATP members. It includes information "about the purposes for which they collect and use farm data." The sector – in the U.S., at least – is emerging as a little-known champion of progressive data policies, and stands poised to reap a high-tech benefit in the not-too-distant future.

We believe that most of these innovative solutions – while interesting and helpful to the discussion and debate – collectively fall short of addressing the larger, underlying problem. Data needs to be coaxed out of its silos – not put in more of them, even if the rationale behind some of these mechanisms are the

right ones. To be sure, there is certainly a realm of personal rights sitting at the heart of this debate – a zone within which people's privacy must be protected – but we believe the problem is best solved not with more and more intermediaries, but with better common understandings and greater public awareness.[45]

Concretely, we believe the problem could be resolved through a new concept of "co-ownership" that recognises the mutually overlapping levels of "ownership" in the broad sense of the word for the different types of data that are generated in the economy, the different ways that data is stored and analysed and the common interest we have in pulling those insights out and getting them into the social arena. A concept of "co-ownership" is particularly important because of the joint stewardship that it implies: an individual "owns" his personal data; but so in some sense does the state which gathers and aggregates it (personal data is stored by the state and is aggregated and analysed in the form of population statistics, land registries and the like. Participation in these systems is not optional. Businesses, citizens and land ownership must be registered, which makes the state a key participant in determining how data is collected, shared and analysed.). Similarly, individuals create data when they use their smartphones, drive their cars or

---

44 See the Ag Data Transparency Evaluator (ADTE), an online tool created to help farmers make decisions about data transfer, usage and sharing. Visit http://www.agdatatransparent.com/the-privacy-and-security-principles-for-farm-data/.

45 See, especially, the excellent article from Solon Barocas and Helen Nissenbaum, in which the authors argue that anonymity and consent clauses have failed for technical reasons to provide genuine anonymity or real consent. The goal now should be ensure that the "outcomes in question can be defended as morally and politically legitimate." Solon Barocas and Helen Nissenbaum. "Big Data's End Run around Anonymity and Consent" in Julia Lane, Victoria Stodden, Stefan Bender and Helen Nissenbaum (eds.). *Privacy, Big Data, and the Public Good: Frameworks for Engagement* (Cambridge: Cambridge University Press, 2014).

'A concept of "co-ownership" is particularly important because of the joint stewardship that it implies.'

share social media. Clearly, that information belongs in some sense to the people who produce it. But it also belongs to the people who provide the phone service, make the cars or manage the social-media network. This is a particularly important point: much data in the modern world is proprietary and commercial; it was gathered based on a business model where a company offered a customer a free service in return for the right to track how

they use it. Beyond that, you don't need much imagination to see a third, very public interest in accessing as much of this data as possible for broad, social goals. We already use smartphone data for catching criminals and investigating crimes. And data aggregated by our increasingly smart cars will have a huge impact on our ability to lower traffic congestion and fight pollution in our cities.

IV. A New Vision: Greater Use and Reuse through Transparency and Portability

The perpetual disagreement and confusion surrounding the need to share data more freely for greater social good essentially derives from one cogent fact: data – regardless of its source or nature – clearly has social utility and economic value. But the value of the data varies widely depending on the use you make of it. To be sure, there is value in the individual points of a database. Data, in that sense, is how we keep track of things, the way we ensure that the detail is not lost in the composite. But the real insight comes not from the ability to figure out which social media user might be best expected to buy which product, but from the aggregation of all of these data points into a larger picture – and the capacity to use that aggregation to drive social insight for the greatest common good.

The emerging awareness of the value of data is recognised in the metaphors cited at the start of this policy brief. But the issue is data is not oil, or currency, or infrastructure. Data is data. It plays a unique role, has unique characteristics and often follows an economic logic of its own. And it originates from a multitude of sources, each of which has differing needs and expectations of the rights those needs bestow and how those rights might best be upheld. Its "non-rivalrous" nature means it is always there to be used even if it has been already used once. A traditional economic asset, by contrast, is depleted when it is consumed. But not data. Its use in one context does not mean it can't be used over

and over again in another situation – at near zero additional cost.

So who then owns the data? The answer is we all do.

Though we own it not in the traditional sense of the word, but in a cascading array of overlapping rights, responsibilities – and even opportunities. See Chart I on page 19 for a graphical description. One way of seeing it might be to compare data "ownership" to the rights parents have over their children. No parent owns their child; and, indeed, children have strong state-mandated protections *vis-à-vis* the parents, should it come to that, including iron-clad restrictions that no human being can ever be bought or sold. But parents absolutely have a say in how a child's life will unfold: which schools will she or he attend? Which playdates will she or he go to? What's more, we exercise and adjudicate those rights through joint decision. It's not up to Mom or Dad. To the contrary, it's for Mom and Dad to make joint decisions together. This is how we must come to think of data – it's an area where we decide together what is best. We have a common and mutual interest in ensuring that data is accessed and stored for the right reasons. But we have an equally strong incentive to make sure that data is used for the common good in the first place. This interest sits alongside the rights and duties of individuals, which have already been so clearly defined in Europe, especially with the forthcoming general data protection
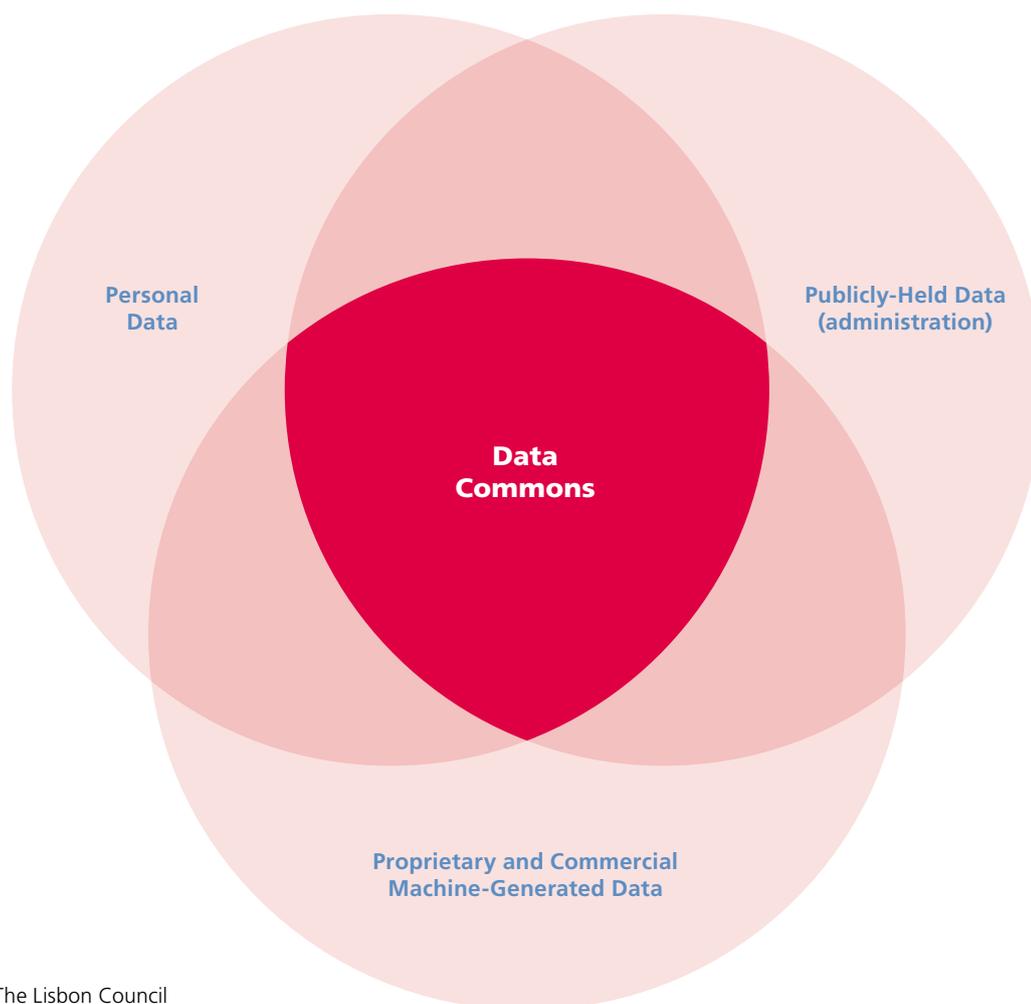
Source: The Lisbon Council

regulation.[46] But this includes a right and duty (and perhaps even an incentive) to make sure that anonymised data is accessible for the social projects on which we need it. What's more, companies themselves must be given appropriate incentives – certainly not to abuse personal data, but to allow the companies themselves to develop and build the models that are already driving so much data-led insight and rewarding so many consumers with better products and services. This means several things concretely:

1) We need a better framework for aggregating and anonymising public data (held by the state), including a better framework for aggregating and analysing data at the European level (involving cross-border data sets) and a built-in system of transparency by default, so people

_____

46 The regulation will take effect in May 2018. See European Commission, *Regulation on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data, and repealing Directive 95/46/EC*, 27 April 2016.

'The value of data varies widely depending on the use you make of it.'

can (easily) see how their data has been accessed and report potential abuse;

2) We also need more open data;

3) We need to avoid moves to retain data within national borders through unjustified localisation requirements;

4) And we must make sure that enough data is available, anonymised, and on the right terms, to make Europe a world centre of society-improving data analytics.

In other words, we need a framework that encourages and incentivises cutting-edge research on the great social, economic and health-related issues of our time as well as encouraging the emergence of more and better data-driven businesses – within EU member states and across the EU as a whole.

What is missing is not more rules and regulations – we have thousands of those, and many of them clash and overlap. What is missing is a simple framework which every European citizen can understand – one which easily conveys the unique status of data as an exceptional commodity that can be many things at once – both public and private. But in whose aggregation we all share a common interest and bear a common responsibility. The sooner we quit seeing the difficult contradictions of data as trade-offs and points of clash, the easier it will be to construct a

broad social framework based on trust and confidence. Obviously, this will need a strong legal base behind it. But as is, too much of the law – and the institutions that draft it – spend their time seeking to claim data jurisdiction as their sole domain, effectively dragging this unique new asset upon which so much modern economic life relies into the institutional power struggles – national as well as European – that characterised political life in less international times. We need to move beyond that into a period of "co-ownership." That period, in turn, can be based on the patchwork of existing laws: contract law, intellectual property law, consumer consent and strong enforcement at the European level. And the state may well need to take concrete action to make this happen – first and foremost, by setting an example with the use and release of more "open data" in easy-to-use, non-proprietary, machine readable formats, but possibly also through legislative initiative to put forward model contracts that set the context and tone for effective, socially beneficial data exchange.[47] But the important thing is that the contractual basis upon which data exchange exists is simple, easy to communicate, broadly understood and widely accepted.

---

47 In principle, a "model contract" should be drawn up in broad consultation between companies and stakeholders. But the state can play a helpful role here. One useful model would be the "take down notice" practice adopted by the U.S. in the Digital Millennium Copyright Act in 1998. The benefit of the "take-down-notice" principle is that it is simple, easy to understand and doesn't need reams of legal notices to explain. But the initiative in developing this concept was taken by the U.S. Congress.

V. Roadmap: Three Areas for Urgent Initiative

The odd thing about data is that it is generated in so many different ways. It comes from many different sources and ultimately serves many uses and needs – and often at the same time. Concretely, and within this context, it might be worth re-categorising how we think about data, and dividing it into three primary sources: 1) Personal data (made up of the key facts of an individual's life, including health data), 2) Public data (the data that is aggregated in public registries or generated through public administration), and 3) Commercial/proprietary data (which includes machine-generated data as well as data generated by a company through its activities and interactions with individuals. This often includes undertakings, like analytics, where "value is added." But it will also include the coming explosion of Internet of Things, smart cars and other machine-generated data.).[48] Oddly, each of these sources is easy to define, but the most important point to grasp is the extent to which they all overlap. The same data can be classified and re-classified in different ways, depending on the vantage point of the user. For example, is the data held in the public registry personal or public? [Answer: it is both.] Did the platform create the data it mines or did the user who generated it? [Answer: both did.] This is why any effective framework for managing data access and aggregation must recognise the "co-ownership" of the data itself, namely, all data has several "owners," each of whom have a clear stake and legitimate interest in determining how it is used. But we as common and collective data-owners have a clear stake as well: we must ensure that all data is ultimately best used for the greatest common good – which means first and foremost that it is put to good use, rendered most accessible and made available within a framework of mutual respect and co-decision – much as parents decide jointly on what might or might not be best for their child.

With that framework in mind, we propose three concrete initiatives to improve and extend the "free movement" of data within Europe: 1) An initiative to make "once only" a reality in Europe, as was first proposed in 2015 in the European Commission's digital single market programme and more recently given a dramatic boost in the form of a proposed European regulation on a single digital gateway. This should include wider use of base registry interoperability to ensure better use of public data in Europe.[49] 2) A ban on data localisation except in extremely clear and clearly defined cases, and 3) A stronger framework for co-decision on the use of public/private/proprietary data and

---

48 See Stacey Higginbotham, "Ericsson CEO Predicts 50 Billion Internet Connected Devices by 2020," *Fortune*, 14 April 2010. The figure has become a standard reference point in technology debate since first touted by then Ericsson CEO Hans Vestberg.

49 European Commission, *Proposal for a Regulation on Establishing a Single Digital Gateway to Provide Information, Procedures, Assistance and Problem Solving Services* (Brussels: European Commission, 2017). A base registry "refers to a trusted and authentic source of information under the control of a public administration or organisation appointed by government," according to the European Commission. In other words, these are the databases which public administration uses to keep track of the most things for which it is responsible for knowing, e.g., the registration of residency, citizenship, property, business, family, automobile, driving permit and building ownership. See European Commission, *Access to Base Registries: Good Practices on Building Successful Interconnections of Base Registries* (Brussels: European Commission, 2016).

'Data is not oil, or currency, or infrastructure. Data is data. It plays a unique role, has unique characteristics and often follows an economic logic of its own.'

the gradual evolution of a "data commons" (this will be discussed on pages 31-34). These initiatives are important because of the coming explosion of machine-generated data. The important thing is not that legal definitions be tightened – we have spent years doing that. The important thing is that common understandings be generated. We need to understand our common interest in building a data-driven economy; and we need to find a framework based on mutual respect for all data-generating, gathering and analysing parties based on a common understanding of co-ownership. It must be a framework that is easily understood, broadly accepted and widely communicated.

Fortunately, we have a lot of ideas in Europe on how data could be made to flow more freely, taking advantage of the EU's "ever closer union" and finding its way into more friendly formats that would better serve citizens – and allow the kind of big data processing that could yield important, valuable insight. We list a three-point programme here. The three items were chosen not because they are exhaustive, but because we believe they could yield the quickest returns.

1) **Once only, and greater European "interoperability" for national databases**

A lot of data in Europe sits with national governments. Every citizen is born and dies. Many of them marry. Still more have children. And some of them buy houses and land. Some learn to drive, or register to live in new neighbourhoods, or change their names.

All of this data is stored – somewhere – in the national or local state administration databases.

But the problem is Europe is not just fragmented into 28, possibly soon to be 27, member states. There is fragmentation within countries, as many national agencies maintain their own databases, some of which have grown up on proprietary ICT systems, which don't talk with the ICT systems in other departments. The modern trend is to move away from this. In that sense, Estonia has led the way with the pioneering X-Road platform and a legal rule that requires the government to gather data "once only" from citizens. This has obvious administrative benefits for citizens – when you want or need to talk to the state, you don't have to enter all of your data multiple times or chase documents at one agency just to hand it over to another. Every citizen has a unique user ID. For example, Estonians no longer carry drivers' licences – the relevant data can be accessed via their national IDs from the national database, if a police person stops them.

But making this possible required state intervention – not just to set up a system of fail proof underlying protocols which allowed basic government data stored on different servers to be exchanged accurately, securely and interoperably with other government agencies. To be blunt, the key moment was when the government required the agencies to use this system – a move which is referred to in shorthand as "once only." This meant the agencies had to make it work; as long as the X-Road was voluntary, its pick up was minimal. But from the day that agencies

'We have a common and mutual interest in ensuring that data is accessed for the right reasons.'

were required to use it, citizens no longer had to spend their time supplying the state with their data every time they needed something; the agencies had to go and get the data themselves. Today, more than 800 Estonian agencies take part.[50] One important caveat: the platform is accompanied in key digital services (e.g. national electronic health records) by a simple, easy-to-use online feature with which citizens can see exactly who has been accessing their data – and why. This is very important for maintaining trust in the system. State administrators must obey strict data access laws. But if, say, a person gets married, the marriage licence granting agency can consult the personal data on the population registry – to make sure the person isn't married already. That check will leave a small notation – including the name of the accessing agency and the reason for the access – on an individual's personal data file, which the individual can later check herself or himself online. If there is a violation or a suspected misuse of the system, Estonians have a strong right to redress. Amazingly, since the system was implemented in 2003, there has not been a single law suit filed alleging unlawful access.[51] Transparency has delivered compliance.

The advantages of such a system are both evident and obscure. The evident part is the enormously simplified administrative procedures for most Estonians. Company registration is made considerably easier. Financial data can also be exchanged in machine readable format, electronically, at the request of the citizen (you can have your tax returns sent to your bank electronically, for example, to support a bank loan). And there is built in transparency; no one is allowed to access personal data without reason or permission; and these reasons and permissions must all derive from national legal acts or be based on the citizen's explicit consent. In a population of just over one million, the system processes roughly one billion requests for information per year – nearly 100 requests per citizen.[52]

One important additional benefit is the boost this makes towards data aggregation, the knowledge Estonians gain about their society and its key trends as well as the potential it opens up for better, more advanced public administration in years to come. Casinos, for one, are part of the system; this allows them to do quick and effective identity checks on individuals who might like to enter an establishment and gamble. Imagine such a system rolled out and easily accessed at Europe's borders. It would lead to vastly more effective border protection – as well as the more effective exchange of important information within the Schengen area.

Other countries have also led on this. At the federal level, Belgium has Magda – a

---

50 Utilities, telecommunication service providers, banks, businesses (through enterprise resource planning APIs), even state-owned casinos also take part.

51 Fines and penalties have been levied in response to several incidents, but this is "pre-court."

52 Interestingly, most of these exchanges are machine-to-machine automatic queries.

'The important thing is that the contractual basis upon which data exchange exists is simple, easy to communicate, broadly understood and widely accepted.'

base registry interoperability system, which allows a version of "once only" to thrive in Belgium. In keeping with Belgium's federated system, the base registries are maintained at the regional level – but Magda makes them interoperable at the federal. Likewise, Spain has an "intermediation platform" which performs a similar function (linking regionally held population registries into a single national system). And the Netherlands has an advanced system of base registries built around iNUP, a commonly agreed data-sharing principle, which has made this low-lying nation a quiet pioneer in the field.

Germany, however, is something of an outlier. It has more than 200 copies of its national registries, few of which can talk to each other. This makes German data into an island in Europe – a situation which is perhaps supported and endorsed in certain legal quarters, but not one that lends itself towards better public administration, more effective border control or advanced data analytics on key social and economic trends.

Put simply, in order to get a grip on the larger issue of broad data aggregation in Europe – and the related problem of "free flow" of data across borders – European governments must learn to communicate with each other more effectively, and to take the burden of

effectuating that communication off the backs of citizens. This would have immensely positive effects on simple matters like cross-border e-procurement and support major European initiatives like the "digital pole" for open data being constructed in Luxembourg.[53] But also on larger, more complicated projects like better border control, where recent years have shown Europe could stand to improve, and where recent elections have shown that European voters would like public administrations to do so. This would – almost as an afterthought – go a long way towards creating the kind of anonymised, behaviour-driven, population-based data aggregation that could yield huge insight into key European trends – and serve as a basis for a host of new and better services and service-provision throughout Europe.[54]

Above and beyond the effort to simplify the procedures in which a European citizen from one member state can access her or his data in another EU member state, and make that data more readily available to authorised services and procedures, is the effort to ensure inter-operability among agencies within individual EU member states. This will require political and technical leadership – not the least being an effort to define common standards and protocols for data exchange and to ensure those standards' uptake by thousands of

---

53 DG Digit and DG Connect, *The Digital Pole in Luxembourg* (Luxembourg: European Commission, 2016).

54 One area where Europe has made important progress in this regard is with the Business Registers Interconnection System, or BRIS. Due to become fully operable in summer 2017, the system is intended to give a boost to cross-border business activity by facilitating access to official information on EU companies. Once complete, it should make basic information more readily available to the public, and in particular to simplify access to information on companies in a member state other than the one where the company is registered.

'The important thing is not that legal definitions be tightened – we have spent years doing that. The important thing is that common understandings be generated.'

local public administrations.[55] This, in turn, will require funding. The European Union structural funds should be used for this. The funds themselves are a common pool of money available to help regions with transition, and what change could be more important to all European regions than the transition to a digital economy? But money is not the only lever in policymakers' tool chest. Common standards for cross-border exchange would dramatically lower costs by taking expensive technical processes off of the table. And it would facilitate the implementation of common standards within countries – i.e., at the local and regional level – where pick up and compliance have been slowest.[56]

2) Stronger moves against data localisation and strengthened cybersecurity

Russia has one of the world's most stringent data localisation requirements. If you are processing data related to Russia's citizens, you must store that data in Russia. And it's not hard to see why a system like this might appeal to a regime like Russia's. It creates business for local Russian providers who might not have had it if market terms were

allowed to apply. And it helps the government to keep an eye on what anyone in Russia might (or might not) be up to.

Oddly, data localisation requirements in Europe take place largely for the opposite reason – at least officially. The German government has a host of data-transfer restrictions – starting with the 2015 Telecommunications Act, which requires that German cell phone data be stored on servers in Germany – which Germans argue is necessary for national security reasons, i.e., to prevent other countries from spying on their citizens. But it doesn't end there. The Data Protection Supervisor of Schleswig-Holstein – a German state of 2.8 million people – has ruled that all local data transfer to the U.S. must be stopped in the wake of the 2015 Court of Justice of the European Union *Maximillian Schrems v. Data Protection Commissioner of Ireland* decision (which held that the U.S. lacked adequate protection against un-notified personal data accessing by the U.S. government).[57] But Germans are not alone in this. At the end of the day, at least 14 European countries have data localisation requirements of one sort or another, including such disparate countries as France (which

---

55 The problem now is there are too many "common standards." There is scope for governments to move here, using their power to guide the market towards highly-effective, truly interoperable, non-proprietary standards, much as they did 25 years ago with the Global System for Mobility, or GSM, agreed at the European level. See Jacques Pelkmans, "The GSM Standard: Explaining a Success Story," *Journal of European Public Policy*, 8(3): 432-453, 2001. Stephen Temple, "Chapter 23 – Could Europe Create Another GSM Success?" in *Inside the Mobile Revolution: A Political History of GSM* (stephentemple.co.uk, 2010).

56 A recent European Commission report showed moves towards inter-governmental data interoperability at the European level was more advanced on a conceptual basis than at the level of implementation, although, defined broadly, 72% of existing national ICT systems were roughly compliant with EU interoperability guidelines. But the situation is vastly different at the local and regional level – across different departments and local and regional institutions – where the implementation level reaches an average of only 28%. See European Commission, *State of Play of Interoperability in Europe: Report 2014* (Brussels: European Commission, 2014).

57 The Schleswig-Holstein ruling has yet to be enforced, and it is unclear what it's larger legal status in Germany is.

> 'One important benefit is the potential "once only" policies open up for better, more advanced public administration in years to come.'

forbids the storing of public administration-generated data on "non-sovereign" clouds) and Luxembourg (where financial data is required to be processed within the country but "exceptionally permitted" within multinational institutions or with "explicit consent.")[58]

The European Commission is fighting back, vowing to put an end to "unjustified" local data storage requirements, though there has yet to be a serious infringement case despite the plethora of possible cases in Europe where the European Commission could make its mark.[59] If and when it does move, it is important the European Commission does two things: 1) First and foremost, it must move even-handedly against all alleged violations, giving and appearing to offer no favour or bias to any individual country, and 2) it must accompany any moves with steps to show that it understands the security questions at hand here and is aggressively pursuing solutions through a stronger European framework. Put simply, any measures to end "unjustified" data localisation should be accompanied by a broad and credible initiative to strengthen the European cybersecurity system, which remains underdeveloped given

the tremendous importance of the issue and the extreme pressure to which it is already being subjected. European policymakers should move beyond merely the "cooperation" of national cybersecurity agencies proposed in the Directive on Security of Network and Information Systems (NIS Directive).[60] Europe's cybersecurity policies should be second to none. And there is scope for using European leverage to bring greater weight – and more real time data – to this problem through a stronger instrument.

Data localisation actually has two important shortcomings in this regard. First and foremost, it usually fails to deliver the security it promises – data sets held on national servers are not necessarily safer than data that is, for example, dispersed in unknown locations with strong encryption behind it.[61] So any country genuinely concerned with data security – and not simply trying to drum up national business for local suppliers – will support a more robust, non-nation-state-based solution in this field. But there is the larger problem of the free flow of data – which, in this context, becomes a classic European single market question. The European Union treaties guarantee free establishment, i.e., the right

---

58   See the list of known data localisation requirements in the Joint Industry Statement on the Free Flow of Data. BusinessEurope, "Free Flow of Data is at the Essence of a True European Digital Single Market: Digitalisation Can Be at the Heart of Europe," 29 November 2016.

59   European Commission, *Building a European Data Economy*, op. cit.

60   European Commission, *Directive of the European Parliament and of the Council of 06 July 2016 Concerning Measures for a High Common Level of Security of Network and Information Systems across the Union*, 19 July 2016.

61   One advantage of the Estonian X-Road system is that no data is stored centrally. Instead, it is dispersed throughout a system of servers controlled by different agencies. The X-Road protocols allow that data to be searched together even if it is stored in different places. Thanks to Andres Kütt, advisor at Estonian Information System's Authority (X-Road), for an excellent discussion of this.

'European governments must learn to communicate with each other more effectively, and to take the burden of effectuating that communication off the backs of citizens.'

for nationals in one country to set up business freely in another. For years, this has been used to justify the free flow of data, which should be a *sine qua non* for cross-border business in the digital age. If the famous third freedom on business establishment is not enough to let information flow where it needs to – a role it has played traditionally and unofficially for years – then the free flow of data needs to be guaranteed explicitly in a fifth freedom guaranteed by the Treaty of the European Union.[62]

In principle, the incoming general data protection regulation should help. The reality of a single set of rules for the transfer of data across borders within Europe – as well as the regulation's "consistency mechanism," which promises that the soon-to-be-christened European Data Protection Board will make uniform decisions out of the sometime patchy European quilt – should provide some important support.[63] But this 2016 regulation itself has been subjected to unfortunate "gold plating" in some EU member states, which are starting to introduce so-called "opening clauses" to build in local requirements for data storage and retention.[64] This is an annoying trend, which all stakeholders should unite to overcome. European policy can and does need some teeth if it is to be effective

in the area. The law was written to provide legal safeguards, but it must be uniformly interpreted and enforced if those safeguards are to provide the level-playing field they were intended to support.

3) New frameworks for sharing proprietary data

Few issues exemplify the problems in data management more than the enormous amount of confusion over the role and status of data being gathered by companies and created by human activity. This set of data derives from almost every aspect of our distinctly modern lives – from your cell phone records, which belong to your cell phone company even though you are the one who made the calls, to the places where your car has been today, which is known to your GPS even though you are (still) the one who drove the car. Put simply, this is big business. In the age of the Internet, the value of many goods and services is no longer in the margin the offerings command at sale but in the knowledge that their use generates. And, as the Internet of Things comes on line, bringing as many as 50 billion new objects on to the Internet; as machines themselves start to generate more and more usage data for their makers and users, the problem of who owns all of

---

62 "Data should be able to flow freely between locations, across borders and within a single data space," says European Commission Vice-President Andrus Ansip in launching the Data Economy initiative. See European Commission, *Building a European Data Economy Press Release*, 12 January 2017.

63 When the general data protection regulation takes effect in May 2018, the Article 29 Working Party, which governs intra-EU cooperation on cross-border data sharing rules, will be rechristened the European Data Protection Board.

64 See the 16 November 2015 letter co-signed by 14 leading startup associations regarding gold plating and the general data protection regulation. The letter was addressed to the "trilogue" negotiating the final provisions of the GDPR and co-ordinated by Allied for Startups, a European umbrella group.

'At least 14 European countries have data localisation requirements of one sort or another.'

this data – and, crucially, who is able fairly and legitimately to profit from it – becomes extreme and acute.[65]

Perhaps not surprisingly, few areas give rise to more confusion in the European context – not least of which because of far-reaching European legal notions of "personal data," and the pervasive tone of dismissiveness in Europe towards large American service providers operating with innovative new business models which are very popular with European citizens. Yes, Facebook is making money off of your data; that's what the company does. It provides you with a very good free service in return for the right to track how you behave when using that service. But the increasingly complex relationships between customers and service providers – who exactly is adding the value? And what value exactly is being exchanged? – are felt in other, more traditional sectors as well. More than is commonly understood, banks are becoming data-driven businesses – the value of the knowledge they are able to aggregate from the data they hold is fast outstripping the profit margins they make on routine financial transactions. And car companies, too, are starting to find as much value in the data their products generate as in the sales of the cars themselves. There's an important principle at work here, too. These companies – including the many startups destined to innovate in these fields – were set up to do business.[66]

They want and need a profit margin if they are to survive. And society wants and needs them to have incentives to deepen innovation and aggressively compete to provide new and better services.

So where does data come in?

Put simply, Europe has created a system which – ironically – turns the traditional power relationship on its head. The GDPR gives extensive and deep rights to individuals over their personal data – declaring, among other things, that any data which can be "connected with a person," i.e., data which someone, somewhere could conceivably match with other data sets to determine who generated the data is "personal," and therefore subject to strict rules governing access, use, reuse and cross-border transfer. In principle, this approach makes sense; Europeans have led the way in establishing crucial human rights in the use and processing of data, and subsequent events – most notably the dramatic revelations of U.S. intelligence contractor Edward Snowden – have shown the need for a rights-driven framework in this area. But the law sets the bar so high that much of what is considered modern data analytics – including the need to combine disparate datasets for new insights – could easily fall outside of it. Take geospatial data, for instance. Is it personal? Not if it has no personal information. But what if the satellite could zoom in on you and

---

65 See Higginbotham, op. cit.

66 Sergey Filippov, "Data-Driven Business Models: Powering Startups in the Digital Age," *European Digital Forum Digital Insight* (Brussels and London: Lisbon Council and Nesta, 2014).

'Any measures to end "unjustified" data localisation should be accompanied by a broad and credible initiative to strengthen the European cybersecurity system.'

take a picture of you in your yard? Wouldn't that be personal?[67] In the end, it will fall to European courts to decide much of this, and that's bad news; legal uncertainty is poor soil for businesses to grow in. And it could wind up slowing European progress in this vital field – or perhaps just driving the important, cutting-edge business elsewhere.

At the same time, European regulators have complained, correctly, that standard service agreements – the contract between the consumer and the company that governs many Internet-based businesses – are too complicated. And they are right. But what is missing is not simply greater legal definition drafted by lawyers. What is missing is a broader social compact regarding what is going on – of where the users' rights begin and end and how those rights might best be exercised for the greatest social good. Society is missing a simple concept, an off-line users' agreement so simple that even children can understand it. The important thing is that the prevailing concept is simple; it can't be the subject of pages and pages of legal text. It should be an idea that anyone can grasp.

We believe that there is a relatively simple way of dividing up and articulating the rights of all users in this case, uniting them behind a simple four-point framework which can be easily communicated. First and foremost, there are two categories of "co-owners" in this concept. There are the "data-producers," which includes the people who

use products (phones, cars, web sites) that produce trackable data. And there are the "data gatherers" that own the phones, cars and websites where the activity is tracked. We call these parties "co-owners," or perhaps more precisely "co-producers" because they produce the data sets together. And we believe they share common, but non-exclusive rights and responsibilities – each to the other – for the use to which their "co-owned" data is put. Here's a four-part scheme, based on four principles, with concrete policy recommendations:

I. **All parties involved in data production are "co-owners" of the data.** Crucially, these rights should be non-exclusive. Data use and reuse should be built into the terms of use contracts by default, and that point should be understood well enough that users don't need to read the terms and conditions to get it. The companies that make the devices and the websites where the data is gathered and produced will reuse it; they should be transparent about this. And they should be required to produce annual summaries – something like a shareholders annual report – to all of their customers, describing their data policies, and, in particular, how they are using data, what their research objectives are, what, perhaps, they have learned. The other owners of the data – the people who produce it – should have the right of portability, allowing them to withdraw their data at will.

---

67   Efrén Díaz Díaz, "The General Data Protection Regulation Expands the Definition of Personal Data," *Open Data Institute*, 21 September 2016.

'The free flow of data may need to be guaranteed explicitly in a fifth freedom guaranteed by the European Union Treaty.'

To be clear, we do not propose a new exclusive right to individual data ownership. "Co-ownership" is about joint control, and neither the "data producer" nor the "data gatherer" has an exclusive right over the other one. To the contrary, it's a joint responsibility, like parenting, in which both co-manage, respecting each other and communicating regularly. Data gatherers should be encouraged to reuse data, including for analytical purposes other than those for which it was gathered. But there must be full transparency about this. Companies should publish annual Data Producer Reports, written for the people whose individual machine-generated data they possess.[68] These reports should tell the "data producers" what is being done with the data, i.e., with whom it is being shared/to whom (if anyone) it is being sold and/or how it is being used for research and analytics. If the data "producers" are somehow unhappy with this, they can invoke their right to portability and withdraw their data. Best practice argues for clarity and transparency. But transparency is not the same as the right to a veto; however, with the help of data portability, it does amount to an opt out for citizens who wish to leave the service.[69]

Concretely, we propose:

- Standard contracts are drafted that include "cross-licensing by default," ensuring data portability for individuals (as envisaged by the general data protection regulation) as well as for machine-generated data.

- Data portability should be provided through direct download and user friendly APIs as much as possible in real-time. This could be encouraged with data portability benchmarking through a set of "quality criteria" inspired by the eight principals of open government data laid down in 2007.[70]

- Stakeholders should unite to pursue ambitious, principled self-regulation in the forms of "seals of approval" over contracts, following the model of the AG Data Transparency Evaluator developed by the U.S. agricultural industry.

- New legislation should require annual reporting from large data gatherers to all data producers about how aggregated data has been used and reused – for what ends, and, perhaps, what has been learned. These should be broad, easily readable, publicly-available reports, produced on an annual basis.

---

68 To be clear, these reports should be broad and anonymised, like shareholder reports – one report for all shareholders, and publicly accessible. The point is to inform "co-owners" how the co-owned data is being used. The goal is not to report to them on how their individual data points were or were not processed or used. Among the many advantages of these reports would the additional incentives it would give "data gatherers" to do more socially useful work with the data they possess.

69 One area of legal uncertainty is the status of the data generated by individuals who are deceased. Do the rights of those individuals pass on to the heirs? The answer to the question is beyond the scope of this paper, but the logic of "portability" does imply that the heirs would have the right to withdraw the data of the deceased if they chose to do so. There may well be "transparency" issues involved here as well; many heirs cannot naturally know the extent of the deceased's digital footprint. They may need new legal rights to learn the full picture, so they can make informed judgments. And those legal rights, in turn, may require new procedures for contacting companies to inform them that the digital "rights" have transferred due to death. This could require legislative initiative, and could take place through the "once-only" process. "Digital wills" may be a concept whose time has come.

70 The eight principles are complete, primary, timely, accessible, machine processable, non-discriminatory, non-proprietary and licence free. They were adopted at a seminar in 2007 and have become extremely influential in the ensuing decade. Visit https://opengovdata.org/ for more.

'Few issues exemplify the problems in data management more than the enormous amount of confusion over the role and status of data being gathered by companies and created by human activity.'

II. **Government should lead.** The innovation we propose should start with more and better open government data. Government should set the tone and take the lead on best practice; all public data should be open by default. When possible, government data should be published through application programming interfaces, or APIs, to promote ease of use. Operationally, this means better and faster uptake for the 2013 public sector information directive. And it also means widening the directive's scope towards "data of public interest," as the European Commission has proposed, and taking concrete steps to populate and popularise a "data commons," where companies, individuals and public authorities could share pools of data. NB: this is not the same thing as "co-ownership." Under "co-ownership," the data producers and gatherers enjoy rights and responsibilities towards one another; they take decisions on the data's use, collectively, collaboratively and transparently. Under the commons, the data exists freely, anonymised and aggregated, non-rivalrously, all rights lifted.

This is a new and potentially exciting idea, whose parameters have only begun to be conceived and described in academic and other literature.[71] The key will be creating an enabling regime and better incentives to make this happen. In that, governments can play a vital role by adopting "best practice" themselves. First and foremost, they should ensure the release of lots of good, quality, usable data to a "commons," where it can be used widely and non-exclusively. But, should they choose to move forward on this, they must also use legislation to create a climate in which private companies are more comfortable sharing "publicly useful data" – perhaps offering them the chance to use the commons themselves in exchange for their willingness to contribute to it. We think this might be an easier challenge than is often thought; data-driven companies more than others understand what public good could arise from greater sharing and aggregation. But they have some legitimate issues to solve before they can join such a system – including the issue of liability in the event of sharing. We will discuss this concern in more detail in the next section.

---

71 The OECD defines a "data commons" as "an area where some data is shared publicly after adequate anonymisation and aggregation." See OECD, *Data-Driven Innovation: Big Data for Growth and Well-Being*, op. cit. See also, European Data Protection Supervisor. *EDPS Opinion on Coherent Enforcement of the Fundamental Rights in the Age of Big Data* (Brussels: EDPS, 2016).

'In the age of the Internet, the value of many goods and services is no longer in the margin the offerings command at sale but in the knowledge their use generates.'

Concretely, we propose:

- Governments should continue the progress in opening up data and services. The European Commission should closely monitor the transposition of the public-sector information directive. And they should work with EU member states to enlarge the scope of the public-sector information directive to include "data of public interest," as proposed in digital single market mid-term review.[72]
- Governments should work to radically improve data quality, moving towards standardised releases in API formats. To be reused, data should be clean, machine readable, standardised and with good metadata. Opening up data should be integrated in data management processes, and not rest on an *ad hoc* process. The European Commission should increase its support for implementing and assessing the progress of the new European Interoperability Framework recommendations in this sense.[73]

III. **Reuse by default.** Anonymised data should be made routinely available to third parties from completely different sources and sectors. Data analytics and other companies need to be able not only to access, but also to aggregate and reuse data even when they did not contribute to the production of this data. Data reuse is particularly important for technological innovation. AI and machine learning requires massive amounts of data from many different implementation of similar processes. However, most European companies typically are not allowed to aggregate and reuse data, mainly due to concerns about loss of control over sensitive commercial data they might have access to. Even when data sets are made available, limited interoperability requires extensive resources. In a typical "big data" project, 50% to 80% of the

costs will go towards preparing the data for analysis. In that sense, increased adoption of common, non-proprietary standards would lower the costs of data-driven innovation and increase adoption. Ideally, defining standards and adopting them should be a market-driven process, but there is a role for government to support and lead by example here, mainly through the publication of high quality open government data and by pursuing the widest possible adoption of cost-effective common standards within public administration (once-only).[74]

What's more, much data held privately can be very useful for addressing societal challenges. Governments and social innovators would benefit from using aggregated data held by companies for addressing societal challenges. We

---

72 European Commission, *Communication on the Mid-Term Review on the Implementation of the Digital Single Market Strategy: A Connected Digital Single Market for All*, op. cit.

73 European Commission, *Communication on European Interoperability Framework: Implementation Strategy* (Brussels: European Commission, 2017).

74 David Osimo, Giorgio Micheletti and Gabriella Cattaneo, *Technical Barriers to Data Sharing in Europe* (Milan and Barcelona: IDB and Open Evidence, 2016).

'What is missing is a broader social compact regarding what is going on – of where the users' rights begin and end and how those rights might best be exercised for the greatest social good.'

should provide appropriate incentives and frameworks to do that. In particular, companies that share anonymised data with the public – putting it out as trend data, or sharing it with government to help them to understand traffic, energy use or other patterns – should not be held liable for data disclosure. This is very important; many companies say they would share more non-personal data for public good – a movement that began in the 2000s around the concept of "open innovation."[75] But they worry about law suits arising from the release of anonymised data sets based on data that came at some level from personal activity. This will require the development of a better agreement on what constitutes "good governance" in data sharing and a recognition that much data gathered is proprietary, including of course the analytics built on top of it. Greater disclosure of real-time data would require a "safe harbouring" provision for companies that do share anonymised data – such as city traffic data, or electric-grid usage – in the public interest; these companies don't want to cooperate on the release of anonymised public-interest data one day only to be sued by a customer for breach the next. It is in our interest to make it easier – and to encourage them – to do so.

But we might in that context remember that "open innovation" began in the private sector, when companies realised that it was in their interest to share some of their intellectual property to spread R&D costs more evenly among themselves and perhaps more easily generate the kind of "market-shaking" insight that benefits all players.[76] Only later did governments follow suit with "open-government" initiatives which brought similar principles to the public sector. This time, we believe the order may best be reversed. The private sector is more ripe and more ready for game-changing initiatives than is frequently thought. As we argue in the previous section, we believe governments should move first and decisively to release more data into the "data commons." And we believe governments should work actively and assertively to resolve existing disincentives through legislation, and make it easier for private-sector companies to join the "data commons" as well. The fact that those same companies would have the opportunity to draw from, and learn from, much larger data sets on a range of topics – fuelled by broader, more enabling rules on use and reuse of anonymised data sets – would serve as an additional incentive not just to take part in the commons but perhaps to undertake the dramatic first-mover initiatives needed to start this radical new way of sharing in the first place.

---

75 Don Tapscott and Anthony D. Williams. *Wikinomics: How Mass Collaboration Changes Everything* (London: Penguin, 2008).

76 Ibid.

'Data use and reuse should be built into the terms of use contracts by default, and that point should be understood well enough that users don't need to read the terms and conditions to get it.'

Concretely, we propose:

- Government and industry should define standard contract clauses to foster data reuse and to create a climate of trust between data gatherers and analytics firms. The European Union could lead on this, offering a shortened, simplified, easy-to-understand model contract available on a voluntary basis. The text should be composed of short, easy-to-understand statement and fit on a small card.

- Governments should fund pilots and innovation spaces to allow trusted data sharing and mutual understanding between the different players. These could lead to information sharing and a new "data commons" of anonymised data sets available for broad, non-rivalrous use.

- The European Commission should extend the forthcoming public-sector information request repository to include "private data requests" that will monitor requests for access to company data by all players to facilitate sharing and the needs not addressed by the market.[77]

- The European Commission and EU member states should extend the public-sector information directive to include rules and conditions for access to specific privately held datasets for reasons of public value creation, providing safeguards and incentives for companies releasing the data, as proposed in the mid-term review of the digital single market strategy.

IV. **Full portability for all data producers; and portability-based access for competitors in some sectors.** Not everyone involved in the data value chain is a producer of data. To be sure, many people are developing businesses based on the use of data that they did not produce – and do not "co-own" – themselves. Access to customers for these important market participants needs to be protected, and in particular, the law must avoid facilitating consumer lock in through overly aggressive use of restrictive data access rules. Thanks to some existing requirements, customers who wish to switch services – taking their data elsewhere – are able to do so. But there's a "use case" that falls somewhere short of those exceptions – cases where companies that wish to provide a service need access to data held by other companies that are unwilling to provide it. To date, there have been examples of regulations in specific sectors, such as the payment services directive II, which

---

Concretely, we propose:

- Forced data sharing as in the case of the payment services directive II should happen exceptionally, based on clear cases of market failure, and on a sector-by-sector basis.

- Data portability should be the preferred method for dealing with these cases, offering consumers easy and easy-to-execute power to move their data as and when they want.

---

77 European Commission, "PSI Request Repository SMART 2016/0088: Draft Technical Specifications," *Annex to the Ex-Ante Advertisement* (Brussels: European Commission, 2016).

'Data gatherers should be encouraged to reuse data, including for analytical purposes other than those for which it was gathered. But there must be full transparency about this.'

requires access be granted to "payment service providers" to data held by other banks or service providers to fulfil a service. And Regulation 715/2007, which grants independent automobile repair shops access to otherwise proprietary car data (and with the arrival of connected cars, there are calls for widening the access requirements to include real-time data). We believe such legislative measures should remain limited to well-demonstrated circumstances of market failure. And data portability should itself be strengthened to provide an effective and agile way for customers to switch providers and allow data creators to grant access to competing firms.

VI. Conclusion

Underlying all of these proposals is one clear concept that will be vital for Europe's success in the next round of modern economic development: sharing data is good for society and necessary for competitiveness. We all benefit from it, so long as it is done effectively, carefully and within the deftly defined confines of the law. And, if the current set of arrangements in Europe has one big drawback, it is the unspoken implication behind much of the legislation that sharing data is bad. To be sure, we need to ensure the safeguards are there, but once they are there we need to move beyond fear and misunderstanding. Otherwise, Europe is destined to be an economic midget in an era when data analytics will define so much of the cutting-edge of global economic development.

The European data framework – first conceived nearly a decade ago and about to take its largest, widest effect with the 2018 implementation of the general data protection regulation – plugged some important holes in the system by setting up basic norms and guidelines for the protection of personal data. The times – and particularly the revelations of U.S. intelligence contractor Edward Snowden – showed that this was necessary. But life has moved on. Individual rights are still very important, but the system is moving towards one where our common responsibility towards one another – and our common interest in developing and contributing to a better data-driven economy – should be clear as well.[78] For that, we need a simpler concept that will

make that transition possible. We need an awareness and commitment to the value of creating a data-driven economy, one where the economic benefits are shared and the differing incentives (government, citizen and corporate) are aligned and activated. We believe a concept of co-ownership built around the four key principles and three areas for priority initiative outlined above comes closest to doing that.

---

78 This is particularly important at the time of fast developing business models, when the challenge of AI creates new opportunity for some – and new risk for others.

References and Further Reading

Barbero, Martina, Jo Coutuer, Régy Jackers, Karim Moueddene, Els Renders, Wim Stevens, Yves Toninato, Sebastiaan van der Peijl and Dimitry Versteele. *Big Data Analytics for Policymaking* (Brussels: European Union, 2016)

Barocas, Solon, and Helen Nissenbaum. "Big Data's End Run around Anonymity and Consent" in Julia Lane, Victoria Stodden, Stefan Bender and Helen Nissenbaum (eds.). *Privacy, Big Data, and the Public Good: Frameworks for Engagement* (Cambridge: Cambridge University Press, 2014)

Bourgeois, Jacques, Cameron F. Kerry, William R. M. Long, Maarten Meulenbelt and Alan Charles Raul. *Essentially Equivalent: A Comparison of the Legal Orders for Privacy and Data Protection in the European Union and United States* (Chicago: Sidley Austin LLP, 2016)

Bozio, Antoine, and Pierre-Yves Geoffard. *Rapport au sécretaire d'état chargé de l'industrie, du numérique et de l'innovation* (Paris: Conseil national de l'information statistique, 2017)

Brill, Julie. *Transatlantic Privacy After Schrems: Time for an Honest Conversation*. Keynote Address at the Amsterdam Privacy Conference, 23 October 2015

Cattaneo, Gabriella, Giorgio Micheletti, Alys Woodward and David Osimo. *Data Ownership and Access to Data: Key Emerging Issues* (Milan and Barcelona: IDC and Open Evidence, 2016)

Ctrl-Shift. *A New Paradigm for Personal Data: Five Shifts to Drive Trust and Growth* (London: Facebook, 2016)

The Economist. "Briefing: The Data Economy," *The Economist*, 06 May 2017

European Commission. *Building a European Data Economy* (Brussels: European Commission, 2017)

----------------. *Exchanging and Protecting Personal Data in a Globalised World* (Brussels: European Commission, 2017)

----------------. *Proposal for a Regulation on Establishing a Single Digital Gateway to Provide Information, Procedures, Assistance and Problem Solving Services* (Brussels: European Commission, 2017)

European Data Protection Supervisor. *EDPS Opinion on Coherent Enforcement of the Fundamental Rights in the Age of Big Data* (Brussels: EDPS, 2016)

------------------. *Meeting the Challenge of Big Data: A Call for Transparency, User Control, Data Protection by Design and Accountability. Opinion 7/2015* (Brussels: EDPS, 2015)

'All public data should be open by default. Government should set the tone and take the lead on best practice.'

-----------------. *Towards a New Digital Ethics: Data, Dignity and Technology. Opinion 4/2015* (Brussels: EDPS, 2015)

European Parliament and Council of the European Union. *Directive of the European Parliament and Council concerning Measures for a High Common Level of Security of Network and Information Systems across the Union* (Brussels: European Parliament and Council, 2016)

European Political Strategy Centre. *Enter the Data Economy: EU Policies for a Thriving Data Ecosystem* (Brussels: European Commission, 2017)

-----------------. *Hearing: Building a European Data Economy*, 28 March 2017. The transcript is available on the European Commission website at https://ec.europa.eu/epsc/events/hearing-building-european-data-economy_en

Federal Ministry for Economic Affairs and Energy (Germany). *White Paper: Digital Platforms; Digital Regulatory Policy for Growth, Innovation, Competition and Participation* (Berlin: BMWi, 2017)

Gellman, Robert. *Fair Information Practices: A Basic History* (10 April 2017). Available at SSRN: https://ssrn.com/abstract=2415020

HM Government. *Personal Data: Review of the Midata Voluntary Programme* (London: Department for Business, Innovation and Skills, 2014)

Hagel III, John, and Jeffrey F. Rayport. *The Coming Battle for Customer Information* (Cambridge: Harvard Business Review, 1997)

Hardjono, Thomas, David Shrier and Alex Pentland. *Trust::Data: A New Framework for Identity and Data Sharing* (London: VisionaryFuture, 2016)

Henke, Nicolaus, Jacques Bughin, Michael Chui, James Manyika, Tamim Saleh, Bill Wiseman and Guru Sethupathy. *The Age of Analytics: Competing in a Data-Driven World* (San Francisco: McKinsey Global Institute, 2016)

Landström, Ola, and Erik Dahlberg. *Data Flows: A Fifth Freedom for the Internal Market?* (Stockholm: National Board of Trade Sweden, 2016)

van der Marel, Erik. *Digital Investments, Data and Growth in Europe: A Framework for Analysis* (Brussels: ECIPE, 2016)

Organisation for Economic Co-operation and Development. *Data-Driven Innovation: Big Data for Growth and Well-Being* (Paris: OECD, 2015)

'Data reuse is particularly important for technological innovation.'

--------------------. "Benefits and Challenges of Digitalising Production" in OECD, *The Next Production Revolution: Implications for Governments and Business* (Paris: OECD, 2017)

Osborne Clarke LLP. *Legal Study on Ownership and Access to Data* (Brussels: European Commission, 2016)

Pentland, Alex. "Reality Mining of Mobile Communications: Towards a New Deal on Data" in *The Global Information Technology Report 2008-2009 – Mobility in a Networked World* (Fontainebleau and Geneva: Insead and World Economic Forum, 2009)

-----------. *Social Physics: How Social Networks Can Make Us Smarter* (New York: Penguin Books, 2014)

Pew Research Center. *Privacy and Information Sharing: Scenarios* (Washington: Pew Research Center, 2016)

Poikola, Antti, Kai Kuikkaniemi and Harri Honko. *MyData: A Nordic Model for Human-Centered Personal Data Management and Processing* (Helsinki: Open Knowledge Finland, 2015)

Ratas, Jüri. *The Next Goal of EU Should Be Free Movement of Data* (Tallinn: Republic of Estonia, 2017)

Reimsbach-Kounatze, Christian. *Maximising the Economic and Social Value of Data: Understanding the Benefits and Challenges of Enhanced Data Access* (Paris: OECD, 2016)

UN Global Pulse, *Big Data for Development: Challenges and Opportunities* (Geneva, United Nations, 2012)

Veronesi, Lorenzo, Gabriella Cattaneo and Giorgio Micheletti. *Industrial Data Platforms: Key Enablers of Industry Digitisation* (Milan: IDC, 2016)

World Bank. *Principles on Identification for Sustainable Development: Toward the Digital Age* (Washington, DC: World Bank, 2017)

World Economic Forum. *Personal Data: Emergence of a New Asset Class* (Geneva: World Economic Forum, 2011)

-----------------. *Unlocking the Value of Personal Data: From Collection to Usage* (Geneva: World Economic Forum, 2013)

Zech, Herbert. "Data as a Tradeable Commodity and the New Instruments for Their Protection," in Alberto de Franceschi, *European Contract Law and the Digital Single Market: the Implications of the Digital Revolution* (Cambridge: Intersentia, 2016)

Acknowledgements

'Governments and social innovators could benefit from using aggregated data held by companies for addressing societal challenges. We should provide appropriate incentives and frameworks to do that.'

- **Katalin Imrei**, policy officer building the data economy, DG-communications networks, content and technology, European Commission

- **Gertrud Ingestad**, director-general, DG-informatics, European Commission

- **Delilah Al Khudhairy**, director strategy and work coordination, Joint Research Centre, European Commission

- **Torsten Körber**, professor of civil law, competition law, insurance law, corporate law and regulation law at the Georg-August-Universität, Göttingen, Germany

- **Lenard Koschwitz**, director European affairs, Allied for Startups

- **Taavi Kotka**, former government chief information officer, Estonia; vice-president of sales, Plumbr

- **Baldur Kubo**, project manager Shareminder team, Cybernetica (Estonia)

- **Andres Kütt**, adviser, Estonian Information System's Authority (X-Road)

- **Ivar Laur**, head of analysis division, Tax and Customs Board (Estonia)

- **Urmo Parm**, IT adviser, Estonian Data Protection Board

- **Antti Peltomäki**, deputy director-general, DG-internal market, industry, entrepreneurship and SMEs, European Commission

- **Jiri Pilar**, legal and policy officer building the data economy, DG-communications networks, content and technology, European Commission

- **Antti Poikola**, researcher, Institute for Information Technology, Helsinki University of Technology; chair, Open Knowledge Finland

- **Celina Ramjoué**, deputy head of unit data innovation and policy, DG-communications networks, content and technology, European Commission

- **Christian Reimsbach-Kounatze**, information economist and policy analyst, directorate for science, technology and innovation, digital economy policy, Organisation for Economic Co-ooperation and Development

'Data portability should itself be strengthened to provide an effective and agile way for customers to switch providers.'

About the Lisbon Council

The Lisbon Council for Economic Competitiveness and Social Renewal asbl is a Brussels-based think tank and policy network. Established in 2003 in Belgium as a non-profit, non-partisan association, the group is dedicated to making a positive contribution through cutting-edge research and by engaging political leaders and the public at large in a constructive exchange about the economic and social challenges of the 21st century.

www.lisboncouncil.net

The Lisbon Council asbl

IPC-Résidence Palace
155 rue de la Loi
1040 Brussels, Belgium
T +32 2 647 9575 F +32 2 640 9828
info@lisboncouncil.net
Twitter @lisboncouncil

Digital Knowledge Economy