STI Policy Note

October 2015

Data-driven Innovation for Growth and Well-being

What Implications for Governments and Businesses?

More data are being generated every week than in the last millennia. With the accelerating digitalisation of social and economic activities, the flows of data – the equivalent of around 50 000 years of DVD-quality video every single day, a figure set to considerably rise as smart devices and the Internet of Things become pervasive – are such that the implications for the economy and society are colossal.

On the positive side, they hold the promise of significantly accelerating research and the development of new products, processes, organisational methods and markets – a phenomenon known as data-driven innovation (DDI). This will result in greater productivity across the economy, as available evidence suggests that firms using DDI have raised productivity faster than non-users by around 5-10%. DDI can and is already helping address social and global challenges, including climate change and natural disasters, health and ageing populations, water, food and energy security, urbanisation, and issues of public governance.

But considerable challenges are also ahead. The growing use of data will accentuate many of today's pressing issues, like questions of privacy and security, or the impact of technology on jobs. And new concerns will arise, for example around automated decision making, data-driven discrimination, and a likely shift in power derived from a new "data divide," based on who owns, collects and analyses the data.

The quick read

In an October 2015 report, the OECD argues that data and data analytics have become an essential driver of innovation akin to R&D, and that governments must redefine 21st Century infrastructure to include not only broadband networks, but also data itself.

Seizing the benefits from DDI requires government action. Policies are needed to encourage investments in data, promote data sharing and reuse, and reduce barriers to cross-border data flows that could disrupt global value chains. They thereby need to strike the right balance between the social benefits of enhanced reuse and sharing of data and analytics, and individuals' and organisations' legitimate concerns about such openness, including the protection of privacy and intellectual property rights.

Governments should also focus on small and medium enterprises, which face severe barriers to the adoption of DDI-related technologies. They must address shortages of data specialist skills, which points to missed opportunities for job creation. And they must anticipate and address the disruptive force of DDI, with far-reaching effects on the economy and overall well-being.

Data: a new infrastructure for innovation in the 21st century

Data are an infrastructural resource – a form of capital that cannot be depleted and that can be used for a theoretically unlimited range of purposes. Physical infrastructure such as roads and bridges enables benefits to 'spill over', for instance, by fostering trade and social exchanges. In the same way, greater access to data also has beneficial spill-overs, whereby data can be used and re-used to open up significant growth opportunities, or to generate benefits across society in ways that could not be foreseen when the data were created.

But some of the spill-overs of data cannot be easily observed or quantified (e.g. socialisation and behavioural change, cultural and scientific exchange, or greater levels of trust induced by transparency). As a result, countries – and governments in particular – risk under-investing in data and data analytics and may end up giving access to data for a narrower range of uses than socially optimal. This risks undermining countries' capacity to innovate, as data and its analysis have become a fundamental input to innovation, akin to research and development (R&D).

In this sense, data are the new "R&D" for 21st century innovation systems. Data and R&D share a number of common properties: both are intangible assets that can be combined with other innovation investments like training, software, organisational change, etc.; both enable the creation of knowledge with positive externalities or spill-overs across society; and both face the challenge of these externalities possibly negatively impacting on incentives to invest. Organisations may well be able to capture the private benefits of their investment in data, but do not yet always see the larger benefits that the data can bring to society.

Although successful innovation requires a bundle of investments in addition to R&D, innovation policy all too often focuses most prominently on R&D, for example through R&D tax credits, or government investments in basic R&D. Governments are thereby not yet sufficiently considering the crucial role of data in shaping or strengthening innovation performance.

Policy considerations

DDI requires the attention of policy makers involved in innovation, just like R&D today. Governments should develop an appropriate innovation policy mix that encourages investments in data (its collection, curation, and reuse), while addressing the low appropriation of returns to encourage data sharing. The combination of intellectual property right (IPR) licences and alternative incentive mechanisms such as data citations, data donation or philanthropy need to be considered further.

Obstacles to the reuse and sharing of data should be examined carefully with an eye to enhancing the benefits that can be reaped from data. Non-discriminatory access regimes, including data commons or open access regimes, should be explored, as a means to support the production of public and social goods without requiring governments or businesses to pick winners (either users or applications).

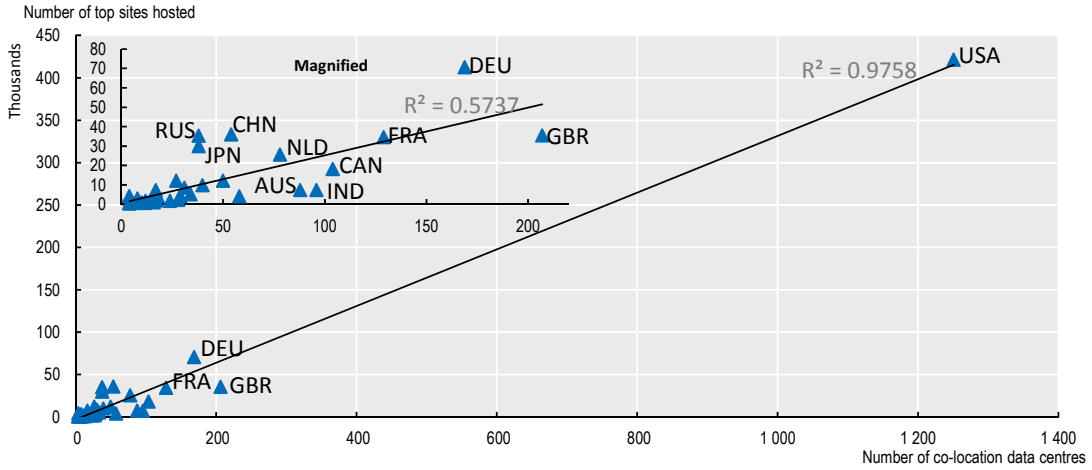
Coherent data governance frameworks should be developed. Access to data should not necessarily be free, nor unregulated: it is critical to strike the right balance between openness and the social benefits of greater access and reuse of data, and the legitimate concerns of those whose privacy and IPRs may be negatively affected.

The open Internet and global value chains

A global data ecosystem is emerging where data and analytic services are traded and used across sectors and national borders. The flows of vast amounts of data have led to the growth of global value chains (GVCs), in which companies increasingly divide up their data-related processes – hosting, storage and processing – across many countries.

Figures on the distribution of data-driven services are not known, nor are the magnitudes of cross-border data flows. But analysis of the world's top Internet sites suggests that data-driven services are disproportionately concentrated in the United States, which alone accounted for more than 50% of all top sites hosted in the OECD area. Canada, Germany, France, Ireland, the Netherlands, Japan and the United Kingdom, as well as China, India and the Russian Federation are catching up as they increase their contribution to global trade in ICT-intensive services.

Figure 1. Top locations by number of co-location data centres and top sites hosted, 2013



Sources: Based on Pingdom, 2013; and www.datacentermap.com, accessed 27 May 2014.

Countries that have the largest numbers of top sites are also those that have the highest number of co-location data centres and in turn, the top locations for data-intensive services. Major exporters of ICT services and top locations for data-driven services are more likely to be larger destinations of cross-border data flows. As a consequence, the leading OECD area importers of ICT-related services are also the major sources for trade-related data; this includes in particular trans-Atlantic flows.

Policy considerations

Barriers to the open Internet, legitimate or otherwise, can limit the effects of DDI in particular in economies where deployment of data-driven services is poor due to failures in co-location and backhaul markets. Barriers to the open internet can be the results of business practices or government policies. They may also have a legal basis such as the protection of privacy and IPRs as well as security. However, these barriers can have an adverse impact on DDI, e.g. by limiting trade and competition.

Governments looking to promote DDI in their countries should take the *OECD 2011 Council Recommendation on Principles for Internet Policy Making* further into consideration as well as ongoing OECD work to develop a better understanding of the characteristics, and the social and economic impact of the open Internet.

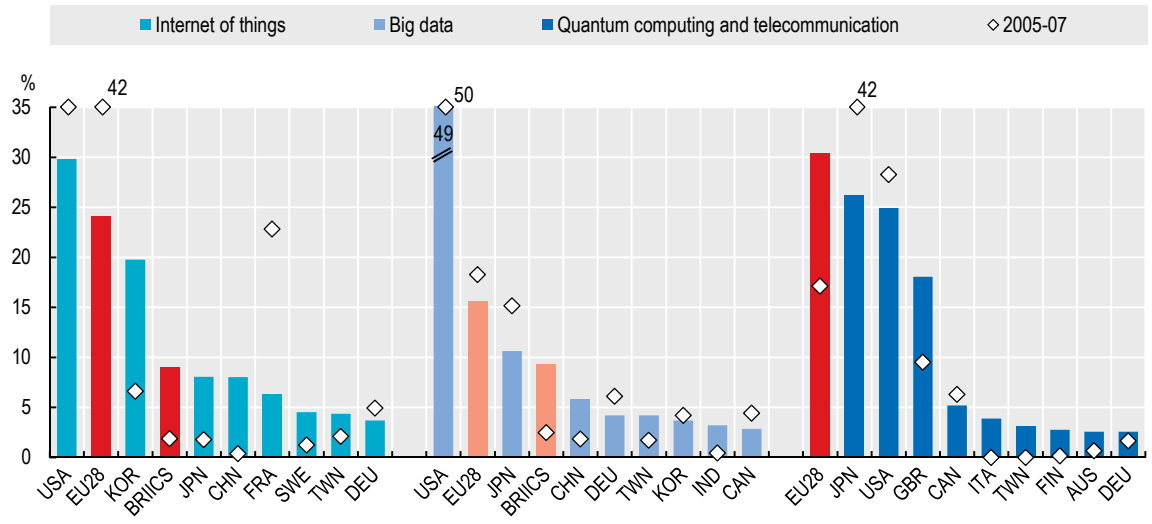
Invest for the future

The generation of data requires upfront investments, including in the development of Internet services as well as in smart devices that generate data via the Internet of Things (IoT). Access to technologies for data analysis is just as critical for realising the potential of DDI. Countries with enhanced capacities to supply and adopt these factors will be in the best position to benefit.

Figures on international patent filings provide evidence that technologies related to DDI are rapidly increasing in importance when it comes to inventive activity. Since 2007, the number of patent filings related to the IoT, big data analytics, and quantum computing and telecommunication have grown at two digit rates; in 2012, the latest year for which data is available, at more than 40% year-on-year. But the provision of DDI-related technologies is concentrated in few economies, with the United States leading in terms of the number of filed patents, followed by Canada, France, Germany, Korea, Japan, and the United Kingdom, as well as China.

Figure 2. Top players in IoT, big data and quantum computing technologies, 2005-07 and 2010-12

Share of IP5 patent families filed at USPTO and EPO, selected ICT technologies



Source: OECD Science, Technology and Industry Scoreboard 2015, forthcoming. OECD calculations based on IPO (2014), Eight Great Technologies: the Patent Landscapes and STI Micro-data Lab: Intellectual Property Database, June 2015.

Policy considerations

Governments aiming to promote the supply of DDI-enabling technologies should consider supporting investments in R&D on key technologies such as big data analytics, cloud and high-performance computing, and IoT, but also on security and privacy enhancing technologies.

Through its national digital economy strategy, Canada for instance foresees investments worth CAD 15 million over three years to support leading-edge research in, and the commercialisation of quantum technologies. France intends to invest EUR 150 million to support R&D in five technologies identified as strategic: IoT, super and cloud computing, big data analytics, and security. Germany's Digital Agenda 2014-2017 also foresees the promotion of super computing, security and big data analytics, the latter particularly through two Big Data competence centres. And Japan's *Declaration to be the World's Most Advanced IT Nation* aims at furthering developments in data analytics, as well as in sensor and robotic related technologies among others.

Address barriers to the adoption of key enabling technologies

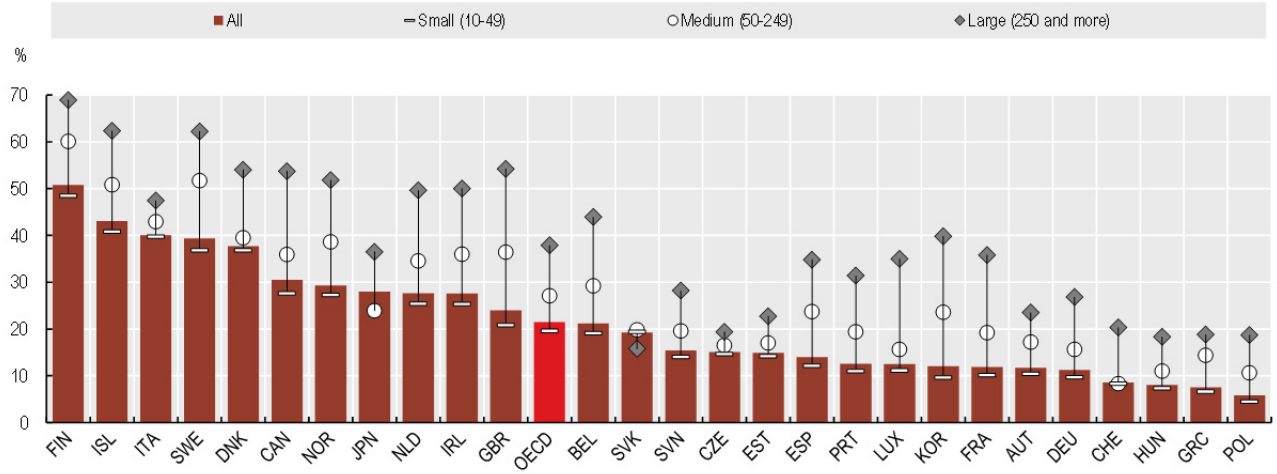
An important condition for leveraging DDI is the adoption of DDI-related technologies such as cloud computing. This is particularly critical for small and medium enterprises (SMEs), which lag in productivity relative to larger firms.

However, the adoption of these technologies by firms is still much below that of broadband networks or websites. This is partly because businesses have difficulty implementing organisational change due to limited resources, including the shortage of skilled personnel. In addition, the lack of appropriate standards and vendor lock-in due to proprietary solutions can also be a barrier to adoption. For cloud computing in particular, privacy and security concerns are also often listed as common barriers.

There are also regulatory barriers preventing effective adoption of some DDI-related technologies. Large-scale IoT users such as car manufacturers who need to control their own devices with their own SIM cards cannot do so in many countries, for example.

Figure 3. The diffusion of cloud computing in enterprises, by country and size, 2014

Percentage of enterprises with ten or more persons employed



Source: OECD (2015), OECD Digital Economy Outlook 2015, OECD Publishing, Paris. Based on OECD, ICT Database; Eurostat, Information Society Statistics and national sources, July 2014.

Policy considerations

Governments should further consider demand-side policies to encourage the adoption of DDI-related technologies. The *Information Economy Strategy* of the United Kingdom, for instance, aims at promoting the use of ICTs in businesses and organisations, especially SMEs, through a series of targeted activities including awareness raising, a web portal with online tools, and a network of digital advice centres offering training, mentoring, and access to voucher schemes.

The promotion of open standards in application programming interfaces (APIs) and data formats can be key, as adoption would boost interoperability and reuse of data and data-driven services, while enhancing competition among service providers.

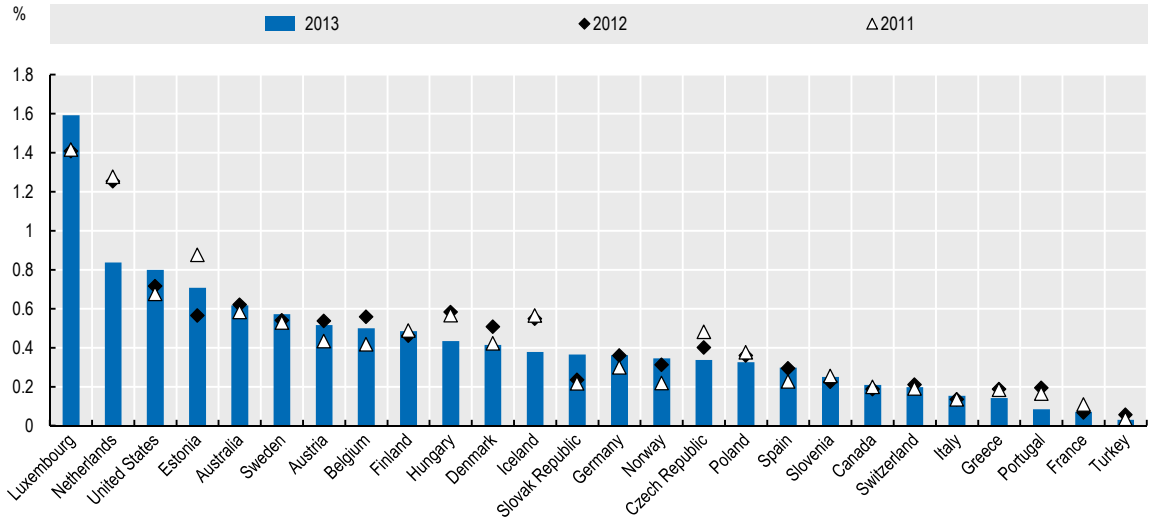
Governments should further support the development of a culture of digital risk management across society, as promoted by the *2015 OECD Recommendation on Digital Security Risk Management for Economic and Social Prosperity*.

Governments need to act if regulatory barriers are preventing adoption. For the IoT, for instance, removing regulatory barriers to entry in the mobile market would allow the million-device customer to become independent of the mobile network and would further competition.

Address the skills bottleneck

The increasing use of data and analytics across the economy has driven demand for new types of skills and jobs. OECD estimates show that data specialists in 2013 accounted for under 1% of total employment in the large majority of countries. In countries such as Portugal, France and Turkey, the share of data specialists was below 0.1%.

Figure 4. Data specialists as share of total employment in selected OECD countries, 2011-13



Source: Based on data from Eurostat, Statistics Canada, Australian Bureau of Statistics Labour Force Surveys and US Current Population Survey, March Supplement, February 2015.

The low availability of specialist skills may prove not only a barrier to adoption of DDI: surveys point to the shortage of skilled data specialists as one of the biggest impediments to making use of data analytics. Demand for statisticians, actuaries and mathematicians is expected to grow fastest. In the United States, these occupations have also seen the fastest growth in relative wages since 1999, suggesting – with the decline in the share in total employment – the country, like others, could be facing a shortage of skills.

Many countries are struggling to develop these specialist skills, however. OECD data reveals that from 7% to as much as 27% of adults in OECD countries still have no experience in using computers, or lack the most elementary skills, such as the ability to use a mouse. Only 6% of people in the OECD have the "highest level" of ICT skills. In countries such as Austria, the United States, Korea, Estonia, the Slovak Republic, Ireland and Poland, the share is 5% and below.

Policy considerations

National education systems, in collaboration with businesses, need to support the development of data-related skills starting with basic ICT skills. Educational needs extend beyond ICT to include science, technology, engineering and mathematics (STEM).

Education systems should also go beyond STEM to support the understanding of human behaviour and social systems as both qualitative and quantitative reasoning are needed to enhance the sense of responsibility of future data-informed decision makers.

The disruption to jobs needs to be anticipated, and benefits seized

Inherently disruptive, DDI may induce the "creative destruction" of established businesses and markets. Evidence suggests that creative destruction is an essential engine of long-term economic growth. With the global recovery still sluggish, business and policy leaders need to harness these forces to facilitate structural shifts to a stronger and more sustainable economic future.

Labour markets will be affected, as data and analytics enable the automation of an increasing number of tasks. This includes a wider range of intellectually demanding tasks, such as diagnosis based on analysis of medical images. It also includes use of a new generation of autonomous machines and robots, no longer restricted to narrowly defined environments, but that can be deployed and redeployed at much faster rates compared to current generation robots. More middle-income jobs may thus be negatively affected – jobs largely held by the segment of the population that "glues" our societies together.

However, certain categories of jobs are less susceptible to computerisation and data-driven automation, and new types of skills and jobs may emerge, like data specialist jobs which allow firms to innovate and compete. OECD data based on business innovation surveys show that innovative firms are significantly more likely to use skills, whether internally or externally, related to data analytics, software development, and product design than non-innovative firms.

Overall, the data-driven economy may provide more jobs that involve:

- solving unstructured problems, including problems that lack rules-based solutions
- working with new information, including making sense of new data and information for the purpose of problem solving or decision making, or to influence the decisions of others
- non-routine manual tasks, carrying out physical tasks that cannot be well described via rules because they require optical recognition and fine muscle control that continue to prove difficult for robots to perform.

While solving unstructured problems and working with new information will be particularly important for high-end jobs, carrying out non-routine manual tasks will become more and more important for low-paying jobs.

Policy considerations

Policy makers may need to address unemployment and inequality induced by skill- and capital-based technological change, and help workers adjust to the emerging data-driven economy.

A double strategy should be considered: promoting continuous education, training and skills development, while addressing the risks of worsening inequality in earnings in labour markets through social and tax policies. This is all the more necessary given the weak global recovery and lingering high unemployment in major advanced economies.

Policy makers should acknowledge that inequality could become a major issue, in particular if access to urgently needed high-quality education to take advantage of the job creation opportunities ahead is limited to a few.

Liability, transparency, and ownership: new areas of uncertainty to be addressed

Data analytics leads to new ways of decision making, through low cost and rapid experiments, often based on correlations and A|B testing, as well as through use of autonomous machines and systems able to learn from previous situations. This can lead to higher productivity as it accelerates the decision-making process, and in some cases even eliminates the need for human intervention.

But data-driven decision making could also unexpectedly lead to false results. This may be because of poor quality data, errors due to the inappropriate use of data and analytics, or caused by unexpected changes in the environment from which data are collected. Recent financial losses caused by unforeseen behaviour of algorithmic trading systems, like Knight Capital Group's loss of USD 440 million in 2012, are illustrations of this last point.

The risk of taking the wrong decisions raises questions about how to assign liability between decision makers, data and data analytic providers. The issue is exacerbated by the challenges associated with the concept of "ownership", which poses specific challenges when applied to data. In contrast to other intangibles, data typically involve complex assignments of different rights across different stakeholders who will typically have different power over the data depending on their role. In cases where the data are considered "personal", the concept of ownership is even less practical, since most privacy regimes grant certain explicit control rights to the data subject that cannot be restricted – see for example the Individual Participation Principle of the *OECD Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data*.

As data and data analytics become more pervasive, users need to be aware of their limitations, or they may unintentionally cause social and economic harm (or costs) to themselves, as well as to third parties. This is especially true given the incentives for the users of data and analytics to minimise the risks to third parties may be low, in particular when the costs of false decisions are mainly borne by third parties.

Policy considerations

Further thinking on the attribution of responsibility for inappropriate decisions and the attribution of liability between decision makers, data and data analytic providers is necessary, given externalised social and economic costs to third parties, including individuals. This also calls for further studies to understand the limitations to the concept of data ownership, in particular in the case of personal data.

It also calls for a careful examination of the appropriateness of fully automated decision making and the consideration of transparency requirements and human intervention in areas where the potential harm of such decisions may be significant (e.g. harm to the life and well-being of individuals, denial of economic or social rights).

However, policy makers need to acknowledge that transparency requirements may extend to the processes and algorithms underlying automated decisions, in which case these requirements may come into tension with existing proprietary intellectual property rights and the processes and algorithms at the core of certain businesses' operations.

Privacy, consumer protection, competition law and taxation: new challenges to regulation

The exploitation of data may raise serious concerns that touch on privacy, consumer protection, as well as competition and taxation concerns. Yet existing regulatory frameworks may be ill-suited to deal with the new challenge of DDI.

Comprehensive data collection can lead to the loss of privacy, with advances in data analytics making it possible to infer sensitive information, including from apparently non-personal data (e.g. meta-data). The misuse of such insights can affect core social values and principles, such as individual autonomy, equality and free speech. Discrimination enabled by data analytics, for example, may result in greater efficiencies, but also limit an individual's ability to escape the impact of pre-existing socio-economic factors. Meanwhile the applicability of core principles on which privacy protection relies (e.g. the definition of personal data and the role of consent) is being challenged.

The approach competition authorities use to assess potential abuses and harms of market dominance and mergers is also being challenged by DDI. These include in particular challenges in defining the relevant market, and in assessing the degree of market concentration, and potential consumer detriments due to privacy violation.

DDI may also challenge the paradigm used by tax authorities to determine where tax-relevant activities are carried out and value is created. Inherent in this is the difficulty in measuring the monetary value of data, determining data ownership, and acquiring a clear picture of the global distribution and interconnectedness of data-driven services.

As a whole, DDI may undermine the social fabric on which democratic market societies are based, as greater information asymmetry leads to shifts in power away: from individuals to organisations (consumers to businesses, citizens to governments); from traditional businesses to data-driven businesses, inducing the risk of market concentration and dominance; from governments to data-driven businesses, which can gain much more knowledge about citizens (and politicians) than governments can; and from lagging economies to data-driven economies. If unaddressed, this could exacerbate existing inequalities and lead to a new digital data divide.

Policy considerations

Governments need to promote the responsible use of personal data to prevent violations of privacy. Efforts to promote privacy-enhancing technologies and the empowerment of individuals through greater transparency of data processing, and data portability through initiatives such as midata (United Kingdom) and MesInfos (France) should be further considered. Governments also need to increase the effectiveness (i.e. resourcing and technical expertise) of privacy enforcement authorities. The adoption of privacy risk management frameworks requires further studies to understand how best to support existing privacy protection principles.

In addition, governments need to assess market concentration and competition barriers through better definitions of the relevant market and the consideration of potential consumer detriments due to privacy violations. This will also require fostering the dialogue between regulatory authorities (in particular in the area of competition, privacy and consumer protection).

Governments need to encourage improved measurement to help better assess the economic value of data, also to prevent base erosion and profit shifting through aggressive tax planning by firms seeking to reduce taxable income artificially or shift profits to low-tax jurisdictions.

Overview of policy considerations

Data-driven innovation may yet be in its infancy, but it is becoming a defining phenomenon of our era. Clinging to the status-quo is not an option. Rather, policy and decision makers need to be pro-active, and understand the contours of this development and how it may evolve.

Through its work on DDI, the OECD has analysed major economic and societal challenges that will require policy makers to consider the trade-offs, complementarities and possible unintended consequences both of their policy actions – and of inaction. The work resulted in an initial mapping across a range of policy issues – e.g. innovation, consumer policy, competition, and jobs and skills – while outlining a number of issues that warrant more in-depth analysis.

More generally, OECD analysis identifies two sets of challenges, or tensions, that need to be addressed concomitantly by policy makers in order to maximise the benefits of DDI, and mitigate the associated economic and societal risks:

1. the need to promote "openness" in the global data ecosystem and thus the free flow of data across nations, sectors, and organisations, and at the same time to address legitimate considerations of individuals' and organisations' opposing interests (including in particular their interests in the protection of privacy and intellectual property rights).
2. the need to activate the enablers of DDI such as funding technological development and the infrastructure to support and help diffuse DDI across the economy, and at the same time to address the effects of the "creative destruction" induced by DDI, in particular with a focus on small and medium enterprises and on labour markets.

These two sets of tensions are not unrelated. A closer look at the policy issues suggests that a move towards more "openness" may further the disruptive effects of DDI and thus lead to more "creative destruction". There is no one-size-fits-all optimal level of "openness"; instead the optimal level strongly depends on the domain and the cultural environment in question.

Addressing these tensions and finding the optimal level of openness is therefore complex and cannot be undertaken in silos; these require governments to invite the democratic participation of all citizens – in addition to stakeholders including civil society, the technical Internet community and business groups in the relevant domains – in order to be resolved. This also calls for a whole-of-government strategy, and international dialogue and coordination on the cross-border issues raised by DDI, to help maximise the benefits of DDI and mitigate associated risks and obstacles.

Further reading

Anderson, G. and J. Oderkirk (eds.) (2015), *Dementia Research and Care: Can Big Data Help?*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264228429-en.

OECD (2015a), *Digital Economy Outlook 2015*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264232440-en.

OECD (2015b), *Inquiries into Intellectual Property's Economic Impact*, http://www.oecd.org/sti/ieconomy/intellectual-property-economic-impact.htm.

OECD (2014), "Unleashing the Power of Big Data for Alzheimer's Disease and Dementia Research: Main Points of the OECD Expert Consultation on Unlocking Global Collaboration to Accelerate Innovation for Alzheimer's Disease and Dementia", *OECD Digital Economy Papers*, No. 233, OECD Publishing, Paris, http://dx.doi.org/10.1787/5jz73kvmvbwb-en.

OECD (2014c), *Digital Security Risk Management for Economic and Social Prosperity: OECD Recommendation and Companion Document*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264245471-en.

OECD (2015c), *OECD Science, Technology and Industry Scoreboard 2015*, OECD Publishing, forthcoming.

OECD (2014a), *Measuring the Digital Economy: A New Perspective*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264221796-en.

OECD (2013a), *Supporting Investment in Knowledge Capital, Investment and Innovation*, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264193307-7-en.

OECD (2013b), Recommendation of the Council concerning Guidelines Governing the Protection of Privacy and Transborder Flows of Personal Data, OECD Publishing, Paris, 11 July, www.oecd.org/sti/ieconomy/2013-oecd-privacy-guidelines.pdf.

OECD (2008), *Recommendation of the Council for Enhanced Access and More Effective Use of Public Sector Information*, OECD Publishing, Paris, 30 April, C(2008)36, www.oecd.org/internet/ieconomy/40826024.pdf.

OECD (2006), *Recommendation of the Council concerning Principles and Guidelines for Access to Research Data from Public Funding*, OECD Publishing, Paris, 14 December, www.oecd.org/sti/sci-tech/38500813.pdf.

Reimsbach-Kounatze, C. (2015), "The proliferation of data and implications for official statistics and statistical agencies: A preliminary analysis", *OECD Digital Economy Papers*, OECD Publishing, Paris, http://dx.doi.org/10.1787/5js7t9wqzvg8-en.

Ubaldi, B. (2013), "Open government data: Towards empirical analysis of open government data initiatives", *OECD Working Papers on Public Governance*, No. 22, OECD Publishing, Paris, http://dx.doi.org/10.1787/5k46bj4f03s7-en.

Directorate for Science, Technology and Innovation Policy Note

This series of Policy Notes is designed to make available, to a wider readership, selected studies that have been prepared for use within the OECD.

Comment on this Policy Note is invited, and may be sent to OECD, 2 rue André Pascal, 75775 Paris Cedex 16, France, or by e-mail to Christian.Reimsbach-Kounatze@oecd.org.

Please cite this note as:

OECD (2015), "Data-Driven Innovation for Growth and Well-Being. What Implications for Governments and Businesses?", Directorate for Science, Technology and Innovation Policy Note, October 2015.