analysys mason

Analysys Mason Report

Data-driven innovation in Japan: supporting economic transformation

31 October 2014

David Abecassis, Nico Flores, Sara Montakhab

Ref: 2001132-417

www.analysysmason.com

Contents

Analysys Mason Limited
Bush House, North West Wing
Aldwych
London WC2B 4PJ
UK
Tel: +44 (0)20 7395 9000
Fax: +44 (0)20 7395 9001
london@analysysmason.com
www.analysysmason.com
Registered in England No. 5177472

DISCLAIMER

This document has been written on behalf of Google Inc. by Analysys Mason, a global consultancy specialising in telecoms, media and technology.

The analysis contained in this document is the sole responsibility of Analysys Mason and does not necessarily reflect the views of Google.

1 Executive summary

1.1 Introduction

Japan is engaged in its most ambitious economic recovery programme for a generation. Trying to break away from the recent cycle of deflation and loss of confidence, the government is focusing on 'three arrows':[1] (i) aggressive monetary policy; (ii) large fiscal stimulus; and (iii) profound transformation of the economy and, to a large extent, society.

A key theme of the 'third arrow' is the use of technology to deliver improved productivity, a world-class business environment and dynamic domestic markets in goods and services that can support global expansion. In its *Declaration to be the World's Most Advanced IT Nation* from 24 June 2014, the Japanese government's IT HQ highlighted the central role of data in this transformation.

In this context, the government is currently reviewing the legal and regulatory framework that surrounds the use of personal data.[2] This is informed by two key considerations. First, there is an awareness that the widespread use of data is an essential enabler of the growth that IT can help produce. Second, at the same time, consumers and citizens are becoming aware, and at times wary, of the fact that companies are using data, some of which they themselves contribute. As a result, policymakers are trying to find the right balance between the undeniable value of data-driven innovation and the concerns of their constituents.

This report provides a neutral perspective on considerations that we view as central to an informed debate about these issues. Specifically, we explore two broad questions:

- What does data-driven innovation bring to businesses? How much might it be worth to the economy as a whole? What are the enablers for data-driven innovation to realise its potential?
- How do policies and regulations affect these enablers, either positively (encouraging the development of services and a know-how to make the most of them) or negatively (putting barriers in place that may reduce the potential of data-driven innovation)? How do they mitigate the risks inherent to the use of data, such as misuse or unauthorised leakage of personal data?

In order to explore these questions, this report starts by setting out an analytical framework for analysing how different types of innovative services driven by data contribute value to the economy. We then present quantitative estimates of the current and potential future value of data-driven innovation in the Japanese economy, and then discuss some of the key policy challenges posed by data-driven innovation.

1 http://lexicon.ft.com/Term?term=abenomics.

2 See IT Strategic Headquarters, *Outline of the system reform concerning the utilization of personal data*, 24 June 2014 (パーソナルデータの利活用に関する制度改正大綱) available at http://www.kantei.go.jp/jp/singi/it2/info/h260625_siryou2.pdf.

The report reflects work conducted by Analysys Mason between July and September 2014. It draws on Analysys Mason's experience,[3] evidence from a telephone interview programme of 100 enterprises in Japan, interviews conducted with innovative providers of data-driven services, and extensive secondary research including case studies mentioned throughout the report.

1.2 Introducing data-driven innovation

In this report, we use the phrase '**DDI service**' (or 'innovative data-driven service', or simply 'service') to refer to an innovative use of data, together with the technical and commercial arrangements that enable economic or social value to be created from this data.

While DDI services often involve a DDI-centric firm performing a service for a customer (which may in turn be another firm or a consumer), they may also be contained within a single firm (e.g. a supermarket chain can analyse its customers' data itself, without recourse to a third party). Similarly, DDI services may themselves directly provide value to end users (for example, a GPS-enabled route-finding application), or they may be only a link in the chain of added value that goes into a product or service (for example, when data is used to enhance industrial processes).

1.2.1 The activities involved in DDI services

In general, DDI services can be thought of as involving some or all of the following activities:

- **Data collection**: Collected data may be explicitly submitted by users or collected in the background (for example, by monitoring users' location). Collected data may not relate to individuals, and hence may not be personal data (e.g. environmental sensors).
- **Data storage**: Stored data may be personal data, pseudonymous (i.e. relating to a single individual, but with the link to the individual removed), aggregate (i.e. relating to groups of people) or entirely non-personal information (e.g. stock prices).
- **Combination and repurposing** of datasets, which may have been originally or primarily obtained for one purpose (e.g. users registering a product for warranty purposes) but may be used for another (e.g. to identify likely buyers of a new product).
- **Analysis** of data, to deliver the service that creates value for end users – for example, the segmentation of a group of consumers.
- **End use**, in which a DDI service is used and generates its value. The beneficiaries of this value may be the persons whose data is being used, other customers of the service, wider society or a combination of these.

[3] Notably a similar study conducted in Singapore in 2013, and published under the title *Data-driven innovation in Singapore*, available at: http://www.analysysmason.com/Research/Content/Reports/Data-driven-innovation-in-Singapore/.

These activities may all be performed by a single DDI service provider, or they may involve co-ordination across multiple firms. When multiple firms are involved, a **transfer of data** typically occurs between these firms (domestically or internationally), and this can take place at any point in the sequence above.

DDI service providers may offer their services directly to consumers or to firms in sectors throughout the economy, such as manufacturing or transport.

1.2.2 Five types of DDI service

The landscape of innovative data-driven services is extremely varied, and fast-changing. In order to reflect this diversity, our analysis focuses on broad categories of services, each based around a different form of value proposition, rather than a series of individual services. We have been able to classify all the examples of DDI services that we have encountered as part of this study[4] into the following five broad categories:

- **The new bazaar: personalised services, at personalised prices.** Just as before industrialisation each product or service was a one-off and was sold at a one-off prices, in recent years technology has enabled a shift towards personalised products and pricing, driven by digital services, on-demand manufacturing and dynamic pricing. Box 3-1 on p.20 discusses Sompo Japan's DoRaRoGu and the partnership between Towers Watson and Vodafone.

- **Services for here and now: real-time awareness of time and space data.** Services of this type aim to help users carry out their 'real-world' tasks in real time by combining location data with other information that may be pertinent at the time. Box 3-2 on p.22 discusses the case of Colopl and KDDI in more detail.

- **Attentive services: using patterns and real-time data to identify or predict behaviours**. These services seek to understand their users' behaviour and circumstances in order to anticipate their needs and proactively adapt the service provided. To do this, they may monitor users in real time and predict behaviour using probabilistic models. Box 3-3 on p.23 discusses the cases of DeNA Life Science and Komatsu's KOMTRAX vehicle management system.

- **Services for people like me: characteristics and correlations to improve business performance**. These services target groups of consumers with common characteristics, relying on algorithms to identify consumer segments. This helps to improve the fit between customers and the products that are marketed to them, which in turn can lead to increased customer satisfaction. Box 3-4 on p.25 discusses the case of Yahoo! Japan.

---

4 Although our taxonomy is not exhaustive, it was arrived at in a 'bottom-up' fashion by considering numerous examples of DDI services; over the course of this project we have not found an example that cannot be categorised. Similarly, although the categories are not strictly mutually exclusive, in the economic analysis behind Section 4 we have relied on a more technical characterisation of the categories in our taxonomy, which has allowed us to minimise any overlaps.

- **Intelligent planning: large-scale pattern analysis to improve complex systems.** Finally, 'intelligent planning' services aim to extract patterns from the analysis of large, often disparate and/or anonymous datasets to help firms or governments prepare and respond to large-scale evolutions or trends. Box 3-5 on p.27 discusses the cases of Fujitsu's Akisai cloud service for farms and food-related industries, and Metawater.

1.3 Economic value of data-driven innovation in Japan

As noted above, the direct customer of a DDI service can be a consumer or another firm. Of these two scenarios, the second one (that is, DDI services being offered to other firms) accounts for the majority of the value created.

In this study, we consider the use and value of DDI services in the following traditional sectors or verticals: manufacturing; trade (wholesale and retail); transport and logistics; financial services; information and communications; and health, education and social services. These verticals are essential to the economy of Japan, both domestically (e.g. retail trade, healthcare and education) and as a major global economic power. Together, we estimate that they represent about 71% of the gross value added (GVA)[5] estimated for 2013 in Japan, as well as 70% of total employment.

Below, we explore the value that DDI contributes to the economy via these verticals first from the perspective of economic value (which we have estimated quantitatively), and in terms of broader notions of public and social value (which we have not).

1.3.1 The economic value of data-driven innovation in Japan

We have approached the question of economic value primarily from the perspective of producers (that is, firms within the verticals listed above), with consumers' benefits flowing from producers' benefits through lower prices.

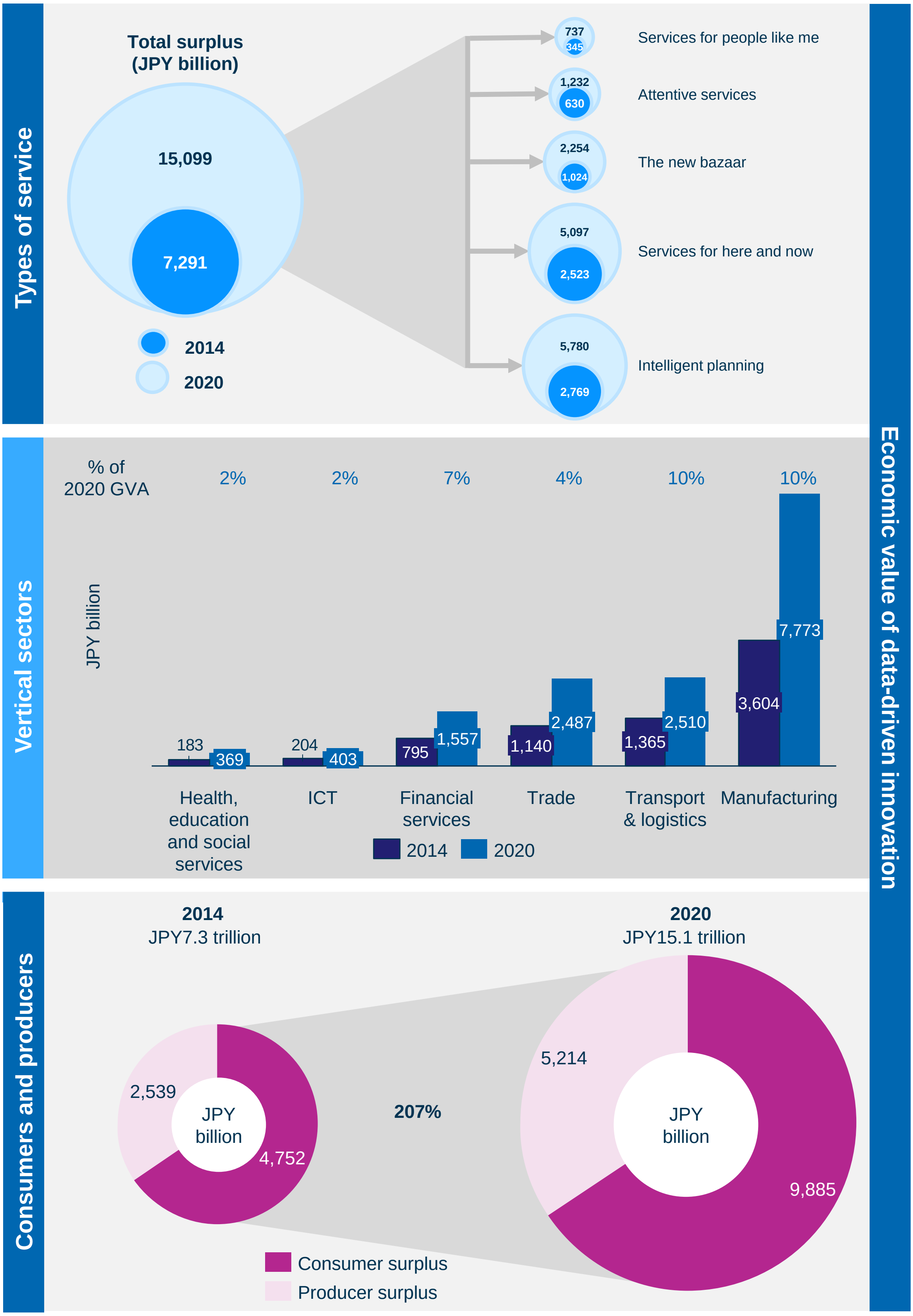
Specifically, we started by estimating the value of cost savings and incremental revenues due to data-driven innovation that firms in each vertical under study have experienced. Once we estimated this amount, we have attempted to estimate what proportion of it is:

- passed by firms on to consumers as lower prices: this leads to an estimation of DDI's impact on 'consumer surplus'; or
- retained by producers as higher profits: this leads to an estimate of DDI's impact on 'producer surplus'.

We have estimated the cost savings and incremental revenues by relying on a wide range of case studies in the public domain, interviews with Japanese firms, and our own analysis. The infographic below shows our key estimates for different verticals and service types.

---

5 Gross value added (GVA) is a measure of GDP that considers output (revenues) minus the costs of intermediate goods and services; in this paper, we have measured the value of DDI services based on the incremental net revenue and the costs savings that use of DDI services brings to each vertical, which is not identical to GVA but provides a useful measure of economic value based on consumer and producer surplus.


Types of service
Total surplus (JPY billion)
15,099
7,291
2014
2020
737
345
Services for people like me
1,232
630
Attentive services
2,254
1,024
The new bazaar
5,097
2,523
Services for here and now
5,780
2,769
Intelligent planning
Vertical sectors
% of 2020 GVA
2%
2%
7%
4%
10%
10%
JPY billion
183
369
204
403
795
1,557
1,140
2,487
1,365
2,510
3,604
7,773
Health, education and social services
ICT
Financial services
Trade
Transport & logistics
Manufacturing
2014
2020
Consumers and producers
2014
JPY7.3 trillion
2020
JPY15.1 trillion
2,539
JPY billion
4,752
207%
5,214
JPY billion
9,885
Consumer surplus
Producer surplus
Economic value of data-driven innovation

Section 4.1.2 in the main body of this report briefly discusses our reasoning for a variety of verticals and DDI service types. Combining consumer and producer surplus across all verticals under study, we estimate that data-driven innovation and associated services already contribute over JPY7 trillion to Japan's economy in 2014 (equivalent to about 3.4% of Japan's overall GVA),[6] with the majority due to consumer surplus. We also estimate that by 2020 the value of data-driven innovation in Japan can exceed JPY15 trillion.

1.3.2 The wider social value of data-driven innovation

Beyond the economic value discussed so far, data-driven innovation has the potential to unlock social and public benefits that are not easily quantifiable. Key examples include the following:

- **Healthcare and education:** One of the findings from our economic analysis is that the economic value of data-driven innovation for health and education is smaller than in other verticals. However, this underplays the potential significance of the use of data-driven innovation in an educational or healthcare context, where social considerations are paramount. These social benefits include, for example, improved educational and health outcomes; more efficient use of public resources (which may not translate into value-added gains) and alleviating labour shortages in elderly care.

- **Better evidence-based government:** Through the publication of more internal governmental information, citizens are able to analyse the workings of governments, thus increasing transparency. This in turn allows more informed debate among the public and can enhance governments' accountability to their citizens. The popularity of this approach internationally can be seen from the Open Government Partnership, which currently contains 62 countries. These include the USA (where open government has been a key objective of the Obama administration), as well as Indonesia, Australia and New Zealand.[7]

Our research and modelling confirm that Japan has the right technological and economic foundations to benefit greatly from data-driven innovation. Traditionally strong sectors such as manufacturing and transports appear best placed currently and, in the future, increased adoption of DDI services should enable more gains.

6 The average of our overall range is between JPY4,614 billion and JPY10,081 billion; all figures in this section are central estimates based on a range of values.

7 See http://www.opengovpartnership.org/.

1.4 Policy considerations related to data-driven innovation

Motivated by the significant potential for economic value creation of data-driven innovation, policymakers around the world have been taking steps to ensure that data-driven innovation can reach its full potential. As we will see, Japan is no exception to this.

1.4.1 The impact of data-related policies on data-driven innovation

Policy action related to data can either enable or potentially constrain the development of data-driven innovation, so that we can thus speak in terms of 'enabling' and 'constraining' policies. These are listed in Figure 1.1 and briefly discussed below.

Figure 1.1: Selected enabling and constraining policies [Source: Analysys Mason, 2014]

Enabling policies	Constraining policies
Positive policies: • Direct funding • Ecosystem development • Open data • Skills development • Technical standards	**Activity-specific:** • Collection • Storage • Combination/repurposing • Analysis • Use
Removal of barriers: • Regulatory clarity • Trust-enabling data protection • International regulatory interoperability	**Non activity-specific:** • Inter-company transfers • International transfers • Security

This classification is by no means intended to suggest that only 'enabling' policies are desirable while 'constraining' ones are always undesirable. As we will see, a key challenge for policymakers is to reconcile the objectives behind both types of policy, so that (for example) innovation can be promoted while also ensuring that privacy is protected.

Enabling policies

Enabling policies include 'positive' policies that help data-driven innovation from a resource standpoint, including:

- *Direct funding* for innovation, for example through seed capital or competitions
- *Ecosystem development* policies aimed at developing a critical mass of firms
- *Open data* initiatives, whereby public datasets are made accessible using standard formats
- *Skills development* to ensure that there is a sufficiently large talent pool of data experts
- *Technical standards* for data exchange to facilitate collaboration and trading.

In addition, the removal of regulatory barriers to data-driven innovation is also an enabling type of policy. Key examples include:

- *Providing a clear and effective regulatory climate*. As with any other type of business, regulatory uncertainty acts as a barrier to investment in data-driven innovation
- *Ensuring that essential data protection safeguards are in place* to allow consumers to trust that their data will be handled adequately
- *Enabling international regulatory interoperability*, so that DDI service providers can work with overseas partners or subsidiaries for delivering their services. This can take the form of multilateral certification schemes such as APEC's Cross Border Privacy Rules system[8] (CBPR), the European Union's Binding Corporate Rules (BCRs) scheme or its 'safe harbour' arrangements.
- *High-level international guidelines*, such as those issued by the OECD,[9] so that at least a minimum of interoperability exists.

Constraining policies

Potentially constraining rules and policies exist in order to protect important public interests, such as privacy, crime prevention, or national security, none of which can be neglected. As such, constraining policies cannot be assumed to be 'negative' or undesirable in overall terms, even if their impact on data-driven innovation is not a positive one.

Although there are important exceptions, in general these rules revolve around the protection of personal data[10] – that is, they are *personal data protection* rules. Many such rules apply specifically to one of the data processing activities involved in DDI which we discussed in Section 1.2.1. This includes rules concerning:

- data collection, which may limit firms' ability to collect data
- data storage, which may require firms to delete data after a certain period
- combination and repurposing of datasets, which may prevent the use of datasets collected for one purpose from being used for another, typically in combination with other datasets
- analysis, which may by itself render data subject to certain restrictions if, for instance, it results in the re-identification of anonymous or pseudonymous data[11]
- end use, such as for example the use of profiling techniques to make 'automatic' decisions affecting consumers.

8 See http://www.cbprs.org/.

9 See http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm.

10 Specifically, personally identifiable information, or data that can be linked to an individual.

11 Pseudonymous data is data that relates, but is not by itself linkable, to a single, real individual; a key example "user ID" numbers used by websites to detect repeat visitors even if the user's identity is unknown. Anonymised data is not only not linkable to an individual but may relate to one or to multiple individuals.

Additionally, other rules apply across DDI to multiple data processing activities, including

- rules limiting the transfer of data between different firms and/or countries
- rules concerning security arrangements that should apply to data.

The restrictions outlined above can often be overridden through consumer consent. For example, if a service provider needs to store data for a longer period than would normally be allowed, then once it obtains user consent the relevant restrictions may no longer apply.[12] Thus, data protection regulations often represent only the 'default setting' under which service providers must operate *unless* they obtain consent from users to operate in a different way. A strict consent regime may place strong limitations in terms of what services can do without user consent, thereby requiring consent – which may need to be explicit and/or detailed – in order to carry out their activities.

The impact of data protection on DDI services

Data protection policies can negatively affect DDI services in two ways. First, to the extent that different types of service rely on different types of data processing, they are differently affected by rules affecting those activities. Thus, for example, 'services for people like me' often apply segmentation techniques to large datasets and automatically classify users (and other consumers) into groups, using information from multiple sources, some of which may be anonymous or pseudonymous. Accordingly, these services would be particularly affected by regulations limiting the repurposing and combination of data. Figure 5.3 on p.61 of this report offers a systematic analysis of the main types of policy affecting each type of DDI service.

Second, and more importantly, certain types of restriction can also have a more indirect, but potentially more profound, negative impact on data-driven innovation by constraining the *processes* of innovation and entrepreneurship through which new services can develop and flourish. Specifically,

- *Strict purpose limitation requirements may hinder innovation*. Finding hidden patterns in data, and extracting value from these, is a key aspect of innovation with data. This process often involves the combination of datasets collected for different purposes – which, because of the nature of the process of experimentation involved in service development, may very well be unrelated to that of the new service being developed.

- *Limits on international transfers may hinder competitiveness*. DDI services are often built by combining technical functionality contributed by different firms, which may often be in different countries – e.g. an IP geo-location service[13] might be provided by a US firm, while an analytics service might be provided by a Japanese-based one. Preventing businesses from relying on best-in-class partners in this way may lead to domestic service providers being globally uncompetitive in terms of service quality, price or both.

[12] However, in some jurisdictions certain requirements e.g. concerning sensitive personal information cannot be waived through consent.

[13] Business-to-business services that allow other services to determine a user's geographical location on the basis of an IP address.

- *Restrictive consent requirements may hinder DDI service growth.* If a service relies on activities (e.g. data retention, inter-firm transfers, etc.) for which regulations require that users give explicit consent, this can act as a significant hurdle against service use. To a significant extent this would likely be due to users who may lack the time or inclination to engage with a privacy policy or statement, but who might otherwise be willing to give consent. In these cases, consent requirements result in a deadweight loss and in an unnecessary impediment to DDI service use and growth.

1.4.2 The way forward: reconciliation and trade-offs

In this context, the role of policy-makers is two-fold:

- they must achieve a balance between competing interests in a way that is as accommodating of these interests as possible, which involves adopting policies that achieve the desired outcomes of both data protection and data-driven innovation where possible; and

- they must sometimes accept trade-offs where reconciling these outcomes is not possible, in which case the impact of these trade-offs should be well understood and carefully considered.

Reconciling DDI development and data protection by focusing on outcomes

Although data protection rules may clash with data-driven innovation in specific cases, the broad policy *aims* of fostering data-driven innovation and protecting other public interests (e.g. privacy) are *not* fundamentally at odds with each other; they are simply about different things. This suggests that, at least in some cases, it may be possible to achieve the desired aims of data protection without unduly hindering data-driven innovation.

Data protection policies seek to prevent certain *harms* to individuals or society, or at least to minimise the *risk* that such harms may materialise. For example, the handling of information such as credit card numbers is protected so as to prevent fraud. More generally, the rationales for rules such as those outlined earlier can be linked to consumer harms such as fraud, illicit discrimination or the invasion of privacy. On this basis, a key question when considering a given type of data protection rule is whether alternative rules can be found that address the same risks to a similar or higher extent but without conflicting with data-driven innovation.

We believe this is often possible. If we consider the harms that policies aim to prevent (i.e. the *outcomes* sought), and then consider whether the associated risks can be addressed through different sets of policies that are compatible with data-driven innovation, it is often the case that the risks at stake can be addressed in ways that do not conflict with data-driven innovation. We consider three examples:

- *Using robust authentication as an alternative to limits on repurposing pseudonymous data.* Traditionally, a common way to avoid some of the obstacles posed by purpose limitation requirements has been to rely on anonymised or pseudonymised data that cannot be traced back to the original data subject. However, in recent years, concerns about the possibility that anonymous or pseudonymous data may be 'de-anonymised' through analytics has led some players to call for pseudonymous data to be subject to purpose limitation restrictions,[14] which might limit much of the combination and repurposing of datasets that is essential to DDI experimentation. However, concerns about re-identification are (to an important extent) related to the risk of fraud and similar abuses, as responses to a recent survey from the Japanese government (MIC) suggest.[15] But *to the extent* that the objective is to prevent such harms, an alternative to imposing consent requirements on pseudonymous data might be to develop policies aimed at ensuring stronger authentication methods, which might render the possession of most types of personal data insufficient to commit fraud. This would not only avoid conflicts with data-driven innovation, but might also be more effective at preventing fraud itself.

- *Using bilateral or multilateral agreements instead of restrictions on international transfers.* As we have seen, limits on international transfers (e.g. rules on data localisation) can severely constrain some DDI business models. Depending on the country in question, the rationale for these policies may include avoiding consumer harm from inadequate or unenforceable legal protection overseas, national security, and (in some cases) industrial policy. Of these, the first is arguably the main and most common. But in respect to this, solutions that are targeted specifically at ensuring that data is adequately protected overseas (for example, this is the first principle in the APEC's Privacy Framework) may be more appropriate and less costly in terms of impact on data-driven innovation than direct limitations or bans on international data transfers.
- *Alternatives to strict consent requirements*. As we have seen, onerous consent requirements can be detrimental to the growth of new DDI services. On the surface, consent requirements give consumers a means to avoid certain harms ensuring that potentially risky activities do not take place without their consent. But as we have seen, if consumers are expected to engage in detail with technical matters for which they lack the time or understanding,[16] this can lead to uninformed consent ('choice fatigue',[17] which defeats the purpose of the policy) or to an uninformed refusal to give consent (which may be a loss to everyone).[18] Thus, explicit and detailed consent requirements may not be a particularly effective tool to prevent privacy-related harms, and it may also be self-defeating. Alternative approaches include placing greater emphasis on substantive restrictions on data use and/or processing (for example, if it is

14 See for example http://history.edri.org/eudatap-issuesheets#defi.

15 A recent survey by the Japanese government found that when respondents were asked to state how sensitive different types of personal information were to them, by far the two highest categories were credit card numbers and account information (with over 70% of respondents assigning the highest level of sensitivity to this type of data). Clearly, the link between personal data and fraud is paramount in Japanese consumers' concerns about data. See page 16 in *Overview of the 2014 Edition of Information and Communications White Paper*, issued by the Ministry of Internal Affairs and Communications (MIC) (平成２６年版 情報通信白書」の概要 -総 務 省).

16 Although there are efforts to mandate simplicity in privacy statements, there is an inherent limit to the extent to which this can be done meaningfully.

17 See, for example, Augenblick, Ned and Nicholson, Scott, *Choice Fatigue: The Effect of Making Previous Choices on Decision Making*, available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.175.2560&rep=rep1&type=pdf.

18 Additionally, there is growing body of research suggesting that too much choice may be neither beneficial to society nor appreciated by consumers. See, for example, Schwartz, B, *The paradox of choice* (2005), and Botti, S. and Iyengar, S. S., "The Dark Side of Choice: When Choice Impairs Social Welfare", *Journal of public policy and marketing* Vol 25(1), 2006.

agreed that pseudonymous data can be combined, repurposed and shared freely between firms that meet certain criteria, then DDI service providers would not need users' permission for this); and lowering the explicitness requirements for consent (for example, by allowing 'deemed' or 'opt out' consent to apply in more contexts.)

Facing trade-offs between innovation, privacy and other policy goals

We have shown how, in some cases, it may be possible to address certain key objectives of data protection through alternative routes that may have a relatively small impact on data-driven innovation. However, it is important to acknowledge the limitations of this approach.

In each of the examples we gave above, it is also possible that other considerations not discussed by us might argue for the permanence of the restrictions in question. For example:

- Concerns related to re-identified pseudonymous data may sometimes be not about the risk of fraud but rather about the potential for embarrassing facts to become known (i.e. a 'pure privacy harm').
- Although explicit and detailed consent requirements may not always be in users' best interests, they may still sometimes be demanded by users. We cannot rule out the possibility that public opinion may veer towards explicit and detailed consent even if (in the light of our analysis) this is not necessarily to anyone's advantage.

In cases like these, it may be that an impact on data-driven innovation cannot be avoided, and that an adequate balance between the interests of DDI development and other public interests (e.g. privacy or national security) must be struck. This balance will likely depend on local sensitivities and public opinion. In any case, we believe that an approach grounded on the specifics of each situation should prevail; the potential downsides or harms (such as impeding DDI development or limiting privacy) should be assessed realistically in terms of the risk of them materialising, and this in turn may depend on the specific types of data, data processing or firms involved. This calls for a flexible approach that avoids all-or-nothing solutions where possible.

A key example of such an approach is the 'legitimate interests' doctrine employed in EU data protection, which allows firms to process data when consent is not practical and the processing "is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject which require protection […]".[19] Firms relying on legitimate interests are required to conduct a structured 'balancing test' between their own and consumers' interests, thereby allowing a flexible, nuanced approach (Box 5-2 on p.71 of the main body of this report provides an overview of the 'legitimate interests' doctrine). We note, however, that recent discussions about updating the 'legitimate interest' provisions have given rise to passionate controversies. To us, this reinforces our view that the trade-offs involved in DDI policy cut across important public interests, and that discussion of these issues calls for an evidence-based and democratic debate.

---

19 See Art 7(1)(f) in EU Directive 95/46/EC, available at http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:31995L0046.

1.5 Conclusions: ways forward for policy

Our research and modelling confirm that Japan has the right technological and economic foundations to benefit greatly from data-driven innovation. Traditionally strong sectors such as manufacturing and transports appear best placed currently and, in the future, increased adoption should enable more gains. Other sectors such as trade and healthcare can also benefit from data-driven innovation in different ways.

Data-driven innovation can create significant value. In the case of Japan, our central estimate suggests that it already contributes JPY7 trillion (equivalent to around 3.4% of GVA) in the sectors we have considered. By 2020, we estimate the economic potential of data-driven innovation to be in excess of JPY15 trillion, provided the right environment is in place.

A balanced approach to regulating and managing data can help make Japan a leader in data-driven innovation. Proactive, positive policies such as open data are important in encouraging DDI development. They are already being pursued by governments and such efforts should continue. However, DDI development can clash with other policies.

To address these challenges, where possible policymakers should seek to minimise the tensions between the interests of data protection and those of DDI development. To do so, they can focus on the key outcomes that are sought under each of these two (and other) interests and consider solutions that satisfy both requirements.

Sometimes, simple policy solutions may not be available. In those cases, policymakers need to consider what fundamental trade-offs might be appropriate, based on the available evidence and informed by local sensitivities. The choices for policymakers may be difficult, but they need not be made in the dark.

Finally, it is important to keep these concerns in perspective. It is clear that data-driven innovation is a positive development which in general has been able to grow without fundamentally falling foul of existing policies. As with any new technology and its associated economic activity, friction is bound to be created as data-driven innovation continues to develop, but nothing suggests that this cannot be addressed successfully.

2 Introduction

2.1 The role for innovation and information technology in Japan's economic future

Japan is engaged in its most ambitious economic recovery programme for a generation. Trying to break away from the recent cycle of deflation and loss of confidence, the government is focusing on 'three arrows':[20] (i) aggressive monetary policy, (ii) large fiscal stimulus, and (iii) profound transformation of the economy and, to a large extent, society.

As part of the 'third arrow', the Japanese government is promoting a wide range of reforms across many aspects of the economy, from agriculture to healthcare, and from nursing to the structure of labour markets.[21] Throughout these reforms, exploring and exploiting technology is a recurring theme to achieve many of the government's ambitions: improved productivity, a world-class business environment, dynamic domestic markets in goods and services that can support global expansion.

In a rapidly-aging society, where population is contracting, technology will also play a major role in supporting an unprecedented demographic transition towards an increasingly elderly population. Robots can be used in nursing, and technology can support greater efficiency in healthcare, for example.

There is a broad recognition, at the highest levels of government, of the central role of technology and innovation in Japan's economic future. This is enshrined in the IT Headquarters' *Declaration to be the World's Most Advanced IT Nation*, from 24 June 2014. A core element of that declaration is the importance of data. Considered broadly, data of all sorts is a fundamental 'raw material' that enables corporations and individuals to innovate, inventing new uses and becoming more efficient.

2.2 Data-driven innovation: an old idea, accelerated by technology

Data-driven innovation (or 'DDI') is now all around us. Mobile phones and other connected portable devices – among other Internet-related platforms – create vast amounts of data on a daily basis. Innovative techniques, services and business models seek to harness this resource (data) to unlock significant potential across many sectors of the economy.

In the eyes of the public, data-driven innovation is often associated with 'big data': huge volumes of data, collected across many devices and environments, and processed rapidly in large data centres by powerful computers running sophisticated 'data analytics' algorithms. Indeed, in the Japanese context,

[20] http://lexicon.ft.com/Term?term=abenomics

[21] See *Japan Revitalization Strategy*, Kantei, 24 June 2014.

the government's declaration refers to "measures to encourage the use of big data",[22] in order to promote "the development of an environment conducive to data-driven innovation".[23]

Consider for example an everyday practice that people have engaged in for centuries: applying for a loan. This involves a financial institution making a decision on whether to lend money to the applicant, at what interest rate, and for how long. This decision is clearly 'data-driven': it relies on information about the borrower, the use that will be made of the funds and the economic environment. The data comes from various sources: some of it is data that is independent from the borrower but is related to the broader environment, such as interest rates; some data is provided by the borrower; and some of the data may be provided, or validated, by an 'information broker' who can obtain data about the borrower (typically with the borrower's consent, implied or explicit). At the end of the process a decision is reached; if the loan is granted, both sides are better off because the borrower receives a (financial) service that he desires, and the financial institution can charge interest commensurate to the risk it takes.

The basic idea that data can be used in the economy to create value is therefore an old one. Communications services are essential to this data being available and communicated efficiently and effectively, as evidenced for example by the early use of telegraph machines in order to transmit stock quotes. Since the creation of the Internet and the widespread adoption of data-capable personal devices, the opportunities presented by data-driven technologies have grown immeasurably. Given the rapid growth in the ability to process enormous volumes of data at high speed, it is not difficult to see that this trend has the potential to significantly affect the global economy.

2.3 Background and objectives of this paper

In this context, the government of Japan is currently reviewing the legal and regulatory framework that surrounds the use of data, in particular 'personal data'.[24] Although the very definition of personal data is the subject of discussions and some debate, it generally refers to information that can be used to identify a specific individual.

Personal information is an important resource for data-driven innovation, but by no means the only one; in particular, data related to goods and machines is becoming increasingly valuable across the economy. With the advent of machine-to-machine communications, the wealth of data unrelated to individuals will grow rapidly, and can be harnessed by businesses to become more efficient and compete ever more effectively in the global economy.

22 「ビッグデータ」の利活用を促進するための取り組み, *Declaration to be the World`s Most Advanced IT Nation,* June 2014, Page 9.

23 データドリブンイノベーションが創出される環境の整備, *ibid.*

24 See for example IT Strategic Headquarters, *Outline of the system reform concerning the utilization of personal data*, 24 June 2014 (パーソナルデータの利活用に関する制度改正大綱) available at http://www.kantei.go.jp/jp/singi/it2/info/h260625_siryou2.pdf.

At the same time, consumers and citizens are becoming aware, and at times wary, of the fact that companies are using data, some of which they themselves contribute. As a result, policymakers are trying to find the right balance between the undeniable value of data-driven innovation and the concerns of their constituents. This tension is clearly visible in the ongoing public debates that seem to oppose proponents of privacy and data protection with companies that stand to benefit from data-driven innovation.

This report provides a neutral perspective on considerations that we view as central to an informed debate about these issues. Specifically, we explore two broad questions:

- What does data-driven innovation bring to businesses? How much might it be worth to the economy as a whole? What are the enablers for data-driven innovation to realise its potential?

- How do policies and regulations affect these enablers, either positively (encouraging the development of services and a know-how to make the most of them) or negatively (putting barriers in place that may reduce the potential of data-driven innovation)? How do they mitigate the risks inherent to the use of data, such as misuse or unauthorised leakage of personal data?

In order to explore these questions, this report starts by setting out an analytical framework for analysing how different types of innovative services driven by data contribute value to the economy (Section 3), before presenting our findings on the current and potential future value of data-driven innovation in the Japanese economy (Section 4).

It is against this background that policymakers can then assess the potential impact of various approaches to data protection. In Section 5, we discuss how different paradigms can contribute to the same objectives, for example deterring misuse of personal data, but often have distinct impacts on the drivers of data-driven innovation and therefore the value it can bring to the economy.

The report reflects work conducted by Analysys Mason between July and September 2014. It draws on Analysys Mason's experience,[25] evidence from a telephone interview programme of 100 enterprises in Japan, interviews conducted with innovative providers of data-driven services, and extensive secondary research including case studies mentioned throughout the report.

---

25 Notably a similar study conducted in Singapore in 2013, and published under the title *Data-driven innovation in Singapore*, available at: http://www.analysysmason.com/Research/Content/Reports/Data-driven-innovation-in-Singapore/.

3 Data-driven innovation and its role in the economy

3.1 Definitions and scope of this study

A main focus of this report is to develop an understanding of the potential for the creation of economic and social value through data. This value can be unlocked by designing **innovative processes** to collect, combine and analyse data, in large part thanks to technology. This requires a broad view of how electronic data, of many types and at different scales, contributes to economic and social activities in developed economies in general and in Japan in particular.[26]

In order to focus our study, the data that we are concerned with is **structured information** about the physical or social worlds, which machines are (or will be) able to organise and combine. Some forms of electronic data are excluded from this definition, for example digital 'content' such as music or films.[27]

In this report, we use the phrase '**DDI service**' (or 'innovative data-driven service', or simply 'service') to refer to an innovative use of data, together with the technical and commercial arrangements that enable economic or social value to be created from this data. While DDI services often involve a DDI-centric firm performing a service for a customer (which may in turn be another firm or a consumer), they may also be contained within a single firm (e.g. a supermarket chain can analyse its customers' data itself, without recourse to a third party). Similarly, DDI services may themselves directly provide value to end users (for example, a GPS-enabled route-finding application), or they may be only a link in the chain of added value that goes into a product or service (for example, when data is used to enhance industrial processes).

The remainder of this section characterises DDI services in terms of:

- a description of the activities involved in providing DDI services
- a classification (or 'taxonomy') of some of the main types of value-creating service
- a simplified model of the way in which DDI services affect the broader economy.

3.2 Stages of data-driven innovation

In general, DDI services can be thought of as involving some or all of the following activities:

- **Data collection**: this may rely on information explicitly submitted by users (e.g. date of birth, address) or collected in the background (for example, by monitoring users' location or online behaviour); the user may not always be aware of data collected in the background, depending on the level and prominence of disclosure. Some data collection does not involve any users per se, but rather machines or sensors in the environment. Data is a broad concept that includes 'personal data' and data that is not personal.

26 For a useful discussion of "Big Data", see Mayer-Schönberger, Viktor and Cukier, Kenneth, *Big Data: A Revolution that Will Transform how We Live, Work and Think* (2012).

27 However, the meta-data that often accompanies digital content satisfies our definition.

- **Data storage**: this may involve storing data for a period of time. The data can be of many types: personal data, pseudonymous data (i.e. relating to a single individual, but with the link to the individual removed),[28] aggregate data (i.e. relating to groups of people) or entirely non-personal information (e.g. stock prices, or weather information).
- **Combination and repurposing** of datasets, which may have been originally or primarily obtained for one purpose (e.g. users registering a product for warranty purposes) but may be used for another (e.g. to identify likely buyers of a new product).
- **Analysis** of data, to deliver the service that creates value for end users – for example, the segmentation of a group of consumers.
- **End use**, in which a DDI service is used and generates its value. The beneficiaries of this value may be the persons whose data is being used (e.g. when a navigation app tells a user what the best way to get home is at a given point in time), other customers of the DDI service (e.g. when data about one user's location alerts other drivers that there is a traffic jam in their vicinity), wider society (e.g. when using traffic data to improve road planning) or a combination of these.

These activities come together as part of a 'DDI value chain' of interconnected activities, which is illustrated in Figure 3.1 below:

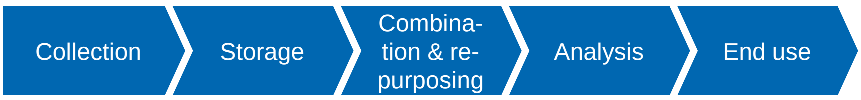


Figure 3.1: DDI value chain [Source: Analysys Mason, 2014]

These activities may all be performed by a single DDI service provider, or they may involve co-ordination across multiple firms. When multiple firms are involved, a **transfer of data** typically occurs between these firms (domestically or internationally), and this can take place at any point in the sequence above. Conversely, all these activities can be vertically integrated within the entity that eventually uses the service, such as a retailer that analyses its own customer data internally in order to improve its sales.

3.3 The five types of DDI services

The landscape of innovative data-driven services is extremely varied, and evolves faster than any study such as this one could ever hope to track. As a result, our analysis considers broad categories of services, each based around a different form of value proposition, rather than a series of individual services. This 'taxonomy' of DDI services was designed with two aims:

- The five types of services reflect the main ways through which data-driven innovation creates social and economic value (this is then developed in Section 4).
- Each service type is characterised by a dependency on different types of activities (e.g. collection or transfer), so that policies or regulations that affect a particular activity can be traced to its impact on a particular service type (this is discussed further in Section 5).

[28] Pseudonymous data is data that relates, but is not by itself linkable, to a single, real individual; a key example "user ID" numbers used by websites to detect repeat visitors even if the user's identity is unknown. Anonymised data is also not linkable to an individual, but may in addition relate to one or to multiple individuals.

We have been able to classify all the examples of DDI services that we have encountered as part of this study[29] into the following five broad categories, which we explore further below:

- the new bazaar
- services for here and now
- attentive services
- services for people like me
- intelligent planning.

3.3.1 The new bazaar: personalised services, at personalised prices

'The new bazaar' refers to services that allow firms to offer one-off, personalised products at one-off, personalised prices (although many online services are free at the point of use).

In the 19th and 20th centuries, the industrial revolution enabled the emergence of mass-produced consumer goods. This led to price-based competition which, together with the emergence of modern retail, led to what is now the universal practice of standardised products and services sold at uniform prices driven by cost and competition.[30]

In recent years, technology has enabled a shift towards more personalised products and pricing. Both competition and transaction costs remain valid drivers for standardisation in many markets for goods and services, but technology is now also enabling 'mass customisation' and the delivery of personalised services (e.g. nano-fabrication, on-demand book printing, and more recently on-demand 3D printing). In addition, and even for mass-produced goods and services (e.g. airline tickets), algorithmic tools mean that optimising the correct price for a given customer is no longer a costly affair. All of this involves extensive analysis of data about users, products, and about the relationship between the two.

Important examples of the types of service that are offered in the 'new bazaar' include:

- online bidding systems (e.g. Priceline.com)
- dynamically optimised pricing (e.g. as used by Austrian DIY retailer bauMax)[31]
- 'porous' paywalls in online newspapers that only charge customers who are likely to have high willingness to pay (a famous example is the New York Times's website, nytimes.com)
- insurance quotes that take into account applicants' behaviour or relationships (see Box 3-1 on the next page).

---

29 Although our taxonomy is not exhaustive, it was arrived at in a 'bottom-up' fashion by considering numerous examples of DDI services; over the course of this project we have not found an example that cannot be categorised. Similarly, although the categories are not strictly mutually exclusive, in the economic analysis behind Section 4 we have relied on a more technical characterisation of the categories in our taxonomy, which has allowed us to minimise any overlaps.

30 But notably, this change did not apply to every sphere; for example, large-ticket items like real estate or large B2B transactions have always been less standardised. For an interesting discussion of the introduction of uniform prices, see Boorstin, Daniel J., *The Americans: The Democratic Experience*, Vintage (1974).

31 Source: http://ro.uow.edu.au/cgi/viewcontent.cgi?article=1686&context=commpapers.

Box 3-1: DDI in practice – the new bazaar[32]

Sompo Japan's DoRaRoGu

Sompo Japan Insurance is a 'non-life' insurer, with a large vehicle insurance business. As part of its portfolio of car insurance products, Sompo Japan is selling a 'pay-as-you-drive' insurance product (DORAROGU), specifically marketed at drivers of Nissan's Leaf electric cars. The system uses telematics embedded into the car to gather data and track the history of driving distances through a telecommunications system especially developed for the Leaf.

The system gives feedback to the driver on speed and distances as well as on driving in a safe and environmentally friendly manner. Additionally, the firm also offers a theft tracking service that allows cars to be located based on GPS data and involves a 'tracking police' team that is deployed to help get the stolen cars back to their drivers. This system is intended to support a customised insurance product by combining this real-time, car-specific information, with traditional demographic data and other risk-assessment methods.

Depending on the driving distance, the insurance fee offered by Sompo Japan can vary by +–10%[33] compared to the other premium insurance fees in the market and helps Sompo have a better control on the cost estimates to be expected from each of its customers as a result of accident/theft claims. Conversely, bad drivers are unable to externalise their excessive claims on good drivers, and are incentivised to improve their driving.

Towers Watson/Vodafone partnership

In March 2013, Vodafone and Towers Watson announced a partnership to deliver an innovative solution to the automotive insurance industry. The solution uses Vodafone's machine-to-machine (M2M) modules and provides real-time data on driving activity for insurance firms that wish to provide a customised premium to participating consumers.

After completion of a trial period (during which the monitoring application is installed on the consumer's smartphone), an M2M module is installed in the insured vehicle. Data on location and speed of travel is combined with other information (such as local speed limits and weather conditions) to present an overall picture of the risk for the vehicle (and driver) in question. The data collected by Vodafone is passed to Towers Watson, which in turn provides a driving risk profile to insurance firms. The insurance firms can then develop a personalised premium for their client by combining this information with demographic data and other credit scoring methods.

As a result, insurers can produce highly individualised estimates for the costs that can be expected from each consumer as a result of accident claims (the 'loss ratio'), identifying more than a tenfold difference in costs between the 10% 'worst' drivers and the 10% 'best'.

---

32 For more information, see https://www.towerswatson.com/en-IE/Press/2013/03/Vodafone-and-Towers-Watson-launch-telematics-usage-based-insurance-service

33 See http://www.sompo-japan.co.jp/~/media/SJcms/news/2012/20121221_1.pdf

3.3.2 Services for here and now: real-time awareness of time and space data

Services of this type aim to help users carry out their 'real-world' tasks in real time by combining location data with other information that may be pertinent at the time. They rely on the ability of data-driven systems to take into account information in real time, analysing it and combining it with data stored in a more static way.

A familiar example is maps, which can be augmented based on a user's location to show a list of nearby shops, or the best route to a destination. Services can also take action on users' behalf, for example by paying for a road toll automatically, saving time and cutting pollution. Services of this type often rely on, or can benefit from, the use of public government data.

Examples include:

- mapping and apps (e.g. Google Maps)
- navigation/routeing apps (e.g. Navitime, Jorudan)
- augmented reality and shop finder apps (e.g. Colopl, Koozyt, Layar or Google Glass)
- fleet management systems (e.g. Hitachi Transport Management System (TMS)).

Box 3-2 on the next page discusses a specific example.

Box 3-2: DDI in practice – services for here and now

Colopl and KDDI's collaboration to provide user insights

Colopl, together with KDDI, is offering a data analysis service targeted at local governments for such purposes as local tourism development and disaster-response measures.

Colopl collects the location based-data from KDDI which in turn gathers data from the AU mobile users who gave their consent for third-party data processing. KDDI extracts only the data needed, anonymises it and sends it to Colopl, which is then able to analyse in detail the behaviour of the people who travel in certain areas.

The analysis comprises information on visitors (number of visitors, where they came from, prefecture, age, gender, how long they stayed in a particular area), their itinerary (number of visitors who did a one-day trip or stayed overnight), their stop-overs (number of visitors who did a stopover per city, village), the lodging facilities (number of visitors who used lodging facilities per city, village and type of lodging facilities), their transportation means (number of visitors per means of transportation), and number of inflow/outflow of visitors per each time period per each touristic area. The analysis also statistically calculates popular tour rankings by gender and age.

At the moment, the analysis does not appear to be happening in real time, but Colopl's product portfolio includes sophisticated "augmented reality" games that do provide real-time, context-specific information to their users, including marketing messages.

Their first customer was the local community of Hida City in Gifu prefecture. According to Mitsu Shimizu (the Head of the Tourism Division of the Hida City Planning Chamber of Commerce and Tourism), the report offered by Colopl is "extremely effective since it allows Hida City to create a more concrete tourism strategy by differentiating the large amount of visitors travelling within a large area".[34]

3.3.3 Attentive services: using patterns and real-time data to identify or predict behaviours

'Attentive services' seek to understand their users' behaviour and circumstances in order to anticipate their needs and proactively adapt the service provided. To do this, attentive services may monitor their users' behaviour in real or near-real time, and use the resulting data to predict future behaviour (often using probabilistic models). This may then cause the system to take action (for example, a thermostat controller might switch off central heating if it expects that its users will be away from home for an extended period) or trigger an alert which might be aimed at affecting future outcomes (e.g. health-monitoring service might provide feedback on lifestyle behaviours or alert a doctor about potential symptoms).

34 See http://colopl.co.jp/news/pressrelease/2014042802.php.

Some examples currently observable in Japan include:

- Health monitoring systems that can track patients remotely and alert doctors when measurements suggest that a condition may be developing; this includes for example Imizuno Home System, tele-health system implemented by Imizu Municipal Hospital (Toyama Prefecture).
- Fraud prevention systems that react to an apparent change in users' monitored behaviour but which may actually signal the presence of fraudulent activity; an example of this sort of services is the Fraud Prevention System developed by SECOM, a security firm.
- Customer retention systems, such as the one developed by IBM Japan,[35] rely on data on consumers' past transactions to predict and prevent 'churn'.

Box 3-3 on the next page discusses two specific examples.

Box 3-3: DDI in practice – attentive services

DeNA Life Science

DeNA Life Science, Inc. (the healthcare subsidiary of Internet company DeNA) is set to launch a direct-to-consumer genetic testing service, called Mycode, in Japan in mid-August 2014.

As part of this service, DeNA will gather genetic data from consumers through special test kits and analyse the samples for genetic markers that indicate risks of up to 283 test items including some types of cancer, lifestyle-related diseases and other genetic predisposition to health problems.[36]

In order to provide a highly reliable and accurate genetic testing service, the firm has developed, in collaboration with the Institute of Medical Science of the University of Tokyo, a Japanese-specific disease risk prediction algorithm that analyses the data from the kits and compares it with the content of hundreds of academic papers.

The intended result is to determine the statistical probability that a person could develop a certain condition or disorder and to help consumers prevent illness. DeNA is also planning to anonymise the genetic information it collects and, together with the Institute of Medical Science of the University of Tokyo, to jointly create a database of Japanese people's genetic information and study ways of making use of it. One possible use is to share it with medical institutions for the development of new drugs and treatments.

This case illustrates well the life-changing potential of certain uses of data, provided legitimate concerns about highly sensitive data around health can be overcome.

---

35 See http://www-06.ibm.com/software/jp/analytics/spss/solutions/cancell_prevented.html

36 DeNA claims that Mycode will be able to assess the risk of 40 types of cancers, such as stomach, lung and esophageal; 25 types of lifestyle diseases, such as diabetes, high blood pressure and heart disease; 87 types of other diseases and 131 physical predisposition factors, such as obesity and skin qualities.

Komatsu's KOMTRAX vehicle management system

Heavy machinery manufacturer Komatsu offers a machine tracking system (KOMTRAX) whose terminals are installed on construction equipment to record and transmit real-time information: the location of the machine, its cumulative hours of operation, the last date a part was replaced, and general operating condition of the vehicle.

As of March 2013, the system has been installed on over 300,000 vehicles worldwide. Data is centralised and analysed by the KOMTRAX fleet monitoring system and can be accessed by the owners (construction, mining companies) anytime and anywhere via a smartphone and tablet device app with access to daily, weekly and monthly statistics.

The system monitors machine-specific data in real time, but also builds upon information centralised in Komatsu's own system. This enables dealers and service personnel to anticipate users' requirements, and plan timely inspections, part replacements and preventive maintenance measures. Komatsu claims this can lead to a reduction in maintenance-related time and cost for the customer. The system also provides the owners with anti-theft measures that rely on connectivity and real-time location information, rather than simply alarms and immobilisation.

KOMTRAX information is also intended to be used to devise more efficient ways of operating the equipment, based on the characteristics of customers' sites. Selecting more appropriate machine types and using this equipment more efficiently leads to lower fuel consumption and reduced environmental impact – for example, Komatsu points to a potential reduction in fuel consumption of 10% by cutting unproductive idle running time in half.[37] Finally, the system is useful for Komatsu itself: it helps the company's representatives understand their customers' needs and future demand better, and respond accordingly.

3.3.4 Services for people like me: characteristics and correlations to improve business performance

'Services for people like me' target groups of consumers who share common characteristics. They rely on analytics algorithms to identify consumer segments, sometimes with a high degree of granularity ('micro-segmentation'), and then identify new customers as belonging to a given category. This helps to improve the fit between a customer and the products that are being actively marketed to them, which in turn can lead to increased customer satisfaction, and potentially can incite users to spend more (for example on e-commerce sites).

Key examples include:

- comparative effectiveness research (CER), which assists clinicians and patients in making better-informed healthcare decisions[38]

---

37 See http://www.komatsuamerica.com/articles/Komtrax_08no4.pdf

38 The USA appears to be leading the way in this field, whilst Japan appears to be catching up. Japan appears to be using insurance data rather than clinical data, and efforts to link the insurance data to the registry data (disease, drugs and medical device registration data) are relatively recent. The Japanese government is watching initiatives in the EU and USA closely, and moving towards linking the medical information from Electronic Healthcare Record (EHR) with the care information by 2025.

- targeted advertising (which is typically sold in terms of a defined target segment characterised by 'user-level attributes'); examples in Japan include Synergy! 360 (analytics) developed by Synergy Marketing or such as that used by Lawson, Gurunabi and Cookpad)
- recommendation services based on 'collaborative filtering' such as those used by Rakuten or Amazon.

Box 3-4 on the next page provides a specific example.

Box 3-4: DDI in practice – services for people like me

Yahoo! JAPAN's use of predictive analytics

Yahoo! JAPAN is using predictive analytics to give customers what it terms "a unique digital experience". This interfaces with Yahoo! JAPAN's marketing solutions for businesses, enabling corporate customers to connect with Internet users in Japan through multiple channels in a way that is as personalised and relevant to end users as possible.

In order to classify users in meaningful segments that can be used by marketing partners, Yahoo! JAPAN uses algorithms to analyse data from all the services it provides. Thanks to a technical architecture that enables both segmentation on the basis of past data, and real-time analytics, Yahoo! JAPAN is able to anticipate certain future behaviours and serve ads relevant to end users.

This has non-commercial applications as well: for example, Yahoo! JAPAN used predictive analytics to forecast the economic growth index for Japan before the government published it and predicted the number of seats each political party won in Japan's 2013 upper house election accurate to one digit.[39]

Yahoo! JAPAN claims that this enables a first-year increase in advertising sales of 4% (out of a total in excess of JPY200 billion), leading to a return on investment of JPY10 billion, through better segmentation.

The New York Times's 'porous paywall'

Online newspapers often see themselves as having to choose between maximising online traffic, or monetising a small group of loyal readers who may be willing to pay a subscription fee.

In 2010, *The Times* of London put all its online content behind a 'strict' paywall that allowed no free access to articles, and around the same period, industry commentators noted that its online traffic may have dropped by up to 90%.[40] Although advertising losses may have been less pronounced than the drop in traffic, the move was widely seen as trading advertising for subscription revenues.

---

39 See http://www.teradata.co.uk/Resources/Videos/Yahoo-JAPAN-Increasing-ROI-Through-Predictive-Analytics-to-Solve-Customer-Challenges/?LangType=2057&LangSelect=true

40 See, for example, http://seekingalpha.com/article/234330-ny-times-shouldnt-make-the-same-mistake-as-news-corp and http://techcrunch.com/2010/11/02/times-paywall-4-million-readers/

At around the same time, the *New York Times* took a different approach. Based on extensive analysis of its readers' browsing data, it concluded that readers who might be likely to pay (a minority) exhibited different behaviour from those who would not pay (the vast majority). In particular, potentially paying users generally read more than 20 articles per month, and found their news by visiting the newspaper's front page rather than following 'deep links' on third-party websites (e.g. social networks, search engines or news aggregators). The newspaper introduced a paywall that was shown only when users had read more than 20 articles in a month (and articles found via third-party websites did not count towards this limit).[41]

Three years on, the *New York Times*'s approach appears to have succeeded in creating substantial subscription revenue, estimated at over USD90 million per annum,[42] with only minimal disruption to site traffic (reported to be around 10%).[43] This was not a matter of luck, but the result of a sophisticated analytics-driven exercise. But for all the technical complexity involved, the idea is simple. In the words of the project's leader, "there is undoubtedly a segment of your audience that considers your content very valuable and is willing to pay for it. It's a matter of finding them and again asking them, respectfully, to support what you do."

3.3.5 Intelligent planning: large scale pattern analysis to improve complex systems

Finally, 'intelligent planning' services aim to extract patterns from the analysis of large, often disparate and/or anonymous datasets to help firms or governments prepare and respond to large-scale evolutions or trends. This can deal with movements of raw materials and work-in-progress through a complex manufacturing supply chain, as well as the impact on mass transit systems of specific disruptions such as weather events or engineering work.

A common (but not universal) feature among services of this type is the reliance on large sets of anonymised data from consumers to benefit service users (who may or may not be among the consumers whose data is used by the system), wider constituencies or even society at large[44] – for example, when location data from drivers in a traffic jam is used to find alternative routes for service users. Likewise, in many cases data is not 'personal' at all, such as for example in the case of 'just-in-time' supply-chain optimisation.

Notable examples include:

- optimising the management of supply chains by forecasting fluctuations in demand, such as used by Lawson, the convenience-store chain
- the analysis of historical traffic congestion data and school locations to aid in urban planning(e.g. Toyota's Big Data Traffic Information Service for the use of local governments and businesses, IBM Japan, Fujitsu, Honda, NTT Docomo Big Data systems under development)

41 See http://www.nytimes.com/2011/03/18/business/media/18times.html?adxnnl=1&ref=business&src=me

42 See http://go.bloomberg.com/tech-blog/2012-12-20-the-new-york-times-paywall-is-working-better-than-anyone-had-guessed/

43 See http://www.journalism.co.uk/news/two-years-of-the-new-york-times-paywall/s2/a552534/

44 That is, positive externalities.

- identifying flu outbreaks on the basis of web search data (e.g. Google Flu Trends)
- conducting 'sentiment analysis' on social issues or consumer products(e.g. BuzzFinder by NTT Communications)

Box 3-5 below provides two examples.

Box 3-5: DDI in practice – intelligent planning

Fujitsu's Akisai cloud service for farms and food-related industries

Fujitsu offers a SaaS[45]-based solution called 'Akisai cloud service' for agricultural production management, targeted towards both agriculture and food-related industries.

As of January 2014, the Akisai user base comprises 160 companies, including Aeon Agri Create, Fukuhara Farm and Eto Sangyou.[46]

The Akisai cloud service uses mobile devices to collect data on day-to-day operations, which are then stored and analysed 'in the cloud' by Fujitsu's own servers and software. This provides farmers with information such as the results of daily on-site operations and planting information, and enables them to visualise metrics indicative of the quality, cost structure and profitability of each individual crop. With access to data accumulated over time, farmers can analyse the results of previous plans and modify their approach accordingly. This brings enterprise-style management to smaller agricultural companies for whom this would have been excessively complex and costly previously: such information was previously not easily available, and relied on a high degree of sophistication both in agriculture and in analytical techniques by farms.

This improves both earnings and efficiency. According to Eto Sangyou,[47] fertiliser costs were cut by up to 30% and overall production costs by around 10%; the system also enabled increased production and sales on the same plot of land by around 1.4 times.

By centrally managing data (such as production schedules, production history, harvest volume and planting information) from the contract farm product suppliers across Japan, the system displays the status of products from the production stage. This enables contract producers to engage in production management, supporting procurement based on a fixed-time, fixed-quantity, fixed-quality, and fixed-price model. The system also allows for time saving of up to 80%[48] and minimisation of manual-input errors.

Food-related companies, such as food processors, wholesalers, retailers, and restaurants, are also able to establish a process for regulating supply and demand adjustments as well as quality control with their contracted producers.

45 Software-as-a-service, meaning its customers do not need to invest in hardware and software themselves, and can use Fujitsu's service online directly.

46 See http://www.fujitsu.com/global/vision/2014/casestudy/aeon-agri-create/.

47 See http://jp.fujitsu.com/solutions/cloud/agri/movies/#case.

48 See http://www.fujitsu.com/global/vision/2014/casestudy/aeon-agri-create/.

Metawater

Metawater has been offering a service called Smart Field Service (SFS) on its Water Business Cloud (WBC) platform since 2013, covering about 100 water purification plants in Japan it now services under contract.

The Water Business Cloud collects data from sensors installed throughout water facilities to monitor the operational status of the infrastructure, variables such as the level, quality and pressure of the water supply and information on daily inspections and crisis response related to water treatment facilities as well as know-how of skilled engineers. This way the firm has been able to streamline the operation and maintenance of facilities and reduce costs. As part of the service it offers, Metawater is sharing this information with local governments and water management enterprises.

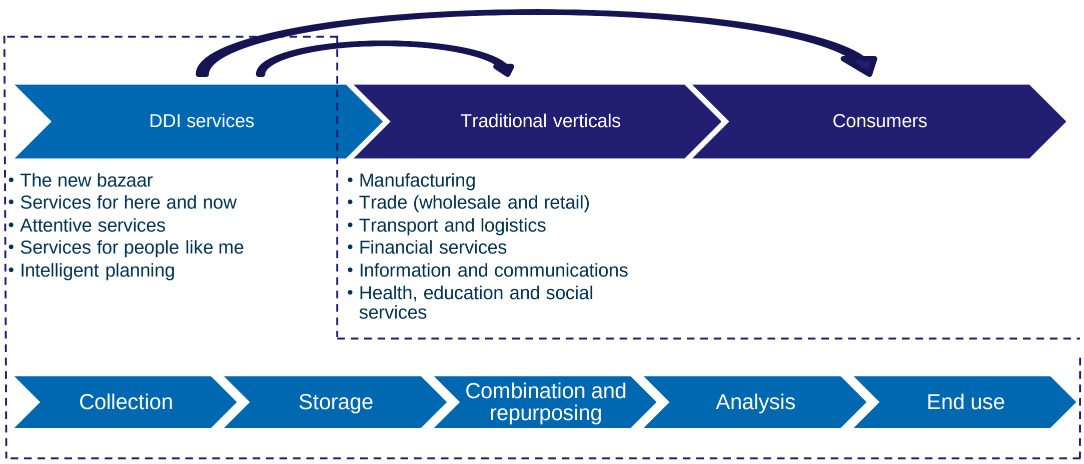
In future, Metawater plans to further optimise operations and maintenance by analysing data in a much more extensive way while developing a platform to cover the entire water treatment business. This will include, for example, the integration of operations and maintenance activities between treatment plants on the one hand, and pipelines on the other hand. This should further reduce costs whilst optimising the quality and safety of water treatment and transport in Japan.

3.4 The role of DDI services in the wider economy

The examples of DDI services above illustrate an important point: there is a wide range of uses for data-driven innovation throughout the economy. The direct customer of a DDI service can be a consumer (in which case we speak of a business-to-consumer service, B2C) or another firm (in which case we speak of a business-to-business service, B2B).

Although B2C services may be very visible to consumers (for example the users of Navitime or Google Maps services), the impact of data-driven innovation on the economy is primarily in the B2B sector, as an enabler for a wide range of economic activities in all sectors of the economy, as shown in Figure 3.2.

Figure 3.2: Mapping of DDI services and their role in the economy [Source: Analysys Mason, 2014]



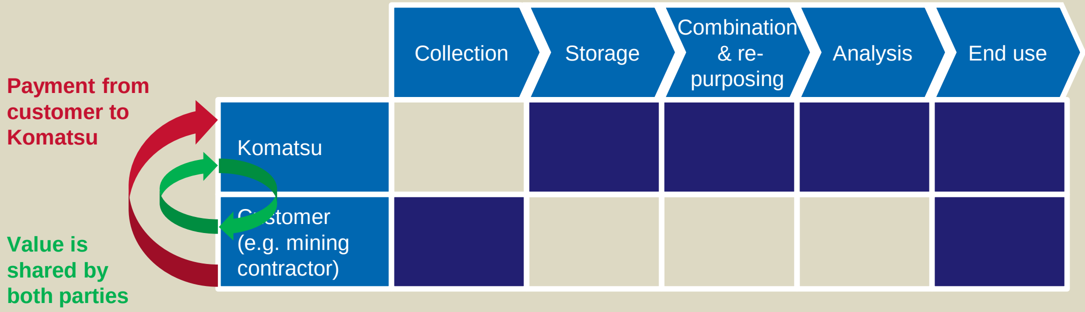
Box 3-6 below illustrates how firms can work together across this DDI value chain based on two examples of DDI services developed in the previous section.

Box 3-6: Roles taken in the DDI value chain

The following figures are visual representations of two of the five DDI service types described above. Participating organisations are shown on the left, with the curved arrows indicating the overall flow of data in the service. These diagrams illustrate how different parties in the DDI value chain perform different tasks.

In the Komatsu example, the heavy machinery connected to the KOMTRAX system is owned and operated by Komatsu's customers, for example a mining company. The data is collected by the machine and sent to Komatsu, which then processes it in order to use it itself (for example to understand demand trends) and to share some of it back with the customer (for example statistics on running time and other worksite-related information).

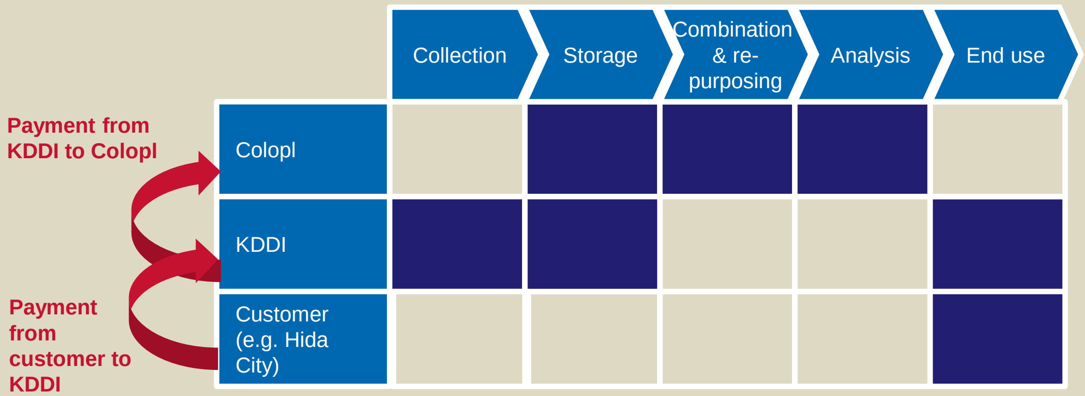
Figure 3.3: Komatsu's KOMTRAX data flow analysis [Source: Analysys Mason, 2014]



In the case of Colopl's collaboration with KDDI, the data is collected by KDDI and shared, apparently in pseudonymised form, with Colopl for processing and analysis. The data is then used by KDDI and its customers, such as Hida City in Gifu prefecture.

This example also highlights how firms that are specialised providers of DDI services and expertise, such as Colopl, can work with other firms so that the benefits of data-driven innovation propagate throughout the economy.

Figure 3.4: Colopl and KDDI's revenue and data flow analysis [Source: Analysys Mason, 2014]



In order to quantify the impact of data-driven innovation on the economy of Japan, we consider the use and value of DDI services in the following traditional sectors or 'verticals':

- manufacturing
- trade (wholesale and retail)
- transport and logistics
- financial services
- information and communications
- health, education and social services.

These verticals are essential to the economy of Japan, both domestically (e.g. retail trade, healthcare and education) and as a major global economic power. Together, we estimate that these verticals represent about 71% of the gross value added (GVA)[49] estimated for 2013 in Japan, as well as 45 million jobs, or 70% of total employment, as shown in the charts below.

Figure 3.5: Contribution to GVA of the six verticals modelled in 2013 [Source: Analysys Mason, Statistics Bureau][50,51]

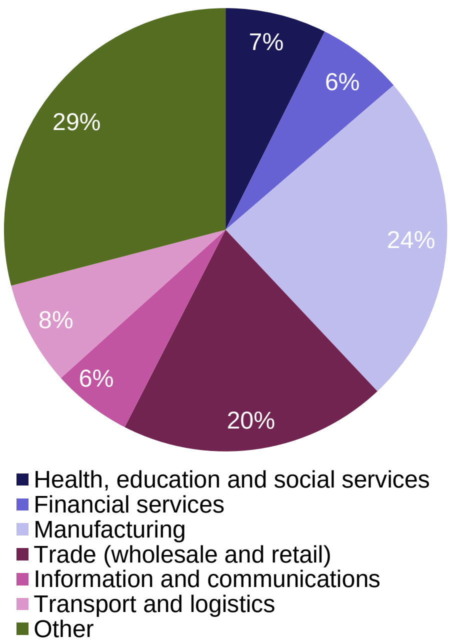
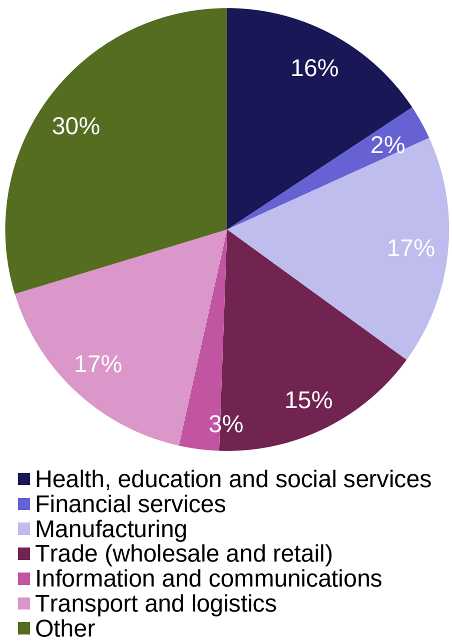


Figure 3.6: Contribution to employment of the six verticals modelled [Source: Analysys Mason, 2012 Employment Status Survey][52]



---

49 Gross value added (GVA) is a measure of GDP that considers output (revenues) minus the costs of intermediate goods and services; in this paper, we have measured the value of DDI services based on the incremental net revenue and the costs savings that use of DDI services brings to each vertical, which is not identical to GVA but provides a useful measure of economic value based on consumer and producer surplus.

50 Compiled from Information Systems Department, Policy Research Institute and Ministry of Finance.

51 Official data is available for 2011 by sector, for 2012 in aggregate and for some (but not all sectors); values shown here are adjusted for expected growth between 2012 and 2013, to show our estimates of the split of GVA by sector in 2013.

52 The relatively lower contribution to GVA of the 'Health, education and social services' vertical appears linked to the classification of public education expenditure; although these expenses do not appear directly in GVA, they support significant employment and DDI has the potential to make this expenditure more efficient, as discussed in the following section.

***

This completes our discussion of our analytical framework (types of DDI service and value chain activities), and our initial view of DDI services and their role in the wider economy. In the next section, we will look at this again from a quantitative perspective, providing estimates of the economic value contributed by data-driven innovation to the Japanese economy.

4 Economic value of data-driven innovation in Japan

As explained in Section 3, the DDI 'universe' contains a multitude of services, used increasingly widely by consumers and businesses. These services, classified into the five categories described in Section 3.3, already create a potentially large amount of value to the Japanese economy and to society more generally.

The purpose of this section is to illustrate this impact in terms of two broad questions:

- What is the economic contribution of data-driven innovation to the Japanese economy today? This question is addressed in Section 4.1. We estimate DDI services' current value to the Japanese economy to be in excess of JPY7 trillion per annum today. Importantly, more than half of the value that we have identified ultimately accrues to consumers rather than firms.
- What is the potential for growth from a wider adoption of innovative data-driven services in the economy? Section 4.2 provides our view on how the economic value of data-driven innovation is likely to grow by 2020, with specific examples of how data-driven innovation can support Japan's ongoing economic transformation. Quantitatively, this value could reach in excess of JPY15 trillion by 2020, provided firms adopt services more broadly and the policy environment remains conducive to such growth.

Also, in Section 4.2.3, we focus on how data-driven innovation contributes to public and social value through a range of examples, showing how Japan's society can benefit from these services beyond the economic sphere.

Together, this shows that DDI services are already contributing significantly to economic output in sectors of the economy where Japan's firms are recognised as world class, such as transport, advanced manufacturing and financial services.

The potential for growth between now and 2020 is real and significant, and provides an illustration of how technology can support the economic transformation sought by the government of Japan. We summarise our findings in Section 4.3.

4.1 The value of data-driven innovation in Japan

As a first step, we have estimated the current contribution of data-driven innovation to the Japanese economy. At a high level, economic value relates to the profits that firms can make from selling goods and services ('producer surplus'), and to the value that consumers derive from purchasing those goods and services ('consumer surplus').[53] Our approach is explained in Section 4.1.1 and illustrated in Section 4.1.2, and the resulting estimate of the current value of data-driven innovation to the economy is summarised in Section 4.1.3.

[53] Consumer surplus stems from prices being lower than willingness to pay for goods and services; for example, if a consumer is willing to pay JPY500 for a coffee, and the price is JPY200, then a surplus is generated equal to JPY300, which contributes to overall welfare; for the purpose of this paper, we have focused on the impact of DDI on prices, rather than on its impact on willingness to pay, which is less tangible.

4.1.1 Approach

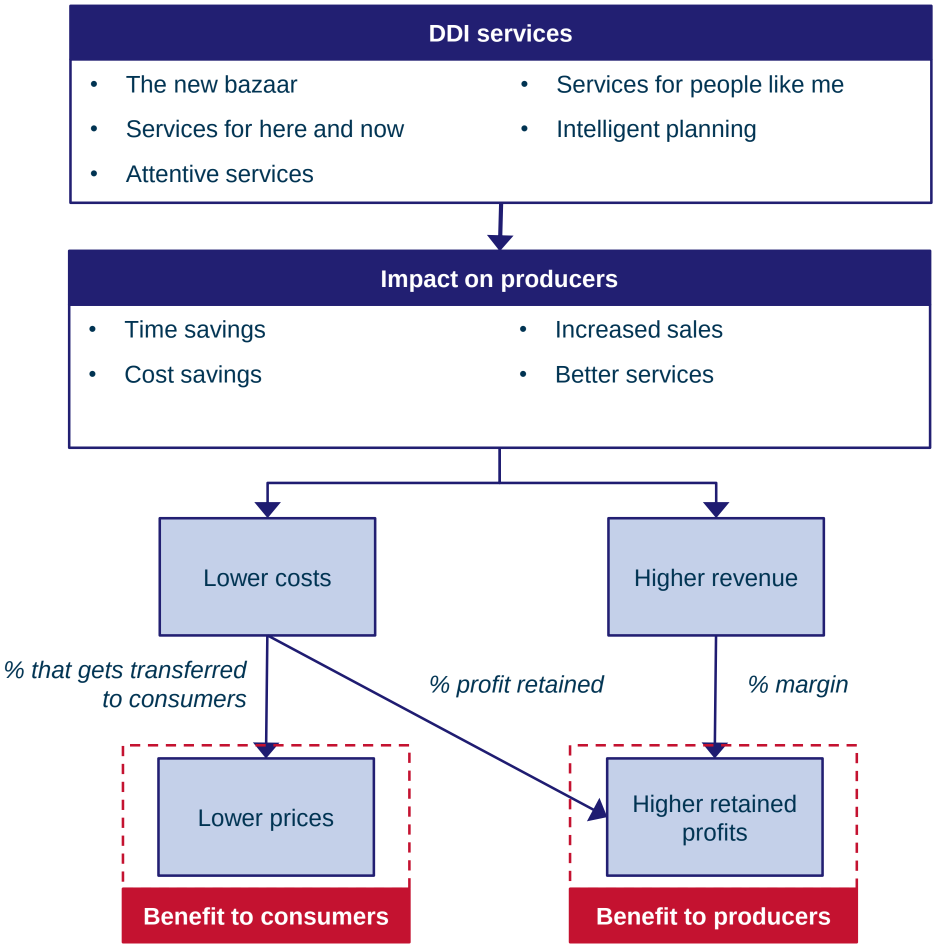
Overview

In what follows, we consider how data-driven innovation contributes economic value to both producers and consumers, by focusing on the impact that DDI services have on the productive economy. Specifically:

- DDI services enable firms to sell more, and produce more efficiently
- these improved economics (higher revenue and lower costs) translate into higher profits
- as most sectors of the economy are highly competitive, firms can use these higher profits to compete more aggressively.

These competitive dynamics typically result in lower prices for consumers, and so the value created is shared between producers and consumers, as illustrated in Figure 4.1 below.

Figure 4.1: Economic value for producers and consumers [Source: Analysys Mason, 2014]



In addition to this flow of economic benefits from producers to consumers, there are further gains for consumers stemming from DDI as a result of better services. For example, thanks to data-driven innovation, consumers can be provided with more customised services, and the pricing of these services can be more adapted to their circumstances; they can also save time thanks to services 'for here and now' and benefit from the 'intelligent planning' that makes transport links more efficient. This quality gain for consumers,[54] over and above the benefit they get from lower prices, is an integral part of consumer value, or surplus; it is, however, very difficult to quantify, and so we have omitted it from the quantitative part of our assessment of the value of data-driven innovation.

Current and future impact

The main objective of this section is to provide a quantified estimate of the contribution of DDI services in Japan's economy in 2014. This forecast is based on expected growth of adoption of the various service types in each of the six verticals considered. Our DDI adoption estimates for 2014 are based on data from publicly available studies of other advanced economies, refined on the basis of a focused primary research programme involving 100 companies, large and small, in all six sectors modelled.

This forms the basis for a forward-looking view of the potential for data-driven innovation to support economic growth to 2020, in line with the government's Revitalization Strategy (the "Third Arrow"). In assessing potential 2020 adoption, we have assumed that companies will be able to progress gradually from 'observers' of DDI services to 'adopters'. We assume that companies that are currently developing DDI strategies/implementation plans will begin to pilot DDI initiatives within the next five years. This approach is moderated by the assumption that the contribution of data-driven innovation to cost and revenue within adopters will not change significantly; we believe this to be relatively conservative given the pace of innovation in this sector.

The impact of data-driven innovation on revenues and costs

As described above, we have approached the question of economic value from the perspective of producers, with consumers' benefits flowing from producers' benefits through lower prices. Specifically, we have attempted to answer the following questions:

- What cost savings and incremental revenues do data-driven innovation and associated services enable for producers (i.e. corporations) in a representative part of Japan's economy?
- What proportion of this value is passed on to consumers as lower prices?
- What proportion is retained by producers as higher profits?

We have asked these questions for every combination of industry verticals and DDI service type.

54 This is an increase in willingness to pay.

To answer the first question, we critically assessed a wide range of publicly available examples and case studies illustrating the impact that specific DDI services have had on established industries to date, both in Japan and around the world. Based on this, as well as on further research (including primary research conducted in Japan), we estimated a range of cost savings and revenue increases that could be enabled by each combination of DDI service type and vertical, and then scaled this to an estimated level of adoption of services of each type of vertical.

Our estimate of how these benefits are then split between retained profits (which are benefits to producers) and price reductions (which are benefits to consumers) relied on an assessment of each vertical's degree of profitability and competitiveness, with profit margins across verticals used as a proxy.

4.1.2 Illustration of our approach in selected cases

We conducted the analysis described above for each of the six industry verticals of Japan's economy mentioned in Section 3.4 in connection with each of the five types of DDI service described in Section 3.3. This means that we analysed 30 (i.e. 6 sectors × 5 service types) different scenarios. Before providing our quantitative estimates of the value of data-driven innovation in the Japanese economy (which we do in Section 4.1.3 below), by means of illustration, we discuss our reasoning in the case of five selected combinations of industry vertical and DDI service type. The selected cases are listed in Figure 4.2 and are discussed below.

Figure 4.2: Economic value of DDI services for Japan's producers, in 2014[55] [Source: Analysys Mason, 2014]

Map of cases discussed below	Health, education & social services	Financial	Manufacturing	Trade	Information and communications	Transport and logistics
The new bazaar			X			
Services for here and now						X
Attentive services		X				
Services for people like me			X			
Intelligent planning				X		

55 Figure 4.3, Figure 4.4 and Figure 4.5 show the average of the ranges calculated using maximum and minimum likely effects and adoptions in each service type, for each vertical.

► *The new bazaar: application to manufacturing*

This first type of DDI service can add significant value to companies in the manufacturing sector. Net benefits are largely associated with cost savings, although more competitive manufacturing processes enable better competitiveness and increased turnover.

The advent of small-scale manufacturing technologies, such as 3D printing, is a business opportunity for small-scale Japanese companies, which represent 11% of manufacturing firms in terms of value of goods produced. Moreover, small-scale manufacturing can reduce transport costs by producing goods on the customer's site. This applies to general manufacturing (19% of all manufacturing),[56] rather than chemical or machinery manufacturing, which are likely to be produced in a factory and transported. Land transport costs are estimated to represent 4% of Japanese manufacturers' costs.[57]

Our assessments of the 2014 and 2020 potential adoption and economic impact of DDI services of the 'new bazaar' type on the manufacturing sector are given below:

		2014	2020
Adoption within vertical sector (%)		15%–25%	45%–50%
	Revenue increase	Related gross margin impact	Cost saving
2014 DDI impact on adopters (%)	0.38%–0.57%	Vertical margin: 3%	0.52%–0.78%
2014 impact (JPY billion)	229–574	7–17	305–763
2020 impact (JPY billion)	745–1,242	22–36	992–1,653

► *Services for here and now: application to transport and logistics*

DDI 'services for here and now' contribute to transport and logistics in two main ways.

First, they increase the potential to serve more customers through fleet management and smart routeing, which can match supply of vehicles to the routes demanded by customers. It is estimated that this type of service can increase road passenger fare-related revenue by 5% where road passenger fares represent 5% of transport revenues,[58] by encouraging customers to use these services when they would otherwise walk or not travel. There is also a potential revenue source from package tracking, which is likely to be less material.

56 General manufacturing includes food, beverage, printing and other types of manufacturing. See http://www.stat.go.jp/data/nenkan/zuhyou/y0804000.xls.

57 Transport is estimated to make up 5% of manufacturers' costs (based on international benchmarking) and land transport is expected to make up around 80% of this. It is assumed that land transport is used for domestically consumed products, so this is the balance between manufacturing exports and domestic consumption. Exports represent 16% of manufacturing turnover (see http://www.stat.go.jp/data/nenkan/zuhyou/y1503000.xls).

58 We assume that this type of service can increase road passenger fares by 5% by encouraging use of transport services other than walking. Road passenger revenues represent 5% of transport revenues (see http://www.e-stat.go.jp/SG1/estat/XlsdlE.do?sinfid=000023619100).

Second, and more significantly, smart routeing can save maintenance, wages and fuel costs by optimising routes to reduce distances travelled and time taken. Savings are expected to be in the order of 10%–25% of road transport costs for adopters.[59]

Our assessments of likely adoption and economic impact of ‘services for here and now’ on the transports and logistics sector are given below:

		2014	2020
Adoption within vertical sector (%)		25%–35%	50%–60%
	Revenue increase	Related gross margin impact	Cost saving
2014 DDI impact on adopters (%)	0.26%–0.76%	Vertical margin: 4%	1.76%–3.28%
2014 impact (JPY billion)	46–187	2–7	311–809
2020 impact (JPY billion)	99–347	4–14	647–1,443

► *Attentive services: application to financial services*

The main value from this service type is linked to fraud prevention and time/cost saving, rather than revenue generation. Fraud is a concern in this vertical sector given the nature of business in finance, banking and insurance. DDI services are being used to minimise fraud risks through the analysis of transactions to discover anomalies which may indicate fraudulent activity, leading to more accurate risk assessment.

We have used both international and national benchmarking to calculate the savings from fraud prevention related to DDI services, to ensure the most realistic cost-saving limits. Fraud represents 0.9% of total costs to the financial industry[60] and 0.3% of costs to the insurance industry.[61] Fraud is considered a cost to finance rather than a loss of revenue, as in other industries.

Our assessments of likely adoption and economic impact of ‘attentive services’ on the financial sector are given below:

		2014	2020
Adoption within vertical sector (%)		30%–35%	50%–60%
	Revenue increase	Related gross margin impact	Cost saving
2014 DDI impact on adopters (%)	0.07%–0.10%	Vertical margin: 11%	0.25%–0.71%
2014 impact (JPY billion)	24–42	3–4	87–265
2020 impact (JPY billion)	43–77	5–8	146–492

59 See http://public.dhe.ibm.com/common/ssi/ecm/en/nic03008deen/NIC03008DEEN.PDF and http://www.critchlow.co.nz/files/5113/5806/1939/Critchlow_Smart_Routing_Infosheet.pdf.

60 This is based on the average of two international case studies. Nationwide used attentive services to reduce fraud losses by 75% (see http://www.sas.com/en_us/customers/nationwide.html). Viseca Card Services reduced fraud losses by 15% (see http://www.sas.com/offices/africa/southafrica/success/Visecacardsvcs.html).

61 The total value of insurance claims in Japan was JPY4,775 billion in 2012 (see http://www.sonpo.or.jp/en/publication/pdf/fb2013e.pdf). The rate of fraudulent claims in Japan is around 2.5% (see http://www.studymode.com/essays/Insurance-Fraud-81268.html).

► *Services for people like me: application to manufacturing*

In the manufacturing sector, the main value of this type of DDI service is related to an improvement in the effectiveness of marketing (e.g. micro-segmentation of customers). Companies can understand their customers better using analytics. For example, an Italian home appliance manufacturer saw a 10% increase in turnover following the implementation of systems to allow rapid extraction of customer data across segments, which enabled it to understand customer product preferences and willingness to pay in different markets.[62] Normal impact levels are expected to be lower than this.

Such services can reduce time spent by marketers on extracting resource planning data by 40%, therefore reducing the time spent on planning duties.[63] Time savings can also be achieved using comparative effectiveness in the biomedical manufacturing industry, for example in the case of drug research.

Our assessments of likely adoption and economic impact of 'Services for people like me' in the manufacturing sector are given below:

		2014	2020
Adoption within vertical sector (%)		10%–20%	35%–40%
	Revenue increase	Related gross margin impact	Cost saving
2014 DDI impact on adopters (%)	0.10%–0.15%	Vertical margin: 3%	0.13%–0.19%
2014 impact (JPY billion)	40–121	1–4	50–151
2020 impact (JPY billion)	153–262	4–8	191–328

► *Intelligent planning: application to trade*

Additional turnover can be generated in this vertical by optimising in-store merchandising, adding insight to marketing campaigns and improving customer service using workforce optimisation. Third-party case studies indicate turnover increases of around 1%, though we assume a more conservative impact of around half of this.[64]

There are also time/cost saving benefits, in particular related to supply chain optimisation. Optimised merchandise planning can make product routes more efficient and reduce inventory. These programmes can realise savings of around 15% of supply chain costs in the trade vertical.[65]

Our assessments of likely adoption and economic impact of 'intelligent planning' on the trade sector are given below:

62 See http://www.sas.com/offices/NA/canada/fr/success/faber.html.

63 *Ibid.*

64 See http://www.infor.com/shared_resources/casestudies/REI.pdf.

65 *Supply chain network optimisation,* Deloitte (2007).

		2014	2020
Adoption within vertical sector (%) – Retail – Trade		17%–28% 20%–35% 15%–25%	45%–52% 55%–65% 40%–45%
	Revenue increase	Related gross margin impact	Cost saving
2014 DDI impact on adopters (%)	0.32%–0.48%	Vertical margin: 2%	0.33%–0.50%
2014 impact (JPY billion)	282–719	5–12	295–753
2020 impact (JPY billion)	825–1,419	13–23	849–1,462

4.1.3 Quantified estimate of the economic value created by data-driven innovation in 2014

Results

Based on the analysis outlined above, by aggregating our estimates across the six verticals we conclude that data-driven innovation and associated services already contribute over JPY7 trillion to Japan's economy in 2014,[66] in the form of profits for producers and cost savings for consumers. By comparison, this is equivalent to around 3.4% of GVA in the selected verticals (a key measure of GDP). This is summarised below in Figure 4.3 (for producers) and Figure 4.4 (for consumers):

Figure 4.3: Economic value of DDI services for Japan's producers, in 2014[67] [Source: Analysys Mason, 2014]

JPY billion	Health, education & social services	Financial	Manufacturing	Trade	Information and communications	Transport and logistics	Total
The new bazaar	1	180	172	11	51	29	**444**
Services for here and now	1	14	435	101	11	173	**734**
Attentive services	3	125	40	11	8	86	**273**
Services for people like me	17	57	33	13	20	12	**153**
Intelligent planning	55	201	428	113	23	115	**935**
Total	**77**	**576**	**1,108**	**249**	**114**	**415**	**2,539**

66 The average of our overall range is between JPY4,614 billion and JPY10,081 billion; all figures in this section are central estimates based on a range of values.

67 Figure 4.3, Figure 4.4 and Figure 4.5 show the average of the ranges calculated using maximum and minimum likely effects and adoptions in each service type, for each vertical.

Figure 4.4: Economic value of DDI services for Japan's consumers, in 2014 [Source: Analysys Mason, 2014]

JPY billion	Health, education & social services	Financial	Manufacturing	Trade	Information and communications	Transport and logistics	Total
The new bazaar	1	63	374	28	48	66	**580**
Services for here and now	1	5	1,009	376	5	392	**1,789**
Attentive services	3	52	78	20	3	200	**357**
Services for people like me	26	14	71	46	11	25	**192**
Intelligent planning	76	85	964	419	23	267	**1,834**
Total	**106**	**219**	**2,496**	**890**	**90**	**950**	**4,752**

Figure 4.5 brings together producer and consumer benefits. We have included a comparison of this value with the total gross value added (GVA) in each vertical for illustration.[68]

Figure 4.5: Summary of economic value of DDI services in Japan, in 2014 [Source: Analysys Mason, 2014]

JPY billion	Health, education & social services	Financial	Manufacturing	Trade	Information and communications	Transport and logistics	Total
The new bazaar	2	243	546	39	100	95	**1,024**
Services for here and now	2	19	1,444	477	17	565	**2,523**
Attentive services	6	177	118	32	11	287	**630**
Services for people like me	43	71	103	60	31	37	**345**
Intelligent planning	131	286	1,393	532	46	382	**2,769**
Total	**183**	**795**	**3,604**	**1,140**	**204**	**1,365**	**7,291**
As % of GVA in vertical	**0.8%**	**4.1%**	**4.8%**	**1.9%**	**1.1%**	**5.9%**	**3.4%**

68 This is a useful comparison between the economic value created by DDI (consumer and producer surplus) compared to the total economic value in each vertical. The flow of value to consumers comes only from price reductions, themselves made possible by cost reductions in each vertical, which makes the comparison relevant.

Figure 4.6 summarises how the economic value of data-driven innovation in Japan is divided between producers and consumers:

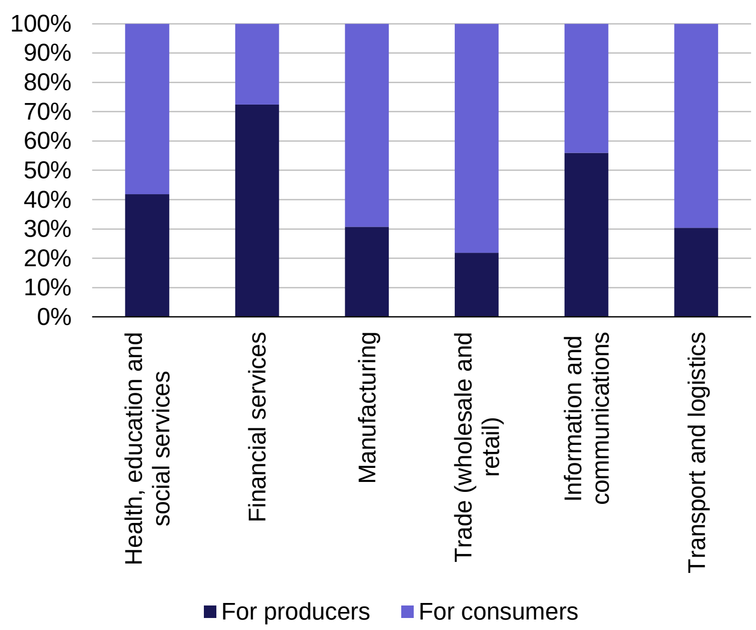


Figure 4.6: Split of economic value of DDI services by consumers and producers [Source: Analysys Mason, 2014]

Key observations

► *The main economic impact is on the manufacturing and transport verticals*

Our estimates suggest that data-driven innovation has particularly high potential in the manufacturing sector in absolute terms, and in the transport vertical relative to its GVA. These sectors are followed by the financial vertical in relative terms, and the trade vertical in absolute terms.

The most valuable type of service today is 'intelligent planning', and the underlying dynamic is the large efficiency gains that data-driven production planning and supply chain management can bring to goods-producing industries under the manufacturing vertical.

Whilst the manufacturing and transport sectors appear to benefit most from data-driven innovation today, in other sectors such as healthcare, education, social services, ICT and trade the value of data-driven innovation appears less well realised. We note however that in the case of many of these areas data-driven innovation creates significant social value, which we have not quantified (but which we discuss in Section 4.2.3); similarly, from our perspective, ICT plays the role of an enabler for value to be created in other industries through data-driven innovation, rather than a direct profit centre.[69]

69 We note that many DDI-specialised companies will be classified in the ICT sector; for the purpose of this analysis, we are considering the impact of GVA at the point of use of DDI services, rather than the transfer of value from various sectors that use DDI to ICT-sector companies that provide DDI services.

► *Consumers are the main beneficiaries*

For all verticals other than financial services and information and communications, consumers retain a greater proportion of economic value than producers, and this conclusion also holds overall. This means that – unlike what may be a common perception – the main beneficiaries of data may be consumers rather than businesses. Notably, this is without even considering consumers' benefits from having access to new or better services (as our estimation only considers cost savings), and before the social benefits discussed in Section 4.2.3 are taken into account.[70]

► *What our approach omits*

This approach captures the contribution of data-driven innovation and associated services to macro-economic aggregates such as GDP, for major sectors of the Japan economy. While this illustrates the importance of data-driven innovation in the economy as a whole, it should be noted that it understates economic value in three important respects:

- The six verticals modelled, and the underlying sectors of the economy they capture, represent 71% of Japan's overall GVA. There is undoubtedly value associated with data-driven innovation in the remainder of the economy, although we believe we have captured the majority of producer benefits in a meaningful way.
- We make no attempt at quantifying the total consumer surplus impact from higher willingness to pay (e.g. better customer service drives a higher perceived utility for a given good or service, which increases economic welfare), and so our estimates reflect only part of the impact of data-driven innovation on consumer welfare.
- Finally, certain forms of 'social value', which we explore in Section 4.2.3, are related to externalities. For example, cleaner fuel helps protect the environment, which has externalities in terms of public health and climate. We do not seek to quantify these externalities in this paper.

All of these considerations mean that our estimates of the value of data-driven innovation are likely to be fairly conservative.

► *Comparison to recent MIC study*

We note that the Ministry of Internal Affairs and Communications (MIC) recently published its annual white paper on ICT in Japan.[71] This includes a very wide-ranging review of the state of the ICT sector in Japan, as well as, for the first time, a detailed study on the use of data by Japanese companies and its expected economic impact. In Box 4-1 below, we highlight the differences in approach and find that the conclusions of both this study and the MIC study are compatible.

70 These benefits are extremely difficult to quantify and we have not attempted to do so here.

71 Available at http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h26/pdf/index.html.

Box 4-1: Comparison between this study and the MIC ICT White Paper

The definition of data used by MIC is broader than the scope of this study and includes for example accounting and point-of-sale data, telephone numbers and the use of telephony and video feeds from security cameras. In contrast, this study focused on innovative services that are data-driven, as developed earlier in this report. This will in some cases use data such as security camera feeds or telephone call records, but only to the extent that it can be processed automatically to create value.

Having conducted an extensive survey programme and analysed survey and economic data since 2005, MIC concluded that data-related activities contributed an estimated JPY60.9 trillion to the turnover of Japanese firms, in excess of 4% of total turnover (total = JPY1,335.4 trillion). Importantly, this refers to total turnover/revenue, and appears to include intermediate goods and services.

Separately, the MIC survey identified a statistical correlation between the use of different types and sources of data on the one hand, and the contribution to real GDP on the other hand. Because the review considers inputs between 2005 and 2012, MIC's findings suggest that types of data that are well established[72] appear correlated with real GDP, whereas others that are emergent[73] are less so.

Whilst MIC's numbers relate to revenue, the estimates shown in this section relate to profits (and indirectly GVA, which is a major component of GDP). GVA and GDP are both materially smaller than revenues in the economy, as they exclude all consumption of intermediary goods and services. Broadly speaking, GVA is about a fifth of turnover in the Japanese economy. Simplistically, this means that 1 yen of value added requires 5 yen of revenue, everything else being equal.

In our own modelling, as shown in Figure 4.5, we find that in the six industries we reviewed (which represent ~70% of value added), the current contribution of data-driven innovation is around JPY7 trillion. This represents around 3.4% of the GVA in these industries and is quite comparable to the contribution arrived at by MIC, despite the differences in scope, approach and measure.

4.2 Supporting Japan's revitalisation strategy: DDI's growth potential to 2020

The importance of data-driven innovation to advanced economies such as Japan goes beyond its current contribution. Indeed, both primary and secondary research suggest that overall adoption of DDI services in Japan remains relatively limited, at around 20–25% of firms.

This implies vast headroom for growth, as services become relevant to more firms, barriers to adoption decrease (e.g. technology costs, manpower skills) and the benefits of DDI services become increasingly apparent. As we will discuss in Section 5, a conducive set of policies and a balanced regulatory environment are critical enablers for greater adoption of data-driven services.

72 Customer data, accounting data, POS data, business diary, voice log, fixed IP telephone, mobile phone, E-commerce log, GPS data, e-mail, blog post, access log, video viewing log.

73 RFID data, weather data, traffic information, security camera, sensor log.

4.2.1 Potential increase from higher adoption in 2020

The contribution of data-driven innovation and associated services to the Japanese economy in 2014 is highly significant, with a conservative estimate suggesting around JPY7,345 billion on the basis of the six verticals we have considered, and with close to 66% of this value accruing to consumers.[74]

We believe that the value related to data-driven innovation will continue to grow strongly in the next few years, based on drivers that are already visible and taking into account Japan-specific aspects such as the government's goals for 2020 as part of the Japan Revitalisation Strategy. These include advancing the robotics industry and completing preparations for Tokyo to host the 2020 Olympics and Paralympics.

On the whole, current research suggests that adoption could double (to 40–50%) by 2020, with manufacturing and trade showing the greatest potential for growth as compared to other verticals such as transport, for which adoption is already comparatively high.

Based on these assumptions for increased adoption in each vertical, Figure 4.7 summarises our estimates of the economic value that DDI services contribute to the Japanese economy in 2020. These figures are expressed in (real 2014) JPY, and therefore exclude the impact of inflation.

Figure 4.7: Summary of economic value of DDI services in Japan, in 2020 [Source: Analysys Mason, 2014]

JPY billion	Health, education & social services	Financial	Manufacturing	Trade	Information and communications	Transport and logistics	Total
The new bazaar	6 (234%)	444 (83%)	1,351 (147%)	90 (129%)	183 (84%)	181 (91%)	**2,253 (120%)**
Services for here and now	5 (171%)	60 (225%)	2,925 (103%)	1,023 (114%)	30 (79%)	1,054 (87%)	**5,097 (102%)**
Attentive services	10 (53%)	326 (84%)	260 (121%)	87 (176%)	22 (102%)	528 (84%)	**1,232 (96%)**
Services for people like me	72 (68%)	160 (124%)	266 (157%)	114 (91%)	58 (89%)	67 (83%)	**737 (114%)**
Intelligent planning	277 (112%)	567 (98%)	2,972 (112%)	1,174 (121%)	110 (140%)	681 (78%)	**5,780 (108%)**
Total	**369 (102%)**	**1,557 (96%)**	**7,773 (115%)**	**2,487 (118%)**	**403 (98%)**	**2,510 (84%)**	**15,099 (107%)**
% of GVA in vertical	**1.5%**	**7.5%**	**9.8%**	**3.8%**	**2.1%**	**10.0%**	**6.5%**

Note: Total % growth from 2014 is shown in parentheses.

74 This is based on vertical-by-vertical assumptions on how cost reductions for producers flow into lower prices for consumers, which are linked to the competitiveness and margins in each vertical; this results in more value flowing to consumers in verticals with higher competition/lower margins than in verticals with lower competition/higher margins.

Looking ahead to 2020, the contribution of DDI services to the economy could be equivalent to some 6.5% of GVA, up from 3.4% of GVA in 2014.[75]

The main areas of growth in terms of service types are expected to be in ‘the new bazaar’ and ‘services for people like me’, highlighting the importance of personalisation, micro-segmentation and powerful granular data analysis in underpinning future economic performance. These are closely followed in growth terms by ‘intelligent planning’, which will remain the largest type of service in value creation terms.

It is clear from this analysis that data-driven innovation and associated services will play an increasingly large role in enabling Japan’s economy to continue growing in a sustainable way, by enabling continued improvements in productivity and helping to sustain its competitive advantage internationally.

4.2.2 Impact of data-driven innovation on competitiveness, labour markets and economic growth

This forward-looking perspective reflects the potential for data-driven innovation in the Japanese economy. In the context of the recent revision of the Japanese Revitalisation Strategy (the government’s ‘third arrow’ of the economic reform), it is useful to explore how data-driven innovation can support some of the specific objectives of the government.

The first point of note is that our modelling of DDI services considers primarily efficiency savings. The impact of revenue increases is moderated by the application of a margin that is typically low, in line with the net profitability of the sectors modelled. In addition to these cost benefits, however, data-driven innovation supports two additional effects that can boost the earning power of Japanese companies:

- By making goods and services more attractive to consumers, data-driven innovation can increase the willingness to pay for these goods and services. This should in principle stimulate demand at a given price level, thereby increasing consumer spending.
- In addition, the development and implementation of data-driven techniques throughout the Japanese economy has the potential to further stimulate exports, in two ways: first, a large domestic market provides a fertile environment to develop and perfect these techniques; second, efficiency improvements can work (together with a lower yen) to make Japanese exports more competitive in global markets.

Furthermore, investing in increasing cost-efficiency is an essential driver of productivity improvements. This is particularly true with regard to technology that can alleviate labour shortages, in an economy where capital is widely available. Three examples are particularly relevant in the context of the Revitalisation Strategy: healthcare, agriculture and rural economies.

75 See the approach outlined in Section 4.1.1 and in Annex B for more information

In the healthcare sector, and particularly in nursing care for the elderly, labour shortages are a structural issue: as the population ages on a massive scale, human resources cannot physically scale productively to ensure that a high level of care is provided to all. In this context, advances in automation and robotics, powered by data and information technology, are critical. Early examples in Japan include Cyberdyne's HAL cyborg-type robot, which aims to assist physically challenged persons to move.[76]

In the agricultural sector, Japan faces a related challenge: with the average age of farmers above 70 years old, productivity at a remarkably low level amongst advanced economy and agricultural land extremely scarce,[77] the sustainability of the sector is in question. Indeed, the entire 'Agriculture, forestry and fisheries' sector represents around 1% of GDP and less than 0.5% of GVA in Japan. Leaving aside strategic considerations around self-sufficiency in food production, it is clear that the agricultural sector can benefit strongly from improvements in productivity and the expansion of opportunity for farmers. As illustrated by the case study developed in Section 3.3.5 (Fujitsu's Akisai cloud service), technology and data-driven innovation can play a role in this transformation, across the whole supply chain (from production to processing to distribution of agricultural products).

The challenges of the agriculture sector translate into economic challenges for rural Japan in general. Besides making agriculture more attractive and more productive, small-scale manufacturing and tourism are also important aspects of the government strategy. Data-driven innovation can have a positive impact on both: productivity in small-scale manufacturing is particularly important as economies of scale are difficult to achieve; in tourism, as shown in the KDDI/Colopl case study in Section 3.3.2, personalised and location-aware marketing can help make touristic experiences relevant and adapted to people's needs.

In areas where Japanese firms are already world leaders, such as manufacturing and transport, data-driven techniques are already common place. In other industries including agriculture, healthcare and tourism, innovative examples are emerging of how data and the data-driven innovation that stems from it can work for the benefit of the Japanese people and economy, in areas of particular challenges.

4.2.3 The wider social value of data-driven innovation

Beyond the economic value discussed so far, data-driven innovation has the potential to unlock social and public benefits that are not quantifiable, such as environmental and health benefits from lower fuel consumption, or successful policy delivery supported by data.

In this section we illustrate the potential for social value through selected examples in education and healthcare, and government policy supported by data-driven innovation. These capture various types of social value, including: environmental benefits, improved service, increased consumer willingness to pay, social equality, community engagement and better use of limited public resources.

76 See http://www.cyberdyne.jp/english/products/HAL/index.html

77 Japan's agricultural land is 12.5% of total land area, or about 46,000km^2, and Japan's average population density is 350 people per km^2; as a result, agricultural land per 1,000 inhabitants is 0.36km^2, compared with 13.02km^2 in the USA, 3.83km^2 in China and 2.08km^2 in Germany. In other words, not only does Japan have 10–30× less agricultural space per capita than the other large economies in the world, but it also has a relatively less productive sector.

Healthcare and education

One of the findings from the economic analysis developed in Section 4 is that the economic value of data-driven innovation for health and education is smaller than in other verticals. However, as we have noted earlier, this underplays the potential significance of the use of data-driven innovation in an educational or healthcare context, where social considerations are paramount. These social benefits include, for example, improved educational and health outcomes, and more efficient and effective use of resources, such as public spending on health and education, which do not translate into value-added gains.

DDI services can be used in the analysis of both education and healthcare operational results,[78] across multiple schools/hospitals and at various institutional levels (i.e. general clinic vs. hospital, school vs. university). This analysis can highlight differences in performance, which is especially useful in addressing social inequality. Within education, data analysis can be used to narrow the achievement gap between different student subgroups. The use of "carefully designed school improvement interventions" designed around data, is considered a robust method of increasing intervention success and reducing educational differences.[79] Over time, this benefit can also lead to improvements in other areas of social equality, such as employment status, workforce participation and wages.

We have explored earlier in this report some of the benefits of DDI services in the healthcare sector, from an economic perspective. Some of the economic benefits, for example in alleviating labour shortages in elderly care, ultimately perform a social function: the decision to have a high and consistent standard of care for all elderly citizens, irrespective of their circumstances and geographic location, is the main driver behind high demand for resources in that sector for example, and is a social rather than economic goal.

Data-driven techniques are being explored across the world as a means to improve the utilisation of scarce resources in healthcare, for example through the development of comprehensive clinical effectiveness research. The objective is partly focused on patients and clinicians (to improve effectiveness and adequacy of treatments) and partly focused on insurers, including in the case of Japan the State itself. On the patient side, such advances can lead to significant gains in quality of life, including reduced time off work, reduced waiting time for operations, and reduced stress. From the perspective of insurers and the State, evidence- and data-based effectiveness research can reduce waste and enable better targeting of scarce funding.

78 This is typically relevant to cases where these results can relate to a range of measured tests. In education, for example, test scores or leaving age could be considered, whereas in healthcare examples include readmission or mortality rates.

79 Kerr, K. and West, M., *Social inequality: can schools narrow the gap?*, British Educational Research Association (2010).

Better evidence-based government

Through the publication of more internal governmental information, citizens are able to analyse the workings of governments, thus increasing transparency. This in turn allows more informed debate among the public and can enhance governments' accountability to their citizens. The popularity of this approach internationally can be seen from the Open Government Partnership, which currently contains 62 countries. These include the USA (where open government has been a key objective of the Obama administration), as well as Indonesia, Australia and New Zealand.[80]

> "*Transparency promotes accountability and provides information for citizens about what their government is doing. Information maintained by the Federal government is a national asset. My Administration will take appropriate action, consistent with law and policy, to disclose information rapidly in forms that the public can readily find and use.*" – Barack Obama, 2009

Internally, the use of data-driven innovation can lead to better option analysis and increased project efficiency, leading to greater overall outcomes for society given the government's limited resources. Data-driven innovation can also support increased governmental policy innovation by highlighting unseen trends and correlations within the overall governmental aims.

4.3 Summary and conclusions

Our research and modelling confirm that Japan has the right technological and economic foundations to benefit greatly from data-driven innovation. Traditionally strong sectors such as manufacturing and transports appear best placed currently and, in the future, increased adoption of DDI services should enable more gains.

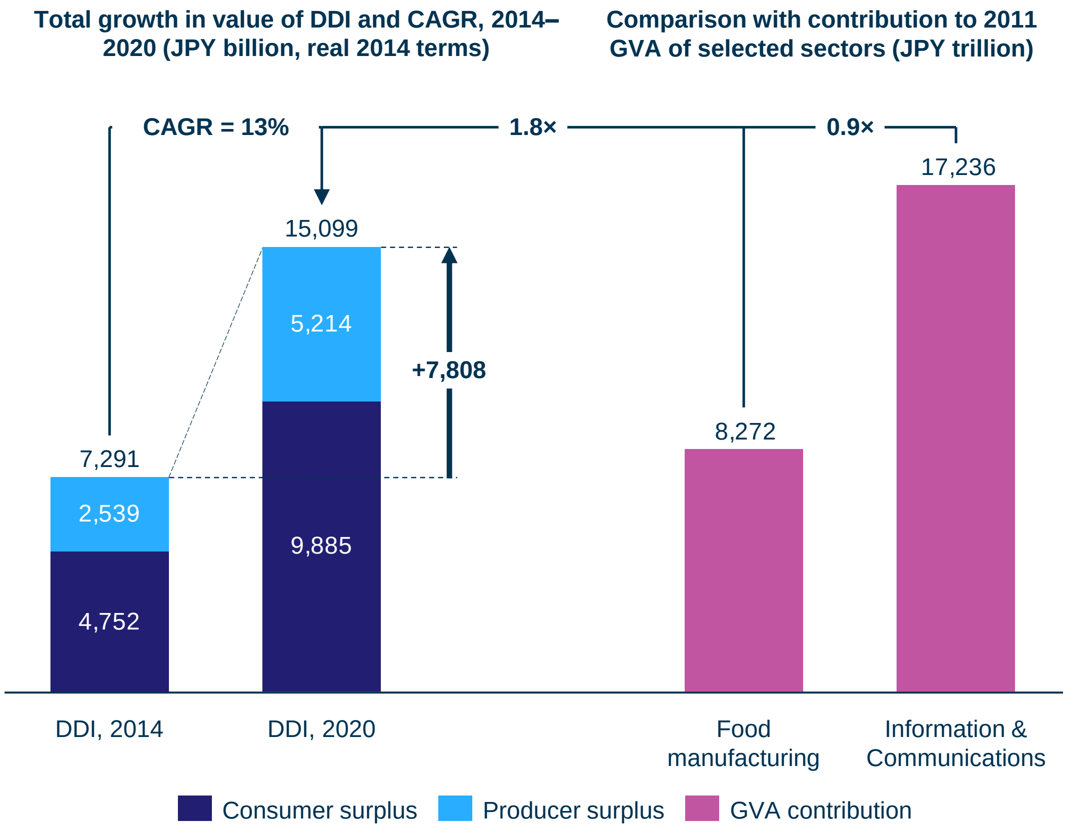
Other sectors such as trade and healthcare can also benefit from data-driven innovation. For trade, which is characterised by a large number of small firms, adoption barriers are higher than for large and more concentrated industries. For healthcare and education, conversely, economic value is by no means the only measure of success, and data-driven innovation has the potential to improve social outcomes. The estimates discussed in this section do not assume that DDI-related public policy would change materially, in either a positive or negative way, between now and 2020. However, policy changes could well have the effect of increasing or decreasing the potential growth in the value of data-driven innovation. In the next section of this report, we discuss how a balanced approach to regulating and managing data can help make Japan a leader in data-driven innovation. Such an approach might well have the effect of enabling the value of data-driven innovation to grow beyond our projections; however, we have not attempted to quantify this.

---

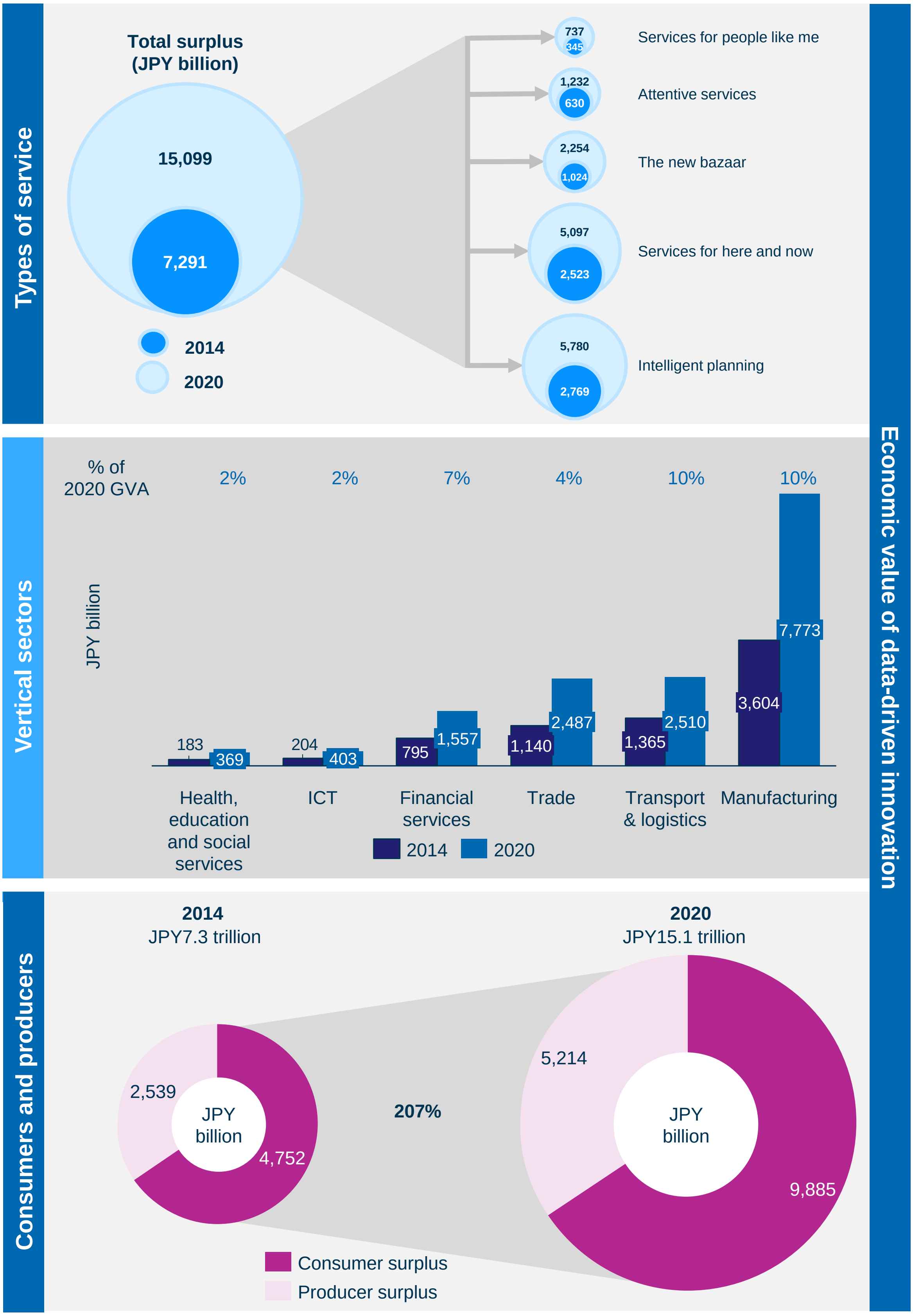
80 See http://www.opengovpartnership.org/.

Our estimates of how much value data-driven innovation represents for the Japan economy in 2014 and in 2020 are provided in the graphic on the following page. In Figure 4.8 below, we show, for illustrative purposes, the value of data-driven innovation compared to the 2011 gross-value added in sectors such as food manufacturing (comparable to DDI now) and the ICT sector (comparable to DDI in 2020).

Figure 4.8: Illustrative comparison of the contribution of DDI to the economy compared to selected sectors [Source: Analysys Mason, 2014]



Notes: GVA = gross value added; Source for GVA contribution: Statistics Bureau of Japan, Business accounts of incorporated enterprises by industry


Types of service
Total surplus (JPY billion)
15,099
7,291
2014
2020
737
345
Services for people like me
1,232
630
Attentive services
2,254
1,024
The new bazaar
5,097
2,523
Services for here and now
5,780
2,769
Intelligent planning
Vertical sectors
% of 2020 GVA
2%
2%
7%
4%
10%
10%
JPY billion
183
369
204
403
795
1,557
1,140
2,487
1,365
2,510
3,604
7,773
Health, education and social services
ICT
Financial services
Trade
Transport & logistics
Manufacturing
2014
2020
Consumers and producers
2014
JPY7.3 trillion
2020
JPY15.1 trillion
2,539
JPY billion
4,752
207%
5,214
JPY billion
9,885
Consumer surplus
Producer surplus
Economic value of data-driven innovation

5 Policy considerations related to data-driven innovation

Motivated by the significant potential for economic value creation of data-driven innovation, policymakers around the world have been taking steps to ensure that data-driven innovation can reach its full potential. As we will see, Japan is no exception to this.

However, achieving this is not straightforward. Indeed, much of governments' action with regard to data is often motivated not by a need to encourage data-driven innovation, but in response to important concerns about risks related to privacy, crime prevention or national security. These concerns can neither be ignored nor neglected.

The challenge thus is to reconcile the tensions between an ambition to foster data-driven innovation, and the need to continue protecting important public interests. But while this need not be impossible, navigating the issues requires careful consideration of the risks that different policies are meant to address and of the ways and extent to which DDI development may suffer if regulatory barriers are too restrictive.

The purpose of this chapter is to take the reader through these issues. We do this by first discussing the kinds of policy that affect DDI development (enabling policies that can help enable data-driven innovation in Section 5.1.1 and data protection and other policies that may act as barriers to it in Section 5.1.2). We also review how these various policies can affect DDI services (Section 5.2), in terms of day-to-day operation, development, growth and international competitiveness (Section 5.3).

In seeking to resolve this apparent tension between data-driven innovation and data protection, it may be constructive to focus on the outcomes sought through existing policies and to consider alternative ways in which these can be secured (we discuss this in Section 5.4). However, this is not always possible: in such cases, difficult choices and trade-offs may be unavoidable, which should be informed by a careful consideration of the benefits at stake (we discuss these cases in Section 5.5).

5.1 What kinds of policy affect DDI development?

Policy action related to data can either enable or potentially constrain the development of data-driven innovation, so that we can thus speak in terms of 'enabling' and 'constraining' policies. In this section we briefly discus the main examples of each type of policy. But before we do, two observations are necessary:

- The economic analysis in Section 4 is not predicated on an explicit assumption of policy change – that is, it does not make any assumptions as to the degree to which policies would be different from what they are today. The 'enabling' policies discussed in this section should be understood as not only permitting the expected growth to take place, but more broadly as setting the stage in which data-driven innovation can flourish, potentially well beyond what our (deliberately conservative) economic analysis has foreseen.

- Our classification of policies into 'enabling' and 'constraining' types is by no means intended to suggest that only 'enabling' policies are desirable while 'constraining' ones are always undesirable. Indeed, the importance for citizens of issues such as privacy protection means that some constraining policies may always be necessary. However, as we will see, policymakers can aim to address the intended outcomes of data protection through alternative policies that minimise potentially negative effects on innovation.

We address the relationship between enabling and constraining policies later in this section. First it is necessary to review the types of policy in question. The main types of policy are listed in Figure 5.1 and discussed in more detail below.

Figure 5.1: Selected enabling and constraining policies [Source: Analysys Mason, 2014]

Enabling policies	Constraining policies
Positive policies: • Direct funding • Ecosystem development • Open data • Skills development • Technical standards	**Activity-specific:** • Collection • Storage • Combination/repurposing • Analysis • Use
Removal of barriers: • Regulatory clarity • Trust-enabling data protection • International regulatory interoperability	**Non activity-specific:** • Inter-company transfers • International transfers • Security

5.1.1 Enabling policies

Governments can rely on a broad range of policies to promote the development of data-driven innovation. These can be broken down into two broad groups: introducing positive policies aimed at helping data-driven innovation from a resources standpoint, and removing regulatory barriers (including absence of clarity in regulations). We discuss these in turn below. As we will then see in Box 5-1 (p.55), the Japanese government is already pursuing many of these policies. Box 5-1 discusses some of these, as well as key initiatives being undertaken by other countries' governments.

Positive policies that help data-driven innovation from a resource standpoint

Key positive policies that governments can pursue include:

- *Direct funding* for innovation, for example through seed capital or competitions open to start-ups, established firms and/or academic groups.
- *Ecosystem development* policies aimed at developing a critical mass of firms and commercial relationships in relevant fields – for example, offering help to relevant firms in innovation 'clusters'.

- *Open data* initiatives, whereby public datasets (ranging, for example, from roadworks updates to companies' tax records to government spending) are made accessible using standard formats and Application Programming Interfaces (APIs) so that firms can develop innovative services that create value from existing data.
- *Skills development* to ensure that firms have a sufficiently large talent pool of experts in relevant fields (including data mining and predictive analytics).
- *Technical standards* for data exchange to facilitate collaboration and trading between firms.

Removing regulatory barriers

The main ways in which governments can remove regulatory barriers to data-driven innovation include the following:

► *Providing a clear and effective regulatory climate*

As with any other type of business, regulatory uncertainty acts as a barrier to investment in data-driven innovation, as investors need to reflect the risk of adverse future regulations when assessing business plans (in addition to the technical or commercial risks inherent in any innovative business). As a result, and in particular, a lack of clarity concerning future regulations on data protection (which we discuss extensively below) may lead to smaller and fewer investments. Thus, ensuring that the scope and key aspects of regulations are clear is a key priority.

Similarly, if consumers are unsure that their data will be treated in accordance with their expectations, they will be less willing to provide it. Accordingly, having clear, enforceable and enforced data protection regimes can help create an environment of trust which is important for a successful DDI sector.

Given this, and in a context in which the Internet has made data processing an increasingly important consideration both for businesses and for consumers, governments around the world have embarked on efforts to update and clarify their data protection policies. Key examples include the European Union's proposed *General Data Protection Regulation*,[81] the White House's *Consumer Privacy Bill of Rights*,[82] and the US National Telecommunications and Information Administration's (NTIA) request-for-comments on *Big Data and Consumer Privacy in the Internet Economy*.[83]

---

81 See http://ec.europa.eu/justice/newsroom/data-protection/news/120125_en.htm.

82 See http://www.whitehouse.gov/sites/default/files/privacy-final.pdf.

83 See http://www.ntia.doc.gov/federal-register-notice/2014/request-comments-big-data-and-consumer-privacy-internet-economy.

► *Enabling international regulatory interoperability*

In certain circumstances, businesses that rely on processing personal data can be prevented from working with overseas partners or end users. This happens as a result of concerns that foreign firms operating in a different regulatory environment may not uphold the principles that are enshrined in law in the country of origin of the data. However, as we will see later in this section, the ability to exchange data internationally is key for domestic DDI service providers who want to offer services globally; for domestic DDI services who wish to rely on best-in-class overseas partners for some of their activities; and for domestic firms in traditional verticals that wish to benefit from the services of an overseas DDI service provider.

National and regional governments have created a variety of arrangements aimed at mitigating these issues. These include:

- 'Safe harbour' arrangements, which allow domestic firms to be bound by the data protection rules that apply in their customers' jurisdictions (e.g. data coming from the EU into the USA is treated in accordance with EU rather than US principles).[84]
- Multilateral certification arrangements, such as APEC's Cross Border Privacy Rules system[85] (CBPR). Companies that apply to and are certified into the scheme undertake to adhere to certain principles irrespective of the country where data is processed.[86] The scheme allows for transfers not only between different countries but also between different firms.
- The European Union's Binding Corporate Rules (BCRs) scheme which, subject to approval by the Data Protection Authorities (DPAs) involved, allows multinational corporations to transfer data in and out of the EU (but within the same corporate group) in compliance with EU law.[87]
- The 'common referential',[88] which seeks to facilitate dual compliance with the BCR and CBPR schemes.
- Certain 'derogations' or exceptions to rules limiting international transfers, based on, for example, user consent or contractual considerations.[89]

---

84 In the case of EU law, overseas rules need not be exactly the same but should offer an adequate level of protection. Article 25 of the EU's General Data Protection Directive (95/46/EC), Adequacy decisions may be limited to a certain type of data (such as financial data or passenger records). See http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML. For information tailored to US firms, see http://www.export.gov/safeharbor/eu/eg_main_018365.asp.

85 See http://www.cbprs.org/.

86 http://www.futureofprivacy.org/2014/05/01/japan-approved-to-participate-in-apecs-cross-border-privacy-rules-system/.

87 See http://ico.org.uk/for_organisations/data_protection/overseas/binding_corporate_rules.

88 See http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp212_en.pdf.

89 For example, Article 26 in the EU's General Data Protection Directive (95/46/EC).

While these schemes all share the key objective of facilitating data transfers by providing cross-border legal arrangements (including accountability and enforceability), they also reflect a variety of degrees of 'interoperability', or lack of 'friction', between jurisdictions. At the end with the least friction, 'safe harbour' arrangements result in the lowest transaction costs for firms, as they incorporate blanket agreements into which firms are allowed to self-certify themselves as compliant. CBPRs require a firm-level self-assessment, which must be certified by an independent third party; once accepted into the scheme different firms in different APEC countries can exchange data. At the more demanding end, schemes such as BCRs not only require firm-by-firm application to Data Protection Authorities (which can easily take a year to process[90]), but limit data transfers to members of the same corporate group (thereby requiring inter-firm transfers to be covered by alternative and potentially more burdensome contractual arrangements).[91]

In the absence of formal arrangements for data exchanges, high-level international guidelines such as those issued by the OECD[92] mean that, when these are reflected in local law, at least a minimum of interoperability exists.

► *Reassessing existing data protection rules*

Finally, governments are reassessing their existing data protection regimes in the light of technological changes and the need to reconcile the demands of privacy with those of innovation. However, as we will see in the rest of this section, this is an extremely difficult endeavour.

Box 5-1: DDI-enabling policies in selected countries

Japan

Japan stated its goal of promoting data-driven innovation in 2013 as part of its *Declaration to be the World's Most Advanced IT Nation*,[93] which aims to form "a society where new industries and services can be created" using IT and data. Launched by the Strategic Headquarters for the Promotion of an Advanced Information and Telecommunications Network Society (IT HQ), which is headed by the Prime Minister, the declaration has been followed up by various government departments; for example, in June 2014, Ministry of Economy, Trade and Industry (METI) announced the creation of a Strategic Council for Data Driven Innovation tasked with encouraging the increased sharing and use of data by firms. It identified as key challenges a lack of relevant

90 According to the UK's Information Commissioner, "a straightforward application could take 12 months to conclude." See http://ico.org.uk/for_organisations/data_protection/overseas/binding_corporate_rules

91 A useful comparison between BCRs and CBPTs is provided at http://www.iispartners.com/downloads/IIS%20CBPR-BCR%20report%20FINAL.pdf

92 See http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm.

93 See Kantei (Prime Minister and his Cabinet) (2014): *Declaration to be the world's most advanced IT nation,* available at http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20140624/siryou1.pdf.

skills and the existence of management principles that lead firms to rely exclusively on internal data.[94] Additionally, in 2014 the IT HQ introduced its *Policy Outline of the Institutional Revision for Utilization of Personal Data*, which is a key step towards providing a regulatory framework in which data-driven innovation can flourish. Key areas of these policies include:

► *Regulatory certainty*

The *Policy Outline* explicitly recognises the existence of a "grey area" of regulatory uncertainty as to the exact definition of personal data and the circumstances under which it can be collected and exchanged. It notes that this, coupled with concerns about potential negative publicity, may prevent firms from using data to its full potential. Accordingly, the government has made it a priority to address these issues in legislation to be introduced.

► *International interoperability*

Japan's *Declaration* explicitly recognises the importance of facilitating international data flows, stating that the government will ensure the standardisation of rules concerning privacy and information security to facilitate "effective international flows of information through international systems", which is deemed essential. It further notes that "international collaboration will be promoted through international negotiations in forums such as the OECD".

► *Open data*

Open data forms an important part of the strategy. As the Declaration states at Section III(1):

> *There are high needs in the private sector, particularly, for data in the possession of government, as highly-reliable basic data. But the data is not used adequately because there are currently restrictions on the secondary use of released data, much data is not machine-readable (possible to analyze and process by software), the existence and location of desired data is difficult to identify, and much of the data that can be used by businesses and others is not released publicly.*

Accordingly, the government's open data strategy aims to allow unrestricted use of public data by third parties.[95] This data should be available on the Internet, meet international standards and be in machine-readable format. Best practice in using public data will be shared, for example through competitions. By international standards Japan scores well, but not at the top, in availability of open data; the Open Data Barometer report ranked Japan as 14th globally, with a relatively high score for readiness, but a more moderate score for implementation and impact.[96]

94 See http://www.meti.go.jp/english/press/2014/0609_02.html.

95 See http://japan.kantei.go.jp/policy/it/20120704/text.pdf.

96 See http://www.opendataresearch.org/dl/odb2013/Open-Data-Barometer-2013-Global-Report.pdf.

The government's priorities seem to be backed by marketplace demand. A recent survey commissioned by the government found evidence of a mismatch between supply and demand of open data: the two categories most demanded by private businesses (map / topographic information as well as transport information) are only the 8th and 9th most commonly offered by local governments.[97]

► *Other aspects*

Other notable aspects of the government's strategy include focus on:

- development of relevant skills, which are seen as a key barrier to data use among firms)
- availability and sharing of data for disaster prevention and relief
- road transportation, with plans to enable real-time exchange of data between roads, vehicles and individuals, such as location data
- cyber security, which can be a key enabler for a safe environment for data-driven innovation. The strategy aims to establish a response mechanism for serious risks, and encourage stakeholders to work together to keep cyberspace secure.

Elsewhere

Initiatives pursued by other governments include the following:

- **UK:** The government's recently published strategy on developing the data economy focuses on educational and professional development, a data-supportive R&D infrastructure, and commercial and academic capability.[98] In addition, the Technology Strategy Board (a government research funding agency) routinely invests in data-related projects,[99] and also provides funding for the Open Data Institute, an initiative founded by Sir Tim Berners-Lee, the inventor of the World Wide Web.

- **USA:** The White House recently published a major report into how the potential of 'big data' can be best leveraged while safeguarding privacy interests.[100] The government's data.gov website is now home to over 91,000 openly available data sets generated by the government, backed by a presidential instruction[101] to government agencies to make data available. Other efforts include the Open Data Initiative which aims to stimulate entrepreneurship, jobs and social welfare by making government and voluntarily contributed corporate data freely available.[102]

---

97 Also, data on disaster prevention ranked third both in terms of demand and supply. See page 16 in *Overview of the 2014 Edition of Information and Communications White Paper*, issued by the Ministry of Internal Affairs and Communications (MIC) (平成２６年版 情報通信白書」の概要 -総 務 省). Henceforth referred to as "MIC white paper overview".

98 See Department for Business, Innovation & Skills (2013): *UK data capability strategy: seizing the data opportunity*, available at https://www.gov.uk/government/publications/uk-data-capability-strategy.

99 See https://connect.innovateuk.org/web/harnessing-large-and-diverse-sources-of-data/funding and https://connect.innovateuk.org/web/data-exploration/overview.

100 See http://www.whitehouse.gov/issues/technology/big-data-review.

101 See http://www.whitehouse.gov/the-press-office/freedom-information-act.

102 See http://www.whitehouse.gov/innovationfellows.

- **Singapore:** Singapore has a wide range of policies aimed at fostering data-driven innovation, including an open data portal with 8,600 data sets for public use and its OneMap service with a one-stop geospatial data platform.[103] Singapore's Infocomm Development Authority (IDA) has embarked on a number of relevant initiatives, which include[104] support for developers needing access to data and the establishment organisations such as the Business Analytics Translational Centre (BATC) and the Business Analytics Centre of Excellence (CoE), in partnership with the private sector.

5.1.2 Constraining policies

The ambition to encourage data-driven innovation development must co-exist with a range of policies, laws and regulations (for simplicity, 'rules') which may restrict certain uses of data for innovative purposes, presenting barriers to data-driven innovation. These rules exist in order to protect important public interests, such as privacy, crime prevention, or national security. None of these can be neglected, nor can data-driven innovation simply be expected to take precedence over them.

Although there are important exceptions, in general these rules revolve around the protection of personal data[105] – that is, they are *personal data protection* rules. But, notably, as data becomes increasingly interlinked, it becomes increasingly difficult to separate personal from non-personal data. Accordingly, personal data protection rules have a broad scope across many types of data.[106]

Figure 5.2 lists the main types of rules, which we have grouped either around the different data processing activities introduced in Section 3 (where relevant), or thematically:

103 See https://www.ida.gov.sg/About-Us/Newsroom/Media-Releases/2013/More-government-Data-to-be-Made-Available-to-Spur-Social-Innovation-and-Create-Greater-Public-Value.

104 Details of recent IDA recent initiatives can be found at http://www.ida.gov.sg/.

105 Specifically, personally identifiable information, or data that can be linked to an individual.

106 In the words of a recent OECD report on DDI, "these developments challenge a regulatory approach that [is based on] the 'personal' nature of the data involved. As the scope of non-personal data is reduced, the difficulty of applying existing frameworks effectively becomes more acute". See OECD (2013), "Exploring Data-Driven Innovation as a New Source of Growth: Mapping the Policy Issues Raised by "Big Data"", *OECD Digital Economy Papers*, No. 222, OECD Publishing.

Figure 5.2: Key types of data protection rules [Source: Analysys Mason, 2013]

Group	Rules	Rationales
Rules linked mainly to specific data processing activities		
Collection	• Firms may be restricted from **collecting data in the background**,[107] i.e. without users actively contributing the data (e.g. through cameras, online tracking, etc.) • Firms may be restricted from **collecting data indiscriminately**, in a way that is not justified by the purpose for which consent is given	• Minimising the amount of data that is collected limits the risk of subsequent harm through misuse or security breaches • Users may prefer to keep information confidential on personal grounds
Storage	• Firms may be required to **delete** data beyond a set period, or after it is no longer needed for its original purpose • Firms may be required to **retain** certain data for at least a minimal period in connection to 'lawful intercept' requirements	• Increased likelihood of security breaches or loss of traceability • Data minimisation / purpose limitation • Law enforcement
Combination and repurposing	• Firms may be prevented from using datasets for **purposes** other than those for which collection was authorised[108] • Firms may be prevented from **combining different datasets** collected for different purposes and /or by different service providers unless explicitly allowed by data subjects	• Possibility of unwanted commercial targeting • Possibility of re-identification (see below)
Analysis	• Special regulations may apply to **pseudonymous** data which, through analytics and combination with other data, may allow individuals to be **re-identified**[109]	• Re-identification may effectively amount to collection of personally identifiable information, with attendant risks and obligations[110]
End use	• Regulations may apply to **automatic decisions** based on algorithmic **profiling** (e.g. on service price, characteristics, availability, etc.) • Rules may limit **unsolicited messaging** (phone calls, email spam, etc.)	• Using sensitive information (e.g. race, sexuality, gender or other factors) to influence decisions may be illegal • Errors may have high personal cost

107 Also referred to as 'passive' data collection.

108 For example, in the USA, the FTC has prevented Facebook from combining datasets created before its acquisition of mobile messaging service WhatsApp with those created by the latter before it was acquired unless it could get "affirmative consent" from end users; see http://www.ftc.gov/news-events/press-releases/2014/04/ftc-notifies-facebook-whatsapp-privacy-obligations-light-proposed.

109 Re-identification can be defined as the process by which formerly anonymous data can be linked to named individuals, thus rendering it personal data. For extensive background, see http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

110 A study of 58,000 volunteers by the University of Cambridge and Microsoft Research was able to show that "easily accessible digital records of behaviour, Facebook Likes, can be used to automatically and accurately predict a range of highly sensitive personal attributes including: sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender." See "Private traits and attributes are predictable from digital records of human behaviour", PNAS 2013 110 (15) 5802–5805.

Group	Rules	Rationales
Rules applying to multiple data processing activities		
Transfer between firms	• **Transfer between firms** (whether for purposes of outsourcing or trade of data) may be disallowed without consent • Firms may be required to **conduct due diligence** of prospective partners • Firms may be required to use 'back-to-back' data protection provisions in **contracts**	• Increased difficulty of subsequent access to their data by consumers, enforcement and redress (especially with long chains of intermediaries) after data has gone through multiple hands • Difficulties in enforcement or in identifying the required restrictions may increase the probability of violation of consent conditions or of substantive conditions
International transfers	• **International transfers** may be disallowed for certain kinds of data • International transfers may be **limited to selected countries** (e.g. under 'safe harbour' arrangements)	• Overseas data protection may be seen as inadequate or incompatible • Extraterritorial enforcement may be seen as administratively challenging • National security concerns • Industrial policy aimed at incentivising domestic data processing
Security	• Users must be notified **(breach notification)** if their data becomes compromised • Service providers that hold sensitive data may need to be **certified** and/or undergo periodic **auditing**	• Illegal or unauthorised use of data • Disclosure of private information

The restrictions outlined can often be overridden through consumer consent. For example, if a service provider needs to store data for a longer period than would normally be allowed, then once it obtains user consent the relevant restrictions may no longer apply.[111] Thus, data protection regulations often represent only the 'default setting' under which service providers must operate *unless* they obtain consent from users to operate in a different way.

It is worth mentioning at this stage that, as we will see in Section 5.3.3 below, consent is not a 'silver bullet'; by requiring an unrealistic degree of engagement from end users, stringent consent regulations may lead to consent being given in a less than meaningful way, or being denied unnecessarily, thereby unduly limiting DDI services' activities.

The rules listed in Figure 5.2 are not comprehensive, and the associated rationales are only posited for discussion and may or may not correspond to the intent behind specific rules in specific jurisdictions. We will return to the question of rationales for data protection rules in Section 5.4.1.

111 However, in some jurisdictions certain requirements e.g. concerning sensitive personal information cannot be waived through consent.

5.2 Impact of data protection on the operation of DDI services

Each type of DDI service relies on different combinations of data processing activities, as discussed in Section 3.2. As a result, they are affected by policies/rules that enable or constrain these activities to varying degrees, as we illustrate in Figure 5.3 below. The economic value of each type of service, estimated in Section 4 and reproduced here, provides an illustration of the value at stake when considering rules and policies that can affect each service type, positively or negatively.

Figure 5.3: Vulnerability of DDI services to key restrictions [Source: Analysys Mason, 2014]

	Regulated activities					***Value in JPY billion***		
	Background collection	**Retention**	**Combination & repurposing**	**Profiling & discrimination**	**Transfers (int'l & domestic)**	**2014**	**2020**	**CAGR 2014-20**
The new bazaar	◑	◑	◕	●	◑	1,024	2,254	17%
Services for here and now	◑	◔	◔	◔	◔	2,523	5,097	15%
Attentive services	●	◕	◔	◕	◔	630	1,232	14%
Services for people like me	◕	◑	◕	◕	◑	345	737	16%
Intelligent planning	◑	◔	◑	○	◔	2,769	5,780	16%

Legend

Activity is not directly relevant to service type (but may be involved in specific ser	○
Activity is not uncommon but is not strictly needed for service type	◔
Activity may facilitate or improve service provision	◑
Service may be difficult to provide without this activity	◕
Activity is typically essential to service provision	●

Based on this, we can say the following:[112]

- The two types of service with highest value, '**intelligent planning**' and '**services for here and now'**, have the lowest regulatory barriers at present. This is mainly because the datasets involved are often pseudonymous, and in general are subject to far lower restrictions than personally identifiable data.
- Services of the '**new bazaar**' and '**attentive services**' types often use behavioural and/or sensitive data, which may be collected in the background, may come from multiple sources, and may be used in applications involving profiling or similar types of analysis.
- '**Services for people like me**' often apply segmentation techniques to large datasets and automatically classify users (and other consumers) into groups, using information from multiple sources. Data from non users may be anonymised.

112 These points are not specific to the Japanese context and are valid in a general sense

As can be seen, a key common theme in DDI services' exposure to regulations includes a reliance on processing non personally-identifiable data, which is generally not subject to the rules that apply to personal data. However, some of the proposals currently being made in connection to concerns about the possibility of de-anonymisation of data might well change this, with potentially significant impact on data-driven innovation. We return to these issues below.

5.3 Strategic impact of data protection on the development of DDI services

In addition to the direct link between data protection and DDI services that we explored above, certain types of restriction can also have a more indirect, but potentially more profound, negative impact on data-driven innovation by constraining the *processes* of innovation and entrepreneurship through which new services can develop and flourish. Specifically, as developed in this section, limits on repurposing can hamstring innovation; restrictions on international transfers can hinder the global competitiveness of local DDI start-ups; and restrictive consent requirements can slow or even prevent DDI services' growth.

5.3.1 Innovation and purpose limitation

Finding hidden patterns in data, and extracting value from these, is a key aspect of innovation with data. Such patterns are often only visible after combining datasets containing different types of data – for example, when IP addresses are combined with geographical locations. It is through this process that unexpected patterns may emerge, or that the existence of such patterns, suspected by innovators, can be confirmed. These patterns can then be exploited by services that create value for their users.

This process often involves the combination of datasets collected for different purposes. Furthermore, and almost unavoidably, the ultimate purpose of the new service being developed (a purpose which may not even be clear during development) could not have been foreseen at the time of data collection and consent. And even if these uses were foreseen before research began, the nature of innovation is such that new patterns with new uses may arise serendipitously during product development.

Purpose limitation rules can thus hinder the combination and repurposing of datasets that is key not only to the functioning of certain DDI services but also to the experimentation involved in their development – and hence to innovation.

5.3.2 Competitiveness and links on international transfers

DDI services are often built by combining datasets from multiple firms and/or combining technical functionality contributed by different firms, often in real time over the Internet. The firms involved may be in different countries – e.g. an IP geo-location service[113] may be provided by a US firm, while an analytics service might be provided by a Japanese-based one. Even when domestic options are available, competitiveness requires the selection of partners with the best price–

[113] Business-to-business services that allow other services to determine a user's geographical location on the basis of an IP address.

performance combination, regardless of their location. For example, the lowest-cost provider of long-term data storage might be in Europe. Partnering with overseas DDI providers can also allow domestic players to leverage large-scale infrastructure quickly, thereby shortening time-to-market.

Importantly, even in cases that do not involve outsourcing the storage of data, relying on such partners may not simply be a matter of 'importing' data from overseas; collaboration often requires deep technical integration (e.g. the use of 'remote procedure calls'), which often implies a two-way sharing of at least some data.

As a result, rules limiting the international transfer of data risk hindering DDI service providers to compete for customers globally and/or to offer a world-class service to their local customers.

5.3.3 Growth and restrictive consent requirements

Consent requirements can vary in terms of the degree of *explicitness* to which users must actively give consent for their data to be collected and/or processed in certain ways (i.e. 'opt in' vs. 'opt out') as well as in the degree of *specificity* of their consent. Thus, for example, if a user has agreed to have his data collected by a supermarket for the purposes of a marketing programme, this may be understood to imply permission to, for example, pass on some of this data to an external analytics partner without having to ask the user explicitly about this (*deemed consent*).

If a given service relies on an activity for which consumer consent is legally required, then the need to obtain consent represents a 'hurdle' that service providers must overcome. If consent needs to be explicit, the hurdle is higher, and if a similar situation applies to multiple activities involved in delivering the service, then users must be required to understand and consent in detail to each relevant activity – an even higher hurdle. All of this means that fewer potential users adopt DDI services, because:

- they disagree with their data being used in certain ways (e.g. international transfers), or
- they lack the time or inclination to engage with a privacy policy or statement.

Both cases may limit take-up, which may render some innovations financially unviable. In the first case (consumers disagreeing with their data being used in certain ways), this may reflect consumers' own decisions as to the appropriate balance between their privacy and the benefits of data-driven innovation, and as such may be deemed to be unproblematic from a policy standpoint (effectively it is a voluntary transaction).

In the second case (consumers lacking the time or inclination to engage with the notice given to them), however, consumers might have been happy to give consent if only this had not required so much of their time. In those cases, consent requirements result in foregone value for both the consumer and the service provider, which is a form of deadweight loss.

5.4 Reconciling DDI development and data protection by focusing on outcomes

In seeking to promote data-driven innovation while also protecting privacy and other public interests, policymakers face a difficult situation. Existing policies can both make DDI services difficult to operate and may hinder their development and growth. At the same time, however, the public interests that these policies aim to protect cannot be neglected.

However, although data protection rules may clash with data-driven innovation in specific cases, the broad policy *aims* of fostering data-driven innovation and protecting other public interests (e.g. privacy) are *not* fundamentally at odds with each other; they are simply about different things. This suggests that, at least in some cases, it may be possible to achieve the desired aims of data protection without unduly hindering data-driven innovation by reconsidering the specific rules and regulations involved.

In this section we explore how such reconciliation may be possible in certain cases. As a starting point, we discuss a conceptual framework that considers the risks and harms of data processing activities.

5.4.1 Risks and harms

Data protection policies seek to prevent certain *harms* to individuals or society, or at least to minimise the *risk* that such harms may materialise. For example, the handling of information such as credit card numbers is protected so as to prevent fraud. More generally, the rationales for rules such as those outlined in Figure 5.2 above can be linked to consumer harms such as fraud, illicit discrimination or the invasion of privacy.

Such linkages implicitly assume a certain 'causal' model describing how things could 'go wrong' if rules are not followed, ultimately causing *harm* to individuals.[114]

This might proceed as follows:

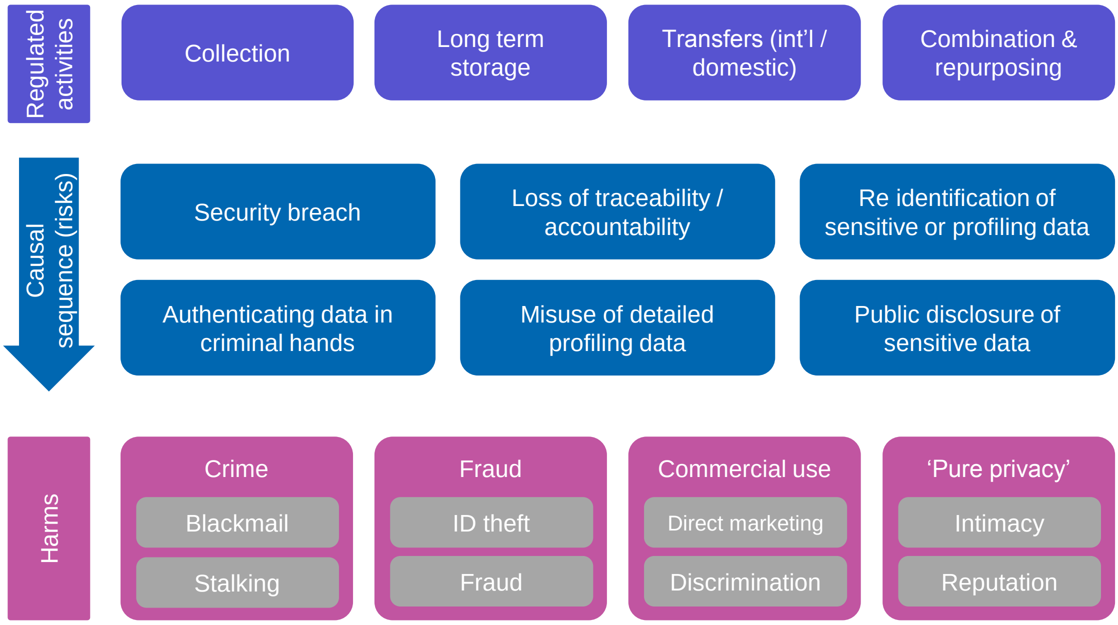
- There might be excessive collection of data, data might be stored for too long, it might be transferred to a third party, or it might be used for purposes not authorised by the user.
- This may then lead to a security breach; a loss of traceability (if data changes hands too many times); or the 're identification' of formerly anonymous but highly sensitive or detailed personal data through analysis of multiple datasets
- In turn, this could in some cases lead to 'authenticating data' like credit card numbers ending in criminal hands; to excessively detailed, non-anonymous data being used by marketers without authorisation; or to sensitive personal data being publicly disclosed.

[114] Our causal account above is by no means authoritative, and policymakers may not assume this type of causal flow. This is not because our account may not fit the rationale behind policies in any given jurisdiction but rather because, in general, the question of harms in data protection is a matter for disagreement and debate. In the words of leading privacy scholar Daniel Solove, "privacy is a concept in disarray". See Solove, Daniel J., "A Taxonomy of Privacy." *University of Pennsylvania Law Review*, Vol. 154, No. 3, p. 477, January 2006; GWU Law School Public Law Research Paper No. 129. Available at SSRN: http://ssrn.com/abstract=667622

- Finally, this could eventually lead to fraud (e.g. ID theft); other crime (e.g. blackmail), commercial misuse (e.g. unauthorised direct marketing); or simply the invasion of privacy (e.g. images of private grief).

This is illustrated in Figure 5.4:

Figure 5.4: Illustrative link between data processing and harms [Source: Analysys Mason, 2014]



Thus, in general,[115] data protection rules can be seen as an attempt to minimize the probability of certain things happening which may carry an increased *risk* of leading to *misuse* of data that would be *harmful* to individuals. Put differently: the regulated activities are not in themselves harmful, but can be the first step in a sequence that has a *risk* of leading to harm.

Based on a similar train of reasoning, in recent years the so-called 'risk-based approach to privacy' has gained prominence among privacy experts as an alternative to an exclusive reliance on rules that restrict data processing practices that may be valuable to data-driven innovation and not in themselves be harmful.[116] As a particular application of this approach, certain authors have proposed that policy should focus less on regulating data *collection* (and the consent requirements involved in it) and more on its *misuse* ("database"), relying on legal accountability and/or technology to minimise the risk of harm.[117]

115 Except for cases where rules explicitly disallow certain uses of data (e.g. rules concerning the use of profiling data for discrimination).

116 See for example: "A Risk-based Approach to Privacy: Improving Effectiveness in Practice". A paper by the Center for Information Policy Leadership (Hunton & Williams LLP), June 2014. Available at http://www.hunton.com/files/upload/Post-Paris_Risk_Paper_June_2014.pdf .

117 For the general approach, see B Wittes: "Databuse: digital privacy and the mosaic", Brookings Institution, 2011, available at http://www.brookings.edu/research/papers/2011/04/01-databuse-wittes. For the technological approach, see C Mundie: "Privacy Pragmatism" *Foreign Affairs*. 12 Feb. 2014 available at: http://www.foreignaffairs.com/articles/140741/craig-mundie/privacy-pragmatism.

We believe the risks-and-harms framework, as described above, is a useful tool when seeking to balance data protection with intent to allow the development of data-driven innovation.

5.4.2 Reconciling DDI and data protection by focusing on outcomes

As alluded to previously, the tension between consumer protection and data-driven innovation does not appear to be intrinsic, and must instead be contingent on the specific rules adopted to achieve a balance between them. On this basis, a key question when considering a give type of data protection rule is whether alternatives can be found that address the same risks to a similar or higher extent, but without conflicting with data-driven innovation.

We believe this is often possible. If we consider the harms that policies aim to prevent (i.e. the *outcomes* sought), and then consider whether the associated risks can be addressed through different sets of policies that are compatible with data-driven innovation, it is often the case that the risks at stake can be addressed in ways that do not conflict with data-driven innovation.

We discuss three key examples below, which we have selected on the basis of their role in the analysis presented earlier in this document: combination, repurposing and the de-anonymisation of datasets; international transfers; and consent requirements. For illustration, in each of these cases we focus on *one* set of risks that existing policies may seek to address; however, we recognise that in each case policies may seek to limit other risks. In principle, a similar analysis may also apply to those other risks; however, we also acknowledge that in some cases this may not be possible, so that 'the best of both worlds' is not achievable and a compromise must be reached. We discuss this last possibility later, in Section 5.5.

Using robust authentication as an alternative to limits on repurposing pseudonymous data

Traditionally, a common solution to the need to observe rules on purpose limitation, while allowing data to be reused in new services, has been to anonymise or pseudonymise[118] data so that it cannot be traced back to the identity of the original data subject.

However, in recent years, some influential research has shown that by analysing and combining large datasets, it is sometimes possible to link formerly anonymous or pseudonymous information to specific individuals, thereby possibly inferring 'sensitive' information that users may not have intended to disclose (concerning e.g. health, lifestyle choices, religion, etc.).[119] While the degree to which this conclusion can be generalised is a matter of debate,[120] this has led some actors to claim

118 Pseudonymous data is data that relates, but is not by itself linkable, to a single, real individual; a key example "user ID" numbers used by websites to detect repeat visitors even if the user's identity is unknown. Anonymised data is not only not linkable to an individual but may relate to one or to multiple individuals.

119 See e.g. Narayanan and Shmatikov: "Robust De-anonymization of Large Sparse Datasets (How to Break Anonymity of the Netflix Prize Dataset)", available at http://arxiv.org/pdf/cs/0610105v2.pdf.

120 See, e.g., A Narayan and EW Felten: "No silver bullet: De-identification still doesn't work", available at https://freedom-to-tinker.com/blog/randomwalker/no-silver-bullet-de-identification-still-doesnt-work/. For a reply, see A Cavoukian and D Castro: "Big Data and Innovation, Setting the Record Straight: De-identification Does Work", available at http://www2.itif.org/2014-big-data-deidentification.pdf.

that "anonymisation no longer works when identities are actively sought",[121] and others to call for pseudonymous data to be accorded similar status to personally identifiable data,[122] in particular subjecting it to purpose limitation constraints.

This would risk hindering much of the combination and repurposing of datasets that is essential to DDI experimentation (see Section 5.3.1), and it could also limit the potential of DDI services that use anonymised information from consumers (who may not be users of the service) to benefit their users or the public in general, such as traffic monitoring apps or public health planning. Notably, this is particularly the case for services of the 'intelligent planning' type, which often rely on the use of extensive datasets of anonymous and pseudonymous data, and which (as we have seen) play a unique role as drivers of both economic and social value.[123]

Concerns about pseudonymous data revolve around the risks involved in any disclosure of private information and may involve concerns about fraud, crime (e.g. blackmail), commercial misuse (e.g. unsolicited marketing) or 'pure' privacy loss (e.g. the disclosure of embarrassing facts). However, among these, concerns about the possibility of fraud loom large. Indeed, a recent survey by the Japanese government found that when respondents were asked to state how sensitive different types of personal information were to them, by far the two highest categories were credit card numbers and account information (with over 70% of respondents assigning the highest level of sensitivity to this type of data).[124] Clearly, the link between personal data and fraud is paramount in Japanese consumers' concerns about data.

To the extent that the objective is to prevent fraud and other similar misuses of data (e.g. identity theft), an alternative to imposing consent requirements on pseudonymous data might be to develop policies aimed at ensuring stronger authentication methods (e.g. through reliable, universally available certificate authorities), which might render the possession of most types of personal data insufficient to commit fraud. This would not only avoid conflicts with data-driven innovation, but might also be more effective at preventing fraud – especially if we consider that determined fraudsters may be able to obtain the relevant information by easier means (e.g. by hacking end users' devices or online accounts).

---

121 EDRi (2012): *An introduction to data protection*. Available at http://www.edri.org/files/paper06_datap.pdf.

122 See for example http://history.edri.org/eudatap-issuesheets#defi.

123 Several commentators have stressed the importance of not limiting data analysis to statistical samples where possible. For example, Savage and Burrows argue that "the sample survey came to enjoy a certain pre-eminence in a situation where the principles of statistical inference had been developed and the technologies for the conduct of surveys invented, and data deriving from routine transactions could not be easily collected, stored and manipulated. This state of affairs existed between about 1950 and 1990, but decreasingly applies." See M Savage and R Burrows "The coming crisis of empirical sociology", *Sociology* 2007 41:885; also See e.g. Mayer-Schönberger and Cukier *ibid*, ch 2.

124 The next category, telephone numbers, was far behind at 56%. MIC white paper overview, p.18.

Using bilateral or multilateral agreements instead of band on international transfers

As we have seen in Section 5.3.2, limits on international transfers (e.g. rules on data localisation) can severely constrain some DDI business models and/or their underlying technologies. Again, it is worth clarifying the desired outcomes of such policies. Depending on the country in question, these may include avoiding consumer harm from inadequate or unenforceable legal protection overseas; national security; and (in some cases) industrial policy.

Of these three rationales for restrictions on international transfers, the first (concern about adequacy of data protection overseas) is arguably the main and most common. But in respect to this concern, solutions that are targeted specifically at ensuring that data is adequately protected overseas (i.e. schemes such as those listed on p.54) may be more appropriate and less costly in terms of impact on data-driven innovation than limitations or bans on international data transfers. Such schemes may explicitly aim at minimising harms (for example, this is the first principle in the APEC's Privacy Framework), and/or or they may involve provisions ensuring that data processing in general is carried out in a way that is compatible with different countries' provisions (e.g. the APEC's CBPR scheme or the EU's safe-harbour provisions).

Alternatives to strict consent requirements

As we have seen in Section 5.3.3 above, onerous consent requirements can be detrimental to the growth of new DDI services. It is thus relevant to ask about the underlying aims of such requirements from the perspective of a risks-and-harms framework.

On the surface, consent requirements give consumers a means to avoid harms such as those outlined in Figure 5.4 by ensuring that potentially risky activities do not take place without their consent. But as we have seen, this is achieved at a potentially high cost to consumers who may be expected to engage in detail with technical matters for which they lack the time or understanding.[125] This can lead to uninformed consent ('choice fatigue',[126] which defeats the purpose of the policy) or to an uninformed refusal to give consent (which may be a loss to everyone).[127]

Thus, explicit and detailed consent requirements may not be a particularly effective tool to prevent privacy-related harms, and it may also be self-defeating. If this is accepted, then the next question is whether other means should be considered. We suggest two approaches:

---

125 Although there are efforts to mandate simplicity in privacy statements, there is an inherent limit to the extent to which this can be done meaningfully.

126 See, for example, Augenblick, Ned and Nicholson, Scott, *Choice Fatigue: The Effect of Making Previous Choices on Decision Making*, available at http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.175.2560&rep=rep1&type=pdf

127 Additionally, there is a growing body of research suggesting that too much choice may be neither beneficial to society nor appreciated by consumers. See, for example, Schwartz, B, *The paradox of choice* (2005), and Botti, S. and Iyengar, S. S., "The Dark Side of Choice: When Choice Impairs Social Welfare", *Journal of public policy and marketing* Vol 25(1), 2006.

- First, greater emphasis may be placed on substantive restrictions on data use and/or processing, which would be set at a level that reflects what users might normally agree to. Provided that services operate within these boundaries, no consent would be needed. For example, if it is agreed that pseudonymous data can be combined, repurposed and shared freely between firms that meet certain criteria, then DDI service providers would not need users' permission for this. (However, as this example illustrates, deciding the precise 'levels' or limits of automatically permitted activity may be a matter of debate.)
- Second, the explicitness requirements for consent may be lowered, or kept low, where explicitness acts as a hindrance to users. In particular, this may mean allowing 'deemed' or 'opt out' consent to apply in more contexts. For example, continued use of a service that offers personalised music recommendations might be deemed to imply that the user has consented for relevant data to be collected and processed for the purposes of providing the service.

As compared to a reliance on strict consent requirements, either of these approaches would achieve a reduction in 'transaction costs' and may also lead to an improvement in privacy. In practice, a solution might involve a mix of both options depending on local considerations.

5.5 Facing trade-offs between innovation, privacy and other policy goals

In the previous section we have shown how, in some cases, it may be possible to address certain key objectives of data protection through alternative routes that may have a relatively small impact on data-driven innovation. However, it is important to acknowledge the limitations of this approach.

In each of the examples we gave, it is also possible that other considerations not discussed by us might argue for the permanence of the restrictions in question. For example, where we considered calls for rules requiring that pseudonymous data be afforded protections comparable to those applying to personal data, we suggested that relevant fraud-related harms might best be addressed through other mechanisms such as enhanced authentication requirements.

But what if the concerns about pseudonymous data have more to do with, e.g., the potential for embarrassing facts to become known (i.e. a 'pure privacy harm') if the de-anonymised data is subsequently compromised? How much importance should be accorded to this concern would depend partly on public opinion (e.g. how concerned would the public be about this?), and partly on the plausibility of the relevant causal link from data processing to harm – i.e. the degree of risk. In particular, if and when it is believed that re-identification is highly unlikely, then the concern might be said to be unjustified. We note that a similar position seems to have been adopted by the European Parliament in its recent adoption of a new General Data Protection Regulation, which includes a qualified presumption that the processing of pseudonymous data should be permissible – see Box 5-2.

Similarly, although we have argued that explicit and detailed consent requirements may not always be in users' best interests, this is not to say that consent mechanisms are undesirable. Consumers may care deeply about controlling their data (as seems to be the case in Japan, where respondents to a recent survey identified consent as the most important condition needed for providing data).[128] To the extent that this is addressable through user-friendly consent policies that are simple and/or rely on 'opt out' mechanisms, this need not have an unduly negative impact on data-driven innovation. However, we cannot rule out the possibility that public opinion may veer towards explicit and detailed consent even if (in the light of our analysis) this isn't necessarily to anyone's advantage – and in a democracy we cannot dismiss this as illegitimate or unimportant. In this case, an informed public debate about the role of consent may be appropriate.

In cases like these, it may be that an impact on data-driven innovation cannot be avoided, and that an adequate balance between the interests of DDI development and other public interests (e.g. privacy or national security) must be struck. Here, we believe that an approach grounded on the specifics of each situation should prevail; the potential downsides or harms (such as impeding DDI development or limiting privacy) should be assessed realistically in terms of the risk of them materialising, and this in turn may depend on the specific types of data, data processing or firms involved. This calls for a flexible approach that avoids all-or-nothing solutions where possible. This can mean high level laws, so that details can be assessed by regulators or sector-specific rules. Alternatively, it can mean laws or regulations calling for firms to carry out case-by-case assessments, which in turn may consist of 'balancing tests' between the conflicting interests and/or a case-specific analysis of potential harms.

A key example of a flexible approach is the 'legitimate interests' doctrine employed in EU data protection (see Box 5-2), which allows firms to conduct a structured 'balancing test' between their own and consumers' interests, thereby allowing a flexible, nuanced approach. We note, however, that recent discussions about updating the 'legitimate interest' provisions have given rise to passionate controversies. To us, this reinforces our view that the trade-offs involved in DDI policy cut across important public interests, and that discussion of these issues calls for an evidence-based and democratic debate.

128 MIC, *ibid*, page 19.

Box 5-2: The 'legitimate interests' doctrine in EU data protection law

Article 7 of the EU's Data Protection Directive (Directive 95/46/EC[129]) deals with conditions under which data may be legitimately processed. These include user consent, contractual reasons and compliance with legal obligations. Condition (f) of the article states that personal data may lawfully be processed if it "is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject which require protection […]". Notably, unlike conditions (a) to (e), condition (f) does not specify grounds under which processing can be considered *a priori* legitimate (e.g. if the user has given explicit consent), but rather calls for a *balancing test* of the interests at stake, including those of the data subject and the data controller (from our perspective, the DDI service provider). Thus the legitimate interest provision allows for a flexible, case-by-case assessment of the interests at stake.

The legitimate interests (LI) doctrine is not without controversy. Its general nature has led to divergent interpretations as to its scope, which has prompted the Article 29 Working Party[130] to issue an Opinion[131] setting out its views on the appropriate scope, providing hypothetical examples of cases where LI would apply (e.g. some cases of simple direct marketing) and others where it would not (e.g. some cases of direct marketing involving extensive background data collection, repurposing and combination).

Reform of LI has been a key point of debate behind the new General Data Protection Regulation (GDPR), during which some voices have argued for LI to have a narrower scope[132] and for users to be given the ability to opt out of cases of data processing that rely on LI.[133] The text approved in draft form in March 2014 specifies that invocation of LI should involve ensuring that the proposed processing is in line with the data subject's reasonable expectations.[134] Notably, it also includes a presumption that the processing of pseudonymous data is within these expectations, subject to a broader consideration of the specific situation.[135] However, it also requires that where processing "permits the controller to attribute pseudonymous data to a specific data subject, the processed data should no longer be considered to be pseudonymous".[136] LI is likely to feature prominently in future discussions as the final text of the GDPR is negotiated between European institutions.

129 See http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:31995L0046.

130 "The Article 29 Data Protection Working Party was set up under the Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. It has advisory status and acts independently." Source: Article 29 Working Party website. See http://ec.europa.eu/justice/data-protection/article-29/index_en.htm.

131 See http://www.cnpd.public.lu/fr/publications/groupe-art29/wp217_en.pdf.

132 See e.g. proposed amendment s 99-100 in the LIBE committee's draft report at http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-%2f%2fEP%2f%2fNONSGML%2bCOMPARL%2bPE-501.927%2b04%2bDOC%2bPDF%2bV0%2f%2fEN.

133 See e.g. http://protectmydata.eu/topics/limitations/.

134 See amendment 100 at:

http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0212+0+DOC+XML+V0//EN.

135 I*bid*, amendment 15.

136 I*bid*, amendment 34.

6 Conclusions: ways forward for policy

Our research and modelling confirm that Japan has the right technological and economic foundations to benefit greatly from data-driven innovation. Traditionally strong sectors such as manufacturing and transports appear best placed currently and, in the future, increased adoption should enable more gains.

Other sectors such as trade and healthcare can also benefit from data-driven innovation. For trade, which is characterised by a large number of small firms, adoption barriers are higher than for large and more concentrated industries. For healthcare and education, conversely, economic value is by no means the only measure of success, and data-driven innovation has potential to improve social outcomes.

Data-driven innovation can create significant value. In the case of Japan, our central estimate suggests that it already contributes JPY7 trillion (equivalent to around 3.4% of gross value added) in the sectors we have considered. By 2020, we estimate the economic potential of data-driven innovation to be in excess of JPY15 trillion, provided the right environment is in place.

The Japanese government has made it clear that it sees data as a valuable resource, which has the potential to support the 'third arrow' of the government's economic policy, an ambitious economic transformation initiated in 2013 and to help solve complex social issues.

A balanced approach to regulating and managing data can help make Japan a leader in data-driven innovation. Proactive, positive policies such as open data are important in encouraging DDI development. They are already being pursued by governments and such efforts should continue. However, DDI development can clash with other policies. In terms of specific services:

- Restrictions on the use of pseudonymous data could have a significant effect on data-driven innovation, particularly affecting the crucial 'intelligent planning' services, with substantial economic and social value at stake.

- Limits on international transfers, combination and repurposing can particularly affect types of service that are small today but which may have most potential for growth – i.e. 'the new bazaar' and 'attentive services'.

Additionally:

- restrictions on combination and repurposing can constrain the process of exploration and unplanned discovery of patterns that is essential to innovation in DDI services
- limitations on international transfers (i.e. data localisation) prevent local DDI service providers from providing a competitive service or achieving sufficient scale
- restrictive consent requirements can limit innovation and adoption, and can also prevent users from enjoying services that they would benefit from.

To address these challenges, where possible policymakers should seek to minimise the tensions between the interests of data protection and those of DDI development. To do so, they can focus on the key outcomes that are sought under each of these two (and other) interests and considering solutions that satisfy both requirements. In particular:

- The need to reduce the potential for malpractice stemming from the combination and repurposing of datasets, and the possible subsequent re-identification of pseudonymous data, may sometimes be best addressable by policy solutions aimed directly at preventing the kinds of misuse and harm that re-identification might lead to; for example, the use of better authentication mechanisms could radically reduce the risk that information could be used in harmful ways.

- The potential harms that may stem from international transfers can be at least mitigated by the introduction of arrangements (e.g. 'safe harbour' agreements or multilateral treaties) that address specific concerns associated with enforceability, consistency of definitions, etc.

- Highly restrictive and explicit consent requirements may not always be the most effective way of pursuing privacy. In particular, introducing rules that reflect a 'default' of what consumers might agree to may make explicit consent less necessary, thereby facilitating DDI development. They may also enhance privacy protection (especially if, in a counterfactual scenario, users are faced with complicated privacy statements and are likely to 'accept' without properly engaging with the terms).

Notwithstanding the above, sometimes simple policy solutions may not be available. In those cases, policymakers need to consider what fundamental trade-offs might be appropriate (e.g. between consumers' demand for confidentiality, the value of data-driven innovation, and technological realities), based on the available evidence and informed by local sensitivities. The choices for policymakers may be difficult, but they need not be made in the dark.

Finally, it is important to keep these concerns in perspective. It is clear that data-driven innovation is a positive development which in general has been able to grow without fundamentally falling foul of existing policies. As with any new technology and its associated economic activity, friction is bound to be created as data-driven innovation continues to develop, but nothing suggests that this cannot be addressed successfully.

Annex A Survey results

A.1 Introduction and summary

Analysys Mason conducted a survey of Japanese businesses to examine their attitude towards the use of the types of innovative data-driven services analysed in this study. The key objective of the survey was to gain insights into the main drivers and barriers associated with the adoption of innovative data-driven services by Japanese businesses of different sizes and industry verticals. These insights complement our secondary research findings by providing anecdotal information from real-world firms in Japan.

The findings of the field research suggest that the majority of Japanese businesses of all sizes and from all sectors of economy do not use innovative data-driven services currently: amongst all the companies contacted to build the sample discussed in this section, only 32% use some form of DDI service, with “services for here and now” the most frequently cited example.

For non-adopters, the main reason behind the decision not to use any such services is the perception that there is no need for such services or that the company would not benefit from them.

The research results presented in the remainder of this section contain a slightly larger (42%) proportion of adopters, so as to understand better some of the drivers for adoption.

A.2 Methodological approach

We commissioned telephone interviews with 100 Japanese companies.[137] Interviewees were mid-level IT executives who we believe have a sufficient visibility over technology usage in their companies.[138] Each interview lasted approximately 20 minutes. We selected to interview companies from the same verticals considered in the main report, which cumulatively represent 71% of gross value added in the Japanese economy.

Figure A.1 shows the segmentation of interviewed companies by company size and industry vertical.

137 The survey was conducted in the Japanese language by native Japanese speakers and managed by Nielsen.

138 Where possible, we tried to talk to an executive reporting to either CIO or CTO in the company.

Figure A.1: Segmentation of interviewed companies by size and vertical [Source: Analysys Mason, 2014]

Number of employees	Health, edu. & social	Finan-cial services	Manufac-turing	Trade	Info. & comms.	Trans-port and logistics	Total
Below 5	1	0	2	1	2	0	**6**
6–20	0	0	0	3	3	3	**9**
21–100	6	6	2	4	3	6	**27**
101–300	2	1	3	10	2	11	**29**
301–3,000	6	3	3	4	0	7	**23**
Over 3,000	1	0	0	1	0	4	**6**
Total	**16**	**10**	**10**	**23**	**10**	**31**	**100**

We designed two different survey instruments: one for companies that use at least one of the innovative data-driven services considered in this study, and one for those that do not use any. This allowed us to ask each group questions most relevant to their motivations for using or not using innovative data-driven services. More specifically, we asked the users detailed questions related to the benefits that the service brings to their businesses and the non-users the questions associated with barriers to service adoption and future plans.

It is important to stress that our sample of businesses is not representative of the entire Japanese economy. Therefore, we were not aiming to obtain statistically significant results that would allow us to extrapolate the findings of the research onto the entire Japanese business universe. As stated in the introduction to this section, the main objective of the field research was to obtain indicative findings and anecdotal information from a diverse set of Japanese businesses that would complement our desk research.

A.3 Key findings

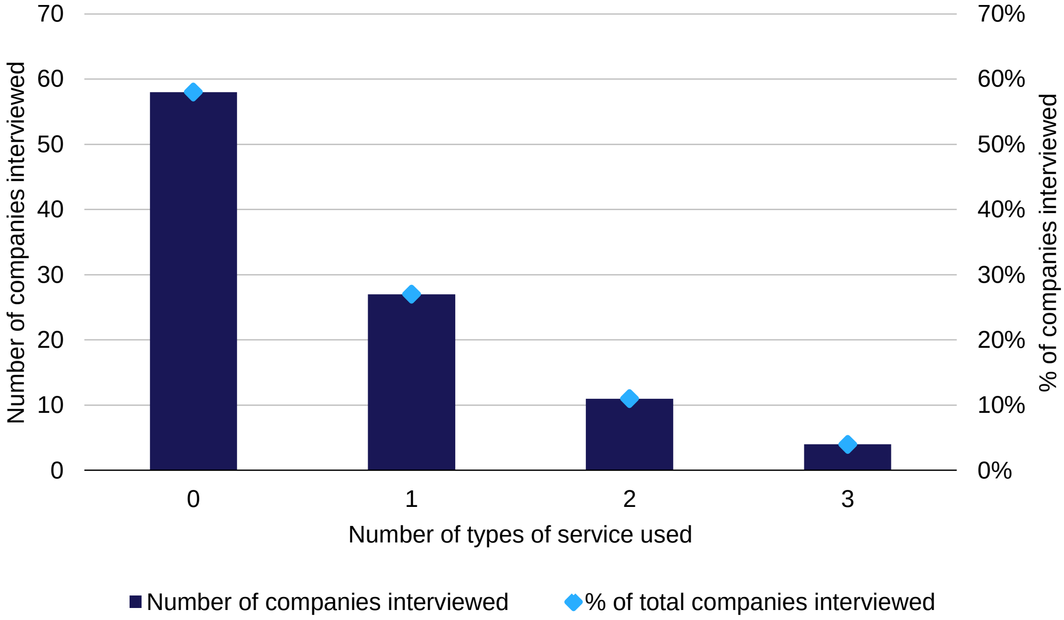
A.3.1 Adoption levels

General adoption levels

We conducted full interviews with 100 companies, but in order to reach at least 40 companies that use innovative data-driven service(s) we had to contact a total of 131 Japanese companies. This suggests an overall adoption level of 32% of companies that use one of the five types of innovative data-driven services we analyse in this study. In the total pool of users, some companies (36%) used more than one type of service. In four cases, companies used three or more types of services.[139]

[139] These respondents were medium-size companies (20–300 employees) in the financial services and transport sectors.

Figure A.2: Interviewed companies by number of services used [Source: Analysys Mason, 2014]



The most frequently used service is the "services for here and now" type (34% of services used). In 75% of cases, the service at stake has been in use for more than three years, which indicates that the users in our sample had enough time to assess the usefulness of the service(s) and understand the way in which the service(s) improve(s) their business.

Adoption level per company size

As stated above, we interviewed companies of different sizes aiming to have a robust representation of SMEs (approximately 80% of total number of interviewees) in our sample. Relative to the number of people interviewed in each group of companies, we found that usage of innovative data-driven services is more prevalent among larger businesses in our sample. Companies with 300 employees and above had an average penetration of 52% of service users, while companies with fewer than 300 employees had an average penetration of users of 34%. This figure is as high as 67% for companies with over 3,000 employees.

Figure A.3 summarises this information.

Number of employees	Number of interviewees	% of companies that use at least one type of service	% of companies that use more than one type of service
Below 5	6	33%	17%
6–20	9	33%	22%
21–100	27	37%	15%
101–300	29	41%	14%
301–3,000	23	48%	13%
Over 3,000	6	67%	17%

Figure A.3: Adoption by company size [Source: Analysys Mason, 2014]

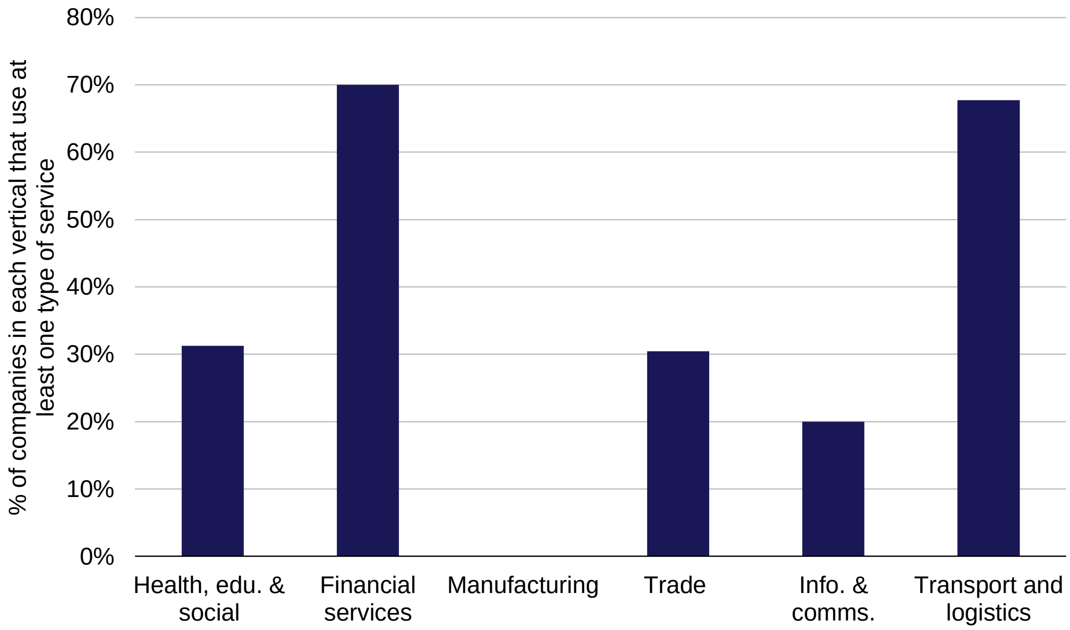
Adoption level per industry vertical

In addition to targeting companies of different sizes, the survey aimed to reach out to companies from different industry verticals, focusing on the six sectors covered in the main report.

Our survey results revealed that, within our sample, most users of innovative data-driven services can be found in the ‘financial services’ vertical, with 7 out of the total of 10 companies interviewed in this vertical claiming to use at least one type of innovative data-driven service considered in this study. This is followed by ‘transport and logistics’ with 68% service users among 31 companies interviewed in this vertical.

Interestingly, while we interviewed the same number of companies in the financial services and manufacturing verticals, the service penetration rates in these two sectors of the economy are on the opposite sides of the spectrum, with very high penetration in the financial services as stated above, and the lowest (0%) penetration in manufacturing. Although this is at odds with the literature review, which suggests manufacturing presents a strong opportunity for DDI, the findings of the primary research suggest a greater degree of variability than in other sectors.[140]

Figure A.4: Adoption by industry vertical [Source: Analysys Mason, 2014]



140 We also refer to the link that we highlighted between company size and adoption; the companies interviewed in the Information and Communications, and the manufacturing sector, were on balance slightly smaller than in other sectors.

A.3.2 Main drivers and barriers to adoption

In the companies that use innovative data-driven services, most interviewees pointed out that the main benefit of the service is associated with the improvement of the quality of the company's core service (42%) as well as market share and revenue increase (31%).

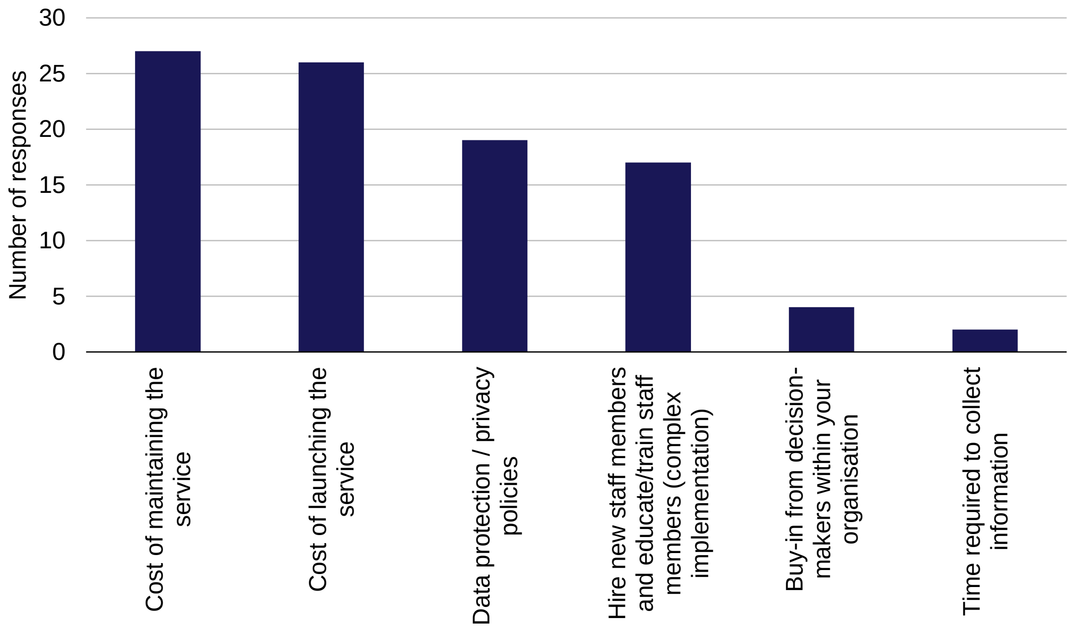
A minority (18%) considered cost savings to be one of the key benefits of the service, which may be related to the fact that most interviewees also saw the usage of innovative data-driven services as quite costly. The cost of launch of the service and the cost of service maintenance were selected as the key barriers to be overcome in order to start using the service in 46% of cases.[141]

On the other hand, 59% of innovative data-driven services that are in use by the companies in our sample have been stated to save employees time; even some of those who did not think that the service contributes to cost-saving thought that it saves time.

In 20% of cases service users pointed out that concerns related to data protection and privacy policies were among the key obstacles they needed to overcome in order to launch the services. In a small number of cases, this obstacle was the only one that needed to be overcome before the service launch.

Figure A.5 shows the number of responses for the various obstacles that need to be overcome for interviewees to start using services.

Figure A.5: Number of responses for the obstacles that need to be overcome for interviewees to start using services [Source: Analysys Mason, 2014][142]



141 The respondents were asked to select any of the 5 barriers to adoption offered or state a different one. A total of 116 selections were made, out of which 53 selections were directly related to the cost of launching or maintaining the service.

142 Please note that in 21 cases users claimed that there was no obstacle that the company needed to overcome in order to launch the service.

We asked the companies that do not use any of the five innovative data-driven services analysed in this report to state the key reasons for not adopting these services. Many non-users stated that they either do not need this type of service or are unsure of the benefits to their business.[143] The cost of the service implementation and its complexity follow as the other two key barriers to adoption as pointed out by the executives from the companies not using the service in our sample (selected as the reason for non-usage in 19% and 17% of cases respectively).

A.3.3 Perspective on future adoption within the sample

Most of the companies from our sample of non-users (76%) said that they were unlikely to use any such service in the foreseeable future. Only 3 out of 58 companies said that they would probably use one such service in the future, with a further 10 being unsure. This is consistent with the key reason for the current non-usage, which is that the need for these services is not clear to the respondent.

This demonstrates a degree of conservatism regarding adoption of DDI services: companies which do not use the service today, because they do not see a need, cannot see a future need either.

Furthermore, the fact that such a large proportion of interviewees said that they would be unlikely to use these services in the future suggests that there may be an information gap, with businesses not fully understanding the range of uses or benefits that can be derived from innovative data-driven services. We note that the government’s IT strategy and its *Declaration to be the World's Most Advanced IT Nation* both seek to address this reticence, particularly in circumstances where material economic gains can be realised from the use of technology and data.

Among current users, only in two cases the respondents stated that a service that is not currently in use is being considered. The motivations that drive this attitude in our sample are quite diverse, varying from the lack of a need for another type of innovative data-driven service (35%) to the high costs of service implementation (18%) and finally to the lack of time to consider and implement another such service (i.e. 15% of current users stated that they have enough ongoing projects).

143 This was selected in 33% of cases as one of the possible answers to the following question: “Why do you not use any of the data or analytics services that we mentioned?”.

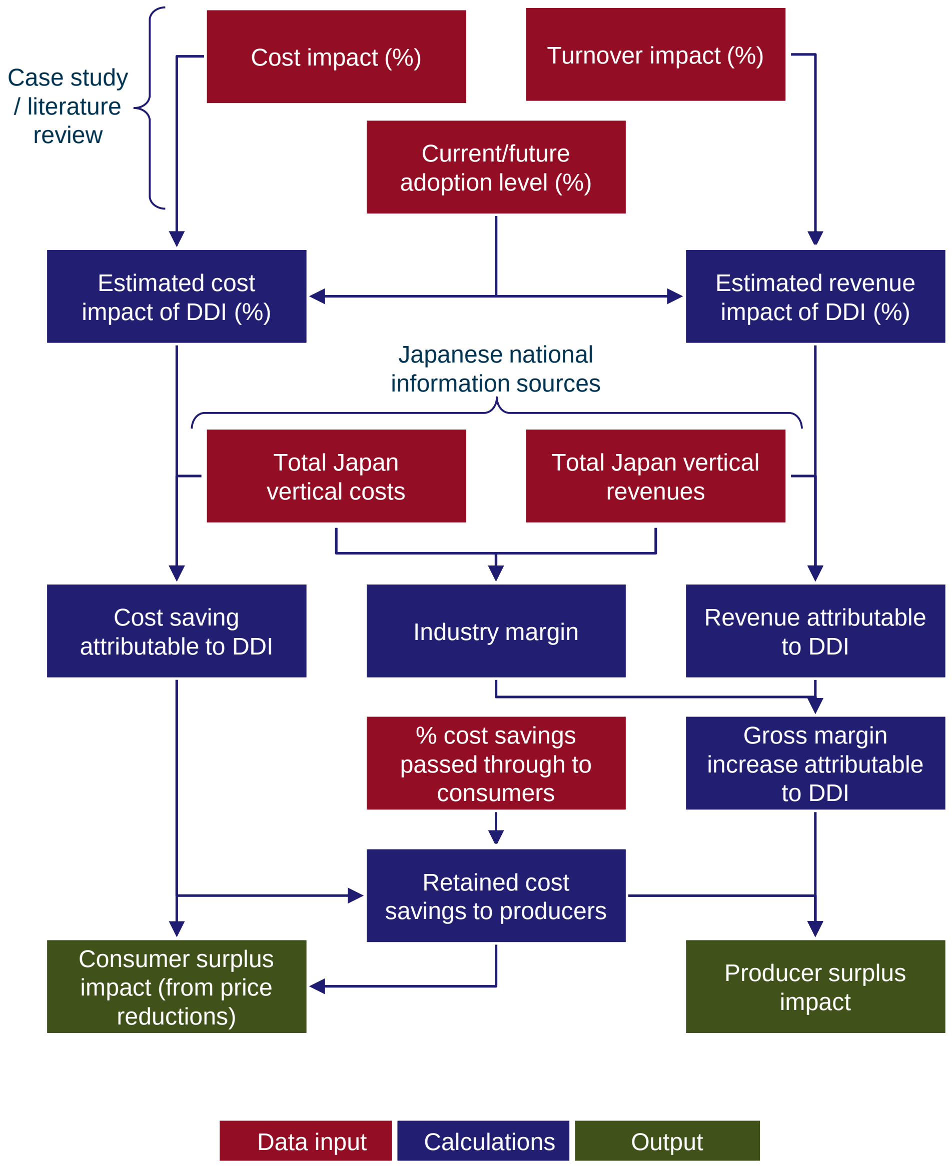
Annex B Summary of approach and methodology

The valuations provided in the main report were arrived at through a model that considers the economic value created by each of our five service types, in each of six industry vertical sectors: 'manufacturing', 'trade', 'transport and logistics', 'financial services', 'information and communications' and 'health, education and social services'.

These vertical sectors are chosen to represent a large part of Japanese gross value added (GVA). The sectors reflect a significant cross-section of uses of DDI in the economy, however we note they are not intended to be fully comprehensive of the areas of the Japanese economy that benefit from DDI services.

Figure B.1 illustrates the methodology for calculating the DDI-related 'producer surplus', and part of 'consumer surplus', of a single service type within a vertical sector. This methodology was used to calculate the benefit of each service type within each vertical sector, thus allowing a calculation of the total benefit to the vertical sector. We note, however, that industry verticals exist outside of the sectors considered within our report (e.g. utilities or construction) and thus the sum of benefits across the vertical sectors we consider shows the majority but not the whole of the value produced by each type of DDI service. Finally, we estimate current (2014) and expected (2020) adoption levels for each DDI service type and each industry vertical. We discuss these steps in more detail below.

Figure B.1: Methodology diagram for calculating the DDI-related economic value of a single service type within a vertical sector [Source: Analysys Mason, 2014]



B.1 Overall calculation of the impact of DDI

The flow chart above translates into the following equation used to quantify the impact of DDI on Japan's economy:

$$Economic\ impact = \sum_{s=1,v=1}^{s=5,v=6} A(s,v) \times [R_i(s,v) \times R(v) \times M(v) + C_i(s,v) \times C(v)]$$

Figure B.2 below defines the variables in the equation, and the sources used. The sources are described in more detail in the remainder of this section.

Figure B.2: Summary of variables and sources used to calculate the impact of DDI on GVA [Source: Analysys Mason, 2014]

Variable	Description	Source	Further details
s	Service type	As defined in the main report	Main report
v	Vertical	As defined in the main report	Main report
A	Rate of adoption (%)	Third party studies, primary research	Section B.2
R_i	Revenue impact (%)	Quantitative international case studies	Section B.3
R	Revenue (JPY)	Japan Statistics Bureau, collated from official ministry data	Section B.4
C_i	Cost impact (%)	Quantitative international case studies	Section B.3
C	Cost (JPY)	Japan Statistics Bureau, collated from official ministry data	Section B.4
M	Profit margin (%)	Calculated as total revenue impact (in JPY), multiplied by the profit margin of the entire vertical (in JPY), i.e.: $M = A \times R_i \times (R - C)$	-

B.2 Rate of adoption

In order to estimate adoption levels for each type of DDI service, we have relied, on a case by case basis, on a combination of:

- A critical assessment of existing international studies in the public domain that assess the real-world impact of data-driven technologies.[144]
- Primary research conducted as part of this study – although not intended to be statistically significant, the results of the primary research inform our estimates.
- Our own judgement and understanding of Japan's economy, informed by secondary research on Japanese-language sources and contact with firms in Japan.

We have followed this approach individually for each combination of DDI service type and industry vertical. Our use of secondary evidence involved:

- considering international-level estimates of adoption of data technologies by different verticals
- considering international-level estimates of adoption of data technologies by different *functional areas* (e.g. finance, human resources, operations) within enterprises
- assessing the international case studies discussed previously and identifying the functional areas within each vertical that are most likely to be affected by each DDI service type
- combining the three points above to arrive at an estimate of the likely impact of each service type within each vertical.

144 Our main sources were (a) M. Schroeck et al: *Analytics: the real-world use of big data*, published by the IBM Institute of Business Value and the Said Business School at Oxford University, 2012, (b) S. Lavalle et al: "*Big data, analytics and the path from insights to value"*, MIT Sloan Management Review, Winter 2011, (c) *How big is big data adoption,* Talend, 2012 and (d) http://www.gartner.com/newsroom/id/2593815.

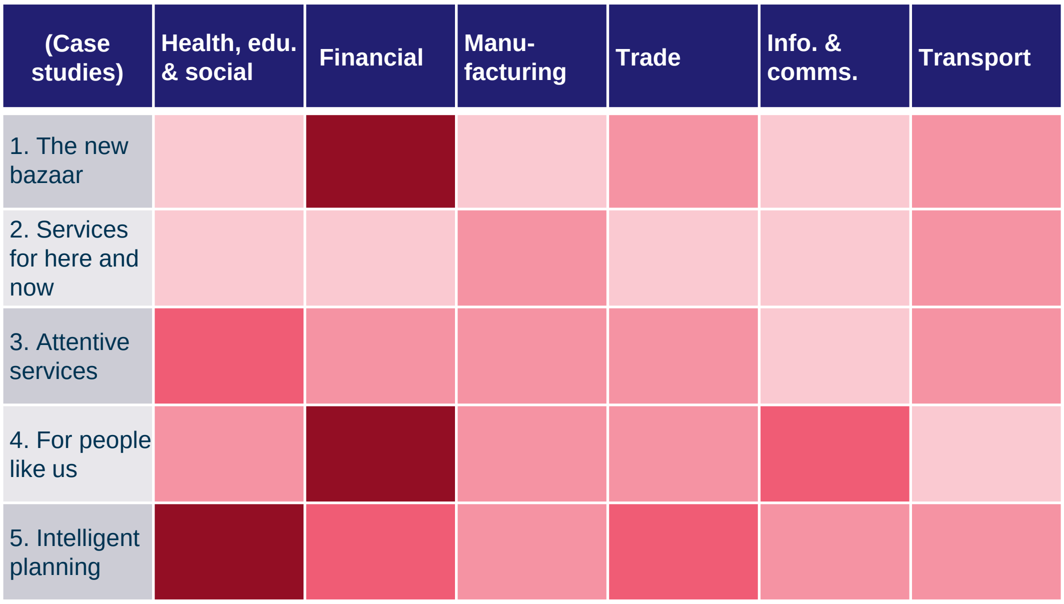
Finally, in order to produce estimates of likely adoption of DDI in Japan in 2020, we have relied mainly on published studies (at a global level) on the level of advancement of different industries' adoption and usage of more advanced data technologies (rather than simply the percentage of businesses adopting any data technologies).

B.3 Cost and turnover impact – inputs and sources

The majority of cost and turnover impact inputs were estimated by critically reviewing a wide range of international case studies and literature around the areas of DDI, data analytics and big data, and then using our own judgement and understanding of the Japanese economy to translate this into likely ranges for the specific case of Japan. Given the differing prevalence of use of the various service types across different vertical sectors, some intersections had relatively more, or less, quantitative case studies than other intersections. In total, 92 quantitative case studies were used.

Figure B.3 shows the relative frequency of case studies within each intersection. This in part illustrates the relative adoption of the service types with the vertical sector, in addition to providing an indication of the degree of confidence that can be given to each intersection (although in practice not all case studies provided equal weight of evidence). The intersections are ranked using four levels, with darker shades denoting a higher number of case studies for a particular intersection.

Figure B.3: Relative frequency of numerical case studies in each service type and sector [Source: Analysys Mason, 2014]



As would be expected, information relating to longer-term process optimisations is frequently more available from case studies – real-time services are often too commercially sensitive for the case studies to include numerical details.

As well as the quantitative international case studies mentioned above, our estimate of cost and turnover impact is supported by over 40 qualitative examples and case studies of DDI activity in Japan. Detailed explanations of selected case studies are included in the main report. The most interesting Japanese examples of DDI, which are used as case studies and examples in Section 3 of the main report, are listed below:

- Sompo Japan's DoRaRoGu
- Navitime
- Jorudan
- Colopl and KDDI
- Koozyt
- Hitachi Transport Management System
- DeNA Life Science
- Imizuno Home System
- SECOM
- IBM Japan
- Komatsu KOMTRAX vehicle management system
- Lawson
- Gurunabi
- Cookpad
- Rakuten
- Yahoo! Japan
- Toyota Big Data Traffic Information Service
- Buzzfinder (NTT Communications)
- Fujitsu
- Metawater.

As noted previously, our analysis only considers the revenue impact from the overall growth in a vertical sector within Japan, rather than the market share gains of a specific company, to avoid including the impact of market share redistribution rather than true incremental growth. This detail, however, is frequently not captured in market reports. International market share abroad has been considered as an increase in the size of the Japanese market and is therefore captured as revenue growth. However, in practice this will ultimately depend on the relative adoption between countries.

Our analysis of revenue considers a net profit impact. Conservatively, we have assumed that the costs of these incremental revenues would be similar to the remainder of a producer's economic activity, in other words that the net margin available would be in line with the average net margin in a given sector. This reduces the net impact of revenue growth significantly.

B.4 Macroeconomic data

Japanese government sources[145] were used to determine the total revenue and costs for each vertical sector, in addition to the relative GVA and GDP contribution of each. Data from government sources has been extracted from the most recent reports available.

For consistency, the latest official data (from 2012) was used, as in many cases more recent industry analysis has yet to be undertaken. Where 2012 data was not available for GVA split by vertical, the 2011 GVA as a proportion of total GVA was applied to the total 2012 GVA. Data was scaled in *nominal* terms to 2014, and then in *real terms* from 2014 to 2020, meaning all the figures shown in our results are in terms of real 2014 prices.

[145] Such as Japan Statistics Bureau, the Ministry of Economy, Trade and Industry, the Ministry of Finance, and the Department of National Accounts.

Further breakdown of revenues of costs was required to apply the impact of DDI as understood from case studies to the correct segment of the vertical. Where official Japanese data (from the Statistics Bureau of Japan or government ministries) was unavailable, further modelling was undertaken with consideration of international benchmarks and industry reports.

B.5 Apportioning of cost savings to producer or consumer surplus

The proportion of industry cost savings that are passed on to consumers was estimated using the margin of each industry, and consideration of the competiveness within Japan (see Figure B.4 below). This is a proxy for the profitability and the competitive intensity in each vertical sector, which are important determinants of the likelihood of cost savings resulting in lower prices.

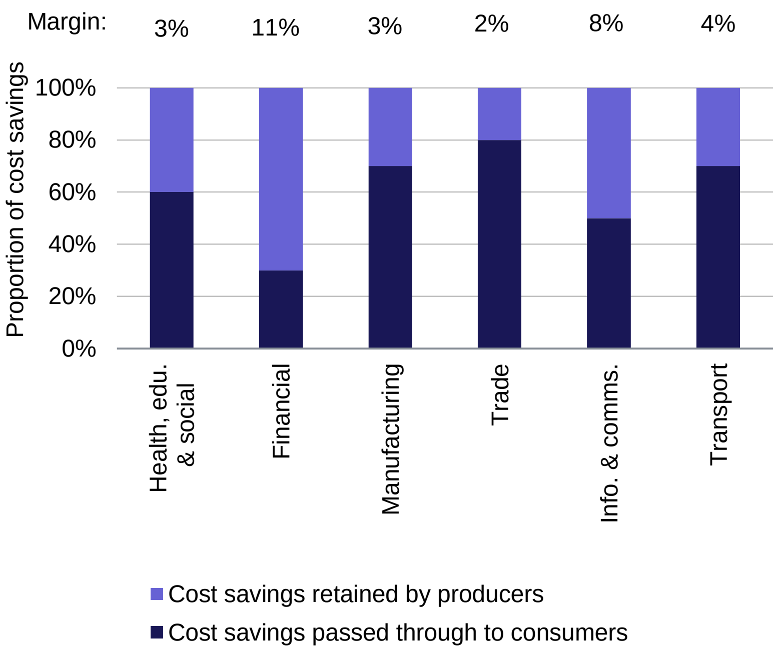


Figure B.4: Comparison of industry sector margins and costs passed on to consumers [Source: Analysys Mason, 2014]