

Report of the UK-French
Data Taskforce

Data Driven Growth

Innovation, Infrastructure,
Skills and Empowerment
in the Digital Age

The UK France Taskforce on Data Innovation was commissioned by Minister George Osborne, Minister Emmanuel Macron and Minister Axelle Lemaire on November 20th to produce a report on the data economy and the opportunities for collaboration between our two countries. The UK and France are widely regarded as world-leading in the area of open data, have pioneered the digitisation of public services and are committed to producing an environment in which data driven innovation can flourish.

Henri Verdier, the Chief Data Officer of the French government and Sir Nigel Shadbolt, Principal of Jesus College, Oxford, Professor of Computer Science at Oxford and Open Data Institute co-founder, were appointed co-chairs. The Taskforce was commissioned to focus on the following questions:

- The Economic, social and environmental opportunities that arise from the use of different data types;
- The role of technology in data innovation and the creation of favorable environments to promote such innovation for the benefits of our citizens;
- The tension between the use of data and increasing fears of the misuse of data assets.

In their work, the Taskforce spoke to governments and international organisations, large companies and SMEs, individual entrepreneurs and startups. The recommendations take into account a wide range of attitudes many of which continue to evolve, around how we use and maintain data. The recommendations are the unanimous view of all members of the Taskforce. They seek to balance benefits and opportunities against legitimate concerns in the rapidly developing world of data innovation.

The Taskforce has developed recommendations that aim to:

- Create the conditions for a thriving data economy ecosystem, in particular building trust in citizens and businesses around data exploitation;
- Strengthen collaboration between UK and French public bodies on open data, and improving the delivery, quality and nature of public services through data;
- Promote data and data science skills in the public and private sector;
- Encourage businesses (especially SMEs) to exploit public and private data in order to strengthen their competitiveness;
- Build our understanding of data innovation models and the economic benefits of data; and
- Support the development of data technology providers in France and the UK.

The co-chairs selected a small team of experts from each country to drive the work of the Taskforce. We are grateful to:

- Patrick Cocquet - Cap Digital;
- Alban Schmutz - OVH;
- Mokrane Bouzeghoub - Allistene;
- Rand Hindi - Snips;
- Yann Lechelle - Snips;
- Clive Humby - DunnHumby;
- Doug Monro - Adzuna;
- Andrew Lawson - Salesforce;
- Stéphanie Finck - Salesforce.

The Taskforce undertook its investigation and developed its recommendations through in-person meetings and discussions via email, as well as workshops and meetings with tech organisations, incubators, SMEs and businesses and government departments in both the UK and France. An extraordinary amount of expertise, information and examples of successes and failures in both countries, as well as in Europe and internationally, sit behind this succinct report.

The launch of the Taskforce's report and its recommendations represents the beginning of an essential and continuing conversation. Now work must begin in earnest designing a roadmap to make these recommendations a reality, to test them and enhance where needed, and deepen the collaboration between governments, tech communities, businesses and consumer advocates on both sides of the Channel.

Nigel Shadbolt & Henri Verdier

UK-FRENCH DATA TASKFORCE ON DATA DRIVEN GROWTH

In this report we highlight four major areas where our two countries can build on their existing expertise and work together to deliver economic growth, achieve efficiencies, increase productivity and enhance social welfare. The first of these areas is **fostering greater innovation with data**. The second focuses on **building robust data infrastructure** to support services and opportunities that benefit everyone. The third concerns **the urgent need to improve data skills and awareness in our societies**. The fourth attends to **the vital question of trust**; trust in the data itself, the way it is used and how its use can be made accountable and proportionate. Today, data touches every aspect of society and our economies. Its benefits should be distributed equitably and fairly.

1

Fostering greater innovation with data

Major innovations are born in ecosystems that are rich in creativity, collaboration, interaction and fair competition. Together, France and the UK could harness their existing data organisations and initiatives to build one of the most creative data ecosystems in the world.

2

Building robust data infrastructure

The strategic importance of data as essential infrastructure for our economies is not widely understood. Just as we invest in roads, rail and other public infrastructure to grow businesses, foster new opportunities and improve lives, we need to invest in data infrastructure.

3

Improve data skills and awareness in our societies

We are facing a data skills shortage. Across Europe we are seeing a rising number of data science job vacancies. We need to strengthen our data skills across sectors already using data in various forms - from health to policing, retail to leisure, transport to energy, the list goes on. Opportunities are being lost every day because companies or administrations aren't fully aware of the value of their data. The USA and China are already training thousands of data scientists with revolutionary educational content. Europe must respond forcefully to this new challenge.

4

Attends to the vital question of trust

We have to think about the society we want. There are concerns around the misuse of data to oppress and control individuals. There are concerns about organisations who have exclusive control of data, and can use it for monopolistic purposes. There are concerns about what is happening with our personal data. A collapse in trust of organisations collecting and using data, including our personal data, is a real risk that could prevent us realising the full economic opportunities on offer.

TABLE OF CONTENTS

INTRODUCTION	2
FIRST PART : THE DATA REVOLUTION.....	5
SECOND PART : RECOMMENDATIONS.....	9
1. INNOVATION	10
2. DATA INFRASTRUCTURE	11
3. SKILLS	12
4. TRUST AND EMPOWERMENT	13
THIRD PART : THE DATA LANDSCAPE.....	14
1. THE DATA VALUE CHAIN	14
Building innovative businesses	15
Creating data we can rely on	16
Using data to make decisions	17
Realising wider economic benefits	18
2. RECOMMENDATIONS	19
Innovation.....	19
Data infrastructure	21
Skills	23
Trust and empowerment.....	25
ACKNOWLEDGEMENTS.....	27

1

The Data Revolution

THE DATA REVOLUTION

We are living through and experiencing a "data-revolution". More and more processes are digitised, completely new digital products and services are arising, all of which leads to the creation of exponentially increasing amounts of digital data. This data is of growing importance to businesses and consumers, governments and citizens. It is becoming cheaper to process, there is constant innovation around the tools for data processing and analysis, we observe the emergence of new data usage policies and an inexorable rise in the demand for people with data skills. Data is now the cutting edge of the digital revolution. As in previous industrial revolutions the data revolution is changing our economy, our society and our politics.

Throughout history, the greatest disruptions have occurred when what was once a scarce resource became an abundance. The transition from hunter gathering to an agrarian lifestyle; the invention of the printing press; the rise of new manufacturing processes and the industrial revolution; and the World Wide Web giving rise to our networked world. In each case what was a scarce resource became an abundance: more food, more literacy, more mass produced products, more data.

These "revolutions" of abundance have invariably seen the emergence of new structures in society, new forms of governance, new sources of wealth, new opportunities and, we should note, new inequalities. This has happened in the past and it is happening now in the data revolution.

Data is sometimes called the "new oil". But this analogy doesn't capture the millennial significance of data and what makes it different from earlier revolutions of abundance. Unlike oil, our data resources aren't finite. Data doesn't diminish over time - as more services and populations go digital, our data resources grow exponentially. It is a non-rival good: consumption by one consumer does not prevent simultaneous consumption by other consumers. It can be produced and maintained by a range of organisations, using a range of tools, and its value increases the more it is used.

There have been many attempts to quantify the value of this profusion of data. In 2015, the European Commission estimated the direct market size of the European open data economy would grow to £325 billion between 2016-2020¹. One 2014 McKinsey report estimated that 'big' data could contribute \$3 trillion USD every year in just seven sectors². It has proven hard to quantify the value of data, because there isn't great value in data itself - the value is in how we use it. The value we stand to gain is from the services built using data, the new products and ideas that give rise to new business models and industries to improve our lives.

Throughout this report, the Taskforce references many kinds of data, and many kinds of data use. The data we collect, use and share in our societies exists on a spectrum³. It might be '**open data**' - data anyone can access, use and share. It might be **shared** between government departments, or between organisations, and licensed for a fee. And it might be **closed** and kept secure, accessible only to a few. It could be 'Big Data' that is openly licensed, or 'Big Data' that is shared. The data economies of the UK and France reflect the diversity, and complexity, of this data spectrum.

The UK and France have proven credentials in supporting activities to make better use of data across a broad range of sectors.

In the UK, successive governments have invested in organisations and initiatives boosting the UK's digital and data economies. Notable examples include the Open Data Institute (ODI), Tech City UK, the Digital Catapult, Tech North and the Alan Turing Institute. 43% of Europe's tech unicorns (private companies with more than \$1 billion valuations) are headquartered in the UK. Data management and data analytics companies make up 12% of digital businesses in the UK and are rapidly growing⁴. And for the third consecutive time, the UK is ranked number one on the Global Open Data Barometer⁵.

In France, a similar dynamic has been implemented that seeks to develop both supply and demand in the data economy. The "Nouvelle France Industrielle (New face of industry)" initiative was launched in 2013. It created opportunities to highlight and reinforce existing initiatives (French Tech, Cap Digital cluster, Mes Infos), launch new ones (including a call for R&D projects, big data challenges) and bring stakeholders closer together across the ecosystem.

In 2015, 61% of French companies recognised that Big Data has become a key driver of growth, as important to them as their existing products and services⁶. The data ecosystem is growing fast and in 2018 the French big data market is expected to reach 652 million euros - an increase of 129% compared to the market in 2014, despite the difficult European economic climate⁷. In 2014, France appointed its first State 'Chief Data Officer', and in 2015 France was ranked number two on the Global Open Data Barometer.

1 <http://www.europeandataportal.eu/en/content/creating-value-through-open-data>.

2 <https://www.ced.org/blog/entry/big-datas-economic-impact>.

3 <https://theodi.org/data-spectrum>.

4 <http://www.techcityuk.com/investors/>.

5 <http://opendatabarometer.org/>.

6 Markess, 2015.

7 IDC, 2015.

This report comes at an important time. Both the UK and France are looking to strengthen their data foundations and to drive greater impact from use of data - in government, across industry, and in the third sector. In the European Union, the General Data Protection Regulation (GDPR) will require significant changes from many organisations that collect, store and process personal data and we need to be ready for that. Whether the UK will need to implement GDPR is unclear, but the Taskforce suggests that both the UK and other European countries would benefit from adopting aligned regulations which permit close interaction, trade and collaboration. The populations of the UK and France are increasingly mobile, and they expect the services they need to be mobile too. And we want to make best use of data, but need the right skills in our workforces to do that.

The report represents an opportunity for the UK and France to deepen connections, and to open up new ways of collaborating and sharing experience. Data can be a common language and common resource. The recommendations in this report bring together the best of both countries' experiences with data.

One such example is open data which has been a priority for both of our governments. Open data promotes transparency and accountability, increased efficiency, innovation and the creation of economic and social value. In France, address data is now available as open data. Address data is part of our everyday lives: they pinpoint where we live and connect us to the services we rely on - from government and from businesses. In France today, anyone can access, interrogate and build services using this data at no cost. What started as a piece of French postal infrastructure has become available everywhere via the web as a part of its essential data infrastructure. In the UK, the Government has committed £5 million in its 2016 Budget to exploring an open address database for the nation.

In this report we highlight four major areas where our two countries can build on their existing expertise and work together to deliver economic growth, achieve efficiencies, increase productivity and enhance social welfare. The first of these areas is fostering greater innovation with data. The second focuses on building robust data infrastructure to support services and opportunities that benefit everyone. The third concerns the urgent need to improve data skills and awareness in our societies. The fourth attends to the vital question of trust; trust in the data itself, the way it is used and how its use can be made accountable and proportionate. Today, data touches every aspect of society and our economies. Its benefits should be distributed equitably and fairly.

To realise the benefits of data we need to foster data innovation. Major innovations are born in ecosystems that are rich in creativity, collaboration, interaction and fair competition. Together, France and UK could harness their existing data organisations and initiatives to build one of the most creative data ecosystems in the world.

This age of data abundance will see transformational outcomes in sectors such as healthcare, transport, finance, food and agriculture, energy and environment, education, security, crime and justice.

In the UK, open data was used to reduce hospital acquired infections by 85% in two years. In Government, the Department for Work & Pensions and the Department for Energy & Climate Change have worked with the Scottish Government and 11 major energy suppliers over recent years to use data to help tackle fuel poverty - using information to provide rebates on energy bills for up to 1.4 million vulnerable households. Data made openly available across public transport has also led to apps that transform our experience of getting around London.

In France, data on electricity consumption has enabled state procurement to save hundreds of millions of euros. In the UK and France we have successful companies making money from data analytics in areas such as predicting retail behaviour, personalising travel options, and drastically reducing the lead time for bringing new products to market. We want to develop a collaborative programme that will seek out and encourage innovative data driven solutions that will save money and make money.

We need to understand data as crucial infrastructure for our emerging economies. The strategic importance of data as essential infrastructure for our economies is not widely understood. Just as we invest in roads, rail and other public infrastructure to grow businesses, foster new opportunities and improve lives, we need to invest in data infrastructure.

Two features of data demonstrate the profound difference between this revolution and previous ones. In the 19th century the widespread extraction and use of raw materials drove the industrial revolution; coal, oil, iron and steel. In the 20th century the widespread distribution of energy via power grids helped drive a technological revolution. In both cases these were "rivalrous" goods; if you used a deposit of fossil fuel someone else could not, and once used it was not available to anyone else. Once used its value was exhausted.

Data is fundamentally different. Potentially your use of data does not deny its use for anyone else; it is not destroyed in the act of use; and in fact repeated use increases its value and can lead to new forms of the resource being generated (e.g. metadata) that make it more valuable. This is why data infrastructure should be as open and accessible as possible.

We need a better basic understanding of data, its opportunities and limitations, across our societies - as well as increased data science skills.

We are facing a data skills shortage. Across Europe we are seeing a rising number of data science job vacancies. We need to strengthen our data skills across sectors already using data in various forms - from health to policing, retail to leisure, transport to energy, the list goes on. Opportunities are being lost every day because companies or administrations aren't fully aware of the value of their data. The USA and China are already training thousands of data scientists with innovative educational content. Europe must respond forcefully to this new challenge.

Connecting data practitioners and scientists to employers is going to be essential in order to benefit from this abundance of data. But, as important is helping leaders of industry and government recognise where the real challenges and opportunities of data lie, and to provide them with the skills they need to commission, work with and support data science activities. Without leaders who can articulate and plan for data transformation, identifying opportunities and planning for risks, we will not exploit the latent potential in the data revolution,

In addition to educating our leaders, building the skills needed to appreciate, exploit and use data must be at the heart of any economic strategy. And we need basic data skills and literacy, basic 'data awareness' to be taught in schools alongside digital civics, cyber safety and coding.

We need to build trust and accountability into everything we do with data. We have to think about the society we want. There are concerns around the misuse of data to oppress and control individuals. There are concerns about organisations who have exclusive control of data, and can use it for monopolistic purposes. There are concerns about what is happening with our personal data. **A collapse in trust of organisations collecting and using data, including our personal data, is a real risk that could prevent us realising the full economic and societal opportunities on offer.**

Both the UK and France have initiated world leading policies and programmes placing citizens in greater control over how their data is collected and used - these should be reflected on and consolidated to inform best practice. Secondly, we need to place transparency at the heart of how organisations collect and manage data about us. The changes ahead that will result from the General Data Protection Regulation⁸ make this a priority. We need codes of conduct and examples of best practice for organisations to help them comply with the GDPR.

We need to make it easier for people to access services they need digitally, across borders, while preserving their privacy. And we need to think ahead to emerging issues that will shape how our personal data is used to make decisions about us - such as greater transparency on the use of algorithms. This already happens in some sectors such as credit scoring, where agencies are obliged to give accounts on how they make credit scoring decisions.

This is the context within which the Taskforce undertook its work. We have identified important areas where the UK and France can collaborate and work together to build on our significant capabilities in the data arena. We believe that our recommendations will benefit not just our two countries, but could also serve to inform developments in the wider European context and further afield. The data revolution is not confined to a particular jurisdiction or geography. If we seize these opportunities we will not only benefit our own economies and societies, others may choose to follow.

8 http://ec.europa.eu/justice/data-protection/reform/index_en.htm.

2

Recommendations

RECOMMENDATIONS

In what follows we provide a summary of our recommendations, these are focused around the major data challenges of innovating with data, establishing effective data infrastructures, securing a supply of data skills and ensuring that the data ecosystem is one we trust and one that empowers us all. A more detailed description of the recommendations is presented later in the report. A plan to implement the recommendations will be presented in the Autumn following discussions between our two Governments. Going forward, the UK and France should establish an annual Data Summit to showcase successes, share experiences and monitor progress implementing these recommendations.



1. INNOVATION

Objective: Encourage collaboration and focus our efforts.

Recommendation : Establish a UK-France Data Challenge Programme to bring UK and French organisations together using data to solve common problems.

Commission a new joint programme featuring two high-impact and high-return challenges each year, with a stated aim of demonstrating clear and measurable improvements. Etalab, Cap Digital, the Open Data Institute and the Digital Catapult may be partners for this work. For example, goals might include economic gains, improvements in efficiency, reduction of friction, and new innovative opportunities in areas that could include jobs, healthcare, transport, finance, food and agriculture, energy and environment, education, crime and justice.

Objective: Scale up our networks.

Recommendation : Create a 'data innovators passport' scheme, to share best practices and extend startup networks between the UK and France.

Build on, provide support for and connect existing startup programmes between the UK and France, through an accelerator passport programme. Startups that are part of the programme would be able to move between incubators including the Open Data Institute, Digital Catapult, Cap Digital, Etalab, Numa Digital, Tech City UK and Scale Up! Startups will be able to learn about respective French and UK ecosystems and markets in order to facilitate their growth.

Objective: Connect data innovators and innovation across regions.

Recommendation : Establish a data twinning programme - cities in regional areas with common industries and/or demographics, and strong startup communities.

Data innovators aren't only in London and Paris. Regional business clusters and incubators are supporting UK and French firms outside the capital cities - like French Tech in Lille, French Tech Culture Avignon-Provence, TechNorth in Manchester and the Leeds Data Mill. The French Tech mission and Tech City UK could be the initiators of this new regional collaboration.



2. DATA INFRASTRUCTURE

Objective : Make using and exchanging data easier.

Recommendation : Develop the use of API's in the data economy, learning from the UK Open Banking experience.

APIs are the nervous system of the 21st century. To that extent, we need to encourage and collaborate on common open APIs in sectors of shared importance. As commerce moves across country borders, how we exchange and access data must be fluid too. The UK has already begun work on an Open Banking Standard for the consumer banking sector. This work should be expanded on.

Objective : Support the basic infrastructure for linked services.

Recommendation : Invest in and share experiences building core data registers, learning from the French National Address Database experience.

Registers provide an authoritative source of core reference data for linking and improving services. France collaborated with the OpenStreetMap community to develop the National Address Database - a commons based on data crowdsourcing, which is freely accessible and usable by anyone. The UK and France should collaborate and share lessons learned from that experience, as the UK explores building its own open address database.

Objective : Make it easier for citizens to get the services they need anywhere.

Recommendation : Continue to invest in and share experiences of e-identification schemes to enable smoother access to public & private sector services in each country.

France and the UK have been working on strengthening pervasive government mechanisms, like e-identification, that stretch across every government service. Privacy should be integrated in the public system by design, building on lessons learnt from projects like France Connect (France) and Verify (UK).



3. SKILLS

Objective : Support senior leaders understanding the power of data.

Recommendation : Support data literacy programmes for leaders inside and outside government.

To realise large scale data transformation, we need to engage at the very top. Training could be developed for 'high potential managers' (business leaders, politicians, government officials, etc.) to inspire and equip them with the data awareness, skills and strategies they will need to transform their organisations. These leaders can come from public and private sector organisations.

Objective : Overcome a chronic data skills shortage.

Recommendation : Embed new data and problem solving skills for existing analytical professions; and create common data apprenticeship schemes and conversion courses to provide a sufficient workforce of data practitioners in the UK and France.

Existing data related disciplines, including management, marketing, engineering sciences, economics and statistics would benefit from curricula that combines basic data training, problem solving and collaborative coding, data ethics and data visualisation. Alongside this, create new entry paths for aspiring data practitioners. There are initiatives in both France and the UK, which aim to upskill developers and data scientists, but generally with a focus on security and high end research. There are sectors struggling to attract and/or utilise data skills. Investment in apprenticeships and conversion courses for aspiring data practitioners are needed. The number of data courses, amount of training and other forms of data education need to increase significantly, so it is important to be able to "train the trainers". These could be coupled with a funded secondment programme placing data practitioners in organisations who make a case for data support on a project.

Objective : Promote basic data literacy.

Recommendation : Develop initiatives to bring basic data literacy into primary and secondary education.

Just as students learn how to be good citizens and be 'cyber-wise', so they need to be 'data-wise' if we're going to make smarter choices about how organisations use our data. Several experiences have proven to be effective and could be easily encouraged and scaled up. For instance, france-ioi.org offers free gamified activities and contests to students through a bottom-up and independent approach. Basic understanding of the potential of data could be included in the curricula of "civic life and ethics" courses for students.



4. TRUST AND EMPOWERMENT

Objective : Empower citizens through their data.

Recommendation : Refine and extend MesInfos and MiData best practices to support citizen data empowerment.

Data ownership, portability, accountability and control are fundamental principles behind MesInfos and MiData and are reflected in the GDPR. How organisations will comply in practice with the incoming GDPR remains unclear. A citizen data empowerment scheme could be built on the experiences of MesInfos (France) and MiData (UK). Research into technology and regulation aiming at helping businesses empower citizens to better control and be able to take decisions regarding their personal data should be encouraged. Work should be undertaken to transform this regulation into an opportunity for companies.

Objective : Plan for future data policy challenges.

Recommendation : Commission research into algorithmic transparency and accountability.

There are emerging questions for policy makers and organisations about decision making using data, for example via algorithms. In some sectors, algorithmic decision making is already a part of our everyday interactions with services. How do we ensure our data driven decision making is fair, and doesn't discriminate? France and the UK should commission research into how organisations could be more transparent about algorithms, data sources and the data analytics process used to make decisions that affect consumers and citizens.

3

The Data Landscape

The Data Value Chain

The data revolution has the potential to drive economic growth and transform businesses, consumer experience and public services. The increasing availability and use of data to generate value creates challenges as well as opportunities.

We must understand how value is created in this evolving landscape to realise the transformational benefits of the data revolution.



BUILDING INNOVATIVE BUSINESSES

Access to data opens the opportunity for creating value through **innovation**. Data can be exploited by organisations as an asset or as a source of externalities. The use of data as an asset brings about positive change inside the organisation while its use as a source of externalities has a broader impact on the economy and the society as a whole. The key to trigger these externalities is to improve circulation of data, by sharing or opening it.

Data is unlocked as an asset when:

- it is used to improve resource allocation, processes, and to lower transaction costs to deliver efficiency gains;
- it is used to improve and personalise services to increase quality;
- it is used as an input for predictive analytics to inform decision-making and improve its accuracy; and
- it is used as an input for social research and behavioural modelling.

Selling access to data creates value in direct revenues to the companies or administrations that own it, but has been proved to be globally inefficient. This is particularly the case with public sector information. In Australia, geographic data was opened in 2002. Taking into account the loss of direct revenues, the annual gain was estimated at 1.7 million dollars¹. In the case of administrations selling data to each other, the transaction costs are proportionally higher and the existing exchanges defy economic rationality².

Data creates value by informing decision-making resulting in concrete financial benefits. For example, the French National CDO has worked with the public procurement department to analyse the power consumption of public buildings³, and identify the best trading strategies during government discussions regarding the liberalisation of the electricity market. The strategy identified by this approach allowed the saving of several hundred million euros.

1 « Re-use of public sector information. Report for Danish Ministry for Housing, Urban and Rural Affairs », Marc de Vries, 2012.

2 Rapport Fouilleron.

3 <https://agd.data.gouv.fr/2015/05/17/analyser-les-consommations-energetiques-des-batiments-publics/>.



CREATING DATA WE CAN RELY ON

The first step when analyzing data-driven value creation is to assess the existing **data infrastructure**. Does the crucial data exist? Does the available data fit desired uses? Some specific datasets are essential as they act as reference sets for processing and combining data from different sources. The company register, the cadaster or the national address database are clear examples of existing data infrastructure. The national address database allows the geolocation of authoritative legal addresses that supports countless applications.

Data infrastructure relates to data production, data identification and data quality. Each of these aspects have a direct impact on the robustness of our data infrastructure. For instance, the data that is made available, whether open or shared, is dependent on who owns, produces and finances that data. Poor visibility of what data is available can make mapping data resources a challenge. Even when produced and discoverable, data can be of poor quality - in terms of accuracy, documentation and the frequency of updates, which limits its ability to be effectively reused.

Once data infrastructure related issues have been addressed, the focus must be on how to make data actionable. Real accessibility requires open and machine-readable formats. Providing access through APIs has the potential to radically increase the value and impact we get from data. Moreover, broad reuse of data requires interoperability through standardisation to make data easy to transfer and to compare. This is a key to achieve economies of scale. In some cases, data needs to be anonymised before being used or shared, to retain useful data while protecting personal data and restraining individual identification.

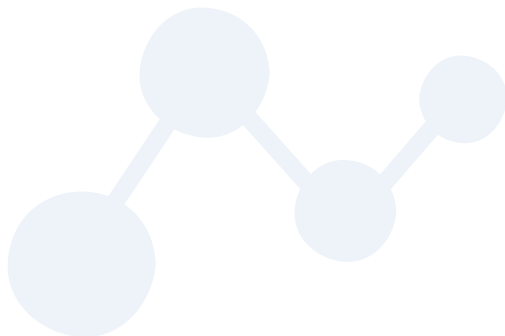




USING DATA TO MAKE DECISIONS

APIs, interoperability, standardisation, anonymisation, etc. The movement towards making good use of data requires very specialised **skills**. This is the reason why data scientists — innovative statisticians and the like who apply mathematics and computer science to data processing — have become key actors in the translation of raw data into concrete action. They are increasingly placed in or very close to decision-making roles, as reflected in the rise of Chief Data Officers in public and private organisations.

The skills within organisations need to be mirrored externally too. An ability to use open data, take advantage of algorithm transparency and general data literacy are needed to ensure effective accountability, consumer protection and active citizenship.





REALISING WIDER ECONOMIC BENEFITS

Data becomes a source of positive externalities when:

- it is used to foster innovation through the creation of new services and business models;
- it creates new markets and enables regulation;
- it is a source of development through the creation of social impact and the improvement of the economic rationality of agents;
- it is a source of transparency through the reduction of information asymmetry;
- it is used to foster collaboration through the creation of contributive commons and the exploitation of economies of scale.

Opening data brings about innovation. In the UK, the Open Data Institute “Open data means business”¹ report identified and analysed 270 companies that use, produce or invest in open data as a part of their business, accounting for an annual turnover of over £92bn, and over 500k employees. In France, the Open Data contest “Dataconnexions”², organised by Etalab since 2012, has allowed the identification of 200 high potential startups. In Spain, a study found that the “infomediary” sector (companies that sell services on top of open data) has at least 150 companies, employs around 4000 people and generates 330-350 million euros annually³.

Data can create new markets and enable regulation. In the mobility and the short term accommodation industries, for instance, data allowed the creation of new business models through companies such as Uber and Blablacar or Airbnb and Bedicas, respectively. In this context, open data can redefine the scope of a market. In France, a project called “Le Taxi”⁴ allowed taxis to compete on equal terms against Uber thanks to open data on the geolocation of taxis and open APIs for booking them. This approach was designed as a new regulation strategy for the sector.

Data can become a factor in international development as the “Data revolution”⁵ movement launched by the United Nations clearly illustrates. In particular, a dataset of cadastral parcels can create wealth. It supports the rise of a middle class in developing countries by allowing taxes to be collected and property titles to support access to loans and investment.

As a final example, the thriving Open Street Map (OSM) community embodies the creation of contributive commons and economies of scale through collaboration at both the national and global level. Crowdsourced data has proven to be competitive with official statistics. Collaboration minimises duplicate efforts across the public and private sector. OSM lead a wide range of projects, from the construction of public goods such as the National Address Database⁶ in France, to crisis management logistic support through disaster mapping in Afghanistan, Nepal or to track Ebola⁷.

Value can be created from data through many different mechanisms, more or less actionable depending on the context, the actors in play and the resources involved. Nevertheless, there is an infrastructure of methods, skills and mindsets that are crucial when aiming to harness the data revolution. The objective of this Taskforce is to identify and initiate concrete collaborations between the United Kingdom and France to mobilize and reinforce these essential levers of change.

1 <http://theodi.org/open-data-means-business>.

2 <https://www.data.gouv.fr/fr/dataconnexions-6>.

3 Rapport annuel du portail espagnol des données ouvertes, « Characterization Study of the Infomediary Sector », juillet 2012.

4 <http://le.taxi/>.

5 <http://www.undatarevolution.org/>

6 <https://adresse.data.gouv.fr/>.

7 <https://hotosm.org/projects/ongoing>.



1. INNOVATION

Together, the UK and France can develop one of the most creative and dynamic data ecosystems in the world.

Data management and data analytics companies already make up 12% of digital businesses in the UK and that proportion is rapidly growing¹. In France, the impact of big data, associated with the emergence of an Internet of Things could reach 3.6% GDP in 2020 (and about 7% in 2025)². As companies offering data solutions, tools and services continue to grow, so every sector will eventually be transformed by data. Companies using data effectively may be up to 13% more productive and 40% more likely to launch new products than their rivals³.

Startups managing and analysing data in new ways are driving data innovation. In both the UK and France, a range of programs and initiatives support data startups and SMEs, both in the capitals - through organisations including Tech City UK, the Open Data Institute and the Digital Catapult in London and Etalab, CapDigital and NUMA in Paris - and in regional areas (Tech North, Leeds Data Mill in the UK and all French Tech metropolises like Aix-Marseille and Digital Grenoble).

Today's networked societies extend beyond country borders. Digital teams are increasingly spread across cities and nations, utilising developer talent worldwide. The UK and France need to continue to reduce friction for businesses finding the data skills they need and growing into new regions and sectors. They need to provide support and incentives to innovate.

To foster more open, collaborative data economies for both the UK and France, we need common goals and ways to communicate and share experiences. The Taskforce recommends:

Objective: Encourage collaboration and focus our efforts.

Recommendation : Establish a UK-France Data Challenge Programme to bring UK and French organisations together using data to solve common problems.

Commission a new joint programme featuring two high-impact and high-return challenges each year, with a stated aim of demonstrating clear and measurable improvements. Etalab, Cap Digital, the Open Data Institute and the Digital Catapult may be partners for this work. For example, goals might include economic gains, improvements in efficiency, reduction of friction, and new innovative opportunities in areas that could include jobs, healthcare, transport, finance, food and agriculture, energy and environment, education, crime, justice, etc.

In 2013, the UK Department for Business, Innovation and Skills (BIS) invested £1.2 million in an Open Data Challenge Series, with seven challenge components. Research conducted in 2015 indicated that for every £1 invested by BIS, there was a return of £5 - £10 to the UK⁴. That challenge series focused primarily on the use of open data to deliver social impact.

In France, two main programs have been working on data awareness: DataConnexions challenges (created by Etalab) and Big Data Challenges (created through the "Nouvelle France Industrielle"). More than 25 Big Data Challenges have been realized by 5 organisations (Cap Digital, Numa, Images & Réseaux, Euratechnologies and Lyon TUBA) to help SMEs respond to public and private sponsors' challenges since November 2014. The French Government is supporting these Big Data Challenges for the next 2 years with a commitment of €2.8 million. France has also launched new Digital Challenges with nearly 100 challenges to be organized by 12 organisations at a cost of €7.7 million until 2019.

1 Tech Nation 2016.

2 Big Data et Objets Connectés, Institut Montaigne, 2015.

3 Nesta, Rise of the Datavores, 12.

4 <https://theodi.org/news/investment-in-open-data-challenge-series-could-see-10fold-return-to-uk-economy-over-3-years>.

A UK-France Data Challenge could involve collaboration using a mix of closed, shared and open data sets to solve problems common to both the UK and France. These common problems might include:

- **Public transport and traffic management in expanding cities;**
- **Improving health outcomes for citizens using data;**
- **Transforming consumer banking institutions.**

Data Challenges bringing UK and French startups and entrepreneurs together will be a catalyst for change in the way we work together. To build on challenges, however, we need to foster a shared, cross-border identity.

Objective: Scale up our networks.

Recommendation : Create a 'data innovators passport' scheme, to share best practices and extend startup networks between the UK and France.

Build on, provide support for and connect existing startup programmes between the UK and France, through an accelerator passport programme. Startups that are part of the programme would be able to move between incubators including the Open Data Institute, Digital Catapult, Cap Digital, Etalab, Numa Digital, Tech City UK and Scale Up! Startups will be able to learn about respective French and UK ecosystems and markets in order to facilitate their growth.

This scheme aims to harness peer support, and setup a virtual network. Any incubator or startup can apply for a data innovators passport if they work with data, creating a virtual network for data-related startups.

Objective: Connect data innovators and innovation across regions.

Recommendation : Establish a data twinning programme - cities in regional areas with common industries and/or demographics, and strong startup communities.

Data innovators aren't only in London and Paris. Regional business clusters and incubators are supporting UK and French firms outside the capital cities - like French Tech in Lille, French Tech Culture Avignon-Provence, TechNorth in Manchester and the Leeds Data Mill. The French Tech mission and Tech City UK could be the initiators of this new regional collaboration.

The French Tech mission and Tech City UK could work with cities interested in sharing their ecosystems resources and opportunities regarding specific themes (transport, tourism, culture, agriculture, etc.).



2. DATA INFRASTRUCTURE

Data is infrastructure. Data infrastructure will only become more vital as our populations grow and our economies and societies become ever more reliant on getting value from data.

According to the UK's Open Data Institute, a data infrastructure "consists of data assets, the organisations that operate and maintain them and guides describing how to use and manage the data. Trustworthy data infrastructure is sustainably funded and has oversight that provides direction to maximise data use and value by meeting the needs of society. **Data infrastructure includes technology, processes and organisation**".

There are data infrastructures that are owned and maintained by public sector organisations. And there are data assets of high value that are owned and maintained within the private sector. Like physical infrastructure - roads, rail, air travel - there's a mixed model of ownership of data infrastructure.

Data should be **fluid and able to circulate without friction where it is needed for services**. This is a prerequisite for the data economy to be able to produce growth, strengthen collaboration and link services. A strong data infrastructure will nourish value creation (use and reuse of data coming from all sectors). The existing data community would be supported and grown.

Data ownership and local storage are important considerations for how we build and maintain a data infrastructure with high level of security in the underlying layers. We need to be able to make choices about where we store data, and who has access to it that enable frictionless provision of services while preserving privacy and necessary control, and keeping the flexibility and innovation inherent in cloud computing services.

Increased demand for and reuse of data is constantly triggering new quality, accessibility and maintenance requirements. The need to identify and map existing data, improve access and facilitate frictionless exchange of data is core to data infrastructure. The identification and constant maintenance of a set of key data assets, that need to be as accessible and usable as possible, should be contemplated as a basic public good. The UK began to map its key (public) data assets as part of its conception of a National Information Infrastructure (NII) in 2013. In France, the Digital Republic law which passed in June 2016 reaffirms the "service public de la donnée", assuring the accessibility, quality and maintenance of core data registers deemed to have important social and economic potential.

The public sector has a responsibility to provide core data infrastructure to its citizens: this is a common good, a basic public service. Some reference data assets may need to be built by the public sector, and will need investment. Significant data assets will exist in the private sector; where possible, these should be as accessible as possible to maximise the value of data infrastructure to all.

Barriers to the development of key data infrastructure to underpin essential services often have to do with a piecemeal approach to data governance. Today, policy makers in both the UK and France are grappling with challenges including:

- Lack of knowledge and visibility of existing key data assets;
- Information systems that, for a majority of actors, are not optimised for data-driven, machine-based approaches;
- Cultural and legal barriers that impede data sharing;
- Lack of funding and resources to plug gaps and invest in new data assets and bodies where needed; as essential infrastructure, data infrastructure does require investment.

Considering this France and the UK should:

Objective : Make using and exchanging data easier.

Recommendation : Develop the use of API's in the data economy, learning from the UK Open Banking experience.

APIs are the nervous system of the 21st century. To that extent, we need to encourage and collaborate on common open APIs in sectors of shared importance. As commerce moves across country borders, how we exchange and access data must be fluid too. The UK has already begun work on an Open Banking Standard for the consumer banking sector. This work should be expanded on.

Interesting sectors for the development of common open APIs could include other regulated commodities such as energy, transport, water, broadband, telecoms; areas where people would benefit from enabling third-party access to their data, such as health or tax records; and sectors that include multiple collaborating actors, such as food supply chains.

Objective : Support the basic infrastructure for linked services.

Recommendation : Invest in and share experiences building core data registers, learning from the French National Address Database experience.

Registers provide an authoritative source of core reference data for linking and improving services. France collaborated with the OpenStreetMap community to develop the National Address Database - a commons based on data crowdsourcing, which is freely accessible and usable by anyone. The UK and France should collaborate and share lessons learned from that experience, as the UK explores building its own open address database.

Objective : Make it easier for citizens to get the services they need anywhere.

Recommendation : Continue to invest in and share experiences of e-identification schemes to enable smoother access to public & private sector services in each country.

France and the UK have been working on strengthening pervasive government mechanisms, like e-identification, that stretch across every government service. Privacy should be integrated in the public system by design, building on lessons learnt from projects like France Connect (France) and Verify (UK).



3. SKILLS

As the global digital economy grows, all leading economies - including the UK and France - are seeing a major increase in the demand for data skills. Access to data skills and talent is a constant challenge for both the public and private sector. Addressing the data skills shortfall should become a strategic priority, in that it has direct and very serious economic implications. Serious resources will have to be invested in order to counter this skills shortfall.

In France, the objective of the Big Data industrial plan roadmap, published in December 2014¹, was to **create or consolidate about 137,000 jobs**, either directly in the computer industry or in technological functions and business functions (for example in sales functions) in all companies, by 2020. At the outset, the project managers of this industrial plan, François Bourdoncle (FB&Cie) and Paul Hermelin (Cap Gemini), identified 5 key sectors at risk with the data economy transformation: insurance, energy, transport, health and employment, all linked to flagship projects. A new observatory of data usage was launched in 2016 through the Nouvelle France Industrielle that will monitor 40 sectors, their transformation in the Data Economy² and will also highlight the necessary skills for private companies.

New masters degrees dedicated to data and knowledge, big data analytics, machine learning, etc. are launched regularly in France and the UK to answer the need in coming years but these still seem insufficient in the face of the three following challenges.

Firstly, we need to help our current leaders to be innovative and pragmatic about how data can transform their organisations.

Secondly, there is a growing demand in both countries for data analysts who are able to combine their technical and analytical skills with soft skills and industry knowledge to turn data into real value for their organisations. While data science skills are increasingly becoming a part of formal education, there's little opportunity to learn soft skills like collaboration, problem-based learning, or working with non-data experts before entering the workforce. Both the UK and France are investing significantly in data science skills and research, such as the UK's Alan Turing Institute and the proposed Institute for Coding. While these are important, they won't result in the large scale transformation every sector of the economy will need. More needs to be done to connect data practitioners with sectors struggling to make use of data.

Thirdly, as more sectors become disrupted by digital technology, we are likely to see a growing need within the workforce of both countries for basic data awareness and literacy, as people are encouraged to make smarter choices about how their data is collected and used, and who has access to it.

France and the UK should:

Objective : Support senior leaders understanding the power of data.

Recommendation : Support data literacy programmes for leaders inside and outside government.

To realise large scale data transformation, we need to engage at the very top. Training could be developed for 'high potential managers' (business leaders, politicians, government officials, etc.) to inspire and equip them with the data awareness, skills and strategies they will need to transform their organisations. These leaders can come from public and private sector organisations.

1 <http://www.economie.gouv.fr/big-data-feuille-route-en-action>.

2 <http://observatoirebigdata.fr/>.

The programmes should involve both in person workshops and a peer network and have a focus on solving specific problems or challenges. Existing public and private organisations are creating new courses and this task force will support programmes that are targeting 'high potential managers'.

Objective : Overcome a chronic data skills shortage.

Recommendation : Embed new data and problem solving skills for existing analytical professions; and create common data apprenticeship schemes and conversion courses to provide a sufficient workforce of data practitioners in the UK and France.

Existing data related disciplines, including management, marketing, engineering sciences, economics and statistics would benefit from curricula that combines basic data training, problem solving and collaborative coding, data ethics and data visualisation. Alongside this, create new entry paths for aspiring data practitioners. There are initiatives in both France and UK, which aim to upskill developers and data scientists, but generally with a focus on security and high end research. There are sectors struggling to attract and/or utilise data skills. Investment in apprenticeships and conversion courses for aspiring data practitioners are needed. The number, range and extent of training, courses and other forms of data education need to increase significantly, so it is important to be able to "train the trainers". These could be coupled with a funded secondment programme placing data practitioners in organisations who make a case for data support on a project.

In France, the EdFab project was launched by Cap Digital³, in 2016 and targets innovation in the areas of education, training and job transformation. The Open Data Institute's Train the Trainer programme trained over 50 trainers in 2015.

Objective : Promote basic data literacy.

Recommendation : Develop initiatives to bring basic data literacy into primary and secondary education.

Just as students learn how to be good citizens and be 'cyber-wise', so they need to be 'data-wise' if we're going to make smarter choices about how organisations use our data. Several experiences have proven to be effective and could be easily encouraged and scaled up. For instance, france-ioi.org offers free gamified activities and contests to students through a bottom-up and independent learning approach. Basic understanding of the potential of data could be included in the curricula of "civic life and ethic" courses for students.

3 <http://www.capdigital.com/edfab/>.

4. TRUST AND EMPOWERMENT

The concept of trust and empowerment lie at the heart of the data opportunity.

In order to gain and maintain the trust of its citizens and consumers we need the real commitment of governments to protect the data of their citizens and their companies. Network security and digital sovereignty for people, organisations and governments are essential to our economic sovereignty but also to the overall confidence of citizens in the digital revolution, and perhaps in governments themselves.

Research commissioned by the UK Royal Statistical Society in 2014 and undertaken by Ipsos MORI indicated that the media, internet companies, telecommunications companies and insurance companies rank lowest in a “trust in data” league table. Only between four and seven percent of citizens surveyed indicated they had a high level of trust in these organisations to use data appropriately, compared with 36% trusting the UK national health service (NHS) and 41% their local GPs¹.

In France, authorities worked on a “Secure cloud” label² with public and private stakeholders in order to tackle the trust issue when using cloud computing services whilst also ensuring a level of security of infrastructure services. The chain of trust is important from the underlying infrastructure to the applications and their use of the data: 74% of French still claim that they do not trust how mobile applications use their data³.

For the data revolution to succeed, organizations must be more transparent with their users about how they are handling and protecting their information — including with whom it is shared and what data they can withhold. These principles are enshrined in the GDPR.

Many data services offered today are “black boxes” that deliver results or decisions based on algorithms and data sources whose behavior and content are opaque to the user. How these systems are built, how their requirements are established and the range of capabilities they give rise to are not known to their users. If elements are explained it is often in a technical language which is not accessible to common users. More and more, citizens, companies, public actors and governments need to explore and challenge this situation.

There will be circumstances in which sharing of our personal data is necessary, to help provide us with services we need (i.e. public health). There will be circumstances in which gaining consent before use of our data isn't feasible (e.g. implicit data derived from user actions during a web session). Nonetheless, we are increasingly seeing citizen centric data governance being placed at the heart of data economies. The citizen has to become an empowered actor who can take decisions regarding their own data. **Citizen empowerment means the possibility of controlling your own data and choosing what to do with it; for instance, to share it with other citizens, associations, research organisations or private companies.**

The French government (Conseil général de l'économie) has been engaging with different public and private actors with a view to commissioning a study on the regulation of algorithms that process content. This work could provide a starting point as we seek to understand the necessity of new policies in our data economies regarding trust and citizen empowerment.

Given the importance of trust and empowerment France and the UK should:

Objective : Empower citizens through their data.

Recommendation : Refine and extend MesInfos and MiData best practices to support citizen data empowerment.

1 <https://www.ipsos-mori.com/researchpublications/researcharchive/3422/New-research-finds-data-trust-deficit-with-les-sons-for-policy-makers.aspx>.
2 <http://www.ssi.gouv.fr/actualite/appel-public-a-commentaires-sur-le-referentiel-dexigences-applicables-aux-prestataires-de-services-securises-dinformatique-en-nuage/>.
3 Baromètre de l'innovation (Odoxa, Microsoft, L'Usine Digitale), 2016.

Data ownership, portability, accountability and control are fundamental principles behind MesInfos and MiData as they are reflected in the GDPR. How organisations will comply in practice with the incoming GDPR remains unclear. A citizen data empowerment scheme could be built on the experiences of MesInfos (France) and MiData (UK). Research into technology and regulation aiming at helping businesses empower citizens to better control and be able to take decisions regarding their personal data should be encouraged. Work should be undertaken to transform this regulation into an opportunity for companies.

There is a real opportunity for French and UK stakeholders (Mes Infos, MiData, Cozycloud, Imperial college, Digi.me, etc.) to collaborate on actions to support the development of personal data technologies and services: personal data management platforms, personal data leakage detection with expert systems, etc.

Objective : Plan for future data policy challenges.

Recommendation : Commission research into algorithmic transparency and accountability.

There are emerging questions for policy makers and organisations about decision making using data, for example via algorithms. In some sectors, algorithmic decision making is already a part of our everyday interactions with services. How do we ensure our data driven decision making is fair, and doesn't discriminate? France and the UK should commission research into how organisations could be more transparent about algorithms, data sources and the data analytics process used to make decisions that affect consumers and citizens.

Some international initiatives have been launched around transparency. INRIA will join the Data Transparency Lab⁴ this summer to pursue and develop new activities. The Open Data Institute was a founding partner in this initiative. The work of the Data Transparency Lab could be explored for insights and practical steps to ensure algorithmic transparency and accountability.

4 <http://www.datatransparencylab.org/>.

ACKNOWLEDGEMENTS

This report was made possible thanks to the support and advice of many individuals and organisations. The Taskforce would like to thank everyone who contributed to this endeavour.

Particular thanks go to Cap Digital, the Open Data Institute and NUMA who kindly hosted meetings of the Taskforce between November 2015 and April 2016.

We are extremely grateful to Salesforce, Access Partnership, techUK, IRT SystemX, and OVH for their time and assistance in facilitating consultation events with a range of tech businesses to help develop the report recommendations.

And our sincere gratitude goes to Cédric Mora (DGE), Fabien Terrailot (DGE), Achille Lerpiniere (DGE), Paula Forteza (Etalab), Vicki Brown (DCMS), Paul Driver (DCMS) and Ellen Broad (ODI) for their guidance and energy in providing the secretariat to the Taskforce.

