DG DIGIT
Unit D1

Study on public sector data strategies, policies and governance – ANNEX Case studies

Data analytics for Member States and Citizens

Date: 22/07/2020
Doc. Version: Final

THE REPORT HAS BEEN PRODUCED FOR THE EUROPEAN COMMISSION BY:

The research presented in the report has been carried out within the scope of the study Data Analytics for Member States and Citizens (Framework Contract DI/07624 - ABC IV Lot 3) commissioned by the European Commission, Directorate-General for Informatics, to Deloitte and the Lisbon Council for Economic Competitiveness and Social Renewal. The project has been carried out within the scope of the ISA² Action 2016.03 – Big Data for Public Administrations. More information is available at https://ec.europa.eu/isa2/sites/isa/files/library/documents/isa2-work-programme-2016-detailed-action-descriptions_en.pdf.

Contact information
DIGIT-DATA-SERVICES@ec.europa.eu

DISCLAIMER

TABLE OF CONTENTS

1. INTRODUCTION AND KEY FINDINGS

The data explosion is affecting all aspects of the society and the economy – and public administration is no exception. Data is a fundamental resource for carrying out all government activities, from regulation to service provision. And governments everywhere and at all levels are looking into the opportunities of data driven innovation, and in many cases experimenting with it. IDC estimates that central government is the fifth largest industry of the of the big data analytics market, covering about 7% of the expenditure, and growing fast. The European Commission itself has set up a data strategy (DataStrategy@EC) and a related Action Plan in 2018, with the objective of transforming the EC in a data-driven organisation. The Action Plan is centred around 5 different dimensions: data, people, technology, organisation, policy. The data strategy highlights indeed that these dimensions need to mature and evolve harmonically to deliver a real transformation on how data is used in the decision-making processes.

But how are data strategies deployed concretely? What can the European Commission learn from leading countries? To understand that the European Commission has commissioned the study **Data Analytics for Member States and Citizens,** which provides policy Directorate Generals of the European Commission and Member States public administrations with a knowledge base and guidance on the adoption of public sector data strategies, policy modelling and simulation tools and methodologies, and data technologies fostering a data-centric public administration.

Specifically, the study covers three domains in relation to data analytics in government:

1. **Data strategies, policies and governance**: initiatives in the public sector both at the strategic level, such as data strategies, data governances and data, management plans; and at organisational level, aimed to create units or departments, and to elaborate new processes and role.
2. **Policy modelling and simulation**: initiatives to improve policy analysis through new data sources, robust and reliable models to perform "what-if" scenarios, predictive analytics and hypothesis testing, and tools allowing policy makers to carry out scenario analysis through intuitive interfaces.
3. **Data technologies**: new architectures, frameworks, tools and technologies to be used by public administrations to gather, store, manage, process, get insights and share data. This domain includes the study of how data are governed as well as data collaboratives, and in particular stresses the joint analysis of governance and technologies.

This report presents the result for domain 1 building on five case studies selected for the in depth analysis, based on the level of ambition and maturity: Barcelona Data Commons, Data Agenda Government in the Netherlands, New Zealand Data Strategy and Roadmap, Secondary use of health and social data (Finland), and Udbetaling Denmark. The cases can be considered pioneers in ensuring a strategic approach to data governance in public administration. These are not economy-wide data strategies – like the digital agendas or data economy strategies, but initiatives focusing on greater adoption of data driven solutions in the public sector. But because of the encompassing nature of the topic of strategy, this report draws lessons also from the parallel reports of domains 2 (modelling) and 3 (technology).[1]

[1] This report looks deliberately only at dedicated, self-standing data strategies that aim to maximize data driven innovation in public administration. Many local, national and supra-national organisations have adopted "data strategies": but the actual content of these strategies can vary substantially. In most cases, there is an overarching digital agenda include a digital government policy, which includes some aspects of data management, for instance base registries or open data. Other data strategies – such as the European Data Strategy - are developed in the context of industrial policy, typically covering the overall economy, beyond government, and are designed to ultimately grasping the strategic opportunities of artificial intelligence.

The results show that data strategies are diverse in nature and objectives. All pursue the dual goal of greater data reuse and protection, with different emphasis on the two.

The first and foremost result of the strategies is putting the topic on the agenda and raising widespread awareness in public administration about the strategic value of data. This increased awareness has been reflected in greater access and reuse of data. Yet when it comes to visible results of these data sharing, the picture is uneven. Pilots, such as in the Dutch and Barcelona cases, show the potential to use data to improve policy and services, and enable quick delivery on different areas such as renewable energy, public order, housing and poverty. But while pilots are a good start, all strategies also point to the limited uptake and implementation of data analytics projects as a major challenge compared to expectations.

The limited visibility of the results goes hand in hand with the reluctance to share data, in a classical chicken and egg problem. Despite extensive co-creation effort, and despite data sharing being the declared goal of the strategy, it remains one of the most important challenges: this is not only due to technical barriers related to legacy and interoperability, but also to the resistance and lack of know-how to share data as well as the low quality of the data.

In other words, policy prioritisation is necessary but not sufficient. Just as in other domains such as research data, or open government data, or business data, top down mandates have to be combined with clear benefits and opportunities for reuse: they should be demand driven– whether internal (municipalities in the Danish case) or external (reusers of health data).

Obviously, one of the expected barriers to data sharing is data protection. Because of GDPR is still relatively new and there is limited knowledge about its practical implementation, several organisations exhibit cautiousness on going ahead with a data project. This continues to be true, but what is clear is that this barriers can be overcome when there is a clear case for it. Again, data protection is a much more powerful barrier when it goes hand in hand with lack of a business case.

The lack of data skills is obviously a problem. But this is not limited to the mere absence of a sufficient number of data scientists in public administration. It refers to the more general problem of data literacy among civil servants, and notably decision makers, which leads to a dangerous cultural gap between the data scientists and policy makers. But this also includes the lack of a policy knowledge by data scientists (policy literacy). Data is not a commodity. Extracting value from it requires not only cleaning and preparation, but plenty of assumption and choices in building analytical models. This processing implies that the results of the analysis are to be carefully considered in light of the limitations and assumptions. Data analytics should not be treated as a black box managed by data scientists where data are the input and decisions are the outputs. The recent example of Covid-19, where modelling tools have been at the root of potentially disastrous decisions by policy makers, is a clear demonstration of the need for distributed data competences in every policy domain and of the need of close collaboration between data scientists, data analysts and decision-makers. And the Danish case shows how important the collaboration between data analysts and frontline workers is.

The report includes a set of policy conclusions for all levels of government when designing data strategies:

1. Start with the problem, not with the technology
2. Analyse permanently user needs.
3. Co-creation is a fundamental component of the strategy.
4. It is not sufficient to consult and co-create with stakeholders: what matters is delivering results.
5. In order to ensure delivery, it is crucial to take a practitioner led approach.
6. Create a data culture across department and institutional level.

7. Because it's a long-term process, expectations need to be managed correctly and hype should be avoided.
8. A robust ethical framework is crucial and can be instrumental to innovation.
9. Monitoring should be present and structured but not drive the process.

The report starts off by providing a basic description of the strategies: their objectives and structure. The second section deals with the fundamental components of its governance, in terms of mechanisms in place. The third section looks at three necessary enablers of all strategies: the safeguards, the skills and the monitoring mechanisms. The final section looks at the results, the lessons learnt and the policy recommendations and is built on the insight presented in this report as well as on the other two reports for modelling and technology.

1 ANNEX 1: CASE STUDIES

1.1 Case Study: NL DIGITAAL: Data Agenda Government

1.1.1 Introduction

The Netherlands is a country in north-western Europe counting over 17 million inhabitants. It finds itself among the top ten best performing EU countries when it comes to digital government. It is one of the best European countries in both the widespread of online services (penetration) and the digitisation level of back- and front-office (digitisation).[2] When looking at the availability, accessibility and government support for reuse of open data, the Netherlands finds itself above OECD average.[3]

By publishing a new digital government agenda in 2018, *NL DIGIbeter. Digital Government Agenda* the government of the Netherlands demonstrates the ambition to hold on to its top performing standard. Regarding the future of digital government in the Netherlands, the responsible Minister Raymond Knops emphasizes that change is a structural thing and that this requires continuous adaptation by the government.[4] He also iterates that the only way forward is that of more digitalisation, and that this has to be accompanied by safeguards for basic rights such as privacy and security. Digital inclusion is another key theme. The theme of data forms a red thread throughout the digital government agenda, with a total mentioning of 92 times.[5] The creation of a national government data agenda is announced under the heading of protecting fundamental rights and public values, pointing to the importance of this theme for the way in which the Dutch government envisions data to be managed and used in the public sector.

1.1.2 Policy context

1.1.2.1 First whole-of-government data agenda

The *NL DIGITAAL: Data Agenda Government*[6] was published in February 2019 and presented to parliament the next month. The data agenda can be characterised as a white paper making explicit the government's policy assumptions on how to better manage personal data, open data and big data rather than having a prescriptive function. It describes how data can be leveraged for policymaking and solving social challenges. It is not a policy, but "lays the foundation for the government's data policy for the coming years".[7] According to the OECD, in developing its data agenda, the Netherlands manifests itself as one of the frontrunner countries when it comes to better managing, protecting and sharing data within the public sector.[8]

The data agenda involves the whole of government, ranging from state departments, provinces, municipalities to water boards and provides a shared vision and concrete actions to be taken to support all public stakeholders involved. Executive agencies are not yet well represented in the agenda. Here lies a chance for more cooperation, as much of the public action happens in executive agencies (e.g. the Tax Agency). Even though the agenda has a clear government focus, it is acknowledged that the government doesn't work in isolation and that B2G data sharing may be useful for public policy development and service provision.

> "*The government must — in consultation with society and the business community — determine where the boundaries are for using data*."[9]

[2] European Commission: 2018.
[3] OECD: 2019a.
[4] https://www.digitaleoverheid.nl/dossiers/wet-digitale-overheid/dossier-berichten/raymond-knops-digitalisering-kent-geen-eindfase/ (accessed on 24 January 2020).
[5] In comparison: "Digital": 170 mentions; "Information": 33; "Technology": 16; "ICT": 3.
[6] In Dutch: NL DIGITAAL: Data Agenda Overheid.
[7] BZK:2019.
[8] OECD, 2019b.
[9] BZK:2019

The agenda covers a period of three years, from 2019 to 2021. The period of three years for this agenda is quite long compared to what's usually the case for coordinating functions. Often, plans are only made for one year. The advantage of having a three-year period for this agenda is that it provides continuity, stability and security for the parties involved instead of a fear that plans may be cancelled due to other priorities. There's also an efficiency gain, because organisations don't lose time in reassessing every year whether they want to continue data-driven activities, but can actually focus on developing the activities themselves. For example, the Data-Driven Approach Learning and Expertise Centre (LED) may set different priorities each year. However, its existence is undisputed, as this multi-annual effort was agreed upon in the data agenda government.

The agenda is not set in stone, but should be considered an evolving document which is updated on an annual basis. For the purpose of the 2020 update, working sessions have taken place in the steering board and sounding board group. The updated version should be published in April 2020. This is going to be a supplement to the existing document, rather than a revision in its entirety. It will be focused more strongly on the impact of data usage on major political issues, such as climate, organised crime and citizen involvement in government.

The data agenda is a policy document providing both an overview of ongoing activities and at the same time putting on the agenda the importance of what is being done with data, thereby fulfilling several functions:

- *Means of communication to show what's happening in the area of data-driven government. Stakeholders expressed the need to know what's happening where and to bring this information together.*
- *Policy document presented to parliament in March 2019 gathering political attention. It presents both actions that already took place and ones that still need to take place, up to 2021.*
- *Create unity and consistency between all layers of government, focusing on shared social challenges that can be solved by data rather than working in siloes.*
- *Concrete guidance as to which aspects are relevant when starting to work in a data-driven way (technical, people etc.); especially useful for smaller local governments.*

1.1.3 Building blocks of the agenda

1.1.3.1 Main themes

The agenda first presents three overarching action points aimed at developing new regulations and agreements about data in the public domain. These entail the development of a model agreement on the use of data collected in the public domain, an inventory of bottlenecks, and a proposed package of measures based on the bottleneck analysis. More specific action points are then presented, divided over five key themes. For each action point, a responsible actor is defined. The five key themes of the agenda are:

1. Problem-solving with a data-driven approach;
   - Two actions are identified to address and learn from the five selected social challenges of energy transition, manure issue, infrastructure and spatial bottlenecks, poverty and the issue of debt, and subversive crime.
2. Focusing on legislation and public values;
   - Eight actions are identified to develop new general principles on a responsible way of dealing with data taking into account legal and ethical frameworks.
3. Improving the quality of government data and using it more efficiently;
   - Twelve actions are identified to ensure government has the right data and is able to share (open) data at the right time and in the right way in order to foster a service-oriented and transparent public sector.

4. Collecting and sharing knowledge about a data-driven approach;
   - Eight actions are identified to address the efficient sharing of best practices.
5. Investing in people, organisations and changes in corporate culture.
   - Five actions are identified to address skills needs and cultural change.

1.1.3.2 Connections to other strategies

The 2018 Dutch Digitalisation Strategy (Nederland Digitaal) – is a Cabinet-wide strategy that considers every aspect of digitalisation. Under the umbrella of this strategy, a new national digital government strategy was developed: NL DIGIbeter, Digital Government Agenda. NL DIGIbeter puts into practice the Dutch Digitalisation Strategy for the public sector. While this agenda focuses on government and its interaction with citizens and entrepreneurs, a separate agenda deals with cyber security issues – namely, the Dutch Cyber Security Agenda.[10]

At the same time as the Data Agenda Government, work was carried out by the Ministry of Economic Affairs and Climate Policy in formulating the Dutch vision on data sharing between companies (De Nederlandse visie op datadeling tussen bedrijven).[11] Another policy document related to the Data Agenda Government is the vision shared by the municipalities on taking a data-driven approach, drawn up by the Association of Netherlands Municipalities (VNG). There's an opportunity to make more explicit connections between the different agendas, perhaps in the next update of the data agenda.

The data agenda was preceded by the *National Open Data Agenda* (2015).[12] The new data agenda provides a holistic approach towards the collecting, sharing, publishing and (re)using of both open and closed data. As such, the open data agenda has become an integral part of the country's digital government agenda, which is also reflected in the governance arrangements for the continuous evolvement and simultaneous implementation of the agenda.[13]

1.1.3.3 Legal framework

The legal framework related to government data in the Netherlands is in full development. On different topics and existing regulation, the Dutch government is raising the bar to enable a better data quality and governance across the public sector.

Key developments when it comes to regulation are the revision of the **PSI directive** (Wet hergebruik overheidsinformatie) and the development of the **Open Government Act** (Wet Open Overheid). The latter must ensure that government information is easier to find, exchangeable, easy to access and easy to archive. The bill has been approved by parliament and is yet awaiting the scrutiny of the senate.[14]

- We're also advancing open data legislation by implementing EU open data directive
- Access to information – wet open overheid
- Revision of Archive law
- Introduction of a baseline on information security

[10] https://www.nldigitalgovernment.nl/digital-government-agenda/ (accessed on 27 January 2020)
[11] https://www.government.nl/documents/reports/2019/02/01/dutch-vision-on-data-sharing-between-businesses (accessed on 28 January 2020)
[12] https://data.overheid.nl/ondersteuning/open-data/beleid (accessed on 12 January 2020).
[13] More details are provided in the section on implementation process.
[14] https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/democratie/democratie-in-actie/wet-open-overheid/ (accessed on 28 January 2020).

The GDPR already existed when the data agenda came out. This shows a trend of awareness on the fact that data quality has to be improved to enable more reuse of government data.

The **Digital Government Bill** (Wet Digitale Overheid) provides the legal framework for the Generic Digital Infrastructure (GDI). The latter is seen as a best practice when it comes to data sharing. It formulates provisions on the base registries and data exchange in those registries. Making data available and shareable G2G and G2B and sometimes B2G is a precondition to be able to use data. If it's not there, you can't use it

The proposed law has as objective to ensure a safe login for the Dutch citizens and businesses to the (semi) government entities. This includes the rules on information security and privacy. In this manner, the Bill will set out the principle that requires government bodies to employ open standards. It will also give supervision and enforcement a basis in law.[15]. In doing so, The Netherlands implements the EU directive on accessibility of government websites and apps.[16] In February 2020, the bill was approved in parliament and now awaits the scrutiny of the senate. The law is expected to get into force by mid-2020 and to be implemented in phases from that moment on.[17]

The data agenda indicates that Investigations are still ongoing into the extent to which new legislation or an amendment to existing legislation concerning privacy is required. The Ministry of Justice and Security has prepared a **Bill Data Processing Public-Private Partnerships** (Wet gegevensverwerking door samenwerkingsverbanden).[18] If it becomes easier for different parties to process data together, this could lead to a more efficient approach to fighting subversive crime as well as fraud. The bill also contains several safeguards for the protection of privacy. The Ministry of Justice and Security seeks to proceed with the bill within the foreseeable future, as soon as it has processed the advice that it has received.[19]

1.1.3.4 Funding

The implementation of the agenda is funded through a number of sources.

The **principal budget** made available by the Ministry of the Interior and Kingdom Relations (BZK) to give an impulse to the data agenda is about EUR 10 million for the whole three-year period.[20] The agenda shows a table indicating eight budget categories based on the standard government budget classification system rather than the content of the agenda itself. In practice, this doesn't lead to discussion as stakeholders find that investments are made where they should. There's no clearly defined process for the allocation of funding. Proposals on this can be made to the steering group, which can then decide who gets what. A complicating factor is that some organisations applying for funding are also represented in the steering board, potentially creating a conflict of interest.

Since the implementation of the agenda is the shared responsibility of the involved government stakeholders, several of them use their own **institutional budget** to finance some of the activities. Sometimes this depends on particular people advocating for their organisation's activities to be included in the data agenda to ensure both visibility and continuity of these projects in their own organisations. Since activities are mentioned in the data agenda, they need to be accounted for in parliament, making it harder for these to die a silent death.

15 https://www.nldigitalgovernment.nl/digital-government-agenda/ (accessed on 27 January 2020).
16 Government of the Netherlands: 2020.
17 https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/wetgeving/wet-digitale-overheid/voortgang-wet-digitale-overheid/ (accessed on 24 January 2020).
18 https://wetgevingskalender.overheid.nl/Regeling/WGK008727 (accessed on 7 February 2020).
19 BZK:2019.
20 In the Data Agenda Government, an amount of EUR 2.5 million per year is stated, but a key stakeholder revealed that in fact a bigger sum has been dedicated.

Finally, there's BZK's "**Innovation Budget** Digital Government" of EUR 18 million, which government organisations can apply to for funding. BZK directs organisations there in cases where there's no budget available from the agenda.

1.1.4 Governance of the implementation process

The political responsibility for digital government in the Netherlands lies with the Minister for the Interior and Kingdom Relations. Sectorial ministers are responsible for ICT in their domains. Central and local government are jointly responsible for implementing the Data Agenda Government. BZK plays a coordinating role in this respect, based on its responsibility for digital government and the safeguarding of fundamental rights. According to the OECD (2019a), BZK has very few decision-making and advisory responsibilities compared to other bodies in charge of the digital government strategy in their respective countries.[21]

1.1.4.1 The central coordinating department

The role of the coordinating department is a fascinating one, in which it seems difficult to do things right. The challenge for BZK is to find the balance between showing vision and not dictating every single detail to the other stakeholders. To deal with this tension BZK has organised working sessions in which they explicitly say that discussions are still open and that's it's possible to react now or later. The Ministry also feels the responsibility as coordinating department of being accountable for the agenda's progress in front of parliament. This explains why BZK has taken care of the implementation of several actions themselves, to guarantee these are well executed and can be presented as such.

> *"Yes, BZK is responsible, but the data agenda has been strongly advised and co-decided by all the partners involved."*

Unlike many other countries[22] the Netherlands has not appointed a National Chief Data Officer (CDO) and is hesitant to make the creation of this position compulsory across the public sector. It rather chooses to strengthen the CIO system at the central level of government, among other things by reassessing the tasks and responsibilities of the central, departmental and executive agency CIOs.[23] More and more public and private organisations across the country, however, choose to have a CDO, and are free to do so, along the creation of other positions, such as Chief Digital Officer, Chief Information Security Officer (CISO) or Data Protection Officer.

1.1.4.2 Decision and coordination boards

Two key groups of actors are the steering board and the sounding board. The steering board consists of directors from several departments: BZK, Education & Science, Economic affairs, Justice and Safety, Infrastructure, Statistics Department (CBS), VNG, Water Board Union, IPO (provinces) represented by the province of North-Brabant or South-Holland, as they share the responsibility for the data agenda of provinces. The steering board has existed for a long time. It used to be the steering board for open data and open government. All national government departments have been invited to join the steering board, but not all have joined due to other priorities. The steering board meets once in every 8 weeks.

It is not clear what the exact role of the steering board is, notably with regards to the ability to take decisions. Then it was decided to create the sounding board consisting of knowledgeable employees able to discuss content, share knowledge and prepare the steering group meetings. It remains somewhat complicated for the steering board, though, to take decisions on policy directions, as the members don't always feel legitimate in doing

[21] For the exact data, see StatLink https://doi.org/10.1787/888934033042.
[22] In many cases, the CDO's responsibilities only include open data, and not closed data.
[23] https://zoek.officielebekendmakingen.nl/kst-26643-656.html (accessed on 6 March 2020).

so. This is especially the case when an action is fully executed and financed by one particular organisation. The steering board does have the ability to block certain developments. If the members turn their back on the agenda, the Minister of Interior Affairs would not have a whole-of government agenda to present anymore.

Last summer, due to shifting responsibilities in the different organisations, there were a lot of new people around the table, who then need to be informed and convinced about the added value of the data agenda for their organisation. Sometimes the impression exists that BZK is showing off other organisations' efforts. Through interaction and exchange of ideas people can be convinced, when they see new ideas pop up and there's budget. Also, the request was expressed for BZK to support the community of involved organisations, e.g. by organising meetings on specific themes, such as the role of the chief data officer, for which several organisations have vacancies. The LED has consequently indicated they would organise this, thereby fulfilling a different role for BZK. Not just asking organisations for updates on the agenda, but also offering concrete help. It's a dynamic process with progress.

Both the steering board and the sounding board, have gotten a better representation of stakeholders and a better turnup than when they started. Stakeholders want to be present and feel they miss something if they're not there.

1.1.4.3 Fostering collaboration between stakeholders

BZK has approached both the creation and the implementation of the data agenda as a process of collaboration rather than enforcement. As such, it employs a mix of instruments to encourage all government stakeholders to play their part in the data agenda: practitioner-led co-creation, attention to communication, a dedicated budget for implementation and facilitating knowledge exchange through the LED.

"*We're a catalyst and sponsor of on ongoing development*."

Co-designing the agenda

The initiative was first created thanks to a group of very dedicated people within BZK making a real effort several years ago. They had the important insight of developing the agenda from the practitioner perspective rather than a policy perspective. Instead of following a typical policy approach of sending out an email with plans and giving the opportunity to react to those, BZK decided to commission the ICTU, an independent consultant and executor within the government to draw up the agenda using a hand-on approach of bringing stakeholders together, holding consultations and having one-on-one meetings to see what needs exist across the different government organisations. This effort resulted in an agenda with a large support base having a storyline that all agree upon. Once the ICTU had drawn up the agenda, an editor and graphic designer made sure the text was comprehensible to all and had an attractive look.

As mentioned before, a critical success factor for the implementation of the Data Agenda Government is the practitioner-led co-creation taking place during its definition. The Data Agenda Government has also been made possible by the Government-wide Policy Consultation on Digital Government (OBDO), in which the following parties are represented:

- The Central Government Chief Information Office (CIO-Rijk)
- The Association of Dutch Provinces (IPO)
- Ministry of the Interior and Kingdom Relations (BZK)
- Ministry of Economic Affairs and Climate Policy of the Netherlands
- Ministry of Infrastructure and Water Management
- Ministry of Justice and Security
- Ministry of Education, Culture and Science
- Ministry of Social Affairs and Employment
- Ministry of Health, Welfare and Sport

- Logius, the digital government service of BZK[24]
- Dutch Water Authorities
- Cooperation Agency of the Association of Netherlands Municipalities (VNG)

Furthermore, the Directorate-General for Public Works and Water Management (Rijkswaterstaat), and CBS were also involved.

An inventory was made of what actions already existed. This was done by sending out emails to organisations asking which data-driven initiatives they were running and whether they had new plans in this area for 2019. During that consultation round, parties were also asked which social challenges they found important enough to include in the data agenda. This has become the list of five social challenges as included in the agenda.

The data agenda is meant to stimulate and entice government stakeholders to cooperate better when it comes to data. It also helps to address bottlenecks. Actors don't get punished for not delivering on an action point. If crucial partners are lagging behind, though, BZK will try and convince them by listening to any doubts, hesitations or practical reasons for not advancing the data agenda.

> "W*e can't impose: 'thou shalt work in a data-driven way'"*

However, since the reporting to parliament is mandatory, stakeholders make an effort to make good on the actions they signed up for.

Connecting data communities and knowledge exchange as a policy lever

Next to financial resources, BZK provides other resources for the implementation of the data agenda: the connection of data communities and the facilitation of knowledge exchange. For example, BZK identifies where resources are already available and whether stakeholders who are already more advanced are willing to help out those that need it.

Using the dedicated budget for the agenda, BZK finances among other things a trajectory focused on two social challenges: 1. poverty and debts and 2. subversive crime. This trajectory is jointly executed by the CBS and VNG providing a design approach to help municipalities use data to deal with these social challenges. The exchange of knowledge and experience between government stakeholders is a crucial part of advancing the data agenda. A lot of interesting data-driven activities are happening across the country, but sharing information about this is not the first thing on people's minds.

The expertise centre for open government was rebranded as expertise centre for the data-driven public sector (LED) and tasked with a broader agenda. The LED, commissioned by BZK, and having a solid budget of its own, plays a facilitating role in interconnecting data communities of open data specialists, privacy officers, information security officers, archivists, big data specialists, etc. For example, based on a question from a municipality, the LED organised a knowledge exchange among municipalities on the use of data for early signalling of people running into debt and poverty related problems. The LED also facilitates community building internally at BZK, by bringing together ministry stakeholders who are involved in data-driven ways of working. Thereby it helps the ministry to practice what it preaches.

While most exchanges take place via email, there are a couple of collaborative tools available to the stakeholders in the agenda.

- *Pleio, is an open source platform, which all government stakeholders can use to share their strategies and plans. It is not very interactive, but really focused on sharing documents. It is not widely used though.*

[24] https://www.logius.nl/english.

- *On 31 October 2019 a Data and Ethics summit took place in Den Bosch, where LED was represented. Here, the website mydataquestion (mijndatavraag) was launched. People can ask questions but also share their knowledge and expertise. So far, several stakeholders have provided input.*

1.1.5 Data quality and sharing

1.1.5.1 Towards a climate of data sharing

Data sharing is not mandatory for Dutch government organisations, except in relation to the base registries. Organisations are encouraged, rather than obligated to share data, which fits the Dutch culture. Municipalities have the right to do things in their own way.

New legislation, such as the GDP and the PSI Directive, raises the bar on data governance and data quality. While data quality is usually sufficient for its primary use, it isn't for secondary use. To improve this, the data agenda proposes an inventory of specific areas where data quality should be improved and a stricter way to manage data.

The data agenda aims to encourage data sharing from government to government (G2G), government to business (G2B), government to science (G2S) and government to citizens (G2C). The central data portal, data.overheid.nl currently offers almost 15,000 data sets. There's also an API portal, developer.overheid.nl for all APIs offered by the Dutch government.

Companies developing navigation systems have real-time information on the whereabouts of cars. This is very interesting information for Rijkswaterstaat.

"Give a little, take a little."

Rijkswaterstaat then delivers data of a higher quality to these companies than they would normally receive. That's a way to help each other.

BZK aims to facilitate the creation of an infrastructure of existing data labs, so organisations know where they can take their specific data and/or data question. This can be the lab holding the appropriate data and/or expertise to generate insights from this particular type of data. For instance, if an organisation wants to perform an analysis concerning mobility data, the National Data Warehouse for Traffic Information[25] is the data lab to go to.

"*Depending on the datasets you need and also the type of operation, you determine which data lab is the best fit to help out*."

In a report commissioned by the Ministry of Economic Affairs and Climate, nine common building blocks of trust frameworks for data sharing were identified for the purpose of stimulating business to business data sharing:[26]

- Standards for messages and data
- Operational agreements
- Legal agreements
- Earning and funding model
- Connectivity
- Governance
- Metadata
- Consent
- Identification and authentication

The recommendation to develop a generic agreement system for data sharing in the business realm also resonates with BZK for the public sector context. The current priority

[25] https://www.ndw.nu/en/ (accessed on 7 February 2020).
[26] Innopay: 2018

is to create a climate of data sharing, where organisations see the value of sharing and reusing data, rather than developing a technological infrastructure.

1.1.5.2 The base registries system

The Dutch government has a base registry system containing 10 interconnected operational base registries[27], of which nine are fully available in 2018:[28]

- The Key register Persons (Basisregistratie Personen, BRP)
- The Key register Addresses and Buildings (in Dutch: Basisregistratie Adressen en Gebouwen, BAG)
- The Key register Large-scale Topography (in Dutch: Basisregistratie Grootschalige Topografie, BGT)
- The Key register Income (Basisregistratie Inkomen, BRI)
- The Key register Kadaster (Basisregistratie Kadaster, BRK)
- The Key register Subsoil (Basisregistratie Ondergrond, BRO)
- The Key register Topography (Basisregistratie Topgrafie, BRT).
- The Key register Vehicles (Basisregistratie Voertuigen, BRV)
- The Trade Register (Handelsregister, HR)
- Key register WOZ (Waardebepaling Onroerende Zaken, or Valuation of Immovable Property)

The most recent data on the use of these registries show that relatively large numbers of users already access the base registries. The figure below indicates the total number of user consultations, downloads or views per base registry for 2018.

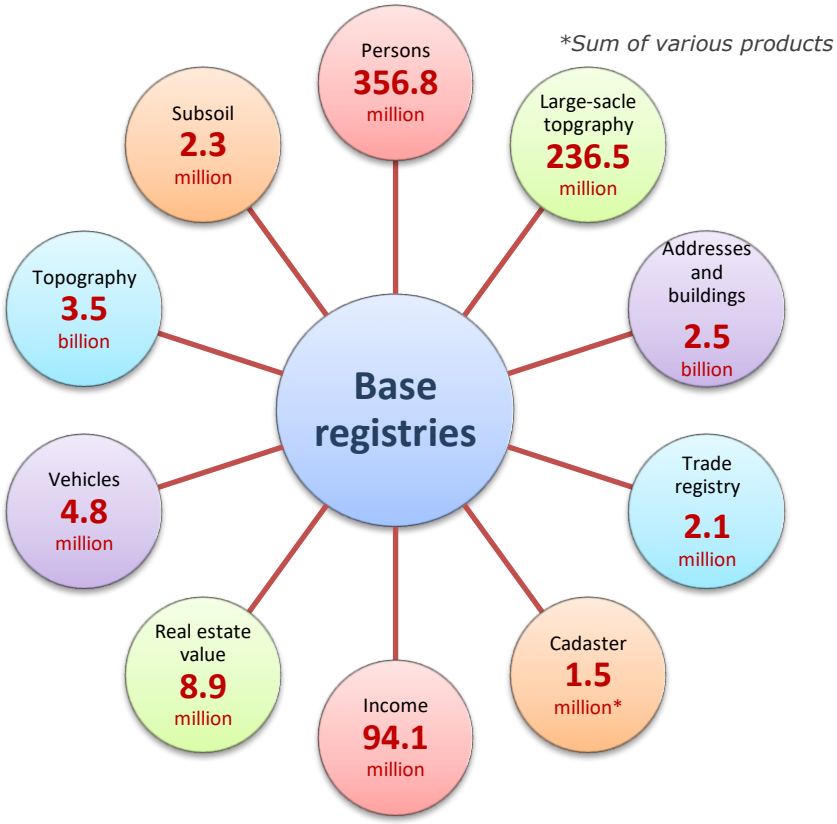

Figure 1 Dutch base registries use in 2018[29]

[27] Common English translations of The Dutch term 'Basisregistratie' include: base registry, key register and basis registration.

[28] Only the BRO is not yet fully available by the end of 2018, as the Base Registry Subsoil Law only came into force on 1 January 2018 (ICTU: 2019).

[29] This figure has been compiled based on data presented in the Digital Government Monitor 2019 (ICTU: 2019).

Several ministries are responsible for the base registries in their domain. The Ministry of the Interior and Kingdom Relations is responsible for the coherence of the base registry system, enabling the enforcement of the once-only-principle. For coordination, OBDO is in place. The system includes descriptions of base registries and responsible ministries.[30]

A connection in the base registry system means a data point in one base registry referring to a data point in another base registry. The referring data point can be a unique administration number (an "identification key"), but also an x / y coordinates.[31] Most connections are made through life events, such as the start of a business or a person moving house. To realise this connectivity, system standards have been agreed upon. For instance, to enable data from one base register to be used in another one.[32]

It is mandatory for all government organisations to use the base registry system with regards to the authentic data.[33] A specific base registry's law dictates which data are authentic and which aren't. This is the so-called primary use of the data. An example is a municipality using data from the Key Registers Persons when registering them as a new resident. In practice, compliance can still be improved. The secondary use of data concerns making new analyses based on existing data.

The system consists of two types of registries: public and closed. Public registries contain open data, which are accessible to all, such as license plate data in relation to a car. Six base registries, the BAG, BRK, BRT, BGT, and BRV are in part or completely available as open data.[34] From 2016 to 2017 the use of this open data has seen an increase of over 40% to 5.2 billion consultations. Closed registries, containing data such as private data concerning a license plate's owner, are only accessible to those who need to access the information for work purposes.

Legacy IT systems are a problem in the Netherlands, especially for big executive agencies, such as the Tax Agency. However, in the context of the data agenda, it doesn't arise as a key problem. This can probably be explained, because there already are several solutions available, forming a layer on top of the old IT systems. Data in the old systems can be made available and an API tool is used to perform the analysis. Also, there's a lot of innovation and pioneering going on in data labs. However, this doesn't mean there's a seamless integration with structural organisational systems.

1.1.6 Safeguarding public values

1.1.6.1 Ethical data use

The Dutch government considers the debate on the ethical use of data a public debate. Research among millennials showed that nudging is acceptable to them as long as they had the impression of having a free choice to ignore financial, legal or other incentives. For them, the use of apple pie scent in the streets is a step too far, because their bodily integrity is harmed. This example shows how generational differences as to what is accepted and what's not are important to take into account in the public debate on ethics and development of data policy.

Together with the Cooperation Agency of the Association of Netherlands Municipalities, BZK has developed general principles for the responsible use of data, after several municipalities had indicated running into difficulties regarding data sharing with companies. Another initiative to help public bodies navigate the legal and ethical data

30 https://joinup.ec.europa.eu/sites/default/files/inline-files/Digital_Government_Factsheets_Netherlands_2019_0.pdf (accessed on 24 January 2020).
31 BZK:2018.
32 https://www.nldigitalgovernment.nl/dossiers/base-registers-and-system-standards/ (accessed on 27 January 2020).
33 https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/basisregistraties-en-afsprakenstelsels/inhoud-basisregistraties/ (accessed on 27 January 2020).
34 BZK:2018.

landscape is the policy letter on artificial intelligence, public values and human rights that BZK sent to parliament in the autumn of 2019.[35]

1.1.6.2 GDPR compliance

Stakeholders find the application of GDPR quite complicated, especially in the social domain. In the social challenge on poverty and debts, for instance, it is not clear which data are allowed to be linked, whether analysis can be done or not, and whether subsequent policy actions can be taken. Municipalities don't always have the same interpretation on how to proceed. The more cautious approach is to not use the data if it's not 100% clear that it is allowed to do so. Municipalities on the more progressive side interpret the grey legal area as a go-ahead until they're stopped. The latter position is supported by BZK and the Ministries of Health and Social Affairs. Public values, such as privacy, good service provision and transparency are balanced on a case-to-case basis. BZK in its coordinating role consults legal advisors to answers concrete questions they receive from public stakeholders and provide clarity on the responsible way to go ahead. In the case of subversive crime, for instance, if there's an ongoing criminal investigation, the police cannot share data with municipalities. Then it falls under the Policing Law. If the data are anonymised for statistical research, it can be shared, because then it's not traceable to specific persons.

How big a challenge it is to provide for an adequate application of the legal framework related to the processing of data, has become very apparent in recent days. On 5 February 2020, a Dutch court has ruled the way in which the government uses the fraud detection system SyRI, which links data from different sources, such as the Tax Agency, The Unemployment Agency and municipalities, as an infringement of Article 8 of the European Convention on Human Rights.[36] According to the court, the SyRI legislation doesn't provide a fair balance between preventing and combating fraud in the interest of economic well-being on the one hand and the violation of privacy on the other. This ruling urges the Dutch government to go back to the drawing board.

1.1.6.3 Personal data management

Regarding personal data management, the Dutch government has launched the policy initiative *Control over data* (Regie op gegevens), which aims to give citizens and businesses more control on what's happening with their data.[37] At the moment, various appointment systems and solutions are being developed to support citizens and businesses in managing their data. Ultimately, this should result in a generic cross-sectoral framework that enables secure, reliable and user-friendly digital exchange of data between governments, private and social organizations.

1.1.7 Human capital and skills

BZK started a data science trainee programme, coming from business operations rather than the policy side. The idea is to place young talented people with data science experience across different government organisations. These people have great skills but learn more about policy and working in a big government organisation. This programme also serves to create awareness in organisations about working with data. Some organisations react in a very positive way, whereas others do not see the added value of data visualisations for their policy work. However, sometimes there are practical problems, such as not having proper work spots, for which it's not always clear who's responsible for providing a solution. The data trainees sometimes find being a sensor for data-driven readiness a rather frustrating experience.

35 https://www.digitaleoverheid.nl/wp-content/uploads/sites/8/2020/01/Dutch-policy-brief-on-AI-public-values-and-fundamental-rights_DEF-T.pdf (accessed on 7 February 2020).

36 For the official ruling in English, see https://uitspraken.rechtspraak.nl/inziendocument?id=ECLI:NL:RBDHA:2020:1878 (accessed on 7 March 2020).

37 https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/gegevens/regie-op-gegevens/ (accessed on 7 February 2020).

The department of Policy information within BZK has several data scientists and has a new vacancy as well. This is an existing department, which has taken more of a data-driven approach over the course of the past few years. e.g. HR analytics and analysis on business process information. There are also data scientists in the personnel administration which serves all state level organisations. They provide analysis for multiple departments already. Furthermore, there's a cooperation with the Statistics Department CBS for their knowledge and expertise. They're government, but their services are not free.

The government has managed to attract data talent coming from the private sector. The main reason for a data scientist to make the transfer from the private to the public sphere is the social impact they can have with their work. Other reasons are the work pressure, which is considered to be lower in the public sector, and the room for personal development.

Because of the scarcity of data scientists, people with a more unusual profile, coming from a different background and having acquired data analytical skills later on, get the opportunity to have a career as a data scientist in government. They get the chance to learn while doing, which they might not get in a commercial environment.

However, just having enough good data scientists isn't the holy grail. Often, public organisations run into problems which have more to do with organisation than with data skills, such as defining which question needs to be answered using data or deciding which organisational entities should be involved. Data scientists may not always be able to be of added value anywhere, because of the need of a special mandate (e.g. working on migration issues in the justice department) and there may be a lack of trust towards a data scientist coming from outside the organisation. She may not get access to the data she would need for the analysis. It might therefore be advisable to train a data scientist in the organisation she will work in, in order to become familiar with existing processes.

The agenda mentions the intention to make the training programme on data-driven work provided by the National academy for digitalisation and informatisation (RADIO) into a standard or possibly mandatory part of every employees' development plan. It is not sure whether this action has been realised at this moment. RADIO offers several online and offline courses on data-related themes, such as data governance, GDPR, and data for policy as well as a dedicated data-webinar.[38] The modules are given at a very basic level, as the developers found the basic knowledge of public servants lower than expected. In 2019, there were over 3000 users of RADIO courses (2800 online and 250 in the physical classroom).[39]

1.1.8 Monitoring and evaluation

The data agenda announces to "monitor the effect of the initiatives set out in this agenda on the government and society" (Ministry of the Interior: 2019). For the moment, this is not done in a data-driven way using KPI's, but by means of a planning spreadsheet detailing all the actions, phases and responsible actors. This file is used to keep the steering group and sounding board group informed on the progress. A challenge in the monitoring process is the use of the classic spreadsheet planning file to present progress to the sounding board and steering board, as it results in a big A3 print with a lot of information which is not very insightful. The planning document mentioned earlier contains cases on progress made per quarter. BZK has asked the responsible stakeholders for the actions about the state of affairs and problems they may have run into. All of this information is included in the planning file.

The implementation of the data agenda is on schedule. Many of the actions have already been realised. This can be explained in part through the fact that during the creation of

[38] https://www.it-academieoverheid.nl/aanbod (accessed on 7 February 2020).

[39] This numbers refer to the total number of courses related to digitalisation. No specific numbers on the data-related courses are available. With thanks to Ms. Marijke Abrahamse, Senior Advisor at RADIO for providing these numbers.

the agenda, several actions were included which were likely to be completed soon or had already been well prepared. At this point, no actions are lagging behind.

Sometimes, BZK finds itself in a difficult position as the coordinator of the agenda when trying to monitor. Stakeholders may feel that they're the ones doing all the work and BZK being the one pushing them and checking up on them. So, BZK tries to have a horizontal approach and discuss problems and look at solutions together with the stakeholders.

No formal evaluation has been proposed, but there will be a letter to parliament on the progress of the data agenda, probably in March 2020. This will contain an analysis of what has been realised and what's lagging behind.

1.1.9 Successes, challenges and lessons learnt

1.1.9.1 Key successes

- *Data on the agenda*

  The topic of data and the urgency of sound data management across the public sector have gotten a more prominent place on both the political and administrative agenda. Questions like 'Do we use our data fully?' and 'Are we careful enough when using data?' are now on the table.

- *Knowledge exchange on best practices: the added value of being data-driven*

  The social challenges building block of the agenda has shown that data allows the government to move from a reactive approach to a proactive and anticipatory one:

  a. Energy transition - Using data from Statistics Netherlands (CBS), municipalities, housing corporations and energy companies a data model is developed to predict which neighbourhoods are most eligible to make the transition from fossil fuels to renewables.
  b. Public order and safety - On King's Day[40], the City of Amsterdam had real-time insight in people's whereabouts through the analysis of mobile phone data, allowing the city to take preventive measures against altercations, such as blocking a street.
  c. Poverty and the issue of debt - Instead of helping a person once (s)he has a EUR 40,000 debt, costing about EUR 100,000, a municipality can approach them before they get into the situation. This prevents a lot of suffering for the people at risk of getting into trouble and is much cheaper.

- *Knowledge exchange on best practices: how to be data-driven*

  Since the introduction of the data agenda, knowledge exchange has been taken to a higher level. More and more organisations know where to go with their questions on starting a data-driven policy project (the LED) and have been brought in touch with each other in case of similar experiences and challenges. Knowledge communities have been created, for instance between a number of municipalities using data to tackle the social challenge of poverty and debts.

1.1.9.2 Key challenges

- *Resistance to data sharing*

  This is due to a feeling of ownership of the data, shyness about the data quality and/or uncertainty on how the data can be shared technically.

- *Low data quality*

[40] This is the King's birthday on 27 April, which is a national holiday in the Netherlands. Traditionally, a lot of people go out into the streets for all kinds of festivities.

This only becomes apparent once data is being used for other purposes (secondary use).

- *Organisational obstacles*

  The infrastructure and people in organisations aren't always ready to work with data scientists. Having a proper work spot with the necessary software and access rights to data is a big challenge. Also, policy makers and data scientists often don't speak the same language, which can result in on the one hand policy makers not knowing what they can ask the data scientists to do nor understanding how they can act on data insights, and on the other data scientist not sufficiently understanding the policy context they work in to provide actionable insights.

- *Uncertainty leading to cautiousness*

  Because the interpretation of GDPR is not always clear, several organisations exhibit cautiousness on going ahead with a data project.

1.1.9.3 Lessons learnt

- It is important to connect different data communities, so they know how to advance and why they want to advance.
- The organisational response to data insights matters. Accompanying different municipalities in using data analytics for debt prevention, showed that policy effectiveness depends not just on the quality of the analysis, but also on the type of actions taken based on the data insights. An aggressive and intimidating communication revealing privacy-sensitive information on family members, such as drug use, may have a deterrent effect, leading families to lose trust in public authorities instead of accepting help. A more compassionate way of communicating with a family at risk led to them being more open to social assistance.
- When reusing data, government should take into account both the purpose for which the dataset was originally created (primary use) and the unintended side effects of its secondary use. For instance, personal data may have to be anonymised in one policy context and not another.
- Ethical data standards aren't universal. The research among millennials illustrates that there are generational insights with regards to data ethics.

1.2 Case study - Barcelona Data Commons[41]

1.2.1 Introduction

Counting more than 5.6 million inhabitants of which over 1.6 million for the city itself, Barcelona is the sixth most populous urban area in the European Union after Paris, the Ruhr area, Madrid, and Milan.[42][43]

The 2014 European Capital of Innovation[44] has continued its journey beyond the concept of smart city integrating the values of openness, fairness, circularity and democracy into its digital strategy.[45]The development of municipal data policies centred around data sovereignty is at the core of this Barcelona Digital City Plan, as this is considered crucial to put citizens' needs at the centre of public policy.

This case study report discusses the governance arrangements underpinning Barcelona City Council's approach to ethical and responsible data management, known as the **Barcelona Data Commons**[46][47]. It also sets out how the city council has dealt with key issues regarding data quality and sharing, the safeguarding of public values, human capital and skills, and monitoring and evaluation. Finally, key successes, challenges and lessons learnt are identified.

1.2.2 Policy context

In 2011, the Barcelona City Council put forward a new IT strategy focused on improving the overall operation and management of the city, fostering economic growth and strengthening citizens' welfare (Ferrer, 2017). The engagement of public and private stakeholders, the creation of new alliances and the promotion of lasting connections was one of the strengths of the smart city strategy, laying an important foundation for future policies. While the 2011 strategy recognises that "Smart government is about citizen participation and even more about citizen engagement" (Ferrer, 2017), it is only after the 2015 elections and with the arrival of a new Chief Technology and Digital Innovation Officer (CTID) that the city council would go a step further and start experimenting with identifying citizens' needs through data and at the same time granting citizens more agency regarding the use of their data.

> "*We are reversing the smart city paradigm. Instead of starting from technology and extracting all the data we can before thinking about how to use it, we started aligning the tech agenda with the agenda of the city*."
>
> Francesca Bria, Chief Technology and Digital Innovation Officer for the City of Barcelona, 2015-19[48]

[41] This case description is based on the findings from desk research, a presentation and two stakeholder interviews. It is illustrated by citations from interviews with and/or presentations by stakeholders. These citations aren't attributed to any specific stakeholder out of considerations of anonymity and confidentiality. The case study report has been submitted to the involved stakeholders for factchecking.

[42] https://www.ine.es/jaxiT3/Datos.htm?t=2861 (Accessed on 17 February 2020).

[43] Demographia World Urban Areas: 2019 http://www.demographia.com/db-worldua.pdf (Accessed on 13 February 2020)

[44] https://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/prizes/icapital/icapital2014_en (accessed on 17 February 2020)

[45] https://ajuntament.barcelona.cat/digital/en/blog/publication-of-the-balance-for-the-barcelona-digital-city-plan-2015-2019 (accessed on 17 February 2020)

[46] https://ajuntament.barcelona.cat/digital/en/blog/ethical-and-responsible-data-management-barcelona-data-commons (accessed on 18 February 2020)

[47] The English version of the official government measure announcing this strategy, the 'Government measure concerning ethical management and accountable data: Barcelona Data Commons', uses the names 'responsible and ethical use of data strategy', 'strategy of responsible data management' and 'strategy for the responsible and ethical use of data' to refer to what is called Data Commons Strategy in the remainder of this case study report.

[48] Graham: 2018

1.2.2.1 From a smart city to a citizen-driven city

By joining the *Decentralised Citizens Owned Data Ecosystem* project (DECODE), Barcelona has been able to make the first strides in applying the idea of a data as a common good and building an effective data commons ecosystem. The DECODE project 'provides tools that put individuals in control of whether they keep their personal information private or share it for the public good'.[49]This three-year project runs from 2017 to 2019 and is funded through the European Union's Horizon 2020 Programme. Not only is Barcelona one of the two pilot cities, together with the city of Amsterdam, where newly developed technology and operating system would be tested, but the Technology and Digital Innovation Office also takes care of project coordination and management. The collaboration of the City Council has allowed to test the feasibility of the proposed solutions into real world scenarios, making a strong impact regarding innovative data policies and approaches tested by cities.

In parallel to running the DECODE pilots, the Barcelona policy on municipal data governance is further formalised. In 2018, the **Responsible and Ethical Use of Data Strategy: Data Commons** is launched.[50]While acknowledging its status of leading city when it comes to data management, the city announces it wants "to go further and make a qualitative and quantitative leap".[51] This ambition stems from an awareness that in the city big data are produced in real time and that the way in which the city responds to this data revolution influences its effects on wealth, inequality and the city council's responsiveness to citizens' needs.

The scope of the Data Commons Strategy is defined in terms of *organisations and systems*, *data* and *environment and ecosystem*.[52] Regarding organisations and systems, it applies to the entire municipal executive structure and its dependent bodies. The latter are external agencies, which are separate legal entities holding information on the city, but find themselves under municipal control, although the City Council does not directly manage them (transport, IT management, economic promotion, etc.). An example is the Municipal Institute for Informatics (IMI). They are responsible for putting the actions and requirements arising from the strategy into practice. These are applicable to management activities that use any IT resources of any kind (unless expressly excluded) that collect, store, use or publish data, including information in paper and electronic mediums.

When referring to municipal data, this not only concerns the data itself, but also metadata and the data-management systems. Five categories of data are distinguished:

1. Management, administrative and technical data;
2. Open data and sets of raw data and open metadata;
3. Official statistics;
4. Open Content (municipal documents);
5. External data.

Regarding the environment and ecosystem, the Data Commons Strategy indicates that its provisions:

- must be applied to all acquisition and technology and service management contracts that work with data;
- are extended to all City Council interactions with the "community" involved, i.e. all the possible reusers or proactive consumers (prosumers) of the data that the City Council makes public as open data: activists, data journalists, NGOs, foundations and associations in a non-economic environment, as well as business people,

[49] https://decodeproject.eu/what-decode (accessed on 20 February 2020)
[50] Ajuntament de Barcelona: 2018a
[51] https://ajuntament.barcelona.cat/digital/en/blog/ethical-and-responsible-data-management-barcelona-data-commons (accessed on 18 February 2020)
[52] Ajuntament de Barcelona: 2018a

emerging companies, companies related to information sciences and major companies in the business world, and finally, academic research institutes.

The strategy initially covers a period of two years: 2018-19, up to the municipal elections of May 2019. It is considered as a continually evolving strategy, in line with the dynamic character of its mother policy, the Digital Transformation Plan (DTP) under the umbrella of the Digital City Plan. Consequently, its general provisions will be updated and put into practice as the programme is applied.

1.2.3 Building blocks of the strategy

1.2.3.1 Main themes

The Data Commons Strategy functions as[53]:

- a roadmap designed by the Commission for Technology and Digital Innovation and its policies for modernising data management and use within Barcelona City Council;
- a conceptual framework for achieving a cultural change in terms of the public perception of data;
- a plan that involves various programmes, covering all the aspects of the DTP that concern data, and in particular the open-data and data-commons strategies, data-driven projects, with the aim of providing better urban services and interoperability based on metadata schemes and open-data formats, permanent access and data use and reuse, with the minimum possible legal, economic and technological barriers within current legislation.

The overall goal of the **Data Commons Strategy** is to:

> "ensure the City Council is able to establish a general infrastructure and harmonised processes in order to manage, use and (partially) disseminate data, which is known as data commons, and that it is also able to promote, implement and supervise projects that capitalise on and enhance data in an agile, uniform way (e.g. in terms of data interoperability and standardisation) through the Municipal Data Office, and ensure the responsible management of data in accordance with current legislation, together with the Data Protection Officer".[54]

> "*Data becomes an infrastructure of interest to the public; a shared resource for the common good*."[55]

The main premise underpinning the Barcelona City Council's approach of data governance is understanding the value of data as a common asset and putting citizens back in control, hence 'data commons' as the key focus point of the ethical and responsible data strategy.

The five complementary pillars of the City Council's approach demonstrate a clear focus on citizens' rights, agency and value creation:[56][57]

1. Understanding data as an urban infrastructure, just as the provision of water and energy are. Data are seen as a meta-utility that will enable the city of Barcelona to support more effective delivery of public services to Barcelona citizens for greater equity, safety and quality of life;
2. Integrating the use of Big Data & Data Analytics to improve public decision-making (Data-driven projects).

[53] Ajuntament de Barcelona: 2018a
[54] Ajuntament de Barcelona, 2018a: 6
[55] https://ajuntament.barcelona.cat/digital/en/digital-transformation/city-data-commons/municipal-data-office (accessed on 12 February 2020)
[56] https://ajuntament.barcelona.cat/digital/en/blog/ethical-and-responsible-data-management-barcelona-data-commons (accessed on 12 February 2020)
[57] Balcells: 2020

3. Treating data as a common asset, and making it available for social and economic innovation processes focused on citizens' needs. This also means that the immense economic value that citizen-produced data represents should be returned back to those that generate that value in the first place: the citizens.[58]
4. Enforcing data & algorithmic transparency (Data Ethics). This not only requires opening up data, but also encouraging the reuse, providing citizens with the tools and knowledge to be able to verify these, and to be informed about automated decisions and their underlying algorithms.
5. Protecting people's privacy and data sovereignty. This is also about shifting agency and control to citizens themselves that have the right to decide what data they want to share, with whom and on what terms.

> "*Our approach is to create communal rights to data, and treat data as a common good. There's a lot of public value that can be unlocked out of this data*."
>
> Francesca Bria, Chief Technology and Digital Innovation Officer for the City of Barcelona, 2015-19[59]

An example of this approach can be found in the data sovereignty clause, which is being developed for city service procurement contracts. This clause obliges city contractors to give data back to the city. While the main legal document is in progress, it has already been tested in a couple of contracts. For example, in the case of the bike sharing system Bicing, usage data collected through the contractor's application is analysed in an ethical way by the city in order to better share the bikes and optimise vehicle transit.[60]

1.2.3.2 Connections to other strategies

The Data Commons Strategy is rolled out under the Digital Transformation Plan (DTP). The DTP is one of the three pillars of the 2016 Digital City Plan 2017-20 (Barcelona Ciutat Digital) [61][62], along with the *digital innovation and the revitalisation of the innovation ecosystem* and the *digital empowerment of citizens*. The DTP focuses on the digital transformation of both the city and the city council itself. The responsible and ethical use of data is a key element of Barcelona City Council's Digital Transformation Plan, especially the Open Data and Data Commons strategies, data-driven projects and interoperability based on open-data formats.[63]

1.2.3.3 Legal framework

Two national laws, one regional law and two city council provisions are especially relevant for Barcelona's Data Commons Strategy.

1. The **Common Administrative Procedure of Public Administrations** Law (Ley 39/2015)[64] stating that by 2020, digital channels must take priority in the provision of public services in Spain. The Barcelona Data Commons Strategy aims to contribute to improving digital public services and adapting them to citizens' needs as announced in the Digital City Plan.
2. The **Legal Regime of the Public Sector** Law (Ley 40/2015)[65], which establishes, in accordance with the EU once-only principle, that public administrations

[58] Decode: 2020
[59] https://sifted.eu/articles/barcelonas-robin-hood-of-data-francesca-bria/ (accessed on 13 February 2020)
[60] Ajuntament de Barcelona, 2019a
[61] Ajuntament de Barcelona: 2016a (English summary); Ajuntament de Barcelona, 2016b (Full document in Catalan)
[62] https://ajuntament.barcelona.cat/economiatreball/en/technology-and-digital-innovation (accessed on 17 February 2020)
[63] https://www.barcelona.cat/digitalstandards/en/data-management/0.1/summary (accessed on 18 February 2020)
[64] https://boe.es/diario_boe/txt.php?id=BOE-A-2015-10565 (accessed on 19 February 2020)
[65] https://boe.es/diario_boe/txt.php?id=BOE-A-2015-10566 (accessed on 19 February 2020)

a. will interact with each other and with linked or dependent public organisations through electronic means that ensure the interoperability and security of the systems and solutions adopted by each of them,
b. will guarantee the protection of personal data, and
c. will facilitate with preference the joint provision of services to interested parties.

3. The **Catalan Law on Transparency, Access to Public Information and Good Governance** (LLEI 19/2014)[66], granting citizens the right to know and request public information, provides the legal basis for Barcelona's open data activities as part of the Data Commons Strategy.
4. The **Transition Towards Technological Sovereignty** city government measure (2016)[67], approving the digital city plan 2017-20 and announcing that public services must be provided through digital channels from the outset, following new guidelines based on citizen guidance and the use of open standards and open software and in accordance with an ethical data strategy that focuses on privacy, transparency and digital rights.
5. The City of Barcelona's **Government Committee Decree 542/18**[68] on the approval of the municipal data governance model, the municipal data office and the data protection officer

1.2.3.4 Funding

Through the Barcelona Digital City Plan (2017-19), the city council has invested EUR 75 million annually in the Digital Transformation Plan.[69] Originally, it had budgeted EUR 65 million for the digital transformation and city infrastructures, EUR 10 million for innovative public procurement, and EUR 10 million for the Digital Technology and Innovation Office.[70]

Parts of this budget, along with funding by individual city departments, have been used to finance the flagship projects of the Data Commons Strategy. The sum of the different relevant budgets for the period 2018-19 amounts to EUR 3,807,000.

Project	**Budget (EUR)**
The Municipal Data Office	680,000
City OS Services	1.132,000
BCN Open Data	56,000
DECODE	545,000
Data City Challenge	180,000
BCN Data Exchange	130,000
Municipal Management Dashboard	235,000
Monitoring gentrification	*Own resources*

66 http://portaldogc.gencat.cat/utilsEADOP/PDF/6780/1395384.pdf (accessed on 17 February 2020)
67 https://ajuntament.barcelona.cat/digital/sites/default/files/pla_ciutat_digital_mdgovern.pdf (accessed on 19 February 2020)
68 https://bcnroc.ajuntament.barcelona.cat/jspui/bitstream/11703/108746/1/GM_DC_542-18.pdf (accessed on 19 February 2020)
69 Ajuntament de Barcelona, 2019a
70 Ajuntament de Barcelona, 2016b: 38

Census/EIAE	500,000
Housing Observatory	349,000

Table 1 Barcelona Data Commons Budget 2018-19[71]

For the activities started or continuing since mid-2019, the budget for the OMD's personnel and recurring annual contracts is part of the city's general management budget. Costs for technological investments for new hardware and maintenance of the city's data lake are shared between the general management's office and the IMI. Funding for specific analytics projects continues to be realised as part of individual city department budgets. No overarching budget for the Data Commons Strategy is reported, but the OMD is aware of the different budget lines available.

1.2.4 Governance of the implementation process

1.2.4.1 The central coordinating department

One of the key actions developed in the Data Commons Strategy is the creation of a **Municipal Data Office** (OMD – Oficina Municipal de Dades), headed by the **Chief Data Officer** (CDO). Following the example of cities such as New York, Boston, Paris and London, Barcelona is the first Spanish city to set up a Municipal Data Office. This body is created "to ensure that data is managed as a real asset during its entire life cycle, which will generate added value".[72] The City Council chose to first appoint a new CDO and then develop the new data commons strategy, so the new CDO's vision would be reflected in the strategy. The council decided to appoint an internal candidate with a good social sensitivity regarding data issues and a proven track record of managing a new structure within the City Council.

> "*There are very few City CDOs, so we're still learning what this role entails*."

The OMD is responsible for the management, quality, governance and use of data controlled and/or stored by Barcelona City Council and all of its associated bodies (both public and private) that provide services to the general public. The OMD represents a transition from a more traditional statistics office and view on data to a transversal office centred on the approach of data as a common good. This transition includes an organisational repositioning of data-related topics, as the OMD is not part of the IT-department but is placed at the top executive level in the office of the city's CEO.

The OMD unites a number of directorates and departments that was each in their own way involved in council data management and use, such as the Statistics Department. As such, it centralises several aspects of data life cycle management. It fulfils both the role of coordinator and implementer of the Data Commons Strategy. As a coordinator, the OMD promotes the necessary change in organisational culture, and facilitates and oversees appropriate data processing throughout the organisation, including liaising with relevant city players from the data sector. As implementer, it is actively involved in data processing itself through the execution and supervision of data exploitation projects needed by City Council departments.

> "*The topic of data governance and science is not a technological one, it's about knowledge on getting value out of data*. […] *We're on a journey from a more classic knowledge production system to a new environment, which is more real-time, big data and predictive.*"

[71] Ajuntament de Barcelona, 2018a
[72] Ajuntament de Barcelona, 2018a: 11

Originally, the OMD was politically dependent on the Chief Technology and Digital Innovation Officer. This is no longer the case since the political change in mid-2019 and the appointment of a new Commissioner for Digital Innovation, E-Government and Good Governance, having a different profile than the former CTID. Since the start of 2020, the OMD's political responsibility lies with the Councillor's Office for Presidency, situated within the mayor's office.[73] Through his former executive position of municipal manager (city CEO), the new councillor is familiar with the OMD's mandate as he has been part of its journey from the beginning. The OMD's position in the executive structure of the city council has not changed and remains at the top level under the Municipal Manager's Office, thereby ensuring its empowerment as a key organisational entity for the whole city council.

The OMD is staffed by around 40 people, making up a multi-disciplinary team of administrative, technical, analytic, graphic and business specialists. It is organised in three units totalling four departments:[74]

1. The Department of Data Integration
2. The Department of Analytics
3. The Department of Opinion Studies
4. The Department of Statistics and Data Dissemination

Another key position which is created through the Data Commons Strategy is that of **Data Protection Officer** (DPO), in line with the new European data protection regulations. The delegate's role is to guarantee citizens' data security and privacy. Several governance boards have been set up to ensure the coordination between the OMD and the DPO is ensured.

1.2.4.2 Decision and coordination boards

To enable transversal decision-making and coordination regarding data governance in the City of Barcelona, in September 2018, three boards were created through mayoral decrees:

1. The **Executive Data Committee** (CED - Comité Executiu de Dades), consists of the Councillor for the Presidency and Strategic Planning (president), the Commissioner for Digital Innovation, E-Government and Good Governance (vice-president), the Municipal Manager (vice-president), the Innovation and Digital Transformation Manager, the Director of the Municipal Data Office (CDO), the Manager of the Municipal Institute of Information Technology (IMI), The DPO and the Executive Secretary of Municipal Management.[75] The CED convenes twice a year and fulfils the following functions:[76]
    a. To approve the actions regarding the data governance of the Corporation.
    b. Determine the work to be carried out directly by the Municipal Data Office and to set its strategic and tactical criteria.
    c. To approve the priorities for the development and definition of data sets, access rights and municipal data use, as well as the high-level requirements for the data infrastructure.
    d. To approve, at the proposal of the Data Protection Officer (DPO), the corporate data protection policies.
    e. Approve the agenda of the Transversal Data Coordination Board.
    f. To monitor the execution of the agreements of the Committee itself.

[73] For more information on the Barcelona City Council's political and executive structure, see https://ajuntament.barcelona.cat/en/municipal-organisation/government-bodies (accessed on 9 March 2020).
[74] Bria: 2018
[75] Ajuntament de Barcelona: 2019b
[76] Ajuntament de Barcelona: 2018b

2. The **Transversal Data Coordination Board** (TCTD - Taula de Coordinació Transversal de Dades) consists of the Municipal Manager (president), the CDO (vice-president), the Executive Secretary of Municipal Management (vice-president), the DPO and representatives of 21 municipal policy departments. So far, it can count on a good representation of stakeholders. The TCDT usually convenes once every four months, but can have extra meetings at the request of a minimum of three members. Its main aim is to coordinate cross-departmental projects and to act as a means of disseminating the projects. The TCTD fulfils the following functions:[77]
    a. The establishment, application and follow-up of the guidelines on the work carried out so much in the OMD as in the various sectoral centres for municipal data storage, analysis and dissemination.
    b. The establishment of operational and restoring criteria to guarantee all aspects of the data life cycle.
    c. Monitoring of cross-cutting and corporate data projects.
    d. The coordination with the Data Protection Officer (DPO) of the aspects related to the privacy management.
    e. Compliance with the rules of intellectual property and authorship of the contents and documents with data.

3. The **Data Protection Board** (TPD - Taula de Protecció de Dades) consists of the Resources Manager (president), the DPO (vice-president), the CDO and representatives of 14 municipal departments dealing with sensitive data and/or security policy. The TPD minimally convenes once every four months, but can have extra meetings at the request of a minimum of three members. The TCTD fulfils the following functions:[78]
    a. Validate proposals for data protection policies at the municipal level, and their coordination, in collaboration with the Data Protection Officer.
    b. Inform proposals for internal regulations that guarantee the correct implementation of the approved data protection policies, as well as those that guarantee the compliance with the provisions of the European Data Protection Regulation and other specific regulations.
    c. Validate models of informative, consent and contract clauses applicable to different processing of personal data.
    d. Participate, when required, in the various municipal projects supporting the field of privacy of personal data.
    e. Collect the queries and doubts of the different municipal areas and analyse or send them to the competent municipal bodies for their resolution.
    f. Establish internal training plans in the field of personal data protection in coordination with the Data Protection Officer.

1.2.4.3 Fostering collaboration between stakeholders

1.2.4.3.1 Co-shaping the data commons strategy

The Barcelona Digital City Plan, which has laid the basis for the development of the Data Commons Strategy, was co-created with the city's innovation ecosystem for rethinking the smart city to serve its citizens.[79] A large-scale participatory experiment was powered by a digital participatory platform, DECIDIM, to tap into the collective intelligence of citizens to create policies that better respond to their needs.

[77] Ajuntament de Barcelona: 2018c
[78] Ajuntament de Barcelona: 2018d
[79] Ajuntament de Barcelona: 2019a

"*If you don't do it with the machine, with the public servants, it's just beautiful ideas which you won't be able to implement. You have to do it with them*."

Francesca Bria, Chief Technology and Digital Innovation Officer for the City of Barcelona, 2015-19[80]

The City Council also involved internal stakeholders to get them on board in treating data as a common good and public asset. The CTID, who was leading the development of the data commons strategy, mainly consulted stakeholders at the local level for their input on setting the main goals of the data sovereignty framework. The Commissionate later took teams from all departments—legal, procurement, HR, technology—through a six-month transformation process, which included various workshops encouraging leadership and ownership on data management and use. The CTID also discussed with various international stakeholders.

1.2.4.3.2 Executive legitimacy, communication and support in data processing and analytics as policy levers

The creation of the OMD and the positions of CDO and DPO functions as an important policy lever for the implementation of the data commons strategy. Through a Municipal Decree of the Governance Committee, the mandates of these new institutional entities have been legally determined. The OMDs legitimacy vis-à-vis other municipal bodies and its continuity is affirmed through its positioning in the Municipal Manager's office, which finds itself at the top of the executive structure and is responsible for the provision of the principal services to citizens. For proper coordination between the various areas of municipal management, data officers are to be defined for each sector, who will be responsible for the data in their own manager's office and for coordination with the council as a whole.

The OMD also has a central role in the promotion, execution and supervision of municipal data analytics projects. It also provides hands-on support to municipal bodies by developing standardised protocols for data processing throughout the organisation and executes a precise mapping of all the existing wealth of data under the umbrella of the City Council, creating a classification for all the data and the promotion of appropriate processing, care and preservation policies for all bodies involved. In doing so, it generates information that helps to resolve the challenges facing the city and functions as a service provider to the various municipal units.

"*It is very difficult for municipal departments to understand what the OMD can mean for their service to the organisation, <u>what</u> question* they *can ask <u>us</u>*. […] *We need to teach them what kind of capacities we have regarding all these new analytical possibilities*"

Producing concrete data-driven outputs has proven to be crucial to guarantee the individual departments' support for and continued involvement in the data commons strategy. For instance, through private-public data sharing, including data from a telecom company the OMD managed to measure the city's beat, to be used for infrastructural purposes. The OMD's efforts are aimed at closely accompanying municipal units during a data analytics project. While the City Council still has a long way to go, the ambition of

[80] https://sifted.eu/articles/barcelonas-robin-hood-of-data-francesca-bria/ (accessed on 24 February 2020)

the OMD is to make sure that the involved personnel understands the change from traditional statistics to data-driven working, is adequately trained and develops the necessary skills to execute such a project by itself a next time.

1.2.5 Data quality and sharing

1.2.5.1 Access to data: the open data portal and the municipal dashboard

Barcelona Municipal Data are increasingly publicly and freely accessible through the open data portal and a new municipal data website. At the same time, a municipal management dashboard is available as an internal city council data dashboard and visualisation tool.

Since 2011, the Barcelona City Council has had a repository of open data, **the Open Data BCN portal**. The data commons initiative aims to make the portal more useful for developers and the general public, as the portal had low standards and there was a lack of decisive political support for opening up data. One of the milestones of the Digital City Plan was the complete reshaping of the portal by February 2017. It now includes more than 462 data sets on the population, health, economy and education, among many others, which can be found in formats that can be reused and downloaded.[81] Barcelona has also opened up data that wasn't collected by the city itself, but which is considered of public value, such as bike sharing data collected by a private contractor.

In addition, a **Municipal Data Website** is being developed, which will provide services to both citizens and city council users. It will give access to three levels of data: raw data, treated data and data-based knowledge through visualisations and reports. For example, on unemployment in the city of Barcelona, at the first level, the entire raw open data set can be downloaded. At the second level, the data could be presented in tables or visualisations per neighbourhood, age or gender. At the third level, a report can be published providing an analysis of the unemployment data in the wider economic context or in relation to other developments, such as the coronavirus.

The **Municipal Management Dashboard** is an application available to city council employees. It includes a data-visualisation tool developed by the Municipal Manager's Office showing real-time information about what is happening in the city in terms of housing, employment and care, coupled with indicators of how council management is progressing in implementing public policies and how citizens receive these actions.[82] Through this portal, data from several municipal sources can be integrated: Statistics BCN, Open Data BCN, BcnRoc, Barcelona Economia, Opinion studies and surveys, Geoportal BCN, CartoBCN and Observatory.[83]

The ongoing work of the OMD regarding the inventory and classification of data and the development of standard protocols for all municipal data is crucial to further improve data accessibility. Another critical element is the underlying data infrastructure.

1.2.5.2 Municipal data infrastructure: The City OS data lake

Through the flagship project for a Municipal Data Lake, the **City OS**, the City Council has worked towards a model of grouped data from derived from numerous sources and in heterogeneous formats in order to create common repositories for management, analysis and secure storage, thereby supporting the management of municipal data as a public asset throughout its a life cycle.

[81] https://opendata-ajuntament.barcelona.cat/en (accessed on 17 February 2020)
[82] Ajuntament de Barcelona, 2019a: 28
[83] Bria: 2018

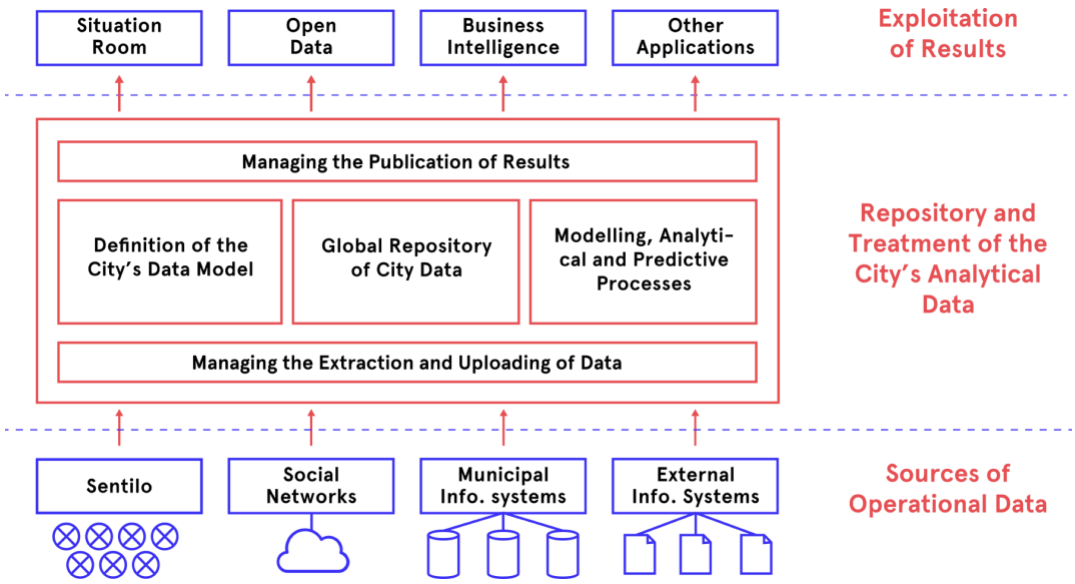

Figure 2 Barcelona City Data Lake[84]

The City OS is a data infrastructure based on open-code, big data and analytics technology for machine learning, which uses single-window access and internal data management. It constitutes a single analytic repository, a data lake, where the data input and consumption or access points are centralised. City OS works with both internal City Council data (adjudication of contracts, subsidies, Municipal Action Plan projects, districts, etc.) and with external agencies under municipal control which hold information on the city, although the City Council does not directly manage them (transport, energy, environment, etc.). The City Council states that this infrastructure provides better data governance, quality controls, more effective privacy and security through the single input and access points and, above all, gives the City Council an overview of this area.[85]

At the start of 2020, the main infrastructure is operational, containing around 500 different datasets. The next step for the OMD, is to make sure that the City OS is structurally fed with municipal data, based on the city's needs. Through a new action plan, the OMD will identify the priority data sets to be included and work together with the relevant city departments to make this happen. This is an incremental process of adding data on a case-by-case basis, which will strengthen the model's analytical power step by step. The city CEO supports this work, for example through announcements in the Transversal Data Coordination Board, thereby pushing the city departments to cooperate.

Sentilo is an open IoT sensor network, which transmits real-time data, which is accessible to city hall and private sector partners. Developed with open-code software, it measures energy, noise, rubbish, weather, parking areas, air quality, water levels and flows of bicycles, people and vehicles. Sentilo generates 3,000,000 records per day using about 1,800 devices and 15,000 sensors involving 30 sensor companies. These data concern the 1,62M city inhabitants, 15.000 lampposts, 36.000 traffic lights, 40.000 garbage containers and 80.000 public parking spots in the street.[86]

[84] Ajuntament de Barcelona, 2018a: 28
[85] https://ajuntament.barcelona.cat/digital/en/digital-transformation/city-data-commons/cityos (accessed on 14 February 2020)
[86] Bria: 2018; Ajuntament de Barcelona, 2019: 8

1.2.6 Safeguarding public values

1.2.6.1 Data Commons for data ethics

The Data Commons Strategy is one of the eight Barcelona **Ethical Digital Standards**, which are promoted through the Cities Coalition for Digital Rights[87] and made available to other cities and interested stakeholders as an open source policy toolkit[88]. These standards have been developed under the umbrella of the 2017 Open Digitalisation Plan for a more open, transparent and collaborative government and include the use of open-source software, open standards, data sovereignty, agile development of digital services and guaranteed privacy, ethics and security by design.[89] The council is a pioneer in this area and has committed to investing more than 70% of the new software development budget into free and open-source software and services based on open standards, open formats, open interfaces and interoperability.

The Data Commons Strategy explicitly mentions 14 values which the Barcelona City Council's management and use of data have to respect and comply with: shared municipal knowledge, the strategic value of data, geared towards results, data as a common asset, transparency and interoperability, reuse and open-source licenses, quality and security, responsible organisation, care throughout the data's life cycle, privacy "by design", security, technological sovereignty, open standards, agile methodology.[90]

By incorporating the aforementioned values into municipal data management and its governance mechanisms, the City Council aims to accomplish its overall vision of data as a common good and ensure data sovereignty.

> "*We believe that data is a public infrastructure—like water, like roads, like the air we breathe—and it should be treated as such. It belongs to the citizens of Barcelona*."
>
> Francesca Bria, Chief Technology and Digital Innovation Officer for the City of Barcelona, 2015-19[91]

The council defines data sovereignty as "the need for an individual to have control, at all times and in all relevant systems, over the collection, storage, use, transfer and publication of their data, whether it be of a technical, scientific, economic, social or personal nature".[92]As such, Barcelona's approach focuses on public leadership and value, privacy and transparency.[93]The focus on individual control, however, might raise the question whether the concept of data sovereignty provides an adequate answer to questions of social revenue, common good and the social and ethical impact of technology. The focus on citizens' individual decision-making capacity forgets the collective dimension of the rights at stake, such as privacy, and opens the door to terribly harmful data relationships.[94]

At the same time, the council not only considers citizens' data sovereignty, but also that of the council itself, as it want to guarantee "the City Council's sovereignty over its data, and in particular, over the (personal) data of the city's residents, which is administered by

[87] https://citiesfordigitalrights.org/barcelona (accessed on 14 February 2020)
[88] https://www.barcelona.cat/digitalstandards/en/init/0.1/index.html (accessed on 14 February 2020)
[89] Ajuntament de Barcelona: 2019a
[90] Ajuntament de Barcelona, 2018a: 13
[91] https://sifted.eu/articles/barcelonas-robin-hood-of-data-francesca-bria/ (accessed on 13 February 2020)
[92] Ajuntament de Barcelona, 2018a: 15
[93] Galdon Clavell: 2017
[94] Galdon Clavell: 2017

the City Council for providing its services. The City Council has to have the ability, at all times, to access its data, process it and make "local" security copies."[95]

1.2.6.2 GDPR compliance

As part of its mandate, the OMD oversees responsible data management, in accordance with current legislation concerning privacy, through coordination with the data protection officer (DPO). From the outset, the DPO was placed outside the OMD in order to be able to function in an independent way, like an external auditor. The Data Commons Strategy prescribes that subsidiary municipal bodies establish and maintain a data-protection programme, in accordance with the municipal policy on privacy.[96] They're required to appoint a data protection officer (sub-DPO) to ensure compliance with all regulations, norms and directives that affect the life cycle of personal data, by means of a privacy policy that applies legal measures to privacy risks. The DPO has authority over the sub-DPOs in each organisation and will coordinate them. Privacy impact assessments, privacy by design strategies and privacy by design technologies are key elements of the City Council's privacy strategy.

1.2.6.3 Algorithmic transparency and accountability

Algorithmic accountability is supported by the transparency of the City Council's open source code IT systems supporting (automated) decision-making.[97] Wherever possible, data-driven projects will be able to check the algorithms using simulations based on city data. Likewise, using open source code or other means, third-party technology suppliers must reveal the underlying logic behind any IT process for (automated) decisions pertaining to any of their systems used by the City Council.

An OMD working group is dedicated to identifying the issues raised by algorithmic determinism relating to automated decisions taken by the City Council, and the necessary measures for ensuring the city council's ethical data principles.

1.2.6.4 Technological tools for data sovereignty

> "*We want to move from a model of surveillance capitalism, where data is opaque and not transparent, to a model where citizens themselves can own the data*."
>
> Francesca Bria, Chief Technology and Digital Innovation Officer for the City of Barcelona, 2015-19[98]

One way in which the City Council has concretised the idea of data sovereignty is through its participation in three pilots testing DECODE technology, enabling citizens to decide what kind of data they want to keep private, what data they want to share, with whom, on what basis, and for what purpose.[99]

1. Over the course of one year, 228 people participate in the Digital Democracy and Data Commons (DDDC) pilot to improve the digital participation platform DECIDIM by improving the user's control over their data as well as the transparency in citizen petitions. Moreover, the process has the goal of discussing alternative visions, networks and practices on citizens' data.
2. The Citizen Science Data Governance (IoT) pilot enables communities to support IoT data gathering and allow them to control what information is shared with

[95] Ajuntament de Barcelona, 2018a: 15
[96] Ajuntament de Barcelona, 2018a: 23
[97] Ajuntament de Barcelona: 2018a
[98] Graham: 2018
[99] Decode: 2019

whom, and under which conditions. This pilot involves more than 1.7 million readings from 24 sensors, generating 3.2 million data points over a period of 93 days impacting 500 people.

3. In the BarcelonaNow pilot[100], a platform for empowering citizens with interactive dashboards and facilitating them the exploration of urban data from different sources is tested: crowdsourced from citizens, open data from the city and data gathered from other DECODE services. The platform has 21 datasets used in 5 open data dashboards and 6 data commons dashboards, generating over 200 visits per month.

The next challenge for Barcelona is to consider under what conditions these projects can be scaled up, so the whole citizenry of Barcelona can benefit. The pilots showed that work remains to be done in terms of technology readiness level and user-friendliness of the DECODE technology.

1.2.6.5 Data clauses in procurement contracts

The idea of data sovereignty through increased transparency is further concretised by integrating new "data sovereignty" clauses into public procurement contracts. These clauses mandate city providers, such as the garbage bin collector and the public biking provider, to give back the data they gather to deliver the service to the city hall in machine readable format. Together with the Data Protection Office, the OMD is developing an instruction manual for the city departments on how to integrate these data clauses in their contracts, thereby ensuring that data are provided in the correct format and in alignment with GDPR. In this way, the city council takes its data commons approach a step further by backing their communication efforts with legal provisions.

The city council not only publishes the anonymised data as open data, but also generates added value by integrating the data into the City OS data lake and making it available for data-driven policy-making. In this way, data are transformed from a private into a public good and placed in the public domain, while safeguarding privacy, ethics and security by design via the use of strong cryptography. An example is the contract that the city has with a big telecom provider, who has to share machine-readable data with the city on a monthly basis, whereas before, the company kept it for its own benefit.[101]

1.2.7 Human capital and skills

> *"Our primary challenge is about having the people who are capable of working with these new data, with these new technologies."*

Both at the central and decentral city levels, activities are undertaken to attract new talent and make sure that Barcelona public servants acquire the necessary data skills. This remains one of the biggest challenges for the Barcelona City Council for proper data governance.

> *"It's not as easy as in the private sector where you put a profile out, someone sends a CV, you interview and that's it. [...] Public servants have their own track to be appointed for which you need the proper legal process to ensure transparency, equal opportunities and so on."*

Three lines of action have been set out to affront this challenge:

[100] http://bcnnow.decodeproject.eu/ (accessed on 20 February 2020)
[101] Graham: 2018

1. Develop a **data science profile** for Barcelona public servants and integrate it in the existing human resources structure. This has resulted in the appointment of one person so far. Subsidiary municipal bodies are responsible for developing a set of requirements and skills for the teams in charge of or working with data. Relevant aspects to be considered include access, analysis, privacy, security, and above all innovative capacity in a context of constant change. For the selection of new data personnel, they are required to coordinate with the CDO. Regarding personnel at the management/executive level dealing with data, the Department of Human Resources must ensure that they have sufficient knowledge to carry out their task appropriately.[102]
2. Make an **inventory of data enthusiasts** within the City Council, who are already working with data and independently learn about the most recent data developments. The OMD offers them training on specific tools to help them on their way to achieve a certain level of expertise in data management and analysis. At the same time, the data office explains to their managers that these tools will support them in decision-making and can improve the efficiency and efficacy in their respective areas of education, culture, health etc. The internal training and encouragement of City Council personnel is thus not only focused on understanding the methodology of a data-driven qualitative focus on city problems, but also seeing its value. Furthermore, the analytical services that the OMD provides are always accompanied by appropriate courses, ensuring the continuity of the data-driven approach.[103]
3. **Outsourcing data science projects** remains an important solution to the shortage of adequately skilled personnel within the city council. Besides the hands-on assistance, this also brings the opportunity of interaction with the private sector and being up-to-date about the latest developments. However, the OMD acknowledges that the core of data analysis has to be developed and executed in-house for the City Council to continue to be knowledgeable on data issues and remain a relevant stakeholder in the city data ecosystem.

In the broader context, Barcelona aims to let the city's digital innovation ecosystem grow.[104] Over the past years, it has organised several renowned events, such as the Mobile World Congress or the Smart City Expo World Congress, where public and private stakeholders can meet, thereby helping local talent to flourish and attracting new talent to the city.

1.2.8 Monitoring and evaluation

Both the Executive Data Committee (CED) and the Transversal Data Coordination Board (TCDT) play a crucial role in the monitoring and management of the strategy's progress and use of resources, as relevant stakeholders meet here to share project information and review overall developments. The CDO regularly coordinates with project managers in order to evaluate the IT resources and the data-management processes that support the aims and mission of each municipal body. During the CED meetings, the CDO gives an account of the status of projects and budgets to his political and management superiors.

No KPIs have been defined for the overall strategy. Monitoring of progress generally happens at the concrete project level. When supporting a city department in a new data project, the OMD defines the milestones together with the department in question. The main evaluation criterion is whether a change has been achieved with regard to the original policy problem, either by solving it or by gathering a new level of understanding of it through the application of data analytics. Furthermore, the Data Commons Strategy sets out that each organisation should have a single review policy for evaluating the investments, the analysis of operations and the other evaluation systems for IT resources, including the projects under development and ongoing activities.

[102] Ajuntament de Barcelona: 2018a
[103] Ajuntament de Barcelona: 2018a
[104] Ajuntament de Barcelona: 2019a

The cross-departmental coordination must lead to the proper management and review of resources and programmes, including the work carried out by the OMD at the central level and ensure coordinated and unified budget management. Both inter-departmental and council projects are monitored, as well as the management of data protection and the regulations concerning intellectual property and authorship protection for the data.

1.2.9 Successes, challenges and lessons learnt

1.2.9.1 Key successes

The Barcelona Data Commons vision started approximately one year before the formalisation of the strategy and led to the redefinition of several previously existing tools (e.g. the City OS data lake and the Open Data BCN portal). This approach presents several successes.

- *Predictive analytics to inform public policy*

  The city's treasury department has used the services of the OMD to help them use data analytics in the development of tax policy. Previous to the COVID-19 pandemic, several scenarios of tax policies were tested to see if they would favour or disadvantage different areas of the city depending on their economic capacity.

  Since several years, the city council has also been using predictive modelling to estimate the occupancy of the city beaches. This model takes into account several variables, with data coming from different city sources. More recently, the OMD has been working on integrating mobile phone data into the model to help verify its accuracy. This is expected to eliminate the need for manual counting and allow for a more detailed analytical model regarding the time frames and places covered.

- *Data for monitoring and business intelligence*

  The City Council managed to gather insight in the evolution of Barcelona's housing market, through an analysis of a wealth of data on topics coming from different city departments and associated organisations, such as house prices, gentrification, house movements and type of use (resident or touristic).

- *Realising a new municipal data infrastructure*

  Even though work remains to be done in terms of collecting data and feeding it into the City OS, this municipal data lake is a fact, providing the City Council with an analytical services platform that allows the integration, rapid and efficient and reliable correlation of information sources to boost the management and transparency of the city. After almost seven years since its initial conception, the City Council is now able to enjoy the benefits of the City OS: extracting data from it and performing data analytics.

- *Involvement of key stakeholders*

  Through the transversal data coordination board, presided by a representative of the municipal manager's office, key stakeholders from the different department areas and institutions have been gathered around the table to stay and continue to provide their input and collaboration in analytics projects.

- *Using private sector data for public purposes*

  A revolutionary step in opening up private data and reusing it for public policy purposes is in the process of being made through the development of a set of clauses and guidelines for tenders. Contractors will thereby be obliged to share the data they generate in providing a public service with the City Council, which will then make sure

the data are processed in a responsible and ethical way, before they're made available as open data if this is possible. Among other projects, the city council uses aggregated and anonymised mobile phone data from a telecom company to gather knowledge on mobility in Barcelona's surroundings (e.g. how many commuters there are in the metropolitan area).

- *Strengthening the city hall's position in the city data ecosystem*

  Through the different projects under the data commons strategy, the City Hall has managed to position itself as a "Data Player" in front of the private stakeholders.

- *International inspiration and facilitation*

  As part of the Cities Coalition for Digital Rights, a global Alliance supported by the United Nations, UCLG and Eurocities joined by more than 60 cities, the ethical digital standards developed in Barcelona are now shared amongst cities in an open source platform.[105]

- *Invigorating the data commons ecosystem*

  A local civil society network emerged from one of the DECODE pilots: the Barcelona Data Commons Network.[106][107] Its aim is to engage the organizations of Barcelona, with principles, such as 'our data, our rules' and 'common data for collective challenges'. The network consisting of citizens, social collectives, research groups, public bodies, cooperatives and SMEs, foundations and other related actors is likely to help spread the adoption and promotion of the City's data commons work.

1.2.9.2 Key challenges

- *Breaking internal silos*

  Evangelising a shared approach and common vision remains a challenge as municipal departments are protective regarding the specific data that they process and don't always see the added value of sharing nor how data from one policy field can be relevant for others. In addition, legacy IT systems sometimes impede breaking silos between the different areas and departments. Interoperability and data standards are therefore among the technological challenges faced by the city council.

- *Attracting and training public servants*

  There's still a deficit of adequately skilled personnel, which understands how to work with data. This points to both the challenge of attracting new talent and that of training current staff. At the OMD, the limited human resources sometimes force the office to say no to helping out city departments on data projects.

- *Increasing data awareness among city managers*

  A lack of knowledge and understanding among city managers of the potential value of data linking and analytics withholds certain departments from requesting the analytical services of the OMD and relevant data from other departments.

- *Data duplicity*

  As the city's data lake grows by incorporating more and more datasets from different sources, the risk of duplicating datasets grows as well.

- *Obtaining standardised private sector data*

  It remains a challenge to convince the bigger companies, such as utilities, mobility and real estate providers to share their data with the city for the common good and in the

[105] Decode: 2020
[106] Decode: 2020
[107] https://datacommons.barcelona (accessed on 26 February 2020)

right format. Private actors are reluctant to share the economic value of their data for free and their data aren't always interoperable with open government data sets.

- *Social awareness*

  The population of Barcelona is yet insufficiently aware of the data infrastructure that is being built for the common good and the role that they can play in this. This is needed to scale up the data commons technology as tested in the DECODE pilots to the level of the whole city.

1.2.9.3 Lessons learnt

- It has proven to be a strategic decision to place the OMD directly in the Municipal Manager's Office, at a high level in the executive municipal structure, and not as part of the IT-department. This positioning has helped to ensure that the OMD as well as its framework and models are empowered to be a key element within the organization.
- When setting up a new central data office, it is important to manage expectations of all stakeholders involved on the pace of the changes ahead. It needs to be clear that this is an incremental and complex process requiring the contribution of all individual departments.
- Through 'data sovereignty' clauses in public procurement contracts, data can be turned into a public good and placed in the public domain, while at the same time preserving privacy, ethics and security.
- Making the financial investment to set up the technological infrastructure is a necessary, but not sufficient precondition for successful data governance. Adequately skilled people who are capable to work with the infrastructure and the (big) data in it are necessary to get value out of the data.
- Even though technological challenges remain, data can be governed and managed as a common good, shifting agency and control to citizens themselves having the right to decide what data they want to share, with whom and on what terms. The terms and conditions for data access and sharing are in this way set by citizens themselves using the DECODE cryptography tools and smart contracts, after which it can be integrated with the existing city hall digital infrastructure.
- Data from the private sector hold great potential for several domains of the public administration, such as environmental policy, mobility and public safety. In order to facilitate access to these data in open formats, legally binding standardisation and interoperability criteria will be of key importance.

1.3 Case Study: New Zealand Data Strategy and Roadmap

1.3.1 Introduction

Stats NZ Tatauranga Aotearoa is New Zealand's official data agency. It is responsible for collecting information from people and organisations through censuses and surveys, publishing insights and data about New Zealand, and supporting others to use the data.

The States Services Commissioner has assigned the Chief Executive of Stats NZ the role of Government Chief Data Steward (GCDS) to support the government's priority to get more value from data.

The GCDS works cross-governmentally, leading by facilitating and enabling a joined-up approach to data, responding to new and emerging data issues, and ensuring that government agencies have the capability and right skills to maximise the value of data.

Cross-government initiatives have been a response to the need to adapt to the digital era, where data creation increases at an exponential rate and there is an increasing expectation and demand for real-time information.

As part of this cross-government role, in 2018 the GCDS produced a new **Data Strategy and Roadmap**. The roadmap was commissioned by the GCDS to provide a joined-up approach to the significant amount of data-related activity underway across government, enabling organisations to collectively work towards, and align their efforts towards generating the maximum impact through data. The strategy and roadmap is supported by a mature set of frameworks and guidance for government agencies, covering: data leadership, statistical capability, data standards, data stewardship, open data and data investment.[108]

1.3.2 Policy context

The New Zealand government has recognised the increased demand for better, more widespread use of data across New Zealand. To forward this agenda, the Chief Executive of Stats NZ was given the Functional Leadership role of Government Chief Data Steward to support the government's priority to get more value from data. At the heart of this work is transparency around how the government uses the data it holds on behalf of New Zealanders.

At the time, it was identified that agencies were operating in an ad hoc way, creating inconsistency in how data is:

- described and recorded
- managed, accessed and stored
- reused
- shared across government
- shared as open data (data that is freely available to be used, reused and redistributed by anyone for any purpose).

Consequently, the agency planned work across government to co-create a framework to support agencies to build capability, and to access and use data effectively and efficiently. This included a Data Strategy and Roadmap, a review of legislation, and a Data Capability Hub.[109]

[108] https://www.stats.govt.nz/about-us/data-leadership (accessed on 13 February 2020).
[109] Data Leadership overview Jan 2019

The Data Strategy and Roadmap was published in 2018; its priorities are furthered by the Data Stewardship Framework and the operational Data Governance Framework, both published in 2019.

These strategies underline a shared direction for New Zealand's data system, and a common language and understanding of data stewardship and data governance.

What is data stewardship?

Data stewardship is the careful and responsible collection, management, and use of data.

The New Zealand government believes that because government stewards and uses data on behalf of New Zealanders, it has a duty to ensure data is managed as a valuable asset and used ethically.

1.3.3 Building blocks of the agenda

1.3.3.1 The Data Strategy and Roadmap

The Data Strategy and Roadmap provides a shared direction for New Zealand's data system and outlines goals and actions which are expected to cover a 3-5 year view from date of publication. It has a broad scope, including the people and organisations that collect and use data; data itself; supporting components such as data access and capabilities; and enablers like public trust and protection mechanisms that make sure data within the system is used in a safe way. It is intended to be a living document that is updated regularly.

The roadmap identifies four focus areas, aimed to generate maximum impact, each with underlying goals. These are expected to be delivered over 3-5 years (from 2018).

1. Invest in making the right data available at the right time

The NZ government collects, holds and uses a broad range of data, and uses it to answer policy questions, design services and invest public money. Barriers to data use include visibility and availability of data, lack of clarity on access rights, inefficient processes for requesting data and charging mechanisms creating inequity of access for some organisations. In response, the government set the following goals:

   a. To provide visibility of key data sets and proactively address gaps
   b. To improve accessibility of government held data
   c. To open up more non-sensitive, non-confidential data to the public

2. Grow data capability and supporting good practice

Data capability refers not only to the technical specialist skills required to work with data, but also the processes, tools and practice that support data use. Data capability can be varied across the NZ government. Key challenges include a lack of data and analytics skills, a shortfall in capacity (not enough practitioners and competition for resources), decision makers without the right skills to use data in decision making, and a lack of

'translators' to bridge the gap between practitioners and decision makers. The roadmap sets two goals in response to these challenges:

a. To take a strategic and coordinated approach to uplifting capability across the public sector

b. To make better use of existing data capability.

3. Build partnerships within and outside government

Collaboration on data projects occurs, but often happens informally. The government sees an opportunity to take a deliberate approach to be more connected across the system, with increased collaboration fostering innovation. The government aims:

a. To co-design the future data system and work together to maximise use and impact of data.

b. To co-design with Māori across the data system

4. Implement open and transparent practices.

Public trust and confidence are seen as essential to allowing different uses, collection methods and enhanced benefits of data to be realised. This trust has to be earned and maintained through continued engagement with the public. The roadmap outlines priorities in building and maintaining this trust:

a. To establish appropriate accountabilities and protection mechanisms

b. To build public knowledge and understanding of how they can benefit from data use.

1.3.3.2 The Draft Data Stewardship Framework

The draft Data Stewardship Framework aligns with the Data Strategy and Roadmap. The objective of the framework is to help the government better manage and safely share New Zealand's data. It will be implemented through a data stewardship toolkit, which will help people and organisations to steward data effectively.

The toolkit provides access to data stewardship guidance, resources, and tools from across the data system. The framework and toolkit set out a multi-layer approach, identifying 3 key layers of the system with different needs, with tools and resources for these layers mapped against 8 elements of the framework.

The 3 key layers:

1. **Data system**: The people and organisations that collect, manage, and use data. Effective system stewardship will sustain a data system that protects privacy and confidentiality, maintains trust, and maximises the value of data.
2. **Organisations**: Effective data stewardship will ensure the data an organisation collects, stores, uses, and releases will be well managed, ethically used, and safely shared.
3. **Datasets and individuals:** Effective stewardship will ensure individuals have the appropriate capabilities to carefully and responsibly manage and use datasets.

The 8 elements of the framework[110]

110 https://data.govt.nz/manage-data/data-stewardship/a-draft-data-stewardship-framework-for-nz/ (Accessed 14 February 2020)

- **Strategy and culture** – a strategy that provides a shared vision and clear direction, and a data culture that enables strategy implementation and sustains good data stewardship practice.
- **Rules and settings** – legislation, policies, principles, and sanctions providing boundaries and guiding how the data system should operate.
- **Roles, responsibilities, accountabilities** – governance structures, role definitions and expectations, and leadership.
- **Common standards** – consistent ways of describing and recording data to make sharing, comparing, and re-using data more efficient.
- **Data capability and quality** – tools, processes, designs, metadata structures, and platforms for managing, storing, describing, and sharing data.
- **People capability and literacy** – skills, knowledge, and services for accessing, managing, analysing, and communicating data and insights.
- **Influence and advocacy** – effective relationships and networks to endorse, promote, and support good data practice.
- **Monitoring and assurance** – assessing environmental trends and developments, measuring stewardship performance, and adapting the stewardship framework to respond to changing circumstances or new information.

Layer		Strategy	Rules & settings	Roles, responsibilities, accountabilities	Common standards	Data capability and quality	People capability and literacy	Influence & advocacy	Monitoring & assurance
Data system / Organisation / Datasets and individuals	What's available	Data strategy and roadmap (GCDS)	Principles for the safe and effective use of data and analytics (OPC, SNZ) Privacy human rights and ethics framework (MSD) Open Data Charter (GCDS)	Data Investment Framework (GCDS)	Data content standards development process (GCDS)			Government Analytics Network Iwi Data Leaders Forum Government data champions network (GCDS)	Government algorithm transparency (GCDS,GCDO)
	What's being developed		New data and statistics legislation (SNZ) Privacy Bill (MoJ) Data protection and use policy (SIA)	Co-designing a Māori data governance approach (GCDS) Advisory group on trusted data use (GCDS)	Data content standards (GCDS) Education sector data standards (MoE)	Data quality framework (GCDS) National Health Information Platform (MoH)			Artificial intelligence action plan (MBIE)
Data system / Organisation / Datasets and individuals	What's available								
	What's being developed	Value proposition assessment tool (GCDS)		Implementing operational data governance (GCDS)		Steady states data flow modelling (GCDS)	Data and analytics consultancy (GCDS) Data capability framework (GCDS)	Establishing a data science campus (GCDS)	
Data system / Organisation / Datasets and individuals	What's available						Supporting iwi data access (SNZ)		
	What's being developed					Data de-identification tool (GCDS)	Redeveloped data.govt.nz (GCDO, GCDS)		

Figure 1: Draft Data Stewardship toolkit[111]

1.3.3.3 The operational Data Governance Framework

The operational Data Governance Framework (oDGF) aligns with the Data Strategy and Roadmap, enabling agencies to demonstrate data accountability and best practice data management behaviours amongst all staff in their organisation. Like other aspects of the Roadmap, the oDGF emphasises the importance of effective data stewardship as a basis for strengthening data capability. It was created in recognition that traditional data

[111] Draft Data Stewardship Framework and Toolkit, May 2019
https://data.govt.nz/assets/Uploads/Draft-data-stewardship-framework-and-toolkit.pdf

governance approaches, based on designated roles and a top-down hierarchy must adapt to meet the needs of a system in which there is increasing volume, movement and importance of data. Effective data governance means that accountability for data must sit with all staff who handle it, rather than being held through hierarchical and formal roles. This increases governance relevance, creates transparency, strengthens data quality, and supports customer trust.

The oDGF is built on two main components:[112]

1. Enabling agencies to cultivate a thorough knowledge of their data

    a. Promotion of comprehensive understanding of data assets emphasising data flow, through maps showing how data flows into, through and out of an agency.

    b. Use of steady states data flow model

    c. Implementation of scaled data flow maps, from line of business to enterprise

2. Helping them exercise proper care of that data

    a. Based on a set of 10 foundational data governance capabilities

    b. Implementation of a human resource competency framework that incorporates the ten data governance capabilities

    c. Embedding data accountability and best practice data management across all data-handling positions, with the goal of evolving beyond the need for traditional data governance roles

    d. Using resultant staff accountability to lift organisational data culture and maturity from the bottom up.

[112] Operational Data Governance Framework paper, published by Stats NZ, 2019
https://statsnz.contentdm.oclc.org/digital/collection/p20045coll1/id/2694/

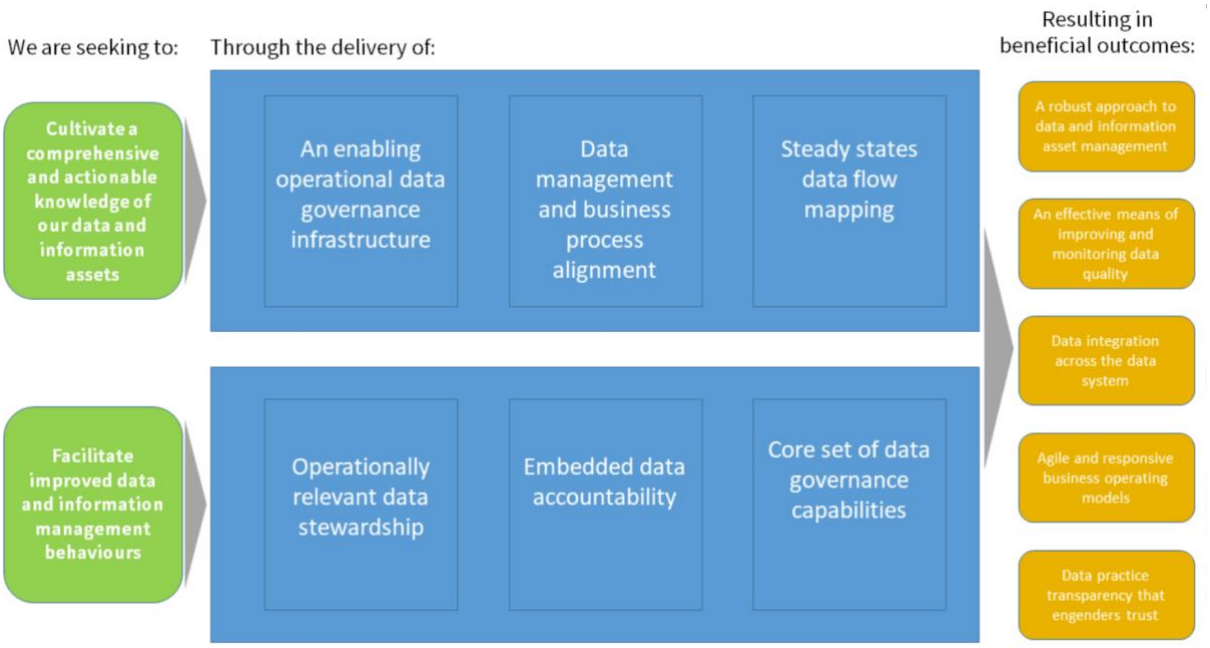

Figure2: An operational Data Governance Framework for New Zealand Government[113]

In designing the oDGF, Stats NZ identified a common disconnect in the way that data is governed. Data is often talked about as an enabler for new opportunities, creating innovation, transformation, ideas of new products and services etc. However, the way data is commonly governed is through compliance mechanisms which are geared toward reducing risk: legal, compliance, privacy and security considerations take centre stage.

The framework proposes a "stewardship" mindset is required for organisations to overcome this disconnect. Staff adopting a proper stewardship perspective will understand that governance is as much an enabling mechanism as it is one for control.

For an individual, a stewardship mindset could be articulated like this:

*"Within the context of my position, I am **accountable** for my organisation's relevant **data assets**, and as such I have a view of where they reside throughout their lifecycle, a measure of their **quality**, understand and can protect against **associated risks**, I am informed of relevant **constraints**, and possess knowledge of how they **contribute value** within my organisation and across the wider system."*

This model helps to eliminate the burden on a data custodian, expands good practice accountability, and supports a more agile way of doing business.

113 https://statsnz.contentdm.oclc.org/digital/collection/p20045coll1/id/2657/

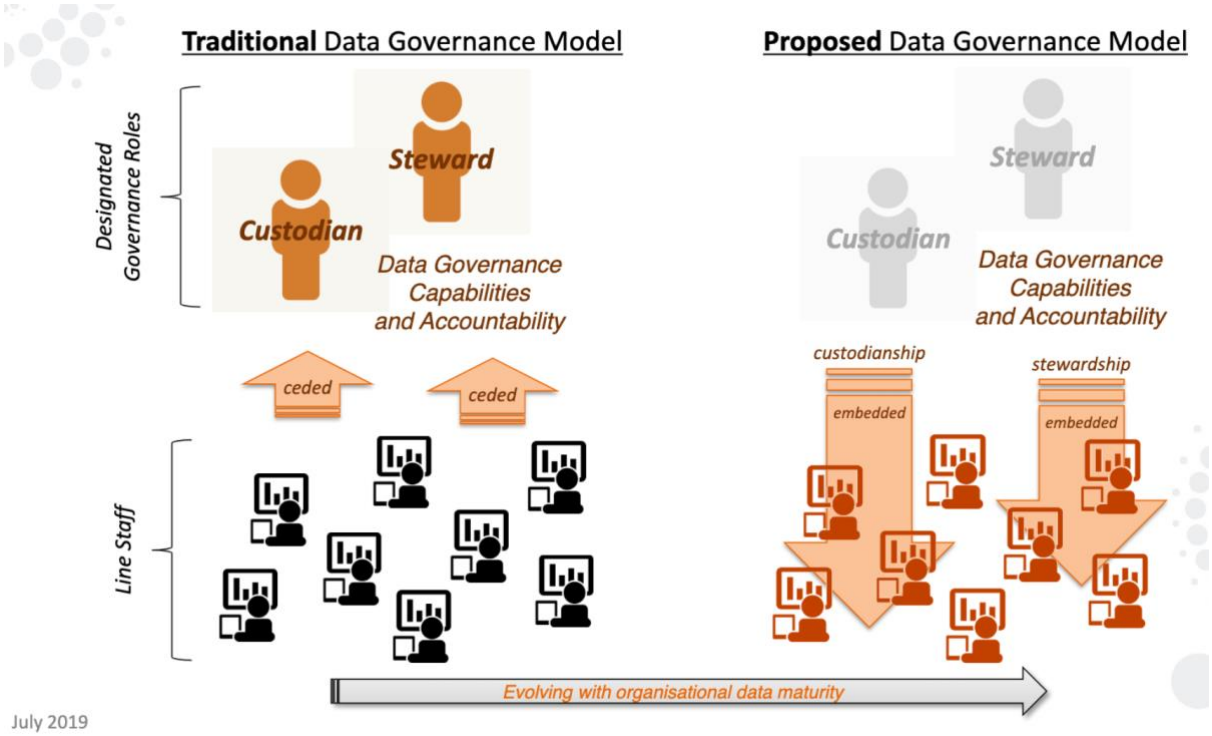

Figure 3: Organisational Data Maturity journey[114]

To measure this journey, they developed a model called KAP of three dimensions, each of them with their own kind of measures:

- Knowledge, with quantitative measures.
- Actitude, qualitative measures for attitude like stories of change and case studies.
- Practice, financial and efficiency measures.

At the moment, their general maturity level is still low, although it varies in an assessment by agency. However, the whole framework is still in its initial phase and it meant a big step for the organization towards an improved data governance.

When implemented, the framework can help promote consistently good practice data management, strengthen data integrity, and build trust. This was the case with a pilot made with Environment Canterbury.

Operational Data Governance Framework: partnership with Environment Canterbury

Stats NZ have worked in partnership with Environment Canterbury (ECan) to deliver a set of data flow maps across multiple parts of their business. These maps will enable ECan to document the way data and information moves through their organisation and

[114] https://statsnz.contentdm.oclc.org/digital/collection/p20045coll1/id/2657/

understand how it reflects and influences their business processes. This knowledge can then inform changes that benefit their organisation.

The pilot had very positive results and a second round is being developed by ECan now.

They included a data quality framework and map flows showed them where to improve their internal process. It also promoted a more collaborative work between business analysts and data analysts. Different teams and roles shared the same room for the first time, which was a way to fight a siloed culture.

1.3.3.4 Connections to other strategies

The draft Data Stewardship Framework and the operational Data Governance Frameworks, both published in 2019, align with the priorities of the Data Strategy and Roadmap.

New Zealand's Strategy for a Digital Public Service, published in 2019, also complements and aligns with the vision of the Data Strategy.

Both of these strands of work form part of a wider series of Public Sector reforms, led by the State Services Commissioner to enable the government to work together to deliver better outcomes and services for New Zealanders.

1.3.3.5 Legal framework

The legal framework related to government data in New Zealand is primarily shaped by the Privacy Act 1993 and the Statistics Act 1975; the Official Information Act 1982 and Public Records Act 2005 are also relevant. The government has recognised that these have not kept pace with changes in the data landscape and are therefore in the process of modernising them.

This includes a review of the Statistics Act carried out in 2018, which found that:[115]

- A range of actors are carrying out governance functions, both in specific domains and across them.
- The establishment of system leadership roles (eg. Government Chief Data Steward; Government Chief Digital Officer, etc.) need to be considered and defined in relation to the overall system governance functions.
- There are potential issues and gaps particularly in the governance approach to dealing with novel questions and responsive accountability.

Officials are currently in the process of repealing and replacing the Statistics Act 1975. Current practice is not expected to significantly change under the modernised Act, rather, modernisation will uphold statistical data needs and usages which are already developed and well established.

115 https://www.stats1975.nz/

1.3.3.6 Funding

Leadership of the agenda is primarily funded through Stats NZ and the GCDS.

A key challenge of this approach is that there is no specific funding for cross-agency initiatives. This is a common issue in governments for cross-cutting work, and is a significant operational barrier to improving use of data. This is compounded by intra-agency incentive models where Chief Executives are measured by agency centred measures and outputs, and minimal consideration is given to benefits at a cross-agency scale.

1.3.4 Governance

The State Services Commission has an overall leadership role in public sector management and supports organisations in exercising their functional leadership responsibilities. A number of cross-cutting agendas in the New Zealand Government are led by Functional Leaders, with responsibility for leading improvements across government for that agenda. For example; there are functional leaders for data, procurement, property, and digital.

As of 2017, the Functional Leader for data has been the Government Chief Data Steward (GCDS), who is also the Chief Executive of Stats NZ. The GCDS holds responsibility for setting the strategic direction for New Zealand's data system, responds to new and emerging data issues, and ensures that government agencies have the capability and right skills to maximise the value of data.

The GCDS works closely with the Government Chief Digital Officer (GCDO), the state sector's digital functional lead. Prior to 2017, a Government Chief Information Officer had held responsibility for data and digital; these 2 new functional leadership roles were created in recognition of the size of the roles and the growing value and need to invest in both digital and data.

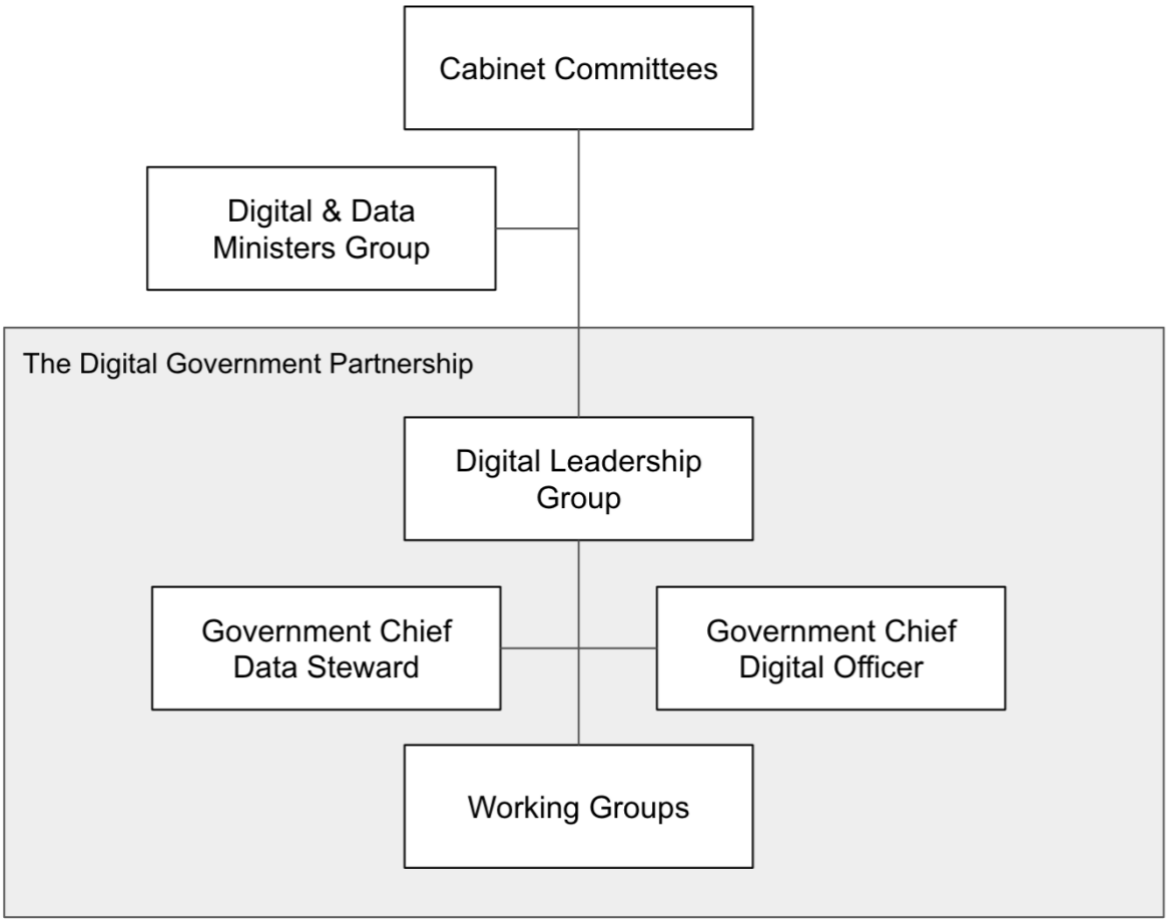

Figure 4: An overview of governance for digital and data in New Zealand

Together, the GCDS and the GCDO use the Digital Government Partnership as a governance mechanism for cross-cutting work on data and digital. The partnership is made up of a leadership group (comprised of public sector Chief Executives) and 4 working groups of senior officials that support New Zealand's digital and data system.

The leadership group typically meets bi-monthly, and is comprised of representatives from:

Department of Internal Affairs, Stats NZ, Government Communications, Security Bureau and Chief Information Security Officer (GCISO), Ministry of Social Development, Inland Revenue, Ministry of Education, Land Information New Zealand, Ministry of Business, Innovation and Employment, Ministry of Health, Ministry of Transport, NZ Transport Agency, Social Investment Agency, and Customs.[116]

Officials under the GCDS and the GCDO provide joint secretariat services for the Digital & Data Ministers– a grouping of eight ministers with digital and data portfolios. The Digital Government Leadership Group is also managed by the GCDO with support from the GCDS.

[116] There have been several recent leadership challenges across both data and digital portfolios. The Digital Government Partnership website notes that the terms of reference for the group will be updated in 2020. https://www.digital.govt.nz/digital-government/leadership-and-governance/digital-government-partnership (accessed 14 February, 2020)

The emphasis of data functional leadership in New Zealand is collaboration and co-creation; the Government Chief Data Steward does not have a hard mandate with which to compel other agencies to follow its leadership. Therefore, the primary levers are co-operation from Chief Executive level down, co-ordinated through the Digital Government Partnership, and through the Data System Leadership group at Stats NZ.

A Cabinet mandate empowers the GCDS to direct agencies to adopt common capabilities and data standards. Agencies are enabled and supported by the GCDS to voluntarily adopt those standards.

1.3.4.1 Co-designing the agenda

The data workstream is designed to 'cut-across' and impact government at a broad level. This requires the building of relationships, capability and capacity amongst stakeholders, and GCDS has adopted an assistive co-design approach to achieve this. For example, the data strategy and roadmap was developed in collaboration with key stakeholders from central and local government, business, community organisations, and NGOs. Stats NZ led multiple workshops and interviews, gaining independent technical guidance and expertise on what was important to stakeholders and where help is needed. Throughout development of the roadmap it was rigorously tested to make sure it was future-focused, robust, practical, and able to be adapted to different data sources and uses.

The roadmap is also considered a living document, designed to be revisited and updated as the data landscape evolves along with needs and priorities. To ensure that the roadmap continues to deliver for the system as well as New Zealanders, any update will involve extensive consultation across government agencies as well as offering stakeholders outside of government to have the chance to influence the final document.

1.3.5 Human capital and skills

A core aspect of the GCDS functional leadership role is to improve skills and capability across the public service. In 2019, the agency started to roll out a **Data and Statistical Capability Framework**. The framework describes the data and analytic capabilities required by those who manage, and interact with, government-held data. It's designed to be incorporated into the role descriptions for people who manage data.

The capabilities are described at five skill levels – from beginner to expert advisor – and all or some of them can be applied to relevant roles. Role filters can be designed to focus on the subset of capabilities for a role or group of roles – for example, a team. The capabilities are linked to many demonstrations of how they can be used. They're also linked to dimensions of the operational Data Governance Framework.

Although primarily used by their skills and capability consulting team, individuals are also encouraged to self-score against it before and after support as a means of benchmarking progress in capability.

The Framework's final goal is to start a conversation about data capabilities and to make people aware that everyone is managing data in one way or another, not only data

scientists or data engineers. It is a powerful tool for managers that can be used to inform workforce planning and to allow users to identify capability strengths and gaps for individuals and teams.

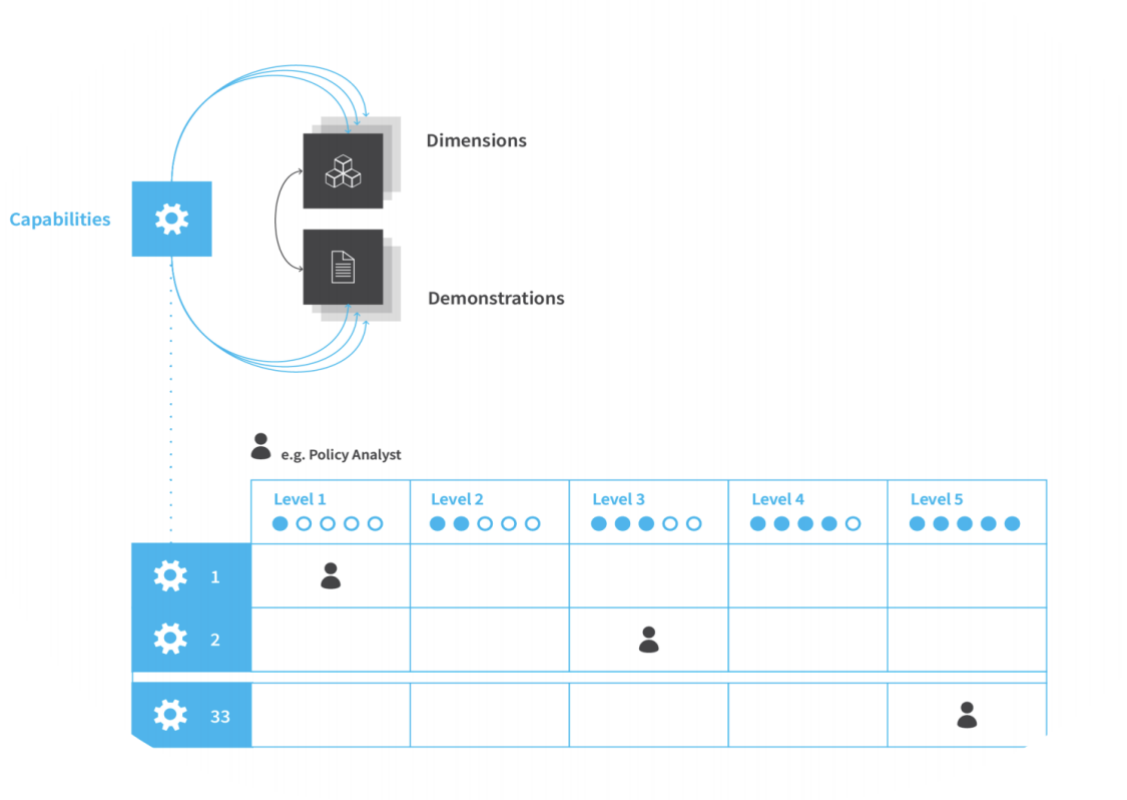

Figure 5: Capabilities framework and skills levels.

GCDS has also developed e-learning modules to support improved 'R' coding capability. These modules have been used by a number of agencies, including Customs, NZ Defence Force, and MBIE.[117]

Capability is about more than people. The GCDS also has a key focus on infrastructure capability, as reflected in the Data Strategy and Roadmap, and there are key work streams within Stats NZ that are focussed on capability issues, such as data content standards, in alignment with this.

1.3.6 Lessons learned

As many other cross-government initiatives, the new Data Strategy and Roadmap have challenged the New Zealand Government. Many are both the successes and the lessons that the Stats NZ team can identify in the processes of providing a joined-up approach and fostering collaboration to generate the maximum impact through data.

Overall positive feedback across government

[117] https://www.data.govt.nz/assets/Uploads/Q4-2017-18-final-dashboard2.pdf

By defining shared goals and coordinating activity across the system, it supported others to succeed and link in existing initiatives by highlighting how they contribute towards the defined focus areas. In practice, evidence includes:

- Departments are using it within their own data strategies. For instance the Ministry of Business, Innovation and Employment has reframed their internal data strategy around the four focus areas to better focus their effort and understand how they contribute beyond their own agency. Stats NZ is using it for its own internal information and data strategy.
- Has resulted in increased engagement with Stats NZ to work towards common standards and data interoperability.
- The Roadmap has also been used to provide advice into the Budget process, and set priority areas of work across government.
- The Strategy and Roadmap has informed what the data system priorities for the system are and ensured that the foundational pieces are established to address needs.
- It has informed the data investment framework, which looks to help guide investment, report on data system progress, and promote an understanding of the scale of what is needed (especially in regards to infrastructure).

A strategy beyond government

The AI Forum[118] and Iwi Leaders Chairs Forum[119] are using the data strategy and roadmap as a part of their work.

The Strategy should better embed te ao Māori

The work that underpinned the original roadmap was undertaken by a third party. Although Māori and iwi were consulted along with other system stakeholders, the ongoing engagement and te ao Māori were not included into the Strategy to the extent befitting Treaty Partners[120].

Articulating value can be a challenge

With many agencies deprioritising work that is very beneficial for the system as a whole, but is not overtly beneficial for their own agency. For this to happen, they are learning how to better engage with the agencies and how to deliver an appropriate message for each level of management, from analysts to CEOs. More used to talking to middle managerial levels, they understand the need to ensure the CEO's support to articulate the strategy's value across government.

Promote universal ownership of the strategy by giving other agencies a clear role in achieving the vision

[118] https://aiforum.org.nz/
[119] https://iwichairs.maori.nz/
[120] Treaty of Waitangi, New Zealand founding document signed by representatives of the British Crown and representatives of Māori iwi and hapū. https://nzhistory.govt.nz/politics/treaty/the-treaty-in-brief

The current Data Strategy and Roadmap has sometimes been interpreted as a declaration of Stats NZ's intended future work, rather than a unifying vision for all agencies to contribute to and engage with. The team is now trying to clearly outline their role as coordinator and enabler rather than owner, to more clearly define the parts all agencies can play in realising the vision of the strategy. This would help system actors become closer and more nuanced partners with jointly set goals and priorities by highlighting how Stats NZ is only one of the contributions to this outcome.

Having a clear understanding of the steps to take in the short and medium-term to reach the goals

The current strategy sets out a high-level future view for New Zealand's data system. A 'lesson learnt' would be to outline material steps that they, as a system, need to take to move toward this vision. This could be difficult given the scope of the GCDS's mandate, but by providing more detail and practical insight into how to achieve the data system of the future, the Strategy can better encourage system-wide engagement and bridge the gap between the current state of the data system and the future vision they are working toward.

Learning means of measuring strategy and metrics to evaluate success

With clearer roles and articulated steps to guide the system on its journey comes improved opportunities for ensuring that the Strategy is successful and that we are moving in the right direction. It can be difficult to measure the success of the Strategy considering its broad and long-term view of the data system. They are currently exploring how to better measure the impact and success of the Strategy to ensure it continues to deliver for New Zealand's data system.

Providing clarity of alignment of the Data Strategy with other key initiatives and strategies

Given its high-level role to guide other data initiatives towards a collective vision, the relationship between the Data Strategy and Roadmap and other existing workstreams can sometimes appear unclear. With no intention of interfering, preventing or delaying other initiatives, but rather to provide coherence to the system, it is important that the role of the Strategy is better articulated and more firmly aligned with existing areas of work. This requires a coordinated activity by all functional leads to ensure a greater system focus on the strategic outcome of achieving public value.

New leadership demands internal transformation

The role of functional leaders as a strategic way to pursue key cross-government initiatives is still very recent in an organisation not used to them. A few but important issues -such as lack of responsibility boundaries, roles superposition on a single person and a soft mandate with no assigned budget- are challenging their leadership and capacity to have a real impact in the organisation. Eventually, the New Zealand government will need to find a way to adapt their governance structure to sustain the work of all functional leads and deepen the transformation sought.

1.4 Case study – Denmark: Welfare Fraud Analytics[121]

1.4.1 Introduction

Denmark is a Scandinavian country counting almost 6 million inhabitants. It finds itself among the top ten best performing EU countries when it comes to digital government. It is one of the best European countries in both the widespread of online services (penetration) and the digitisation level of back- and front-office (digitisation).[122] One of the government's most notable accomplishments is the creation of the Basic Data Registries, which provide public sector data access through web services and APIs, thereby supporting various government operations. However, when looking at the availability, accessibility and government support for reuse of open data, Denmark has made impressive advances over the last two years, but still finds itself among the weaker performing OECD countries.[123]

Supported by the Basic Data Programme[124], which started in 2013, since 2015 the Danish government has been using data analytics for the detection and prevention of fraud and error regarding welfare benefits.

This case study report discusses the governance arrangements underpinning the Danish approach to the data-driven detection of fraud and error in the provision of welfare benefits. It also sets out how the Danish government has dealt with key issues regarding data quality and sharing, the safeguarding of public values, human capital and skills, and monitoring and evaluation. Finally, key successes, challenges and lessons learnt are identified.

1.4.2 Policy context

1.4.2.1 The Danish welfare state reform

Denmark has an extensive welfare benefits system of over 30 different benefits, including child benefits, student grants, unemployment benefits, maternity leave benefits and old-age pensions. In the last quarter of 2019, just over 1 million Danes were receiving public benefits.[125] For 2018, the OECD calculated that public social spending in Denmark amounted to 28% of GDP, making the country to have the fourth highest public social spending-to-GDP ratio of 36 OECD countries.[126] Contrary to the majority of OECD countries, where pensions and health services account for the largest expenditure area, the largest share in Denmark is devoted to income support for the working age population.[127]

[121] This case description is based on the findings from desk research and two stakeholder interviews. It is illustrated by citations from the interviews. These citations aren't attributed to any specific stakeholder out of considerations of anonymity and confidentiality. The case study report has been submitted to the involved stakeholders for factchecking.
[122] European Commission: 2018a
[123] OECD: 2019a
[124] For more information, see The Danish Government and Local Government Denmark, 2012.
[125] https://www.statbank.dk/AUK01 (accessed on 26 March 2020)
[126] OECD: 2019b
[127] OECD, 2019c: 104

In 2007, as part of the implementation of the Structural Reform, which was a major reform of the Danish local government system, the responsibility for the administration of the majority of the numerous welfare benefits in Denmark was placed with the 98 municipalities.[128] This decentralised approach has partly been abandoned with the legal creation of the Udbetaling Danmark[129] (UDK) organisation at the end of 2010 and the definition of its tasks and cooperation with the municipalities in April 2012.[130] Between October 2012 and May 2015, Udbetaling Danmark progressively assumed the responsibility for the collection, disbursement and control of ten municipal areas: Family benefits, Maternity/paternity benefits, State pension, Disability pension, Housing benefits, the Labour market exit benefit, international health tasks.[131] In June 2013, UDK also took over the responsibility for international pension and social security tasks from the former Danish Pensions Agency. The ATP group, a self-governing institution that is also a pension provider and investor of pension funds, is the factual administrator of Udbetaling Danmark.[132] The responsibility for sick leave benefits, child care benefits as well as the cash benefits (*kontanthjælp*) for the unemployed remains with the 98 municipalities.

1.4.2.2 Digitising the welfare state

Supported by the Basic Data Registries, a large part of the Danish welfare system works in an automated way. The IT systems of Udbetaling Danmark, are automatically updated when base registry data, such as a person's address or income changes. In case this kind of personal changes has consequences for any welfare benefits, Udbetaling Danmark will inform the citizen(s) involved about these consequences and give them the opportunity to contest the decision. Without a reaction from the citizen the benefit payment will be changed accordingly.

> "*If you live by yourself in Denmark, you can get a higher pension, because you are alone to pay all your bills. If our department receives a notification that somebody moved to your address, we will write to you telling we have received this information about someone moving into your house, and that this will change your benefits. Do you agree with that? And if people do not disagree, we will just regulate the pension*."

The high level of digitisation of the Danish public sector does not only have implications for the way in which public services are provided to citizens, but also provides opportunities for the control function in the welfare system. In June 2014, after several municipal welfare areas had already been transferred to UDK, the National Association of Municipalities (KL) and the national government decided in an agreement on the municipalities' finances for 2015 to strengthen control efforts in municipalities and Udbetaling Danmark.[133] Following a report commissioned by the Joint Public steering group for effective case management and control (Den fællesoffentlige styregruppe for effektiv sagsbehandling og kontrol - ESK) about the options for consolidating control efforts, they

[128] https://english.sim.dk/responsibilities-of-the-ministry/economics-of-municipalities-and-regions/structural-reform/ (accessed on 23 April 2020)
[129] English translation: Payout Denmark
[130] Udbetaling Danmark was established by law on 22 December 2010, see https://www.retsinformation.dk/eli/lta/2010/1594. Its tasks in relation to the municipalities were established on 11 April 2012 , see https://www.retsinformation.dk/eli/lta/2012/324 (accessed on 23 April 2020).
[131] https://www.atp.dk/en/our-tasks/processing-welfare-benefits/udbetaling-danmark (accessed on 23 April 2020)
[132] https://www.atp.dk/en/our-tasks (accessed on 23 April 2020)
[133] Udbetaling Danmark, 2015

concluded that a centralised approach would be the most effective and efficient way to seize the opportunities offered by both better access to citizen data and developments in machine learning and data mining in the fight against wrong benefit payments. Also considering that many municipalities wouldn't be able to invest in new IT systems and skilled personnel, it was decided to set up a central data analytics unit, **The Data Mining Unit** (Den Fælles Dataenhed - DFD), to serve not only the benefit areas under the responsibility of Udbetaling Danmark, but also those left under the responsibility of municipalities. The necessary investments in IT infrastructure and people could thus be shared by all the public institutions involved.

1.4.3 Building blocks of the strategy

1.4.3.1 Main themes

In the 2015 **Strategy for the Data Mining Unit** (Strategi for Den Fælles Dataenhed), the main goal of DFD is defined as: "to help strengthen the control of social benefits and grants". [134] Its "core product" is to improve fraud & error detection through cross-referencing of data and data analysis.[135]

The 2015 strategy sets the direction for the data unit's efforts in its first year of existence, but is found to be still very relevant in 2020. DFD is in the process of updating its strategy to incorporate an even stronger focus on preventing fraud and error than is the case in the original strategy. The 2015 strategy already mentions:

> *"The work of the data unit must be viewed as an extension of the idea that the control of social services should be moved then as far forward in the case as possible, so that errors are avoided or detected as early as possible, cf. the ESK analysis as well as Udbetaling Danmark's strategy for comprehensive control."*[136]

In line with Udbetaling Danmark's 2011 **Strategy for Comprehensive Control** (Strategi for helhedsorienteret control), the DFD's focus isn't only on fraud, but also on mistakes made by the authority or citizen. A municipality, for instance, might misinterpret a citizen's personal situation and unrightfully start a benefit payment. A citizen might misunderstand the complex rules regarding welfare benefits and apply for a benefit thinking he's eligible. The Danish government takes into account that welfare recipients "are not always the strongest citizens and it's very hard for them to understand the rules"[137].

The DFD's strategy consists of four pillars:[138]

1. High quality data analysis and deliveries
   The data unit must ensure increased and better data analysis, including increased data interconnection for use the control of the payment of social benefits and grants. The prioritization lists from merging registries are the cornerstone of the data unit's deliveries and make a significant contribution to streamlining and targeting the authorities' current control efforts.

[134] Udbetaling Danmark, 2015: 3
[135] Source: PowerPoint Presentation made available by DFD for the purpose of this case study report.
[136] Ibid.: 5
[137] Source: interview with informant
[138] Udbetaling Danmark: 2015

2. Credible partner for government control efforts
   The ongoing development of the data unit's work and deliveries must be done in close cooperation with the authorities that receive deliveries from the data unit, especially municipalities and Udbetaling Danmark. The collaboration is mutual, with all parties participating and responsible for developing the data unit.

3. Strengthened cooperation with foreign authorities
   In order to obtain data on recipients who are affiliated abroad, it is necessary to establish a targeted cooperation with foreign authorities.

4. Special initiatives and collaboration with authorities
   Focus areas for collaboration with other authorities are selected in light of strengthening the overall control efforts and are thus assessed on the basis of specific needs and practical opportunities. The cooperation can, for example, be about implementation of data sources, exchange of information, information campaigns and joint control actions.

1.4.3.2 Connections to other strategies

The DFD strategy contributes to accomplishing some of the goals of two other government strategies:

1. Udbetaling Danmark's 2011 "**Strategy for Comprehensive Control**"[139], which focuses on setting up a framework for cooperation with municipalities on repayment cases, in light of the transfer of five municipal benefit areas to Udbetaling Danmark.

2. The **Digital Strategy 2016-2020** by the Government of Denmark, Local Government Denmark and the Danish Regions contains a focus area *Better use of data and quicker case processing*. Within this focus area, initiative 2.3 *Correct Payments and Better Data on the Individual*, focuses on efficient and effective prevention of incorrect payments and fraud through evidence-based control processes, harmonisation of concepts, high data quality and smooth cross-government collaboration.

1.4.3.3 Legal framework

Amendments to the Act on Udbetaling Danmark

The Data Mining Unit (DFD) was established by Act No. 523 of 29 April 2015, which was an amendment to the Act on Udbetaling Danmark. Changes in section 12 and the introduction of sections 12a-h were meant to strengthen cooperation on the control efforts on the payment of social benefits through the coordination of information on beneficiaries from benefits systems and public registers. The act includes provisions on the purposes for which UDK's may obtain and link certain non-sensitive data without the citizen's or employer's consent and the obligations of other public authorities to make this data available. It also establishes the authority of the Minister of Employment to lay down further rules on cooperation between the municipalities and UDK for the purpose of UDK's

[139] Udbetaling Danmark, 2011

provision of register co-ordinations in the municipal benefit areas. In the course of the years, some more amendments have been made to the articles regulating UDK's rights to access, exchange and process data in support of benefit control. The most recent update of the Act on Udbetaling Danmark dates from 16 January 2019.[140]

1.4.3.4 Funding

The initial investments to set up DFD were a joint effort by municipalities and Udbetaling Danmark. Specific fraud investigations are financed by DFD's clients on a cost-based hour price. In the first four years of its existence (2015-19), the total annual budget amounts to EUR 3.4 million.[141]

1.4.4 Governance of the implementation process

1.4.4.1 The central coordinating department

1.4.4.1.1 Organisation

As shown in Figure 3, the **Data Mining Unit (DFD)** has been created as a standalone unit in Udbetaling Danmark (UDK). Following a number of legal amendments to UDK's tasks and responsibilities, DFD can provide services to the municipalities and the Danish Agency for Institutions and Educational Grants (Styrelsen for Institutioner og Uddannelsesstøtte - SIU). UDK falls Under the political responsibility of the Ministry of Employment (Beskæftigelsesministeriet) and is administered by the ATP group, a self-governing institution that is also a pension provider and investor of pension funds.

[140] https://www.retsinformation.dk/eli/lta/2019/49 (accessed on 28 April 2020).

[141] An informant confirms an amount of DKK 25 million as the average yearly amount received from all client together. The conversion from DKK to EUR was made based on the exchange rate on 1 January 2019, as mentioned on https://www.xe.com/currencycharts/?from=DKK&to=EUR&view=5Y. (accessed on 23 April 2020).

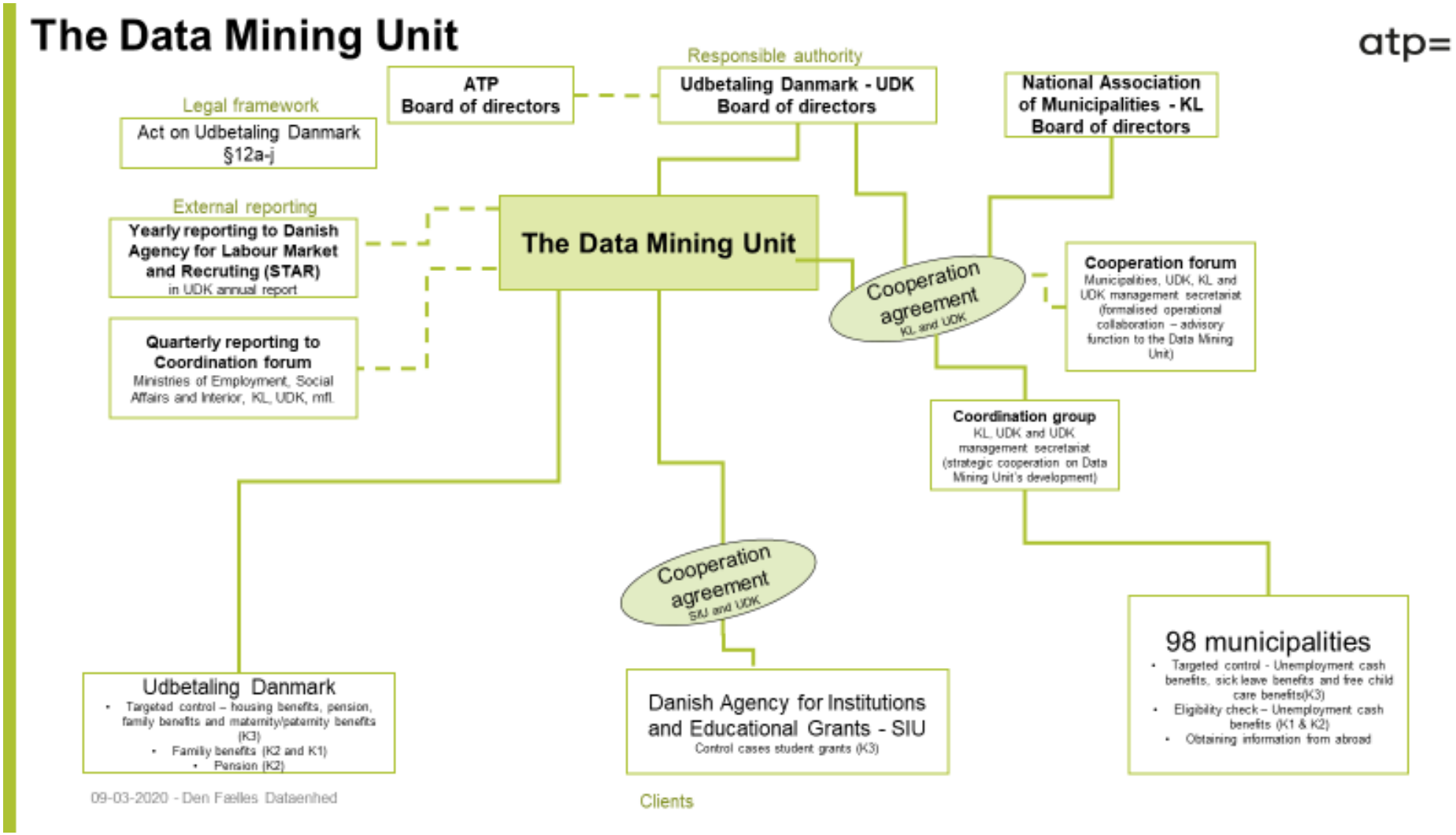

Figure 3 Governance framework of the Data Mining Unit [142]

[142] The original figure was made available by DFD and translated from Danish to English by the author. The translation has been approved by Udbetaling Danmark.

DFD functions as a service provider to a number of clients in the Danish public sector:

- The 98 municipalities, which are responsible for sick leave benefits, free child care benefits as well as the cash benefits for the unemployed (*kontanthjælp*);
- Udbetaling Danmark, which is responsible for a number of public benefits, such as pension benefits, housing benefits, family benefits and maternity/paternity benefits;
- SIU, which provides student grants supporting living costs related to both youth education programmes and full-degree higher education programmes[143];
- ATP's security schemes.

DFD is staffed by around 25 people divided over two sections:

1. The Customers and Products Section, which employs business specialists who focus on client relations and the integration of client needs in the services that are delivered.
2. The Systems and Data Section, which houses the data scientists and software developers, who work on the technical aspects of the fraud detection system.

Two external data specialists collaborate with the Systems and Data Section to make sure the data coming from nine different sources is imported into the DFD data warehouse in a correct and timely fashion.

1.4.4.1.2 Tasks and responsibilities

The core tasks of the Data Mining Unit are:

- Data processing and analysis - to connect data from public sector registries relating to recipients of benefits in a data warehouse and to provide priority lists and relevant information to the authority administering the benefit;
- Coordination - to establish systematic cooperation with relevant authorities, including in other countries and obtain information on citizens' conditions abroad;
- Monitoring - to calculate the effect of better data interconnection and use of data including from foreign authorities.

These three tasks are interdependent and all essential to contribute to an efficient and effective use of data mining and machine learning against incorrect benefit payments.

With the advent of the DFD in 2015, the detection of potential fraud and error cases in the payment of social benefits does no longer solely rely on tips from citizens or the intuition of investigators in the different authorities administering the benefits.

"We identify possible fraud cases with data. That is what is new."

[143] https://www.su.dk/english/state-educational-grant-and-loan-scheme-su/ (accessed on 27 April 2020)

The priority lists produced by DFD are used by the responsible authorities to initiate control cases in the relevant cases and identify erroneous payments and fraud. After having completed the investigation, the caseworker reports back to DFD for reinforced learning and algorithm development.

The DFD's new strategy is aimed at detecting potential fraud and error in the earliest stage possible.[144] Instead of having to ask a citizen to pay back a benefit (s)he incorrectly received, the Danish government wants to discover irregularities in the application phase, before any payments are made. Currently, the DFD supports control activities (Kontrol) in three stages of benefit payments:

K1 Prevention

Detect irregularities in new benefit applications.

This allows caseworkers to prevent citizens from receiving benefits that they're not entitled to.

K2 Early intervention

Detect irregularities with regard to recent changes in personal circumstances of benefit recipients (e.g. change of address).

This allows caseworkers to correct erroneous changes in benefit payments in an early stage (usually one month), thereby preventing a potentially fraudulent situation.

K3 Targeted inspection

Detect irregularities in existing benefit payments.

This allows caseworkers to start an investigation into suspected fraudulent citizens, provide evidence in court and recover up to several years of illegitimate benefit payments.

DFD is expanding its activities in stage K1, with the intention to prevent citizens of arriving in stage K2 or K3. For this purpose, the DFD is running some tests to develop a web service for new applicants. This will avoid Danish authorities the trouble of getting back money that has been paid out while there was no right to receive it. It also avoids financial problems for citizens being confronted with having to back large sums of money.

> "*The idea is to help the citizen as soon as possible to identify: now you get a payment that you shouldn't have had. So, they don't all of a sudden have to pay back a lot of money*."

At the start of its activities, DFD's focus was on supporting its clients' K3 controls to recover illegitimate benefit payments, which in some cases had been going on for years. In 2017, the first K2 analysis were a fact, which by now account for the majority of analytics-based revenue.[145] For municipalities, the disbursement of unemployment cash benefits and the retrieval of incorrect payments form a heavy administrative task,

[144] This is a developing practice, which is in the process of being formalised in an updates DFD strategy.
[145] This is discussed in more detail in the section 'Monitoring and evaluation'.

which is why a large majority[146] has decided to use DFD's K2 analysis of this benefit. K2 analysis of pensions under the responsibility of UDK has been enabled in 2015 by an amendment to the Act on Udbetaling Danmark. The procedure of adjusting pension rates to match the level of income for the beneficiary was restructured from a yearly check-up to an automatic monthly check-up on the basis of the Danish National Income Registry (eIndkomst).[147] Thus, errors made by the beneficiaries who may not have been aware of their obligation to inform about changes in circumstances on a continuous basis can to a large extent be avoided and cases of fraud can be detected at an earlier stage.

As demonstrated in Table 2, most DFD analysis still concern K3 controls. The goal is to increase the success of K1 and K2 controls, so K3 ultimately becomes an irrelevant category.

	K1 - prevention	K2 – early intervention	K3 – targeted inspection
Objective	Prevent fraud and error	Stop fraud and error in an early stage	Correct fraud and error
Type of analysis	Eligibility check	Check on changes in citizens' personal circumstances	Targeted control of ongoing benefits
Benefit claim stage	During application	Ongoing case management	Multiple benefit payments
Types of benefits	Udbetaling Danmark • Family benefits Municipalities • Unemployment cash benefits	Udbetaling Danmark • Family benefits • Pension Municipalities • Unemployment cash benefits	Udbetaling Danmark • Family benefits • Pension • Housing benefits • Maternity/paternity benefits Municipalities • Unemployment cash benefits • Sick leave benefits • Free child care benefits SU • Student grants

Table 2 Data analytics and stages of welfare controls

1.4.4.1.3 Methods

DFD deals with relatively simple and more complex cases. The simpler cases rely on experience-based criteria coming from caseworkers. An example of a simpler case is the detection of so-called 'father hotels' enabling illegitimate payments of single parent child benefits. It is very easy to cheat with single parent child benefits in Denmark, as citizens can change their address in the automated e-movement system. Some fathers

[146] 92 out of 98 municipalities
[147] European Commission: 2018

change their address to that of a friend's, enabling their children's mother to receive extra benefits. This type of fraud is easy to spot with data, by looking at mothers receiving single parent benefits, the father's address, the size of the residence and number of rooms, the number of other fathers living there and whether their partners are receiving extra benefits. However, one of these fathers might really be living there and is as such not committing fraud. To adequately deal with this kind of false positives, the role of the caseworker is essential.

> *"Data can point out a case and say 'something is here', but there has to be a human that looks at it and figures out if there's something or not."*

The more complex cases are handled with machine learning techniques. In 2017, DFD put together a roadmap for Machine Learning initiatives.[148] The unit aims to improve precision (less false positives) when identifying potentially fraudulent cases and be able to find fraud cases that would otherwise not have been found (new true positives). It does so by using both supervised and unsupervised learning models. An example of supervised learning is the prediction of EU citizens forgetting to inform authorities that they emigrated from Denmark or pretending to still work in Denmark to receive family benefits while in reality they live abroad. One of the key data sources for the detection of this type of fraud is the Income Registry (eIndkomst), where citizens can register working hours themselves. DFD performs this analysis on a monthly basis to catch the irregularity as early as possible. Other relevant data to catch this type of fraud is Health Contact Data.

> "*We have seen a case where a guy from Turkey lives here and his whole family has moved back to Turkey. We can see that he has not been visiting a Danish doctor for the last three to five years. He's maybe 70 years old and that is odd*."

A project using unsupervised learning is aimed at detecting data outliers to identify fraud with sick leave.

1.4.4.2 Decision and coordination boards

To enable decision-making and coordination on the use of data analytics for the detection of welfare fraud in Denmark, two boards were put in place.

1. The **Cooperation Forum** (Samarbejdsforum - SF) consists of the DFD, between 10 and 15 representatives of municipalities, UDK, KL and the UDK management secretariat. The SF must help ensure that the data unit's tasks and operational priorities are arranged in such a way that it strengthens the municipalities' control activities as best as possible. During the board meetings, it is decided which types of fraud are considered a priority and should be the focus of the DFD's work. The SF also approves the budget. Depending on the topics on the agenda, different municipalities may sit around the table. The KL

[148] European Commission: 2018b

coordinates beforehand with the municipalities who will be represented at the next meeting.

2. The **Coordination Group** (Koordinationsgruppen - KG) consists of DFD, KL, UDK and the UDK management secretariat. This group is centred on the coordination of the involved authorities at the strategic level regarding the development of DFD. It prepares the meetings of the SF and meets four times per year.

Coordination between DFD and the SIU takes place in separate meetings, as the SIU's stake in the DFD is quite small and given its responsibilities there's no clear need for coordination with the municipalities.

A third relevant board is the **Coordination Forum** (Koordinationsforum), where several authorities get together to coordinate welfare control more generally. Among other authorities, the Ministries of Employment, Social Affairs and Interior, the SIU, KL and UDK are represented. This board meets every quarter and receives a new DFD progress report as material for the meeting.

1.4.4.3 Fostering collaboration between stakeholders

As previously mentioned, establishing systematic cooperation with relevant authorities is one of DFD's core tasks. Close cooperation between DFD and its clients is crucial to harness the potential of data and machine learning in fraud detection. The DFD needs its client's policy-specific knowledge to develop an analytics programme that generates relevant results. The Danish authorities on the other hand need the DFD's infrastructure and expertise in working with data.

1.4.4.3.1

1.4.4.3.2 Strategic and operational collaboration

The collaboration between DFD and UDK's control unit is facilitated by both units being housed in the same organisation. This enables close cooperation at both the strategic and operational level. Data scientists and caseworkers are in regular contact, both in joint workshops and individual meetings between employees of both departments. Coordination is especially relevant when new analytical criteria have been developed.

> "*If the data mining unit has made new criteria for us, they follow up close about how the cases are being handled in our department and data workers who have developed the citizen criteria, they come into our department. They sit next to one of my co-workers and just look into what they're seeing. What do they think about the case? What could have been better? Why is it a wrong case, and what can they do to adjust the criteria? So, every time we have new criteria in the cases, they make sure that it's the right cases that we are looking into. Sometimes they can make an adjustment, another criterion or some other citizens that we look into instead*."

An example demonstrating the value of close cooperation between DFD and its clients is an initiative in relation to housing benefits in 2016 that the DFD developed together with Udbetaling Danmark's Control Unit.[149] Based on the Danish National Income Registry the initiative identified beneficiaries having a zero-household income in a period of at least six months. In some cases, this pattern is due to the beneficiary having an unregistered income. In other cases, the reason is that another person lives unregistered on the address concerned, thus contributing to the household income. Both scenarios can lead to incorrect payments. Well over a hundred cases were identified for further examination by the initiative, resulting in end of payment and claim of recovery in 1 out of 5 cases.

Strategic and operational collaboration with the municipalities is shaped through interactions in the Cooperation Forum and the Coordination Group.

For a successful collaboration between the Data Mining Unit and the caseworkers, DFD provides a data sheet along with an identified potential fraud case, giving insight in why the algorithm identified the specific case and pointing to data that requires further investigation.

> "*Often when you deliver a case, the caseworker will say 'that looks perfectly normal. That's why we deliver a data sheet with it that says something about why the case has been identified and what data is necessary to have a look at. So, we sort of explain it to the caseworker with objective data around the case.*"

The introduction of data-driven detection of potential welfare fraud and error, doesn't mean that the human-based identification of suspicious cases has disappeared. It should rather be seen as a complementary method, which delivers different kind of cases to the caseworker. For the UDK Control Unit, following up on the notifications coming from other channels, such as citizens and municipalities remains a priority over handling the priority lists coming from DFD.

> "*I don't think that you can compare the cases we get from the data analysts and the cases we get from citizens and others, because the cases from DFD are based on objective data and the cases we get from other citizens or from the municipalities, are more like personal observations. These have nothing to do with the CPR or the Income register in Denmark. The personal observations about how people behave is something that DFD cannot see in its systems."*

1.4.4.3.3 Central service provision, public savings and relieved administrative burden as policy levers

[149] European Commission: 2018b

It is unlikely that the individual municipalities or the SIU would have made the leap to data-driven fraud analytics if this service wasn't provided through a centralised unit. The necessary investments to recruit skilled analysts and set up the IT infrastructure, including the unit's data warehouse and the necessary machine-learning algorithms, could only be realised by joining contributions of all Danish authorities using DFD's services.

This kind of competences and analytics firepower it is not possible for the municipalities to develop and maintain, so there is much to gain from this being in a central function supporting local case management.

The promise of public savings has been a big motivator for stakeholders to join the programme, as the premise is that more fraud and error can be detected through analytics than through human investigation. In addition, caseworkers are likely to work in a more effective and efficient way, as they no longer solely rely on tips from other authorities and citizens. Especially the latter group doesn't always turn out to be the most reliable source for detecting potential fraud.

> "*They're not always quite as good, the cases we get from other citizens. But we also get cases from the municipalities that we have a close working relationship with. Many of them know their citizens very well. If they observe something that could have an influence on a UDK payment, they contact us and it's often good observations they have.*"

.

It is not mandatory for municipalities to use the services of the DFD, but they all do out of efficiency gain.[150] Section 12 of the Act on Udbetaling Danmark does specify that organisations (including municipalities) are obliged to provide Udbetaling Danmark, and thus DFD with the relevant data for the fulfilment of its tasks.

Relieving the administrative burden on caseworkers in providing welfare benefits to citizens who aren't entitled to receiving them, has been an important motivator to join the DFD initiative. This is especially the case for municipalities.

1.4.5 Data quality and sharing

1.4.5.1 Data Mining Unit data warehouse

The Data Mining Unit has its own data warehouse, which it feeds with the data related to citizens who are receiving benefits from its clients. DFD receives the data from several Danish authorities. Between 2015 and 2018, the DFD has moved from three to nine data sources.[151] DFD started its activities by integrating three Basic Data Registers into its data warehouse:

[150] All clients are on board for receiving DFD support in K3 control cases. K2 cases are accepted by UDK and 92 municipalities and K1 cases are in the pilot phase.

[151] PowerPoint presentation "Architecture and data sources" shared by informant.

- Civil Registration System (CPR), which contains personal data, such as name, address, civil state and family liaisons;[152]
- Central Business Register (CVR), which contains data such as a business' CVR number, address and business type;[153]
- Building and Housing Register (BBR), which holds data on the identification, location, purpose, year of construction, technical conditions, layout and electric installations of buildings in Denmark.[154]

Regarding data sharing within the public sector, the Danish Basic Data Programme has been of critical importance. This programme, which was initiated in 2013, has led to the creation of digital registrations. The Civil Registration System, the Central Business Register, and the Building and Housing Register are among Denmark's digital resources.[155]

Over the years, DFD has added the following six sources to its data warehouse:

- National Income Registry (eIndkomst), which contains data on earnings and received payments, not only on salaries, but also on state education grants (SU), pension and social benefits such as cash benefits;[156]
- SU data, which is data on beneficiaries of student grants, provided by the SIU;
- R75 - Tax base registry, which contains personal tax data, e.g. the amount of taxes paid, car ownership and bank savings;
- Health contact data, which reveals citizen check-in information at the municipal level with the time precision of one day. The five Danish Regional Authorities compile these aggregated data based on time and location data they extract from the use of the yellow health card (sundhedskort). This card entitles Danish residents to medical treatment and must be carried when seeing a doctor and going to municipal offices. It can also be used in other situations, such as visits to libraries, post offices and stores.[157]
- MOMS data, which is the Danish VAT register;
- STAR data, which is data shared by the Danish Agency for Labour Market and Recruitment (STAR) on visits to public job centres. If a citizen receives cash benefits or unemployment benefits (s)he must register as a job seeker with the job centre.

As the data warehouse only contains data on people who already receive benefits, and not those who are in the stage of applying for benefits, these data are less useful in K1 analysis, which are aimed at identifying fraud or error in benefit applications. Therefore, DFD uses an API to obtain base registry data on new applicants directly from the source organisation.

DFD is not the only user of these data sources. Its clients, such as the control unit of UDK, also consult the data in the first stages of its fraud investigations. When a UDK caseworker starts working on a case, regardless of the channel (s)he received it from, (s)he will consult the data warehouse to perform a first check on the beneficiary.

> "W*hen you put in the number of a citizen, it will automatically look into*

[152] https://econ.au.dk/the-national-centre-for-register-based-research/danish-registers/the-danish-civil-registration-system-cpr/ (accessed on 27 April 2020).
[153] https://datacvr.virk.dk/data/?language=en-gb (accessed on 27 April 2020)
[154] https://econ.au.dk/the-national-centre-for-register-based-research/danish-registers/the-central-register-of-buildings-and-dwellings-bbr/ (accessed on 27 April 2020).
[155] http://grunddata.dk/english/ (accessed on 24 March 2020)
[156] https://skat.dk/skat.aspx?oid=1814027&chk=202455 (accessed on 28 April 2020)
[157] https://international.kk.dk/healthcard (accessed on 18 April 2020)

the CPR and say how many kids you have, who you're married to, where you live. It will look into the eIndkomst, which is the data on how much you earn in Denmark and if you own a house. So, we have this work tool that will easily look into some of the systems that DFD is also working with.

1.4.5.2 Collaborative tools

DFD uses collaborative tools to communicate priority lists to its clients:

- For the 98 municipalities and the SIU, a SharePoint site is in place, which can be accessed through NemID, the Danish public signature system. Municipalities and the SIU can download cases that a caseworker will take up. A new web interface replacing the SharePoint site is expected to be in use in the course of 2021, as the Danish government has decided that the municipalities should communicate through one database. DFD would like to tap into the new system, so its activities can be integrated into the normal workflow of municipalities.
- As the DFD is housed in the UDK organisation, it uses the same IT system as UDK's control unit. This permits the use of software robots writing directly into the control unit's database to create cases in the normal pay-out system.

1.4.6 Safeguarding public values

The Danish approach to safeguarding public values in the use of data analytics in the fight against welfare fraud and error is essentially of a legal and procedural nature. The DFD Strategy mentions two prerequisites for the activities of both UDK's control unit and the Data Mining Unit itself: legal security of the citizen and IT security.[158]

Critics claim that the social and political debate on data ethics in Denmark - not just in the area of fraud detection – is insufficiently present.[159] According to these critics, public values such as transparency and proportionality are in dire need of being discussed with regard to the ambitious digitisation efforts demonstrated by public authorities. Within Udbetaling Danmark, there is an awareness about this type of reservations.

"*There has been discussion in other countries, but also in Denmark about the surveillance of the data of people. Should we have cameras all over in streets to prevent violence? We also have to talk here in Denmark about whether it is okay to look into so many people in so many registries?* […] *In the end, it's politics that has to*

[158] Udbetaling Danmark, 2015:6
[159] https://dataethics.eu/is-scandinavian-digitalisation-breeding-ground-for-social-welfare-surveillance/; https://algorithmwatch.org/en/automating-society-denmark/; Arent Eiriksson, 2019

decide how to proceed in Denmark and we are just following the law. But we get close to some people and some people also think that it it's too close."

1.4.6.1 Non-sensitive data in the data warehouse

By law, the Data Mining Unit doesn't have access to sensitive data, such as criminal records or health care data to use in its analysis of potential fraud cases. The unit creates data interconnections on a weekly basis and only uses non-sensitive personal data in this process. In its workflow, DFD pays attention to having the highest possible data quality, for example by reserving the right to refuse register records if there is any doubt about the quality and reliability of the data used.[160]

The IT security around the data mining unit is ensured by the unit's solutions and systems being established and operated in accordance with ATP's overall IT security policy. This includes securing user access, logging all queries in the system and separating data. Furthermore, DFD ensures that solutions and processing of data in its systems and solutions take place in accordance with applicable legislation in, for example, the Personal Data Act. With the advent of the GDPR, the topic of data protection has gotten a more prominent place on the agenda, but hasn't led to a factual change in the operations of the DFD.

"*The change that has happened is that there is a lot more focus on it to make sure that when you don't need data, you delete it. We of course only have the data that the law permits us to have.*"

The results from DFD analysis that aren't qualified as a potential fraud case are overwritten at the next round of analysis.[161] While DFD states that registry merging is solely carried out on data that is legally authorized to be linked, the Danish legal thinktank Justitia argues that the interconnections between non-sensitive data can create new information that is of a private nature, as the result of the data linkage may give rise to suspicion of a criminal offense in the form of social fraud.[162] Limits to what information Udbetaling Danmark, and thus DFD, may use in its controls are sometimes met. A proposal on UDK having access to data on citizens' electricity consumption was withdrawn after extensive criticism.[163]

1.4.6.2 Citizens' rights integrated in the control investigation

The law doesn't permit caseworkers to investigate a potential benefits fraud case without notifying the citizen involved. They're obliged by law to notify citizens within 10 days after taking their case from the DFD priority list, even if a first check reveals that there is no question of fraud.

[160] Udbetaling Danmark, 2015: 6
[161] Arent Eiriksson, 2019 : 10
[162] Ibid.: 11
[163] Ibid.

When a caseworker takes up a case and contacts the citizen, (s)he will mention the use of data by the authority. No information is communicated on the use of data mining or machine learning techniques.

> *"Some cases we will close immediately after we receive them. We just do a few check-ups and see if everything looks right. And if it does, we will write to the citizens saying that we have looked into your data and also according to the GDPR and telling them that we're not investigating further and they are receiving the benefits from us.* […] *Then we have other cases, where we have to ask the citizens for more information so we can look into and see if they're receiving the right benefits or we have to regulate them."*

Only when there's a very strong assumption of fraud and an in-depth investigation is opened by the caseworker and, then additional, more sensitive information may be obtained from public organisations and banks without the citizen's consent and from other private actors if the citizen consents. Still, the caseworker has to provide beneficiaries with the opportunity to share information to clear up the situation first.

> *"Half of the citizens don't answer […] then you can ask additional data, such as bank data directly"*

The authority has to prove that the citizen should not have received the social benefit.

> *"We have the burden of proof. There are two steps in it. First, we have to ascertain that this happened the citizen received the benefits incorrectly. We also have to prove that this was against better knowledge. And that can be difficult."*

Before making the final decision, the authority must provide all information on which the decision is based, in order to give the citizen possibility to comment or give more relevant information. When the decision is made, the citizen has a possibility to complain to the Council of Appeal on Health and Safety at Work (Ankestyrelsen).

Compared to other countries, control efforts in Denmark have access to a lot of public data. At the same time, they have limited possibilities in their investigations, as they always have to consult with citizens and need their permission to gather additional information from private actors.

> *"I was in the UK to see how they do things and then I was in Norway at*

a conference where many European countries were participating that are working with fraud and data. I can see that in Denmark we have so much more data were everything runs automatically. […] *In the UK you're allowed to do all kinds of things you're not allowed to do here. The social workers in the UK are like the police.* "

The law in Denmark is very clear on needing the citizen's cooperation to get certain data. While public authorities are obliged by law to share data with UDK for the fulfilment of its responsibilities, this is not the case for private entities.

In some cases, fraud cannot be established without the citizen's cooperation:

For example, you can get maternity leave benefits here in Denmark if you have been working during the last four months before you go on maternity leave and you have to work a certain number of hours. If we have a case saying that a beneficiary has not been working in a firm, which happens to be her uncle's, that she has never been working and doesn't speak Danish. […] *Then we would ask her to prove that she has been working there and she may provide us with a contract.* […] *Then we can look into her bank account and see she has not received any payment from this employer. We would then ask her how she received the payments and she would say I've received them in cash. And then we don't have a case because we have to prove that she did not receive them in cash and we cannot do that."*

From a human rights perspective, it can be considered a positive development that Udbetaling Danmark has begun to make data interconnections at the start-up of cases and the ongoing control of whether changes in the circumstances of the citizen are of significance to the benefits. Here's the focus namely on payment of correct benefits (K1 and K2), and not on whether the citizen already has overpaid (K3).[164]

1.4.7 Human capital and skills

The Data Mining Unit consists of 25 highly-skilled specialists, coming from different backgrounds. Several employees hold PhDs. Others have been recruited right after

[164] Arent Eiriksson, 2019: 25

finishing their university education. Recruitment took place via traditional job advertisements. DFD has grown significantly over the past couple of years, as it only had five data scientists in 2016.[165]

"*Machine learning is quite new. If you want people having learnt this, you need young people from the university*"

DFD management acknowledges that special efforts have to made to retain these data talents. Maintaining a modern management style granting enough freedom to the employees to think for themselves and be part of the decision-making process is key.

"*You can't use modern technology with an old-fashioned leadership*."

Working for the public good and generating financial gains for the government is an important motivator for the data scientists working at the DFD. To help to prevent citizens from cheating the welfare system finances by all the Danes.

The success of data-driven fraud prevention and detection doesn't solely rely on the proper functioning of the Data Mining Unit (DFD). The role of the caseworkers as users of the analytics results produced by the DFD is crucial to turn policy information into policy action.

The number of caseworkers; their capacity to work with the analytics results and the time available to them to work with the analytics results (compared to other channels for suspected fraud).

1.4.8 Monitoring and evaluation

Monitoring the uptake and effects of DFD's activities has been an integral part of its strategy from the start. The development of a reporting system on the DFD's results is mentioned as one of the key tasks for the unit, alongside the development of the analytics programme itself.

Since 2016, the DFD has registered a number of data points for its K2 and K3 analysis. It receives these numbers from its clients on a weekly basis. For the K2 analysis, focused on early intervention, these are Udbetaling Danmark and 92 out of 98 municipalities. For the K3 analyses, focused on targeted fraud investigation, these are UDK, all 98 municipalities and the SIU. Data on K1 (anticipatory checks) are not yet available, as the DFD has started conducting this type of analysis only recently. Table 3 shows which data points are collected from each client to monitor the uptake and results of the analytics services provided by DFD.

Data point	Explanation
Number of cases transferred to priority list for first check	These are the cases generated by the Data Mining Unit which are followed up on by a caseworker. Only when the caseworker withdraws a case from the system, it becomes

[165] European Commission: 2018b

	visible which citizen is concerned and the legal term of informing the citizen within 10 days comes into force. The caseworker performs a first check, informs the benefit recipient and other affected citizens (e.g. co-habitants) of what the data says and provides them with the opportunity to respond.
Number of cases closed after first check	These are the cases where the irregularity in the data is cleared up, either by the caseworker's desk research or a reaction from the citizen. There might be an error, but there is no fraud. The intervention of the caseworker can lead to changing the behaviour of citizens, e.g. stopping a benefit, before committing fraud.
Number of cases created as control cases after first check	These are the cases where the caseworker suspects that fraud has taken place, for instance if there has been no reply from the benefit recipient or if other involved citizens provide relevant information. These cases require additional research by the caseworker, as the case has not been excluded through the first check and the burden of proof for fraud lies with the authorities. At this stage, additional data, such as bank data may be requested.
Number of control cases completed with proceeds	These are the cases where the caseworker is able to prove that fraud has been committed and repayment of the benefit is ordered.
Number of control cases closed without proceeds	These are the cases where the caseworker has been unable to prove that fraud has been committed.
Value of repayment claim	This is the amount of money that the citizen has received in illegitimate benefit payments and they have to pay back to the authority.
Value of stop or down regulation	This is the amount of money that the authority would have continued to pay out for another 12 months if the fraud hadn't been discovered.
Wondering ratio	These are the cases requiring additional controls as a percentage of the total number of cases taken up by caseworker for a first check.
Total revenue	This is the combined value of the repayment claim and stop or down regulation.

Table 3 Key performance indicators of welfare fraud analytics

When looking at the development in the revenues over the years, it becomes clear that the total revenues have multiplied by more than 5 times, going from EUR 10.9 million in 2016 to EUR 61.9 million in 2019. Figure 4 shows that in 2018 the role of biggest revenue generator shifted from analysis aimed at targeted inspection (K3) to that aimed at early intervention (K2). This might point to the start of a trend of more irregularities in benefits being detected early on in the payment process.

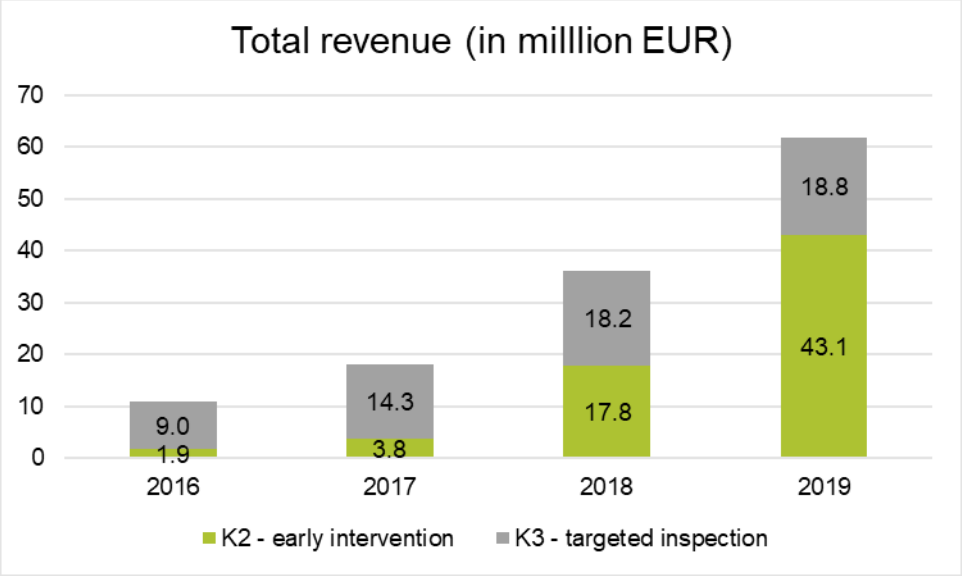

Figure 4 Total revenue data-driven approach to welfare fraud[166]

The reporting data suggest that a data-driven approach towards fraud and error in the welfare benefits system is lucrative for the Danish government. The budget made available to the DFD through contributions of its customers has remained roughly the same throughout the years, EUR 3.4 million per year and has been exceeded by the revenues from the first year of operation.

> "*There are forty caseworkers in the Holistic Control Unit in Udbetaling Danmark. That's a kind of limit on how many cases we can look into.* […] *I think we could get much more cases from DFD, but then we would have to be much more people to look into them*."

Yet, it is likely that even more value could be generated if the number of caseworkers handling analytics-based cases could be augmented. For instance, in the Holistic Control Unit of Udbetaling Danmark, about half of the 10 to 12000 cases they handle per year come from DFD. The cases coming from other channels, such as the municipalities or citizens, are still the priority cases for the Holistic Control Unit and are not threatened by the arrival of the DFD.

The reporting of the key figures is sent to the 98 municipalities on a monthly basis. This forms an important basis in the decision-making of the Cooperation Forum on

166 This figure is based on data made available by the DFD for the purpose of this case study report. The original amounts were converted from DKK to EUR based on the exchange rate on 1 January of each of the reported years, as found on https://www.xe.com/currencycharts/?from=DKK&to=EUR&view=5Y. (accessed on 23 April 2020).

DFD's budget. They also serve in Udbetaling Danmark's year report and the quarterly reporting to the ministries connected to Udbetaling Danmark (see Figure 3).

As part of the Digital Strategy 2016-2020, an evaluation of the Data Mining Unit (DFD) was carried out in 2016. It was published and approved by the Joint Public Steering Committee for Error Payments and Control in November 2017.[167] Recommendations were formulated to improve the unit's performance in six areas:

1. Establishing new key figures and improved presentation of data
2. New and updated KPIs
3. Governance and process for criteria development
4. International collaboration
5. Increased focus on support for control steps 1 and 2
6. Strengthening of legal framework for data processing and use

The evaluation takes into account that at the time of research DFD is only in the start-up phase. The report therefore suggests that the evaluation forms the basis for decisions on further optimization of the inter-institutional collaboration on the Data Mining Unit.[168]

1.4.9 Successes, challenges and lessons learnt

1.4.9.1 Key successes

- *Government savings*

  In 2019, the Welfare Fraud Analytics Programme has saved the Danish Government EUR 61.9 million[169] in benefits payments.

- *Avoidance of financial problems for citizens*

  Because fraud and error cases can increasingly be caught in an early stage, less citizens experience the problem of having to pay back benefits they were not entitled to. Communication to citizens explaining the rules can be better targeted.

- *Detection of new fraud patterns*

  The analytics system is able to detect welfare fraud, which isn't discovered by human detection, enabling the Danish authorities to discover new fraud patterns.

1.4.9.2 Key challenges

- *Time needed to set up the data infrastructure*

  It has taken the DFD three years to build up a completely functional data unit including the necessary amount of data acquired from the relevant authorities, the IT architecture, security, knowledge about fraud cases and the skilled personnel.

[167] https://digst.dk/forenkling/afsluttede-initiativer/fejludbetaling-og-kontrol/evaluering-af-den-faelles-dataenhed/ (accessed on 24 April 2020)
[168] Deloitte, 2017
[169] The amount of DDK 445 million has been converted into EUR based on the exchange rate of 0.13363 on 1 January 2020 https://www.xe.com/currencycharts/?from=DKK&to=EUR&view=1Y

- *Finding skilled data personnel*

  It was hard to find people capable of working with machine-learning, as they are in high demand. Young talents were recruited right from university.

- *Retaining data talents*

  In order to ensure that their highly skilled employees are motivated to stay at the DFD, an appropriate management style is necessary.

- *Insufficient human capacity to process analytics results*

  Many potential fraud cases generated by the analytics system are never investigated by a caseworker, as there are not enough caseworkers available to process all the potential cases.

1.4.9.3 Lessons learnt

- Don’t underestimate the time necessary to create a proper functioning data unit. Don’t expect data-driven results in a short amount of time.
- Create an attractive and stimulating working environment by making data scientists part of the decision-making process on what to use data analytics for.
- Make sure analysts and end users collaborate closely to ensure that policy knowledge is leveraged for data analytics and that end users understand how to use analytics results for their work.
- Using data analytics for prevention and early intervention on welfare fraud and error rather than targeted inspection, can contribute to not only alleviating the administrative burden on authorities, but also more respect of citizens’ human rights.

1.5 Case study: FINDATA

1.5.1 Introduction

Finland has a long history of collecting extensive data in registers but making use of the data has been difficult and inefficient. In 2019 a new **Act on Secondary Use of Health and Social Data** entered in force in Finland. With the new enabling legislation, Finland has become the first country in the world to successfully enact a law on the secondary use of well-being data that meets the requirements of the European General Data Protection Regulation (GDPR).

The new legislation **enables and expands the use of social and healthcare data from the traditional areas of scientific research and statistics** to those of management/control of social welfare and healthcare, development and innovations, knowledge management, education, authorities' planning and forecasting tasks, and steering and supervision of work. The new Act facilitates the establishment of a **new central data permit authority** in Finland, knows as Findata. The latter organisation has been established and operates within the National Institute for Health and Welfare, but as a separate entity.

Findata is the **one-stop-shop** responsible for streamlining and securing the secondary use of social and health data. It guarantees a flourishing ecosystem around the secondary use of social and health data streamlining the processes for the issuing of research permits and data collection and ensuring that data is being used in secure environments, thereby maintaining the trust that the general public have in authorities and the public sector.

This case study reports on how the provisions of the **Act on Secondary Use of Health and Social Data** have been implemented in practice, describing the main building blocks of the strategy and the main themes, the governance of the initiative, the Finnish legal framework around the use of the health and social data, as well as key elements linked to data sharing & quality, safeguards for sound use of the data, required skills, etc. In addition, key lessons learnt are explored in terms of success and obstacles faced in the execution of the strategy. In that sense, it also describes in detail the main characteristics of Findata, how it works, the governance, and how it guarantees access to the social and health data. A description of Findata users, data sources and data lakes is also provided.

1.5.2 Policy context

To support electronic healthcare and social welfare services in Finland, the **Act on Secondary Use of Health and Social Data**[170] (hereinafter "Act" or "new Act") was proposed in 2016 and entered into force in 2019. The Act was based on the "**National health-sector growth strategy**"[171], which aims to make Finland an internationally renowned pioneer in health business and in well-being. In fact, the health sector in Finland has grown and become more international at a much quicker rate than many other industries in recent years.

"Finland is an internationally known pioneer country in research and innovation, investments and new business in the health sector[172]."

The Act provides new options for the use of data. Most importantly, it enables and expands the use of social and healthcare data from the traditional areas of scientific research and statistics to new domains that will be further analysed in section 1.1.3.1.

It not only opens up the national social and healthcare registers but also data from the operational client and patient systems in primary care, specialist healthcare and

[170] Act of Secondary Use of Health and Social Data; https://stm.fi/en/secondary-use-of-health-and-social-data

[171] *National health-sector growth strategy. https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/75145/MEE_guidelines_8_2016_Health_sector_growth_strategy_17062016_web.pdf.*

[172] Health sector strategy; https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/75145/MEE_guidelines_8_2016_Health_sector_growth_strategy_17062016_web.pdf

social services (e.g. through the Kanta electronic data management system of health and well-being[173]).

The Act is aligned with the EU's General Data Protection Regulation (GDPR) and is one of the first implementations of the GDPR for the secondary use of data in Europe. In addition, the Act guarantees an individual's legitimate expectations as well as their rights and freedoms when processing personal data.

Central to the implementation of the secondary use of health and social data is **Findata** – the **Data Permit Authority.** The new data permit authority operates directly under the Ministry of Social Affairs and Health and is a separate legal entity functioning as part of the National Institute of Health and Welfare (THL). It grants permits to allow the secure and easy use of social and health data for the purposes laid down by the new Act on the Secondary Use of Health and Social Data. As such, Findata is a one-stop-shop for data, centralising the decision procedures and access to the data. The sensitive data is handled in a safe and secure environment. Access to data is controlled, and only the results of the analytics can be used externally.

"For data users, "one-stop-shop" principle will significantly reduce bureaucracy and streamline and speed up access to various data sources[174]".

1.5.3 Building blocks of the strategy

The main drivers that led Finland to issue the new **Act on Secondary Use of Health and Social Data** have been the high bureaucracy and difficulties in making use of the data, as well as the separation of services with a silo approach, the existence of different protocols and practices, the lack of secure and safe environment for the data.

In order to overcome those existing barriers and leverage on the potential of the huge amount of collected data, the new Act entered in force to facilitate:

- **the elimination of overlapping and administrative burden related to the processing of permits:** previously, obtaining the permits and data was a difficult and expensive process in terms of time. Indeed, the entity requesting the secondary use of data needed to approach each authority and data source separately. Today instead of having to apply for separate permissions from several different data owners, a single central operator (Findata) issues and grants research permits[175];
- **smoother and faster processing of permits**: thanks to new permit authority (Findata), retrieving combined health and social data from different sources is easier, faster and possible with just one permit application;
- **smoother collation of data from different registers:** Findata aggregates in one single place all the data coming from data controllers in scope (see section 1.1.5.2);
- **easier and more efficient use of valuable social and health materials in research and development activities:** companies (health technology and life-science) can access to valuable data and material for R&D purposes;
- **clearer knowledge management by service providers:** the knowledge-management opportunities in social welfare and healthcare sectors increased thanks to easier access to comprehensive data sources.

1.5.3.1 Main themes

The objectives of establishing a new centralised body (Findata as Data Permit Authority) devoted to the implementation of the secondary use of health and social data have been mainly:

- **enable efficient and secure processing of personal data collected during the provision of social and health care** as well as personal data

[173] *Kanta; https://www.kanta.fi/en/citizens*

[174] Anne Pitkäranta, Research Director at HUS Helsinki University Hospital, https://www.helsinkibusinesshub.fi/a-new-act-will-streamline-the-secondary-use-of-finnish-health-and-social-care-data/

[175] Findata's permit service; https://lupa.findata.csc.fi/

collected for the purpose of steering, supervision, researching and collecting statistics on the social and health care sector, in full compliance with GDPR prescriptions;
- **allow the collected personal data to be combined** with the personal data held by Social Insurance Institution of Finland, Population Register Centre, Statistics Finland and Finnish Centre for Pensions;
- **secure the legitimate expectations, rights and freedoms of individuals** when processing personal data.

The secondary use of health and social data means that the data generated during health and social services are also used for other purposes, in addition to the primary purposes for which they were originally saved. Health and Social data were initially only used for traditional **scientific research and statistic in the health care domain**. Thanks to the Findata approach, it is possible now to activate new data usage, such as:

1. **Development and innovations activities (R&D):** not only researchers but many diverse worldwide health technologies and life-science companies benefit from this new approach. Thanks to the access to social and healthcare data reserves, these companies can start to see opportunities in Finland and expand their R&D activities to the country.
2. **Knowledge management:** thanks to Findata, each organisation can improve its knowledge-management opportunities in social welfare and healthcare sectors with easier access to comprehensive data sources and new services around high-quality registered data.

New opportunities for knowledge management

Information, including that on customers' well-being and the use of services, can be used to support the management of social welfare and healthcare services. There has previously been no clear legal basis for the collation of data, which is required for knowledge management. Knowledge management is one of the grounds for the secondary use of data. The necessary collation of information for the purpose of management from the service provider's own registers is possible without authorisation by a permit authority. From the perspective of social welfare and healthcare reform, both promoting knowledge management by service providers and the expansion and timeliness of national monitoring data are very important areas of development. Amended legislation facilitates the better use of data to support decision-making while respecting the privacy of individuals.

Source: Ministry of Social Affairs and Health[176]

3. **Planning and forecasting of the activities and initiatives performed by social and health care Finnish authorities**: to transform the Finnish authorities in a data-driven organisations, the data collected from Findata can be used as a basis for the planning of central initiatives and programmes.
4. **Governance and supervision of social and health care organisations:** the governance and control of organisations by social and health care **Finnish authorities** based on personal data and statistics and/ or on data received from case-studies, such as the National Institute for Health and Welfare or the Population Register centre.
5. **Education:** higher education institutions, such as biomedical campus universities, can benefit from the data stored in Findata using data for the development of projects, publications, preparation of seminaries and other materials.

1.5.3.2 Connections to other strategies

The preparation process and issue of the new Act, followed by the implementation through Findata was guided by two main national strategies in the healthcare domain:

[176] Seconday use of health and social data; https://stm.fi/sote-tiedon-hyodyntaminen?p_p_id=56_INSTANCE_7SjjYVdYeJHp&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_count=4&_56_INSTANCE_7SjjYVdYeJHp_languageId=en_US

- **Health Sector Growth Strategy for Research and Innovation Activities**[177]. This the growth strategy was prepared by three ministries (Ministry of Employment and the Economy, Ministry of Social Affairs and Health, Ministry of Education and Culture), the Finnish Funding Agency for Innovation (Tekes), and the Academy of Finland, in cooperation with research and innovation funders and organisations in the health sector. The strategy contains key recommendations for the systematic development of research and innovation activities, as well as for increasing investments and achieving economic growth in the health sector. The core of competitiveness is formed by university hospitals and the development of centres of expertise arising around them from the perspective of research and business partnerships. With the help of regional specialisation and nationally integrated operating models, Finland is able to improve its competitiveness.
- **Information to Support Well-being and Service Renewal: eHealth and eSocial Strategy 2020**[178]. The strategy has been prepared in active cooperation with citizens, social welfare and health care organisations and different ministries as well as the Association of Finnish Local and Regional Authorities and the Ministry for Social Affairs and Health. The objective of the strategy is to support the renewal of the social welfare and health care sector and the active role of citizens in maintaining their own well-being by improving information management and increasing the provision of online services. To achieve these ends, it is essential to make active use of information related to social welfare and health care services and to refine it into knowledge that will support both the service system and individual citizens.

Both of these strategies also form part of a wider series of Public Sector reforms, to enable the government to work together to deliver better outcomes and services for citizens.

1.5.3.3 Legal framework

The Finnish legal framework around the use of personal and health data is schematically presented in the figure below. It is composed of legislative reform measures, new acts and related supporting EU regulations.

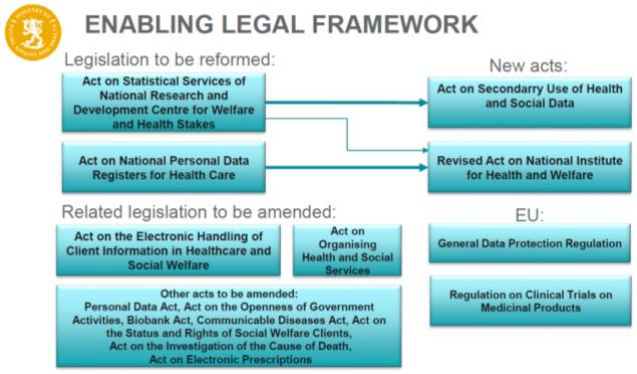

Figure 5- Legal framework for the use of health and personal data[179]

Legislative reforms

The following two are old acts that have been reviewed in new Acts:

- **"Act on Statistical Services**[180] **of National Research and Development Centre for Welfare and Health (STAKES)"**. STAKES is subordinated to the Ministry of Social Affairs and Health. It compiles statistics on the clients using social welfare and health care services. The Statistic Act was adopted in

[177] Health Sector Growth Strategy for Research and Innovation Activities; https://tem.fi/documents/1410877/3437254/Health+Sector+Growth+Strategy+for+Research+and+Innovation+Activities+26052014.

[178] Information to Support Well-being and Service Renewal: eHealth and eSocial Strategy 2020; http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/74459/URN_ISBN_978-952-00-3575-4.pdf.

[179] Source: Secondary Use of Health and Social Data in Finland, Joni Komulainen Ministrial adviser Master of Laws.

[180] Statistics Act; https://www.stat.fi/meta/lait/statistics-act-2802004_en.html

Helsinki on the 23rd of April 2004. It lays down provisions for the procedures and principles concerning the collection of data and the designing and production of statistics that shall be applied by state authorities in their statistics compilation.

- **"Act on National Personal Data Registers for Health Care (556/1989)**[181]**"**. This Act entered in force in 1989. For the purposes of this Act, national personal registers of health care mean national central registers containing information on a person's state of health, illness or disability, or on treatment measures or comparable measures applied to him or her.

New Acts

The previous two old acts have been translated in the following acts that represent the new acts of the legal framework:

- **"Act on Secondary Use of Health and Social Data**[182]**"**. It entered in force in 2019, and its purpose is to facilitate the access to social and health data for different scopes (for more details see 1.1.3.1 Main themes).
- **"Revised Act on National Institute for Health and Welfare 561/2019**[183]**"**. It entered in force in May 2019, and it repeals the Act on National Personal Registers of Health Care (556/1989).

Legislation to be amended

There are still several legislations to be amended due to the new approach to the use of secondary data. In particular, the most relevant acts to be amended are:

- **"Act on the Electronic Handling of Client Information in Healthcare and Social Welfare**[184]**"**. Under this Act, entered in force in 2007, public healthcare organisations are obliged to enter patient records in a nationally centralised archive. Deployment of the centralised archive is mandatory for private healthcare organisations, if they have an electronic system for long-term storage of patient records. The aim of the Act is to further the data security of patient information processing, patients' access to information, and provision of healthcare services with better patient safety and efficiency.
- **"Act on Organising Health and Social Services (1326/2010)**[185]**"**. It was issued in Helsinki on the 30th of December 2010. The objectives of this Act are to 1) promote and maintain the population's health and welfare, workability and functional capacity, and social security; 2) reduce health inequalities between different population groups; 3) ensure universal access to the services required by the population and improve quality and patient safety; 4) promote client-orientation in the provision of health care services; and 5) improve the operating conditions of primary health care and strengthen cooperation between health care providers, between local authority departments, and with other parties in health and welfare promotion and the provision of social services and health care.

Additionally, among others, also the following acts necessity adjustments: Personal Data Act[186], Act on the Openness of Government Activities[187] and Biobank Act[188].

[181] Act on National Health Registers of Health Care 556/1989: https://finlex.fi/fi/laki/smur/1989/19890556
[182] Act of Secondary Use of Health and Social Data; https://stm.fi/documents/1271139/1365571/The+Act+on+the+Secondary+Use+of+Health+and+Social+Data/a2bca08c-d067-3e54-45d1-18096de0ed76/The+Act+on+the+Secondary+Use+of+Health+and+Social+Data.pdf
[183] Act repealing the Act on National Personal Registers of Health Care 561/2019; https://finlex.fi/fi/laki/smur/2019/20190561
[184] Laki sosiaali- ja terveydenhuollon asiakastietojen sähköisestä käsittelystä; https://finlex.fi/fi/laki/smur/2007/20070159
[185] Health care act; https://www.finlex.fi/fi/laki/kaannokset/2010/en20101326_20131293.pdf
[186] Personal Data Act; https://www.ilo.org/dyn/natlex/docs/ELECTRONIC/73914/75774/F350179689/FIN73914%20English.pdf.
[187] Act on the Openness of Government Activities; https://www.finlex.fi/en/laki/kaannokset/1999/en19990621_20150907.pdf
[188] Biobank Act; https://www.finlex.fi/fi/laki/kaannokset/2012/en20120688.pdf

EU regulation

At European level, the Finnish framework is supported by GDPR and Regulation on Clinical Trials on Medicinal Products[189].

1.5.3.4 Funding

The project was funded by the **Ministry of Social Affairs and Health and by returns on endowment capital and capital investments.** State's supplementary budget allocated 2.5 million euros for the year 2019 towards launching the operations of the data permit authority and the construction of a data-secure environment.

The provision of Findata services is fee-based: costs are based on fees defined by the Ministry of Social Affairs and Health regulation nr 1500/2019[190].

In particular, Findata can charge a **fee for the picking, delivery, combination, pre-processing, pseudonymisation and anonymisation of data**. These activities are crucial for providing a secure operating environment. The compensation for picking and delivering data from other data repositories is based on the regulation and it is applied to the controllers who deliver data. A controller or personal data processor, who discloses data to Findata in compliance with the **Act**, is entitled to the compensation. Upon request, a party entitled to compensation must deliver to Findata an estimate of costs that the processing of data causes. Based on the information received, Findata creates a cost estimate for carrying out the data request and will deliver the cost estimate to the permit applicant. Then, Findata charges the permit holder the costs and it pays to the controllers who provided the data.

1.5.4 Governance of the implementation process

1.5.4.1 Fostering collaboration between stakeholders

The new Act is a good example of government policies enabling legislation to promote digitisation and the development of ecosystems. **Cross-sectoral cooperation between several ministries integrated it into the political decision-making process**, which ensured support from the government and from across the political spectrum.

"Unique cooperation between public authorities, companies and associations was the key to success[191]."

Experts from ministries, authorities, companies and associations from across the private and public sectors worked together to prepare the implementation simultaneously with the legislation process. It was a unique way of working and something carried out for the first time at the national level. The traditional preparation process follows the pattern of preparing the legislation first, having it passed by Parliament and then starting the implementation phase. The new way took a lot of courage, some risk-taking and a strong belief in the ability to succeed. The reward is worthwhile, as it speeds up development and innovation and reduces the length of the implementation phase.

Alongside the legislation work, Sitra has played an essential role in the Isaacus project (2015-2018) to pilot[192], test and develop service models and processes, metadata descriptions, data lakes and cooperation models with authorities and stakeholders. The basic principles and an operating plan for Findata were prepared.

189 REGULATION (EU) No 536/2014 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 16 April 2014 on clinical trials on medicinal products for human use, and repealing Directive 2001/20/EC https://ec.europa.eu/health//sites/health/files/files/eudralex/vol-1/reg_2014_536/reg_2014_536_en.pdf

190 Ministry of Social Affairs and Health regulation nr 1500/2019 on Findata fees; https://storage.googleapis.com/thl-production/2020/01/findatan-maksuasetus_1500-2019.pdf.

191 SITRA; https://www.sitra.fi/en/blogs/new-legislation-will-speed-use-finnish-health-data/

192 A Finnish Model For The Secure And Effective Use Of Data; https://www.sitra.fi/en/publications/a-finnish-model-for-the-secure-and-effective-use-of-data/#abstract.

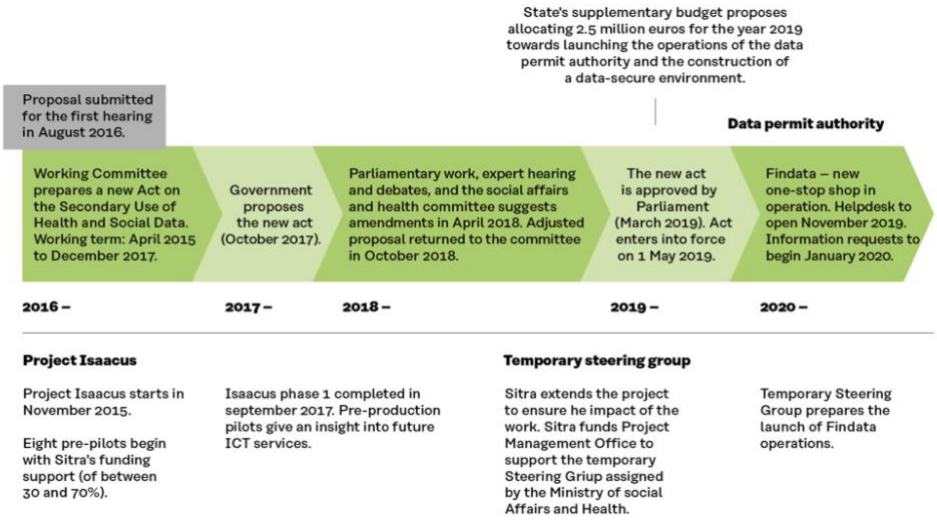

Figure 6 - The process of developing the new one-stop shop body in Finland[193]

The project was also looking at the international dimension, with an active dialogue and an exchange of information and best practices between the Nordic countries, EU member states and several other countries. The exchange of experiences and ideas was crucial to the success of the work.

In 2019 a temporary steering group was put in place to prepare the launch of Findata operations. In the summer of 2019 the Findata director and staff were recruited. The Findata website has been opened since August 2019. From the 1st of November 2019, the Findata help desk (website, e-mail and phone service) is available. Currently, it Findata is open to receive data requests for anonymised statistical data (since the 1st of January 2020). The system also collects data permit applications for individual-level data from the 1st of April 2020.

As regards future developments, Findata aims to guarantee a secure remote use environment for customers and improve data set descriptions, data management and methods.

1.5.4.2 Supervision and control

The new data permit authority operates directly under the **Ministry of Social Affairs and Health** and is a separate legal entity functioning as part of the National Institute of Health and Welfare (THL).

Findata's operations are supervised by the **Parliamentary Ombudsman and the Data Protection Ombudsman**, among others.

The **National Supervisory Authority for Welfare and Health Valvira** monitors Findata's data secure user environments. In addition, Findata must give an annual report to the Data Protection Ombudsman regarding the processing of health and social data and the related log data.

1.5.4.3 Governance of the operations

To steer the operations of Findata and to develop the cooperation, the Ministry of Social Affairs and Health organises a **steering committee** every threeyear and elects a **chair person** for the commitee. The members of the steering group have been choosen from the **Ministry of Social Affairs and Health**, the **Finnish Institute for Health and Welfare**, the **Social Insurance Institution of Finland**, the **Finnish Centre for Pensions**, the **Populaton Register Centre**, **Statistics Finland**, the **Finnish Institute of Occupational Health**, the **Finnish Medicines Agency Fimea**, and **representatives of social welfare and health care service providers**.

The task of the steering committee is to process and make proposals to the National Institute for Health and Welfare and the Ministry of Social Affairs and Health on:

[193] Source: https://www.sitra.fi/en/publications/a-finnish-model-for-the-secure-and-effective-use-of-data/

- the annual action plan of the Data Permit Authority and the associated budget;
- the report on operations and financial statements as applicable to the Data Permit Authority;
- the joint development of controllers and the resources allocated to the task;
- the resources allocated for the development of information systems and cooperation;

Additionally, the steering commeetee is responsible to:

- set goal indicators for the processes of the Data Permit Authority and initiate external audits on the processes;
- if necessary, make a proposal to the National Institute for Health and Welfare and the Ministry of Social Affairs and Health on the improvement of the Data Permit Autority's operations;

The Ministry of Social Affairs and Health establishes a **high-level expert group** for Findata. The task of the group is to create guidelines on anonymisation, data protection and data security for the Data Permit Authority's operations. The expert group must have an expert on each of the following fields: artificial intelligence, data analytics, data security, data protection, suitable research, statistics and statistical service as well as a representive of the Data Permit Authority.

1.5.5 Data quality and sharing

1.5.5.1 Access to data: the new "one-stop shop" approach

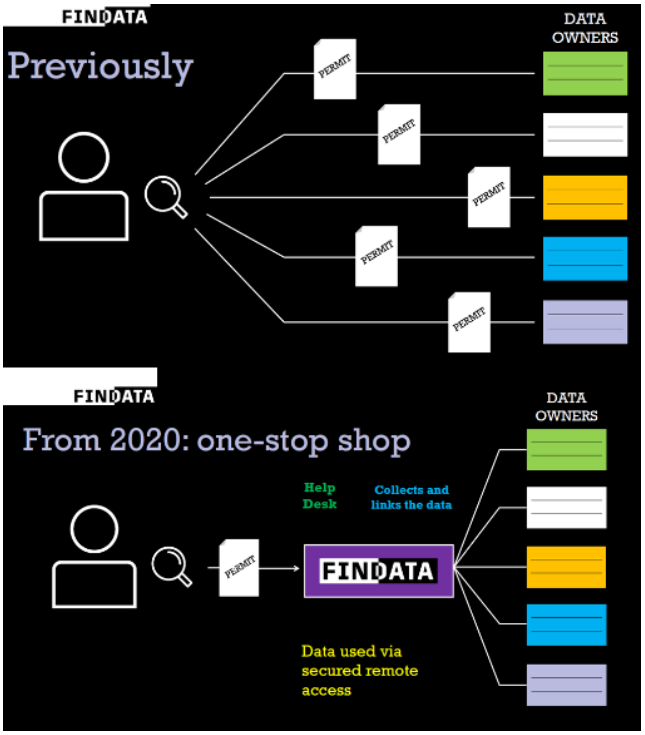

Figure 7 - The evolution of the data provisioning[194]

As described in figure 4, previously, obtaining the permits and data was a difficult and expensive process in terms of time. Indeed the user needed to approach each authority and data source separately. Today instead of having to apply for separate permissions from several different data owners, a single central operator/ service operator (Findata) issues and grants research permits[195], including ethical evaluation. After granting the permission to use data, the service operator collects relevant data from different registers and edits, combines and anonymises the data before distributing it to the user. As also depicted in the figure 5, Findata ensures that data handling and transfer of data

Figure 8 - The evolution of the data provisioning

[194]Source: Implementation of the national Social and Health Data permit authority Findata. Johanna Seppänen, PhD, Director

[195]Findata's permit service; https://lupa.findata.csc.fi/

occurs in a secure environment and that the process meets all the requirements defined by Finnish law. To do this Findata uses a data management system including a secure remote user environment with associated tools. Additionally, a data description system serves as a centralised place for saving the metadata of available materials. The solution includes, among other things, a metadata editor for editing and updating data descriptions.

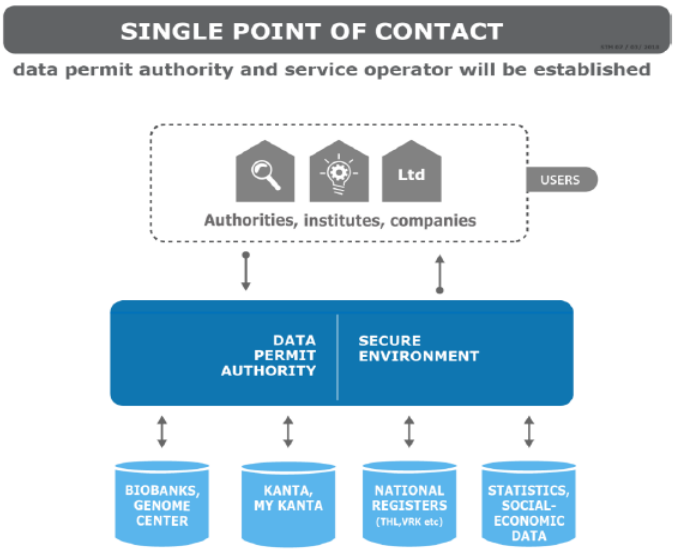

Figure 9 -Access to data[196]

There are two different levels of data and different ways to access related datasets:

1. **Individual level data.** The data of this level can be used for scientific research, statistics, education, authorities' steering, supervision, planning and forecasting. This data is available in a remote access environment for a set period. The data has been anonymised or pseudonymised. A data utilisation plan is required for access to data sets.
2. **Statistical level data.** The data of this level can be used for the aforementioned purposes and, in addition, for development and innovation and knowledge management. This kind of data are directly delivered to customers.

In addition to the citizens, the user who can access Findata also include:

- **Authorities:** among others Social Insurance Institution of Finland (KELA), Social Welfare Office and National Supervisory Authority for Welfare and Health – Valvira.
- **Institutes:** research institutes, universities and biomedical campuses.
- **Companies:** pharmaceutical companies, health technology and life-science companies.
- **Professionals:** Healthcare and social welfare professionals, professors and PhDs.

1.5.5.2 Data sources and data lake

Findata is responsible for data permits and data requests when the data is combined from the following controllers:

- Social and health care operating units;
- Finnish Institute for Health and Welfare (does not apply to data collected for statistical purposes);
- Social Insurance Institution of Finland Kela (benefits and prescriptions);
- Data saved in Kanta Services;

196 Source: Finland – Most advanced ecosystem for healthcare innovation. Nora Kaarela, Head of Industry, Health & Wellbeing, Invest in Finland, Business Finland.

- Finnish Centre for Pensions (work and earnings data, benefits and the bases for them);
- National Supervisory Authority for Welfare and Health Valvira;
- Finnish Medicines Agency Fimea;
- Finnish Institute of Occupational Health (occupational illnesses, exposure tests);
- Regional state administrative agencies (matters related to social welfare and health care);
- Population Register Centre (individual's basic details, family relations, places of residence and building information);
- Statistics Finland (to the extent that access is required to data covered by the Act on Establishing the Cause of Death 459/1973).

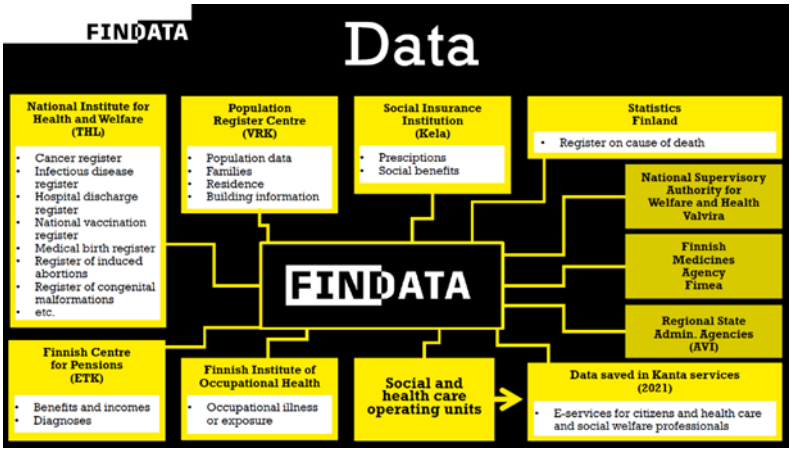

Figure 10 - Data controllers and related data[197]

Health and social data are stored in various national and local databanks. There is a large variety of different kinds of patient record, well-being, social wellness and other data available. The usage of a unique national person ID-number makes it possible to combine personal records.

The following architecture represents the databases in scope and two types of data lakes:

- Local data lakes: County hospitals, local social and health care providers etc. have enormous data in various systems. In many places, the data is now gathered into data lakes.
- National data lakes: Social care data, Patient data and Prescription data, Personal Health Record and social data are stored in national data lakes.

[197] Source: Implementation of the national Social and Health Data permit authority Findata. Johanna Seppänen, PhD, Director

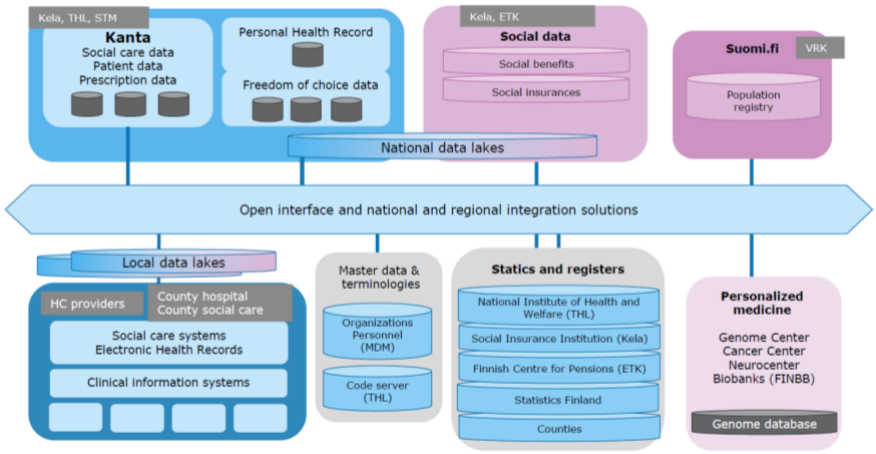

Figure 11 - Data sources and data lakes[198]

1.5.6 Safeguarding public values

Findata, is responsible for ensuring the ethically sustainable use of data. Findata makes decisions on data permits concerning data held by other controllers, and it is responsible for the collection, combination, pre-processing and disclosure of data for secondary use, in accordance with the Act. Furthermore, the data permit authority maintains a data request management system to forward and process data requests and permit applications. Findata also maintains a secure hosting service for receiving or disclosing personal data and a secure operating environment, in which the permit holder may process the personal data he/she has been disclosed on the basis of data permit. It also supervises compliance with the terms and conditions of the permit it has issued. The data permit may be revoked if the permit holder fails to comply with the law or the terms and conditions of the permit. Lastly, the data permit authority is responsible for the pseudonymisation and the anonymisation of personal data.

1.5.7 Human capital and skills

In 2020 Findata counts around 11 people. A first investment has been made to hire profiles with legal expertise and administrative skills. Additionally, several ICT profiles have been hired. The hiring process is still on-going, and in 2021, Findata expects to have 20 professionals. The hiring strategy aims to guarantee a good mix of skills that can enable the use of the new technology and methods, analytics skills and achieve a good understanding of research practices.

1.5.8 Monitoring and evaluation

To monitor the customer experience, feedback are collected from clients to analyse customer satisfaction and to collect any suggestion on the improvement of the service. Several KPIs were also created to monitor the grants, the number of clients and the time spent for the delivery of the service (from the onboarding process until the deployment of the service).

1.5.9 Key successes

Easier and faster access to data

Thanks to Findata, retrieving combined health and social data from different sources is easier, faster and possible with just one permit application, removing the need to approach each authority and data source separately. Previously, obtaining the permits and data has taken as long as up to two or three years. The Act guarantees the provision of a permit within just three months. For exceptionally complex data requests that can cover several data registries, the data permit authority can extend the time it takes to obtain a permit by a maximum of three extra months. In addition, the data is provided with little delay, no later than within 60 business days after the permit has been approved.

198 Source: Secondary Use of Health and Social Data in Finland, Joni Komulainen Ministrial adviser Master of Laws.

Openness in the preparation of the Act

Additionally, the Act can be considered as a success story for open and interactive preparation. The drafting of the Act on the Secondary Use of Health and Social Data proceeded in stages as an open administrative process. At different stages, researchers, developers, directors and companies participated in thematic workshops that set goals and sought solutions to detected problems. Both the workshops and the work of the working group that drafted the Act were documented in Innovillage's open workspaces[199], where everyone could view and comment on them. During the process, there was active cooperation between data controllers and ministries. Openness and interaction were key guiding principles.

Trust and privacy of citizens

The trust and privacy of citizens are guaranteed thanks to a safe and secure environment and access control of users. The individual level data sets are anonymised. The compliance with the Act is constantly supervised by the Data Protection Ombudsman, Valvira and Findata. The data permit may be revoked if the permit holder fails to comply with the law or the terms and conditions of the permit.

1.5.10 Key obstacles

Lack of customer-centricity thinking

Findata, at the moment, is not responsible for potential customers, to help them through the process, maintain active contact and ensure a smooth customer and service process. An investiment have to be done in this sense for example hiring new resourses dedicated to the customer care.

Lack of a good mix of technical skills

A good mix of skills that can enable the use of the new technology and methods, analytics skills and achieve a good understanding of research practices are important. Indeed, technical problems were caused either by new technology, such as the support for the functionality of early-stage data-lake software versions or by new requirements.

[199] Innokyla; https://innokyla.fi/en/home

1.6 Special analysis: mapping data strategies in the Member States

The crisis hit heavily the Europe's economy and exposed its structural weaknesses. Aiming at looking beyond short-term recovery, the European Commission launched in March 2010 the Europe 2020 Strategy to exit the crisis and prepare the EU economy for the challenges of the next decade. The Digital Agenda for Europe, one of the seven flagship initiatives of the Europe 2020 Strategy, proposes actions that will get Europe on track for smart, sustainable and inclusive growth. It also set out the key enabling role of Information and Communication Technologies (ICT) in successfully attaining the objectives of the Europe 2020 Strategy.

The Digital Agenda aimed to maximise the social and economic potential of using ICT, especially of the internet within both economic and societal activity – from business environment to citizens wellbeing. It was focused on finding more effective use of digital technologies that will enable Europe to address its key challenges and to provide a better quality of life for citizens. Areas where ICT can make a significant difference are diverse, such as health care, transport, environment, and access to public services.

The main focus areas of the Digital Agenda are the digital single market, interoperability and standards, trust and security, fast and ultra-fast internet access, research and innovation, and digital literacy, skills and inclusion.

After 10 years from the European Digital Agenda, the digital technologies have transformed the economy and society, affecting all sectors of activity and the daily lives of all Europeans. Data has become the centre piece of this transformation and to what will follow. Data-driven innovation will bring enormous benefits for citizens, through improved personalised medicine, new mobility and through its contribution to the European Green Deal.[200]

Currently, individuals generate increasing amounts of data, and how these data are further collected and used must place the interests of the individual first, in accordance with European values, fundamental rights and rules. Citizens will trust and embrace data-driven innovations only if they are confident that any personal data sharing is in strict compliance with EU data protection rules. The increasing volume of non-personal industrial data and public data combined with technological change in data's storing and processing, is an important potential source of growth and innovation that also should be tapped. Citizens will be able to make better decisions based on insights gleaned from non-personal data.

The data should be available to all – whether public or private, big or small, start-up or giant. This will help society to get the most out of innovation and competition and ensure that everyone benefits from a digital dividend. The digital Europe should reflect the best of Europe - open, fair, diverse, democratic, and confident.

The European Commission Digital Strategy[201], launched in 2018, sets out a vision to meet the challenge brought by new technological developments and to evolve towards a digitally transformed, user-focused and data-driven administration. The strategy formulates principles to underpin the development of digital solutions to support the effective and coherent use of data by the Commission in compliance with data-protection regulations. It also aims at building a data ecosystem that will include a set of interoperable data repositories, corporate base registries, associated tools and frameworks for data sharing and reuse between DGs and external stakeholders. The ecosystem will be a set of interconnected and interacting elements for the collection, acquisition, management, storage, curation, sharing, reuse, publication, protection, archiving and preservation of the Commission's data. It aims addressing the issues of data fragmentation and data silos across the Commission and to ensure the

200 European Commission, C*ommunication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A European Strategy for Data* (Brussels: European Commission, 2020).

201 European Commission, *Communication to the Commission - European Commission Digital Strategy: A Digitally Transformed, User-Focused and Data-Driven Commission* (Brussels: European Commission, 2018).

emergence of semantically interoperable data repositories, master data and base registries. This initiative is consistent with the Public Sector Information Directive and the Infrastructure for Spatial Information in Europe Directive.

Interoperability and technology neutrality of the digital transformation of the government and public administrations are important aspects to be considered. When it comes to enhance the re-use of emerging joint solutions and promote their implementation in more domains to avoid sectoral duplication of service infrastructures, the Member States use various approaches. The most popular building blocks seem to be related to the eID / eSignature (Austria, Croatia, Czech Republic, Finland, Hungary, Portugal, Slovakia), as well as eInvoicing (Belgium, Finland, Poland, Portugal, Slovakia, Slovenia, Sweden). But some countries, such as Finland recognise that the use seems limited. Few member states mention the eIDAS Node – and their success. But already before the Tallinn Declaration, the Netherlands, Austria and Germany have managed to connect their identification and authentication systems via their (technically compliant) eIDAS-nodes, making it possible to use Austrian or German eID to access the Dutch national eID infrastructure, and vice versa.

Digital transformation implies development and use of different ICT solutions, in general, all countries claim that they use open source software (some, like Finland, stress that the legacy is somehow heavy but they keep promoting these standards, while others like Latvia, claim that the demand for open source is rapidly increasing). Many countries state that their open data portal is built with open solutions. In countries such as Hungary, Italy and Portugal, procurement rules have been modified to suit open source criteria. In these countries, the administration must prefer open source solutions when buying new software and the administration has to justify itself when they choose differently, i.e. demonstrating that an open source solution doesn't exist. Estonia reports that the majority of digital government solutions are available as open source, including their flagship initiatives such as X-road, RIHA, Personal Data Usage Monitor, Estonian eID software.

Many member states make their ICT solutions available to citizens and businesses by making the source code open. In Finland for instance, several notable ICT solutions developed for the public administration have been based on open standard software and common creative 4.0 licence such as Suomi.fi services, and X-Road. In Portugal, many public ICT solutions are available for re-use, such as the eID middleware and the Open Data Portal. But in general, member states didn't built any centralised platform where public administration ICT solution source codes can be found. They rather publish terms and conditions on how this can be done, such as in the Czech Republic, Portugal and Romania. The major exception is Estonia, with its own repository "Koodivaramu" after having first used Github to publish software solutions. Yet even in this case, the legal framework is still been adjusted in order to facilitate reuse and define clear licences. In Italy, Developers Italia is, among others, a site for developers of Italian digital public services that provides source code. In Spain, it's an institution, the Centre for Transfer of Technology, which is in charge of making available to citizens and businesses the ICT solutions developed by the public administrations.

The following sections provide an overview of how data strategies are addressed in the member states. In contrast with the case studies, which aimed at extracting insight from leading experiences, this chapter aims to systematicall map what is being done in the member states.

1.6.1 Data strategies

Across Member States digital strategies, the numbers of priorities and focus areas varies from four (more general) to more than 11 priorities (with different levels of details for specificity). Often the digital strategies concentrated on three areas of interest – citizens, business and public services. Both "only-once" and "digital by default" principles are important elements in development of digital (or eGovernment) strategies in all Member States. Interoperability is another relevant aspect of the transformation process of the public administration, and is reflected in many cases into national interoperability frameworks or similar initiatives. Usually these frameworks put in place at national level the principles stated by the European

Interoperability Framework.[202] Alongside digital strategies, countries are developing additional strategies that consider the new advancements in data and new technologies fields. One strategy often encountered is the eGovernment strategy that focuses on digitalisation of public administrations and development of eServices for both citizens and business. In some cases, this strategy is included into the broader digitalisation one, but there are countries that provide it separately. In general, these strategies propose principles to reduce the data fragmentations and break the silos across public services, and to improve transparency and efficiency of national administrations.

In the majority of the Member States, provisions relate to data and use/re-use of data in the national context are often embedded into the digital transformation of governments. However, there are four countries (Ireland, Italy, Malta and Netherlands) that have also developed stand-alone data strategy or action plans. In **Ireland**, the strategy aims to build a data ecosystem that will improve the governance, the management and the re-use of public data in a secure, efficient and transparent way, to the benefit of citizens, businesses and policy makers. Delivering efficient and effective public services, with the least burden possible for data collection from citizens, businesses and other public bodies, is another goal set up by the Irish data strategy. The strategy is backed-up by the Data Sharing and Governance Bill that aims to provide a generalised legal basis for the sharing of data between public bodies while also setting out appropriate safeguards under which such sharing should take place. The implementation of the strategy will allow the government to better respond to service demands by reducing the administrative burden for both citizens and businesses, improving the transparency of personal data processing and improving the process for policy formulation and evaluation.

Through the Code for Digital Administration, **Italy** has introduced the National Digital Data Platform and the mandatory provisions for strategic datasets. The initiative targets six problematic areas, such as compliance of the data sharing, system interoperability, data standardisation and valorisation, users' engagement and democratisation. The platform adopts the concept Software as a Service (SaaS) simplifying the process the public administration needs to follow for the management and sharing of public data to make them better and more valuable. The platform is a data hub, with open data as a service. It has a secure and standardise framework to gather, control and distribute data. It also includes a catalogue and ontology service, and classifies data by organisations, roles and group access. The platform offers the opportunity to connect via their web API for the administrations that already have commercial tools for data visualisations, business intelligence or data science/analytics.

The goal of **Malta**'s National Data Strategy is to ensure the proper level of governance and management of all the data and records used to support the normal administrative functions of the Government (excluding only the processes and data used by the units responsible for national security). When it comes to its objectives, the strategy aims to implement the "Once Only" principle, reduce the administrative costs and burden as well as eliminate the duplication of efforts. It also strives to deliver the products and services in an equitable, efficient, timely and effective manner. Alongside the strategy, a national data infrastructure is put in place that includes also a register of registries and a national data portal, with two sections – the open data portal, and the internal data sharing service. The register of registries provides a comprehensive classification scheme for datasets, while the data portal is the one-stop shop for data discovery and consumption. In 2016, the Persons Register and the Address Register were set up as the common data that required in almost all the internal and customer facing processes within the public administration.[203]

The **Netherland**'s Data Agenda Government (NL DIGITAAL in Dutch), explained in depth in section 1.1 sets out how data can be used (even) better to improve policy-making and resolve social issues. It also pays specific attention to the protection of public values and fundamental rights. The agenda focuses on five main themes: problem-solving with a data-driven approach, legislation and public values, improve

[202] European Commission, *European Interoperability Framework* (Brussels: European Commission, 2017).
[203] https://mita.gov.mt/en/nationaldatastrategy/Documents/OPM%20Circular%20No%2031_2016.pdf

the quality of government data and use it more efficiently, collect and share knowledge about a data-driven approach, and invest in people, organisations and changes in culture. The motto of the data agenda is "Making data work for us" and intent to take advantage of the enormous opportunities the data offers in the country's benefit. The Data Agenda Government focuses on how data is used within our society and addresses the proper and responsible use of data by the government. In particular, it looks at the social opportunities and the public value, as smart application of data and technology can improve the quality of living. The implementation of the agenda is a joint responsibility of central and local governments. The agenda emphasis on the fact that data is only the mean to an end, and to make to most of it a join approach between citizens, businesses and government is needed. The legal and ethical frameworks are no forgotten either, as algorithms and artificial intelligence have an enormous potential in solving social problems. However, algorithms' transparency and technology neutral solutions are important factors in further developments in these areas. To be able to contribute to a well-functioning administrative and democratic system, the data needs to be shared at the right time and in the right way. It is also important to improve the quality and usability of government data to achieve the right goals. This will help to create a transparent government and a properly functioning data system. Sharing the knowledge about data-driven approaches available in the public sector, and breaking the silos is also part of the Data Agenda Government. It offers the opportunity to benefit from others experiences and reduce duplications, and encourages to brining forward new ways of collaborations (e.g. secure multiparty computation for analysing sensitive data, sharing costly and complex facilities, such as a data lab, with laggards and smaller organisations etc.). To efficiently benefit from the data-driven approach, both the corporate culture and the employees need to change. The agenda purpose is to help training the staff in basic data skills alongside with raising awareness about the possibilities of data.[204]

In 2012, the Danish government has launched a basic-data initiative.[205] The public authorities in **Denmark** collect and re-use various core information (basic data) about individuals, businesses, real properties, buildings, addresses and other. Basic data is considered an important contribution to modernising the public sector. Re-use of high-quality data is an essential basis for public authorities to perform their tasks properly and efficiently across units, administrations and sectors. Citizens and businesses are provided with better and more efficient service, while the employees in the public sector are less burdened by repetitive and routine tasks. To achieve its goals, the initiative concentrate on five parallel processes: ensure the re-use of data and to prevent double registration and shadow registers by releasing the basic data to the public, enhance the quality of data by expanding and including other necessary data (redundant registries will be phased out), make linking data possible by standardised technological requirements, improve the distribution of common public-sector data by putting in place a common infrastructure (data distributor), and ensure efficient, effective and coordinated development and use of basic data by establishing a cross-institutional basic-data committee.

Cybersecurity strategies are another set of strategic documents developed by Member States. This type of strategies started being developed since 2012 (Belgium and Cyprus), while most recent ones set up in 2018 and 2019 that are either revised/updated versions of earlier ones (Netherlands, Ireland, Lithuania, Portugal) or new ones (Estonia, for the next period 2019-2022).

The Artificial Intelligence (AI) strategies are also more often present in the Member States future development planning. In 23 countries, the AI strategies have been already published, either as such or under different public statements, as the next step of technological progress. Using AI services and applications in the public sector is one of the aspects encountered in many of the strategies developed by the Member States.

[204] An experiment conducted by the Education Executive Agency in Netherlands showed that chatbot can make services more efficient, but also more user friendly (NL Digitaal).

[205] https://en.digst.dk/data-and-it-architecture/basic-data/.

Austria AI strategy[206] aims to establish the public administration as a lead user and provider of AI-supported public services for citizens and thus to push ahead the modernisation of the administration. Cyprus aims to improve the quality of its public services through the use of digital and AI-related applications[207]. In Czech Republic, the elaboration of AI pilot projects in public administration and health care and the development of a binding plan for public administration data availability plan, including data standards, in the view of AI use are some of the short-term objectives of the AI strategy.[208] One of the priority areas identified by the strategy for AI development in Denmark aims to ensure that the public sector uses AI to offer world-class services for the benefits of citizens and society.[209] Estonia's strategy encourages the use and development of AI applications in both the public and private sector.[210] Finland aims for the leading position in artificial intelligence, with wide adoption of open data and framework conditions adequate to prosperous development of AI. By using the AI applications, it also strives to provide high-quality public services and improve the efficiency of the public sector.[211] In Germany, the strategy sets up a framework for a holistic policy for future development and applications of AI. The strategy intend to integrate AI in society in ethical, legal, cultural and institutional terms in the context of a broad societal dialogue and active political measures by developing various initiatives (e.g. new Skilled Labour Strategy, the Digital Work and Future Fund) and by recognising the need of developing AI expertise in the public administration.[212] In terms of data infrastructure, the action plan of the Hungary's future AI strategy will establish a National Data Asset Agency to ensure a responsible and efficient data use in the public sector.[213] In Ireland, the main areas covered by the strategy[214] will include: societal opportunities and challenges of AI, enterprise development and deployment of AI, research, development and innovation, human capital aspects, data issues, digital and connective infrastructure, public sector use of AI, as well as ethics, governance, standards and regulatory framework. Two of the key objectives of the National Strategy on Artificial Intelligence in Italy are to develop a data infrastructure to fuel AI developments and to improve the public services through a wider adoption and use of AI applications.[215] Latvia outlines policy actions to promote the adoption and development of AI in the public and private sector in its national strategy. In Luxembourg, the strategy looks for the simplification process of the public administrations by offering more efficient and qualitative services to citizens using new technologies AI-based.[216] The Strategy and Vision for AI in Malta 2030 aims to support for increased adoption of AI in the public sector.[217] The Strategic Action Plan for Artificial Intelligence of Netherlands envisages making optimal use of AI in the performance of public tasks and to offer solutions for societal challenges.[218] The Spanish RDI strategy in AI has a dedicated focus on the public sector. The strategy recommends the creation of a National Data Institute for a better data governance and improved quality of public services. AI applications based on natural language processing (e.g. chatbots) could serve a better interaction between public services and citizens.[219]

The Directive on open data and the re-use of public sector information provides a common legal framework for a European market for government-held data (public

206 https://www.oecd.ai/dashboards/policy-initiatives/2019%2Fdata%2FpolicyInitiatives%2F24233.
207 https://ec.europa.eu/knowledge4policy/ai-watch/cyprus-ai-strategy-report.
208 https://www.mpo.cz/assets/en/guidepost/for-the-media/press-releases/2019/5/NAIS_eng_web.pdf.
209 https://ec.europa.eu/knowledge4policy/ai-watch/denmark-ai-strategy-report.
210 https://ec.europa.eu/knowledge4policy/ai-watch/estonia-ai-strategy-report.
211 https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf.
212 https://www.oecd.ai/dashboards/policy-initiatives/2019%2Fdata%2FpolicyInitiatives%2F24114.
213 https://www.kormany.hu/en/ministry-for-innovation-and-technology/news/ai-action-plan-and-strategy-to-be-developed-this-year.
214 The strategy is expected to be finalised in the first quarter of 2020; https://ec.europa.eu/knowledge4policy/ai-watch/ireland-ai-strategy-report.
215 The strategy is currently in public consultation phase; https://ec.europa.eu/knowledge4policy/ai-watch/italy-ai-strategy-report.
216 https://digital-luxembourg.public.lu/sites/default/files/2019-05/AI_EN.pdf.
217 https://ec.europa.eu/knowledge4policy/ai-watch/malta-ai-strategy-report.
218 https://ec.europa.eu/knowledge4policy/ai-watch/netherlands-ai-strategy-report.
219 https://ec.europa.eu/knowledge4policy/ai-watch/spain-ai-strategy-report

sector information). Two key pillars of the internal market - transparency and fair competition – are the building blocks for the framework. The adoption and implementation of this directive across the Member States show a mix picture. In twelve countries specific measures linked to the PSI Directive have been adopted (Belgium, Cyprus, Germany, Greece, Hungary, Ireland, Italy, Luxembourg, Malta, Romania, Spain and Sweden). In other twelve countries their previous framework for access to documents has been adapted to fit the new requirements (Bulgaria, Croatia, Czech Republic, Estonia, Finland, France, Latvia, Lithuania, Netherlands, Poland, Portugal and Slovak Republic. The remaining three countries (Austria, Denmark and Slovenia) use a mixed approach, with new measures specifically addressing re-use and legislation predating the Directive.

1.6.2 Data protection and privacy

Over the recent years, data has been the centre of the digital technologies transformation that has transformed the economy and society, affecting all sectors of activity and the daily lives of all Europeans. And this phenomenon is far from coming to an end. Data is the new lifeblood of economic development. It is the source of new products and services that are brought to the market and drive productivity and resource efficiency gains across all sectors of the economy. It will also allow for more personalised products and services and will enable better policy-making and upgrade government services.[220]

In this context, effective data governance needs to ensure that data is consistent and trustworthy and doesn't get misused. It's increasingly critical as organisations face new data privacy regulations and rely more and more on data analytics to help optimise operations and drive business decision-making.[221] The data is an essential resource for start-ups and small and medium-sized enterprises (SMEs) in developing products and services. Their availability is crucial for training artificial intelligence systems, with products and services rapidly moving from pattern recognition and insight generation to more sophisticated forecasting techniques and, thus, better decisions. Without effective data governance, data inconsistencies in different systems across an organisation might not get resolved. For example, different ways to register the same data across various databases could complicate data integration efforts and create data integrity issues that affect the accuracy of business intelligence (BI), reporting and analytics applications. In addition, data errors might not be identified and fixed, further affecting BI and analytics accuracy.

In the same time, the right to privacy and to the protection of personal data are fundamental rights in the EU and need to be properly implemented. The application of the principle of "Privacy by Design"[222] needs to be effectively put in place within the developments of products and services under various ICT technologies. With digital transformations of the public administrations towards a more "user-centric" approach, the data security and privacy protection issues become even more essential. To be able to fulfil these needs, the digital/ICT strategies of the EU Member States include from the beginning principles concerning data protection and privacy, as well as the ones covering the cyber-security of the infrastructures.

The concern for personal data protection is not a new issue as Member States have had in place acts and regulations regarding this matter since the 90s. Some of the countries were early adopters, such as France and Germany (1978), Luxembourg (1979), Ireland (1988), Netherlands (1991) and Belgium and Spain (1992). However, technological advancements and the increasing importance of data in daily lives raised awareness for better and more secure legal frameworks. With the General Data Protection Regulations (GDPR)[223], the EU brings forward a set of principles and rules

220 European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A European Strategy for Data (Brussels: European Commission, 2020).

221 Margret Rouse, "What is data governance and why does it matter?", TechTarget, last accessed on 02 March 2020.

222 The privacy and data protection are embedded throughout the entire life cycle of technologies, from the early design stage to their deployment, use and ultimate disposal.

223 O.J. of EU, Regulation (EU) 2016/679 of the European parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free

that aims to ensure the protection of natural persons with regard to the processing of their personal data, respecting their fundamental rights and freedoms. The Member States embedded this new set of regulations in the national legislation frameworks by only amending the existing regulations (two countries: Hungary and Portugal), by amending existing regulations and setting up new laws (seven countries: Austria, Czech Republic, Denmark, Germany, Finland, Netherlands, Slovenia and Sweden) or only by setting up new legislation (the rest of the countries).

In **Croatia**, one of the strategic objectives of the digital transformation is to ensure secure data exchange for all public sector bodies through a central interoperability system (Government Service Bus).

In **Denmark**, the security and confidence safeguard the sense of security in an ever-more digital society and improve information security in the public sector and enhance the digital competences of citizens and of businesses. Data protection will, as far as possible, be incorporated into the design and development of public IT solutions from the very outset. Investing in security comes at a price. However, in an ever-more digital world, data protection is an increasing necessity to ensure that people and businesses have confidence in digital public services.

The Digital Agenda of **Estonia** includes a detailed approach on two important focus areas - development of information society and increasing cyber security. The strategy assesses that the strength of the public sector lies in the systematic development of the state information system, based on agreed principles such as distributed service-based architecture, high security objectives for data and data exchange, web-based, e-service orientation, and strong authentication tools.

The Digital Strategy of **France** defines the standards required to ensure personal data protection and security. Industrial development should not disregard the importance of the right of the citizen to decide on the communication and use of its personal data and the privacy and protection of such data should always be at the forefront.

In **Germany**, the strategy mentions that the eGovernment must secure and ensure adequate data protection in order to gain the trust of both businesses and citizens. In addition, the strategy also puts the principle of data minimisation at the core, alongside the need of state-of-the art technological solutions, with appropriate organisational measure and reliable IT service providers.

The Public Service Data Strategy (2019-2023) of **Ireland** sets out a detailed vision on how to approach data use and management within the public services. The strategy aims to put in place measures to improve data governance, management and re-use in a secure, efficient and transparent manner, for the benefit of citizens, businesses and policy makers.

In **Italy**, the digital transformation strategy of the Public Administration includes, in the Triennial Plan, the principle of trust and security. The principle states that the protection of personal data, the protection of privacy and IT security must be integrated from the design phase. When it comes to digital security for the public administration, it is important to increase the level of security of data and digital communications to enable new levels to the services for citizens and business. The ultimate goal is to protect the privacy, integrity and continuity of the Public Administration services, a truly critical infrastructure for the country.

In **Malta**, the development of the National Data Portal will contain the Catalogue of Registers and Datasets and provide mapping facilities that enable the discovery and access to the required central or line-of-business records as authorised on the basis of the Role-based access control mechanisms. The identified Official Registers and the Datasets included in the catalogue will be the only source of data and records required as part of the processes used by the public administration functions. The protection of personal data legislation has been recently updated to answer the requirements of the EU wide General Data Protection Regulation.

movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance).

The **Netherland Digitaal** includes a bill prepared by the Ministry of Justice and Security aimed at making data processing easier for collaborating parties. If it becomes easier for different parties to process data together, this could lead to a more efficient approach to fighting subversive crime. The bill also contains several safeguards for the protection of privacy. The strategy pays specific attention to the protection of public values and fundamental rights.

In **Poland**, the digitisation of back-office processes in government administration aims to improve the functioning of government administration through the creation, development and dissemination of standards and good practice in key scopes. This includes IT security policies and processing and protecting personal data, good practice concerning IT systems procurement and system elements: modern, open electronic documentation management systems, tested ERP systems, service and data interfaces between systems within a single institution and between institutions, common e-service platforms, increasing the qualifications of IT staff and officials.

One of the axes of the **Portugal**'s National Action Plan for Open Administration is developed around the open data concept, promoting the availability and reuse of information generated by the public administration. Additionally, the national ICT strategy 2020 includes as one of the guiding principles "Data security, resilience and privacy, in order to ensure the protection of information held by the Public Administration."

In **Romania**, the National Interoperability Framework stated under the principle of security and confidentiality that public services need to define a common security and confidentiality framework and set up the public services processes in order to ensure a safe and reliable data exchange between the public administration institutions and between the public administration and the citizens and the business environment. The issue of data protection and confidentiality is again mentioned in the open administration principle, within the context of the system interoperability area.

In **Slovakia**, when it comes to services for citizens and businesses, the key factor of success is the security of electronic services and systems with respect to information of citizens' personal data. In this respect, one of the eGovernment layers is focused on security. The layer aims at providing a secure digital environment with the focus on the protection of citizens' identity and their personal data.

In Digital **Slovenia** strategy, the principle behind development is to provide a high level of personal data protection and communication privacy, taking account of proportionality to the objective. Its implementation aims to promote trust in cyberspace and protect the privacy of communication and information. In the development of ICT solutions, with data protection by design and by default at the centre, it becomes important to assess the impacts on privacy, the use of personal data anonymisation techniques, and the use of encoding methods.

One of the strategic goals of the ICT strategy of **Spain** is to implement the smart corporate management of knowledge, data and information in order to capitalise on this asset and improve the efficiency of the Administration to the benefit of citizens, while ensuring the protection of their digital identity. As the information is a vital resource in an organisation, its production, storing, protection, sharing and/or publishing should be done in compliance with the highest standards, making it available at the right time and accessible from any place, with the required security and privacy assurance.

One of the focus areas of **Sweden**'s digital strategy (Sustainable Digital Sweden) is digital security that aims to increase digital security and to entail people, companies and organisations having trust and confidence in the use of digital services and being able to use them easily. Sweden has an objective to be the best in the world at utilising the opportunities created by digitisation. For this reason, the government established the Swedish national digitalisation council in spring 2017.

All the digital strategies across the EU Member States are complemented by separate National Cybersecurity Strategies. These strategies are plans of actions designed to improve the security and resilience of national infrastructures and services. They are a high-level top-down approach to cybersecurity that establishes a range of national

objectives and priorities that should be achieved in a specific timeframe. And, apart from tackling cybersecurity risks, the strategies also build on collaboration between stakeholders.

1.6.3 Transparency and public sector information

Digital transformation plays an important role in supporting governments to engage with citizens and increase the government's activity transparency. At the same time, transparency of data, processes and decision-making within government is an important area that can be enhanced through digital transformation efforts. It will enhance the sense of accountability and trust between citizens and state in ways that are essential to the wellbeing of populations, and deliver higher quality services to citizens, proactively responding to their needs. Moreover, transparency opens up access to the information related to the underlying processes and decisions made by governments, and encourages the feedback and performance evaluation of the public services and their delivery from the citizens' side. Digital technologies contribute to both strengthening trust in public institutions and reinforcing civic engagement. In France, the Digital Economy Law (*Loi Lemaire*) specifies that algorithms used in public decision-making should be transparent to the public (except where they are involved in detecting fraud).

The principles of digital government create opportunities for citizen-driven activity and a resulting service landscape tailored to their needs. In the same time, these efforts are matched by a need to explore approaches to transparency and privacy rights that empower citizens and ensure that they are fully aware of how their information is being used. Also, by adopting a citizen-centred attitude, public sectors are well placed to embrace innovation and rapidly normalise emerging technology where it can add most value. A way to support governments in having a clear strategy on issues of digital security in the context of the other digital government activities is to develop national Digital Security or Cyber Security strategies. In this context, ethical frameworks for the (re-) use of data also play an important role in governmental policy developments.

Transparency of the public administration services is one of the principles included in the development of the national interoperability frameworks. It is also one of the principles frequently used in the set-up of the national digital transformation and government strategies. The digital strategies include often the provisions mentioned by the European Directives[224] in place at the moment of their set up related to the open data and the re-use of public sector information.

In **Austria**, transparency is a way of involving all the parts affected by the development of new technologies the process and, therefore to increase the adoption rate of the new developments. The disclosure of non-personal administrative data - such as Open Government Data - is a possible way of increasing participation in a common process of value creation between politics, administration, citizens and business. For example, in May 2011, the Vienna city administration opened its non-personal data to the public. Other counties, cities, municipalities and federal ministries followed and new ones are constantly emerging. The Austrian administration, in cooperation with partners from industry, and with the help of the cross-organizational platform "Cooperation OGD Austria", has developed the necessary technical framework for user-friendly open government data in Austria.

In this sense, the Oesterreich.gv.at is the government platform linking a large number of public authorities and providing numerous services to citizens and businesses.[225] The Austrian Open Government Data Portal[226] includes more than 27000 datasets from 1214 organisations. These datasets have been used to create over 500 applications for different platforms (iOS, Linux, Symbian, Microsoft, Android etc.). One example is @OpenDataATAssistant[227], a Facebook bot that will help the user to explore the content of the Austrian open data portals: data.gv.at and opendataportal.at. Another

[224] Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information, Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information, Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast).

[225] https://www.oesterreich.gv.at/oeservices.html.

[226] https:// data.gv.at/.

[227] https://www.data.gv.at/anwendungen/opendataatassistant-facebook-bot/.

one is OpenTrees.org[228] application created to explore the combined urban tree datasets of 150+ cities around the world, including Linz and Vienna. The application includes a number of visualisations, including by species, rarity, health or maturity.

In **Belgium**, one of the five goals of the digital strategy of is the digital government[229], where the transparency is seen as part of the operational efficiency of the digital transformation. The government will use new technologies, such as social media and big data, with a clear objective – to provide better services at lower cost to citizens and businesses, and to streamline the government management via digitalised services and processes. Transparency is part of a better government process. Therefore, the federal government will ensure that data, held at federal level, is accessible in a user-friendly manner through a single open data portal, ensuring both the privacy and security of the aforementioned data.

The single point of entry for the federal portal[230] provides compact information about a large selection of the services available to both citizens and businesses. At the same time, the FedWeb portal[231] offers all the information need it on the federal administration/government and its available services. The open data portal[232] includes more than 10000 datasets with numerous tools and application developed at national and regional level – 67 apps covering different domains, such as economy and finance (22), population (23), transport (13), environment (9), health (7) etc.

In **Bulgaria**, the digital transformation of the public administration is one of the main objectives of the eGovernment strategy. The process implies the integration of information processes through the development of systems and services with effective electronic document exchange and archive. The public administration will provide integrated e-services to citizens and businesses, increasing transparency and efficiency of its processes as well as saving resources. It aims to achieve also a better transparency level in the decision-making process.

A platform for access to public information was created and maintained by the administration of the Council of Ministers233. It is a unified, central, public web-based information system that provides electronically the entire process for submitting and reviewing an application for access to information and publishing relevant information. A unified portal for access to the administrative services is also available[234], providing information to the administration, citizens and businesses. The open data portal[235] hosts more than 10000 datasets from 485 organisations, and has over 1300 registered users.

In the eGovernment strategy of **Croatia**, the "open government" includes closer approach to the public, transparency, responsibility, efficiency and better cooperation with citizens. An e-consultation portal for the Public Administration at government level was developed to address the shortcomings of the current system. The portal is accessible through platform of e-citizens. The initiative aims to facilitate public participation in policy-making and legislation, as the Croatian legal framework requires for all public authorities (from national, regional or local level). Moreover, a joint Content Management System of the Croatian Government will be further developed, with the aim to unify all the websites of public sector bodies to ensure a standard approach to all their websites, and access to the same data in all of them. This will provide an easier access to information for citizens, businesses and public administration employees.

The State Administration Portal[236] offers access to all information in one place. It facilitates access to executive information through a central online information solution. The portal presents unambiguously, simply and modern the structure,

228 https://www.data.gv.at/anwendungen/opentrees-org/.
229 http://digitalbelgium.be/en/5-priorities/digital-government/.
230 https://www.belgium.be/en.
231 https://fedweb.belgium.be/fr/services_en_ligne.
232 https://data.gov.be/en.
233 https://pitay.government.bg/PDoiExt/indexExt.jsf.
234 http://unifiedmodel.egov.bg/wps/portal/unified-model/home.
235 https://data.egov.bg/.
236 https://www.gov.hr/.

function and role of all state administration bodies. HITRO.HR[237] is service of Croatian Government for the quick communication of citizens and business entity with the state administration. The e-citizens portal[238] is also available, aiming at modernizing, simplifying and accelerating communication between citizens and the public sector and increasing the transparency of the provision of public services. The open data portal[239] was established in 2015 and hosts over 600 datasets form 86 publishers.

The Digital Strategy for **Cyprus** aims using the ICT as a catalyst to increase productivity and economic growth. The approach will have a direct impact on the modernization and productivity increase of the public sector, the increase of transparency and the promotion of democracy and culture. Through the use of ICT the government aims at becoming smart, sustainable and innovative, more effective and efficient and friendlier to citizens and businesses. It also intends to decrease bureaucracy and costs, by becoming paperless. Furthermore, by providing public services electronically the government will offer better services to businesses and citizens and reduce business's cost.

The Cyprus' open data portal hosts more than 1000 datasets from approximately 90 organisations. The applications on the portal cover different areas, such as health (e.g. PharmaCY Near Me), environment (e.g. Simply, a Cyprus Air Quality Service), transport (e.g. Cyprus bus by motion), government and public sector (Chatterbox.Network) etc. The Cyprus Government portal[240] is a single-entry point for all the necessary information about online public services and government websites for citizens, administration and businesses. It contains extensive information on the topics relevant each of the users' categories. In addition, a point of single contact[241] for the business community was also set up. The portal offers the necessary information for the establishment of businesses in the services sector and the cross-border provision of services within the internal market.

Openness and transparency are principles included in the development of the eGovernment strategy of **Czech Republic**. Within their context, public authorities need to share information and data with each other and to give citizens and businesses access and control to their own data. The public authorities need to involve stakeholders (for example, businesses, researchers, and non-profits) in the design and delivery of services and work with, facilitating the users' monitoring of the administrative processes. The strategy also aims at creating *a public data pool* that will contain the published public data of the administration. It will also be the basic method for sharing public information between public bodies and for the sharing of public data between public and private sphere in the Czech Republic. The public services will use a common approach for the creation and delivery of their services to reduce, and eventually eliminate, redundancies and duplications.

The Czech POINT[242] is a project whose aim is to reduce excessive bureaucracy in the citizen - public administration relationship. The gateway to all public services for citizens, businesses and administrations is the public administration portal[243]. The citizens' portal[244] allows users to receive and send data messages, manage their Basic Registers data, store and manage documents and papers, submit any claim to any office. Portal for Data Boxes[245] was launched in June 2011, and provide a more comprehensive service to users of Data Boxes, which serve as a secure repository of official electronic communications with public authorities.

The **Denmark**'s Digital Growth strategy aims to analyse the opportunities for full digitalisation of physical planning and planning data. This digital transformation will give companies, investors and people a better basis for construction, conversion and

237 https://www.hitro.hr/en/homepage.
238 https://pretinac.gov.hr/KorisnickiPretinac/eGradani.html.
239 http://www.data.gov.hr/.
240 http://www.cyprus.gov.cy/portal/portal.nsf/citizen_en?OpenForm&access=0&SectionId=citizen&CategoryId=none&SelectionId=home&print=0&lang=en.
241 http://www.businessincyprus.gov.cy/mcit/psc/psc.nsf/index_en/index_en?OpenDocument.
242 The Czech Post Verification Information National Terminal, https://www.czechpoint.cz/public/statistiky-a-informace/co-je-czech-point/.
243 https://portal.gov.cz/index/.
244 https://obcan.portal.gov.cz/prihlaseni.
245 https://www.datoveschranky.info/.

data-driven innovation and will ensure that planning processes are transparent and efficient for this purpose. The aim of the digital strategy for 2016-2020 was to make Denmark stronger and more secure digitally, but to also deliver a user-friendly and simple digital public sector. Both citizens and businesses need to have easier access to data about themselves held by a specific authority. Furthermore, they should have a greater insight into their cases, applications, data and relationships with the public authorities. The public digital services are adapted to the current situation and needs of the individual citizen or business. The user meets personalised, relevant content on public-sector websites or portals that they are browsing. The Agency for Digitisation is an agency within the Ministry of Finance, established in 2011, to be in charge of the government's digitisation policies. The agency is responsible for the implementation of the government's digital ambitions and the use of digital welfare technology in the public sector.

The data distributor platform[246] provides access to the users to their public basic data from the Danish authorities (the platform includes the basic data register). The citizen portal[247] is another platform developed that provides extensive and detailed information on public services and other issues related to living in Denmark. The business portal[248] is another platform developed to ease the access to administrative services for the enterprises and entrepreneurs.

The digital strategy of **Estonia** uses of ICT to improve the efficiency of public administration and the quality of public services, including the implementation of e-democracy in the exercise of public authority. The public services need to be simpler and more efficient for users. The aim was to provide services such that the users' needs are met integrally and rapidly, as proactively and "invisibly" as possible (i.e. with the minimal need for intervention by the user). The ICT solutions streamline the procedural processes and promote multi-channel service delivery capabilities, contributing to both the efficiency of national and local administration and increase the availability of services. The public services and ICT support services will be developed integrally in order to do avoid the need to submit data repeatedly to institutions, while the various agencies must re-use and cross-use data already submitted by persons, within the constraints of the purpose for which they were submitted, and as long as the persons do not prohibit such use.

The eGovernment Portal[249] is the gateway to government information and e-services for both citizens and business. It was established in 2003, and was recently redesigned to better fit the users' needs. Open data portal provides a single point of access for general public to unrestricted public sector data with the permission to re-use and redistribute such data for both commercial and non-commercial purposes. The portal hosts over 200 datasets from 48 organisations.

In **Finland**, the openness of public administration and the process transparency have increased through active utilisation of online services. A citizen or organisation can initiate an issue and then follow its progress electronically. Citizens also have knowledge of information the public administration has gathered on them and of any issues being processed. The transformation of the public administration through the use of ICT was one of the sub-sections of the 2003-2007 Government's strategy. The following strategy – A renewing, human-centric and competitive Finland (2007–2015), went further in developing public digital services in a customer-oriented and economical manner to fit the users needs (citizens, organisations) using existing the information. In Finland, citizens see services as seamless concepts that correspond to their living situation.

The main entry point to the web services for citizens, organisations and companies is Suomi.fi. The portal provides extensive information on services available to citizens and businesses. Demokratia.fi is a portal that provides up-to-date information about how matters are prepared at the local, national and EU level. It helps citizens to keep track of and support the most recent initiatives, and to take part in discussions and

246 https://datafordeler.dk/, in Danish only.
247 https://www.borger.dk/ in Danish only.
248 https://indberet.virk.dk/ in Danish.
249 https://www.eesti.ee/en/.

consultations in the legislative initiatives. The open data portal[250] hosts over 1700 datasets from 791 publishers. These datasets are used to build various applications (around 40), also available on the portal (websites, maps, tools, visualisations etc.) For example, Coronavirus in Finland[251] is a visualisation tool that uses open data from Helsingin Sanomat, which is updated whenever there is a new infection. The application is available only in Finnish, and shows the number of infections in Finland at the current moment (depending on the data availability) and, e.g. the number of people who have recovered from the virus.

The International Digital Strategy for **France** calls for opening up public data, a mean of creating new services for citizens, improving the functioning of administrations and meeting the democratic demand for transparency of public power. The strategy aims at improving access to administrative documents, widening the obligations to disseminate public data, and opening up the re-use of data collected within the framework of an industrial or commercial public service. These actions bring France to the forefront in terms of open data, by providing a controlled balance between openness and protection of personal data. The strategy promotes an open, diversified and trusted digital world, where the governance is legitimate and effective if it is also transparent, democratic and inclusive, ensuring real representation of all parties (government, citizens and businesses).

In 2011, the Prime Minister announces the creation of the single portal for state public information "data.gouv.fr", considering the application of the provisions governing the right to re-use public information. The department in charge with this development was Etalab, a department of the interdepartmental digital directorate (DINUM). Etalab coordinates and promotes the action of the State and the organizations placed under its supervision in matters of inventory, governance, production, circulation, exploitation and opening of data, and in particular source codes. The open data portal hosts around 35000 datasets from more than 2500 organisations. The datasets from the portal have registered over 2000 re-uses, within different areas, such as transport, health, elections etc. The API catalogue[252] is a platform primarily intended for service creators, the API consumers. For this, the catalogue facilitates the discovery, understanding and access to APIs and their producers. The public administration put at disposal to citizens and businesses a simplification portal[253] for the public services. There are 2476 partner administrations on the portal and more than one million files have been submitted. The overall processing time was reduced by 50%.

Freedom of information and Open Data are essential for the transparency and accountability of government action and they are one of the goals of the eGovernament strategy in **Germany**. The strategy states that the open information about policy-making and administration should be non-discriminatory, targeted to its audience and user-friendly for the various target groups. An important attention will be paid to data protection and security, without reducing the transparency of the government activities and decisions. When it comes to data availability, they should be provided in standardised and machine-readable formats. The central access for citizens, companies and administrations to the electronic advertisements and job offers of the federal, state and local administration, is provided through Bund.de portal. The portal also includes access to information on the authorities and institutions of the federal administration.

The public authorities from the federal, state and local governments offer access to the administrative data via GovData portal[254]. The portal hosts over 38000 datasets from federal, national and local level organisations. Majority of the datasets are under free to use/re-use licence. However, for 698 datasets the use is restricted, and many of these datasets[255] can be found under environment (362), regions and cities (268) and government and public sector (276) categories.

[250] https://www.opendata.fi/en.
[251] Note that the tool does not give an accurate view of the Coronavirus situation in Finland because since 14.3, health authorities are no longer testing all patients presenting with symptoms in Finland.
[252] https://api.gouv.fr/ available in French.
[253] https://www.demarches-simplifiees.fr/ available in French.
[254] https://www.govdata.de/ available in German.
[255] A dataset could be found in several categories, therefore there is overlapping between the categories.

The main goals of the National Digital Strategy (2016-2021) of **Greece** is to support essential reform actions in providing integrated services to citizens and businesses and in improving the effectiveness of Public Administration through the use of ICT. It aims at making a proper use of ICT resources in public administration in order to reduce waste, and improve re-use and share of solutions. The actions taken within the scope of the strategy intend to substantially upgrade the services to citizens and businesses. It also plans to review the relationship between the government and the public administration considering the technological advancements. In this respect, the ERMIS[256] is the central public administration portal, providing information and online services to citizens and businesses. The portal provides comprehensive information to citizens and businesses on all their transactions with the Public Administration (physical or electronic). It also offers selected services of electronic transactions divided in two categories: with and without immediate result receipt. Another portal developed to provide the best technological solutions for an efficient and citizen-friendly Public Administration is the Secretariat-General of Public Administration Information Systems[257]. The portal provides access to services to both public administrations and citizens and businesses. The Ministry of Digital Governance, within the context of the Transparency Program initiative, set up a transparency portal[258], where all government institutions are required to upload their acts and decisions on the Internet with special attention to national security issues and sensitive personal data.

The third pillar of **Hungary**'s Digital Renewal Action Plan (2011) aimed to provide a more efficient and secure public service to the users. The goal was to deliver simpler, more transparent, more secure, cheaper and more efficient services to citizens and businesses. The National Digital Strategy (2014-2020) brought forward the digital state objective from the government side. The digital state approach had in view an effective public administration, with low operating costs and high-quality services. It also aimed to strengthen the businesses and citizens trust in the public administration by increasing the system 's transparency. The eGovernment portal is active since 2003, and offers citizens and businesses the point of single contact in Hungary. The establishment of an open data portal is a more recent initiative (2019), and the site currently hosts around 50 datasets from only six organisations.

The personal data portal (accessible via MyGovID[259]) was set up to provide access to the government services in **Ireland** for citizens and businesses. The initiative builds on Data Sharing and Governance Bill that put in place a legal framework for the safe and secure sharing of data across Government. Embracing openness and transparency by publishing open datasets, the government helps improve public accountability and drive innovation and entrepreneurship. The government portal[260] combines the websites of Irish government departments and is a trusted source that makes interactions with the government more user-focused. The aim of the portal is to present information in a clear, understandable and accessible manner. Since 2001, the citizens' information portal[261] provides information and advice on social services, operating under the aegis of the Department of Employment Affairs and Social Protection. Open data portal was established in 2014 and includes over 10000 datasets from 114 publishers. These datasets have been used to build apps, websites and visualizations. They've been featured in articles, and written about in news reports and blog posts.

The Digital Growth Strategy (2014-2020) for **Italy** aims at establishing an open structure where the various actors of the public administration contribute to their area of expertise. Therefore, the public administration transforms into a user-centric system by developing a single platform, hosting its opened data, and provides the available services to businesses and citizens. To accompany and support the public administrations in its digital transformation process, for transparency, simplification and efficiency, the "Italy Login - citizen's home" project was created. The goal of the

256 http://www.ermis.gov.gr/portal/page/portal/ermis/.
257 https://www.gsis.gr/ggpsdd available in Greek.
258 https://diavgeia.gov.gr/ available in Greek.
259 https://www.mygovid.ie/en-IE/Home/Index.
260 https://www.gov.ie/en/.
261 https://www.citizensinformation.ie/en/.

project is a "home" on the Internet for every citizen, a single sign-on for all services of the public administration.

The Public Digital Identity System (SPID) is the solution that allows the access to the online services of the Public Administration for citizens and businesses. The service facilitates this access based on a single Digital Identity (username and password) that can be used on computers, tablets and smartphones. The identification can also be used to access the information on the eGovernment portal for business[262], avoiding identification duplications for entrepreneurs.

In the eGovernment strategy for **Latvia** one of the basic principles of the ICT transformation is the access to information. The process of modernisation of the public administration basic activities through digitalisation is based on the service quality improvement, decision-making transparency and higher standards and responsibility from the administrative side. Using ICT solutions properly would provide the optimisation of the public administration resources and reduce the burden on citizens and businesses, as well as improve of the transparency and operational efficiency of public administration itself.

The state and local government portal[263] was developed to ensure the access to public services for both citizens and businesses. The aim of the portal is to offer quick and convenient access to these services. A national open data portal[264] was created in 2017, to gather in one place and to disseminate for public use the data collected by government institutions and organizations to support the development of innovations. The portal hosts currently 366 datasets from 73 publishers.

The data openness is the forth objective of the **Lithuania** Information Society Development Programme 2014 – 2020. By making the data of state and municipal authorities and agencies available to the public and business, the programme encourages their use for innovative solutions and e-service creation, and also creates favourable conditions for businesses to implement and use ICT to improve their efficiency and competitiveness.

In 2004 (redesigned in 2015), an eGovernment gateway[265] was put in place to provide access to administrative and e-public services for everybody – the public sector, citizens and businesses. The portal is designed to be easy for the user to understand, and lists on its home page three portal user groups: citizens, business entities, and service providers (public services). An open data portal is currently under-development, and the beta version of this portal has gathered 314 datasets from seven organisations. The portal will become fully functional by the end of 2020.

The implementation of the Digital Lëtzebuerg action plan, and the digitisation of administrative procedures and digital transformation of the public administration are important objectives for **Luxembourg**. Their Data-driven innovation strategy puts more emphasis on the use of Big Data, analytics and artificial intelligence in terms of process efficiency, demand forecasting, inventory management, real-time end-to-end transparency and route planning. In order to be effective, the transparency and the trust between the stakeholders of the chain are essential. An important role is attributed to the data analytics that allows optimising the underlying logistics services and processes related to activities.

The general information on various subjects, together with a thematic directory of links that guide users to sites offering detailed information on the topic(s) of their interest are provided through the official portal of the Grand Duchy of Luxembourg[266]. When it comes to open data[267], the Luxembourgish portal hosts more than 1000 datasets from 143 organisations. The open datasets are found in over 100 developments of applications and APIs (with a wide range of transport/mobility apps).

262 http://www.impresainungiorno.gov.it/web/l-impresa-e-l-europa/doing-business-in-italy.
263 https://www.latvija.lv/en.
264 https://data.gov.lv/eng.
265 https://www.epaslaugos.lt/portal/en.
266 https://luxembourg.public.lu/en.html.
267 https://data.public.lu/en/dashboard/.

In the Digital **Malta** strategy, under the Government (Public Administration) theme, the document states "better application of digitisation will result in reduced bureaucracy and transparency. Information sharing across government systems will be promoted, as will be the re-use of public sector information by third parties." The Gov.mt[268] is the main gateway for governmental information and services. And the Malta Information Technology Agency (MITA)[269] is the central driver of Government's Information and Communications Technology (ICT) policy, programmes and initiatives in Malta. The Servizz.gov[270] is the one-stop shop portal in local communities and the online guide to government services for the citizens and businesses. The open data portal[271] was become functional in 2018 and currently includes 200 datasets from three publishers.

The Netherlands Data Agenda Government covers several action points linked to the transparency issues. The Government is taking the initiative to formulate wide principles for the responsible handling of data and to draw up a Code of Good Digital Administration for the benefit of inter-governmental collaboration. It is also starting an intergovernmental collaboration on a 'transparency lab'- a place where applications are developed and tested, aimed at increasing governmental transparency about data, source codes and algorithms.

The open government data portal, Overheid.nl serves as the central access point for all information relating to government organisations. The portal provides information about services for persons and businesses by themes, life events and location. The portal includes up to 13000 datasets from 176 organisations. The business portal[272] posts information and advice from the government related to the needs of the businesses. It covers legislation, tax rules, subsidies, entrepreneurial events and industry information.

The Poland's eGoverment strategy aims to provide high-quality e-services adequate to real needs of the citizens and entrepreneurs, delivered by modern IT solutions, developed and maintained with the cooperation of all actors at various levels of public administration. It intends to connect distributed institutions and change complicated procedures into coherent and simple services. A priority task is to allow the widest possible range of public services to be provided digitally, enabling citizens to handle their business remotely.

The government portal of Poland[273] is the main gateway to the government information and services for citizens, businesses and public administration. It provides relevant information, as well as press and multimedia materials, and a catalogue of all services, information and public administration entities. The open data portal[274] offers access to public data from various categories, such as education, the environment, budget and finance, culture, security, sport and tourism, job market and others provided by public intuitions (ministries, agencies, local governments). The portal hosts more than 1300 datasets from 133 providers and have been used in 29 application developments, including in health (e.g. Home quarantine, an application that facilitates and facilitates the implementation of mandatory quarantine at home) and transport (e.g. Railway stations in Poland, which provide information about the location of all train stations in Poland, with distinction from premium stations).

The Public Administration Digital Transformation Strategy of **Portugal** intends to increase system transparency. For this, the information produced by the public administration needs to be made available in a transparent way in order to develop innovative services performed by the administration itself and by the civil society. The strategy focuses also on the accessibility of electronic services, promotes proximity and context cost reduction, and the participation of citizens and companies in the design of new services, in order to improve their usability.

268 https://www.gov.mt/en/Pages/home.aspx.
269 https://mita.gov.mt/en/Pages/The-Agency.aspx.
270 https://servizz.gov.mt/en/Pages/default.aspx.
271 https://open.data.gov.mt/index.html.
272 https://ondernemersplein.kvk.nl/ available in Dutch.
273 https://www.gov.pl/.
274 https://dane.gov.pl/.

The Government Portal[275] is the institutional website of the Portuguese Government. The portal publishes information on various issues related to the government activities such as the Government Program, the Government composition, political communications or on-going public consultations. The gateway to public services for citizens, businesses and public administrations is the ePortugal portal[276]. The portal aims to facilitate interactions between citizens and companies and the State, making them clearer and simpler. Simplex+ programme[277] is a collaborative and nationwide simplification programme launched by the Portuguese Government to co-create new online public services, optimise existing ones and de-bureaucratise the relationship between public institutions and civil society. The open data portal[278] aggregates, references and hosts open data from different Public Administration bodies and sectors, being the central catalogue of open data in Portugal. It includes over 2000 datasets from 89 organisations.

The first field of action of the Digital strategy of **Romania** aims to increase the public administration's efficiency and reduce the public sector costs by a digital transformation. It intends to reform the way the government works, and to provide access to information, engage the citizens and provide services to internal and external beneficiaries to both benefit of government and citizens. Increasing the transparency and efficiency of the public administration through digitisation of the public services are part of the goals of this field of action. The strategy aims to improve the business environment and the governance of electronic public services' implementation.

The central gateway to access the public services is the electronic single point of contact portal[279]. It provides access to public services for citizens and businesses. When it comes to public data availability, an open data portal[280] was developed that currently hosts over 1800 datasets from 102 organisations. The datasets have been used to build eight applications (e.g. apps, websites and visualizations).

The Digital Strategy of **Slovakia** aims to transform the public services towards more user-based needs services. It implemented the "once is enough" principle, ensuring good coordination between different public administration services to achieve this goal. The services are provided to the public automatically on the basis of events and not only on the basis of the service beneficiary impulse.

The Information Society Division[281] provides up-to-date and comprehensive information on strategies, legislation, standardisation and other activities concerned with the information society. It operates under the Office of the Deputy Prime Minister for Investment and Information Technology. The open data portal[282] provides access to over 2000 datasets from 89 organisations.

Digitisation and optimisation of internal operations for a flexible, rational, efficient, transparent and open public administration is one of the strategic objectives in **Slovenia**'s governmental digital transformation strategy. The strategy aims also to improve the communication and harmonisation between citizens and the public administration using the digital channels. Setting up the Government Cloud was the next step to the simplification and streamlining of development of the public sector infrastructure, to improve the national authorities' operation. It put in place a unified and consolidated information system for the public administration, with reference architecture and standards.

To avoid duplication of content on national websites and the main government portal, the Slovenian government approved the development of a new centralised portal[283] and it launched in 2019. The portal contains extensive information on the government and the public authorities for both citizens and businesses. The Ministry of Public

275 https://www.portugal.gov.pt/pt/gc21.
276 https://eportugal.gov.pt/en/inicio.
277 https://www.simplex.gov.pt/.
278 https://dados.gov.pt/en/.
279 https://edirect.e-guvernare.ro/SitePages/landingpage.aspx.
280 https://data.gov.ro/en/.
281 http://www.informatizacia.sk/o-nas/.
282 https://data.gov.sk/.
283 https://www.gov.si/.

Administration made available a Stop bureaucracy portal[284] in order to implement the principles of better regulation to achieve a regulatory environment with minimal burden to its users, improve the competitiveness of the economy, accelerate the market openness, improve standard and increase transparency. The Open data portal[285] is the single national online publication point for open data for the entire public sector. The portal hosts around 4000 datasets from more than 250 organisations.

The public system transparency is one of the main principles for the Digital Transformation Plan for the General Administration and Public Agencies in **Spain**. The strategy aims to transform the public services in more user-oriented format and take in to consideration citizens', businesses' and public servants' needs. It aimed providing more accessible, usable, simple and secure services to all its users. The digital tools helped to improve the administration's activity, by building on a holistic view of its common actions and needs, ensuring robust evolution and avoiding redundancy. The public sector in Spain offers a wide variety of e-services, giving citizens and businesses the chance to complete administrative procedures by electronic means.

The general access point[286] provides the access to government information, allows doing paperwork and accessing the information need to accomplish their objectives for both citizens and business. The portal offers access to the Citizen's Folder[287], a one-stop service where a user can access and browse all the information the administration possesses about him or her, either citizens or businesses, including open proceedings and registry entries across all public organisations. The eGovernment portal[288] is a centralised information point about the current situation of eGovernment in Spain: news and events, reports, studies and gazettes, legislation, organisation and strategies. The portal includes also a catalogue of services grouped in four categories: promotion of the Digital Administration and Citizen Services, internal management, infrastructures, and regulation, guides and relevant reports. The open data portal[289] is the platform that hosts the National Catalogue of Open Data, the single-entry point through which Spanish public administrations make their data available to citizens, researchers, re-users and other administrations for consultation, download and reuse. In the portal, more than 27000 datasets can be found, and they have been used in 298 applications in 22 different areas. The top five areas for these applications are transport (56 applications), public sector (46), culture and leisure (32), environment (26) and economy (22).

One for the main objectives of the Digital Strategy of **Sweden** (A Sustainable Digitalised Sweden) is to achieve the digital leadership, where activities will be improved, developed and enhanced through governance, measurement and follow-up. The objective focuses on area such as clearer central government leadership, simplification through digital transformation, continuous analysis of digital maturity, resource-efficient society's governance through digital transformation and strengthened local and regional engagement.

Since 2012, Sweden has adopted the Government Strategy for Collaborative Digital Services in Government Administration that puts citizen at the centre. The strategy stated that digital services must be developed in a user-centric way: simple and secure to use, and easily accessible to everyone.

The Government portal[290] is well structured and provides documents and records, information about current government bills, initiatives and ministerial activities, and accounts of how the decision-making process works in Sweden. The Swedish Business Link to Government[291] portal provides a comprehensive single-point for entrepreneurs and enterprises to access relevant and official eServices and information from three public authorities: the Swedish Companies Registration Office, the Swedish Tax

284 https://www.stopbirokraciji.gov.si/en/home.
285 https://podatki.gov.si/ in Slovenian.
286 http://administracion.gob.es/pag_Home/en/index.html#.XnsZy2RKifU.
287 https://sede.administracion.gob.es/carpeta/clave.htm in Spanish.
288 https://administracionelectronica.gob.es/pae_Home/?idioma=en.
289 https://datos.gob.es/en.
290 https://www.government.se/.
291 https://www.verksamt.se/en/web/international.

Agency and the Swedish Agency for Economic and Regional Growth. The Swedish Open Data portal[292] is the Swedish portal for the re-use of public sector information, originally developed by the Swedish Innovation Agency (VINNOVA). The portal hosts around 1400 datasets from 116 organisations. The environment registers the highest number of datasets (458), followed by education, culture and sports (223) and energy (188).

1.6.4 Digital Skills

Every Member State considers digital skills development essential for future progress. Most digital strategies aim at improving these skills at both public administration level and citizen level as such. Often the digital skills development plans are included in the main body of the strategy (either digital or eGovernment) as one of the priority areas. Others, such as Netherlands, Luxembourg and Ireland, set up at national level additional initiatives related to digital education.

With the Strategy for Digital Growth (one of the thematic priorities is "*Digital skills for all*"), Denmark aims to further strengthen the digital skills of Danes via a range of specific initiatives, with some aimed at employed and unemployed people, and other at children and young people.

"*Renewal by default*" is one of the principles for a human-centric, thriving and balanced data economy of Finland linked to the digital skills development through continuous learning. The principle states that in a data-driven society, the education systems should provide opportunities for individuals' lifelong learning. Furthermore, the individuals should also actively learn new skills and gain know-how that will help them to adapt and contribute to the data economy.[293]

Digital transformation is not solely about the citizen experience but is also about helping public servants be more effective in their jobs, equipped to respond to the changing needs of the organisations they work for and the communities they serve. Investing in the user experience of public servants, and their skills, is important in contributing to their increased job satisfaction (and by extension well-being) as well as positively impacting on citizens accessing those services who get a better, more effective, experience as a result.[294] Clear and solid leadership is essential for open data policies to thrive.[295] Within public institutions, the leadership can provide the political support and commitment needed to prioritise the data's release and the financial and human resources support for further sustainable developments. The data and technological developments bring forward the need for more specialised job in data coordination at governments and/or public administrations level. The institutional governance model is also a core element of good data governance, as it provides clarity in terms of leadership and accountability.[296] The leadership helps to implement and steer policy design and implementation, increasing the continuity and sustainability needed to deliver results across political terms. It also co-ordinates the different administrative actions covering the entire government data value chain in the public sector and monitors the implementation of central policy and technical guidelines supporting data governance. Some countries have formalised leadership roles by attaching them to existent administrative structures (Austria, Czech Republic, Estonia, France, Ireland, Poland), while others have followed different leadership models, less hierarchical and shared by different individuals, that respond more to the culture within their public sector (Belgium, Denmark, Finland, Portugal, Sweden).

292 https://oppnadata.se/#noscroll.

293 EU2019.FI - Principles for a human-centric, thriving and balanced data economy, Finland's Presidency of the Council of the European Union, 2019.

294 Welby, Benjamin (2019), "The impact of digital government on citizen well-being", OECD Working Papers on Public Governance, No. 32, OECD Publishing, Paris, https://doi.org/10.1787/24bac82f-en.

295 OECD (2018), Open Government Data Report: Enhancing Policy Maturity for Sustainable Impact, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/9789264305847-en.

296 OECD (2019), The Path to Becoming a Data-Driven Public Sector, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/059814a7-en

Base on countries' approaches, the formalised leadership could be the chief data officer position, as in Estonia, France and Poland[297], or chief digital officer, as in the case of Austria, Ireland, Malta, and Czech Republic. The chief data officers are expected to make a measurable impact upon how public institutions create, store, manage, use and share the data and to strengthen evidence-based policy making. The data governance arrangements vary in their form and articulation with the ICT governance arrangements. For example, the Chief Data Officer in France has also become the government's Chief Information Officer, accumulating the two functions.

On the other hand, Denmark has adopted an alternative model for data governance, where data policy, standards, guidelines and interoperability frameworks and platforms fall under the responsibility of the government. To ensure the coordination and the commitment to the national strategy across the public sector, Denmark set up the Steering Committee for Cross Government Co-operation, after an agreement reached between the government, Danish regions and the local governments of Denmark.[298] The Committee is a cross-governmental coordinating body aiming to create a common ground in the work on digital government. The responsibilities of the Committee are to determine the overarching principles and coherent framework conditions for digital government, coordinate initiatives across the public resources for better use of it, deciding on resource allocations, and determine models for digital government operations and maintenance of projects.

In Portugal, the Agency for Administrative Modernisation is responsible for the approval of ICT projects (over EUR 10000) in line with the norms and guidelines defined by the eGovernment Network. Moreover, the Agency's team is in close contact with the focal points at institutions relevant for the implementation of the digital government projects to monitor their roll out.[299]

The ICT Strategy Commission in Spain is an inter-ministerial body at the highest political level, which is responsible for the definition of the strategy, but also with the services to be shared.[300] The commission determines the priorities for the investments, reports on draft laws, regulations and other general standards with the purpose to regulate ICT matters for the general state administration. It also promotes collaboration with the autonomous regions and local authorities for the implementation of integrated inter-administrative services.

2 ANNEX 2: FULL LIST OF CASES

Title	Geo	Sector
Data Analytics Framework (IT): the Italian strategy for data analytics, part of the digitization plan	Italy	General
Public Service Data Strategy (IE): stragtegy for a data ecosystem 2019-20123	Ireland	General

297 OECD (2018), Open Government Data Report: Enhancing Policy Maturity for Sustainable Impact, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/9789264305847-en.

298 OECD (2016), Digital Government in Chile: Strengthening the Institutional and Governance Framework, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/9789264258013-en.

299 Ibid.

300 Ibid.

NL DIGITAAL: Data Agenda Government (NL): cross departmental collaboration for a data driven approach	Netherlands	General
Information policy (FI): national digital strategy for human centric technological solutions	Finland	General
Barcelona Data Commons (ES): citizen centric approach to data governance	Spain	General
Mayor's Office for Data Analytics in NYC (US): civic intelligence centre aggregating data from different agencies to address civic problems.	United States	General
FINDATA and KANTA	Finland	Healthcare
City Data Analytics Programme	UK	General
Denmark data strategy	Denmark	General
Tools for Innovation Monitoring (TIM)	EU	General
Media Monitoring and Analysis for Policy Support - MMA	EU	General
Big Data Cooperational System	Italy	Home and justice affairs
Big Data Analysis for HR efficiency improvement in Ministry of Public Administration of Republic Slovenia	Slovenia	General
Risk Assessment and Horizon Scanning (RAHS)	Singapore	General
Public Policy.ie	Ireland	General
Global Pulse	United Nations	General
OpenFisca	France	Economic/Fiscal/Financial
Transport for London data analytics	UK	Transport
The Big Data for Law project	UK	Justice/Law
Flanders Education approach to benchmark data	Belgium (Flanders)	Education
New Zealand Data Governance Framework	New Zealand	General
Geodata Strategy of the National Land Survey Authority in Sweden	SE	Geographics
Colombia Digital Government Policy	CO	General
Mexico Open Data Policy	Mexico	General
Korea Open Data Policy	KOR	General
Estonia's X-tee platform	Estonia	General
Udbetaling Danmark	DK	General

3 ANNEX 3: DATA COLLECTION ACTIVITIES

3.1 NL Digitaal - Data Collection Activities

3.1.1 List of consulted stakeholders

Interviews

Marieke Schenk, Coordinator Data-Driven Approach Learning and Expertise Centre (LED) (In Dutch: Leer- en Expertisepunt Datagedreven werken - LED), ICTU

Marcel Hopman, Programme Manager Data Agenda Government, Ministry of the Interior and Kingdom Relations (BZK), Government of the Netherlands

Presentation

Désirée Geerts, Head of Information Society Unit, Ministry of the Interior and Kingdom Relations (BZK), Government of the Netherlands

NL DIGITAAL: Data Agenda Government, Presentation at the Lisbon Council High-Level Roundtable and Working Lunch on Data for Policy, 31 January 2020, Brussels.

3.1.2 List of consulted documents

BZK (2019), NL Digitaal*:* Data Agenda Government (Data Agenda Overheid)*, Government of the Netherlands, https://www.nldigitalgovernment.nl/wp-content/uploads/sites/11/2019/04/data-agenda-government.pdf (accessed on 23 January 2020).*

BZK (2018), Monitor Generieke Digitale Infrastructuur 2018, Government of the Netherlands*, https://www.digitaleoverheid.nl/wp-content/uploads/sites/8/2018/09/GDI-Monitor-2018-v1-def.pdf (accessed on 5 February 2020).*

European Commission (2018), eGovernment Benchmark 2018*. Securing eGovernment for all FACTSHEETS, Luxembourg, Publications Office of the European Union, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=55490 (accessed on 24 January 2020).*

Government of the Netherlands (2020), *Wet digitale overheid, Digitale overheid, https://www.digitaleoverheid.nl/dossiers/wet-digitale-overheid/ (accessed on 24 January 2020).*

Government of the Netherlands (2018), NL DIGIbeter*. Digital Government Agenda, https://www.nldigitalgovernment.nl/wp-content/uploads/sites/11/2019/02/digital-government-agenda.pdf*

ICTU (2019*),* Monitor Digitale Overheid 2019, Government of the Netherlands, *https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2019/06/30/rapport-monitor-digitale-overheid-2019/rapport-monitor-digitale-overheid-2019.pdf (accessed on 5 March 2020).*

Innopay (2018), *Generiek afsprakenstelsel voor datadeelinitiatieven als basis van de digitale economie. Onderzoek naar het bevorderen van datadelen in het MKB, in opdracht van het Ministerie van Economische Zaken en Klimaat, Rijksoverheid, https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2018/12/30/generiek-afsprakenstelsel-voor-datadeelinitiatieven-als-basis-van-de-digitale-economie/Onderzoek+datadelen+MKB.pdf (accessed on 7 February 2020).*

OECD (2019a), *Government at a Glance 2019*, OECD Publishing, Paris, https://doi.org/10.1787/8ccf5c38-en.

OECD (2019b), *The Path to Becoming a Data-Driven Public Sector*, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/059814a7-en.

3.2 Barcelona - Data Collection Activities

3.2.1 List of consulted stakeholders

3.2.1.1 Interviews

Pau Balcells Alegre, Programme Manager - Oficina Municipal de Dades, Barcelona City Council

Marius Boada Pla - Chief Data Officer, - Oficina Municipal de Dades, Barcelona City Council

3.2.1.2 Presentation

Pau Balcells Alegre, Programme Manager - Oficina Municipal de Dades, Barcelona City Council

Data for Policy in Barcelona, Presentation at the Lisbon Council High-Level Roundtable and Working Lunch on Data for Policy, 31 January 2020, Brussels.

3.2.2 List of consulted documents

Ajuntament de Barcelona (2019a), *2015-2019 Report: Barcelona digital city. Putting technology at the service of people*, https://ajuntament.barcelona.cat/digital/sites/default/files/pla_barcelona_digital_city_in.pdf

Ajuntament de Barcelona (2019b), *Disposicions organitzatives – Decrets de l'Alcaldia. DECRET D'ALCALDIA S1/D/2019-2328 d'11 de novembre, de modificació dela composició del Comitè Executiu de Dades*, Gaseta Municipal, https://w123.bcn.cat/APPS/egaseta/cercaAvancada.do?reqCode=downloadFile&publicacionsId=18811 (accessed on 9 March 2020).

Ajuntament de Barcelona (2018a), *Barcelona City Council Digital Plan. Government measure concerning ethical management and accountable data: Barcelona Data Commons*, https://www.barcelona.cat/digitalstandards/en/data-management/0.1/_attachments/barcelona_data_management_0.1.en.pdf

Ajuntament de Barcelona (2018b), *Disposicions organitzatives – Decrets de l'Alcaldia. DECRET D'ALCALDIA S1/D/2018-02267, de 28 de setembre de 2018, de creació del Comitè Executiu de les Dades i que determina la seva composició*, Gaseta Municipal, https://bcnroc.ajuntament.barcelona.cat/jspui/bitstream/11703/111141/1/GM_S1_D_2018-02267.pdf (accessed on 24 February 2020).

Ajuntament de Barcelona (2018c), *Disposicions organitzatives – Decrets de l'Alcaldia. DECRET D'ALCALDIA S1/D/2018-02269, de 28 de setembre de 2018, de creació de la Taula de Coordinació Transversal de Dades, i que determina la seva composició.*, Gaseta Municipal, https://bcnroc.ajuntament.barcelona.cat/jspui/bitstream/11703/111123/1/GM_S1_D_2018-02269.pdf (accessed on 24 February 2020).

Ajuntament de Barcelona (2018d), *Disposicions organitzatives – Decrets de l'Alcaldia.* DECRET D'ALCALDIA S1/D/2018-02271, de 28 de setembre de 2018, de creació de la Taula de Protecció de Dades., Gaseta Municipal, https://bcnroc.ajuntament.barcelona.cat/jspui/bitstream/11703/111121/1/GM_S1_D_2018-02271.pdf (accessed on 24 February 2020).

Ajuntament de Barcelona (2017), *Barcelona City Council Digital Plan. A government measure for open digitisation: free software and agile development of public administration services*,https://ajuntament.barcelona.cat/digital/sites/default/files/LE_MesuradeGovern_EN_9en.pdf

Ajuntament de Barcelona (2016a), *BARCELONA CIUTAT DIGITAL. A Road Map Towards Technological Sovereignty*, https://bcnroc.ajuntament.barcelona.cat/jspui/bitstream/11703/99399/1/BCN_Digitaleng.pdf (accessed on 20 February 2020).

Ajuntament de Barcelona (2016b), *Mesura de govern: Transició cap a la Sobirania Tecnològica. Pla "Barcelona Ciutat Digital"*, https://bcnroc.ajuntament.barcelona.cat/jspui/bitstream/11703/98713/3/Pla_Ciutat_Digital_MdGovern.pdf (accessed on 12 February 2020).

Balcells Alegre (2020), *Data for Policy in Barcelona*, Presentation at the Lisbon Council High-Level Roundtable, https://lisboncouncil.net/component/downloads/?id=1480 (accessed on 12 February 2020).

Bria, Francesca (2018), Gestio Etica I Responsable De Dades: Barcelona Data Commons, *Slideshare*, https://www.slideshare.net/francescabria/barcelona-municipal-data-office (accessed on 18 February 2020).

DECODE (2019), *Final report on the Barcelona pilots, evaluations of BarcelonaNow and sustainability plans*, https://decodeproject.eu/file/702/download

DECODE (2020), *Common Knowledge: Citizen-led data governance for better cities*, https://decodeproject.eu/file/705/download

Ferrer, **Josep-Ramon (2017),** Barcelona's Smart City vision: an opportunity for transformation, *Field Actions Science Reports* [Online], Special Issue 16 | 2017, http://journals.openedition.org/factsreports/4367 (accessed on 17 February 2020).

Galdon Clavell, Gemma (2017), *¿Soberanía tecnológica? Democracia, datos y gobernanza en la era digital*, CCCBLAB. http://lab.cccb.org/es/soberania-tecnologica-democracia-datos-y-gobernanza-en-la-era-digital/ (accessed on 14 February 2020).

Graham, Thomas (2018) *Barcelona is leading the fightback against smart city surveillance*. Wired https://www.wired.co.uk/article/barcelona-decidim-ada-colau-francesca-bria-decode (accessed on 13 February 2020).

3.3 New Zealand: List of consulted documents

OECD (2019a), Government at a Glance 2019, OECD Publishing, Paris, https://doi.org/10.1787/8ccf5c38-en.

OECD (2019b), The Path to Becoming a Data-Driven Public Sector, OECD Digital Government Studies, OECD Publishing, Paris, https://doi.org/10.1787/059814a7-en.

Data Strategy and Roadmap For New Zealand, 2018

https://www.data.govt.nz/assets/Uploads/data-strategy-and-roadmap-dec-18.pdf

Fact sheet: Operational data governance

https://www.stats.govt.nz/about-us/data-leadership

Stats NZ's annual report Pūrongo ā-tau o Tatauranga Aotearoa for the year ended 30 June 2019

https://www.stats.govt.nz/corporate/stats-nzs-annual-report-purongo-a-tau-o-tatauranga-aotearoa-for-the-year-ended-30-june-2019

Data Leadership overview / empowering agencies to use data more effectively

https://www.data.govt.nz/assets/Uploads/data-leadership-overview-jan-2019-12488.pdf

Strategy for a digital public service

https://www.digital.govt.nz/digital-government/strategy/strategy-summary/strategy-for-a-digital-public-service/

Data Leadership quarterly dashboards

https://www.data.govt.nz/about/government-chief-data-steward-gcds/data-dashboard/

Data and capability framework factsheet

https://www.stats.govt.nz/assets/Uploads/Data-leadership-fact-sheets/Fact-sheet-data-and-statistical-capability-framework-july-2019.pdf

Broadening our role as a national statistical office – New Zealand's journey so far. Note by Statistics New Zealand for the United Nations Economic and Social Council, Conference of European Statisticians, June 2019

https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2019/ECE_CES_2019_28-1906411E.pdf

3.4 Denmark – Data Collection Activities

3.4.1 List of consulted stakeholders

3.4.1.1 Interviews

Annika M. Jacobsen, Head of Department – Data Mining Unit (Den Fælles Dataenhed), Udbetaling Danmark-ATP.

Sanne E. Molin, Head of Department – Comprehensive Control (Helhedsorienteret Kontrol), Udbetaling Danmark-ATP.

3.4.2 List of consulted documents

Arent Eiriksson, Birgitte (2019), *Analyse: Udbetaling Danmarks Systematiske Overvågning*, Justitia og forfatteren, http://justitia-int.org/wp-content/uploads/2019/07/Analyse-Udbetaling-Danmark-systematiske-overva%CC%8Agning.pdf (accessed on 27 March 2020).

Deloitte (2017). *Evaluering af Den Fælles Dataenhed*, https://digst.dk/media/17806/rapport_evaluering_den_faelles_dataenhed_2017.pdf.

European Commission (2018a), eGovernment Benchmark 2018. Securing eGovernment for all FACTSHEETS, Luxembourg, Publications Office of the European Union, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=55490 (accessed on 24 January 2020).

European Commission (2018b), *Fraud and error in the field of EU social security coordination*. *Reference year: 2016,* Luxembourg: Publications Office of the European Union.

The Danish Government and Local Government Denmark (2012), *Good Basic Data for Everyone: A Driver for Growth and Efficiency. The eGovernment*

Strategy 2011-2015, https://en.digst.dk/media/18773/good-basic-data-for-everyone-a-driver-for-growth-and-efficiency.pdf.

The Government, Local Government Denmark and Danish Regions (2016), *A Stronger and More Secure Digital Denmark. Digital Strategy 2016-2020*, https://en.digst.dk/media/14143/ds_singlepage_uk_web.pdf.

OECD (2019a), *Government at a Glance 2019*, OECD Publishing, Paris, https://doi.org/10.1787/8ccf5c38-en.

OECD (2019b), *OECD Social Expenditure database*, http://www.oecd.org/social/soc/OECD2019-Social-Expenditure-Figures-Data.xlsx (accessed on 16 April 2020).

OECD (2019c), *Society at a Glance 2019: OECD Social Indicators*, OECD Publishing, Paris, https://doi.org/10.1787/soc_glance-2019-en.

Udbetaling Danmark (2015), *Bilag 1: Strategi for Den Fælles Dataenhed*, 27 March 2015.

Udbetaling Danmark (2011), *Strategi for helhedsorienteret kontrol. Herunder samarbejde med kommunerne om tilbagebetaligssager*.

3.5 Findata - Data Collection Activities

3.5.1 List of consulted stakeholders

Mrs. Jaana Sinipuro from SITRA

Interviews

Mrs. Jaana Sinipuro from SITRA.

Presentation

- How the legislation for the secondary use of social and health care data and implementation was prepared the Isaacus project.
- Implementation of the national Social and Health Data permit authority Findata. Johanna Seppänen, PhD, Director.
- SECONDARY USE OF HEALTH AND SOCIAL DATA IN FINLAND. 22.11.2019. Joni Komulainen, Ministrial adviser, Master of Laws.
- A FINNISH MODEL FOR THE SECURE AND EFFECTIVE USE OF DATA.
- Finland – Most advanced ecosystem for healthcare innovation. Nora Kaarela, Head of Industry, Health & Wellbeing, Invest in Finland, Business Finland.

3.5.2 List of consulted documents

Secondary use of health and social data; https://stm.fi/en/secondary-use-of-health-and-social-data?p_p_id=56_INSTANCE_7SjjYVdYeJHp&p_p_lifecycle=0&p_p_state=normal&p_p_mode=view&p_p_col_id=column-2&p_p_col_count=3&_56_INSTANCE_%C2%AD7SjjYVdYeJHp_%C2%ADlanguageId=en_US

Act of Secondary Use of Health and Social Data; https://stm.fi/en/secondary-use-of-health-and-social-data

Findata; https://www.findata.fi/en/

Johanna Seppänen, *Secondary use of Finnish Social and Health Data - a new Act and Data Permit Authority*; https://www.himsseuropeconference.eu/sites/himsseuropeconference/files/sponsors/himss-eu-19/presentations/seppanen_secondary_use_of_finnish_social_and_health_data_-_a_new_act_and_data_permit_authority.pdf

Jaana Sinipuro et al, *A finnish model for the secure and effective use of data*, Sitra, 2019; https://www.sitra.fi/en/publications/a-finnish-model-for-the-secure-and-effective-use-of-data/

Joni Komulainen, *Secondary use of health and social data in Finland- How to securely fully utilize the health and social data for research, development and innovation activities, education and knowledge management duties*, 2019;

Hannu Hämäläinen, *How the legislation for the secondary use of social and health care data and implementation was prepared – the Isaacus project*, eHealth Network, 2019.