Examples of implementing data stewardship in countries

This Addendum provides examples of how NSOs in different countries are pursuing their role as data stewards, and the work that European Union and OECD have undertaken in this area. It describes different data governance and data stewardship models and frameworks applied in countries and the role of NSOs in those models. The role of the Chief Data Steward or Chief Data Officer, and its placement in the government structure are exemplified by centralised, distributed and federated hybrid models (see Chapter 4 in the main document).

Table of contents:

Argentina

The Argentinian National Institute of Statistics and Censuses (Instituto Nacional de Estadística y Censos - INDEC) is a decentralized public body that operates within the Ministry of Economy, which exercises the direction of all official statistical activities carried out in the country.

INDEC is responsible for:

- implementing a statistical policy for the Argentine State
- giving structure to and leading the National Statistical System (NSS)
- designing methodologies for statistical production
- organizing and running statistical infrastructure operations
- producing basic indicators and social, economic, demographic and geographic data.

INDEC's role does not currently extend into the Data Steward space. There are several entities within the National Public Administration that regulate public and private information, data protection, access to data and the way it is produced. There is no formally constituted central Data Steward role, with the function of coordinating the national data ecosystem.

The functions that make up the Data Steward role are split across the agencies that have primary responsibility for stewardship of government data. INDEC does work with these agencies in implementing the Open Data plan, but there is no single regulatory framework or agency that could be described as a national or federal Data Government model or framework.

Australia[1]

The Australian Bureau of Statistics (ABS) modernisation is set against a backdrop of an increasing number of Australian Government data initiatives including establishment of an Australian Public Service (APS) Data Professional Stream in 2020, release of an Australian Data Strategy in 2021 and the passing of legislation in 2022 for a national data sharing scheme (the Data Availability and Transparency Act 2022).

In Australia, the NSO contributes to or partners in data stewardship activities, but it is not their main mandate or function. The ABS contributes to and partners in data stewardship activities, but it is not mandated or our main function. The Australian Government's Department of Finance has policy responsibility for data.

There are three foundations of a data stewardship governance role:

- **National data strategies.** The Australian Data Strategy was released in December 2021 and sets the Australian Government's whole-of-government vision for data. The Strategy was delivered jointly by the then Minister for Employment, Workforce, Skills, Small and Family Business and the Minister for Superannuation, Financial Services and the Digital Economy and Women's Economic Security. The Department of Prime Minister and Cabinet (in its role as policy lead for data at the time the Strategy was developed) led its development. The Australian Statistician co-chaired the Working Group which helped shape the strategy and the ABS contributed extensively.
- **Open data movement.** The Australian Prime Minister and Cabinet Open Data webpage identifies multiple aspects to Australia's open data history including the Government's commitment to open data and data-driven innovation in the Australian Government's Public Data Policy Statement released in December 2015. Responsibility for Australia's open data portal, data.gov.au, transferred to the Australian Bureau of Statistics in 2021.
- **Quality assessment and assurance of data.** The *Essential Statistics for Australia* initiative (2010-2014) originated with a 2004 strategy review for the ABS to set out to identify, in a highly consultative way, the key national datasets that are the essential indicators of the state of the

[1] Based on a paper to the CES (June 2022): https://unece.org/sites/default/files/2022-06/ECE_CES_2022_31-2237439E.pdf

nation, regardless of which organisations produce them. The initiative led to the identification of essential statistics (and their data sources and data custodians), an assessment of their quality, and identification of essential statistical infrastructure underpinning the datasets and statistics. The ABS was also instrumental in the establishment of guidelines for Data Integration Projects involving Commonwealth Data for Statistical and Research purposes in which quality assessment and data assurance are fundamental. Many of the guidelines have now been codified in legislation via the Data Availability and Transparency Act (2022).

Three types of collaboration between the NSO and the wider data ecosystem which can help foster data stewardship:

- **Coordination**. The Australian National Data Advisory Council was established in 2019 and has now been codified in law by the Data Availability and Transparency Act (2022. Its role is to advise the (Australian) National Data Commissioner on using and sharing public sector data including on ethical data use, balancing data availability with privacy protection, trust and transparency, technical best practice, industry and international developments and community expectations. The Australian Statistician is a member of the Council.
- **Facilitation of others with the stewardship of their data**. The ABS is leading the whole of Australian Public Service uplift in data capabilities. The Australian Statistician is the Head of the Australian Public Service Data Profession Stream, working closely with the Australian Public Service Commission and the APS Digital Profession Stream to lift the data capabilities of the APS workforce to generate deeper insights to inform decision-making in policy development, programme management and service delivery. The Stream is underpinned by a Data Professional Stream Strategy and is nearing the end of its second year. The ABS has been working across the system to develop the foundations and specific offerings, such as improved graduate recruitment outcomes, to encourage the uplift of both generalist and specialist data skills – for junior through to senior public sector roles.
- **Partnerships with other data providers to develop new products or enhance existing ones.** The ABS accesses over 100 datasets for statistical and research purposes. These datasets are predominantly public sector data assets including birth and death registrations, taxation and welfare data. Data sharing arrangements are an essential part of data partnerships, and the ABS makes uses of a range of data sharing arrangements – via legislation, memorandums of understanding and licencing arrangements. For many data sources, including taxation data, legislation enables the ABS to receive data for the purpose of the Census and Statistics Act 1905 (the Act), the main legislation for collecting, compiling, analysing and disseminating statistical data. The Act governs all ABS statistical releases and specifies confidentiality requirements. In addition to legislation, memoranda of understanding are also put in place specifying terms and conditions, including review points. Licencing arrangements are used for private sector data sources.

Data stewardship can be part of transferring methods and capability available in the NSO to other parts of the National Statistical Service (NSS) and even the data ecosystem as a whole or in reverse, bringing needed methods and capability from outside the NSO, NSS or data ecosystem to within the NSO.

- **Standards and infrastructure.** Standards and infrastructure developed for official statistics have wider applicability and their dissemination and application across all data actors in the Government can be facilitated through a data stewardship approach. ABS' statistical and data integration capability is complemented by a data access service known as the ABS DataLab. The DataLab allows sophisticated analysis of detailed microdata in a secure controlled environment. The DataLab currently services approximately 400 active projects and 4,000 registered users across government but also Australia's research sector, which is a strong support of this service. Use of the ABS DataLab is growing at about 30 per cent per annum.
- **Data science, machine learning and artificial intelligence.** Building technical capability is one of five objectives of the ABS Methodology Strategy and within it, uplifting data science and machine

learning capability for the ABS is a component. The uplift strategy aims to deliver efficiencies via automation of current ABS processes and enhancing the ability to make effective use of new and alternative data sources. Machine learning applications have been used by the ABS to predict dwelling occupancy status on Census night using administrative data, clustering of automatic grouping of similar dwellings through analysis of smart meter data to inform COVID-19 impacts, an "Intelligent Coder" for automated coding of free text responses to standard classifications, and an Automated Image Recognition model for classification of aerial images in the maintenance of the ABS Address Register.

- **More integrated data - a common approach to data handling across the data ecosystem**. The Australian Climate Service is a $AUS210m investment to bring together Australia's expertise and information to help local communities build resilience to climate change. The Service is a partnership between the ABS, the Bureau of Meteorology, the Commonwealth Scientific and Industrial Research Organisation (CSIRO), and Geoscience Australia. The ABS brings critical social and economic information to the Service, enabling a comprehensive picture of the vulnerability of geographical locations to help prevent or prepare for natural hazards, and the resilience of communities to a changing climate. ABS provides detailed information on who is potentially exposed to hazards and how vulnerable they may be. The Service is currently developing a data strategy for the four key "domains" of built, natural, social and economic. ABS is leading work on the social and economic domains with Geoscience Australia. These strategies will inform priorities, identify data gaps, and be used to guide investment decisions.

Four areas where there is potential for NSOs to expand their role in providing better data access and understanding of data in the ecosystem:

- **Geospatial visualization of the information.** In May 2021, the Digital Atlas of Australia was introduced as part of the Government's Digital Economy Strategy to position Australia as a world-leading digital economy and society by 2030. The Digital Atlas will be an interactive, secure and easy-to-use online platform that brings together a range of trusted national datasets for the first time. It will harness the power of near real-time and historic location-based data to put powerful insights at users' fingertips. By using location as the connecting thread between currently disparate datasets, the Digital Atlas will be the next generation of the Australian Government's location-based data infrastructure, based on a modernised Australian National Map. The ABS is a key partner in delivery of the Digital Atlas.
- **Facilitating more use of microdata**. The ABS' data integration capability is complemented by a data access service known as the ABS DataLab. The DataLab is also being trialled as a platform for data sharing across the Australian Public Service, initially with the Department of Finance using the ABS DataLab to enable secure data sharing and sophisticated data analysis. Increased demand for this service is expected with the newly enacted Data Availability and Transparency Act 2022 which seeks to better enable data sharing across the data ecosystem.
- **Data platforms and dashboards.** Dataplace, an initiative of the Australian Office of the National Data Commissioner, is a new, whole-of-government platform to manage data sharing requests for the Australian Government. Dataplace will make it easier to discover and request access to data, including under the Data Availability and Transparency Act 2022. Dataplace is under development and expected to be available from mid-2022.
- **Building data literacy.** The Australian Public Service Data Professions stream, mentioned earlier, is a new approach to building data literacy in Australia. The initiative is developing a data culture within the Australian Public Service workforce.

Future plans:

ABS is tapping into source data from all tiers of Government – Federal, State/Territory and Local. We are evolving how we integrate data with our existing data assets about people, locations, businesses and the physical environment to deliver faster, better and more useful insights. We are increasingly securing access to administrative and transactions datasets from the public and private sectors to produce timely,

frequent, and high-quality statistics; reduce business and household survey response burden; and support policy development.

- **Multi-Agency Data Integration Project.** The Multi-Agency Data Integration Project (MADIP) contains high-value, person-centred and regularly updated datasets that aim to comprehensively cover the Australian population. It is a secure, person-based research data asset that combines a broad range of information on health, education, government payments, income and taxation, employment and population demographics that can be used to answer complex social and economic policy questions. The data is provided by a range of agencies including the Australian Taxation, Education, Employment Services, Social Services, Health, Home Affairs and Services Australia agencies. There are currently over 200 projects across government and academia using MADIP.
- **Business Longitudinal Analysis Data Environment.** The Business Longitudinal Analysis Data Environment (BLADE) integrates ABS business surveys data, business taxation data from the Australian Taxation Office (ATO) and other administrative data sources. BLADE includes data for businesses that have been active in the Australian economy at any time from 2001-02 to the current fiscal year. Over the last few years, the data available in BLADE has been updated at high frequency and new data sets added. We are continually examining new data sources to ensure BLADE remains relevant and contemporary including improving the quality of location information at the most granular level to better support place-based policy and program development and research. BLADE and MADIP can also be linked to enable linked employer-employee analysis.
- **Justice Spine**. ABS is developing a 'Justice Spine' in partnership with the Australian National Indigenous Australians Agency. It is a longitudinal national data asset linking police recorded criminal offenders in Australia's criminal courts with adult prisoners in the corrective services systems. The dataset will show how persons move and interact within and across the justice system nationally, something that is currently not possible. The dataset will have potential to be linked to other Commonwealth and State/Territory held datasets for deeper analysis of criminal offenders. It will be available to approved policy makers and researchers in late 2023 and will provide critical insights into patterns of offending and recidivism.
- **National Disability Data Asset**. A National Disability Data Asset (NDDA) is under development and will deliver a significant new enduring national asset comprising a collection of linked, de-identified data from across multiple Commonwealth, State and Territory service systems on people with disability and their pathways through services. The NDDA will be enabled by the development of a new data integration infrastructure, known as the Australian National Data Integration Infrastructure (ANDII). The ABS, Australian Institute of Health and Welfare and Department of Social Services, working in partnership, are leading the development of ANDII and NDDA in collaboration with the Commonwealth, states and territories and the disability sector.
- **Australian Immunisation Register**. This data integration project linked the Australian Immunisation Register (AIR) dataset and deaths data to MADIP to generate insights to inform the Australian COVID-19 Vaccine and Treatment Strategy. This project enabled timely analysis of vaccine uptake and outcomes across socio-demographic cohorts and geographic areas.
- **Tracking the labour market**. The Labour Market Tracker Project integrated job-related data, including ATO's Single Touch Payroll data to MADIP and BLADE to enable close to real time monitoring of the labour market and the Australian economy during the COVID-19 pandemic.
- **Private Sector data**. Over the past 12 months, the ABS has focussed on securing long term access to private sector transactions data for statistical purposes. Our activities include extending and enhancing existing data agreements with banks and establishing new agreements with major retailers as well as exploring a range of data relationships in other areas. This work has enabled the:
    - release of the monthly household spending indicator using bank transactions data. This new indicator covers 68 per cent of household spending compared to the long-standing Retail Trade Survey with coverage of about 30 per cent, and
    - development of more frequent, less costly, and more detailed statistics on household income and expenditure using bank, large retailer, loyalty scheme and credit agency data.

A project is underway that builds on work undertaken by the ABS during the COVID-19 pandemic when aggregated, de-identified data from Australian banks helped measure the economic effects of COVID-19 in Australia.

We are leading a consortium of a dozen Commonwealth Government agencies in a collaboration with a major Australian bank to pilot a new analytics environment containing the bank's de-identified transactions data to be used to answer public policy questions.

Belgium

Like most NSOs worldwide, also Statistics Belgium has been impacted by the rapid transformation of official statistics. Statistics Belgium embraces the principle of data stewardship to anticipate the challenges accompanying this transformation and grasp the opportunities it brings. It sees it as one of its key adaptations to prepare for the future. Statistics Belgium's vision of data stewardship is still evolving and progresses to a data stewardship that highly emphasizes a "connecting" role.

Figure 1 below visualizes this connecting role. Connections need to be established over two dimensions. Firstly, a horizontal dimension between data holders and users. This dimension primarily refers to collaboration in the statistical system, but its essence is the data that flow from the (external) data holders to users (over Statistics Belgium). Secondly, vertical integration is needed between the technical dimension of the data production and data governance elements. This dimension is more internal to Statistics Belgium, but it may or should refer to the Belgian statistical system as a whole.

A more concrete approach is shown in Figure 2 where stewardship initiates and drives the streams between three interconnecting work domains. The central domain, the ecosystem, triggers the need for innovations in terms of new data products, methodologies, and the adoption of new data collection technologies. The results of these innovations are fed back directly to the ecosystem through experimental data and statistics, but eventually indirectly via consolidation in a statistical organization's governance system and regular statistical production. Initializing, managing, and strengthening these workflows is a more dynamic aspect of stewardship alongside the structural elements (ecosystem, innovation, and governance).

Both figures hint at a comprehensive interpretation of data stewardship, implying an organizational assignment rather than personal responsibility; Statistics Belgium as data steward, not an individual member of staff trying to encompass the whole field. Nevertheless, an individual steward was assigned recently, partially to enhance the adoption of the principle but to reinforce the dynamic elements of data stewardship as described above.

Figure 1
Data System/Data Steward Relationship

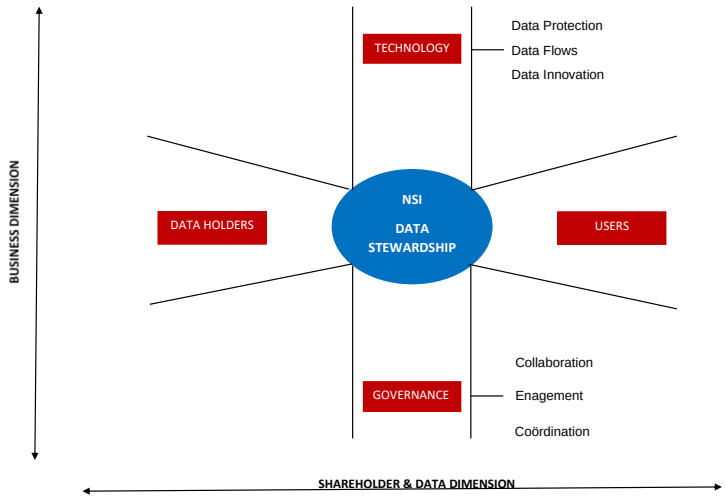
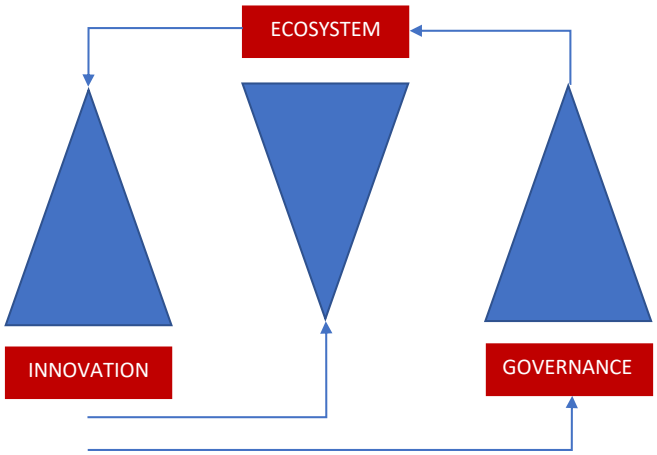


Figure 2
Structural and Dynamic Elements Data Stewardship



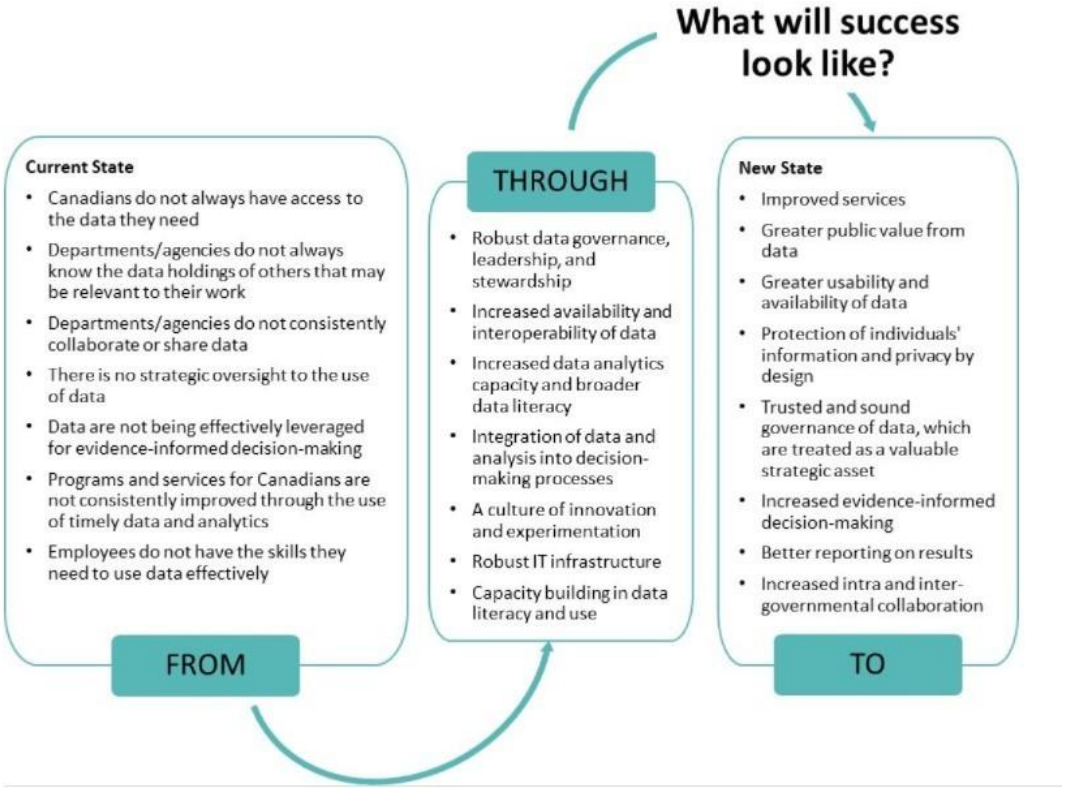
Canada

The Government of Canada (GoC) has been prioritising innovation and increasing horizontality to facilitate a whole-of-government approach to the strategic use of data for public good. One such mechanism has been the creation and evergreening of the Data Strategy Roadmap for the Federal Public Service.

In 2018, the Data Strategy Roadmap for the Federal Public Service was published, a collaborative response by the Privy Council's Office (PCO), The Treasury Board (TBS) and Statistics Canada (StatCan) to a call from the Clerk of the Privy Council to develop a data strategy. While not a national data strategy concerned with all public and private data stores, this federal data strategy underpins the strategic use of data across the GoC, enables the transition to a digital government, and ensures that the entire public service can best leverage data and insights for evidence-based decision making and better outcomes for citizens. See the figure below for details about the current- and end-states that the Roadmap seeks to address and develop.

The Roadmap is an evergreen document currently being revised and updated to improve alignment between all levels of government, to address recent digital developments, and to account for and describe the recent creation of a Chief Data Officer (CDO) of Canada role. The CDO will be responsible for providing GoC-wide leadership for the management of data and information; developing the administrative frameworks and standards for governance and quality; enabling federal organization to leverage data; and oversight for information management and data. The evergreening of the Roadmap, the creation of the CDO and various intergovernmental work occurring around data governance and stewardship all work together to further develop Canada's digital, data, and information capacity.

Figure 3
Current state and desired end-states



Ecuador

Among the limitations that the INEC of Ecuador has had to access data (from external administrative sources), is the obsolescence of the Statistics Law, which dates from 1976 and consequently, by not contemplating access to administrative records, restricts to this NSO to be able to access more information. With this, the INEC manages data from its own statistical operations and protects the databases of those sources that have agreed to provide their information for statistical purposes.

However, it does not exercise a role of data stewardship at the national level, since this falls mainly on the National Directorate of Public Registries (DINARP), which has a solid and up-to-date legal framework that allows it to consolidate the information of public institutions for its interoperability, but not integrated or treated for statistical use. This has determined that the INEC generate internal legal and technical regulations to protect the information, and resolve national legal limitations, through inter-institutional cooperation, internal regulations, support to source entities to strengthen their information for statistical purposes, and allow the access to statistical products with sensitive information, through closed processing and analysis environments, such as the Information Processing Laboratory, where analysts from international organizations and academia have been able to carry out studies of various kinds.
Thus, the INEC has given value to the information, and has allowed access to it for use.

Estonia

Introduction

Estonia has been one of leaders in digital transformation and building of information society since 1990-ies. Since then, over 300 registers and information systems for administrative processes were implemented. Transformation from paper based public services towards electronic was achieved a decade ago and more than 800 million digital signatures have been issued over the last 20 years in a country with population of 1.4 million.

At the end of 1990-ies rapid development of e-government started and many cornerstones were put in place. To name just some, ID-card and digital signature were introduced, technical infrastructure X-road for data exchange was implemented to allow webservices, paper records were transferred to digital. In private sector banking went digital and access to internet was made common and affordable if not free to access.

Importance of quality and digital data have always been considered important, but information systems and electronic services have been the primary concern at the strategic level. More than 1000 data services were developed between registers that allow transactions and activities between information systems. Still in every IT (Information Technology) strategy there are concerns about the number of new information systems / registers and the volume of data requested for e-service processes from people and businesses.

Model and its development

At the beginning, the Estonian model for governing data was not data orientated but rather IT focussed.

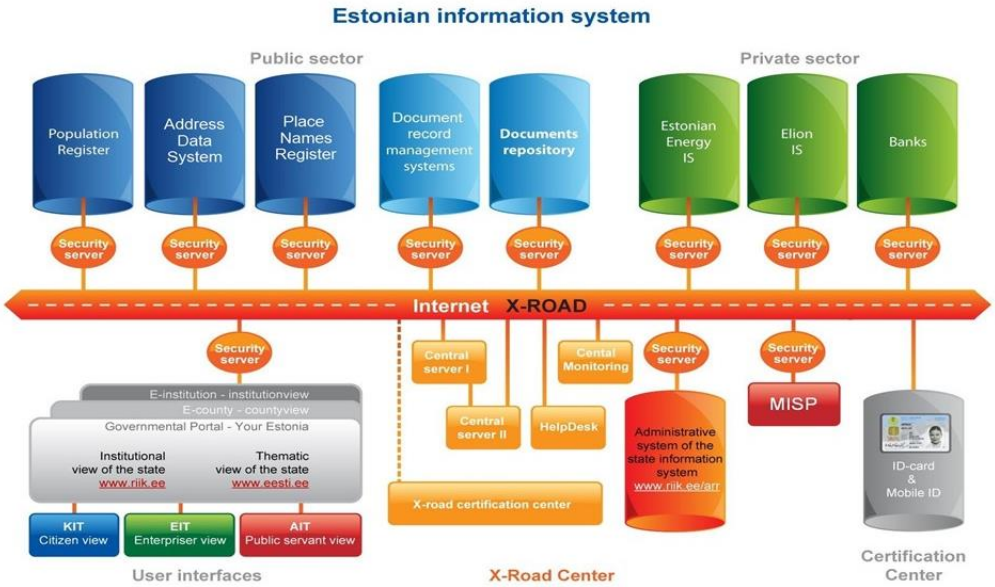
Developments in data governance at the country level started in the middle of 1990ies. State Chancellery was responsible until 2001 when the Ministry of Economic Affairs and Communication took over. Two legal acts: Database act (1997) and Public information Act (2001) addressed organisation's documented information or information systems and not yet data as such.

The next major change started in 2007, when concept of State Information System was introduced, and development of the State Information Systems Administrative System (RIHA) started. This was unique at the time that all state registers, business process or service providing information systems of agencies, as well of local municipalities were catalogued. This was a very promising solution supported by a responsible agency. The State Information System Agency was also responsible for the development of X-road and kept a catalogue of X-road services together with its technical infrastructure. Quite soon there were 1000s of services, connecting most public sector agencies. Social and legal services were most used.

The State Information System is a legal concept and an information system with applications and services. As a legal concept it states that all official registers and information systems providing services to people, businesses or other agencies need to be legally regulated. It gives accountability of what data public authorities have and what they do with it. In 2020 there were 817 such databases in central government and 558 in local authorities. There is an obligation for registers to reuse data other agency's databases already have, for example master data on addresses is taken form the addresses information system. Another obligation is to make use of official and valid classifications.

As an information system RIHA is a catalogue of registers that acts like a guide helping to get an overview of the whole state information system. Without RIHA, it would be very difficult to see what kinds of data institutions collect and reuse it to reduce duplication.

Figure 4
The original X-road infrastructure (2002)



Major change came in 2007 that merged data and information. On the one hand regulation of databases was integrated into the concept of public information, which was widened to cover also open data. On the other hand, an attempt for better data governance was taken by establishing an administration system for the state information system. Statistics Estonia took responsibility for the correct use of classifications in information systems and the State Information System Authority started to implement the once-only principle.

There was a step-by-step change in organisational model of IT-governance. Special IT-agencies were established under the ministries, providing services for all agencies in the area of a ministerial governance. Currently there are six such IT-centres. The competence of IT-centres has been built gradually and now includes competence of business requirements and data.

Third major change in data governance model was started in 2018 when Statistics Estonia added data governance to its strategy and the Ministry of Economic Affairs and Communication appointed a Chief Data Officer.

At the moment, the state information system model is under revision, including its architecture, catalogue of registers and X-road services and their functions. New approaches from open data movement to big data (other types of data than structured) are in development. RIHA has been in use for more than 15 years and now there is a plan to integrate RIHA with the Estonian Open Data Portal to create a single access point to both types of data.

Role of the statistical community

Before 2018 the role of Statistics Estonia in data governance manifested itself in two tiers. First, in 2008 the coordination of the classification system was trusted to Statistics Estonia. Use of classifications in state registers was monitored. This played a substantial role in harmonization of use of classifications and helped to carry out register-based censuses (2000, 2011, 2021). However wider harmonization of the code lists used in registers gave not so good results.

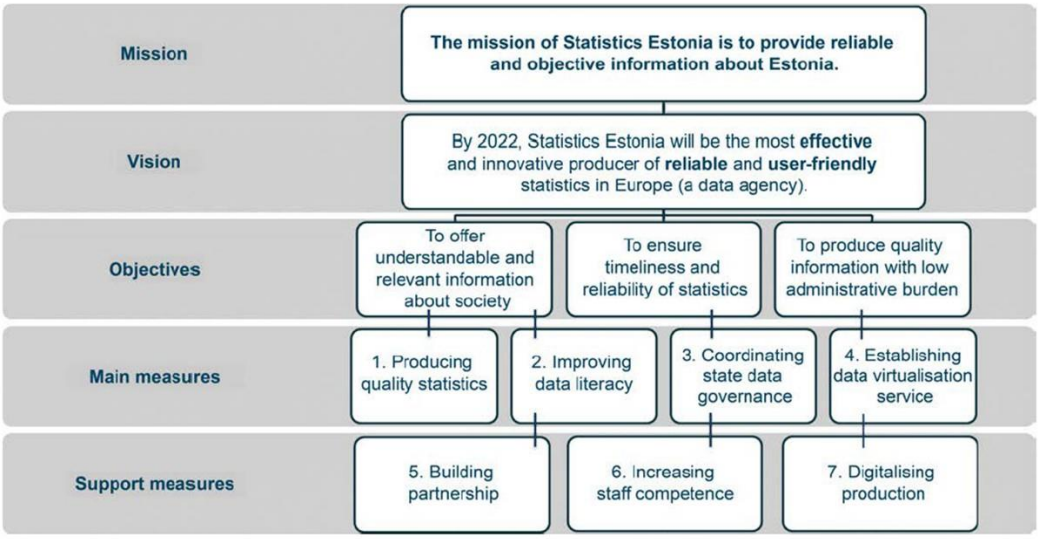
The second tier has been the use of administrative data in the production of statistics. The database register at first and then RIHA from 2008 on have been the source for finding potential new sources for statistics. Statistics Estonia has developed a data gate to capture data safely via X-road from administrative sources and has integrated it to its applications and data architecture. Statistics Estonia used administrative data sources in 67% of its statistical activities in 2020. All together data from more than 100 sources were used.

Use of administrative data brought up a question of data description and quality. These are being solved on the case-by-case basis. In this process technical and business understanding of data by IT-centre or service providing agency slows data delivery.

Statistics Estonia has a strong legal mandate to use data from administrative sources for official statistics. Question of secondary use of data collected for official statistics as well as widening of data capture form administrative sources not explicitly needed for particular purpose of official statistics has been asked.

In 2018 Statistics Estonia published its new strategy with the vision of becoming a data agency. It also asked for a wider mandate in the growing data ecosystem since at that time there was a need for leadership in data stewardship in the public sector.

Figure 5
Statistics Estonia's strategy 2018-2022



Amendments to the Official Statistics Act were made in 2019. Two new legal concepts were introduced: data governance and data sharing services. Firstly, co-ordination of data governance in public sector became the role of the NSO. Secondly, possibility of secondary use of administrative data collected for the purposes of official statistics as well as use of registers for wider analytical needs was permitted.

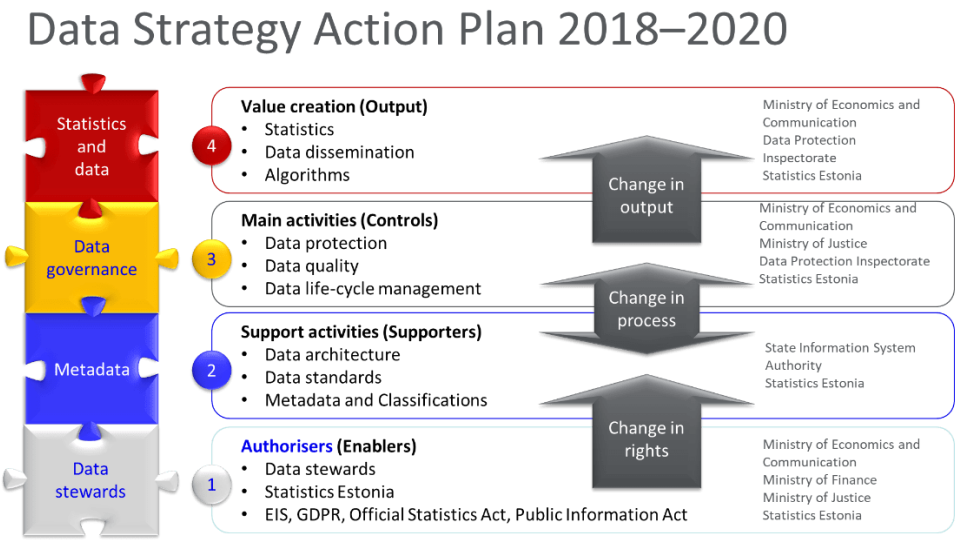
Co-ordination of developments in data governance stared 2019. Focus areas were data description, data quality and framework model to implement data governance, train data stewards and measure its effectiveness. As a first step, a better overview of registers and administrative databases was needed, especially data warehouses or data lakes that have emerged in agencies and serve as a source of analytical outcomes.

The main result of this overview was that there are many data sources in agencies not well known outside the agency that hold important data; dealing with obsolete databases and data needs improvement. Senior management in agencies was enthusiastic about data governance but did not know where and how to start. In majority of registers some data stewardship was part of the daily work but not always called stewardship.

The first tasks of the NSO were to standardise data descriptions and data quality. Statistics Estonia created guidelines for both and at the same time developed software tools for implementing in agencies. The aim of the tool is to allow agencies to create their own data catalogue and exchange metadata as well as open data between their information systems and state data catalogue or open data portal. Main focus was use of data dictionaries and business vocabularies for better semantic understanding of data.

Together with the Ministry of Economic Affairs and Communication and Information System Authority (RIA) an action plan was drawn up, which made co-ordinated developments on guidelines, tools, training and policies possible.

Figure 6
Data governance action plan 2018-2020



The Corona-19 pandemic slowed implementation of data governance in agencies, particularly in 2021. The development of data description standard and co-ordinated development of IT tools continued to allow both semantically enriched data description and automatic exchange of data and metadata between them.

Statistics Estonia took its own data governance into focus. For better use of collected and administrative data the data virtualisation application was implemented to facilitate use of R. Statistics Estonia has integrated metadata and data preparation system since 2012 and its metadata part is in the process of renewal. Colectica has been chosen as the new metadata system that addresses data stewardship and elaboration on data sets and other types of information object ownership.

Estonia's Digital Agenda 2030 and Statistics Estonia's role in digital government

Estonia prepared in 2021 new development program for information society. This a wider document approved by government and parliament explaining and setting objectives how information and communication technology, i.e. digital solutions help to achieve the objectives of 'Estonia 2035'. This "Estonia's Digital Agenda 2030"[2] includes a vision and an action plan concerning the development of the Estonian economy, state and society with the help of digital technology in the next decade.

The agenda has three sub-objectives for society:

- The Estonian economy is innovative and knowledge-based, using new technologies and business models as well as flexible forms of work.
- The needs of all people are considered when shaping the living environment and the foundations of high-quality spaces and principles of inclusive design are consistently followed when making decisions to ensure the accessibility and convenience of spiritual, physical and digital space for everyone.
- Estonia is an innovative country which values the creation and use of knowledge and where social life is organised by means of new human-centric and efficient technologies.

To implement the vision, more specific goals have been set in development plan. The development of digital government, i.e. the use of digital solutions in the public sector is substantial part on the plan, since no other development plan includes the general development of digital government, and the public sector also leads and sets the direction for the development of the Estonian digital society.

[2] https://www.mkm.ee/media/6970/download

In Estonia public sector has been in lead in digitalisation. The vision states: Estonia, empowered by digitalisation. Next leaps in the development of digital government are:

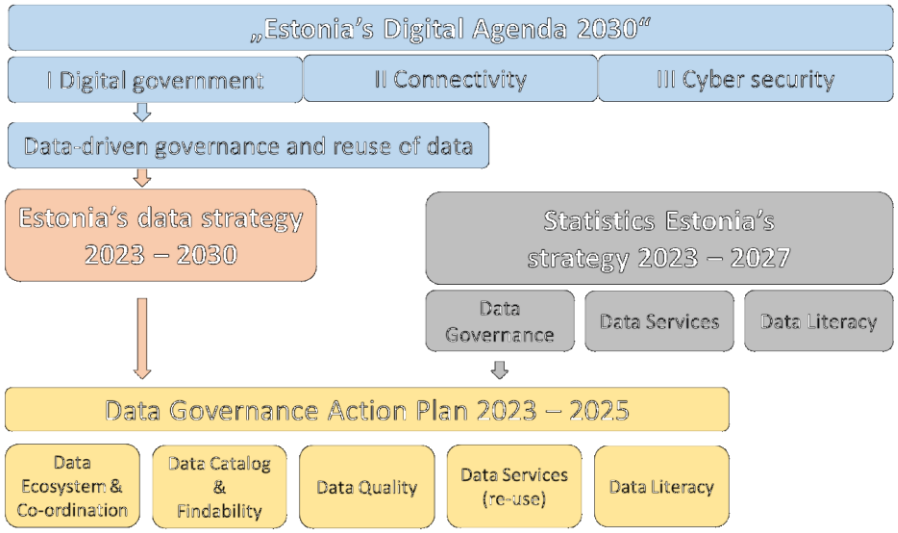
- Switch to life and business event based and proactive services
- AI-powered government
- Human-centric digital government
- Green digital government

To achieve these aims two pillars or enablers has been set up to ensure the sustainability of digital government:

- Introduction of the management and user-centricity of public services
- Data-driven governance and reuse of data

Data-driven governance and reuse of data are direct connection between "Estonia's Digital Agenda 2030" and strategy of Statistics Estonia.

Figure 7
State and Statistics Estonia's strategies



The Ministry of Economic Affairs and Communication started Estonia's first data strategy. It is wider than data governance, covering all types of data, their use for administrative purposes as well data privacy matters. Data governance of registers, information systems in public sector and their repositories is an important pillar. It also addresses privately held data sources and academic data repositories.

Strategy of Statistics Estonia for 2023-2027 has three pillars: data governance, data services, and data literacy. All of them are also addressed in Estonia's data strategy. Co-ordinating bodies of data governance agreed upon Data Governance Action Plan for 2023-2025. This action plan will integrate general governance view, improving data stewardship and addressing data description, data quality and data services issues, and providing and developing tools and information systems to support activities in agencies and country level view of data use.

Finland

Finland's data governance model fits within the federated hybrid model. In the 'Government report on information policy and artificial intelligence'[3] the Government of Finland looked at information policy not only from the viewpoint of information management, but also from the perspectives of the conditions

[3] Report on ethical information policy in an age of artificial intelligence (5 December 2018, PDF 1.7MB)

placed on the use of information, the value of the information involved, ethical principles and financial impacts. The report:

- sets out the new information policy
- proposes extensive use of AI (Artificial Intelligence)
- identifies the levels of AI competence required
- explores possibilities of economies in platforms and data

The measures included in the report cover the collection and merging of information, information disclosure and storage, as well as information security and data protection. The measures examined in the report also include ethical issues, how to secure expertise, regulatory issues and policy-level participation in the EU and international forums.

The report constitutes the knowledge base and policy framework on which a roadmap with prioritised actions can be built in the future. Information policy is connected to all other policy areas because any issue will be investigated, and decisions made based on comprehensive information. This new policy area lays the groundwork for the development of targets for information management in the public and private sectors.

Following on from this policy, a new 'Public Administration Information Management Act' (906/2019) covering data governance, data security and data management came into effect in 2020. This Act contains provisions on the organisation and description of data management, interoperability of databases and information systems, implementation of technical interfaces and implementation of data security for the public administration sector. This Act established the Information Management Board under the auspices of the Ministry of Finance.

The function of the Information Management Board is to promote the implementation of information management and data security procedures laid down in the Public Administration Information Management Act and to ensure that the requirements of the Act are met. The Information Management Board is not a general authority for information management; its tasks are limited to the scope specified in the Public Administration Information Management Act.

The Information Management Board has prepared various general guidelines and recommendations to support data governance and data management in public organisations:

- the handling of classified documents (in English[4])
- data security and risk management
- the metadata of case processing
- technical interfaces and viewing connections
- the documentation of the data governance model used by the organisation
- the responsibilities of directors in data governance and data management

Statistics Finland has a place on the board representing the National Statistical System. The Statistics Finland representative provides expertise on data governance and data management in statistical work. In the statistical community there are already several good examples of how to improve metadata, process descriptions, total quality, common classifications, standards etc. that may be useful to share with other public organisations.

The Board also includes representatives from ministries, government institutions and municipalities. Their expertise covers areas such as basic registers, data security, ICT (Information and Communication Technologies) expertise on interoperability of databases and case management.

All public organisations are expected to follow these recommendations, but it is important to understand that the recommendations themselves are not obligatory. Instead, they are best practice examples of how

[4] https://julkaisut.valtioneuvosto.fi/handle/10024/162846

to implement the requirements in the Act, and every organisation must implement these recommendations in a way that best suits their own operating environment and practices.

The Board also organises seminars and workshops on the implementation of their recommendations. The participation rate so far has been very good (around 300-500 participants per workshop). The purpose of these seminars and workshops is to share knowledge on best practice, give examples how different organisations have applied the recommendations and to support participants in understanding what is expected and what the minimum requirements for good information management are.

The other responsibility of the Board is to monitor the implementation of the Act in public organisations. In the first year, the Board collected information on the descriptions of the responsibilities of directors in Information Management and how data security and data management training is conducted in the organisations. In the following year, the survey focussed on documentation of the data governance model and how public organisations are implementing it.

France

The National Institute of Statistics and Economic Studies (Insee) Coordinates the activities of the Official Statistical Service (SSP). The SSP is composed along with Insee, of the Ministerial Statistical Offices (MSOs) which carry out statistical operations in their areas of expertise. A Ministerial Statistical Office Charter sets out the principles of coordination by Insee of the works produced by the various MSOs, as well as their duties in the six categories of missions: production, study, dissemination, role towards supervisory administrative bodies, harmonisation of classifications and concepts and contribution to international statistics. The coordination of MSOs takes place at political /strategic and operational levels

Over the past years, France has set up an advanced legal framework to get access to administrative documents and reuse public information, in particular data and source codes. Thus, the code of relations between the public and administrations (CRPA) has been modified by the French digital Bill. More recently, the French government has strengthened the national data policy. A report of the French parliamentary mission “For a public data policy” was published in 2020.5 The report addresses the need for even more open data while striking the right balance with protection, the issue of data quality and accessibility, the need for sharing data between government departments, and the need to use private sector data on a large scale (data in the general interest). It calls for an adaptation of the recruitment of digital talents policy and for an instillation of an open data and open-source culture in public services.

Since 2021, a government circular[6] defines the data policy as a strategic priority for the state in its interactions with all its partners. The different administrations must constantly seek the best circulation of data, algorithms and source codes. All ministers are requested to provide their strategic roadmap. As Chief Officer for data, algorithms and source codes, the head of the DINUM relies on a network of appointed ministerial data officers to coordinate the action of the different ministries, with the support of its Etalab Department.

Insee collaborates regularly with the Interministerial directorate for digital technology (DINUM), which oversees the state digital transformation for the benefit of citizens and state agents. Digital transformation is a government top priority and raises the important issue of the need for digital competencies. In February 2021, Insee and the Dinum were asked to lead a joint project to define a typology of needs in the different ministries, to clarify the possible role of the official statistical service, and to evaluate the recruitment processes. Etalab department of Dinum coordinates the design and implementation of the state strategy in the field of data.

[5] https://www.mission-open-data.fr/uploads/decidim/attachment/file/36/Mission_Bothorel_Rapport.pdf
[6] https://www.legifrance.gouv.fr/circulaire/id/45162

Germany

Destatis has the ambition to act as a data steward to promote quality standards of official statistics in the German data ecosystem. In addition to being a data producer, Destatis has broadened its role to being a data competence center, a data service provider and data manager. In its role as data service provider Destatis has established the data platform Dashboard Germany. In its data manager and competence center roles there are also considerations of creating a data hub holding the metadata of all publicly held and selected privately held data.

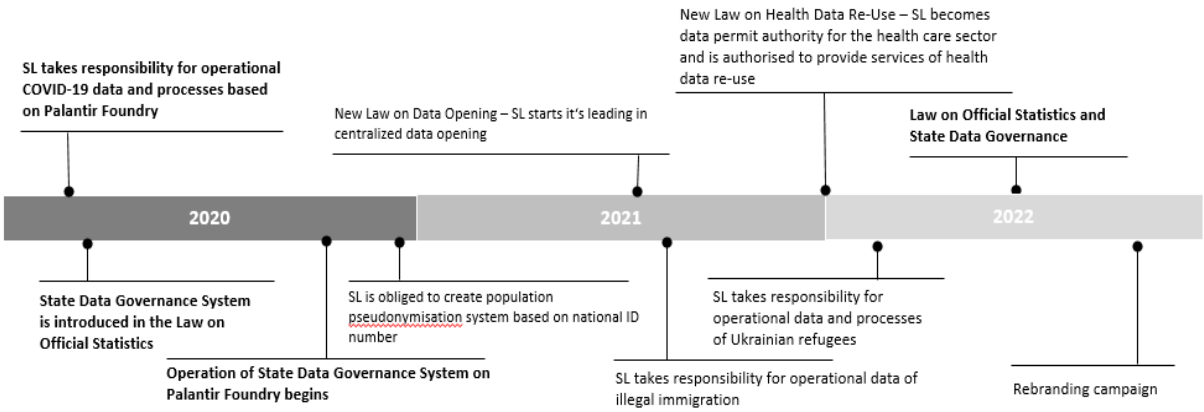
The data strategy of Destatis outlines the strategic orientation and further development of all areas of the Federal Statistical Office. In this context, the data stewardship approach is understood as the operationalisation of the data strategy (further development of the Federal Statistical Office containing the above outlined roles).

Lithuania

So far public sector data in Lithuania is fully decentralized. This results in a great diversity of administrative sources in technological and functional terms. The high fragmentation of data sources and its negative impact on being able to use the data became apparent with the onset of the COVID pandemic. In the case of Lithuania, collecting and processing operational data was a key problem in pandemic management. And it was the COVID pandemic that inspired decisive and rapid decisions in the field of state data management. Statistics Lithuania (SL) became a central actor, demonstrating its ambition to take on a fundamentally new role as a data steward. Statistics Lithuania has taken the lead to make administrative data more valuable and available for different purposes, exchange and re-use. Therefore, we have started to implement our role as a state data steward addressing the evident problems: fragmented, scattered state data, inflexible and slow official statistics, and ineffective initiatives of data exchange, reuse and opening.

The figure below presents a short timeline of strategical and tactical steps of SL in taking the role of national data steward. COVID pandemic was crucial here, and decision taken in right time accelerated all the following changes – from building the State Data Governance System on the powerful platform to the rebranding campaign due to the new name of institution.

Figure 8
Timeline



During the COVID pandemic, SL got a chance to prove its competence and strength – the responsibility for pandemic data management together with full support from the Government. In a few months, changes in the Law on Official Statistics were made, licence of data operational system Palantir Foundry was purchased, and the State Data Governance System was launched. Subsequently, within a month, SL started to produce operational daily based COVID pandemic data.

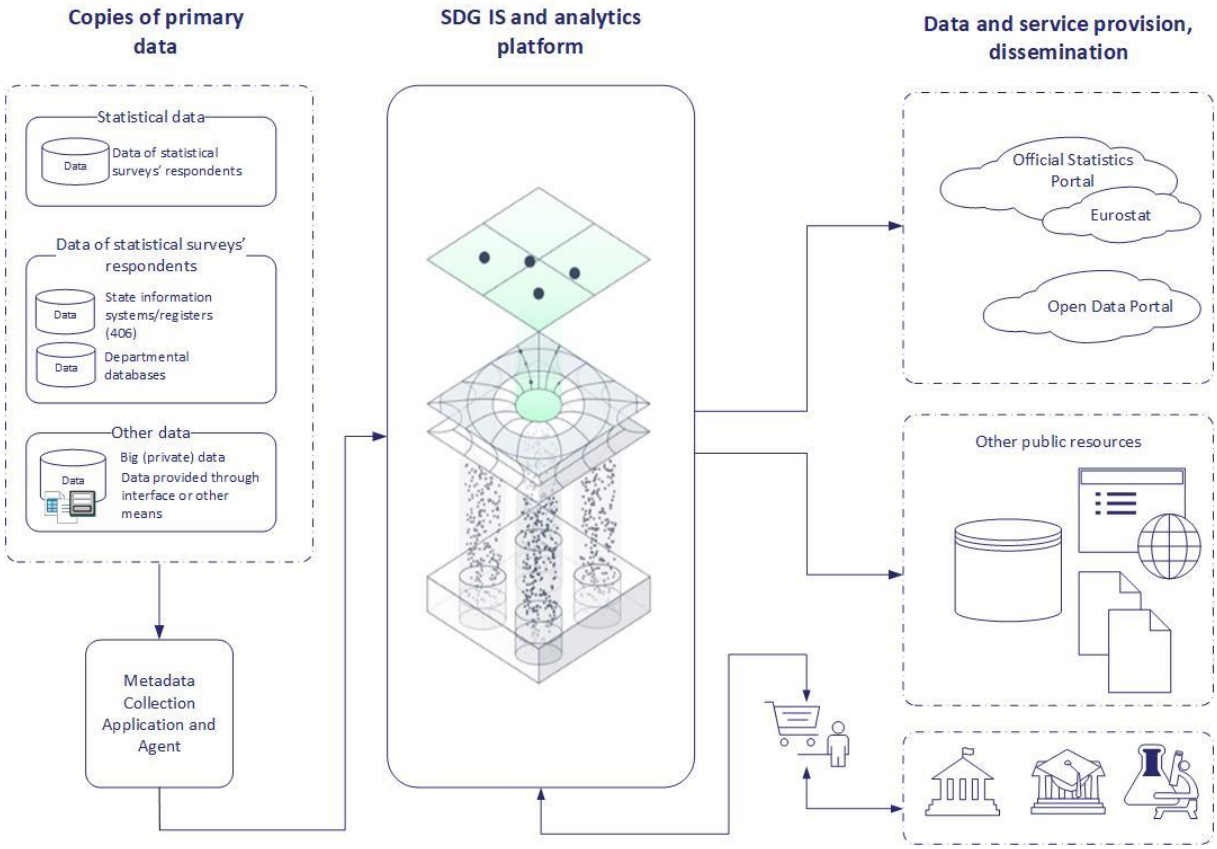
Moreover, SL took responsibility of managing the data on illegal migration crisis last summer as well as the data on Ukrainian war refugees. It gave SL a huge political support and enhanced reputation of SL among public authorities and decision-makers. This allowed to move step by step towards the goal of a unified state data architecture based on our concept.

In short, during the two years of the state of emergency, SL put a strong emphasis on the cutting-edge technology, revised and adapted legislation, and highly qualified staff. And at the current moment, SL does have the technology – multifunction data operating platform for the new ecosystem, does have adopted the main Law coming into force 2023, and does have the team of programmers and analysts able to build processes in the new data ecosystem. The staff and the new organizational structure are the last step in this transition, together with rebranding campaign.

Concept

The simple diagram (see figure below) reflects the concept of the state data ecosystem built by SL – from source to product. It is a visualized idea of the State Data Governance System (SDG IS). The big data operating platform (Palantir Foundry) implements an entire data ecosystem – from data collection to data production also covering infrastructure.

Figure 9
The state data ecosystem



The objectives of the State Data Governance Information System (SDG IS) can be summarized as follows: standardised and fast data collection, data consolidation, information production, data exchange and re-use, and analytical spaces for users. It is necessary to note that we aim to consolidate the primary data. This will allow the centralized production of statistics of new generation – much more operational and detailed. This ecosystem is already up and running from November 2020. All operational data on COVID-19 and migration crisis have been already built in this ecosystem, while integration of main administrative data sources and transferring production of official statistics are in progress.

After creating the technological base for the state data ecosystem, SL realised that becoming a state data steward is not just a technological challenge. This is primarily a legal issue. It was obvious that in addition to technology it was also necessary to expand legal role of SL to make clear regulation of this new activity.

Legislation – dualistic approach

According to the new Law on Official Statistics and State Data Governance, SL takes two distinct roles in the state: the producer of official statistics, as before, and the steward of state data. For this reason, the Law establishes a new name of the institution: the State Data Agency, which participates in the formation of state policy 1) in the field of management of official statistics; and 2) in the field of state data governance.

Table 1
Comparison of the scope of SL activity and its regulation

• State policy of one area • EU regulation • Lithuanian Department of Statistics • One program • One council • Coordination of official statistics • Statistical data • Statistical services • Fragmented activity	• State policy in two areas • EU and LT regulation • State Data Agency • Two programs • Two councils • Coordination of official statistics and state data • Statistical and state data • Statistical and analytical services • Centralized activity

The areas of official statistics and state data governance are strictly separated. This means that each area has its own regulation. The field of state data governance is organized on a "mirror" basis, i.e. all processes of data management should be treated the same, following the fundamental principles, which assure values of European democracy.

State data include data from administrative sources, while data provided by respondents do not fall within this definition. This means that state data can be used for official statistics purposes, as has always been the case, but there is no reverse link – there is no extended use of respondents' data (as it is prohibited by EU regulations). We restrict statistical data defining it just as respondents' data, while definition of state data includes all administrative data, but excludes statistical data. In this way we open administrative data collected to official statistics for other purposes according to the Law.

The state data governance shall be implemented by the Agency through the State Data Governance Programme, which provides analogous information as in the case of the Official Statistics Programme. Issues and problems related to state data governance shall be resolved by the State Data Governance Council. State data are subject to the same confidentiality obligations as official statistics, except the purposes provided in the Law.

The purposes of using state data are expanded. After the official statistics, it can be used for monitoring of country development, for decision-making, research and innovations, legitimate needs of state and municipal institutions, crisis data management, education, governmental projects, and finally data opening and re-use.

The services of the State Data Agency are expanded as well: internal state data for decision makers, centralised opening of public sector data, environment for health data re-use, environment for data analysis for public institutions. In addition, the Agency shall provide to other state institutions the resources of data platform to build their own data sandboxes.

Conclusion

Currently, SL has submitted draft implementing legislation to the Government for approval. The draft State Data Governance Programme will soon be submitted to state authorities for approval. The expectations and needs of the public sector institutions and decision-makers already now suggest that

the scope of the work will be at least equal to that of the Official Statistics Programme, and in the long run, due to greater legal flexibility and faster results, it may even exceed it.

Table 2
Comparison of official statistics and state data governance regulation

Role: **Producer of official statistics**	*Role*: **Steward of state data**
• *Fundamental principles:* professional independence, impartiality and objectivity, quality, confidentiality and protection of statistical data, adequacy of resources and cost effectiveness	• *Fundamental principles:* professional independence, impartiality and objectivity, quality, confidentiality, integrity and availability, lawfulness
• *Planning*: Official Statistics Programme	• *Planning*: State Data Governance Programme
• *Advisory body*: Statistical Council	• *Advisory body*: State Data Governance Council
Data use purposes	
• Production of official statistics • Confidential data for scientific purposes	• Official statistics • Monitoring and analysis of state economic and social development • Decision-making in public administration • Research and development, innovations • State and municipal institutions which are authorized to perform public services • Prevention and management of epidemic and other critical situations • Educational purposes of higher education institutions • Research and projects carried out by international organizations according to obligations of Lithuania • Opening and re-use of state data

Mexico

The statistical and geographic information production system in Mexico is regulated in the Constitution and by legislation. The National Institute of Statistics and Geography (INEGI) is an autonomous body, completely independent from the federal government.

As specified in legislation, INEGI plays a dual role in the national information system:

- as a direct producer of official information and
- as the coordinator that sets norms and standards to produce official statistics by different government agencies.

The Mexican system can be classified within the Federal Hybrid model: it has direct control of the production cycle of census and survey information to produce statistical information, as well as of the statistics derived from the use of administrative records and other sources. INEGI also issues norms and standards to be followed by government agencies that generate information considered to be of national interest.

The Institute also has the authority to establish agreements with companies and non-governmental organizations to obtain data that can be reused in the generation of statistical information.

INEGI is the institution that, by law, assumes the role of custodian of all data and information used to produce statistics and establishes the responsibility for their careful handling, adhering to the principles of quality, relevance, truthfulness, opportunity, confidentiality and independence.

Each institution that provides information or generates statistics of national interest has a coordinator who is responsible for implementing the policies and standards determined by INEGI.

As can be seen, the role of Chief Data Steward does not formally exist in Mexico; however, it is INEGI that carries out the activities that this role implies within the country.

Those responsible for coordinating information within the different government agencies have the role of linking the data, but not of custody or direct administration of the data. By this definition they could not be considered to be true Data Stewards.

For this reason, INEGI is currently reviewing its standards and policies to strengthen the data ecosystem through the figure of Data Stewards across the entire Mexican system.

INEGI has adapted the Generic Statistical Business Process Model (GSBPM) to improve information management throughout the entire cycle, in addition it has adapted its norms and standards to strengthen data management and the generation of information products and corresponding metadata.

In addition, INEGI has developed a new information governance and architecture strategy that facilitates standardization, transparency, confidentiality and the quality of information. It has also implemented systems with improved security schemes to better manage information and preserve its integrity.

One of the main challenges anticipated in the near future will be to incorporate information from companies and organizations to regularly produce statistical information. To get access this large pool of information it will be necessary to improve the regulatory framework to provide these companies and organizations assurances that their information can be processed by INEGI without affecting any other legal obligations they may have.

A clearly defined data governance framework and agreement on the role and responsibilities of Data Stewards will be key to moving the Mexican information system to a new level, involving more intensive use of administrative records and alternative sources of information.

New Zealand

In New Zealand, all government agencies are autonomous, with Chief Executives appointed by a Public Service Commissioner. While being constitutionally separate from the Executive Branch of Government, they support decision-making of individual ministers.

This is a highly decentralised model of government administration, which has several benefits but also some limitations in terms of systemic approaches to data management. In recognition of this in 2017, the New Zealand Public Services Commissioner created the role of Government Chief Data Steward (GCDS), recognising that strong system focussed leadership was required to help New Zealand realise data as an asset. The Public Services Commissioner appointed the Chief Executive of Stats NZ to that role. In this leadership role, the GCDS recognises and champions the benefits of agencies coming together and leveraging their data assets to deliver value for New Zealanders.

In 2022 the role of the GCDS was strengthened with the authority to direct agencies to collect specific data to fill data gaps, in compliance with the New Zealand Privacy Act 2020, and to set system-wide tools to better foster the trusted and ethical use of data.

In conjunction with the role of the GCDS, Stats NZ serves as the lead agency for data within the New Zealand government data system, by facilitating and enabling a joined-up approach to data-related opportunities and challenges. In this capacity, Stats NZ, in conjunction with the GCDS, supports agencies to maximise the potential of their data and ensure it is used effectively, while maintaining the trust and confidence of New Zealanders.

To do this, the GCDS develops data policy and principals, and has published a Government Data Strategy and Roadmap[7] to provide clear guidance on how the data environment should operate, while ensuring a data-driven future for New Zealand. This is accomplished by:

- driving the response to new and emerging data-issues
- guiding best practice and safe innovation

[7] https://www.data.govt.nz/docs/data-strategy-and-roadmap-for-new-zealand-2021/

- setting common data standards to ensure the consistent collection of data
- allowing for integration, comparability and production of meaningful insights and partnering with agencies and (indigenous) Māori to build capability
- helping agencies build the skills, processes, tools and services for the successful collection, management, use and dissemination of data.

New Zealand's central government data governance currently more closely reflects the distributed model, as individual government departments are operationally autonomous. There are two key governance bodies supporting the role of the GCDS, the Information Group and the Digital Government Leadership Group.

The Information Group

The Information Group is the key governance body for the Government Data Strategy and Roadmap - responsible for setting, implementing, monitoring, and reviewing the progress made against the Strategy. In its governance of system-level participation by government agencies, the Information Group supports the GCDS, through delivery of things like implementation planning, progress reporting, identification of risks and facilitating engagements and communications. The group is chaired by the Deputy CE, Data System Leadership, at Stats NZ.

The Digital Government Leadership Group

The Digital Government Leadership Group (DGLG) is co-chaired by the Government Chief Data Steward and the Government Chief Digital Officer. The role of the DGLG is to support the Government Chief Digital Officer and the GCDS to develop and improve the digital and data system across the public service and ensure the public service is aligned with the *Strategy for a Digital Public Service* and the *Government Data Strategy and Roadmap*, and other relevant strategies.

These governance bodies operate under the premise that treating data as a valuable asset does not mean simply gathering more data or applying it indiscriminately. It means properly stewarding and managing, and being intentional about, use and the generation of value from data.

Through its government data leadership role, Stats NZ administers the ongoing development of the government's *Data Investment Plan* which guides and informs strategic system investment in data, including the way that critical data is managed within baseline, and how new investments are considered. The *Data Investment Plan* is a cross-agency initiative and is being developed in collaboration with government agencies through a phased approach.

The Plan has been endorsed by New Zealand's Cabinet, is governed by the Information Group, and will be monitored through a planned annual *Health of the Data System* report.

Māori data governance

The Treaty of Waitangi (New Zealand's founding document) promised that the Crown would uphold the authority that indigenous Māori tribes had over their lands and taonga (prized and valued objects, resources, ideas and techniques). Data about Māori has been identified as a taonga under the terms of the Treaty, and as such the Crown has an obligation to work with Māori Treaty partners when considering governance of data about Māori.

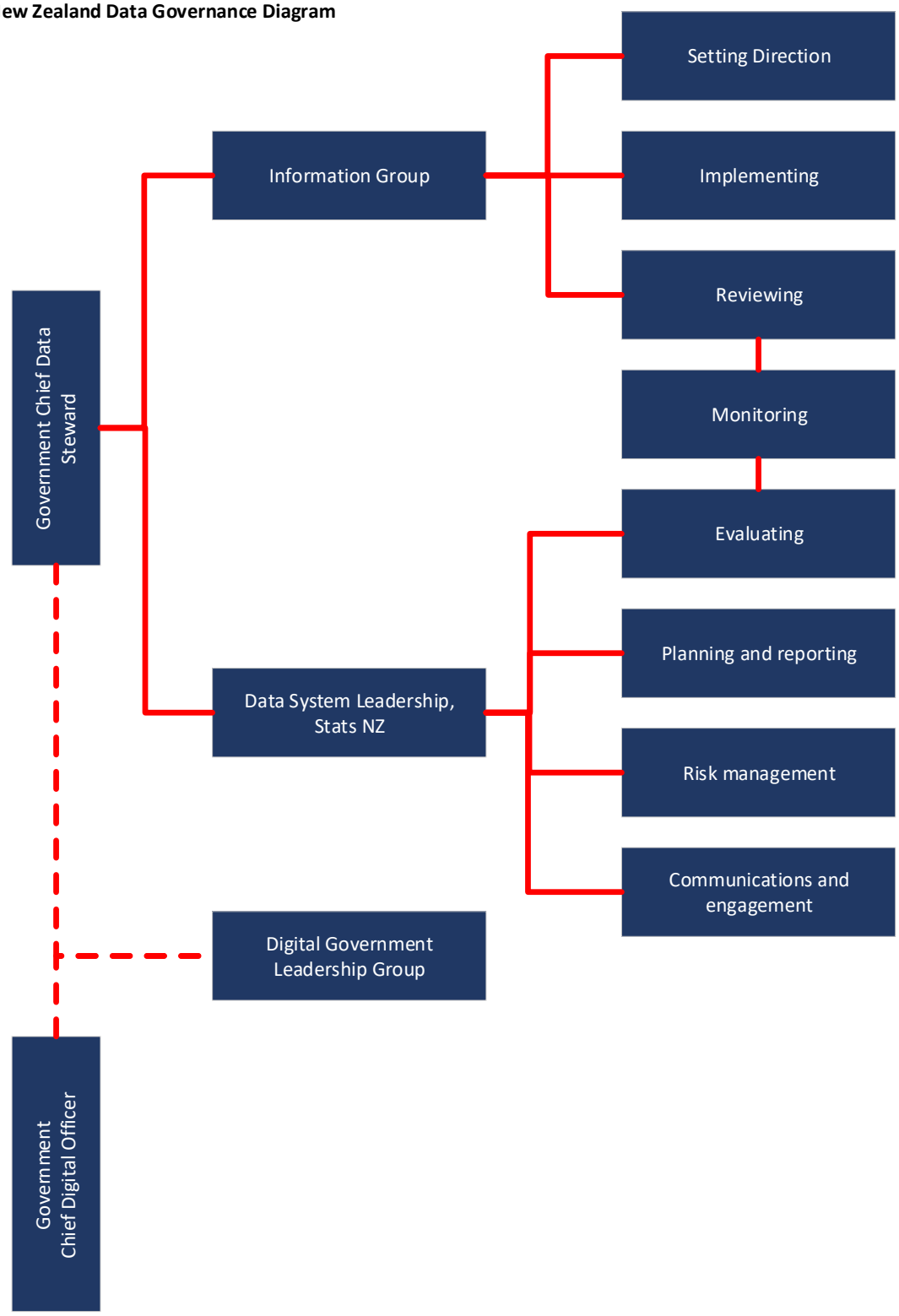
Governance of the New Zealand data system was not designed in partnership with Māori. And the government does not currently consistently reflect a *te ao Māori* (Māori world view) lens across the wider official data system that supports both Māori and government aspirations for data. This has resulted in challenges including a lack of trust and confidence by Māori in the official data system, inadequate meaningful Māori participation in that system, including at governance levels, and missed opportunities for Māori to add value to the official data system through te ao Māori insights and innovations.

The Government Chief Data Steward has recognised the opportunity for government to work in collaboration with Māori partners, co-designing a Māori Data Governance model for the official

government data system. This would employ a Treaty of Waitangi-based co-design process that appropriately reflects the obligations inherent to a Crown and Māori partnership.

Figure 10

New Zealand Data Governance Diagram



The Māori Data Governance design work sits under the auspices of a formal relationship agreement, Mana Ōrite, forged between the Data Iwi Leaders Group (DILG) representing Māori interests, and Stats NZ. The work to date has produced two reports:

1) Tawhiti Nuku, Māori Data Governance Co-design Report on the outcomes of the co-design process and its recommended next steps
2) Māori Data Governance Co-design Review, which focuses on the process of co-design.

Work is progressing through an Ohu raraunga (data working group), including government and Te Ao Māori membership, to continue the momentum on refining, testing, and implementing the Māori Data Governance model.

New South Wales (Australia)

New South Wales (NSW) is one of the 6 states and 10 federal territories that make up the Federation of Australia. Each state is a self-governing political entity with incomplete sovereignty (having ceded some sovereign rights to federation) and has its own constitution, legislatures, departments, and certain civil authorities (e.g. judiciary and law enforcement) that administer and deliver most public policies and programmes.

Data governance policy in NSW is currently under review which will include new definitions of data roles and responsibilities. This summary describes the current state which is one of transition. Data Custodianship Policy in NSW is administered by the Department of Customer Service. This policy took the approach of describing functions rather than labelling roles due to the variation in the use of labels for these roles in different state agencies, and as a result uses the term 'custodian' and does not use the term 'steward'. Custodianship is defined as formally assigning rights and responsibilities for data and information assets, including capture and management on behalf of the NSW Government. Despite this, from the documentation available it would appear that their model of data governance best fits the distributed model, due to the absence of a role equivalent to a central data steward.

In NSW the custodianship role and its associated responsibilities belong to the government agency which acts on behalf of the State of NSW i.e. each government agency acts as the custodian of the data and information assets and products held in their care. The Department of Customer Service has a co-ordinating role but can only provide guidance to the autonomous agencies. The current Data Reform program has established an NSW Data Leadership Group (NDLG). This group consists of Chief Data Officers from each government Cluster. Cluster CDOs have a functional role (i.e. they perform this role in addition to their substantive role). At present, NSW does not have an NSW Government Chief Data Officer but are working toward a hybrid governance model through the NDLG.

In their guidance role the DCS published a Data Governance Toolkit built on a common understanding of the benefits, obligations and best practice. The aim of this was to ensure a consistent approach to data governance across NSW Government agencies and provide practical and consistent guidance on the key components of an effective data governance program, as well as to create a shared understanding of what good data governance looks like. Compliance with the Toolkit is not mandatory, but following the guidance in the Toolkit is intended to:

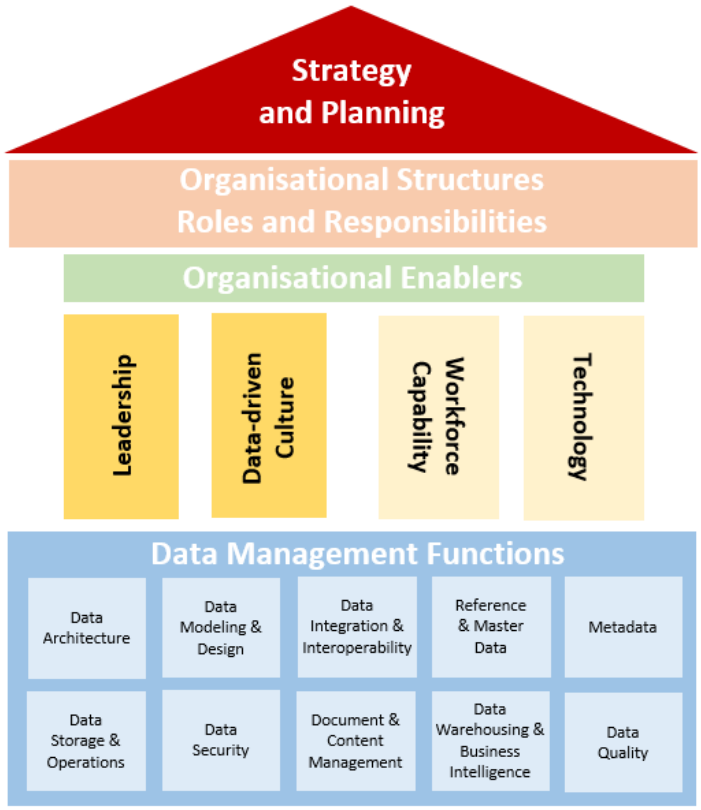
- support agencies to maximise the value of data while reducing data-related risk;
- assist agencies in meeting their legislative and regulatory obligations;
- ensure data is managed in line with national and international standards;
- facilitate better interoperability between agencies; and
- build data governance maturity at both the departmental and all-of-government levels.

<u>The Data Governance Model</u>

The Model defines four interconnected tiers of data governance activities, each of which is critical to effective data governance in agencies. The four tiers are:

1) **Strategy and planning** – agencies clearly define the data governance program's values, vision and mission and compose a business-aligned strategy for governing and managing data as an organisational asset.
2) **Organisational structures & roles & responsibilities** – agencies ensure accountability and decision-making authority for data-related activities to be appropriately assigned and formalised at all levels of the organisation.
3) **Organisational Enablers** – agencies ensure the organisational environment is an enabler of good data governance. This means ensuring there is a strong motivation (or 'will') to achieve good data governance by having sustained buy-in and investment from senior leadership, as well as fostering a strong organisational data culture. It also means ensuring the organisation has the requisite capability (or 'skill') to achieve good data governance, both in terms of workforce capabilities, as well as appropriate tools and technologies.
4) **Data Management** – agencies ensure their data governance program has oversight of core data management functions (e.g. data quality, storage, security, business insights etc.).

Figure 11
Data Governance Model[8]



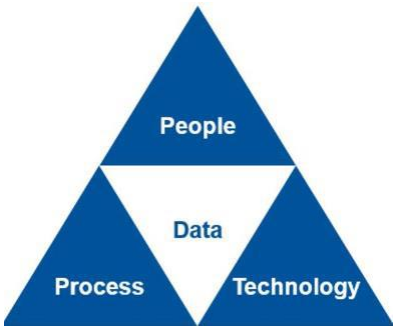
Interpreting the Model

Each component of the Model includes a high-level summary of what the component is, why it is important, what good practice looks like (i.e. the goals), how to achieve good practice and, where appropriate, provides references to useful resources and relevant standards. The level of detail for each component has been kept to a high level and the practical elements of the framework will gradually be expanded with input from agencies once the model is in use.

[8] https://data.nsw.gov.au/data-governance-toolkit-0

The Model also aligns with Gartner's (2017) 'Golden Triangle' of 'People, Process and Technology' with Data at the centre, which recognises that effective data governance is an ongoing effort executed by people, enabled by repeatable processes, and supported by technology. Each component of the Model therefore encompasses a mix of accountabilities relating to people, processes and technologies to support the implementation of that component.

Figure 12
Golden Triangle - People Process and Technology



Source: Gartner 2017

Norway

Statistics Norway coordinates a system of official statistics with 10 other producers of such statistics (however, we count for 85% ourselves). This coordination comprises

- A national programme for official statistics and quality control of all such statistics.
- Statistical confidentiality, sharing of pseudonymized microdata under specific conditions.
- Collaborates with more than 30 owners of administrative registers and use more than 100 such registers for statistics production. Has developed quality reports for all these registers in collaboration with the owners.
- Collects all data from the municipalities to central authorities on behalf of these, using the same data for statistics on public services.
- Collaborates actively with many other government bodies, also the national data steward (The Norwegian Digitalisation Agency) on developing digitized public services, also to the benefit of official statistics.
- Continuous development of methodology and technology, new data platform for statistical production.
- Large data base (StatBank Norway), API (Application Programming Interfaces) service with open data, free use of data, classifications and code lists.
- Standardisation.
- Intention of extended collaboration, statistical literacy.

Slovakia

Within the national statistical system, the Statistical Office of the Slovak Republic (hereinafter referred to as the Office) performs the role of data governance body. Members of the national statistical system consult the Office on the methodology of national statistical surveys, use of national statistical classifications and codebooks and on methodology of data collection, which the Office will use as an administrative source of data for statistical purposes.

Switzerland

The Swiss Confederation is a federal republic composed of 26 cantons.

The Swiss implementation of data management corresponds to the federal hybrid model. The data governance model defined by the central federal administration is only binding for the administration itself. The cantons are free to design their own structures. However, there are institutions and roles in the federal administration that promote nationally applied standards. The Swiss government is working towards establishing the once-only principle (see Appendix 3) while keeping data storage and ownership decentralised. This requires interoperability to achieve sharing of data across administrative units to allow the reuse of data.

In a first step, a nationwide metadata catalogue was developed that contains standardised descriptions of the data as well as information on the origin, use, legal restrictions and quality of the shared data (www.i14y.admin.ch). In addition, further instruments to promote interoperability such as directories of public authority services and electronic interfaces (API) are here.

The next step is to standardise and harmonise the administration's data collections. The role of the Swiss Data Steward coordinates and supports this work. For fundamental topics that affect all data collections, there are topic-specific data stewards. Examples of such cross-sectional topics are statistics or geodata.

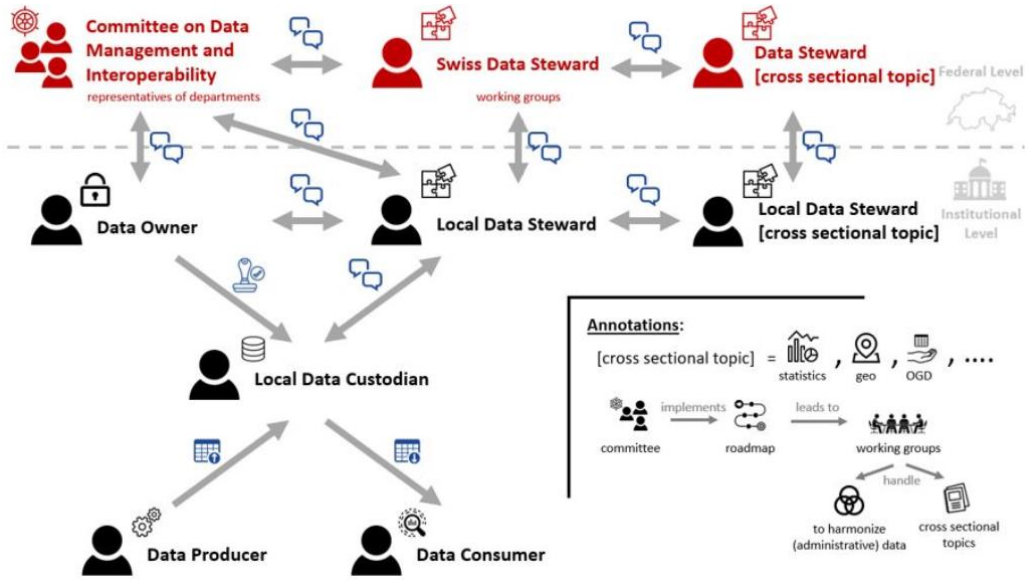
National standards are developed and defined in thematic working groups. All participating agencies from the Confederation, cantons and municipalities, but also from the private sector, are invited to participate. The review and updating of the standards will become a permanent task.

The coordinating and mediating role of the Swiss Data Steward is of vital importance in the Swiss data governance model:

- He is centrally located in the Swiss National Statistical Institute (NSI) also known as the Swiss Federal Statistical Office (FSO) and is responsible for:
  - Coordinating the standardization and harmonization process.
  - Identifying and describing the data requirements of the various users.
  - Managing the content of metadata (data catalogue).
  - Validating the quality assurance of metadata and data in the administrative areas using data analyses.
-

The following diagram shows the data governance model and roles at the federal level:

Figure 13
Data governance model and roles at the federal level



United States

Through its decentralized statistical system, the United States has been the careful steward of vital data for well over two centuries. While some of the language may differ, the concepts and values identified in the Data Stewardship report are consistent with the principles governing U.S. statistical activities. More recently, the United States government has expanded the reach of data stewardship beyond just the statistical agencies, as many other Federal agencies amass large collections of data through oversight, enforcement, regulation, and other activities. The expansion can be seen in the United States' Federal Data Strategy, with a vision to accelerate the use of data to deliver on mission, serve the public, and steward resources while protecting security, privacy, and confidentiality.

The importance of data stewardship in the United States was further codified by the enactment of the Foundations for Evidence-based Policymaking Act of 2018. This new law establishes three key actors across the Federal government: Chief Evaluation Officer, Chief Data Officer, and Statistical Official. These officials coordinate data activities not just among statistical agencies but across the wide range of Federal agencies. The Data Stewardship report speaks of a hybrid model of stewardship. Such a model accurately represents the many actors in the United States system: statistical agencies are housed within, and work with, broader departments characterized by subject matter; the statistical agencies also work together in activities coordinated by the Chief Statistician of the United States; and all Federal activities must align with the Federal Data Strategy. Given these various levels, data stewards in various parts of the Federal government follow coordinated principles to ensure proper data governance.

Staff of United States statistical agencies are responsible for the proper stewardship of statistical data, often reporting through a central steward. Among the many activities that fall under this stewardship, two deserve special notice. First, data are often collected voluntarily from businesses and households under a pledge of confidentiality. Such data collection represents a trust relationship that is vital; any violation of the pledge of confidentiality will jeopardize future data collection. Second, statistical agencies make certain confidential microdata available to qualified researchers under secure conditions. Again, this represents a trust that cannot be violated. Agency data stewards ensure that all such research meets strict standards, and that results do not violate confidentiality.

European Union: The new data governance legislation of the European Union

The European Union (EU) has undertaken several actions to improve data sharing and data governance in the EU Member States.

The following paragraphs describe the 2020 data strategy and the Data Governance Act. This combination of strategy and legislation aims to put in place a federal hybrid model where the European Union has a legislated centralised structure and framework for data governance, and the EU Member States make their own decisions on how to implement these regulations, and even go beyond in the context of their national data governance. Under this structure public organisations have the autonomy to decide how to manage their own datasets within the regulations.

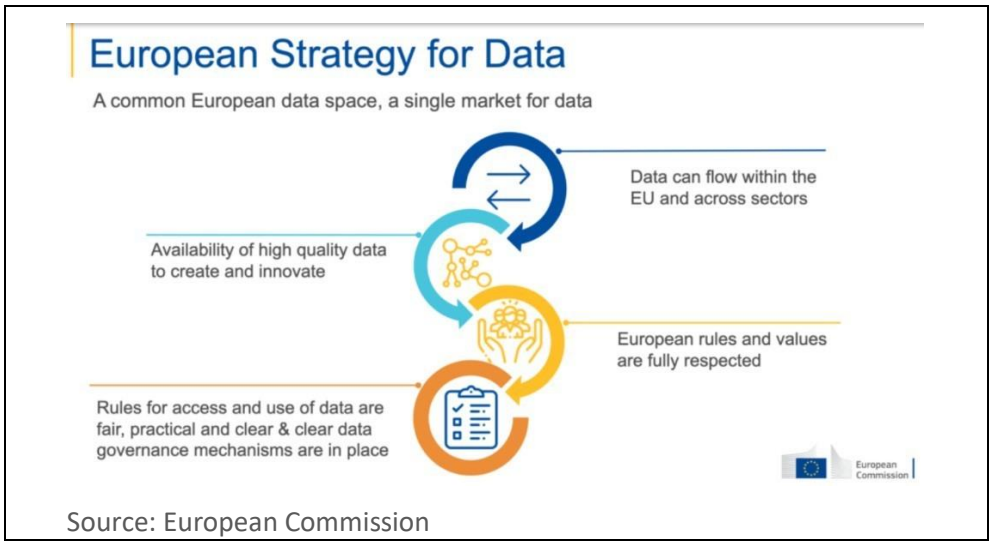
The 'European Strategy for Data'[9] was published in 2020. The objective of this strategy is to make sure that the EU becomes a leader and role model for a society empowered by data.

The EU will create common European data spaces to ensure that more data become available for use in the EU. The aim is to create a single market for data, to unlock unused data and support the flow of data freely within both the European Union and across sectors for the benefit of businesses, researchers and public administrations. Figure 14 describes the aims of the strategy. Figure 15 shows how the common European data spaces will work.

[9] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN

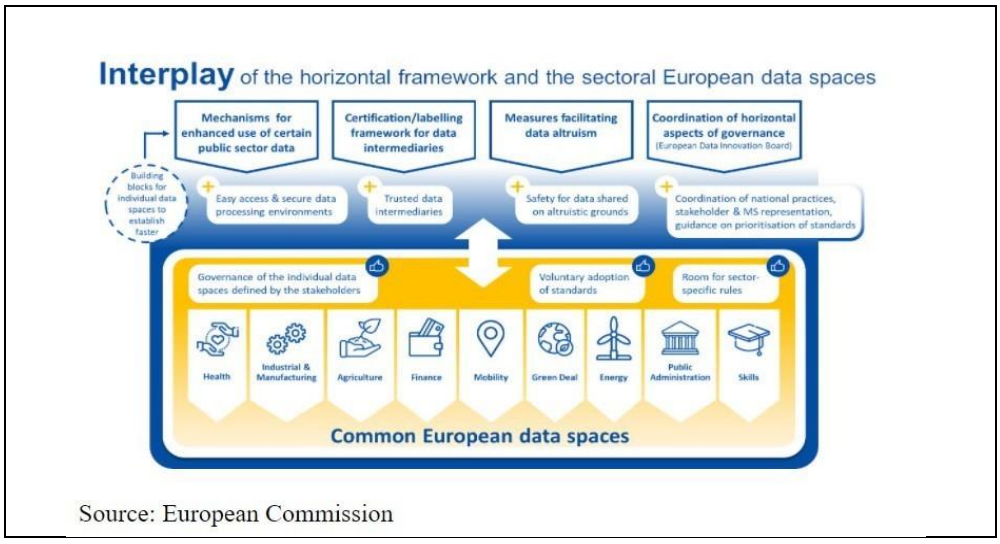
To fulfil the aims of the strategy, the EU is based on existing legal frameworks and building new ones. In 2022, the European Union adopted the Regulation (EU) 2022/868 on European data governance (DGA) which is one of the key pillars of the 2020 European Strategy for Data.

Figure 14
Aims of the European Strategy for Data



Source: European Commission

Figure 15
Common European data spaces



Source: European Commission

The regulation seeks to increase trust in data sharing, strengthen mechanisms to increase data availability and overcome technical obstacles to the reuse of data. As mentioned above, the Data Governance Act will also support the set-up and development of common European data spaces in strategic domains, involving both private and public players, in sectors such as health, environment, energy, agriculture, mobility, finance, manufacturing, public administration and skills (see Figure 15).

The Data Governance Act entered into force on 23 June 2022 and, following a 15-month grace period, will be applicable from September 2023.

The EU will boost the development of trustworthy data-sharing systems through 4 broad sets of measures:

i. Mechanisms to facilitate the reuse of certain public sector data that cannot be made available as open data. For example, the reuse of health data could advance research to find cures for rare or chronic diseases.
ii. Measures to ensure that data intermediaries will function as trustworthy organisers of data sharing or pooling within the common European data spaces.
iii. Measures to make it easier for citizens and businesses to make their data available for the benefit of society (data altruism).
iv. Measures to facilitate data sharing, in particular to make it possible for data to be used across sectors and borders, and to enable the right data to be found for the right purpose.

With respect to the general EU legal framework supporting the EU Strategy for Data, there is as well the Directive (EU) 2019/1024 on open data and the re-use of public sector information as well as the Data Act that is a key measure for making more data available for use in line with EU rules and values aiming at making an important contribution to the digital transformation objective of the Digital Decade. [10]

OECD work on data governance in the public sector

The OECD has developed a model for data governance in the public sector as means to highlight the core elements countries can take into consideration when designing and deploying data projects and initiatives. As presented in the 2019 OECD Report *The Path to becoming a data-driven public sector,*[11] the framework aims to bring greater clarity and structure to the definition and implementation of the concept of data governance at the national level across OECD member and partner countries.

The model is based on the extensive OECD work on digital government and government data and additional research carried-out by the OECD Secretariat. Earlier versions of the model can be found in previous OECD digital government reviews, namely the 201 7 OECD Digital Government Review of Norway, the 2019 OECD Digital Government Review of Sweden, the 2019 OECD Digital Government Review of Peru, and the 2019 OECD Digital Government Review of Argentina.

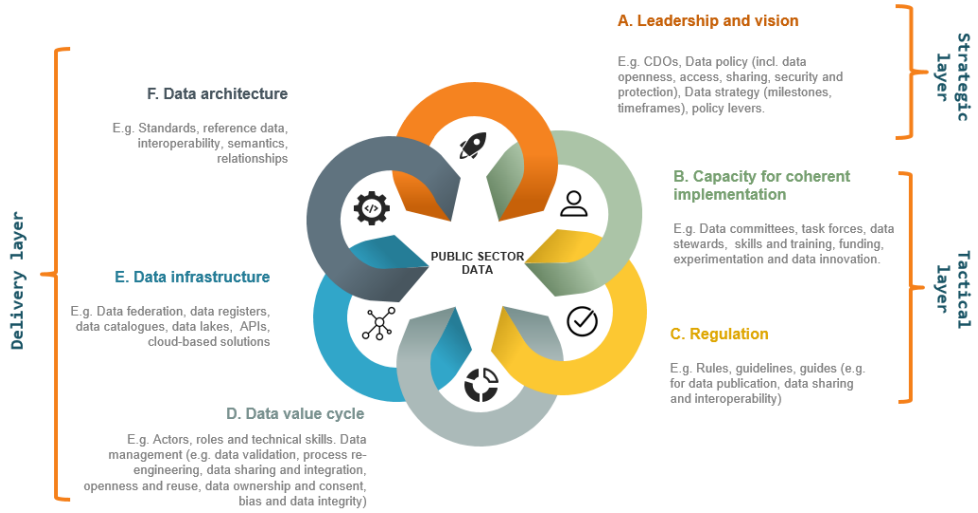
As described in the 2019 Report, we cite, the model intends to highlight the equal and strategic relevance and the value of all organisational, policy and technical aspects for the success of data governance. It identifies a range of non-exclusive data governance elements and tools, and organises them in six groups (a – f). These six groups are then arranged under three core layers of data governance (Strategic, Tactical and Delivery) using the three traditional data governance categories as guidance (Strategic, Tactical, Operational) as discussed and/or presented in Ghavami, 2015; DAMA, 2017; and the BARC's 9-Feld-Matrix [see Grosser (2013) and BARC (2019)]. The model is also based on additional research including Ladley (2012) and Sen (2019):

- **Strategic layer [including (a) Leadership and Vision]:** Some of the data governance elements in this layer include national data strategies, and leadership roles. It is worth noting that the model considers data strategies as an element of good data governance. This argument lays on the fact that data strategies enable accountability in relation to responsibilities and can help define leadership, expectations, roles and goals. The strategic layer also highlights how the formulation of data policies and/or strategies can benefit from open and participatory processes, thus integrating the inputs of actors from within and outside the public sector towards greater policy ownership.
- **Tactical layer [including (b) Capacities for Coherent Implementation and (c) Legal and regulatory frameworks]**. It enables the coherent implementation and steering of data-driven policies, strategies and/or initiatives. It draws upon the value of public sector skills and

[11] For more information see: *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.*

competences, job profiles, communication, coordination, collaboration as instruments to improve the capacity of the public sector to extract value from data assets. It also highlights the value of formal and informal institutional networks and communities of practice as levers of public sector maturity and collective knowledge. This layer also comprises data-related legislation and regulations as instruments that help countries define, drive and ensure compliance with the rules and policies guiding data management, including data openness, protection and sharing.

Figure 16
The OECD model for data governance in the public sector



Source: *OECD (2019)*

- **Delivery layer [including (d) the integration of the data value cycle, (e) data infrastructure, and (f) data architecture]**. The delivery layer allows for the day-to-day implementation (or deployment) of organisational, sectoral, national or cross-border data strategies. It touches on different technical and policy aspects of the data value cycle across its various stages (from data production, openness and re-use), the role and interaction of different actors in each stage (e.g. as data providers), and the inter-connection of data flows across stages. In this light, each stage is inter-connected but has specific policy implications in relation to the expected outcomes. For instance, data sharing initiatives (e.g. the production of good quality, standardised and inter-operable government data) can contribute to data re-use by external actors in latter stages (e.g. as open government data). The adoption of technological solutions (e.g. cloud-based data hosting services, APIs, data lakes) takes place in this layer for it supports of those policy goals defined in the strategic layer. It also relates for instance to the need for reengineering legacy data management practices and processes or retrofitting and adapting legacy data infrastructures. Data interoperability and standardisation also take place at this level.

The OECD underlines that the elements used to exemplify the plethora of policy instruments, arrangements, initiatives and/or tools that can be used by countries to deploy their data governance frameworks is not exhaustive. Thus, OECD countries might opt for adopting different data governance elements and tools that better fit into their national context and public sector culture in line with the proposed three layers and the six underlying categories presented in the model.

More broadly, the OECD has been working on developing common principles for data governance. In 2021, the OECD Council adopted the *Recommendation on Enhancing the Access to and Sharing of Data*[12] to define a set of guidance principles for data governance, including across sectors.

Also in 2021, the OECD launched the *Good Practice Principles for Data Ethics in the Public Sector*[13] as an action-oriented tool to support countries in the implementation of this emerging area of work.

* * * * *

[12] For more information see: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0463
[13] For more information see: https://www.oecd.org/digital/digital-government/good-practice-principles-for-data-ethics-in-the-public-sector.htm