Cooperation During Crisis:

A Cross-National View on National Statistical Organisations and Data Stewardship in Times of Crisis

by

Florian Henning[a], Crystal Sewards[b], Paul Grooten[a], Leanne Houben[a], Matjaz Jug[a], Courtney Cameron[b], David McNamee[b], Dominic Parent[b]

[a] Statistics Netherlands (CBS)
[b] Statistics Canada (StatCan)

Date: 21 July 2021
Version: Final

EXECUTIVE SUMMARY

The COVID-19 pandemic has demonstrated the importance of well-designed and evidence-based policies for tackling such massive crises, and the key role of data and statistical information therein. Good data governance and data stewardship are central to effectively meeting the extraordinary demands that societies face in crises like COVID-19.

This white paper will demonstrate that national statistical organisations (NSOs) are in a unique position to take a leading role for data stewardship and prepare societies for effective crisis response as partners in a data ecosystem in charge of co-creating trusted smart statistics. They occupy this position owing to their special mandate, their responsibility for providing these statistics and their specific experience in dealing with and managing various data sources.

Chapter 1 describes a number of **key developments that set today's stage** for the role of NSOs in data stewardship. These are:

- the urgent public need for high quality, relevant and smart statistical information;
- an ever-evolving data landscape that is becoming increasingly rich and diverse but also complicated to navigate;
- a complex data ecosystem with high interdependency between diverse actors;
- the crucial role of trust and social license (the ongoing acceptance of an NSO's practices and operating procedures by its employees, stakeholders, and citizens), as well as acknowledgement of the changing roles of NSOs.

Chapter 2 discusses the imperative for NSOs to **establish the necessary trust and social license** to fulfil their role in future crises and highlights several key areas for attention:

- Enshrining data protection, confidentiality and privacy as starting points;
- Leveraging the FAIR data principles of **f**indability, **a**ccessibility, **i**nteroperability and **r**euse;
- Basing data governance on defined quality standards and "fit-for-purpose" design;
- Guaranteeing that high quality statistics are available on a continuous basis;
- Ensuring transparent communication about statistical data products and their provenance.

Chapter 3 elaborates on the structure of the environment in which the **unique position of NSOs** is mobilized for data stewardship. In particular:

- NSOs—given their mandate to act as stewards of national statistical systems and use data to measure sociodemographic/socioeconomic phenomena—possess the data foundations and data stewardship expertise and knowledge to effectively manage data;
- There are many examples of NSOs providing support in the form of data collection, data standardization, data hosting and data output to ensure that the health, infrastructure and economic needs of citizens and businesses were being met during the pandemic.

Chapter 4 explores the **conditions necessary for NSOs to fulfil a leading role** on data stewardship within statistical data ecosystems. These ecosystems need to be agile and resilient in order to swiftly and effectively respond to crises. The chapter discusses several prerequisites to this end:

- Having a well-functioning and flexible data governance organisation not only at the intra-organisational level (within NSOs) but also at the national and international levels;
- Having clear systems, roles, responsibilities and legal authorities and defining the roles, duties and decision-making bodies, such as the role of a chief data officer as the senior

authority on data;

- Having an integrated data strategy and framework to leverage the strengths of each partner
- Having standards to facilitate interoperability and alignment between partners in the statistical ecosystem and create agility and resilience to crises.

Chapter 5 details the task of **planning and preparing infrastructure and institutions** that are resilient in the case of crisis. Tasks for NSOs include:

- defining the common goals in a national data strategy is the first step in the planning process;
- determining which problems need to be solved to create an effective roadmap;
- investing in partnering in data ecosystems, in terms of both organisational structure and processes as well as in terms of organisational culture;
- shifting towards a culture of co-creation with NSOs to guide data-driven policy responses to crises, which requires the development of appropriate channels for organisations to come together and exchange information;
- establishing a culture of constant innovation and adaptation—key requirements for NSOs to be prepared and to continue to function well in crisis situations—by constantly and critically inspecting their own operational methods, taking an outside-in perspective to open innovation and comparing these perspectives with changes, challenges and newly emerging possible solutions from their environment ;
- maintaining their core organisational DNA and key values that made them trusted partners in the first place

Chapter 6 distils the aspects covered in the previous chapters into a **set of recommendations** for how NSOs can achieve this resilience and provides concrete strategies in the areas of needed products and services of NSOs, metadata, technical solutions and communication strategies.

- **Products and services**: NSOs need to make sure their output addresses society's need for high quality, relevant, timely and smart statistical information in times of crisis. Within a changed landscape of data collection that is as rich as it is diverse and complex, an integrated, whole-of-government strategy for innovation and data ingestion is essential. NSOs need to provide services targeted at maximum transparency while protecting privacy and confidentiality.
- **Metadata**: The FAIR principles provide a suitable strategic framework. Data catalogues, data glossaries, metadata management solutions and application programming interfaces (APIs) are key instruments to this end. It is important to embed all these solutions as much as possible into accepted international standards in order to maximize interoperability.
- **Technical solutions:** Technology plays a crucial role in crisis scenarios. Key strategies in this area are to pursue partnerships to create new capabilities, scale up existing capabilities such as the cloud, invest in Agile technology innovation, and to invest in skills and open source communities.
- **Communication strategies:** Communication must start internally within NSOs in order to keep the organisation functional during a crisis. Externally, communication strategies need to be targeted at statistics stakeholders, NSOs' cooperation partners and the general public. For both internal and external communication strategies, the emphasis and depth of NSOs' communications must adapt to the specific target group.

The paper concludes with a discussion of the key points of the paper and an outlook on its relevance and applicability for future crisis situations.

Chapter 1. Introduction and background

The COVID-19 pandemic has been a "tough teacher" globally and has revealed strengths and weaknesses of national and international systems through their efforts to use the power of data and information to fight the pandemic. The crisis caused by the pandemic has demonstrated the importance of well-designed and evidence-based policies for tackling such major crises and the key role of data and statistical information.

National statistical organisations (NSOs) have a crucial role to play in providing appropriate crisis responses. Good data governance and data stewardship are a central prerequisite to effectively meeting the extraordinary demands that societies face in a crisis situation like the COVID-19 pandemic.

This white paper argues that NSOs, as partners in a data ecosystem in charge of co-creating trusted smart statistics, are in a unique position to take a leading role for data stewardship and prepare societies for an effective crisis response. They occupy this position owing to their special mandate and responsibility for providing these statistics and their specific capabilities to fill this role.

This paper is structured as follows. Chapter 1 introduces the paper and describes a number of key developments that set the stage for the role of NSOs in data stewardship. Chapter 2 explores the strategies for establishing the necessary trust and social license[1] to fulfil this role. Chapter 3 elaborates on the structure of the environment in which the unique capabilities of NSOs for data stewardship are positioned. Chapter 4 explores the starting conditions required by NSOs to fulfil this leading role in data stewardship. Chapter 5 lays out the requirements for planning and preparing the necessary infrastructure, while Chapter 6 distils the aspects covered in the previous chapters into a set of strategies for implementing and communicating the leading role of NSOs in data stewardship. Chapter 7 concludes.

The goal of this paper is to highlight the needs, opportunities and potential role of NSOs as they relate to data governance and data stewardship in times of crisis. The paper acknowledges that the way in which these needs are addressed, and what role, if any, NSOs assume, is highly context-dependent and may vary depending on national contexts, including legal frameworks and the mandates of NSOs. The examples of strategies and measures taken presented in this paper serve to illustrate some of the roles NSOs have assumed in their own countries and may not be applicable to other jurisdictions.

Contextual diversity, but common challenges

Organisational context and culture can vary significantly across NSOs and national and international data ecosystems. As a result, there is no single one-size-fits-all approach, and any strategy to address the role of NSOs in providing data stewardship expertise and services in response to an emerging need or crisis must take into account the specific context.

However, when there is a need for rapid response, such as in times of crises, the role of NSOs

---

[1] Social license can be defined as "perceptions of local stakeholders that a project, a company, or an industry that operates in a given area or region is socially acceptable or legitimate". Source: Raufflet E., Baba S., Perras C., Delannon N. (2013) Social License. In: Idowu S.O., Capaldi N., Zu L., Gupta A.D. (eds) Encyclopedia of Corporate Social Responsibility. Springer, Berlin, Heidelberg.

across the globe is influenced in particular by the following four challenges.

Challenge 1. The urgent public need for high quality, relevant and smart statistical information
Challenge 2. An ever-evolving data collection landscape that is increasingly rich but also diverse and complex to navigate
Challenge 3. A complex data ecosystem with high interdependency between diverse actors
Challenge 4. The crucial role of trust, social license and acceptance of NSOs in changing roles

While by no means being the only forces influencing the role of NSOs in crisis data stewardship, these four aspects appear particularly relevant in this light. They are explored in more detail in this chapter and are a recurring theme throughout the rest of the paper.

Data stewardship and data stewardship services and expertise form a central focus of this paper. There are variations on and different interpretations and definitions of stewardship. Many organisations are currently working on defining data stewardship and data steward roles that will work within their organisational models. For the purposes of this paper "data stewardship" is defined as "the collection of practices that ensure an organisation's data is accessible, usable, safe, and trusted."[2]

Sometimes, the term data stewardship is used synonymously with data governance; however, this is not accurate because the terms denote connected, but distinct, functions and roles. Data governance can be defined as "the exercise of authority, control and shared decision making (planning, monitoring and enforcement) over the management of data assets."[3] Data governance generally focuses on strategic-level policies and procedures, while data stewardship concerns the tactical-level coordination and implementation of these policies in practice.

CHALLENGE 1: THE URGENT NEED FOR HIGH QUALITY, RELEVANT AND SMART STATISTICAL INFORMATION

Our societies need data-driven and evidence-based policies. In a situation such as the COVID-19 pandemic, the urgency for fast and actionable statistics for policy makers and for society at large has never been greater. The COVID-19 pandemic has tested our ability to harness quality and timely data in a responsible way, including cutting across institutional silos. The pandemic has demonstrated that the value of good quality data is more important than ever in making the lives of citizens better.

NSOs fulfil a critical function in supporting fact-based, efficient crisis response management. They possess the sophistication, tools, expertise and data classifications to allow for protecting and responsibly sharing data to drive data value and for responding to the pandemic and any emerging needs. Chapter 2 explores this topic in more detail.

The ability of NSOs to fulfil this function depends on conditions, such as governance structures and processes that must be in place before a crisis strikes. Hence, the reaction of NSOs to crises, to a large degree, also means data governance must be established before a crisis emerges.

---

[2] Informatica. (2021). What is Data Stewardship? Retrieved from www.informatica.com/resources/articles/what-is-data-stewardship.html.
[3] DAMA International. (2017). DAMA Dictionary of Data Management (2nd Ed.). Basking Ridge, NJ: Technics Publications.

CHALLENGE 2: AN EVER-EVOLVING DATA COLLECTION LANDSCAPE THAT IS INCREASINGLY RICH BUT ALSO DIVERSE AND COMPLEX TO NAVIGATE

The landscape within which NSOs collect data and produce statistics has changed drastically. As a previous joint Statistics Canada (StatCan) and Statistics Netherlands (CBS) vision paper[4] on advanced data collection points out:

> The digital transformation, data revolution and emergence of "big data" all influence the way NSOs collect data. Data are everywhere, generated by everything and everyone. These data are stored in numerous locations and devices. The enormous increase in computing power and cost-efficient storage capacity have created never-before-seen analytical performance capabilities.

Datafication of society makes many new data sources available. This is both an opportunity and a challenge for NSOs when dealing with crises such as COVID-19. On the one hand, the existence and (potential) availability of ever-more data sources provide possibilities for fast, phenomenon-oriented statistics. On the other hand, they result in a data landscape that is increasingly diverse and with a complexity that makes it difficult to navigate, both technically and organisationally.

Data discovery and exploitation efforts of NSOs have increased as a result of the pandemic. Ingestion, cleaning and classification of diverse and complex data assets has driven an increase in the use of modern tools and techniques for data ingestion (such as data collection through connection, crowdsourcing, web panels, web scraping, waste water testing, the use of artificial intelligence in natural language processing), data cleaning (such as the use of artificial intelligence and machine learning), storage and access (such as the protected cloud), and output (such as flash indicators, hosting services, COVID-19 data hubs).

The increasing heterogeneity of data, combined from multiple and diverse sources, creates its own challenges when it comes to quickly producing statistics in response to crises. It becomes increasingly essential that the corresponding metadata is well governed, so that data are findable and interpretable with ease (Chapter 4 describes this in more detail).

While statistical crisis response requires a faster course of action, it is critical that it does not result in shortcuts in data protection and security. The need for a quick response strengthens the need for well-governed data protection; not only because the necessary mechanisms and procedures need to be in place before crisis strikes, but also because the ability of NSOs to provide data-driven crisis response is highly dependent on their continued role as a trusted data organisation (Chapter 3 addresses these issues).

CHALLENGE 3: A COMPLEX DATA ECOSYSTEM WITH HIGH INTERDEPENDENCY BETWEEN DIVERSE ACTORS

The changing data landscape also means that the ecosystem in which NSOs operate, consisting of other actors that provide, process and use data and resulting statistics, is increasingly important. Where NSOs in the past were "self-sufficient" in terms of their (primary) data collection, processing/analysis and distribution chains, this is no longer possible in today's data landscape to respond swiftly to crises such as COVID-19, and NSOs are increasingly dependent on external

---

[4] Salemink, Dufour, van der Steen. (2019). *Future Advanced Data Collection*. Conference of European Statisticians: Workshop on Statistical Data Collection 'New Sources and New Technologies'. Geneva, Switzerland.
[4] Statics Canada. (2021). Statistics Canada's Trust Centre. Retrieved from www.statcan.gc.ca/eng/trust.

partner organisations. This creates a situation where NSOs are partnering with other organisations (including data owners), as access to data for timely statistics becomes something that needs to be continuously managed and negotiated, instead of a one-off data delivery (Chapter 5 discusses the engagement with the data ecosystem in more detail).

Within these data ecosystems, NSOs can function as platforms that enable broad cooperation between municipalities, scientific institutes, governmental bodies and private businesses. The development of data service centres for specific economic branches or sector-wide data hubs (e.g., national facilities) with various public and private data complements the linking patters of NSOs. Examples of data collaboratives are possible in some countries where NSOs were well positioned to steward data (store, clean and standardize data and provide linkage platforms, business intelligence, output) related to statistical output and modelling include personal protective equipment (PPE) supply chains, contact tracing and data related to vaccines and economic recovery.

In a context of new requirements from the environment and a steady stream of innovations and changes in methodology and technology, quality, metadata, standards, confidentiality, privacy and acceptability, NSOs must have the ability to constantly learn and innovate in order to be ready to cope with crises. Data governance requires, in part, an approach to future-proofing that includes significant investment in innovation and involvement in national and international networks for learning and innovation. This approach should embrace extensive collaboration not only between NSOs but also with governments, end users, academic institutions, research organisations and the private sector (Chapter 5 explores the topic of innovation and co-creation within partner ecosystems in detail).

CHALLENGE 4: THE CRUCIAL ROLE OF TRUST, SOCIAL LICENSE AND ACCEPTANCE OF NSOS IN CHANGING ROLES

It is essential for NSOs to maintain social license in order to effectively carry out their role. First and foremost, to ensure they maintain social acceptability, NSOs must maintain relevance, measure what is most important to citizens and drive trust by being open and transparent in their processes. They must be able to provide relevant information on ongoing and new phenomena in a timely manner and at increasing levels of detail and granularity.

Furthermore, it is essential that citizens understand data, its importance and why NSOs require access to it. NSOs need to show that good quality data make the lives of citizens better and that good quality data are benchmarked and certified by authorities such as NSOs. This need has never been more evident than during the COVID-19 crisis. Also, in order to continue to build trust and maintain social license, governments (NSOs included) must have a robust understanding of their citizens and must clearly and consistently communicate why good quality data lead to better outcomes. One way to achieve this is by offering numeracy and data literacy training.

The role of NSOs as a trusted party is essential for them to respond to crises, not only as a provider of trusted smart statistics but also as a trusted data partner, which becomes increasingly important during times of crisis response. Developing conditions like privacy-preserving techniques strengthens this position, conditions that are fundamental developments for the future (Chapter 2 discusses this in more detail).

Maintaining social license between NSOs and citizens and businesses is a circular relationship where the more NSOs adapt and adjust to the ever-changing needs of the data ecosystem and stakeholder requirements, the more this drives further collaboration, which in turn will generate

examples and use cases that increase both awareness of the significance of the work of NSOs and trust in NSOs as providers of essential information (this is explored further in Chapter 5).

One of the results of the data revolution has been the creation of statistical outputs by other organisations—these organisations leverage the availability of new data sources, tools and techniques to create output. NSOs are needed, now more than ever, to continue being an authoritative voice on data stewardship issues including standards, data quality, data science, statistical methods and confidentiality and privacy protection. NSOs must also continue to measure what is most important to citizens by leveraging their data as a strategic asset, acquiring new data assets to drive new insights, leveraging new statistical methods and ensuring they bring the right talent into their organisations. In addition, NSOs must work to bring in the right data and the right amounts of data.

Moving away from a classic survey-based approach of statistics production to using a greater heterogeneity of data sources means that data quality issues take on new dimensions and will have to be addressed in different ways. As well, data reusability and the power of data linkage to produce new outputs—made possible through the use of standardised metadata—become critical to finding data, combining data and growing data value to produce phenomenon-based statistics (these topics are discussed further in chapters 2 and 4, respectively).

Chapter 2. Trusted smart statistics and social license

The previous chapter argues that national statistical organisations (NSOs), as partners in a data ecosystem in charge of co-creating trusted smart statistics, are in a unique position to take a leading role for data stewardship and prepare societies for effective crisis response. This chapter will explore the strategies for establishing the necessary trust and social license to achieve this.

FAIR data principles

The FAIR data principles can serve as an important vehicle to achieve social license. The FAIR principles were established in 2016 in order to facilitate the capacity of computer systems to find, access, interoperate, and reuse data with none or minimal human intervention. They provide guidelines to improve the **f**indability, **a**ccessibility, **i**nteroperability and **r**euse of data (Wilkinson et al., 2016).

The FAIR principles fulfil an important function for social license in the context of crisis response because they are key to the forces explained in Chapter 1, serving as important enablers to successfully co-create data-driven information in complex data ecosystems.

The FAIR principles also have great significance for social license, since they directly facilitate output legitimacy by enabling NSOs to stay relevant by generating public value quickly and making their information products broadly available, (re)usable and actionable in smart (machine-readable) ways.

Guaranteeing confidentiality and privacy by design

It is imperative for the development of trust that society is guaranteed that the data necessary to produce statistics is secure and that confidentiality and privacy are safeguarded. To this end, data governance needs to be based on privacy by design. Data protection, confidentiality and privacy need to be enshrined as the starting points and functioning as principles that are embedded from the outset into NSOs' processes of data collection, storage, processing and analysis/statistical publication.

Access should be granted based on roles and policies. As users move in and out of roles, they automatically gain and lose access to data as appropriate. Transparency regarding the mechanisms and policies used to secure data will increase public trust. Regular reports should be made to the public regarding real or suspected breaches along with the corrective actions taken.

Ireland's Central Statistics Office – Concept of five "safes" for data hubs

In the example of Ireland's Central Statistics Office, which was able to create a national data hub for research microdata for policy and research purposes in response to the COVID-19 crisis within one week, building their data governance principles around the concept of five "safes" was instrumental to ensuring timely collaboration with the relevant organisations in the health sector (Economic Commission for Europe, 2020). The fives safes are:

- Safe people: All analysts are officers of statistics and have research or statistical credentials.
- Safe projects: Projects have clear public benefits.
- Safe settings: Researchers access microdata via the Researcher Data Portal.
- Safe outputs: Outputs are checked by NSO staff for statistical disclosure.
- Safe data: Data is de-identified (pseudonymised).

Data quality

Trusted smart statistics are also defined by their quality. Producing high quality statistics starts with the quality of the underlying data, which is a function of good data governance. To this end, data governance needs to be based on defined quality standards, which are safeguarded by processes designed to control quality spanning from the outset of data collection, through storage and processing, analysis and all the way to the publication.

Furthermore, quality can be defined as "fit for purpose", a satisfying means to the desired end. In a world that is changing ever more quickly, standing still means falling back; in other words, NSOs cannot rest solely on traditional methods of collecting and processing data. Instead, in order to stay up-to-date with new emerging societal phenomena and unexpected crises such as pandemics, NSOs need to substantially invest in innovative ways of making statistics. Only then can a continuous production of high quality data be guaranteed.

Continued access

This means that NSOs need to guarantee that high quality statistics are available on a continuous basis. This is particularly important in order to respond to crises, when phenomenon-based facts and information are needed quickly to develop appropriate crisis response and policies.

Importantly, continued access is also dependent on trust in NSOs. The more NSOs can count on social acceptability of data collection and statistics production, the more they are able to guarantee the continued production and availability of relevant statistics that in turn will generate more understanding and cooperation from society at large.

Transparency

Transparency about their data products and provenance is a key requirement for NSOs as providers of trusted smart statistics. This is even more so given the background of so-called "wicked" crises (complex multidimensional and borderless problems) such as a global pandemic, which intensify

the need for increased collaboration with the partners across society, and across borders.

Users, such as policy-makers responding to a crisis, need to be able to quickly understand and assess the data products they use to inform (policy) decisions. When it comes to resilience to crises, data sharing (internal and external) is an important capability, and good metadata governance and the FAIR principles play a relevant role to this end (this is addressed more extensively in Chapter 4).

Clear and transparent communication about data, statistical products, processes and methods (including possible limitations) is a prerequisite with regard to the FAIR principles. Users of official statistics come to NSOs because it is the mission of NSOs to provide high quality, unbiased and transparent information to the public, which gives users the necessary information to assess official statistics in order to use them appropriately in a given context. Independence and a non-profit basis are important principles for NSOs in order to guarantee the necessary information is available. Operationally, such transparency of NSOs rests upon the requirement for good governance of aspects such as data provenance and process metadata (discussed further in Chapter 4).

NSOs must engage citizens and ensure they tell data stories—explain to citizens what their data are being used for – for instance by using case studies describing how data are collected, cleaned, standardized, anonymised, analysed and used to drive policy-making decisions and services contribute to better outcomes for citizens. A lack of communication fosters distrust. Communicating to citizens about the benefits of data collection, the methods and processes used to protect their data and their rights in terms of how the data will be used and shared fosters trust and shows that the organisation collecting the data respects the citizens. This communication becomes a collaborative exercise and fosters the virtuous cycle of trust because, as each data collection exercise demonstrates benefits for citizens, they are encouraged to participate in these activities.

Statistics Canada Trust Centre

Statistics Canada developed a trust centre[5] on its website to provide Canadians with clear information on how their data is used, and protected, to deliver value to them. The agency has also developed an ethical framework for data collection activities based on the principles of necessity and proportionality. The development of the StatCan Trust Centre and communication of the necessity and proportionality framework will open up more sources of data, allowing the agency to produce more timely insights and increase transparency and trust. In today's digital and data-driven world, governments, citizens and business must work together to build and maintain this foundation of trust.

Given that today's NSOs operate in complex data ecosystems, their approach to innovation and developing advanced data collection capabilities must involve collaboration with external partners (this is addressed in more detail in Chapter 5). NSOs, as stewards of national statistical systems, are responsible for providing a trusted, credible and independent source of statistics and data. Producing high quality, relevant and timely data and statistics on what is most important to citizens, organisations and institutions is what drives the mandates of most NSOs. This trust, adherence to quality frameworks and drive for relevance is expertise that has been vital to

---

[5] Statistics Canada. (2021). Statistics Canada's Trust Centre. Retrieved from www.statcan.gc.ca/eng/trust.

responding to the global COVID-19 crisis.

The pandemic has had major effects on travel, global markets and international trade; it has also magnified existing socioeconomic differences. Countries with less-developed statistical infrastructures have had a more difficult time assessing the impact of the crisis and the impact of recovery policies than those that have the necessary health and economic data. Also, in a world where, more than ever, large volumes of data come from more providers, how can the knowledge and expertise of NSOs be leveraged, and leveraged where it is needed most?

The next chapter explores this expertise in data management and data stewardship and how it can be leveraged to ensure a healthy data ecosystem and healthy national statistical system to ensure countries are able to respond to emerging needs such as pandemics.

Chapter 3. The unique position of national statistical organisations for data stewardship

The COVID-19 pandemic has emphasised the need for governments to provide data and expertise to ensure that citizens receive meaningful insights and information. This chapter elaborates in greater detail on what sets national statistical organisations (NSOs) apart, and has led many governments around the world to look to their statistical offices for leadership during the pandemic.

Data foundations

NSOs, given their respective mandates to be stewards of their national statistical systems and use data and information to drive insights to measure sociodemographic and socioeconomic phenomenon possess the data foundations and data stewardship expertise and knowledge to effectively manage data. More specifically, NSOs already have the following frameworks and expertise that can be leveraged to provide data, information, insights and visualizations to respond to the emerging needs of citizens, businesses and other institutions:

- Legislative frameworks and legal authority to collect, process and disseminate data and information (e.g., *Statistics Act* in Canada);
- Well-established data quality frameworks and modernised frameworks which respond to a rapidly evolving data ecosystem;
- Long history of data protection and leading edge methods to protect privacy and confidentiality;
- Modern statistical methods for data integration and linkage, analysis, categorization via metadata and visualization;
- Growing data science expertise grounded in statistical and mathematical theory;
- Data ethics and growing experience in the ethical use of artificial intelligence and machine learning to ensure methods are without bias;
- Trust frameworks and development of new frameworks such as proportionality and necessity;
- Data visualisation and telling the data story—from data to insights to policy to services—for better outcomes;
- Access to resources and networks within the national statistical system and a long history of data partnerships and collaboration both domestically and internationally;
- Collaborative work spaces (protected cloud, data collaboratives, data trusts);
- High data maturity and high data literacy;
- Statistical and data standards and classification systems to give data meaning and drive interoperability;
- Subject matter expertise on socioeconomic and sociodemographic phenomenon;
- Data collection expertise—including traditional surveys, pulse and panel surveys, web scraping—and ingestion and cleaning of administrative data (structured and unstructured);
- Well-established data dissemination vehicles, including web presence, data hubs and data hosting services;
- Expertise in modelling and forecasting;

- Statistical registers reference data and master data expertise;
- Emerging entrepreneurial spirit and culture of continuous learning and development of innovation ecosystems.

Sharing expertise

Although most NSOs around the world have the tools, frameworks and expertise to leverage data to provide insights, many are modernising in order to better meet the challenges and developments summarised in Chapter 1. The COVID-19 pandemic required many NSOs to respond rapidly and accelerate their modernisation journeys. There are many examples globally of NSOs providing support in the form of data collection, data standardization, data hosting and data output to ensure the health, infrastructure and economic needs of citizens and businesses were met during the pandemic. Many NSOs have created data portals and data hubs to ensure data, information and insights are made available to citizens on the subjects that are most important to them and delivered in as close to real time as possible.

Statistics Canada: Modernisation in the face of COVID-19

Statistics Canada was well on its modernisation journey before the pandemic hit and, therefore, well prepared to respond to the pandemic. The agency prioritized and scaled up statistical capacity in critical areas and adapted processes and products to meet the urgent demands for data and insights. The agency now produces monthly flash estimates of high-frequency indicators such as gross domestic product (GDP) to respond to the need for more timely estimates rather than estimates that are perfectly accurate. Statistics Canada has also created a COVID-19 data hub and produced data visualization tools such as the Canadian Statistical Geospatial Explorer and released over 100 analytical papers on the impact of the virus on the nation's economy and society—these have all extended the agency's traditional reach significantly over the past year.

New Zealand's COVID-19 data portal[6]

Adopting a customer-centric agile approach, New Zealand's statistical agency, Stats NZ, proactively engaged with key government and private sector clients to understand and meet their COVID-19 data needs. Stats NZ found that a majority of clients, while appreciating the accuracy and coherence of their traditional data products (e.g. GDP), were prepared to sacrifice these for better accessibility to timelier and more frequent data. Stats NZ's COVID-19 data portal is a compendium of primarily non-official indicators relevant for understanding the economy in the context of COVID-19. It has improved access to higher frequency and timelier indicators sourced from public and private sector sources. Stats NZ ensures that all indicators on the portal are up to date, can be presented graphically and are easily accessible via download or an Application Programming Interface (API), but the agency does not vouch for their accuracy or the validity of their concepts and methods.

---

[6] Statistics New Zealand. (2021). COVID-19 Portal. Retrieved from https://www.stats.govt.nz/experimental/COVID-19-data-portal.

The key to moving forward, leveraging lessons learned and ensuring that governments do not revert back is ensuring that trust and social license are maintained. As discussed in Chapter 2, this will mean consulting and engaging with citizens and businesses proactively, deliberately and consistently when collecting data and strengthening governance frameworks on ethics, trust and privacy.

Stats NZ Data Ventures

Data Ventures is the commercial arm of Stats NZ. It is a trusted intermediary that pulls datasets from, for example, telecom and debit card companies to map population density in real-time so the government can make crisis management decisions.

Early in the pandemic, Data Ventures began accessing Global Positioning System (GPS) data from telecom companies in order to get hourly people counts—anonymized and aggregated to a small geographical level—to monitor commuting and other mobility changes during lockdown. It also accessed card spend data by terminal, broken down into small areas by industry, to gain economic insights. These powerful datasets enabled national crisis management centres to do pandemic risk modelling and informed their lockdown strategies. The use of New Zealanders' sensitive personal information was positively reported in the press and had wide societal acceptance[7]. The work by Stats NZ and the rest of the country's public sector created a surge in public sector trust from the previous year (between December 2019 and December 2020), with an 18 point rise to 69% of New Zealanders saying they trust the public sector[8].

New roles for NSOs

Although NSOs all over the world were called upon to support government responses to the pandemic, some experimented with new and innovative ways of producing powerful insights and products to respond to the ever-evolving needs of citizens, businesses and other organisations. Going forward, countries must ensure that the advancements made to operational methods throughout the COVID-19 crisis are maintained and built upon. Strategic partnerships, sharing platforms, data and insights "right here and right now" must become the new reality. Governments must continue to build data foundations to ensure there is wholly transformational change in the way the public sector responds to the ever-evolving needs of citizens. New data ownership and stewardship models grounded in laws and legislation will remove barriers to strategic use of data and drive new insights.

The pandemic has changed the view and role of NSOs across governments and countries. As a result of the pandemic, the need for a centralized, data-mature organisation with a strong mandate and governance may be required to ensure the proper stewardship of data. This makes NSOs with their historical ties to being data providers well suited to taking on this new role in a whole-of-government approach. NSOs now find themselves in the midst of a 21st-century data revolution, evolving from the more traditional role of a data producer into that of a modern

---

[7] Data Collaboratives. (n.d.) StatsNZ Data Ventures. Retrieved from https://datacollaboratives.org/cases/stats-nzs-data-ventures.html

[8] Early, Catherine. (2021). Trust in public services soars in New Zealand due to handling of COVID-19. *Global Government Forum.* Retrieved from https://www.globalgovernmentforum.com/trust-in-public-services-soars-in-new-zealand-due-to-handling-of-COVID-19

service provider.

Internationally, countries have been working together to define stewardship models and new and emerging roles for NSOs. Once such example is the work being done as part of the United Nations Economic Commission for Europe (UNECE) Task Force on Data Stewardship[9]. The mandate of this new task force is to clarify the terms related to data stewardship and public data governance as well as the tasks of NSOs that the mandate may include in different settings.

---

[9] United Nations Economic Commission of Europe. (2021). *Terms of Reference for the Task Force on Data Stewardship*. Conference of European Statisticians: Workshop on Statistical Data Collection 'New Sources and New Technologies'. Geneva, Online.

Chapter 4. Agility through standardisation

Flexibility is key in times of crisis

Agility and resiliency are key requirements of national statistical organisations (NSOs) in order to have the ability to fulfil their important role concerning data stewardship in times of crisis and beyond. NSOs must produce statistical output in a timely fashion in order to deliver rapid response. An important prerequisite for this are standards, which enable response to ad hoc demand for statistics in a flexible way that builds on interoperability between partners in the statistical ecosystem.

Insight into the situation is essential for managing the crisis and requires data from various sources. Which sources will yield the best results is highly dependent on the nature of the crisis and is difficult to predict. Sound data governance can help ensure that the necessary data sharing takes place quickly and effectively. Moving away from a traditional survey-based approach to the production of statistics with a greater heterogeneity of data sources, data standards are the key to interoperability, data reuse and data linkage. Data reuse and data linkage result in the production of phenomenon-based statistics and the measurement of things that were not previously measurable.

Metadata is one area that is critically important in this regard. The ability to combine diverse data sources swiftly to produce relevant statistics requires a search-and-find infrastructure that starts with good metadata management[10]. Metadata management is achieved through the systematic use of human readable, machine actionable metadata standards, vocabularies and ontologies. Metadata is linked to data through the use of persistent identifiers (PIDs) and made available through a catalogue. An essential part of this are, thus, well-described data according to common standards and interoperable data models so that users and machines can identify and recombine the relevant datasets as needed.

Data quality frameworks are fundamental to data and to the statistical output of NSOs. Applying these frameworks to ingested, cleaned and standardized data is one of the strengths NSOs bring to a partnership agreement. NSOs also bring privacy protection expertise to the partnership, including leading-edge techniques for masking, de-identification and anonymization of data to ensure security and permit limited sharing. The power of the FAIR principles comes from the issues that gave rise to their existence; datasets are bigger, more numerous, more varied, and more complex than ever before. These datasets are also more often made available to people outside the originating organisation than before. FAIR allows this multifaceted collaboration to occur while maintaining trust and at the same time driving interpretability and understanding of the data. This facilitated by ensuring that people and institutions embed the required information to work with a dataset into a larger package.

Related to this is the requirement for NSOs to have a good data catalogue that is complete and up-to-date. This is a prerequisite to ensuring that available data can be found and used, and also identifying relevant data gaps that need to be addressed by targeted and focused data scouting efforts.

---

[10] Salemink, Dufour, van der Steen. (2019). Future Advanced Data Collection. Conference of European Statisticians: Workshop on Statistical Data Collection 'New Sources and New Technologies'. Geneva, Switzerland.

Statistics Netherlands: General data catalogue

At Statistics Netherlands (CBS), for instance, a general data catalogue provides an overview of all secondary data with key descriptors. Potentially relevant data sources (or types of new data) are continuously added to the catalogue wish list so that key areas for necessary data scouting can be identified in a systematic manner and data scouting activities can be targeted and designed in an organisation-wide approach that optimises synergies.

Statistics Netherlands: Central Data Ingestion Office as a single point for data ingestion

Organisational interoperability becomes important to bring together diverse data into statistical products. CBS has set up the Central Data Ingestion Office program to facilitate a standard process for ingesting diverse data sources and making them available to CBS statistics producers. By establishing standard operating procedures and structures that allow this process to happen in a seamless manner, this program will greatly enhance the ease, speed and safety of getting from external data sources to new statistics.

Partnerships and standards governance

Not only are standards (e.g., definitions and permitted use) for data sharing important but so too are agreements on the quality of the data and how it can be measured so that the receiving party knows what to expect. Therefore, the definition of the metadata to be delivered is very important. To achieve this, it is critical to have a well-functioning and flexible data governance organisation that can act quickly and is free of bureaucracy.

Addressing this need is something that must happen not only at the intra-organisational level (within NSOs) but also at the inter-organisational, national and international level. This paper has already established that, in terms of an appropriate crisis response, NSOs need to consider the broader ecosystem and interorganisational partnerships within said ecosystem.

This raises the question how to organise this collectively, so that the appropriate standards are set and governed at an appropriate level. On a general level, the Data Management Body of Knowledge (DMBoK) model (Association, 2017) distinguishes between three types of operating models for data governance that can serve as a framework for this discussion. These types of models are located along a spectrum of centralised and decentralised governance, with a "federated" model in between.

a) Centralized   b) Federated   c) Decentralized
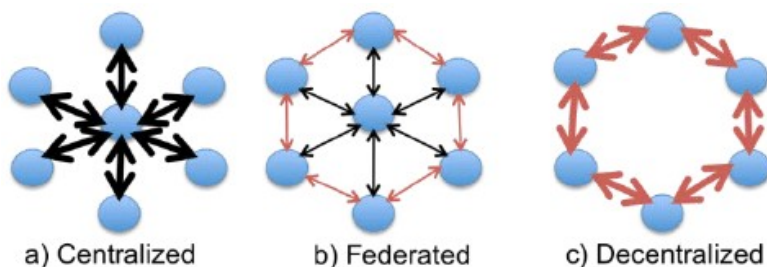
A decentralised model has the risk of leading to a fragmented approach that can result in slower crisis response, while also having the advantage of decision-making taking place closer to local problems. At the other end of the spectrum, a centralised model can focus on a more unified approach, but might also take longer to respond to crises because information needs to "travel" further to a central decision-making instance. Previous research on the governance of interoperability in multi-organisational government information networks has indicated that there is no one-size-fits-all answer to what that the level and model of organising need to be because the choice of model is highly context-dependent[11]. In practice, most situations will, therefore, likely require a hybrid approach, possibly leaning closer to one end of the spectrum—somewhere between a centralised and decentralised approach. Whatever the appropriate governance model for a given country might be, it is clear that based on their unique role, NSOs cannot afford to stand outside of this decision-making organisation and, in fact, are well suited for taking on a leading (broker) role in the networks.

In the federated approach, data governance is centrally led by a core team with local satellite locations. The core team determines the program, the processes, standards and systems and ensures good communication with the satellites. Business programs are responsible for the application of and compliance with data governance at the local level. All business programs are represented in the core team, and cooperation is essential so that the agreements are supported by all departments. This form of organisation is known as highly collaborative and organised. This federated model interacts with the four challenges presented in Chapter 1 as follows:

- The public's great need for high quality, relevant and smart statistical information in order to cope with the crisis:
  - ➤ Quality is often under pressure when there is a crisis. The question is, therefore, whether the highest quality is always needed for all the questions that exist. However, it is important that the quality level is made transparent by means of clear definitions of how quality can be measured quickly and adequately. The fit-for-purpose principle is important: can the data be used for the requested purpose? And if, given the circumstances, it is necessary to deliver the data anyway, how can it be provided with an explanatory leaflet that enables the receiving party to use the data properly?
- The complexity of the situation, formed by an ever-changing data landscape:
  Typically, in a crisis such as the COVID-19 pandemic, new collaborations arise between parties that previously had no connection. In such cases, how can NOS support the process of making agreements as effectively and as flexibly as possible? The use of standards and templates is essential for this, where the focus should be on the collaboration and not on bureaucracy.
- A diverse statistical ecosystem with a high degree of interdependence between the actors:
  - ➤ Knowing what can be expected of you and what information you have in-house is a must for external parties in order to estimate what the value of an organisation can be in times of crisis. Metadata is crucial in such situations. In order for NSOs to operate flexibly, some of the information that should be made immediately available includes how quickly the metadata can be made clear, what is available, where these data come from, what frequency they are available and how they can be used. An updated general data catalogue is essential in this respect.
- The crucial role of trust and social license in order for NSOs to fulfil their special role:
  - ➤ The mandate, image and trust that an NSO enjoys in a country is essential not only for the

---

[11] Henning, F. (2015). Living up to standards : interoperability governance and standards adoption in government information networks. Uitgeverij Boekenplan.

usability of the data but also for the willingness of other organisations to share their data. Clear and transparent data governance is very important in this respect. The quality of the data, compliance with data principles, compliance with regulations (e.g., the General Data Protection Regulation, GDPR[12]) and having transparent escalation and appeal procedures are important for the trust that NSOs have. As a starting condition, it is important to check what level of trust an organisation currently has in society. In addition, it is good to examine the resources at their disposal in order to maintain or increase that trust.

Governing strategic data partnerships

Defining roles, responsibilities and decision-making bodies such as the proposed role of a senior authority on data such as a chief data officer is key for good data governance. There is also the question of legislative and judicial authority over data ownership, data stewardship and data sharing rights and privileges. Clearly defined barriers, challenges and mitigation strategies are also important considerations. These are preconditions for creating data partnership agreements to respond to new and emerging needs such as pandemics and natural disasters. Having clear authorities, roles, responsibilities and legal authority will ensure that trust and social license are maintained and that timely and relevant insights are produced.

Strategic partnership agreements that include multiparty, cross-jurisdiction, cross-sector actors can include varying requirements but typically have one common goal. Each actor has a specific role (or roles) to play but may also have individual goals and demands of the partnership. Public sector organisations may be assigned data ownership or data stewardship rights and data sharing privileges. Private sector organisations may be providing collaborative work spaces, data analytics or modelling and visualization support. It is important that each organisation in the partnership have these goals and demands clearly defined as part of the partnership arrangement or agreement.

Having a strong and integrated data strategy and framework is key to leveraging the strengths of each partner. NSOs often play a leadership role in these types of partnerships, leveraging expertise to identify key questions or concepts for measurement, identify data sources, apply quality frameworks and standards, provide modelling and analytics support and release data and insights that tell the data story—all while protecting privacy and maintaining trust.

The goal of any strategic partnership is to create a common output that meets the needs of each organisation in the partnership. NSOs enter partnership arrangements with other organisations for several reasons, most often to measure phenomena that cannot be measured without the use of partners. For example, the COVID-19 pandemic has created the requirement for governments to produce data and information related to the state of the personal protective equipment (PPE) supply chain, secondary health conditions (related to the pandemic) and current labour market conditions. Measuring these phenomena requires answers to many questions, including:

- What is to be measured (clear definition)?
- What data or information is required?
- Where will these data come from or how will they be created or acquired?
- Where will these data reside?
- Who will be the owner or steward?

---

[12] A regulation in EU law on data protection and privacy.

- How will the data be cleaned, standardized and protected?
- Can the data be shared?
- What are the required outputs (measurements, data, statistics, visualisations)?
- Who will be responsible for data or information releases?
- How will the public be engaged and who tells the data story?
- How long will the data or information be retained?

The Swiss statistics data steward model

An example of the above models in practice is how Switzerland is addressing the establishment of a once-only principle[13], which requires interoperability for data integration and exchange across organisational units. As a first step, Switzerland is developing a national metadata catalogue containing standardised and harmonised descriptions as well as information about provenance, usage, legal restrictions and quality of shared data. This required an effective data governance, "including a role model and processes for standardization and harmonization, taking into account the federalist structures in Switzerland"[14]. As a result, Switzerland opted for a more federated governance model, with as "key concept to keep data storage and ownership decentralised in the individual administrative and statistical units, but to make the data available and reusable by sharing it with other units"[15].
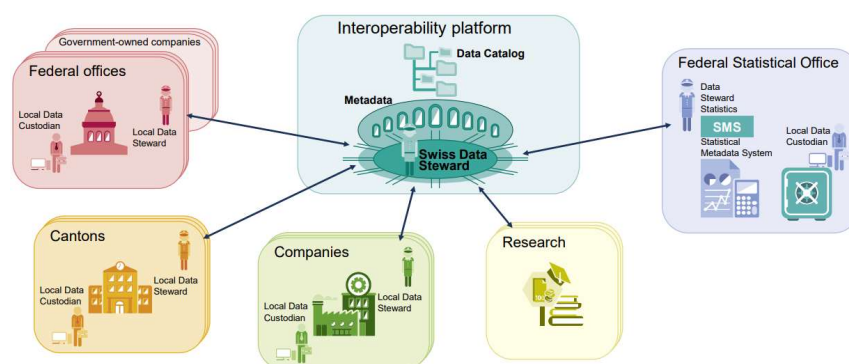
In this model, a data steward for statistics (centrally located in the Swiss NSO) is in charge of coordinating with international bodies and enabling the use of common international classification standards as well as developing guidelines for implementing the once only principle in the field of statistics and defining standards and guidelines for nomenclature and metadata quality assurance of the statistics producers. In addition, local data stewards are situated in the public administration offices. They are responsible locally for data standardisation and harmonisation within their domain and for identifying the data needs of the various users in the statistical field and developing additional specifications to ensure interoperability.

---

[13] The Once-Only Principle aims at the reuse of data that has already been collected once by public authorities, and not asking respondents for it multiple times. This can contribute to administrative efficiency, reduction of administrative burden and better data quality. Interoperability is a prerequisite for making this possible.

[14] Bundesamt für Statistik (2020). Swiss Data Steward responsible for standardization and harmonization of data for nationwide public administration. Simplification for companies and citizens through the Once-Only Principle. Geneva: Conference of European Statisticians, Sixty-eighth plenary session

[15] idem.

A recent discussion paper by Statistics Ireland underlines a similar argument:

> As the primary role of a data steward is data quality, including capturing/documenting meta information and meta data (including definitions, rules, logical models, code sets and classifications), identifying data custodians/owners, and documenting all elements of a data quality framework, NSOs are well positioned to assume the role of Data Stewards....A secondary role for NSOs would be the certification of data stewarded by other organizations but which forms part of the national statistical system; this certification is based on the same standards as those applied to the NSO stewarded data. NSOs could also provide training and knowledge transfer to ensure all data in the national statistical system follows the same rules, standards, legal and regulatory framework to enable sharing and discoverability. [16]

This argument underlines the argument that, for any chosen form of organising crisis response, NSOs are in a key position to have a central role regarding data stewardship.

The need to be prepared

In conclusion, this chapter has argued that standards are indispensable not only for well-governed standard operations but also to enable NSOs to respond to crises and respond quickly. Creating agreed-upon templates increases organisational agility and ensures that data can be found, linked and processed for a timely statistical crisis response, which is particularly important in an inter-organisational ecosystem of data partners. Setting standards, including an agreement on which standards to follow and implementing them, is a complex undertaking, but NSOs are in a unique position to support this process.

A well-standardised, trusted and interoperable system must be in place before a crisis strikes. Therefore, standardisation is part of the larger tasks of planning and of preparing infrastructure and institutions to ensure resiliency in times of crisis (Chapter 5 discusses this task in further detail).

---

[16] Economic Commission for Europe (2020). Emerging data stewardship role of the Central Statistics Office of Ireland and its deployment in the COVID-19 crisis. Geneva: Conference of European Statisticians, Sixty-eighth plenary session: https://unece.org/fileadmin/DAM/stats/documents/ece/ces/2020/ECE_CES_2020_29-2007757E.pdf

Chapter 5. Planning for adaptation and innovation

A changing ecosystem

As the previous chapters have described, through the COVID-19 pandemic, the world has witnessed a shift in the culture of collaboration with national statistical organisations (NSOs). The crisis has highlighted that data-driven and evidence-based policy responses to the pandemic need to be a "team sport".

The pandemic has, for instance, served as a catalyst for opening up access to private data sources that previously were not available. One reason behind this increased data access as a result of the COVID-19 crisis is that NSOs partner organisations realise both the necessity of working together to fight the pandemic and the public value their data hold.

New private data sources at CBS

As an example, in the context of COVID-19 crisis, Statistics Netherlands (CBS) has received access to financial transaction data from a range of payment service providers in order to measure the economic impact of the pandemic. These data help show the sectors of the economy where spending is falling and those where it is rising. In this way, CBS aims to understand the financial impact on different sectors. CBS now receives financial transaction data weekly, enabling it to have a clear picture of the very latest developments and to provide policy-makers with information they can use in their policy deliberations.

NSO partners in the data ecosystem are not limited to private parties. Many collaborations in the COVID-19 crisis also involve research institutions such as universities as well as other governmental bodies, at national and international levels. Several countries have facilitated the need for high quality accessible data and statistics by starting up data collaboratives, data hubs and data trusts.

Examples of NSOs stewarding data hubs and collaboratives are found later in this chapter. The stewarding and hosting by NSOs of data that were meant for non-statistical purposes has been an interesting development globally and could prove to be an important change in the way governments view NSOs. Data trusts, and data trusts held by NSOs, could by leveraged to further expand this stewardship role.

> "A data trust is a legal structure that provides independent (third-party) stewardship of data, with informed consent from the data subjects" and "Data Trusts also guarantee that the people that have data, use it, and are affected by its use have a mechanism to ensure it is shared equitably. A successful Data Trust is a set of principles that facilitate the interaction between citizens, government, and third-parties, promoting open and democratic access to data. A Data Trust should be a model for stewardship and management of data, presided over by a neutral governing body or board." [17]

---

[17] Wiley, B. & McDonald, S. (2018). What is a Data Trust?. *Centre for International Governance Innovation. Retrieved from* www.cigionline.org/articles/what-data-trust
Hardinges, J. (2018). Defining a Data Trust. *Open Data Institute.* Retrieved from www.theodi.org/article/defining-a-data-trust

Data strategy: A tool for managing and leveraging data assets to ensure optimal use and value while maintaining public trust

A data strategy is a plan designed to improve all the ways data is acquired, stored, managed, shared and used. Data strategies have become invaluable tools for creating a centralised vision and foundation for using data as a strategic asset. Many countries are developing national or federal data strategies to drive their data agenda.

The first step in any planning process is determining what problem(s) are to be solved, often framed as "what question(s) are to be answered?" NSOs are using data strategies to plan and manage projects as they lay out the framework for solving or attempting to solve data problems. The basic outline of any data strategy would include the answers to the following questions:

1. What question(s) are to be answered?

2. What data already exist in the ecosystem, within the NSO?

3. What data are missing and can the gap be filled?

4. How are the data governed and organised in order to provide answers (analysis)?

Once developed, a data strategy will provide the basis for planning and is often presented as a roadmap or action plan. Data strategies also often include roles, responsibilities and accountabilities for implementation as well as performance indicators. In the ever-evolving data ecosystem, it is important that data strategies are "evergreen" and updated regularly to remain aligned with changing user requirements.

Challenges for ecosystem partners: Invest in co-creation

Against the background of increasing awareness for the importance of partnering in data ecosystems is the importance of the readiness and willingness of partners in order to increase awareness in terms of their organisational structure, processes and culture.

It is, for instance, a missed opportunity if private organisations base their strategies on partnering with NSOs only in a traditional client–supplier-based approach, which focuses on the commercial aspects of selling data or services. Sharing data for the benefit of data-driven and evidence-based policy responses to crises reflects the organisation's corporate social responsibility, which it can use in corporate communications and to connect with customers in new ways. Organisations can also learn from joint research and development projects, such as those on innovative data analytics or data-protection mechanisms that NSOs have in place. Moreover, participation in data ecosystems widens their network and the potential for useful spin-offs in the future.

Participation in data ecosystems is by no means a responsibility that is limited to private parties. Government organisations also need to learn to better approach crises like the COVID-19 pandemic, not from an organisation-centric approach that revolves around "stove piped" structures but from a point of view that holds the exponential value of collaboration networks and a whole-of-government approach at its core.

A culture shift in this direction is well underway, but this alone is not enough to ensure a robust data ecosystem. In the technical domain, there are already many solutions and techniques to support secure data sharing. Many promising initiatives are being carried out on privacy-preserving techniques, but there are still many opportunities to incorporate them into operating procedures. Such developments can be accelerated and scaled up on a national basis if there are channels

through which organisations can come together and exchange information. NSOs have a key role to play in this regard because of their long-standing experience with secure data processes and their trusted position. Effective communication is critical for this role, in terms of the need for both a collaborative data ecosystem and sharing information and knowledge on how to achieve it. Effective communication will help other actors who have less experience—or who might not be fully convinced of their added value to the ecosystem, or those who need to identify the necessary changes to their culture, structure and technology in order to join the collective efforts to combat a crisis.

Challenges for NSOs: Innovate and adapt, yet keep their DNA

Constant innovation, thus, becomes a necessity for NSOs and means they need to constantly and critically review their ways of operating and compare them with changes, challenges and newly emerging solutions in their environment. This innovation is not only a matter of organisational culture or mindset but it also requires a flexible organisational structure that can swiftly adapt to new challenges as they emerge and adopt appropriate solutions. At the same time, it is essential that, while being open to necessary innovative "mutations", NSOs retain the central organisational DNA and key values that contributed to their success in their trusted role in the first place.

For NSOs to fulfil their role as leading trusted data guardians, they need to continuously invest in ensuring they remain relevant and reliable. They need to stay focused on their contribution to society as an anchor, building upon a commitment to the highest standards of quality for statistics and related services and providing the most up-to-date and accurate figures possible to give insight into important social phenomena and thereby contribute to government policy. These objectives need to be the basis for new endeavours, be they access to a new data source or setting up data hubs for collaborating with other organisations.

The principal commitment of NSOs to privacy and data protection is critical for their role in society, and strict application of privacy rules needs to remain the guiding principle when designing any new endeavours, no matter the urgency or pressure exerted by a given crisis.

Transparency is another component of this DNA. Procedures of NSOs need to remain completely transparent, so that everyone knows exactly which statistics they produce and how. Within a crisis context such as the COVID-19 pandemic, the urgency to move fast is high and developments can cascade very quickly, but it must always be traceable, at all times, which steps were taken, how were they taken and why. With increased access to new data sources in the wake of a crisis, it remains important to be fully transparent on which data sources are being used.

CBS: Open data to fuel collaborative data ecosystems

Being a leading partner in a collaborative data ecosystem also means that NSOs ensure their data adhere to principles that make such collaborative endeavours possible in the first place. The FAIR principles (Findable, Accessible, Interoperable, Reusable), mentioned earlier in this paper, are a key vehicle for ensuring this. As an example, CBS makes all its statistical output available free of charge as open data, a service that has seen a heightened demand during the crisis. The number of requests to its open data portal has been steadily rising, showing that the portal is relevant to not only policy-makers but society at large.

New tasks and agreements

In order to remain strong together in a crisis, NSOs and organisations must take a critical look at existing tasks and responsibilities to identify any necessary changes. This requires organisations and people to be flexible.

CBS: Task reorganisation

In the Netherlands at the start of the COVID-19 crisis, there was a major shortage of test material, which meant that not all deceased persons in the Netherlands were tested for COVID-19. As a result, the number of people that actually died from the coronavirus was expected to be higher than the number of coronavirus deaths that had been registered. CBS then accelerated publication of statistics on the causes of death and switched to weekly publication from quarterly publication. CBS employees whose work was affected during the pandemic were assigned to work on the "causes of death statistics", in order to publish these important figures. CBS figures showed that, in the first nine weeks of the COVID-19 pandemic, almost 9000 more people died than one would have expected during this period, whereas the number of officially recorded COVID-19 deaths in this period was 5400. This means that an additional 3600 people also likely died from the disease. These deceased persons were not tested for the virus and therefore remained outside the official COVID-19 death statistics.

In addition to shifts in existing work within the NSO, collaborations with other partners can also lead to changes. An example is the response of international agencies, led by the United Nations Statistics Division and the International Civil Aviation Organization, collaborating to provide policy-relevant, high-frequency indicators for international trade, air travel and maritime transport[18]. In both cases, good agreements must be created. When external parties are involved, the requirements are more far-reaching and must be formally laid down in a cooperation agreement.

Innovation

As stated earlier, modern NSOs are anything but isolated entities and operate in complex data ecosystems in order to adequately serve their societal function, which also holds true when it comes to innovation and developing advanced data collection capabilities. Open innovation (together with external users and data providers) built around an outside–in approach provides the opportunity to learn from partners, find synergies, combine resources and better understand user needs through direct engagement. For NSOs, this means taking the roles of partners, facilitators and integrators and combining their expertise with that of others to create fit-for-purpose statistics.

---

[18] Big Data UN Global Working Group. (2019). Dashboards on Trade, Travel and Transport. Retrieved from https://marketplace.officialstatistics.org/ttt-dashboards

Ireland: National Data Hub

Another example is provided by the Central Statistics Office of Ireland (CSO), which was able to create in one week a research microdata national data hub for policy and research purposes in key domains in response to the COVID-19 crisis: "Key datasets were identified, acquired and made available for research purposes in around a week based on the researcher access and administrative data governance systems of CSO"[19].

Statistics Canada: personal protective equipment (PPE) innovation

The PPE supply chain is a complicated ecosystem with many actors across jurisdictions and sectors. Many NSOs are taking novel approaches to responding to this need and leveraging lessons learned to create the infrastructure required to respond to any emerging needs. One such novel approach is the creation of data sharing platforms for ingestion, cleaning, and standardisation of disparate data sources. Models are then applied to the data to produce timely and relevant outputs on the state of the PPE supply chain nationally and for each jurisdiction.

Statistics Canada mobilised expertise in health modelling, data repositories and secure cloud computing to build a tool that brings together health and procurement agencies across government so they can make informed decisions about supply chains for PPE such as gloves, gowns, respirators and masks. The agency also implemented standardised coding and conducted regular surveys to provide near real-time demand and supply information to better manage PPE.

The agency recently moved PPE data from a "tools approach" to a "data platform" (data hub), increasing data value and setting the stage for future data-sharing opportunities. Scalability was incorporated in the model as well. Statistics Canada is currently leveraging existing people, processes and technology implemented for PPE data to rapidly scale up data and insights for testing and vaccines supplies.

---

[19] Economic Commission for Europe. (2020). Emerging data stewardship role of the Central Statistics Office of Ireland and its deployment in the COVID-19 crisis. Geneva: Conference of European Statisticians, Sixty-eighth plenary session

Australian Bureau of Statistics (ABS) – Partnerships

The ABS has explored administrative and transactions data from the public and private sector to inform official social and economic statistics during COVID-19[20]. The ABS is currently working to explore:

- Bank transactions data: Aggregated, de-identified transactions data from major banks to generate public value by informing official ABS estimates of business sales, household consumption and Gross Domestic Product (GDP), and assist in understanding the evolution of the Australian economy.
- Energy consumption data: To understand how electricity usage is changing for businesses and households at an area level, and to understand power as a consumption item. Early research is also being done on using electricity data to understand COVID-19 impacts[21].
- Mobility data: Samples of aggregated, de-identified data on mobility and location from a commercial provider based on mobile device GPS to assess the value of these types of data to provide information on short-term population movements.

This exploration also made use of publicly available device-based location data published by Apple and Google and pedestrian counter data published by the City of Melbourne. No data were sourced from the Australian Government's COVIDSafe app.

The partnership between the ABS and the Australian Taxation Office (ATO) was critical for providing these timely insights on the social and economic impacts of the COVID-19 pandemic across Australia and complement ABS's regular labour market statistics[22] [23].

UK Office for National Statistics (ONS) – Partnerships

1. The ONS Secure Research Service (SRS) team has been working with Health Data Research (HDR) UK to provide metadata for their Health Data Research Innovation Gateway[24]. Using a direct onboarding system, ONS supplies descriptive elements that allow researchers to understand their COVID-19–related datasets ahead of the accreditation or project application process. ONS provides metadata in a structured format, using standard conventions and controlled vocabularies, which results in a high degree of searchability and comparability. Applications made via the HDR UK Gateway are treated as high priority by the SRS team and subject to a double check by the SRS Statistical Support team to ensure they adhere to UK Statistics Authority (UKSA) guidance, that the research team is accredited for accessing data in the secure environment, that the research is feasible and that the public benefit of the research is clearly stated. After the checks by the SRS team are completed, the application moves to the review stage and then to the owners of all SRS-held datasets listed in the application for approval. The application is then submitted to the UKSA for submission to the Research Accreditation Panel (RAP) for approval or asking revisions[25].

---

[20] Australian Bureau of Statistics. (2021). COVID-19. Retrieved from https://www.abs.gov.au/covid-19

[21] Australian Bureau of Statistics. (2020). Using electricity data to understand COVID-19 impacts. Retrieved from https://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/4661.0Main%20Features12020?opendocument&tabname=Summary&prodno=4661.0&issue=2020&num=&view

[22] Australian Bureau of Statistics. (2021). Australian Bureau of National Statistics Home Dashboard. Retrieved from www.abs.gov.au/websitedbs/d3310114.nsf/home/ABS+Media+Statements+-+ATO-ABS+partnership+delivers+COVID+jobs+statistics.

[23] Australian Bureau of Statistics. (2020). New Australian labour market information. Retrieved from www.abs.gov.au/websitedbs/d3310114.nsf/home/ABS+Media+Statements+-+new+Australian+labour+market+information

[24] Health Data Research Innovation Gateway. (2021). Retrieved from https://web.www.healthdatagateway.org/search?search=&publisher=Ncs%20%3E%20Office%20For%20National%20Statistics&tab=Datasets

[25] Health Data Research UK. (2021). Retrieved from https://www.hdruk.ac.uk/

2. Early in the COVID-19 pandemic, it was apparent monitoring mortality alone was not sufficient for fully understanding the impact of the virus on society. The ONS collaborated with the UK's Department of Health and Social Care (DHSC) to estimate wider categories in which COVID-19 may have indirect impacts—not just deaths due to COVID-19 but the effects that reprioritising healthcare resources may have had on existing health services and the potential positive and negative health effects of government's non-pharmaceutical interventions, such as mandatory closing of businesses. The ONS, DHSC and other departments mobilised economists, modellers and other analysts to produce innovative methods to estimate these wider impacts swiftly and as robustly as allowed by evidence available at the time to ensure decision makers had the best evidence to inform the government's response to the pandemic. As more data on the impacts of COVID-19 became available, the ONS and DHSC expanded these analyses to include morbidity impacts and framed them around the latest mortality projections to update these estimates and again provide them to the Scientific Advisory Group for Emergencies (SAGE). The ONS recognised the importance of presenting clear and consistent analyses to the public too and supported each paper's release on SAGE's website with an ONS media statement outlining key findings and signposting to relevant statistics on the ONS website. In this way, the ONS collaborated proactively with ministerially led departments and senior experts to inform the UK's response to COVID-19 at pace while maintaining the transparency and quality of statistical outputs[26].

---

[26] https://www.ons.gov.uk/news/statementsandletters/ourfutureanalyticalworkprogrammeinresponsetoCOVID19au-tumn2020

Chapter 6. Communication and implementation

The previous chapters have described in detail our vision of the special role that national statistical organisations (NSOs) should take concerning data stewardship during, and outside of, crises like the COVID-19 pandemic. This section discusses how, in the face of the changing landscape of data collection, NSOs need to have effective strategies for a foundation of social license, for strategic and lasting engagement with the wider data ecosystem and partnerships and for an appropriate and solid foundation of institutional and technical starting conditions.

Whereas the previous chapters have laid out this vision, in this chapter, the focus is on recommendations for how it can be achieved and on providing concrete approaches and strategies for working towards this vision. This chapter will focus, in turn, on the products and services needed by NSOs, metadata, technical solutions and communication strategies.

Strategies on products and services

For NSOs to take a leading role as data stewards in times of crisis and beyond, they must enrich their portfolio of products and services accordingly. This section summarises a set of strategies and priorities for NSOs in order to fulfil their societal role, now and in the future. Naturally, there is no one-size-fits-all approach, and any strategy for developing the necessary products and services needs to take the specific context into account. Nevertheless, on a general level, the following strategies for developing products and services are recommended. These strategies are grouped by the four challenges presented at the beginning of this paper.

CHALLENGE 1. AN EVER-INCREASING PUBLIC NEED FOR HIGH QUALITY, RELEVANT, TIMELY AND SMART STATISTICAL INFORMATION

Strategy 1: Tailor statistical output for maximum relevance
Statistical output needs to be relevant and timely. In an environment of fast-paced technological developments of datafication, user expectations are evolving continuously, and in line with their experiences and the possibilities around them. NSOs need to measure the phenomena that are important to citizens in order to match society's demand for statistical information. This also is important for maintaining social license, as the legitimacy of the work of NSOs is largely determined by their output.

Strategy 2: Be transparent about necessary prioritisations between timeliness and quality of output
The potential trade-off between timeliness and level of quality needs to be addressed openly and transparently—in times of crisis, prioritisation may be necessary for faster availability of statistics, even though this might require foregoing some of the normal review steps before publication. The key strategy here is to be transparent about these choices, given that, in crisis situations, NSOs cannot necessarily rely on data that was obtained with the utmost rigour but must instead have multiple levels of acceptable data in mind. Providing disclaimers that state that the accuracy of emergency "flash" indicators is not the same as "regular" output releases.

Strategy 3: Leverage the FAIR data principles as cornerstones for statistical data and services
The FAIR principles (Findable, Accessible, Interoperable, Reusable) described earlier in this paper

constitute a central strategy in order to create a data landscape capable of combining data quickly and flexibly (for more info see, Go FAIR Foundation[27]).

Strategy 4: Develop a culture of continuous innovation and learning
NSOs need to constantly innovate in order to adapt to ever-changing demands. This is an institutional issue as much as a cultural issue: NSOs need to have a mindset of continuous improvement grounded in research and review of current processes and practices and must remain open to the potential of emerging solutions in their environment. This adaptation must be centred on an organisational architecture that is flexible, adaptable and responsive to a constantly evolving environment with new challenges and solutions. NSOs do not need to do this alone but should seek involvement of national and international networks to ensure continuous learning and innovation and use open innovation strategies built around an outside–in approach to learn from external partners, find synergies and combine resources and to better understand user needs through direct engagement. The UNECE Task Force on Data Stewardship, for instance, is a good example of collaborative network to learn from international perspective.

Strategy 5: Establish an integrated data strategy, including a data collection strategy
As this paper has explained, NSOs are confronted with a data landscape that is increasingly rich, but also highly diverse and complex, which requires fundamental shifts in their strategies and tools for data collection. This means that on an institutional level, NSOs need to have an integrated data strategy that considers the opportunities and challenges posed by this changed data landscape, and enables a flexible approach that focuses on phenomenon-based statistics and new data sources in addition to traditional topic-based primary data collection. A key aspect of such an integrated data strategy is the capability to reuse and combine interoperable and standardized data (e.g., DDI-CDI, DCAT, Schema.org, SDMX, Apache Arrow, Apache Parquet). Data governance in this strategy must be based on defined quality standards for data and metadata throughout the entire statistical process.

Strategy 6: Invest in innovative techniques for data ingestion
Foundational to the statistical process is the leveraging of new and novel technologies for ingestion, cleaning and classification of diverse and complex data. Examples of these techniques include techniques data collection through connection and data virtualisation, crowd-sourcing, web panels, web scraping, waste water testing. An earlier vision paper of Statistics Canada and Statistics Netherlands (CBS) on Advanced Data Collection has described these capabilities in detail[28].

Strategy 6: (Further) develop and expand the data scouting function (also referred to as data discovery)
The need for new data sources driven by the needs and challenges detailed in this paper require NSOs to further develop their data scouting function (i.e., a dedicated function for coordinating data ingestion). The identification of things that need to be measured and what data sources can be leveraged to measure these phenomena starts with an understanding of what data assets an

---

[27] GO FAIR is a bottom-up, stakeholder-driven and self-governed initiative that aims to implement the FAIR data principles. It offers an open and inclusive ecosystem for individuals, institutions and organisations working together through implementation networks (INs). See www.go-fair.org.
[28] Salemink, Dufour, van der Steen. (2019). Future Advanced Data Collection. Conference of European Statisticians: Workshop on Statistical Data Collection 'New Sources and New Technologies'. Geneva, Switzerland.

organisation holds—a robust data inventory or catalogue. Identifying data gaps which will further drive data scouting or data discovery, are directly related to the content of the catalogue or inventory. Data scouting also requires harmonised and coordinated processes in order to maximise efficiency and effectiveness, as well as ensuring that new data sources are acquired according to applicable rules and regulations. NSOs therefore need to invest in the required data scouting capacity, knowledge and strategy development. The strategy of developing a "Central Data Ingestion Office", as at CBS, is just one example of how to facilitate a standard way of ingesting data sources and make them available to statisticians.

CHALLENGE 3. A COMPLEX DATA ECOSYSTEM WITH HIGH INTERDEPENDENCY BETWEEN DIVERSE ACTORS

Strategy 7: Ensure that data is well-described and catalogued
As mentioned earlier in this chapter, NSOs need to have a good metadata-driven data catalogue that provides an encompassing, high quality overview of the data that is available and can be (re)used. This is also a relevant product with regard to transparent engagement and collaboration within the wider ecosystem of partners.

Strategy 8: Ensure a central role for NSOs in whole-of-government data strategies
Government-wide approaches leverage the full capacity of interorganisational collaboration for maximum public value. This rests on standardized approaches to data governance. A central leading role, such as a government chief data steward and chief data officer are key to developing and managing government-wide data standards and processes. NSOs are in a key position to support or provide these leading roles.

Strategy 9: Steward sharing platforms
Partly related to point ix, NSOs are in a key position to establish and facilitate platforms for broad data collaboratives and participate in data ecosystems. Complex phenomena can often not be addressed by governments alone and require co-creation and shared responsibilities with many different actors beyond the public sector. The creation of data sharing platforms and data trusts for ingestion, cleaning and standardization of disparate data sources are a relevant strategy in this regard, and NSOs are in a natural position to facilitate and lead such efforts.

CHALLENGE 4. THE CRUCIAL ROLE OF TRUST AND SOCIAL LICENSE FOR NSOS

Strategy 10: Ensure maximum transparency of processes
Social acceptance of the work and contribution of NSOs is driven not only by the relevance of their output but also by how it is produced. Clear and transparent communication, elaborated on further in this chapter, is, therefore, important for society to understand NSO data products, their relevance and why they need access to a wide array of data assets to produce trusted smart statistics.

Strategy 11: Invest in research to implement new techniques for privacy-preserving analytics
Innovation in the statistical and technological domains creates new possibilities for exchanging and analysing sensitive data in highly secure manners. Techniques such as multiparty computation, homomorphic encryption and synthetic data are domains that are highly relevant for the future statistical products and services of NSOs. Investing in privacy protection will ultimately have a beneficial effect on the trust of citizens, and therefore on public engagement and social licence.

Strategies on (meta)data – Metadata first!

Generally, outside times of crisis, many organisations have the guiding principle of "no data without metadata". However, when decisions need to be made, and made quickly, leveraging data assets already in the ecosystem, which may or may not include high quality metadata, can ensure that statistical output matches societal needs. This has been the reality for many NSOs in response to the COVID-19 pandemic; ensuring that data are representative of the population, and are of acceptable quality to measure given phenomena and help ensure timeliness of response, has been as important, or more important, than high quality metadata.

Applying the FAIR principles can offer solutions but also has an impact on how to deal with metadata. Below is an overview of the strategies, which are then discussed in detail:

1. **Findable**: How can NSOs make data "findable and searchable" and ensure users understand the functionality—both within and outside the NSO's organisation?

2. **Accessible**: Access to data is often via a data catalogue and then by searching in a metadata management solution. APIs must be made available for querying metadata.

3. **Interoperable**: Interoperability of data is supported by the use of international standards for southbound and northbound connections. The "European Interoperability Framework" (EIF)[29] is an example of what can be done to promote data interchangeability.

4. **Reusable**: Metadata include clear descriptions of, for example, where the data come from and how they were obtained as well as comply with or support domain-relevant standards.

Strategy 1: Be findable: Create a national data catalogue

For data to be more accessible, NSOs need to make the data more findable. This is by way of the metadata associated with the data, not by the data themselves.

It is, therefore, important to have an integrated overview of what data sources the government has, which could be shared in a secure and controlled environment on the basis of legitimate interest and legal foundation. Having a national data catalogue (NDC) can be of great importance during a crisis such as the COVID-19 pandemic. The catalogue would identify and index the various government sources and be available to all without restrictions (only a few exceptions, for classified and highly sensitive data sources, which would remain classified).

Here are some of the benefits of an NDC:

- **Transparency and accountability about what data the government has**
  An NDC can answer questions about, for example, who offers what data for which purpose and for which target group.
- **Knowledge sharing**
  Data sources are uniquely identified, including information on the responsible party and version; therefore, the catalogue also functions as a "referral index", ensuring that interested parties can contact the source holder to gain access.

---

[29] The European Interoperability Framework, part of communication COM(2017)134 from the European Commission, was adopted on 23 March 2017. The framework gives specific guidance on how to set up interoperable digital public services.

- **Efficient government**
  The once-only principle is an e-government concept designed to ensure that citizens, institutions and businesses are required to provide some standard information to authorities and government services only once. An NDC allows authorities to know what data are already available, at what level of quality and from whom. This promotes the reuse of data and prevents duplicate requests, leading to a reduced administrative burden.
- **Supporting data-driven work**
  For innovative data-driven policy (based on AI, data scouting for statistics purposes, etc.) a lot of data is necessary. This requires insight into all existing data and the source of the authentic data (system of record). Information about the quality (e.g., random or integral) and timeliness can be essential.

In addition, providing a harmonisation canvas is important because, to understand the data, it is important to define the concepts (with glossaries, ontologies, etc.). There are different levels of "understanding", ranging from technical interoperability to semantic and organisational "understanding" of data. These multi-level approach is also reflected in the EIF, mentioned above.

Strategy 2: Be accessible: Enable machine-to-machine interaction

In addition to human consultation and searching of data sources and datasets, machine-to-machine (M2M) interaction will be enabled by supporting international metadata standards and the FAIR principles. This requires the development of an extensive API that allows the metadata of data sources and datasets to be entered, modified, searched, consulted and archived. In an API strategy, the API must follow industry standards and meet the requirements set by a national chief data officer so that the metadata information shows whether integration is feasible and to avoid duplication as much as possible. Of course, this will be accompanied by an adequate authentication and authorization system pointed to in the metadata, including information on supported standards and contact information for obtaining access. For the benefit of social license, a culture change in the area of making data available is important. Where possible, moving from a "no, unless..." to a "yes, provided that..." attitude is recommended. The provision of open data can also help improve transparency.

Strategy 3: Be interoperable: Make data interchangeable (by adhering to standards)

Are consistent standards used or do NSOs need rules and controls that support the transition from one format to another? These are some of the questions that come into play when analysing data interoperability. When exchanging data, the infrastructural and technical support of transitions is important (even if the use of standards is actively promoted). Although it is not strictly necessary to support standards within systems, it is advisable to support standards throughout the lifecycle in order to avoid possible conversions. On the input side (southbound) and output side (northbound), standards and the transitions between formats/models are extremely important. Only then can data exchange, in times of stress, run smoothly and quickly. It should be noted, however, that repeated conversions can lead to data loss issues, especially when new datasets are used.

Strategy 4: Be reusable: Provide a data glossary

For research purposes, linking data sources and datasets is necessary to add value. In addition to having link IDs, it is important to know what certain variables mean (semantics). Information about the quality is also needed in order to know whether linking is possible at all. Given these concerns, it is important to ask questions like: What is the underlying population, what is the coverage, how representative are the data, are the same definitions being used? Differences can be acceptable as long as they are made explicit so that acceptable solutions can be found.

A data glossary to define important concepts is a valuable too. On the basis of the description of the concepts used in the data glossary, it is possible to assess whether linking the data can be carried out while maintaining the quality of the data. For example, during COVID-19, different countries reported in very different ways. Some countries reported only those deaths that occurred in hospitals or deaths after a positive COVID-19 test. Others reported all deaths where COVID-19 was even a possible cause, and some also reported on excess deaths in the population. Conceptual and semantic metadata and standardization can be useful in arriving at a certain level of understanding so that communication is unambiguous in times of crisis.

Strategies on technology

Technology is an important enabler for many requirements identified in this paper. Timely and reliable access to trusted data, faster time to market, data-driven work, more collaboration between various organisations and improved dissemination and communication are all affected by our technological maturity. This section provides an overview of technologies that play an important role in crisis scenarios and some proven examples of strategies for how to approach their implementation—both before and during the time of crisis.

TECHNOLOGY: A KEY ENABLER FOR STATISTICS IN CRISIS SCENARIOS

In its Top 10 Data and Analytics Trends in 2021[30], Gartner reported that "…these data and analytics trends can help organizations and society deal with disruptive change, radical uncertainty and the opportunities they bring." The pandemic has of course been a major driver for technology innovation, and NSOs can expect this trend to continue. The next paragraphs map some of the trends to phases of statistical process described by the generic statistical business process model (GSBPM)[31] and discuss the impact of new requirements and technology trends on the four "operational phases" of collection, processing, analysis and dissemination.

---

[30] https://www.gartner.com/smarterwithgartner/gartner-top-10-data-and-analytics-trends-for-2021/
[31] The GSBPM comprises eight phases: specify needs, design, build, collect, process, analyse, disseminate and evaluate. See https://statswiki.unece.org/display/GSBPM/Generic+Statistical+Business+Process+Model.

Collect and connect: Advanced data acquisition techniques and technologies

Increased demand for timely access to new data sources requires new data collection channels, for example, crowdsourcing platforms, mobile applications and data streaming. Crisis response often requires new types of secondary data, such as sensors, medical data and geospatial data. These data sources are often distributed, too sensitive or simply too big for direct data ingestion but can be accessed using "connect mode"[32]. Shifting data and analytics to the "edge"[33] will create opportunities for new data ecosystems, also in line with the "data minimisation" requirements of privacy regulations such as the General Data Protection Regulation (GDPR).

Process and integrate

Crisis scenarios require fast insights on unpredicted events and a phenomena-based approach, which require even more data integration than in standard statistical production. A siloed approach to data management and processing cannot support this. Technologies such as data warehouses and data lakes can be combined with data integration tools and data virtualization techniques to create a "data fabric"[34] architecture which can significantly reduce the time required for data integration and remove data siloes.

Analyse and discover

Variability of new data sources requires advanced analytical methods that have to be developed and regularly tested by data scientists and data engineers. Data science technology is now needed not only in the data discovery phase or ad hoc analysis but also in regular statistical production. Data science and advanced analytics platforms can be used to automate analytics tasks, for example, regularly updating machine learning models. The idea of "composable data and analytics" architecture that enables reuse of components from multiple data, analytics and AI solutions can increase flexibility and encourage collaboration on methods and algorithms to speed up development of these advanced methods.

Disseminate and interact

In times of crisis, the traditional "user needs assessment cycle" often is not fast enough. Users of statistical information (policy-makers, researchers, journalists, the general public) have to respond much more quickly and they often struggle to fact check information or even have difficulty finding information on official statistics websites and open data portals in the first place. The COVID-19 cri-

---

[32] Access and connection to sources APIs independently from their location (local/remote/cloud environments); see https://statswiki.unece.org/display/DA/CSDA+2.0+-+X.+Capabilities
[33] A distributed computing paradigm that brings computation and data storage closer to the location where it is needed. See https://www.cloudwards.net/what-is-edge-computing
[34] Data Fabric Architecture is Key to Modernizing Data Management and Integration. See https://www.gartner.com/smarterwithgartner/data-fabric-architecture-is-key-to-modernizing-data-management-and-integration/

sis revealed (and revived the focus on) the need for near real-time information on public dashboards and data visualisation. In the future, technologies such as voice assistance and chatbots could help in the fight against false information that often spreads quickly in a crisis.

Cross-cutting capabilities

All phases of the statistical process are supported by capabilities such as metadata management, security and privacy-preservation controls.

Metadata management technologies play an important role in discovery of data sources, automation of processing (metadata driven) and understanding of statistics. Data catalogues, automated metadata harvesting and semantic and graph technologies play an important role as well.

Security and privacy protection will remain major focus points for protecting data and keeping public trust high. In emergency situations, trust is an intangible yet very important asset that enables NSOs to work with respondents and partners to acquire new data and produce important insights.

The 2018 GDPR and the COVID-19 crisis created increased demand for new privacy protection techniques to support privacy by design and to protect confidentiality. Privacy-preserving computation technologies such as homomorphic encryption, multiparty computation, trusted execution environments, distributed learning and new output privacy techniques are offering new opportunities for providing insights from very sensitive data while assuring strong protection of privacy. There are also developments in technology that provide additional trust to citizens (for example, transparency and self-sovereign identity, or SSI)[35].

The impact goes two ways: while technology providers offer new technical capabilities such as advanced AI systems, statistical organisations can help governments set important standard baselines for things like ethical use of AI.

STRATEGIES FOR INNOVATION AND MODERNISATION (GSBPM: SPECIFY **N**EEDS, **D**ESIGN, **B**UILD, **E**VALUATE)

In a crisis, NSOs not only face technology gaps but also need to understand other key enablers so that innovative technological solutions can be designed, implemented and effectively used. These enablers include partnerships, processes, policies, and most importantly: people with the right skills.

The implementation of new technologies should be supported by strategies that can address key challenges blocking statistical organisations from moving from a reactive to a proactive position. For example:

- How to mobilize sufficient resources and skills;
- How to speed up time to market for technological implementations;
- How to productionise methods and move from R&D to production;
- How to sustain the pace of innovation and keep technology up to date.

---

[35] With self-sovereign identity, the user has the ability to create and manage unique identifiers and store identity data. In a self-sovereign identity approach, users can create and manage their own identities without relying on a central authority.

This paper does not seek to provide an answer to all challenges; however, experience gained before and during the COVID-19 crisis can be used to identify some successful strategies and prepare NSOs for a possible future crisis.

Strategy 1: Pursue partnerships to create new capabilities

Statistical organisations alone will never have sufficient resources to design and implement new methodological and technological capabilities. Partnerships with other public organisations, the private sector and academia can be used to overcome this gap. The data revolution triggered the need for collaboration not only to develop new technology but also to partner on data acquisition, methodology development and even data dissemination. One example of this approach is collaboration on the UN Global Data Platform[36], which helped deliver COVID-19-related data and technology for global monitoring via the United Nation's Dashboards on Trade, Travel and Transport.[37]

Strategy 2: Scale up existing capabilities: Cloud and beyond

Even when technical capabilities are developed, a crisis often requires significant scaling up or repurposing of capabilities to fit new priorities. Modern hybrid cloud environments can provide flexibility to quickly scale up storage, compute for additional data or users and add new capabilities such as dashboards or high-performance processing. An example of a successful response to completely new needs for data sources and analytical environment was the implementation of the ONS COVID-19 hot spot Joint Biosecurity Centre (JBC) platform. The national statistical office quickly deployed a cloud environment for a new JBC and provided analytical capabilities needed to inform local and national decision-making in response to COVID-19 outbreaks.

Strategy 3: Agile technology innovation: XOps (DevOps, DataOps, MLOps)[38]

The COVID-19 crisis gave us good insight into technologies needed for pandemic response; however, the future will most likely bring a new, perhaps completely different, crisis. Organisational agility and innovation will be critical to responding quickly. Agility—initially used in software development and that resulted in short iterative development cycles and value creation—is now expanding to new areas, with DevOps principles being reused and adapted for DataOps, MLOps, etc. These methods can help address the gap between experimental innovation and stable production environments and speed up not only delivery of new technology but also new data products. Further, it can be scaled up to make end-to-end statistical processes much more agile. An example of

---

[36] UN Big Data. (2021). Retrieved from https://unstats.un.org/bigdata/un-global-platform.cshtml.
[37] Big Data UN Global Working Group. (2019). Dashboards on Trade, Travel and Transport. Retrieved from https://marketplace.officialstatistics.org/ttt-dashboards
[38] Panetta, K. (2021). Gartner Top 10 Data and Analytics Trends for 2021. Retrieved from www.gartner.com/smarterwithgartner/gartner-top-10-data-and-analytics-trends-for-2021/.

this approach is the OECD's smart statistics[39] approach, where research-focused data science environments integrate with production-focused statistical environments via an algorithm bank and sandboxing to deliver complete and integrated data lifecycle.

Strategy 4: Invest in skills and open source communities

Technology is changing at a faster pace than ever before. How can statistical organisations (and their partners) sustain this pace and remain innovative? Investment in skills is of course most important way, but organisations cannot rely solely on the skills in their organisations—they need to create new communities and connect to wider communities of practice. Examples are international statistical communities such as the HLG-MOS community on machine learning in official statistics[40] and the open-source community on privacy-preserving techniques OpenMined.org[41]. These communities can create active networks that will maintain innovation momentum, will help define standards and reference architectures[42] and can be mobilized in times of crisis not only for exchange of experience between participating organisations but also to help collectively develop new techniques and technological solutions.

Strategies on communication

The most important factor of this section is the maintenance of trust and social license. In other words: how to act in a changing situation, what actions to take, what results to achieve and how to communicate to the public and stakeholders in order to maintain trust and social license.

Communication is not only directed externally. In a crisis situation, the first thing to do is make sure your own organisation is functioning, thus ensuring that delivery of the required output, continuing to cooperate with stakeholders and partners and, where relevant, taking on extra work. The motto is to keep the people calm, create trust, take stock of the organisation's needs and provide for them as much as possible. Showing the organisation's staff the way forward allows them to do the same for others. Accordingly, implementation guidelines for communication are:

Strategy 1: Create calm and be realistic
Regularly state as clearly as possible what is going on. Explain concretely and explicitly what is being done to improve the situation. This gives people something to hold on to. Do not give them the feeling that they are standing alone.

Strategy 2: Be open about dilemmas
It is important to remain open about dilemmas if they arise. What considerations are made, what choices are made and what are the consequences of these choices?

---

[39] Organisation for Economic Cooperation and Development. (2018). Which Strategies for NSOs in the Digital Era? Towards 'Smart Data' Strategies. Statistics and Data Directorate Committee on Statistics and Statistical Policy. Retrieved from
https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/CSSP(2018)7&docLanguage=En
[40] Project wiki page: https://statswiki.unece.org/display/ML/Machine+Learning+for+Official+Statistics+Home
[41] Open-mined community: https://www.openmined.org/
[42] An example is the Common Statistical Data Architecture (CSDA), commissioned by the High-level Group for the Modernisation of Official Statistics in 2018.See https://statswiki.unece.org/display/DA.

Strategy 3: Empathise on an emotional level
Connect with, and attempt to share, the perception of employees. Use personal stories to show that everyone is affected and communicate that you are diligently mindful of the position of the citizenry. Make sure to use positive words, show that you are listening and that you are adequately informed. Put into words what is going on, without being negative or presenting it as catastrophic.

Strategy 4: Reinforce the "we" feeling
Show what everyone is doing it *together*. Emphasise cooperation, partnership and togetherness. Strengthen common pride, say what has already been achieved together, and what is planned to be achieved together.

Strategy 5: Extra tasks
Call on staff to do something extra where necessary and then celebrate. Put them in the limelight when they succeed.

Strategy 6: Focus on specific colleagues who need more attention
Take stock of who needs extra attention and provide it where possible.

Strategy 7: Widen the focus (over time)
Make even more explicit what is going on, what people are struggling with, what is being worked on, what is important, what has been achieved together. Provide more in-depth information where necessary.

In addition to internal communication, it is also important to communicate externally.[43,44,45] This may involve stakeholders, cooperation partners or the general public. Depending on the desired target group, the emphasis and depth of the communication will shift. A story for a group of enthusiastic scientists will take a completely different form from a story for a group of critics. A story for the entire population will be shaped differently still. However, despite these differences, it is possible to identify global and common goals or best practices:

Strategy 8: Adapt communication to the target group
As explained above, adapt communications to the target group, both in form, language, content and choice of media. Make a deliberate decision about which communication channels to use.

Strategy 9: Put the emphasis on improvement
Indicate what is to be achieved, why it is necessary, who or what is needed to do so and, if possible, when it is expected to be achieved. If necessary, make connections with other projects. Talk about (preliminary) results. Be transparent in what is being done.

Strategy 10: Emphasise trust and safety
Demonstrate being a leader and ensure trust is gained. Ensure ideas, plans and results are well founded and result in faith and confidence.

---

[43] Handreiking communicatieaanpak coronavirus, Nationaal Kernteam Crisiscommunicatie, Rijksoverheid, 23 april 2020 (in Dutch)
[44] Handreiking communicatieaanpak coronavirus, Nationaal Kernteam Crisiscommunicatie, Rijksoverheid, 25 juni 2020 (in Dutch)
[45] Handreiking communicatieaanpak coronavirus, Nationaal Kernteam Crisiscommunicatie, Rijksoverheid, maart 2021 (in Dutch)

Strategy 11: Be aware of the consequences
Remember that communication contributes to behavioural change.

Strategy 12: Combine
Communicate with partners jointly in order to maximize transparency and efficiency.

Strategy 13: Choose the right words
Use words that actively contribute to solving the problem and inspire the right response.

Strategy 14: Monitor
Monitor reactions to communications, including on social media.

Statistics Netherlands corporate news[46]

Statistics Netherlands (CBS) normally releases mortality data every two weeks. The COVID-19 crisis has created a strong need within society for more up-to-date figures on mortality. This has prompted Statistics Netherlands to supplement the existing biweekly counts with estimates and more accompanying details, allowing the organisation to produce reliable death counts over the previous week before all data were available. CBS has now adapted this approach to be able to present the previous weeks' figures further in advance.

CBS works closely with the National Institute for Public Health and Environment, which registers all deaths confirmed as COVID-19 deaths. The number deviates from the number of excess deaths because CBS reports total number of deaths including for those who were not tested but who may have died from COVID-19.

The importance of context-tailored strategies

This chapter has provided a range of recommendations for approaches and strategies to support the vision that NSOs should be taking a leading role in data stewardship during, and beyond, times of crisis. Recommendations have been provided for products and services of NSOs, metadata, technical solutions and communication strategies.

Regarding *products and services*, NSOs need to ensure that their output addresses society's need for high quality, relevant, timely and smart statistical information in times of crisis. Within a changed landscape of data collection that is as rich as it is diverse and complex, an integrated strategy for innovation and data ingestion is essential. This landscape is marked by a complex data ecosystem with high interdependency, which requires whole-of-government approaches and services targeted at facilitating collaboration beyond just the public sector. NSOs need to provide services targeted at maximum transparency and protect privacy and confidentiality in the process.

Regarding *metadata*, the FAIR principles provide a suitable strategic framework. To ensure that data is searchable, findable and understandable, data catalogues are a key tool and function as portals for accessing data. Accessibility can also be enhanced by metadata management solutions

---

[46] Vos, S. (2020). How CBS reports on mortality during the coronavirus crisis. *CBS Corporate News*. Retrieved from www.cbs.nl/en-gb/corporate/2020/16/how-cbs-reports-on-mortality-during-the-coronavirus-crisis.

and APIs to facilitate machine-to-machine interaction. Moreover, it is important to embed all these solutions as much as possible into accepted international standards in order to ensure maximum interoperability. Finally, regarding reusability, it is essential to accurately describe the data with process, quality and conceptual metadata, for instance by means of a data glossary.

Technology plays an important role in crisis scenarios, and this chapter has explored key technological trends and their impact on the statistical process and has looked at non-functional aspects of technology such as the development process, flexibility, scalability, time to market and skills needed by NSOs to respond to crisis situations. Key strategies to achieve this are designed to pursue partnerships for creating new capabilities, scale up existing capabilities such as the cloud, invest in agile technology innovation (for instance with XOps approaches) and invest in skills and open source communities.

Finally, even when all of the above areas are adequately addressed, *communication* is essential in order to ensure that statistical products, services, data, metadata and technological solutions reach the intended stakeholders, and in a suitable manner. It is important not only that NSO stakeholders are external as well as internal, but also that communication starts internally within NSOs in order to keep the organisation functional during a crisis. Externally, communication strategies need to be targeted at statistics stakeholders, NSO cooperation partners and the general public. For both internal and external communication strategies must be tailored to the audience and the emphasis and depth of NSO communications must be adapted to the target group.

This principle of no one-size-fits-all approach can be generically applied across all situations and holds for all strategies provided in this paper: each crisis, and each NSO's situation, is different, and any strategy needs to be tailored to the context of a given crisis. However, whatever a crisis may look like, NSOs need to take a central role in helping overcome it, and the strategies laid out in this chapter can provide them with a number of tools and solutions to take this leading role for society.

Chapter 7. Conclusion

The COVID-19 pandemic has dramatically highlighted the importance of data and evidence-driven policy responses, and revealed both the strengths and weaknesses of national and international systems to harness the potential of data and statistical information to fight the crisis.

It has become clear that good data governance and data stewardship are the most important pre-requisites for addressing the extreme demands societies face in crises such as the COVID-19 pandemic. In this regard, the pandemic also demonstrated that national statistical organisations (NSOs) can, and should, have an essential role to play in such crisis response. They are in a unique position because of their special mandate and responsibility for providing the necessary evidence base and their set of specialized capabilities making them well positioned to fill this role.

This paper, therefore, aimed to learn from this recent experience by taking a cross-national view on the experiences of NSOs during the COVID-19 crisis to distil the lessons learned about the role that NSOs can have for national crisis response in general and to explore various strategies for effectively fulfilling this role.

The key argument of this paper is that NSOs, as partners in a data ecosystem in charge of co-creating trusted smart statistics, are in a unique position to take a leading role for data stewardship and prepare societies for effective crisis response.

The paper discussed the key factors and challenges facing NSOs, the background against which the actions they need to take to innovate in fulfilling their role were discussed. As the COVID-19 pandemic has demonstrated, there is an urgent public need for high quality, relevant and smart statistical information in order to design evidence-based policy responses. At the same time, NSOs need to navigate an ever-evolving data landscape that is increasingly rich but also diverse and marked by a complex data ecosystem with high interdependency between diverse actors. Strong attention to social license constitutes an important cornerstone and further prerequisite of this playing field.

The issue of social license is an important one for NSOs. Not only does it enable them to perform their function in the first place, but it also serves as an important reminder to NSOs that they cannot assume such a leading role by default. NSOs need to "earn" this role by having and growing the public trust needed for the task at hand. Output legitimacy is key in this regard. In addition to their mandate, an NSO's social license is essentially defined by the relevance and trustworthiness of their output. At the core of this are the principles of data protection, confidentiality and privacy. The FAIR data principles (Findable, Accessible, Interoperable, Reusable), data quality standards and transparent communication about statistical output are additional strategies to this end.

NSOs also have the necessary "input legitimacy" that makes them especially well-suited for a central role on data stewardship in crisis response. They already have a public mandate to steward national statistical systems and leverage data for measuring sociodemographic and socioeconomic phenomena. As a result, they also have the data foundations and data stewardship expertise to effectively manage this data. The pandemic has shown in numerous examples (some of which were illustrated in this paper) how NSOs have helped to ensure that the health, infrastructure and economic needs of citizens and businesses were met during the pandemic, by means of data collection, data standardization, data hosting and data output.

However, being ready to assume such a leading role in future crises is not guaranteed, but depends on flexibility, combined with careful planning ahead of time, and proper implementation strategies to have the necessary capabilities in place in order to be prepared. This paper has enumerated many of them.

For statistical data ecosystems to be agile and resilient in order to swiftly and effectively respond to crises, certain starting conditions need to be met. First and foremost, a well-functioning and flexible data governance organisation is needed, defining clear authorities, roles and responsibilities and legal authority, such as the potential role of a senior authority on data such as a chief data officer. A related requirement is to have a strong, whole-of-government (and beyond) data strategy and framework in place, including the necessary standards to enable organisational flexibility and interoperability between partners.

Investing in these partnerships is a vital catalyst to quickly mobilise and leverage the necessary knowledge and (data) resources when a crisis emerges. This goes beyond merely organisational structure, and requires a cultural shift towards co-creation within this ecosystem. Moreover, investing in a culture of innovation is a key ingredient to be future-proof, but this requires critical self-reflexivity, inspired by looking beyond the borders of one's own organisation and being open to ideas from the outside.

In the previous chapter, this paper provided a number of recommended strategies for how this can be achieved by providing concrete approaches. These strategies span the areas of statistical products and services of NSOs, metadata, technical solutions, and communication strategies.

Together, they provide the basis for customizing a toolkit that can, and needs to be, tailored to individual country contexts and help NSOs to provide the much-needed data stewardship role which they are uniquely suited for. This will ultimately help their societies overcome not only the current pandemic, but also future crisis situations. NSOs must remain vigilant and strive for relevancy by constantly adapting to the needs of society and the ever-evolving data ecosystem—this will ensure that the data foundations and expertise of NSOs can be leveraged to respond to any evolving need such as health crises, natural disasters or economic shifts.