

Statistical Commission
Fifty-third session
1–4 March 2022
Item 3(w) of the provisional agenda
Items for discussion and decision: Open data

Background document
Available in English only

Local-Level Statistics as Open Data: A User-centric Approach

Prepared by the Working Group on Open Data

Local-Level Statistics as Open Data: A User-centric Approach

[Workstream 3: Local Level Open Data workstream]

Background

The United Nations Statistical Commission's Open Data Working Group provided a background paper on *Local-level statistics as Open Data*¹ for the Commission in March 2020. The paper described the benefits of releasing more local-level statistics as open data and laid out the basic requirements regarding geographies, content, confidentiality, and visualization. The paper highlighted National Statistical Offices can play an important role as trusted providers of local-level statistics by making them part of the open data assets of a country. However, the paper did not show concrete examples of how local-level statistics are being used and combined with other open data.

The Working Group can now present such examples collected from members of the group and through outreach after a side event during the 52nd Session of the Commission on *The open data revolution: the power of open data for engaging the public with statistics that matter to them* in February 2021². The side event also helped the group identify a number of challenges.

National Statistical Offices can increase the use, reuse, and value from open data by strengthened data governance and stewardship and by engaging with users. Through examples from different countries, it is demonstrated how a user-centric approach helps Offices publish with purpose and make local-level statistics a valuable part of the open data ecosystem of a country in line with how the Open Data Policy Lab³ describes the third wave of open data.⁴

The country examples are categorized into five use cases given in *Table 1* with more detail provided in the subsections that follow.

Table 1: Country use cases of making local-level statistics a valuable part of the open data ecosystem

<i>Use Case Category</i>	<i>Country examples</i>	<i>User audience</i>	<i>Is geospatial included?</i>	<i>Disaggregation level</i>	<i>Statistical type</i>
National data portals with user stories	Affordability (New Zealand)	Citizens	Yes	Suburb	Person and household level data
	Smart Property (New Zealand)	Citizens	Yes	Property	Property data
	Regional Images (Netherlands)	Citizens	Yes	Municipality	Social environment, healthcare
	My Local Stats Publication (Malaysia)	All levels of society and policy makers	n.a.	State & District	Person, household level

¹ https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3v-Local-level_OD-E.pdf

² <https://www.unescap.org/events/2021/open-data-revolution-power-open-data-engaging-public-statistics-matter-them>

³ The Open Data Policy Lab is an initiative of The GovLab, a Governance Lab based at NYU's Tandon School of Engineering, seeking to improve people's lives by changing how we govern. For more information see <https://opendatapolicylab.org/team/>

⁴ <https://opendatapolicylab.org/third-wave-of-open-data/>

<i>Use Case Category</i>	<i>Country examples</i>	<i>User audience</i>	<i>Is geospatial included?</i>	<i>Disaggregation level</i>	<i>Statistical type</i>
					data and establishment
	The Agricultural Analytical System (Malaysia)	All levels of society and policy makers	n.a.	State	Palm oil, rubber, poultry and fisheries
Smart reuse through dashboards and apps	Open Data Portal (Sierra Leone)	All levels of society and policy makers	Yes	District	Household level Data
	Netherlands	Policy makers, citizens	Yes	Municipality	Person / employee
	Smartphone Applications (Malaysia)	All levels of society and policy makers	Yes	State & District	Person, household level data and establishment
	Low Income Households (Netherlands)	Citizens	Yes	Municipality	Household level data
	Indigenous Story Maps (Australia)	Citizens	Yes	Indigenous regions	Person
Developing services by engaging with end users	COVID-19 response (Sierra Leone)	All levels of society and policy makers	Yes	District	Household level data
	GovHack (New Zealand)	Government, citizens and industry	n.a.	n.a.	n.a.
Hackathons and communities for developers	Open Govt Ninjas (New Zealand)	Developers	n.a.	n.a.	n.a.
	Virtual Datathon (Malaysia)	Higher Learning Institutions in Malaysia and the public	n.a.	n.a.	n.a.
	Netherlands - annual open data user meeting	Citizens, journalists, policy makers	n.a.	n.a.	n.a.
	LocalFocus (Netherlands)	Citizens / journalists	Yes	Municipality	General statistics
Working with media	Data Journalism (Sweden)	Citizens	No	Municipality	

Use Case 1: National Data Portals with User Stories

In countries where a National Data Portal exists and includes a user community, the National Statistics Office can encourage users of their open data to provide user stories or the Office may provide such examples.

Statistics New Zealand strives to provide relevant, real-world stories about innovative projects that use open data to promote the benefits of releasing open data and inspire new uses of available data. These are published on an ad hoc basis on data.govt.nz and are discovered through networking and interaction with open data groups and local initiatives. There are some interesting examples with a focus on the local level: “Affordability in NZ”⁵, “Smart property decisions”⁶ and “Where do your rates go?”⁷.

In the Netherlands, the Dutch government have a data register where you can find a number of *Impact stories*⁸ using data from the register, such as the use of energy transition data as a tool for achieving climate change goals. In addition, there is a list of helpful applications⁹ including “Regiobeeld”¹⁰ (Regional images) which provides insight into the state of affairs and future developments regarding healthcare and welfare by region. The website is part of the “right care in the right place” initiative and uses data from Statistics Netherlands (Centraal Bureau voor de Statistiek – CBS) among other sources.

Use Case 2: Smart reuse through Dashboards and Apps

National Statistics Offices can respond to emerging user needs by setting up dashboards, map visualizations and applications enabling the easy reuse of open data from the Office and other agencies in the wider data system. This also contributes to the Office’s leadership role in a National Statistical System. The Open Data Sierra Leone Portal¹¹ is one such example; an open data management platform for the Government to communicate with the public and get feedback from citizens on Government activities (Figure 1).

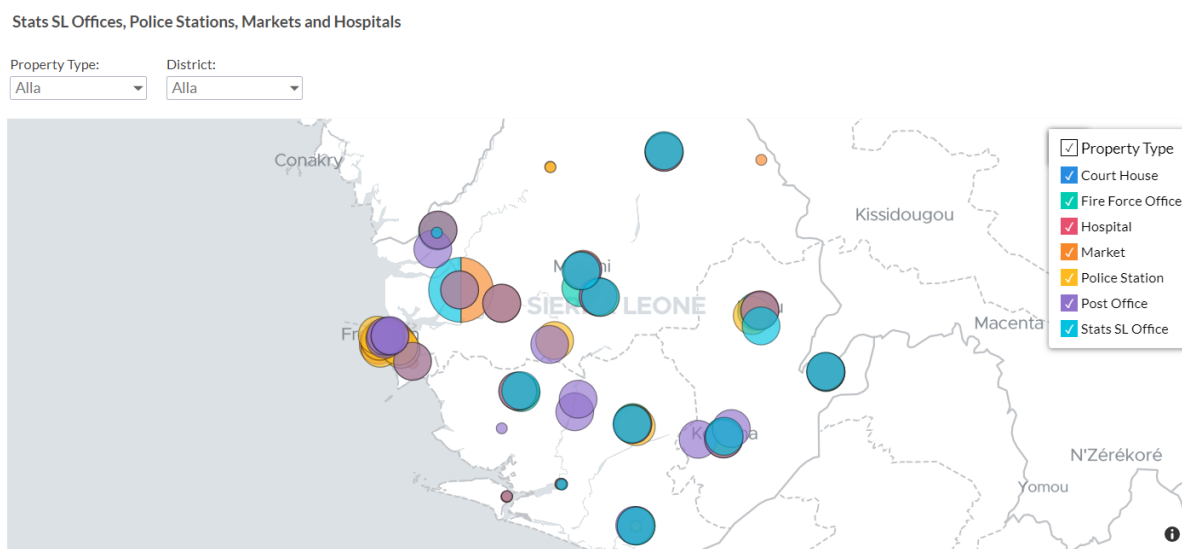


Figure 1: Example from Sierra Leone Dashboard

⁵ <https://www.data.govt.nz/use-data/showcase/affordability/>

⁶ <https://www.data.govt.nz/use-data/showcase/homes-co-nz/>

⁷ <https://data.govt.nz/use-data/showcase/where-do-your-rates-go/>

⁸ <https://data.overheid.nl/impact-0>

⁹ <https://data.overheid.nl/community/toepassingen>

¹⁰ <https://www.regiobeeld.nl/> (only in Dutch)

¹¹ <https://www.opendatasl.gov.sl/stats-sl>

Statistics Sierra Leone also provides a range of metrics measuring their performance and the Sustainable Development Goal indicators.

Statistics Netherlands, in collaboration with the University of Oxford, has developed a new visualization method to gain insight into mobility, called a “doughnut map”. This method has been applied to data on employee jobs, available as open data at CBS StatLine. This data does not contain information about the actual commutes, such as the commuting route, mode of transport or the frequency of commuting, instead this visualisation shows how connected the Netherlands is. For example, it shows how many people are prepared to travel a significant distance between municipalities for work and which cities attract most employees from far away. While it is difficult to convey the connectedness between municipalities from a table, the doughnut map illustrates this connectedness.

Visualization methods like the doughnut map (figure 2) are handy tools for policy-makers, enabling them to analyze and improve the infrastructure between municipalities. The doughnut map can, for instance, help explain where people who are stuck in certain traffic jams come from and go to and show possible impacts of new bicycle or train connections. The data can be further enriched with properties of the companies where people are employed. Furthermore, data from several years can be analyzed, which enables research on the effect of new infrastructure, new residential areas, and new business parks.

Policy objectives relating to climate change mitigation, wellbeing, accessibility, and air pollution demand that data should not only be accessible to data specialists, but to a wide range of stakeholders. Visualizations such as the doughnut map help to make data on complex flows insightful for a wide audience.

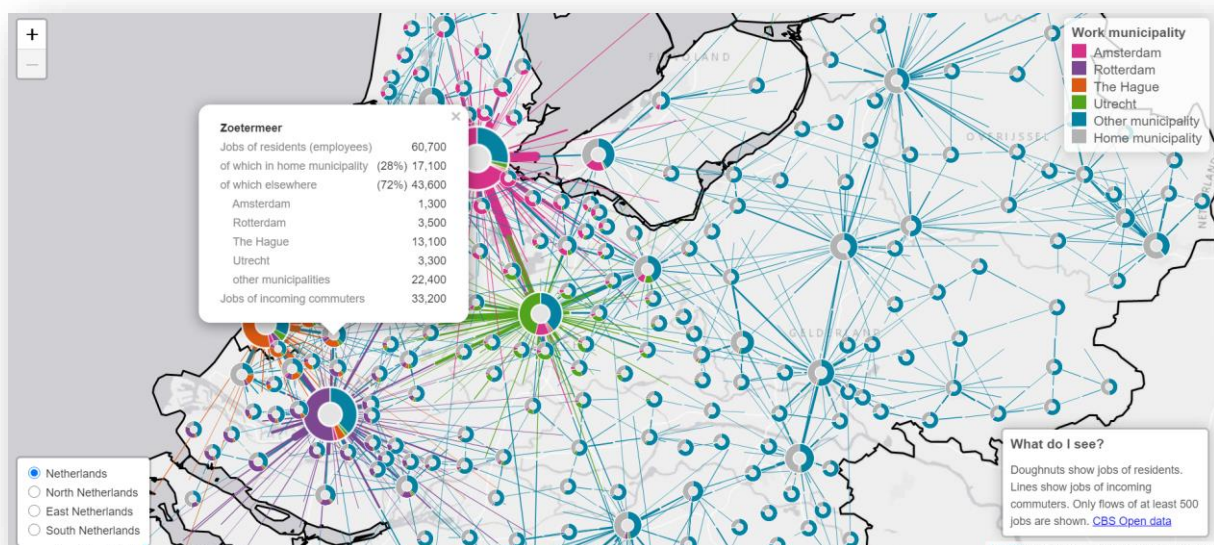


Figure 2: Example from Statistics Netherlands

The Department of Statistics Malaysia provide statistics at the district level in infographic form allowing easy comprehension by all levels of society. The information is available through a mobile application interactive system known as MyLocal Stats¹². MyLocal Stats assists users to access statistics at the district and local levels for the planning of local development. MyLocal Stats 1.0

¹² https://www.dosm.gov.my/v1/index.php?r=column/cthree&menu_id=V2o0d240dzI0Z2IvWkVhei90L3k3dz09

provides the latest statistics on the number of establishments by district and key sectors, obtained from the Economic Census. In addition, data on birth and death as well as population estimation by district are included.

Malaysia also has the Agricultural Analytical System, an initiative of the Department of Statistics Malaysia under the Public Sector Data Analysis 2.0 project. This platform displays agriculture statistics that focus on four products: palm oil, rubber, poultry (chicken and duck) and fisheries (marine and aquaculture). The platform was created to facilitate access and analysis of national and local-level agricultural data¹³.

Indigenous Data Sovereignty is a global movement concerned with the right of Indigenous peoples to govern the creation, collection, ownership and application of their data. Indigenous Data Sovereignty is outlined in the United Nations Declaration on the Rights of Indigenous Peoples (UNDRIP). Open data, especially at the local-level, is an important tool for encouraging this work.

In New Zealand, Indigenous Data Sovereignty specifically relates to Māori Data Sovereignty; the inherent rights and interests that tangata whenua (people of the land) have in relation to the collection, ownership, and application of Māori data.

Figure.NZ, a non-profit charity, provide some useful examples of open data at the local level for and about Māori through their platform [Pātaka Raraunga](#). Here you can also find links through to other agencies' data tools for Māori such as [Te Whata – iwi data tool](#) – a platform which allows the user to explore social, economic, cultural and environmental information broken down by iwi (tribal groups associated with a distinct territory). And [Whenua Viz](#) which uses data from Landcare Research with property and legal data to help Māori landowners and managers find out more about the physical characteristics, constraints, and potential of their land.

For some Māori, Māori data is a taonga (treasure) and iwi-Māori are kaitiaki (guardians) over their taonga.

Australia is committed to returning data and insights to our survey and Census respondents and educating Aboriginal and Torres Strait Islander communities on how they can use it.

¹³ <https://drsa-agriculture.netlify.app/index.html>

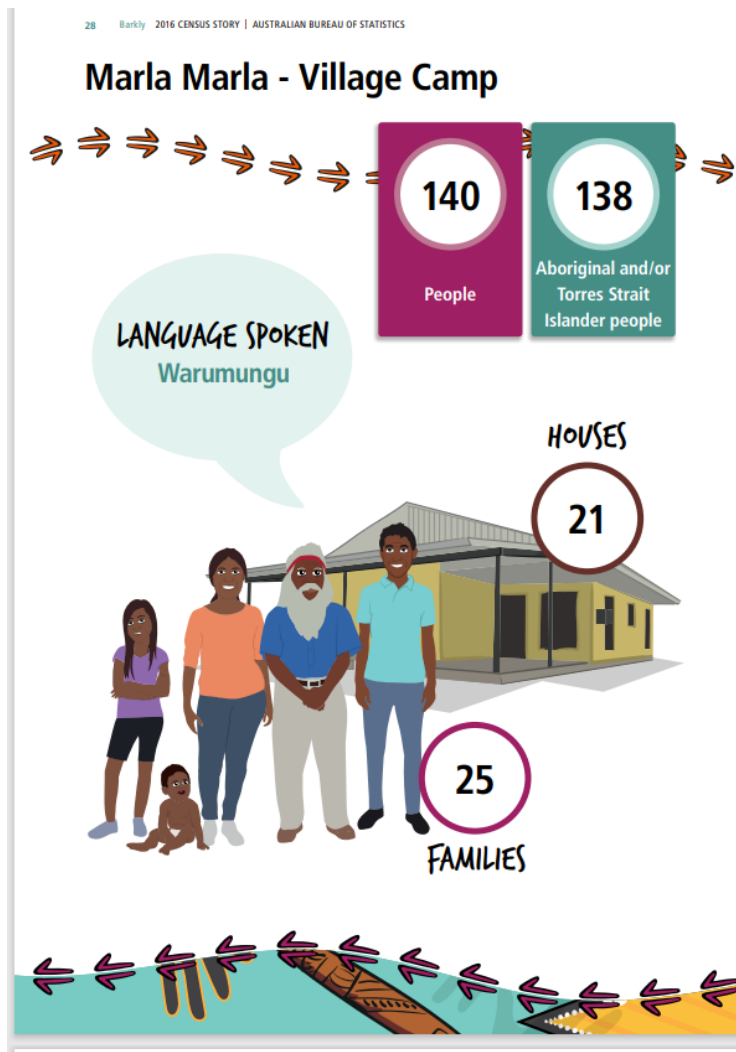


Figure 3: Example from Australian Bureau of Statistics

Following the 2016 Census of Population and Housing, Census storybooks were published for 15 geographic regions with a high proportion of Aboriginal and Torres Strait Islander people including communities, town camps, outstations and regional towns.

The storybooks explore insights the Census provides on a range of important topics including Housing, Employment, Income, Education and Language (figure 3 presents the storybook for the Marla Marla – Village Camp region).

Presented in a visually appealing and culturally sensitive format, the story books are simple and accessible to varying levels of data literacy and reading skills.

The storybooks are downloadable from the ABS website as printable resources.

Use Case 3: Developing Services by Engaging with End Users

Statistics Netherlands has engaged with users to develop a dashboard¹⁴ providing insight into low-income populations on the municipal level. This project first started when a single municipality approached CBS. They wanted to know which households used a variety of municipal welfare benefits and, in particular, which households didn't despite being eligible based on their income level. Moreover, they wanted insight into several characteristics of these households (age, household type, income type, neighborhood etc.). They rightly suspected that non-users could be identified by these characteristics, thereby enabling a more targeted, data-driven approach in increasing the use of income-based welfare benefits the municipality has on offer. Since then, several more municipalities have joined, ultimately resulting in the current dashboard. It has been developed in close collaboration with the end users, which are mainly municipal researchers and policy makers.

¹⁴ https://dashboards.cbs.nl/v2/Armoedescan_ENG/

There is a general page with basic information on low-income populations for all Dutch municipalities, defined by Dutch national policy. The user selects income limits, and the population size is displayed as well as its attributes and the spatial spread across neighborhoods. There is also a local page giving insight into the use of municipal social benefits by low-income populations. These benefit regulations can vary across municipalities both in eligibility requirements and in the extent of the benefits. As benefit participants are not registered nationally, only municipalities who have provided their registrations are included.

Statistics Sierra Leone created a webpage to meet the demand for COVID-19 related data, analyses and reports together with data from other institutions who partner with them and/or have had their work accredited by Stats SL. The Statistician General and Stats SL staff have been fully involved in the national response against COVID-19, providing technical support to the various pillars of the Emergency Operational Centre. The Statistician General has also been a member of the presidential group established in June 2020 – Scientific and Technical Advisory Group on Emergencies (STAGE – C19). Stats SL has been providing district maps to coordinators who are responsible for the district response.

Use Case 4: Hackathons and Communities for Developers

Statistics New Zealand advocates a number of local initiatives in support of open data. Two such initiatives include GovHack and Open Government Ninjas. GovHack is an annual hackathon run in both Australia and New Zealand using open government data. It aims to find innovative cyber solutions for governments and commercial businesses.

The Open Government Ninjas¹⁵ is an open data developer community that encourages the New Zealand government to adopt open data policies. Their forum is a place for people who are interested in open data to communicate, discuss potential problems and solutions as well as connect those within government with technically informed private citizens. This group played a key role in developing the Declaration on Open and Transparent Government and supporting the New Zealand Data and Information Management Principles.

Department of Statistics Malaysia organized a Virtual Datathon 2020 (DvD 2020) competition in support of open data which focused on the involvement of higher learning institutes students. The competition was a continuation of the Datathon 2018 competition with the theme “Towards the 2020 Population Census”. It aimed to encourage students from higher learning institutions in Malaysia and the public to use data sets issued by the Department and other open government data for the purpose of analysis and problem solving.

Use Case 5: Working with the media

In the Netherlands Local Focus is a data-driven news organization. Via an API, they automatically turn the statistical data that are available locally into a news article aimed at a specific municipality or region (figure 3)

Statistics Sweden has worked together with an automated news service, provided by a company run by data journalists, which generated automatic texts together with municipality statistics. This resulted in an all-time high in media coverage of the yearly population statistics, as it was very easy for local papers to include a pre-fabricated article on the population trends for their municipality.

¹⁵ <http://groups.open.org.nz/groups/ninja-talk>

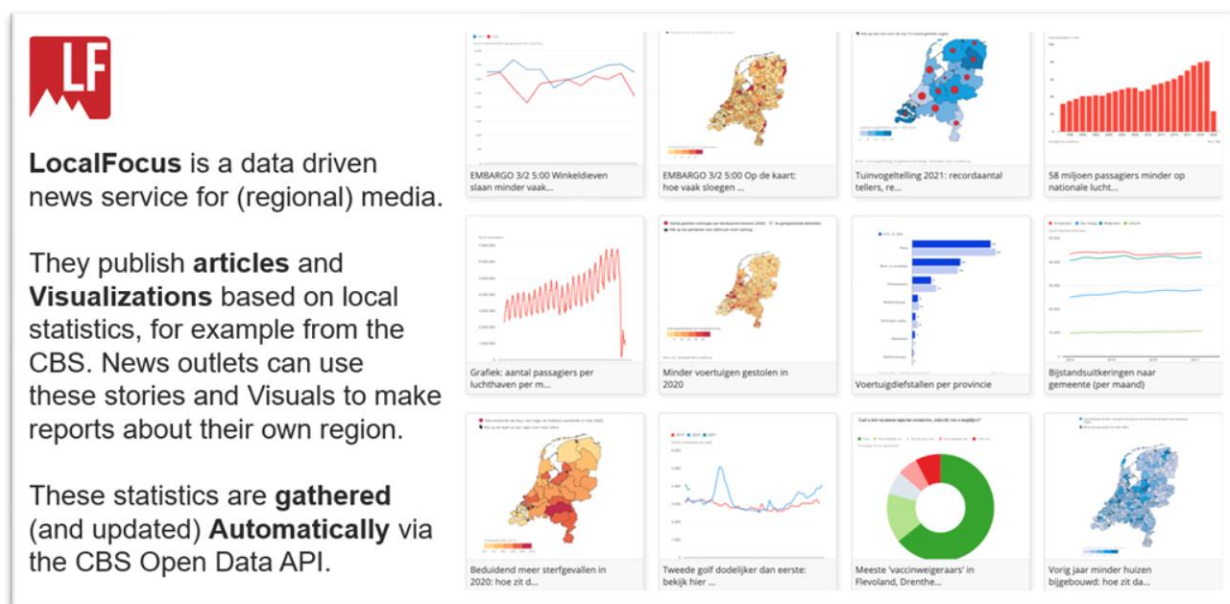


Figure 3: Example from Statistics Netherlands

Challenges

The following challenges have been identified by members of the Working Group and articulated at the Commission's side event on *"The open data revolution: the power of open data for engaging the public with statistics that matter to them"*, mentioned earlier.

Operating environment

Although this can be complex, the complexity should be seen as a benefit not a disadvantage. In New Zealand the operating environment for open data involves legislation such as the Privacy Act and Statistics Act; a declaration on Open and Transparent Government by the Prime Minister; Data and Information Management Principles which all public sector agencies must abide by, and government-wide commitment to the Open Data Charter and a licensing framework, guiding agencies when releasing copyright works and non-copyright material for re-use.

The National Statistics Office must strive to operate consistently with legislation, Prime Ministerial declarations and government-wide principles and commitments, let alone professional commitments such as the Fundamental Principles of Official Statistics. However, the operating environment should be viewed as empowering not constraining as it provides an opportunity for the Office to demonstrate leadership within the wider government sector and deliver on political and policy priorities.

Measuring impact

As with all aspects of statistical business, measuring the impact of local-level open statistics is a challenge. Impact is often measured in terms of frequency counts, such as increase in the number of hits on a webpage, rather than changed behaviors. Some National Statistics Offices have commissioned independent evaluations into the impact of Population Censuses¹⁶, which estimate, amongst other things, the economic value of investing in a Population Census. However, the Working Group are not aware of any studies of the impact of statistical activities beyond Population Censuses, let alone the specific case of local-level open data.

¹⁶ <https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Value+of+the+Australian+Census> and <https://www.stats.govt.nz/assets/Research/Valuing-the-Census/valuing-the-census.pdf>

Multiple audiences

Another challenge with local-level open statistics is the huge diversity of audiences. This challenge, however, can also be a benefit as users can be identified and case studies developed. In the case of New Zealand and the Netherlands, local councils were a main audience and beneficiary of local-level open statistics.

Privacy and confidentiality

One key challenge of local-level open statistics is real or perceived privacy and confidentiality issues. The term ‘open’ is commonly misunderstood to mean all data, even personal information for individuals and businesses, is made freely available. Calling local-level statistics ‘open’ can lead to misinformation and even disinformation.

National Statistics Offices are well placed to manage privacy and confidentiality, both have available expertise and tools for managing both. Such tools could include an explicit choice not to add labels like ‘open’ when releasing statistics, deferring instead to common practices like issuing a licence like Creative Commons. Another good strategy for addressing privacy and confidentiality challenges is to start your open data journey in a less sensitive area of statistics. For example, in New Zealand and the Netherlands, local-level statistics have been opened up for infrastructure (buildings, car spaces). Disseminating local-level data for more sensitive areas, such as person-level statistics, will attract more interest in confidentiality and privacy concerns, real or perceived.

Technology

Technology can be a major challenge, covering issues of insufficient funds to acquire and use the necessary software and visualization tools, a need to modernize websites to make it visitor navigation friendly, broadband/internet bandwidth and data archiving systems. These challenges are not specific to open data or local-level statistics. Coupling open data ambitions with other transformational activities, such as modernization programmes or major investments in Population and Agricultural Censuses, can be a good strategy for National Statistics Offices who find it easier to secure funds for specific activities, like a Population Census, and use these for the wider benefit of the Office, like local-level open statistics.

Standards and good metadata

The use of standards sounds logical but is a challenge. When it comes to combining and reusing open data, national and regional governments often use different standards and outside the country it is an even bigger challenge. There are also various groups of users with different requirements regarding standards. The challenge lies in the fact that, as a supplier, you want to provide all groups with the right data, so that they can easily work with it. Good and adequate data stewardship can offer a solution here, through which agreements about standards and minimum requirements for metadata can be established. If the standards can be combined or at least concurred, there are many possibilities for combining international, regional and local data.

The way forward: Creating an “Open Data Culture” acting as Data Steward

National Statistics Offices are well placed to encourage and help create an “Open Data Culture”, building on their core business of providing official statistics and, in some cases, their role as data stewards. Data stewardship is the careful and responsible creation, collection, management, and use of data. There is always a balance needed to ensure data use is safe while allowing the freedom to explore and experiment to create innovative new value from data. Building on the foundations of trust and transparency, National Statistics Offices can accelerate the release of open data through ease of accessibility, re-use and integration.

Increasing the number of datasets with local-level statistics through user engagement helps National Statistics Offices to stay relevant, as such statistics often can be used and re-used together with other open data, such as geospatial data. For National Statistics Offices looking for ways to increase their role in the national data ecosystem, where different actors can exchange, produce and use data¹⁷, putting out more local-level statistics as open data can be one path to follow. When used securely – protecting privacy and confidentiality – data can provide rich insights about us and our communities (figure 4).



Figure 4: Illustration inspired by Statistics Canada¹⁸, originally showing the transformation from data to information performed by a National Statistical Organisation – here illustrating that this process also results in open data that can start the consumer’s information, innovation or knowledge process.

The concepts of “Data Stewardship” and “Data Governance” are getting a lot of attention both in global official statistical communities and at national level. In a recorded World Bank event December 2020, *Reimagining National Statistical Systems: Risks and Opportunities for Official Statistics in the Public Data Agenda*¹⁹, Professor Paul Cheung, ex-Director of the UN Statistics Division, talks about how the Public Data Agenda must help preserve and strengthen a national system of official statistics. NSOs must strive to be the champion of this system, ensuring its high quality and trust-worthiness. He states that data stewardship is a strong pillar for National Statistics Offices, but only for those who have the pre-requisites and that gaining relevance through knowledge generation a key priority.

Depending on the national context where the National Statistics Office is operating, there are different paths to go. All National Statistics Offices share the same challenge; how to harness the potential from all kinds of data generated in society – and how to guard, grow and give it back to society as official statistics of high quality. Engaging with end users or working through intermediaries such as developers and media is a way to understand how to generate knowledge for society.

¹⁷ Approaches to data stewardship, Background Paper for the UN Statistical Commission 2021

¹⁸ Rancourt, Eric. ‘The Scientific Approach as a Transparency Enabler Throughout the Data Life-cycle’. 1 Jan. 2019 : 549 – 558.

¹⁹ <https://www.worldbank.org/en/events/2020/12/10/reimagining-national-statistical-systems-risks-and-opportunities-for-official-statistics-in-the-public-data-agenda>