15 December 2020

English only

Economic Commission for Europe

Conference of European Statisticians
Expert Meeting on Modernizing Statistical Legislation
3 and 5 November 2020 (online meeting)

Summary of discussions and conclusions of the Expert Meeting on Modernizing Statistical Legislation

I. Attendance

1. The 2020 UNECE Expert Meeting on Modernizing Statistical Legislation took place as an online meeting on 3 and 5 November 2020. The meeting was attended by the representatives of the following countries: Afghanistan, Albania, Argentina, Armenia, Australia, Belarus, Belgium, Bosnia and Herzegovina, Brazil, Chile, Colombia, Croatia, Estonia, Finland, Georgia, Greece, Iran (Islamic Republic of), Ireland, Israel, Italy, Kazakhstan, Kyrgyzstan, Lao People's Democratic Republic, Latvia, Liechtenstein, Lithuania, Malaysia, Malta, Mexico, Micronesia (Federated States of), Montenegro, Myanmar, Netherlands, North Macedonia, Poland, Portugal, Republic of Korea, Romania, Saudi Arabia, Slovenia, Spain, Thailand, Turkey, Ukraine, United Kingdom of Great Britain and Northern Ireland, and Uzbekistan.

2. The Expert Meeting was attended by the representatives of European Union, European Commission – Eurostat, European Data Protection Supervisor (EDPS), European Free Trade Association, Food and Agriculture Organization of the United Nations, Office of the United Nations High Commissioner for Human Rights (OHCHR), Organisation for Economic Cooperation and Development (OECD), Statistical, Economic and Social Research and Training Centre for Islamic countries (SESRIC), United Nations High Commissioner for Refugees (UNHCR), United Nations Economic and Social Commission for Asia and the Pacific (UN ESCAP), United Nations Economic Commission for Latin America and the Caribbean (UN ECLAC), United Nations Economic and Social Commission for Western Asia (ESCWA), United Nations Resident Coordinator Office in Kyrgyzstan, United Nations Statistics Division (UNSD) and United Nations Statistical Institute for Asia and the Pacific (UN SIAP).

3. Aston University, King Abdulaziz University, Barclays Bank PLC and Sensiblecode.io and one independent expert also participated in the Expert Forum.

II. Organization of the meeting

4. The Expert Meeting consisted of two sessions:

(a) Session 1: The right to privacy and the right to live in an informed society (3 November 2020) – Session Chair: Ieva Zaceste (Latvia)

(b) Session 2: Data access use and exchange through a legal lens (5 November 2020) – Session Chair: Nicola Shearman (United Kingdom)

5. Sessions were held in English only via Webex.

6. The meeting was organized by the organizing committee consisting of: The Expert Meeting is organized by a committee consisting of the United Kingdom (chair), Albania,

Armenia, Latvia, Poland, Eurostat, the Organisation for Economic Cooperation and Development (OECD), the United Nations Statistics Division (UNSD) and the United Nations Economic Commission for Europe (UNECE). The outcome of the deliberations during each session are summarized below.

7. All meeting documents are available at: http://www.unece.org/index.php?id=53890.

III. Opening and adoption of the agenda

8. Lidia Bratanova, the director of the UNECE Statistical Division, and Nicola Shearman, the chair of the organizing committee and the Expert Meeting, opened the meeting and welcomed the participants. The first Expert Meeting on Modernizing Statistical Legislation is a continuation of work carried out under the Conference of European Statisticians (CES) on strengthening the legislative and institutional framework of official statistics. Fundamental Principles of Official Statistics (FPOS) were adopted by UNECE in 1992, United Nations Statistical Commission in 1994 and by United Nations General Assembly in 2014. In 2016, CES endorsed the *Generic Law on Official Statistics* and in 2018, the *Guidance on Modernizing Statistical Legislation*. The documents have served as a basis for updating statistical legislation in a number of countries and a generic law adopted by ECLAC.

IV. Session 1: The right to privacy and the right to live in an informed society (3 November 2020)

9. Ieva Zaceste (Latvia), the chair of the session, introduced the topic and emphasized the importance of NSOs being able to navigate privacy and data protection laws. Use of new sources has become an integral part of statistical production, allowing to improve data quality, reduce response burden, avoid duplication of data and capture socioeconomic phenomena in an entirely novel way. NSOs already use extensively administrative data and are starting to use big data and privately-held data to serve many stakeholders: government institutions, international organizations, NGOs and private persons. Although confidentiality is a fundamental principle of official statistics, NSOs also need to carefully consider the right to privacy and data protection to ensure proportionality and existence of necessary checks and balances.

10. During the session, Edina Harbinja of Aston University explained the main concepts and introduced main legal instruments on privacy and data protection in the European context. Jekaterina Macuka, the Head of the Data Protection Supervisory Authority in Latvia, explained the lawfulness of personal data processing in the context of official statistics from the perspective of the General Data Protection Regulation (GDPR). Massimo Attoresi from European Data Protection Supervisor's (EDPS) Secretariat presented the elements of the European Union (EU) legal framework of on statistics, privacy and data protection and reflected on challenges from a data-driven world. Marc Titus Cebreros from the United Nations Office of the High Commissioner on Human Rights (UNOHCHR) gave an overview on the provisions on the right to statistical information given by the International Human Rights Law. Gabriel Gamez of the United Nations Statistics Division (UNSD) presented considerations on principles of protection of data collected for statistical purposes in the context of the international Covid-19 response. Meeting participants shared their perspective on the key challenges for official statistics within the existing legal framework and discussed how NSOs can address them using available tools and increase the trust of data providers and data subjects.

11. The following points were made throughout the session:

(a) In the European law, the right to privacy is referred to as the right to respect for private life. This right consists of a general prohibition of interference, which can only be justified subject to some public interest criteria. The right to respect for private life is embedded in the Universal Declaration of Human Rights (UDHR, Article 12), the European Convention on Human Rights (ECHR, Article 8), Charter of Fundamental Rights of the EU (Article 7) and the OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data.

(b) Unlike the right to privacy, the right to personal data protection is an active right, putting in place a system of checks and balances protecting individuals when their personal data are processed. Essential components of personal data protection are independent supervision and respect for the subject's rights as each processing of data is interaction with a person's private life. The key instruments are the Treaty on the Functioning of the European Union (Article 16), Charter of Fundamental Rights of the European Union (Article 8, Protection of personal data), General Data Protection Regulation (GDPR) and Privacy and Electronic Communications Directive. Also relevant is the Modernized Convention for Protection of Individuals with regard to Automatic Processing of Personal Data.

(c) As with other fundamental rights in the EU, any limitation of the exercise of the right to privacy and personal data protection should be provided by law and respect the essence of these rights and freedoms (Article 52 of the Charter of Fundamental Rights of the European Union). The international human rights framework emphasizes that data collection process should not lead to adverse discriminatory effects and be used as a tool directly or indirectly exposing members of the population to harm of their human rights.

(d) The public's right to access to statistical information stems from several human rights, e.g. the right of participation in public affairs (Article 25, International Covenant on Civil and Political Rights (ICCPR)), right of access to information (Article 19, ICCPR), right to benefit from the scientific progress (Article 7 UDHR and Article 15, International Covenant on Economic, Social and Cultural Rights (ICESCR)). The citizens cannot be expected to be engaged in the political process in a beneficial manner without the necessary information being made accessible to them. Special treaty provisions oblige Member States to produce statistics on and for human rights, e.g. Disability rights convention (Art. 31, CRPD) or Belem do Para Convention (Art. 8(h)). This obligation includes due disaggregation, independent and impartial collection, broad participation by stakeholders in the production process and a number of quality criteria, including accuracy. Recognition of this obligation is reflected in the SDG target 17.18 and treaty body practice and interpretation. The treaty bodies have often recommended the Member States enhancing and strengthening the capacity of NSOs to produce these data. There is an increasing awareness that official data and statistics are a public good and essential to human rights analysis, reporting and implementation (e.g. General comment No. 25 (2020) on science and economic, social and cultural rights). This is reflected in the corresponding constitutional, legal and policy guarantees being established in many parts of the world.

(e) In the European law, article 10 of ECHR ensures the right to receive information by individuals. The EU Regulation 2019/1700 establishes a common framework for European statistics relating to persons and households, based on individual data collected from samples. Article 9 specifies the data sources that can be used, which include administrative records and other sources, methods or innovative approaches insofar as they allow to produce data that are comparable and compliant with the applicable specific requirements. Key legal instruments for the protection of data collected for statistical purposes include article 338(2) of the Treaty on the Functioning of the European Union and regulation 223/2009 on European statistics. Microdata access is regulated by the EU Regulation no. 557/2013 on access to confidential data for scientific purposes. Other sources of principles are the European Statistics Code of Practice and UN Fundamental Principles of Official Statistics.

(f) Other important elements of the EU legal framework are the Public Sector Information (PSI) Directive from 2003 (amended in 2013) and Open Data Directive adopted in 2019, which will be implemented in members states' legal systems next year. The PSI Directive encourages re-use and sharing of data, which creates tension with GDPR recommending data minimization. The Open Data Directive defines high-value datasets that should be available free of charge, which include statistics (article 14). However, data protected by statistical confidentiality will be excluded from open data principles as well as personal data or data whose sharing would undermine protection of privacy and integrity of the individual. This Directive also brings about tension with the GDPR data minimization principle as it includes a principle "as open as possible as closed as necessary".

(g) GDPR applies to the processing of personal data, including processing for statistical purposes. Personal data is any information relating to an identified or identifiable natural person ("data subject"), including pseudonymized data if indirect identifiers still exist or attributes allow for re-identification, considering technology available during processing. GDPR does not apply to data of public organizations, companies or any legal entities. In the language of GDPR, NSOs are "data controllers" because they are in possession of identifiable data and process them. GDPR applies on all controllers processing data of individuals who are in the European Union, irrespective of where the controllers are established.

(h) GDPR defines the principles of the processing: lawfulness, fairness, transparency, purpose limitation data minimization (necessity & proportionality), accuracy, storage limitation, integrity and confidentiality and accountability. It gives data subjects a number of rights: the right to erasure, the right of access, the right to be informed, the right to restrict processing, the right to rectification, the right to data portability, the right to object, rights related to automated decision making. Further protection is delegated to the statistical law but within the limits of the GDPR.

(i) GDPR gives the tools to balance between the right to privacy and personal data protection and the right to information, as it allows for derogations from the principles and the data subjects' rights (Art. 89). There is a clear basis for derogations when processing data for statistical purposes. These derogations need to be regulated by national laws which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject. These measures may include data minimization principles, mechanisms to prevent risks related to pseudonymization and safeguards in the form of technical and organizational measures. A different implementation of these laws in member States, in particular derogations from data subject's rights in different countries result in problems for the statistical institutions which need to navigate different national laws.

(j) Processing of personal data as well as transmitting data from other controllers requires a lawful basis. Two bases applicable in the context of official statistics are public interest/official authority (Article 6(1)(e), GDPR) and legal obligation 6(1)(c). The legal basis for public interest has to be laid down by national law or the EU law. The data subject must be advised which legal basis is used for processing of their data. Data collected for a different purpose can also be processed for statistical purpose if there is a legal basis. The legislation and sub-legislation of the country should ensure the legal basis for personal data processing for statistical purpose. A public authority, such as an NSO cannot use consent as a legal basis. Consequently, consent for participation in a survey, which is sought for ethical reasons, does not have to meet a GDPR consent standard because consent does not provide the lawful basis for statistical data processing. Moreover, the data subject cannot invoke its rights given by GDPR in this context. Communicating this to the respondent when asking for their consent may be challenging.

(k) Transparent regulation, such as a statistical programme, is needed on the purpose of statistics, categories of personal data processed, controllers (authorities/companies) from whom personal data is collected and algorithms used for

analyses of collected data. Transparency of algorithms and methods and data sources is needed to justify the necessity of transmitting data and ensure the lawfulness of personal data processing.

(l) NSOs need to have specific policies, procedures and technical infrastructure in place to ensure data protection and make them known to the public. A vital protection for the citizen preventing surveillance or profiling is guaranteeing the "one-way valve" – that data acquired for statistical purpose cannot be used for any other purpose. Even if the statistical law obliges all national and local authorities and private bodies to provide their data, the necessity of data acquisition should be carefully assessed even when data protection safeguards and independent monitoring are in place. The more data is in control of NSOs, the higher the risks are, e.g. of surveillance and profiling or of NSOs being a target of hacker attacks. It is also important to consider judicial redress in the context of penalties for breaches of confidentiality.

(m) There is a public interest ground for opening the collection of various types of data for statistical purposes, which can be used further for scientific purposes. Access to confidential data for research and scientific purposes needs to be carefully regulated by law, based on the nature of data and risks for individuals.

(n) Although sometimes, we refer to a "balance" between the right to privacy and the right to information, in the light of Charter of the Fundamental Rights of the European Union, the fundamental rights are not up for balancing. When a measure imposes such a comprehensive limitation on the exercise of a fundamental right that it calls this right into question, that measure is incompatible with the Charter and automatically disproportionate.

(o) One of the key challenges for official statistics is the tension between ensuring adequate and sustainable access to new data sources to produce innovative statistical products and services versus the right to privacy and the access to personal under conditions laid down by the law. It may be a political choice what tasks NSOs will get, but these political choices need to be within the existing legal framework, fundamental rights, international human rights and accords. More in-depth thinking on privacy and data protection is also needed on the implications of NSIs becoming data stewards in relation to principles of official statistics and the existing legal frameworks.

(p) Another key challenge is the low level of trusts between data subjects and data controllers such as NSOs. There is a lack of awareness on all levels that individual data is used to provide information at an aggregate level. Communication, partnership, transparency, accountability and continuous advocacy about the role of NSOs are key to build trust with data providers especially as they are often users of statistics as well. A key tool for increasing transparency are the Fundamental Principles and regional and national codes of practice for official statistics. Federation of infrastructure, use of privacy-preserving computation techniques like differential privacy, and minimizing the data to satisfy the necessity should also increase the level of trust.

(q) Increased information needs during the Covid-19 pandemic and similar events are a valid argument for accessing more data sources, but the same rules on the rights to privacy and data protection are in force unless there are changes in national and international legislation. Data protection authorities decide on the necessity if there is a high public interest and legal basis while maintaining the balance between the intrusion and necessity and proportionality. E.g. in the EU, there was a big discussion whether contact tracing applications should have centralized or decentralized architecture and the European Commission considered decentralized architecture enough.

(r) Under the EU data strategy (19/02/2020), it is planned to create a single European space where personal and non-personal data are secure and can be accessed by the EU businesses and the public sector. This initiative intends to unlock the potential of re-using public sector information, or privately-held data while upholding the rights to privacy and

personal data protection. The corresponding communications mention explicitly that not enough private sector data are available for use by the public sector to improve the scope and timeliness of official statistics. A key enabler for this will be reinforcing data governance, infrastructure and technologies, including the legislative framework. This strategy includes the concept of data intermediaries and NSOs might have a role in this.

(s) The United Nations personal data protection and privacy principles aim to harmonize standards across United Nations System Organizations, facilitate accountable processing of personal data for purposes of implementing the organizations' mandates and ensure respect for the human rights and fundamental freedoms of individuals, including the right to privacy. After the outbreak of the Covid-19 pandemic, the UN organizations were prepared a joint statement on Data Protection and Privacy in Covid-19 response; however, the discussions in the international context are less mature than on national and supranational level and further discussions are needed. In particular, the participation of international organizations in development of global data platforms and partnerships extending beyond UN system should be considered, as they increase the complexity of ensuring accountability for proper data use, security and governance.

Conclusions

12. The session provided an overview of the legal framework in place and explained the fundamental concepts so that the statistical community can speak competently about privacy, confidentiality and data protection.

13. NSOs should have the access to data they need within the existing framework and respecting the rights to privacy and personal data protection. Protection of privacy and personal data does not prevent the fulfilment of the right to information and there are tools to ensure balance. A legal basis for accessing and processing data for statistical purpose should be put in place through national legislation and sub-legislation. NSOs need specific policies, procedures and technical infrastructure to ensure data protection.

14. Transparency about the purpose, data sources, safeguards and methods is a key aspect of the trust-building dialogue between statisticians, data protection officers, policymakers and the society. Raising awareness of the value of official statistics as a public good, input to human rights realization and a basis for evidence-based decisions remains critical and helps justify the necessity of personal data processing. This is especially relevant in the context accessing new data sources.

15. NSOs have fundamental principles of official statistics in their DNA and can provide significant input for the new data ecosystems. Further thinking on privacy and data protection may be needed in the contexts of NSOs becoming data stewards and international data sharing.

V. Session 2: Data access, use and exchange through a legal lens (5 November 2020)

16. Nicola Shearman (United Kingdom), the chair of the session, introduced the topic of the session. Data is "the fuel" of national statistical systems, enabling the production of official statistics which are used to inform decisions and drive policymaking. NSOs are increasingly turning to administrative and non-traditional data sources to improve the coverage, efficiency and quality of the regular data production, meet new information needs and reduce response burden and the cost of statistical production. Mandates, safeguards and assurances are key to establishing productive relationships with the holders of these data, but in addition to appropriate legislation, a culture of data sharing based on ethical data use is also needed.

17. The session included presentations and two rounds of questions to the presenters.

18. Anahit Safyan (Armenia) shared the experiences of Armenia on using administrative data for improved production of official statistics:

(a) In 2018, Armenia introduced a new law on official statistics fully based on the *Generic Law on Official Statistics*. The new law defines official statistics, introduces a national statistical system (NSS) including all producers of official statistics, identifies Armstat as being the main producer, gives administrative registers a clear status of data source for official statistics and gives their custodian the responsibility to make the data available and consult Armstat in case of revision or a new data collection. The legislative process lasted two years and built on much longer efforts on improving access, which started from encouraging national and local governments to establish administrative registers and databases for their own use.

(b) To operationalize the law and strengthen its coordinating role in the NSS, Armstat uses MoUs on data delivery and cooperation on data quality with most admin authorities, including provisions on data quality and confidentiality, following a template from Statistics Norway. Training is an important tool for building the culture of using public data. A high-level seminar for ministers was organized to establish the necessary understanding of the importance of using admin data for producing official statistics. Training sessions on statistical classifications and methodology were held for the tax authority and several ministries.

(c) A key milestone in getting access to administrative data was getting access to tax data on households and enterprises in 2018. Now registries from the Ministry of High-Tech Industry, the central bank and State Revenue Committee are used for trade and services and wages and labour statistics. The first combined census will be based on the population register recently updated with data of the electronic border management system and 25% sample collection of additional data using tablets. The census is planned to take place in 2021 after being postponed from 2020 due to Covid.

(d) During the Covid-19 pandemic, Armstat provided methodological support to the tax authority in the interpretation of separate definitions of NACE classification, needed in the implementation of economic and social assistance programmes. This service allowed to significantly improve the tax register quality, prove the importance of the classification and promote more effective collaboration between administrative registers, in particular tax and social register.

(e) Two main issues in further operationalization of the new law are forming an integrated and coherent NSS and a further increase of the use of administrative data for statistical purposes. The government takes action to establish and improve admin registers and databases in line ministries and introduce unique identifiers in all registers. To support this process, Armstat developed a concept note for the establishment of a digital integrated administrative registers system as strategic national infrastructure. The key messages are that effective use of admin data brings benefits for both statistics and administrative authorities and contributes to building public trust by fostering professional independence, high professionalism, scientific principles and transparency.

19. Mart Mägi (Estonia) presented maturity levels in increasing data access and corresponding considerations to be taken into account in the statistical legislation:

(a) In Estonia, 70% data used in official statistics come from administrative data sources through a multitude of administrative databases and thousands of data feeds. This is possible thanks to data governance, allowing to understand where data can be found and what they mean. Statistics Estonia uses interoperability network allowing to move data very quickly between databases and facilitating data linking.

(b) Data governance is essential in improving access and use of new data sources, and NSOs should play an active role in its centre. NSO should support the harmonization of data descriptions in public organizations and set standards for describing data and

interoperability of data, including metadata. A central repository of metadata is essential to make maximum use of public data, enable finding and linking them. In Estonia, data flows are completely separate from the metadata flows. Statistics Estonia also follows the developments related to the establishment of the government open data portal.

(c) In a future-oriented law, a legal obligation to answer to surveys should be replaced with an obligation to provide data. In Estonia, a solution allowing for voluntary direct transfer of lowest level microdata from organizations' ERP systems to the statistical office has been developed. The solution, based on XBRL data exchange and using a state ontology, allows for any survey question to be calculated on the side of the public service, reducing response burden and providing much better information on the economy.

(d) Before getting access to personalized, sensitive data from private data holders, it is important to consider using data pre-processed on the side of the data holder – privacy by design. With a solid methodology used by the data holders to anonymize the data and the right for NSO to audit the implementation, this approach can alleviate privacy and security concerns and trust issues. Statistics Estonia does not have a legal basis yet for accessing personalized mobile data, but during the first wave of Covid-19, it produced mobility analysis cooperating with all mobile operators. This exercise has produced more accurate results than the mobility analysis produced by Google because of Statistics Estonia's access to all microdata. Data from multinationals are needed, but a huge value can be added by NSOs by linking them to other sources. The same approach could be used with banking or any other sensitive data. Upcoming e-privacy regulation is a concern because proper access to mobile and credit card data is needed.

20. Adrienne Harrington and Paul Morris (Ireland) provided an overview at the legal environment and internal governance structures at Central Statistics Office of Ireland (CSO) in Ireland, and the CSO efforts in establishing productive relationships with data providers.

(a) Legal environment is constituted by data protection legislation, national and European statistical legislation and freedom of information legislation. Confidentiality and data protection have always been at the core of all CSO work, but GDPR significantly increased the emphasis on data protection, oversight mechanisms and governance structures. Data protection officers in CSO have new roles and responsibilities, and powers and resources of the data protection commissioner are enhanced. Data Sharing and Governance Act providing a legal basis for sharing of data by public bodies was passed in 2019, but it is still under implementation.

(b) Key elements of the CSO extensive internal governance arrangements are the management board, data custodians, data protection officers, a data management policy, the Confidentiality and Data Security Committee, the Data Office and the Administrative Data Centre. Data custodians are statisticians within the office who have individual responsibility for data they manage. Data management policy creates the CSO data management framework and provides guidelines for ownership, management, storage, retention, protection, destruction and security of CSO data. It is a living document subject to continuous review and was adjusted during the Covid-19 pandemic. The Confidentiality and Data Security Committee coordinates all obligations relating to statistical confidentiality and data security and is chaired by the CSO Data Protection Officer who is also a member of the CSO management board. The Data Office is responsible for compliance, data management policy, overseeing data protection impact assessments and provides data management training, including obligatory training for new staff. The administrative data centre, set up in 2009, is core to the management of administrative data within CSO. Data from the providers flow to the administrative data centre, which releases to individual statisticians on the basis of proportionality and necessity.

(c) CSO has the legal right to access the public sector data and to be consulted in system changes, but these rights have to be augmented with some incentives to be effective in practice. Data providers are also users of the CSO and their own data, so they have an

interest in improving their quality. CSO data services, involvement in public sector data strategy, secondments, bespoke analysis on policy issues and methodological support are central to data access discussions and building productive relationships with data providers.

(d) CSO data services include facilitating access to the CSO datasets for researchers through research microdata files or providing dashboards and guidance to the public sector, e.g. on the percentage of public service records that have identifiers allowing for linkage. Provision of data services is limited to the CSO mandate, so it does not include providing support in econometric analyses or analytics.

(e) CSO maintains two types of liaison groups: with analysis departments and with data departments. Liaison groups with analysis departments include all the government agencies and research organizations, who are also data owners. The groups hold wide-ranging discussions about research and statistical activities in CSO and the various bodies and highlight the value of linked data in the spirit of "help us to help you". Liaison groups with data departments discuss operational details, the legal basis for developing MoUs, addressing GDPR requirements, and proposed quality improvements.

(f) During COVID-19, CSO received for the first time medical datasets under the condition that it facilitates safe research access for epidemiological purposes. The access was enabled by a special provision in the statistical law, allowing CSO to access medical records released by the Department of Health. This allowed for producing a number of reports and recommending quality improvements. In terms of private data sources, the Department of Health paid for CSO to get micro aggregates from one of the mobile operators, also on condition of providing data access for researchers. New opportunities emerged in relation to using aggregates from credit cards microdata from traffic sensors to monitor the impact of restrictions as real-time data were needed. These relationships need to be built on to last beyond the pandemic.

21. In the discussion, it has been mentioned that internationally the statistical community is much less equipped to access data from national and international private data providers than in the European Union. Countries which have a new legal act on statistics should follow the technical, methodological and legal developments to make sure that other legislation enables future access to innovative data sources.

22. Jon Hussey (Barclays Payments) presented the collaboration between Barclays Payments with ONS within their data science campus, showing how can payments data can be leveraged to support producing more agile and granular economic indicators:

(a) As a British universal bank and one of the largest payment processors in Europe, Barclays can see 1/3 payments made in the UK, which gives them a broad and deep understanding of the payments environment and consumer behaviour. The collaboration started in 2017 when ONS sought proactively commercial organizations that could work with ONS within the data science campus looking at new data sources and privately held data sources in a collaborative approach. ONS and Barclays gradually developed the thinking on how to partner and use payments data to improve the understanding of the economy.

(b) In 2017, a data hackathon took place, in which teams comprising Barclays and ONS representatives competed to leverage all publicly available information and data: ONS indices and anonymized data on values and volumes of payments activity published online by Barclays. The winning entry showed that certain payments data, aggregated in a safe, anonymized way, can help to supplement the country economic indicators, including ONS GDP calculations. New data source resulted in more timely outputs, increased granularity of data, analysis of regional or subregional level trends, and introduction of additional variables not readily available from traditional sources, such as spending from foreign versus domestic cards or online versus in-person transactions.

(c) During Covid-19, Barclays has supported the UK pandemic response by establishing a weekly data feed to ONS. The extracts are fully anonymized and indexed to

protect clients' privacy and commercial interests. The output showing the impact on spending across industry, geography and other dimensions has been actively used by decision makers across multiple government departments. Barclays also supported ad-hoc ONS publications that rely on data that cannot be collected during the pandemic like the international passenger survey, where the Barclay index of foreign-issued card spending has been used to model the missing data.

(d) Key learning from this collaboration is that data privacy remains the primary focus for Barclays and ONS. Clarity and standardization of anonymization/aggregation approaches and legal considerations are of paramount importance. Collaboration takes time and trust. Quick response to the Covid request was possible due to the pre-existing relationship. Collaboration scope must be clearly defined and working on a single, high-impact goal greatly improves delivery times and chances of success. Collaboration requires diverse and dedicated teams, and the work could not be completed without the unique expertise of both sides.

(e) The collaboration provided the opportunity to learn and to share data and skills and get a different view how to apply statistics and collaborate on areas were payments data can make a significant impact and decrease reliance on survey or regulator data. Almost every large company already discloses key statistics within annual reports, e.g. number of accounts, companies, changes in volumes and a significant amount of information in totals to regulators. There are key datasets of significant impact for social and national interest that can be disclosed in a very safe way. National or international rules standardizing the principles of data sharing and recognizing private new data sources are needed to assure companies that they are applying the right ethics and privacy measures.

(f) The main incentives for such a collaboration included the fulfilling the responsibility resulting from understanding the power of the data and its impact on the society, wanting to support the United Kingdom economy and society, showcasing innovation in technology and database solutions and showing that Barclays' data can be trusted. Key risks considered were data leakage, ensuring that all internal approvals are in place, protecting data along the pipeline, impact on commercial data-related activities going on in the bank and misinterpretation of the data due to their specific characteristics based on the coverage of different areas in the economy.

23. Albrecht Wirthmann (Eurostat) presented the overview of the European Commission's legislative initiatives on the access to privately held data and how statistical community participates in these initiatives to ensure that interest and needs of official statistics are met:

(a) In the statistical community, the work started in 2017 with the ESS position paper on access to privately held data of public interest. Further, a concept of trusted smart statistics was introduced and a strategy and an action plan enabling the use of such data were put in place. The trusted smart statistics introduced a number of principles, such as using new data sources for multiple purposes, producing certain statistics combining several sources, introducing a layered organization of data processing workflow to decouple the technical aspect from the statistical aspects and to cope with the technological development and heterogeneity of the data, using modular methodological frameworks, pushing computation to the source, and using data without sharing them using privacy-preserving techniques. Pushing computation to the data source helps to balance risks related to confidentiality but requires transparency and involving statisticians in developing the methodology for pre-processing to make sure that results from different data sources fit together.

(b) Actions supporting the use of privately held data sources cover legislation and principles, communication and stakeholder engagement, creating sustainable business relations and enhancing technology and statistical methodology. Principles for legislation guiding the use of privately held data sources include confidentiality, strict professional standards, respecting business interest and confidentiality of certain data sources and

applications, data minimization, proportionality to the benefit to the society, not overburdening certain actors, equal access to all NSOs, transparency to the public, access modalities respecting data providers' interest and free data access.

(c) Several legislative activities of the European Commission contribute to enabling data access. Open Data Directive, amended in 2019, mentions statistics as one of the high-value data sets which are associated with important benefits for the society, the environment and the economy, suitable for creating data value services, new jobs and having several beneficiaries. A business-to-government (B2G) expert group representing data users and producers from different domains developed recommendations for sharing data from private organizations to governments. The recommendations proposed establishing national governance structures supporting B2G data sharing, creating and recognizing data stewards' functions in public and private sectors, encouraging the creation of data stewards' networks as a community of practice, encouraging collaboration between public and private should in testing environments and via public and private partnerships and creation of a regulatory framework governing the data sharing mechanisms. The recommendations were taken up in European Data Strategy aiming to create a single market for data enabling innovative processes, produces and services, based on four pillars: governance, enablers, competences and Common European Data Spaces (CEDS). CEDS are intended as data sharing spaces. The legislative framework will create a governance structure handling data sharing across member states and sectors, covers secondary use of public sector data, a labelling scheme for data sharing services, governances for data donations and management of personal data spaces.

(d) The European Commission is currently is preparing a data act which will cover the B2G data sharing following the recommendations of the group, data stewards function and implementation of the data stewards network. The data act is expected to be published as a proposal in the second half of 2021, and the work to incorporate interest of official statistics is ongoing. Digital services act regulates the relationship with platforms. A proposal covering the needs of official statistics concerning the e-Privacy regulation, especially the use of metadata for compiling and producing official statistics has been submitted to the German presidency and is currently being discussed.

(e) It is important that the statistical community sees itself as part of a bigger development in increasing the use of new data sources for the common good. The statistical community can actively take part in the new developments and emphasize its role, taking up coordinating activities in the context of data stewardship, so that the new data ecosystems eventually allow using new data sources for producing official statistics in a more agile way.

24. Simon Whitworth (United Kingdom) presented how the Ethics Office of the United Kingdom Statistics Authority (UKSA) helps researchers and statisticians from across the statistical system to think about the ethics of their research and statistical production.

(a) Digital economy act introduced in 2017 provided a legal framework for ONS, to access to data held by crown bodies, other public authorities and undertakings for statistical and research purposes and enabled secure data shares with Scotland, Northern Ireland Wales to support their statistical needs. The legislation included provisions on privacy and security of data and obliged the head of Data Ethics to observe appropriate ethical standards.

(b) UKSA produced six high-level ethical principles with the National Statistician's Data Ethics Advisory Committee (NSDEC). NSDEC is an advisory committee assessing research projects from across the research and statistical community and providing the National Statistician with advice and assurance that use of data is ethical and for the public good. It provides an external perspective thanks to the majority of independent experts in data law, data protection, research and statistics, data science, government departments and independent chairing. Full information about meetings and deliberations of the Committee is available online.

(c) UKSA identified that to make sure that have ethical principles have an impact, researchers and statisticians applied data ethics support. To provide such support, UKSA developed a self-assessment tool and provides training, online guidance and oversight of the self-assessment process. The self-assessment tool breaks down high-level ethical principles into individual components that researchers can use as a framework to get an ethical score for their research, identify the risks and come up with strategies to mitigate. The tool is mostly used by various government departments, academia and commercial researchers who want to access government data, but also charities and local government. This allows UKSA to be seen as a thought leader in the data ethics space and a trusted user of data that takes ethics seriously. Example of projects supported are the COVID-19 infection survey, linkage of hospital episodes records with census data, mortality data and primary care data to determine the level of relative risk of hospitalization and risk, environmental and socio-economic impact assessment from geological disposal facilities or a big academic project "Data for children". For some projects, support was provided already in the design phase.

(d) As per the strategy published in July, UKSA will build upon the achievements, aspiring to be recognized world-leaders in the practical application of data ethics for statistics and for research. A centre of excellence is planned to continue to develop the self-assessment tool, provide guidance and support NSDEC. New activities will include online training for users beyond government, new ethics user support service helping to take ethical guidance into account at the research design stage, new guidance on cross-cutting ethical issues and expanding domestic and international collaborations.

25. Julien Dupont (OECD) presented challenges to existing national and international ethical frameworks resulting from new data environment and lessons learned from the OECD experience:

(a) Existing national and international ethical frameworks are challenged by new data environment, digitalisation of economies and societies, artificial intelligence, increasing computing capacity. These challenges are accelerated due to the Covid-19 crisis. The demands for statistics and technical and legal risks for NSOs and international organizations are also increasing.

(b) New data sources are increasingly heterogeneous. The OECD statistical products rely on established sources, such as data from NSOs, but some studies are based on a combination of data from established sources and new data sources. Processing personal data, including data collected by the OECD or on behalf of the OECD, follows the OECD data protection rules, which define the principles of processing and the rights of individuals. The rules are surveyed and controlled by the OECD data protection commissioner and data protection officer.

(c) To address the challenges, OECD has renewed the statistical framework by introducing quality by design approach with security, ethical and legal risk at the centre of quality management. Revised quality objectives are intrinsic data quality (accuracy relevance, coherence, completeness etc.), timeliness and punctuality, accessibility, reproducibility and transparency, and security related to data confidentiality, protection and integration. It is a dynamic framework needing ongoing review and development. Existing principles, such as the United Nations Fundamental Principles, *Generic Law on Official Statistics*, European Statistics Code of Practice and the OECD Recommendation on Good Statistical Practice should be reaffirmed, safeguarded and examined for the need of updates concerning access to new sources and promoting responsible use of personal data.

(d) Better positioning official statistics as a trusted source of information is also needed to respond to the new challenges. NSOs have extensive experience with statistical confidentiality, quality management and combining data sources. Most NSOs already work with external data providers. This experience should be more visible in international cooperation and reflected in the modernized statistical legislation. Countries like Armenia, Norway, Cost Rica already embed data quality and quality management in the law.

Consistency with other legislation, support for national and international cooperation in collecting, processing, and sharing personal data, promoting the value of existing statistics, ensuring that public is well-informed and that the approaches adopted are implemented with full transparency and accountability are also of key importance.

Conclusions

26. The session provided an opportunity for countries and organizations to share their efforts in improving data access and use, through modernizing statistical legislation, building trust with data providers and creating a culture of data sharing.

27. Modern legislation is needed to address the challenges of new data ecosystems and provide safeguards supporting the trust and relationships established based on the mutual benefits of collaboration. Strong data governance in NSS guarantees compliance with the principles and allows for taking the full advantage of new data sources needed by NSOs.

28. NSOs have extensive experience with statistical confidentiality, data ethics, quality management and linking data sources, and their active involvement brings improvements across the whole national data ecosystems. The statistical community can actively take part in the bigger developments in the increasing use of new data sources for common good while working towards using new data sources for producing official statistics in a more agile way.

29. National or international rules standardizing and harmonizing the principles of private data sharing are needed to assure companies that they are applying the right ethics and privacy measures and will facilitate data sharing globally.

30. Privacy, quality and collaboration by design are key approaches to consider in the new data ecosystems.