



Economic Commission for Europe**Conference of European Statisticians****Seventieth plenary session**

Geneva, 20-22 June 2022

Item 8 (d) of the provisional agenda

**Topics of common interest to statistical and geospatial communities –
joint session with the Regional Committee of United Nations Global Geospatial
Information Management for Europe (UN-GGIM: Europe)
Developments in data stewardship****Data stewardship developments in Australia****Prepared by the Australian Bureau of Statistics***Summary*

The document describes the Australian data ecosystem developments and the modernisation of the Australian Bureau of Statistics, set against a backdrop of an increasing number of Australian Government data initiatives. The paper is structured around the four pillars of better governance, better collaboration, better methods and capability, and better access and understanding of data stewardship. The paper concludes with Australian developments in data integration and accessing new data sources.

The document is presented to the Conference of European Statisticians' session on "Developments in data stewardship" for discussion.



I. Introduction

1. Over the past few years, the Australian Bureau of Statistics (ABS) has been modernising by responding to the need for more timely data (most prominent during the COVID-19 pandemic), reducing the burden on businesses, modernising its systems to reduce risk, working across Government to provide integrated data to inform policy, and exploring opportunities to secure private sector datasets for public policy purposes.
2. The ABS' modernisation is set against a backdrop of an increasing number of Australian Government data initiatives including establishment of an Australian Public Service (APS) Data Professional Stream in 2020, release of an Australian Data Strategy in 2021 and the passing of legislation in 2022 for a national data sharing scheme (the Data Availability and Transparency Act 2022).
3. This document presents Australian data ecosystem developments. The paper is structured around the four pillars of better governance, better collaboration, better methods and capability, and better access and understanding put forward in the 2020 UN Statistical Commission data stewardship report¹. The paper concludes with Australian developments in data integration and accessing new data sources.

II. Better governance

4. In 2021, the UN Statistical Commission High-Level Seminar on Data Stewardship and the Role of National Statistical Organisations (NSOs) in a Changing Data Landscape recognised national statistical organisations have responded to the ever increasing data demands and rapidly changing data ecosystem by modernizing and integrating new data sources into regular data production, including private sector data. The Seminar also noted that many National Statistical Organisations have gone even further than reshaping data programmes to adapt to new circumstances – many also strengthened their role as coordinators² and custodians of data principles and data quality.
5. The 2021 Seminar built on a High-level Forum on Official Statistics discussion during the 51st (2020) session of the UN Statistical Commission on whether the concept of data stewardship is an approach NSOs can use to realize opportunities presented in the rapidly changing data ecosystem. The Seminar was held in conjunction with the launch of a 2020 review of data stewardship practices which identified, from case studies, three types of data stewardship governance arrangements:
 - (a) Data stewardship is an expanded mandate and function for the chief statistician and therefore the NSO (Ireland, New Zealand)
 - (b) The NSO contributes to or partners in data stewardship activities, but it is not their main mandate or function, and no single entity holds this role (Argentina, Colombia, Finland and Mexico)
 - (c) The chief data steward position and associated activities are carried out in another part of Government either politically appointed (France and Mexico between 2013 and 2018) or not (Estonia and the city of Los Angeles) and do not formally overlap with the NSOs mandate or functions.
6. Australia was not one of the country case studies in the 2020 review but falls more in the second type of data stewardship governance arrangement. The ABS contributes to and partners in data stewardship activities, but it is not mandated or our main function. The Australian Government's Department of Finance has policy responsibility for data.
7. The 2020 review identified three foundations of a data stewardship governance role, described below.

¹ Microsoft Word - UNSC52_background document_3a_Data stewardship_rev.docx

² Data stewardship is often posed in the context of coordinating official i.e., publicly funded bodies.

A. National data strategies

8. National data strategies are a foundation to support the role of data steward. They set the ground rules, expectations and aspirations for how data are to be organised and used within Government.

9. The Australian Data Strategy was released in December 2021 and sets the Australian Government's whole-of-government vision for data. The Strategy was delivered jointly by the then Minister for Employment, Workforce, Skills, Small and Family Business and the Minister for Superannuation, Financial Services and the Digital Economy and Women's Economic Security. The Department of Prime Minister and Cabinet (in its role as policy lead for data at the time the Strategy was developed) led its development. The Australian Statistician co-chaired the Working Group which helped shape the strategy and the ABS contributed extensively.

B. Open data movement

10. The open data movement is often associated with the creation of open data portals (data.gov).

11. The Australian Prime Minister and Cabinet Open Data webpage identifies multiple aspects to Australia's open data history including the Government's commitment to open data and data-driven innovation in the Australian Government's Public Data Policy Statement released in December 2015. The ABS was an active member of the UN Statistical Commission Working Group on Open Data along with Statistics Sweden, Statistics New Zealand and the Netherlands (CBS). Responsibility for Australia's open data portal, data.gov.au, transferred to the Australian Bureau of Statistics in 2021.

C. Expertise, policy and practices in quality assessment and assurance of data

12. The *Essential Statistics for Australia* initiative (2010-2014) originated with a 2004 strategy review for the ABS to set out to identify, in a highly consultative way, the key national datasets that are the essential indicators of the state of the nation, regardless of which organisations produce them. The initiative led to the identification of essential statistics (and their data sources and data custodians), an assessment of their quality, and identification of essential statistical infrastructure underpinning the datasets and statistics. The ABS was also instrumental in the establishment of guidelines for Data Integration Projects involving Commonwealth Data for Statistical and Research purposes in which quality assessment and data assurance are fundamental. Many of the guidelines have now been codified in legislation via the Data Availability and Transparency Act (2022).

III. Better collaboration

13. The 2020 review of data stewardship practices also identified three types of collaboration between the NSO and the wider data ecosystem which can help foster data stewardship:

A. Coordination

14. The Australian National Data Advisory Council was established in 2019 and has now been codified in law by the Data Availability and Transparency Act (2022). Its role is to advise the (Australian) National Data Commissioner on using and sharing public sector data including on ethical data use, balancing data availability with privacy protection, trust and transparency, technical best practice, industry and international developments and community expectations. The Australian Statistician is a member of the Council. It has met eight times between March 2019 and September 2021.

B. Facilitation of others with the stewardship of their data

15. The ABS is leading the whole of Australian Public Service (APS) uplift in data capabilities. The Australian Statistician is the Head of the Australian Public Service Data Profession Stream, working closely with the Australian Public Service Commission and the APS Digital Profession Stream to lift the data capabilities of the APS workforce to generate deeper insights to inform decision-making in policy development, programme management and service delivery. The Stream is underpinned by a Data Professional Stream Strategy and is nearing the end of its second year. The ABS has been working across the system to develop the foundations and specific offerings, such as improved graduate recruitment outcomes, to encourage the uplift of both generalist and specialist data skills – for junior through to senior public sector roles.

C. Partnerships with other data providers to develop new products or enhance existing ones

16. The ABS accesses over 100 datasets for statistical and research purposes. These datasets are predominantly public sector data assets including birth and death registrations, taxation and welfare data. Data sharing arrangements are an essential part of data partnerships, and the ABS makes use of a range of data sharing arrangements – via legislation, memorandums of understanding and licencing arrangements. For many data sources, including taxation data, legislation enables the ABS to receive data for the purpose of the Census and Statistics Act 1905 (the Act), the main legislation for collecting, compiling, analysing and disseminating statistical data. The Act governs all ABS statistical releases and specifies confidentiality requirements. In addition to legislation, memoranda of understanding are also put in place specifying terms and conditions, including review points. Licencing arrangements are used for private sector data sources.

VI. Better methods and capability

17. Data stewardship can be part of transferring methods and capability available in the NSO to other parts of the National Statistical Service (NSS) and even the data ecosystem as a whole or in reverse, bringing needed methods and capability from outside the NSO, NSS or data ecosystem to within the NSO. Three areas were considered in the 2020 review of data stewardship practices.

A. Standards and infrastructure

18. Standards and infrastructure developed for official statistics have wider applicability and their dissemination and application across all data actors in the Government can be facilitated through a data stewardship approach.

19. ABS' statistical and data integration capability is complemented by a data access service known as the ABS DataLab. The DataLab allows sophisticated analysis of detailed microdata in a secure controlled environment. The DataLab currently services approximately 400 active projects and 4,000 registered users across government but also Australia's research sector, which is a strong support of this service. Use of the ABS DataLab is growing at about 30 per cent per annum.

B. Data science, machine learning and artificial intelligence

20. Building technical capability is one of five objectives of the ABS Methodology Strategy and within it, uplifting data science and machine learning capability for the ABS is a component. The uplift strategy aims to deliver efficiencies via automation of current ABS processes and enhancing the ability to make effective use of new and alternative data sources. Machine learning applications have been used by the ABS to predict dwelling occupancy status on Census night using administrative data, clustering of automatic grouping of similar

dwellings through analysis of smart meter data to inform COVID-19 impacts, an “Intelligent Coder” for automated coding of free text responses to standard classifications, and an Automated Image Recognition model for classification of aerial images in the maintenance of the ABS Address Register.

C. More integrated data - a common approach to data handling across the data ecosystem

21. The Australian Climate Service is a \$AUS210m investment to bring together Australia’s expertise and information to help local communities build resilience to climate change. The Service is a partnership between the ABS, the Bureau of Meteorology, the Commonwealth Scientific and Industrial Research Organisation (CSIRO), and Geoscience Australia. The ABS brings critical social and economic information to the Service, enabling a comprehensive picture of the vulnerability of geographical locations to help prevent or prepare for natural hazards, and the resilience of communities to a changing climate. ABS provides detailed information on who is potentially exposed to hazards and how vulnerable they may be.

22. The Service is currently developing a data strategy for the four key “domains” of built, natural, social and economic. ABS is leading work on the social and economic domains with Geoscience Australia. These strategies will inform priorities, identify data gaps, and be used to guide investment decisions.

V. Better access and understanding

23. The 2020 review into data stewardship practices identified four areas where there is potential for NSOs to expand their role in providing better data access and understanding of data in the ecosystem.

A. Geospatial visualization of the information

24. In May 2021, the Digital Atlas of Australia was introduced as part of the Government’s Digital Economy Strategy to position Australia as a world-leading digital economy and society by 2030. The Digital Atlas will be an interactive, secure and easy-to-use online platform that brings together a range of trusted national datasets for the first time. It will harness the power of near real-time and historic location-based data to put powerful insights at users’ fingertips. By using location as the connecting thread between currently disparate datasets, the Digital Atlas will be the next generation of the Australian Government’s location-based data infrastructure, based on a modernised Australian National Map. The ABS is a key partner in delivery of the Digital Atlas.

B. Facilitating more use of microdata

25. As mentioned earlier, ABS’ data integration capability is complemented by a data access service known as the ABS DataLab. The DataLab is also being trialled as a platform for data sharing across the Australian Public Service, initially with the Department of Finance using the ABS DataLab to enable secure data sharing and sophisticated data analysis. Increased demand for this service is expected with the newly enacted Data Availability and Transparency Act 2022 which seeks to better enable data sharing across the data ecosystem.

C. Data platforms and dashboards

26. Dataplace, an initiative of the Australian Office of the National Data Commissioner, is a new, whole-of-government platform to manage data sharing requests for the Australian Government. Dataplace will make it easier to discover and request access to data, including

under the Data Availability and Transparency Act 2022. Dataplace is under development and expected to be available from mid-2022.

D. Building data literacy

27. The Australian Public Service Data Professions stream, mentioned earlier, is a new approach to building data literacy in Australia. The initiative is developing a data culture within the Australian Public Service workforce.

VI. Looking ahead – data is the new black

28. NSOs are being encouraged and inspired to broaden their mandate and play the role of data stewards at different levels and with different arrangements to ensure the efficient utilization of all data sources while safeguarding data quality, confidentiality and security.

29. In this section, we present opportunities ABS is tapping into source data from all tiers of Government – Federal, State/Territory and Local. We are evolving how we integrate data with our existing data assets about people, locations, businesses and the physical environment to deliver faster, better and more useful insights. We are increasingly securing access to administrative and transactions datasets from the public and private sectors to produce timely, frequent, and high-quality statistics; reduce business and household survey response burden; and support policy development.

A. Multi-Agency Data Integration Project

30. The Multi-Agency Data Integration Project (MADIP) contains high-value, person-centred and regularly updated datasets that aim to comprehensively cover the Australian population. It is a secure, person-based research data asset that combines a broad range of information on health, education, government payments, income and taxation, employment and population demographics that can be used to answer complex social and economic policy questions. The data is provided by a range of agencies including the Australian Taxation, Education, Employment Services, Social Services, Health, Home Affairs and Services Australia agencies. There are currently over 200 projects across government and academia using MADIP.

B. Business Longitudinal Analysis Data Environment

31. The Business Longitudinal Analysis Data Environment (BLADE) integrates ABS business surveys data, business taxation data from the Australian Taxation Office (ATO) and other administrative data sources. BLADE includes data for businesses that have been active in the Australian economy at any time from 2001-02 to the current financial year.

32. Over the last few years, the data available in BLADE has been updated at high frequency and new data sets added. We are continually examining new data sources to ensure BLADE remains relevant and contemporary including improving the quality of location information at the most granular level to better support place-based policy and program development and research. BLADE and MADIP can also be linked to enable linked employer-employee analysis.

C. Justice Spine

33. ABS is developing a ‘Justice Spine’ in partnership with the Australian National Indigenous Australians Agency. It is a longitudinal national data asset linking police recorded criminal offenders in Australia’s criminal courts with adult prisoners in the corrective services systems. The dataset will show how persons move and interact within and across the justice system nationally, something that is currently not possible. The dataset will have potential to be linked to other Commonwealth and State/Territory held datasets for deeper

analysis of criminal offenders. It will be available to approved policy makers and researchers in late 2023 and will provide critical insights into patterns of offending and recidivism.

D. National Disability Data Asset

34. A National Disability Data Asset (NDDA) is under development and will deliver a significant new enduring national asset comprising a collection of linked, de-identified data from across multiple Commonwealth, State and Territory service systems on people with disability and their pathways through services. The NDDA will be enabled by the development of a new data integration infrastructure, known as the Australian National Data Integration Infrastructure (ANDII). The ABS, Australian Institute of Health and Welfare and Department of Social Services, working in partnership, are leading the development of ANDII and NDDA in collaboration with the Commonwealth, states and territories and the disability sector.

E. Australian Immunisation Register

35. This data integration project linked the Australian Immunisation Register (AIR) dataset and deaths data to MADIP to generate insights to inform the Australian COVID-19 Vaccine and Treatment Strategy. This project enabled timely analysis of vaccine uptake and outcomes across socio-demographic cohorts and geographic areas.

F. Tracking the labour market

36. The Labour Market Tracker Project integrated job-related data, including ATO's Single Touch Payroll data to MADIP and BLADE to enable close to real time monitoring of the labour market and the Australian economy during the COVID-19 pandemic.

G. Private Sector data

37. Over the past 12 months, the ABS has focussed on securing long term access to private sector transactions data for statistical purposes. Our activities include extending and enhancing existing data agreements with banks and establishing new agreements with major retailers as well as exploring a range of data relationships in other areas.

38. This work has enabled the:

- release of the monthly household spending indicator using bank transactions data. This new indicator covers 68 per cent of household spending compared to the long-standing Retail Trade Survey with coverage of about 30 per cent, and
- development of more frequent, less costly, and more detailed statistics on household income and expenditure using bank, large retailer, loyalty scheme and credit agency data.

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

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

VII. Conclusion

41. The Australian data ecosystem is a vibrant mix of partnerships, collaborations, data sources, capability uplift and legislation. ABS is constantly transforming and refreshing its approach to statistics and data. Data provided by the ABS has proved invaluable in informing

the policy response to the COVID-19 pandemic and, looking ahead, the pace and scale of change within the ABS is likely to be unprecedented.

42. ABS are focused on rising to the challenge of providing the information demanded of us and playing our part in Australia's ever evolving data ecosystem.
