15 October 2019

Office of the National Data Commissioner
By online submission

DATA SHARING AND RELEASE LEGISLATIVE REFORMS – BSA COMMENTS

BSA | The Software Alliance is grateful for this opportunity to make a submission on the Office of the National Data Commissioner's Data Sharing and Release Legislative Reforms discussion paper[1] (**Discussion Paper**).

BSA is the leading advocate for the global software industry before governments and in the international marketplace. BSA's members[2] are at the forefront of data-driven innovation that is fuelling global economic growth by helping enterprises in every sector of the economy operate more efficiently. As global leaders in the development of data-driven products and services, BSA's members have unique insights to share on how the Australian government can "unlock the full potential of public sector data"[3].

Our submission focuses on:

a. the distinction between "data sharing" and "data release";
b. the role of the National Data Commissioner; and
c. specific "data sharing" framework elements – accreditation, commercial uses, and Data Sharing Agreements.

A. "Data Sharing" vs "Data Release"

BSA is of the view that the distinction between "data sharing" and "data release" [4] is appropriate. We also agree with and commend the intention of the legislative reforms outlined in the Discussion Paper (**Legislative Reforms**), which is to create a framework that will enable government agencies to make available sensitive data that is not subject to the "open by default" policy set out in the Public Data Policy Statement (**PDP Statement**)[5], instead of being restricted to keeping such sensitive data entirely "closed" or in-house.

[1] As made available at: https://www.datacommissioner.gov.au/resources/discussion-paper.

[2] BSA's members include: Adobe, Akamai, Amazon Web Services, Apple, Autodesk, AVEVA, Baseplan Software, Bentley Systems, Box, Cadence, Cisco, CNC/Mastercam, DataStax, DocuSign, IBM, Informatica, Intel, MathWorks, Microsoft, Okta, Oracle, PTC, Salesforce, ServiceNow, Siemens PLM Software, Sitecore, Slack, Splunk, Symantec, Synopsys, Trend Micro, Trimble Solutions Corporation, Twilio, and Workday.

[3] As described on page 1 of the Discussion Paper.

[4] As described on page 3 of the Discussion Paper.

[5] As described on page 19 of the Discussion Paper.

We **recommend that** any new legislation adopted as part of the Legislative Reforms should aim to provide clarity and incentives for agencies to make as much public sector data available as possible. This would include:

- codifying the PDP Statement and having clear definitions for key concepts in the PDP Statement (e.g., "anonymised data") to provide the legislative backing and mandate for the Australian government's "open by default" policy, and encourage as much data as possible to be publicly shared under the "data release" framework; and

- incorporating appropriate protections and limitations of liability for government agencies that, in good faith, share public sector data, whether under the "data release" framework or the new "data sharing" framework.

B. Role of the National Data Commissioner

BSA agrees with and commends the overarching mandate for the National Data Commissioner to be "a champion and advocate for greater data sharing and release"[6] with oversight for the "data sharing" regime.

In addition to the duties and functions outlined in the Discussion Paper[7], we **recommend that** the National Data Commissioner should also:

- have responsibility for overseeing both the "data release" and "data sharing" frameworks, to allow for a coherent and integrated approach in implementing the frameworks, and to ensure that government agencies are making as much data as possible publicly available under the "data release" framework;

- undertake a review of existing barriers to government data sharing (including technical barriers that may arise due to government agencies using outdated technology), and explore opportunities for government agencies to use "tiered access" approaches to data protection[8], differential privacy frameworks[9], homomorphic encryption[10], and other cutting-edge technologies and data governance processes that can facilitate greater access to data while safeguarding the privacy of individuals;

- relating to the above, have a role in the Australian government's technology procurement policies, to advise government agencies on the modern-day systems required to collect,

---

[6] As described on page 13 of the Discussion Paper.

[7] At pages 13 and 39 to 45 of the Discussion Paper.

[8] A "tiered access" approach would enable agencies to make available public versions of otherwise sensitive datasets by, among other things, stripping out personal information and adopting "data minimization" strategies. For example, an eligible researcher's project might earn approval for access to sensitive public data at a highly secure research data center that requires expert review of all output. Another researcher's project may need only access to a data query tool that runs an analysis, checks for disclosure risk without ever showing individual records.

[9] See Kobbi Nissim, Thomas Steinke, Alexandra Wood, Micah Altman, Aaron Bembenek, Mark Bun, Marco Gaboardi, David O'Brien, and Salil Vadhan, *Differential Privacy: A Primer for a Non-technical Audience* (February 2018), available at: https://privacytools.seas.harvard.edu/files/privacytools/files/pedagogical-document-dp_new.pdf. ("Differential privacy is a strong, mathematical definition of privacy in the context of statistical and machine learning analysis. It is used to enable the collection, analysis, and sharing of a broad range of statistical estimates, such as averages, contingency tables, and synthetic data, based on personal data while protecting the privacy of the individuals in the data. . . . Computer scientists have developed a robust theory for differential privacy over the last fifteen years, and major commercial and government implementations have now started to emerge.")

[10] Homomorphic encryption is a form of encryption that allows a computational analysis of encrypted data, ensuring that the data remains confidential. Use of homomorphic encryption could, for instance, enable the sharing of aggregated medical data to facilitate AI research without risking patient confidentiality. See Jean Louis Raisaro, Jeffrey Klann, Kavishwar Wagholikar, Hossein Estiri, Jean-Pierre Hubaux, and Shawn Murphy, *Feasibility of Homomorphic Encryption for Sharing I2B2 Aggregate-Level Data in the Cloud*, AMIA Jt Summits Translational Science (May 2018), available at https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5961814/#.

store, and share data in a manner that is consistent with the overall objectives and safeguards in the "data release" and "data sharing" frameworks; and

- have the power to adjudicate disputes between applicants seeking access to public sector data under the "data sharing" framework and the relevant Data Custodians (as defined in the Discussion Paper), and to require that the Data Custodians share data where the applicant's application has merit.

To facilitate the discharge of the National Data Commissioner's duties and responsibilities, as contemplated above, we also **recommend that** any new legislation adopted as part of the Legislative Reforms should:

- include a requirement for government agencies to undertake regular internal audits of their data assets, and to make the audit reports available to the National Data Commissioner for evaluating the extent to which they are complying with the "data release" and "data sharing" frameworks; and

- maximise the opportunities for industry and citizen engagement and collaboration, by incorporating mechanisms through which the public can provide input to the National Data Commissioner on issues relating to "data release" and "data sharing".

C. Specific "Data Sharing" Framework Elements – Accreditation, Commercial Uses, and Data Sharing Agreements

i. Accreditation

The Discussion Paper contemplates that "the Office of the National Data Commissioner will accredit organizations from the public, private and research sectors as 'Accredited Data Service Providers' to assist Data Custodians to make decisions about data sharing under the Data Sharing and Release legislation"[11]. There is, however, some ambiguity as to whether this effectively amounts to an accreditation requirement for any service provider to the Australian government (e.g., a cloud services provider) that handles and/or processes data. We understand that this might not be the intent of the Australian government, and accordingly **recommend that** the scope of entities that need to be accredited under the "data sharing" framework should be clearly defined in any new legislation adopted as part of the Legislative Reforms.

ii. Commercial Uses

We commend the position taken by the Australian government that under the "data release" framework "[o]pen data can be used by anyone for any purpose, including by companies to analyse and develop products and services", and the Australian government's explicit recognition that "releasing open data for commercial use ensures we are making useful data available fairly to all entities, encouraging competition and innovation in the system".[12]

With respect to the "data sharing" framework, however, we note that the Australian government has not finalised its position on the commercial use of public sector data[13]. We **recommend that** the Australian government should adopt the same principle/rationale for the "data sharing" framework that underpins the "data release" framework, i.e., that making public sector data available for commercial use would encourage competition and innovation in the system. To the extent that there are concerns

[11] At page 44 of the Discussion Paper.

[12] At page 26 of the Discussion Paper

[13] As indicated on page 27 of the Discussion Paper.

over misuse of public sector data under the "data sharing" framework, these can be addressed through the accreditation of data users, strict requirements for compliance with the Purpose Test and Data Sharing Principles (as defined in the Discussion Paper), and other appropriate safeguards, including those that exist in existing laws to address concerns about privacy, intellectual property, consumer protection, and competition in the market[14].

iii. Data Sharing Agreements

We note that the Australian government has developed a template Data Sharing Agreement which it will be piloting with some government agencies. To ensure that the template appropriately balances the rights and obligations of the Data Custodian, the Data Service Provider, and the Data User, we **recommend that** industry inputs should also be sought on the template.

Additionally, we **recommend that**, to further encourage "data sharing":

- inconsistent data sharing and use terms and conditions adopted by different government agencies should be eliminated, by requiring all government agencies to implement the template Data Sharing Agreement; and

- the template should be aligned with models that are already in use internationally, such as the Linux Foundation's Community Data License Agreement[15] and the Open Use Data Agreement[16].

D. Conclusion and Next Steps

Thank you again for the opportunity to make a submission on this important topic of public sector data sharing.

BSA and our members would be delighted to further engage with the Australian government to respond to any questions and to explore ways in which BSA and our members can work with the Australian government and other stakeholders to develop effective and balanced regulatory policies for public sector data sharing.

If you require any clarification or further information in respect of this submission, please contact Mr Darryn Lim at [redacted]

BSA | The Software Alliance

[14] As also noted on page 27 of the Discussion Paper.

[15] The Linux Foundation Projects, *Community Data License Agreement*, available at: https://cdla.io/.

[16] Microsoft, *The Open Use of Data Agreement*, available at: https://github.com/microsoft/Open-Use-of-Data-Agreement.