Data Republic: Submission to National Data Commission

Data Availability and Transparency Bill exposure draft and Explanatory Memorandum

Data Republic Pty Ltd

Level 5, 11 York Street, Sydney NSW 2000

1. About Data Republic

An introduction

Data Republic is an Australian business founded in 2015 to provide a secure and controlled platform for high value data collaboration.

We work with governments, banks, airlines, retailers, telcos and insurers with sensitive data pertaining to citizens and businesses. Our customers engage us because they know that sharing data needs to be a conscious and auditable process. When sensitive data is made available for access by others, there are risks which need to be mitigated through the process of collaboration such as data breaches, re-identification attacks and secondary use of data.

With five years of experience in data sharing and collaboration for large enterprises and governments, we have learned a great deal and continue to do so. We have codified our learning in our products and market advice offerings, and always look forward to assisting organisations with some of the challenges in establishing and executing a data sharing project.

We aim to be an Accredited Data Services Provider (ADSP) to the Australian Government Departments that will participate in the data sharing scheme.

Accreditation Framework Discussion Paper

Data Republic is supportive of the Accreditation Framework proposed by the National Data Commission (NDC) and has provided a separate submission of commentary on the Framework.

2. Executive Summary

Data Republic is supportive of the Data Availability and Transparency Bill ('Bill') and congratulates the National Data Commission (NDC) on its efforts with the Bill. Our feedback focuses on four areas:

1. Whilst the Five Data Sharing Principles meet the needs of the Bill, the amount of trust engendered by them can be enhanced by NDC guidelines about the legal and audit related controls in place as under each Principle (which are often key concerns in a data sharing scheme). Furthermore, to avoid the overzealous and damaging use of the 'Data Principle', the NDC should provide example commentary of how to balance different dialed levels of each Principle to meet various project needs but retain data utility.
2. The inclusion of consent requirement in the Bill is an important factor in ensuring the trust of the data subject in the Data Sharing scheme. To ensure standardisation across Australia and familiarity of experience with the consumer, the management and acquisition of this consent should be aligned with an open source CDR Consent

Management Protocol. Data Republic also states the likelihood that a large set of applicable datasets will fit into the category of it being unreasonable or impracticable to gain consent (such as those in the health space). In these scenarios, if there is not clear guidance on the topic to complement the bill, highly conservative approaches will dominate, even when the outcome of data application is highly positive and valuable for citizens. To improve the value of the data sharing scheme, we recommend providing clear detailed guidance on optionality for data sharing with these types of data assets (e.g. de-identification).

3. A framework for Data Sharing Safeguards is designed effectively in the Bill, but its implementation needs to be supported through clear example driven guidance to truly enable a trusted Data Sharing Scheme.
4. A data sharing scheme is only as valuable as the use cases it supports. Data Republic recommends the NDC support the Bill with guidance on the ideation, evaluation and prioritisation of data sharing use cases.

We look forward to the passing of this Bill and to becoming an ADSP, supporting the implementation of the data sharing scheme and the value derived by the Australian public as a result.

3. Applying Data Sharing Principles is an art form

From the Five Safes to the Seven Controls

Data liquidity is often stifled by trying to strike a balance between risk and reward, with the typical outcome often being a risk averse ‘no’ in the face of a data sharing opportunity. This is the result of the stacked game of compliance, privacy, security, consent, social licence and regulation against data application to social good outcomes. As such, frameworks are required to equalise this, turning these factors into enablers for data sharing rather than inhibitors.

Clear data sharing frameworks such as the Five Safes used by the Australian Bureau of Statistics, and as adopted by the Bill as the ‘Data Sharing Principles’, work to provide guidance to data custodians to navigate pressure from risk, compliance and legal teams that often work against deriving utility out of sharing data. Data Republic utilises its own framework developed from the ‘Five Safes’ to the ‘Seven Controls’ which regard seven levers of control that can be dialed up and down by customers participating in data sharing and collaboration on Data Republic’s platform. Data Republic found that it needed to add two more dimensions to form the Seven Controls: legal and audit, from the existing Five Safes: people, use, security, data and output. The original Five Safes framework was focused on internal (or low sensitivity) data access, but when multiple external parties are involved (especially with or between corporates) it was noticed that there was an opportunity to, and need to, establish further trust with the legal and audit controls. These controls are centered around clarity of roles, responsibilities and recourse, and around monitoring, identifying and reporting respectively, and give organisations the confidence of defensibility of themselves in a worst case scenario.

Given the existing use of the Five Safes principles by the ABS and other governments globally, Data Republic notes it is unlikely that the Bill will deviate from these five principles. As such, we instead suggest making the role of legal arrangements and auditability of activity clearer as part of the existing five data sharing principles. An example of this is to identify the relevant legal components of each safe (e.g. data permitted use agreements under 'Safe Projects'). This should provide participating organisations with more trust in the defensibility of the data sharing principles.

The trap of 'Safe Data'

The original Five Safes framework developed by the UK Office of National Statistics only included four safes, and was missing 'Safe Data' as it was developed to provide safe access to highly detailed data (and so did not need the safe). The framework was extended to include 'Safe Data' to be more broadly applicable (which it achieved) however in doing so it introduced a new issue around the way users approach the balancing of their controls.

It is often the case that Data Republic must expand the range of possibilities in the minds of its customers on how to best utilise all of the Seven Controls, to avoid the trap of 'Safe Data'. This trap occurs as it is typically easy to filter (removal of rows or columns), aggregate (removal of granularity) or perturb (addition of noise) data with immediate and measurable impact in reducing key risks with data sharing (such as re-identification risk). The downside is that this immediately reduces the utility of the data being shared and so can reduce (or entirely mitigate) the outcomes of the sharing exercise.

Data Republic recommends that the NDC provide guidance in conjunction with the Bill on tangible examples of combinations of safe dials that seek to avoid the sinking of controls in the other four Principles. An example of this is as below where the four non-data Principles are dialed up further in order to enable only a small amount of data processing. This retains the data utility and minimises the impact to the final outcomes of the data sharing project.

Principle type	Dial level
Projects	MEDIUM ● Clear permitted use of data, with no secondary purposes allowed
People	HIGH ● Police checks ● User signs legal agreement
Settings	HIGH ● Technical environment setup ensuring only approved data is only added or removed ● All activity logged for auditing
Outputs	MEDIUM ● All outputs are checked for compliance

Data	MEDIUM ● Subject name, address and immediate identifiers removed, but otherwise no further data processing. ● Detailed data allowed

Recommendations

- Make the role of legal arrangements and auditability of activity clearer as part of the existing Five Data Sharing Principles
- Data Republic recommends that the NDC provide guidance in conjunction with the Bill on tangible examples of combinations of safe dials that seek to avoid the sinking of controls into the Data Principle

4. Use of consent

Optionality for impractical consent

The Bill explains that under the project principles that the “sharing of the personal information of individuals is done with the consent of the individuals, unless it is unreasonable or impracticable to seek their consent”. Data Republic considers it a likely scenario that government data assets will exist that contain personal information with no applicable consent for data sharing. It is impractical to gather that consent now, but there is clear value and applicable ethical use cases that would benefit the Australian public.

Whilst the bill consultation paper mentions “other safeguards outlined by the data sharing principles can be dialled up to protect privacy”, it is not apparent as to how this can be applied. In the absence of guidance on alternatives for impractical consent scenarios, data sharing scheme users will default to not sharing the data assets available to them, keeping their value locked up away from valuable research or service provision. An example of this is with health related datasets. The possible benefits from research on these is monumental, benefiting the public from potential new treatments and more. Yet if a clear consent framework was not in place when the data was collected which made allowances for sharing under the circumstances of this Bill, then the data may not yield the value it can. In this scenario, a clear dialing up across the Five Data Sharing Principles is the alternative, yet a complex one. How can a data sharing scheme user, especially one not entirely technical, be expected to appreciate when enough has been done to balance the benefits of the project outcomes against the lack of consent?

Data Republic recommends that guidance be provided, specific to this Data Sharing scheme, for a set of tactical options to still gain utility out of this data, whilst maintaining the ethical use of that data in the context of the given use case. This may include (but not limited to) the following:

- De-identification or Pseudonymization of data assets (and to what standard)
- Enhancement of technical security of the data (and to what standard)
- Enhancement of contractual security of the data (and to what standard)

Through working with customers in scenarios where consent obligations are unclear, Data Republic has successfully provided advice in the past on how to best navigate the various options available to companies, typically in a combination of the above three recommendations (with de-identification as a minimum). We have found that this approach means that value can still be created rather than avoiding the data share, whilst meeting the needs of compliance teams. Typically those driving of data sharing projects go on a learning curve to upskill their data sharing maturity. When this upskilling does not occur the opportunity is undermined from the outset, as they are unable to liaise effectively with more technical stakeholders (such as legal, risk, compliance). This recommendation seeks to avoid that.

The need for a universal consent framework

Earlier this year Data Republic made a submission under the 'Inquiry into Future Directions for the Consumer Data Right'. In that submission Data Republic provided recommendations around the development and implementation of a Consent Taxonomy and Consent Management Protocol as a critical layer in a Consumer Data Right (CDR) supported Australian society. In particular, Data Republic strongly advocated for the creation of an open source Consent Management Protocol built on a standardised taxonomy of use cases to enable both the codification and implementation of dynamic consumer consent at an API level.

With the opportunity CDR presents to create this Consent Management Protocol in the corporate space, it is similarly integral that the government under this Bill, adopts the same Consent Management Protocol. This would result in a ubiquitous, corporate and government aligned, default infrastructure layer for consent management that would drive consistency, adoption and education of consumers through repetition of experience.

Recommendations

- Data Republic recommends that guidance be provided, specific to this Data Sharing scheme, for a set of tactical options to still gain utility out of this data, whilst maintaining the ethical use of that data in the context of the given use case
- With the opportunity CDR presents to create this Consent Management Protocol in the corporate space, it is similarly integral that the government under this Bill, adopts the same Consent Management Protocol as in CDR

5. Data Sharing Safeguards as a value enabler

Data Sharing Safeguards are often considered only as a requirement to ensure the safety of the data custodian and data subject against the misuse of data. However, if wielded correctly, it is much more than that. The framework proposed in the bill, if applied correctly by users, can foster trust between the data subject, data analyst, data custodians, data recipient and regulators, and only from that mutual trust can data liquidity truly exist. If these parties do not work collaboratively together within this shared framework, points of failure

exist which make a data sharing scheme brittle. For example, if there is no trust between the data subject and data custodian about the methods used to ensure the safety of their data, then public scrutiny may force a regulator's intervention (perhaps after a data breach or the misuse of data within existing regulatory frameworks), essentially inhibiting the data custodian's ability to derive value from their data assets, and by extension so to the public (if the value was also in their interest). In this way, Data Sharing Safeguards can be an enabling factor for a data driven government (and broader data driven economy), helping government departments and organisations to overcome the dilemma of ethical data use through clear guard rails.

Lead by example

It is imperative users of the data sharing scheme are appreciative of the safeguards being set up and understand that it supports this foundation of trust, as otherwise they will be reticent to participate in the Data Sharing Scheme, reducing the value it is aiming to create. Data Republic has noticed that even with the existence of clear frameworks, a data sharing scheme user is always in search of applicability to their scenario in order to gain more confidence that the data sharing activity being conducted meets the necessary standards. This necessitates the development of example driven guidance on the framework in order to simplify interpretation and implementation of these safeguards in real scenarios. Only through a practical playbook of real data sharing scenarios, can a potential participant in the data sharing scheme properly interpret how they are to apply the Bill in their own context.

In 2019, the Infocomm Media Development Authority (IMDA) of Singapore released its 'Trusted Data Sharing Framework', with the aim to 'guide organisations through the data sharing journey and outline key considerations for organisations to take into account when planning data partnerships'. With content similar to that considered in the Bill, the framework at the time of publishing lacked specific guidance on implementation with practical examples, ultimately resulting in enterprises struggling to action the framework in their context. Data Republic has provided advice to IMDA and the Personal Data Protection Commission (PDPC) in Singapore, to add practical example driven advice to help with user interpretability of the Personal Data Protection Act (PDPA) in the context of Data Sharing between private sector bodies. Data Republic similarly calls on the NDC to do the same for this Bill and develop practical examples for various data sharing scenarios on how to correctly utilise the data sharing scheme.

A high level approach to explaining these scenarios should be included in the data sharing safeguards. A potential structure exists below for this explanation:

1. Explain the use case: Choose a real use case brought up by a potential user of the data sharing scheme that was raised in during the various consultation periods of the bill
2. Explain how the Bill's safeguards are applied for the scenario as in the example below

Bill Safeguard	Scenario specific details
Accreditation framework	• Create or use an existing generic accredited

	provider, and explain how the provider meets the accreditation standards
Data Sharing Purposes	• Explain the purpose of the agreement, and the steps taken to ensure the purpose meets the requirements of the bill and passes the test of the ethics committee
Data Sharing Principles	• Explain how each of the principles is applied in the scenario, and to what degree is each dialed up or down to manage risk for this scenario
Data Sharing Agreements	• Explain what legal agreement templates were used and why • Clarify the drafting process • Provide an example agreement(s) for this scenario
Transparency mechanisms & oversight	• Explain how transparency was ensured and maintained throughout the setup and execution of the scenario • Detail in what way the NDC was involved for oversight purposes.

3. Explain what various stakeholders were involved throughout the process and for what purpose
4. Explain the various stages of setting up and orchestrating the data sharing exercise

The breadth of scenarios developed should meet a reasonable number of common viewpoints with diversity in use cases in order to ensure at least one similar style scenario that the data sharing scheme user can relate to.

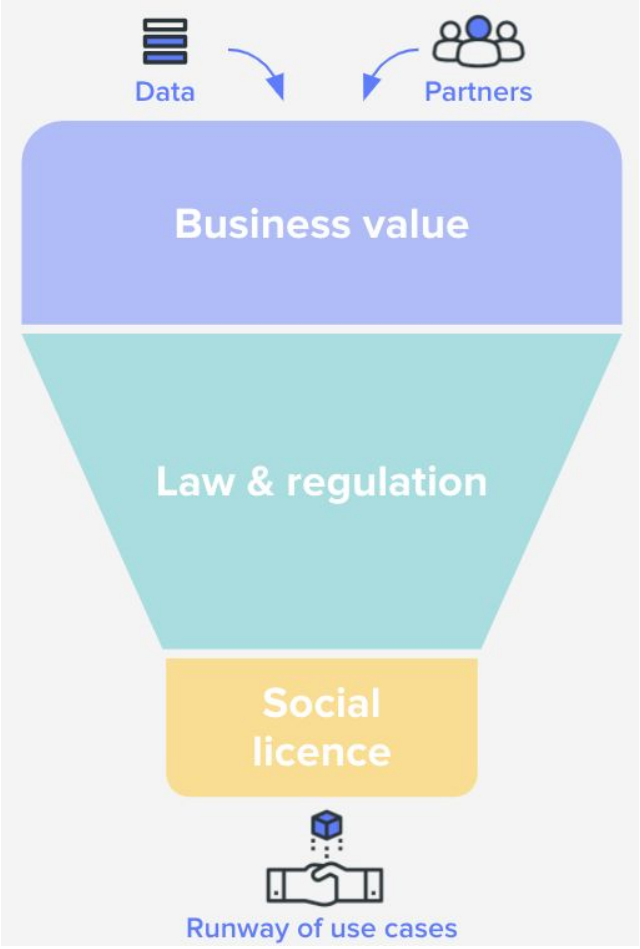
Recommendations

- The NDC should develop practical examples for various data sharing scenarios on how to correctly utilise the data sharing scheme so that scheme users can properly interpret how they are to apply the Bill in their own context

6. Coming up with data sharing ideas

Value derived from the data sharing scheme is only proportionate to the use cases that it can enable. As such, Data Republic recommends that the NDC support the ideation of use cases to maximise the benefit from its Bill.

Working with customers to develop data sharing use cases for the last five years, Data Republic has developed a framework to help consider use cases that we recommend. See the use case funnel below.



1. Top of funnel: Consider what Data assets are available and who is the recipient (and/or analyst)
2. Application Value: As there is typically always value to be derived from data sharing, it is best to consider application value by developing a suite of use cases in various tiers of value to the public
3. Law and regulation: What use cases require particular care with regards to regulation, and which are more clearer in interpretation. This includes what permutations of the project principles should I consider in order to make this project more easily and defensibly compliant without trading off application value
4. Social licence: Typically in consideration of an ethics committee or ethics framework, what are the expectations of the data subjects (and the Australian public) on the use of the data in question and the public good being derived
5. Runway of use cases: Ranking the use cases in consideration of these factors, devise a runway of use cases. Top of the list will be often be a combination of quick wins and long burn but immensely valuable data sharing projects

Recommendations

- To support users of the scheme to make the most out of it, the NDC should provide a helpful framework for use case ideation, evaluation and prioritisation (such as Data Republic's use case funnel)

7. Closing remarks

Thank you for the opportunity to submit on the Data Availability and Transparency Bill exposure draft and Explanatory Memorandum. Data Republic is supportive of the Data Availability and Transparency Bill ('Bill') and congratulates the National Data Commission on its efforts with the Bill. The four main recommendations made here are around devising the most complementary set of guidelines (for data sharing principles, safeguards, consent and use case ideation) and infrastructure (for consent) such that the Bill is made a success. With these guidelines and infrastructure as guardrails, the Bill will prosper to support a more data driven government and economy. Data Republic looks forward to becoming an ADSP to the

Australian Government Departments participating in the data sharing scheme, and so to becoming a part in making this Bill a success.

Please direct any follow-up questions or queries to enquiries@datarepublic.com.

Kind Regards,
Data Republic