Data Republic Submission to National Data Commission

Accreditation Discussion Paper

Introduction to Data Republic

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 or application by others, there are risks which need to be mitigated throughout the process of collaboration.

With five years of experience in working on data collaboration we have learned a great deal and continue to do so. Data Republic have codified some of our learning within our professional services offerings, and always look forward to assisting organisations with some of the challenges in establishing a data sharing project.

We aim to be an Accredited Data Services Provider (ADSP) to the Australian Government Departments who seek to avail of the oncoming Data Availability and Transparency Act.

Data Availability and Transparency Bill

Data Republic is very supportive of this Bill and commends the National Data Commission for its work on the Bill. We are submitting a separate submission on the Bill itself.

Aims of Accreditation Processes

Any Accreditation process needs to be carefully considered as it will, in effect, provide a barrier to the organisations and individuals that seek to support or draw insights from government data sharing. That is good thing in terms of protection of government data, but if over-done or expensive it could hinder government departments as they seek the required services to help them to share data.

Senior public servants, who will be leading data sharing programmes within departments, have a range of concerns about the actors in a data sharing project. They will also seek advice, services and technology to help them manage a range of risks implicit in data sharing projects. Accreditation would ideally serve their needs by providing a level of assurance and by delivering a suitable choice of suppliers.

Some of the assurances that a department might seek include:

User Accreditation

- Is this person or organisation likely to **behave as required** by my data sharing agreement with them?
- Do I have grounds for **trusting** them to have access to departmental data?
- Are they **aware of obligations** under the data usage legislation including the Privacy Act, and Data Availability and Transparency Act?
- Will they be capable of making appropriate determinations using data (i.e. **are they skilled**) so that departmental data is not identified as a source of erroneous insights?

Data Services Provider Accreditation

- Are they **skilled** in the area of de-identification of person-related data, and can they assist me with assessment of re-identification risk management?
- Will they enable me to **match the citizen data** in my custodianship, with citizen data held by other departments, authorities and institutions without requiring me to send personal contact data field outside my IT environment?
- Will they be able to support the process of data sharing based on their prior **experience**?
- Do I have grounds for **trusting** them to manage access to departmental data?
- Are they **aware of obligations** under the data usage legislation including the Privacy Act, and Data Availability and Transparency Act?
- Do they have **systems capability, security and control** that I can use to protect departmental data during sharing projects?
- Do they allow me to **control user privileges to access to both data sharing approval processes and data** within their system?
- Will they ensure that any **retained outputs from data access** projects are in line with the details set out in my data sharing agreement?
- Can I adjust the way that I **manage risks specific to each project** in the data sharing solution?
- Having executed one project with the services provider, can I **repeat and scale my operational approach** to sharing in order to ensure efficiencies?
- Where I require my data sharing agreement to become legally enforceable, do they offer a **ready-use legal framework for data sharing** purposes?

Accreditation can be designed to address specific needs and make data sharing easier, rather than simply following an existing, earlier pattern of approval used within government.

Data Republic recommends that Accreditation is designed on a 'fit for purpose' basis, with the needs of government departments firmly in mind. We suggest that Accreditation is made into a digital process as swiftly as possible in order to control Accreditation costs; and encourage participation so that effective choice can be offered to departmental leaders.

The National Data Commission has suggested a segmented approach, with two types of ADSP (data integration and data sharing specialists). We believe this segmentation might be found to usefully have many additional segments (over time). We further suggest a modular approach to assessment, that can be aligned to the services scope of the different ADSP types (and Accredited Users).

For example, in the case of ADSPs, the assessment could be built around topic segments resembling the below:

Assurance area	**ADSP Type 1 data integration**	**ADSP Type 2 data sharing**
Skill	RDBMS, uniqueness measures, de-identification	Data sharing project orchestration
Experience	Previous client project work in data integration, referees	Previous client work in controlled data sharing, referees
Trustworthiness	Prior breach, disputes, directorships, referees	Prior breach, disputes, directorships, referees
Aware of legislation	Knowledge scores across relevant staff	Knowledge scores across relevant staff
Systems capability	Type 1 data access can be via Type 2	Systems architecture and process maps
Information security	Type 1 data access can be via Type 2	Infosec survey responses

User access and privilege control	Type 1 data access can be via Type 2	Controls documentation
Output and retention control	Type 1 data access can be via Type 2	Extract check process, isolation of analytical systems from networks/web
Project risk management through seven safes framework	Type 1 data access can be via Type 2	Ability for departments to consciously apply ***Seven Controls**** ***framework*** to manage risk appropriately for each project
Scalability and repeatability	Type 1 data access can be via Type 2	Workflow for data sharing project approval and execution
Legal framework to apply	Type 1 data access can be via Type 2	Legal framework ready to use and written for data sharing
Auditability	Type 1 data access can be via Type 2	Records retained and available throughout project approval process, data provision, analytical phase and any approved data extracts (for retention by the user)

Seven Controls framework is based of the Five Safes Framework, with two additional controls that Data Republic has added to manage external data collaboration: *(i) Legal terms (ii) People (iii) Use (iv) Security (v) Data (vi) Output (vii) Audit.*

Data Republic has also considered the questions posed in the Accreditation discussion paper

1. What is considered to be an appropriate level of Australian ownership for an organisation to be eligible for accreditation?
51%

2. Should individuals acting on behalf of an Accredited Data Service Provider be accredited individually? If so, what might be appropriate arrangements?
If such individuals will have access to departmental data then they should either be individually accredited or become the responsibility of their ADSP (in a legal sense).

3. Are there circumstances when it should be mandatory to use an Accredited Data Service Provider for a data sharing project?
Given the risk to the reputation of the Commonwealth and its duty to citizens, all sensitive or high value data sharing exercises should use a means of securing data through an ADSP. We agree with the current NDC thinking that there are broadly three types of data: open data that can be made freely available and can be downloaded; sensitive, but under the right settings, sharable data than can be accessed but not retained by others; and highly confidential departmental data that should not be shared.

4. What would those circumstances be?
At a Minimum: Any data sharing exercise which involves data at a person level.
At an Ideal level: Any data sharing exercise which involves data which, if released to unforeseen parties, could provide insights into Australia and its citizens that damage the nation or its people.

5. Are there elements of data capability that should be given more or less weight in the accreditation process, i.e. making elements mandatory or optional?
ADSP is potentially a broad category and would ideally be considered at a more disaggregate level before designing Accreditation requirements.

For example:
ADSP Level 1 – able to de-identify person level information and assess the re-identification risk remaining in the attribute data, and can enable matching between citizen databases held by

different data custodians, without requiring the personal contact details to leave each custodian's IT environment.
ADSP Level 2 – able to securely store and provide access to person-level information that has previously been de-identified.
ADSP Level 3 – able to aggregate and provide access to data with cells combining data from more than five persons.

Levels of skill and requirement are another dimension to be considered.
For example:
ADSP Type A – data collaboration platform with governance workflow and isolated analytical environments for data access.
ADSP Type B – platform able to facilitate a pre-agreed data share, without approval workflows and other controls and audit facilities.
ADSP Type C – consultant able to support departmental work in planning a data sharing project.
ADSP Type D – data preparation specialist able to assist departments with readying data for sharing.

6. What elements would be most useful to Data Custodians to support their decision-making process when considering sharing and access to data?
Data custodians might like to think about how sensitive the data is that they wish to share and the state it is in, in order to select an appropriate ADSP partner(s).

In our experience, those leading data sharing exercises are interested in a clear statement of the project outline and scope, a very specific definition of the permitted use of their data and level of detail and coverage of the data requested, and an understanding of who will analyse the data and who will see the results. This is because the range of specialist review and approval of a data share may extend to:
- **Legal and privacy compliance**
- **Information security teams**
- **Ethics and sustainability teams**
- **Commercial and reputational risk stakeholders**

7. Should the accreditation process recognise other frameworks, standards or processes that have assessed an element of data capability? If so what standards/processes might be appropriate to recognise?
- **CAIQ style infosec assessments**
- **International standards met such as ISO 27001 and SOC2 accreditation**

8. Are there any elements of data capability that should be captured in order to understand an accredited entity's ability to keep data safe?
Data management would ideally be considered across several segments so that departmental leaders can focus on the aspects that they are more concerned about.

Preparation of data
- **Data transformation**
- **Data quality and correction**
- **Data de-identification**
- **Re-identification risk assessment and mitigation**

Approval workflow
- **Ability to negotiate an agreement in a recorded environment**
- **Align approval of projects to your responsibility matrix**
- **Control of data access to project details that have been agreed**

Access and control of data
- **Data sharing technology and security**

- **Control over application of unforeseen data to a project (that could change data risks)**
- **Data access mechanisms that remove the possibility of retention by the user**
- **Extract checks and controls in the case that retention of any derived data is permitted by the sharing agreement**

Legally binding data sharing agreements

- **Ready to apply data sharing legal frameworks**
- **Ability to seek redress if an agreement is breached**

Audit and record keeping

Ability to review data sharing approvals
Ability to know where data is being accessed at any point in time
Review of queries made against data in the case of a possible misuse

9. What is a reasonable period of time to assess an application?
4 to 8 weeks

10. Are there further ways we can streamline the accreditation process?
Consider segmented and modular approaches to accreditation (i.e. accredited to do X, but not Y or Z).

Ensure an online accreditation process, with score-based assistance to assessors.

11. Do the timeframes to renew accreditation, every 5 years for Accredited Data Service Providers and every 3 years for Accredited Users, seem reasonable?
Yes

Question:
12. Is it appropriate to notify parties to Data Sharing Agreements of an accredited entity's suspension?
Yes.

13. Is there any information that must, or must not, be made publicly available through the registers of accredited entities?
Detailed responses to accreditation questions may include commercially sensitive information that should not be shared.

14. Is there any information that should be made available to Data Custodians through the registers of accredited entities?
Yes, suggest segmented approach to accreditation so that data custodians in departments are better equipped to select appropriate capability to their requirements.

15. Is charging a fee for accreditation, such as a renewal fee, reasonable?
Cost recovery is reasonable. The difficulty with a large fee is that data custodians might be deprived of options to move forward with data sharing projects if too few, or if only a narrow range of specialisms among organisations is available among organisations that are accredited.

Many thanks for your consideration of our submission.

Data Republic