



PRIVACY IMPACT ASSESSMENT


Cloud DataLab

June 2020



TABLE OF CONTENTS

Executive Summary	4
1 Description of the project	8
1.1 Background	8
1.2 Does the project involve personal information?	9
1.3 Scope and purpose of the PIA.....	10
1.4 PIA methodology and consultation.....	10
2 Technology and cloud services	11
2.1 Service provider	11
2.2 Cloud DataLab solution design.....	11
2.3 Hosting and data centres	12
2.4 Security testing	14
2.5 Security certification	14
3 Information flows.....	15
3.1 Information flows – microdata	15
3.2 Information flows – data about Users	16
4 Privacy impact analysis	17
4.1 Handling personal information	17
4.2 APP compliance.....	18
APP1 – open and transparent management of personal information	18
APP2 – anonymity and pseudonymity.....	19
APP3 – collection of solicited personal information.....	19
APP4 – dealing with unsolicited personal information.....	20
APP5 – notification of the collection of personal information	21
APP6 – use or disclosure of personal information (Not fully compliant – action required).....	22
APP7 – direct marketing	22
APP8 – Cross border disclosure	22
APP9 – adoption, use or disclosure of government related identifiers.....	24
APP10 – quality of personal information	25
APP11 – security of personal information	25
APP12 – access to personal information	27
APP13 – correction of personal information	27



5 ABS Response and next steps	27
Appendix A: Glossary and Acronyms	28
Appendix B: Application of the Five Safes	30
Appendix C: Information flows	31
Microdata.....	31
Data about Users	31

EXECUTIVE SUMMARY

The ABS DataLab is a secure analytics environment that provides safe access to de-identified microdata. Access is provided to authorised researchers for approved projects so they can conduct research aimed at informing the development of social, economic, and environmental policy priorities. The DataLab is an important part of a suite of controls designed to minimise the risk of inappropriate disclosure of information through providing a safe environment to access microdata.

The ABS Cloud DataLab project represents an evolutionary shift towards scalable and flexible architecture to meet future demand as the number, scope and complexity of research projects increases. It will replace the existing DataLab with a Cloud hosted platform. The Cloud DataLab is being designed to manage the same types of data as the DataLab, for the same types of researchers, and using largely the same set of processes and controls currently in place for the DataLab. The key difference will be hosting the microdata and access platform in the Cloud.

Microdata is the unit record data that provides detailed information about people, households, businesses or other types of entities. The microdata stored in the DataLab for authorised researcher access is “unidentified” data, as it has had all names, addresses and other direct identifiers removed as well as some other changes made to the data to reduce the risk of re-identification. Further controls are applied to the unidentified microdata (through the Five Safes Framework) before it is accessed in the DataLab, which means that researchers only access “de-identified” data (that is, data which is no longer about an identifiable individual or an individual who is reasonably identifiable¹).

While the Cloud DataLab will be used for an existing function (that is, secure access to de-identified microdata by researchers), the move to cloud infrastructure for microdata access represents a significant change in how the ABS delivers that function. For this reason, following the Office of the Australian Information Commissioner (OAIC) guidelines for conducting Privacy Impact Assessments (PIAs)², the ABS has prepared this PIA for the Cloud DataLab.

This PIA analyses possible privacy impacts of moving to a Cloud DataLab and identifies as well as recommends options for managing, minimising or eliminating privacy impacts. This PIA explores:

- Technology - Potential privacy impacts of a Cloud hosted DataLab (section 2)
- Information flows - How the ABS collects, holds, manages, and discloses microdata and information about users of the Cloud DataLab (Users) (section 3)
- Compliance with the Australian Privacy Principles (APPs) (section 4)
- ABS response and next steps (section 5)

A detailed PIA on one of the key person-centred microdata sources in the DataLab, the Multi-Agency Data Integration Project (MADIP), was published in November 2019. The MADIP PIA Update³ includes information on the ABS DataLab and the ABS’ application of the Five Safes Framework, as

¹ See definitions of unidentified and de-identified in Appendix A: Glossary and Acronyms.

² OAIC Guide to undertaking privacy impact assessments <https://www.oaic.gov.au/privacy/guidance-and-advice/guide-to-undertaking-privacy-impact-assessments/>

³ MADIP PIA Update, November 2019

<https://www.abs.gov.au/websitedbs/d3310114.nsf/home/statistical+data+integration+-+madip+consultation>

international best practice for managing disclosure risk, in providing researchers with access to microdata.

This PIA examines two changes in processes for storing and accessing microdata that will be introduced in the move to a Cloud DataLab. It covers the storage of microdata as well as the collection, storage and use of personal information relating to Users (both authorised researchers and ABS staff) of the Cloud DataLab in a cloud platform.

This PIA has assessed the handling of personal information (as defined under the Privacy Act) in the Cloud DataLab and confirms that:

- The unidentified microdata stored in the Cloud DataLab may include some personal information.
- Authorised researchers access only de-identified data which does not contain personal information. This is because the significant suite of protections in place in the Cloud DataLab and access arrangements make the individuals to whom the data relates not reasonably able to be identified.
- Cloud DataLab access processes involve collection, storage and use of the personal information of Users.

As access to de-identified microdata by authorised researchers does not involve access to personal information, further assessment of the compliance of this activity against the APPs is not required.

In terms of the storage of microdata, analysis against the APPs found the Cloud DataLab to be compliant, but with two best practice recommendations relating to APP1 and APP11. It recommends that the online materials be updated to provide information about the storage and security of microdata in the Cloud DataLab, including use of cloud infrastructure. It also recommends that ABS implement any outcomes arising from security assessments to assure the continued security of microdata.

This PIA found the Cloud DataLab processes for collecting, storing and using the personal information of Users of the DataLab in the Cloud to be compliant with ten of the thirteen APPs, but with a best practice recommendation made to improve transparency around the management of User information (APP1). For three APPs (APP5, APP6, APP11), action is required to ensure transparency around the collection, use and disclosure of the personal information of Users as well as the security of this information.

A summary this analysis and recommendations is provided at Table 1.

Table 1: Summary of recommended actions

APP	Compliance	Commentary	Best practice recommendations
Storage of microdata in the Cloud DataLab			
APP1 – open and transparent management of personal Information	Compliant, but further action recommended	ABS is committed to open and transparent management of the microdata that is stored in the Cloud DataLab. Relevant information is provided on the ABS Privacy pages on the ABS website. ABS plans to update information on the website about storage and security of microdata.	<i>R1: Update online materials to provide information about the storage and security of microdata in the DataLab, including use of cloud infrastructure.</i>
APP11 – security of personal information	Compliant, but further action recommended	The ABS regularly reviews and takes reasonable steps to ensure the security of microdata.	<i>R5: Implement any outcomes arising from security assessments to assure the continued security of microdata in the Cloud DataLab.</i>
Collection, Storage and Use of personal information of Cloud DataLab Users			
APP1 – open and transparent management of personal information	Compliant, but further action recommended	The collection, use and disclosure of personal information about Cloud DataLab Users should be clearly explained to them before they use the system.	<i>R2: Clearly describe the collection, storage, use and disclosure of personal information about Users in collection notices and/or user agreements.</i>
APP5 – notification of the collection of personal information	Not fully compliant, further action required	A formal collection notice (APP5 notice) should be developed to inform Users of the Cloud DataLab how their personal information will be collected, stored, used and disclosed.	<i>R3: Create an APP5 notice to Cloud DataLab Users to make them aware of how their personal information will be used, including that it will be stored on a Cloud-based service, used by an off-shore service provider, and disclosed to Data Custodians. Take additional steps to ensure all Users are made aware of the collection notice.</i>
APP6 – use or disclosure of personal information	Not fully compliant, further action required	While there is no secondary use of personal information about Cloud DataLab Users, there is a disclosure of personal information to Data Custodians as part of the primary use. This use could be clarified in the APP5 notice.	<i>R4: Ensure the APP5 notice notifies Users that their personal information may be disclosed to Data Custodians in order for them to approve access to microdata.</i>

APP11 – security of personal information	Not fully compliant, further action required	The ABS regularly reviews and takes reasonable steps to ensure the security of personal information of Cloud DataLab Users.	<p><i>R6: Implement any outcomes arising from security assessments to assure the continued security of personal information of Cloud DataLab Users.</i></p> <p><i>R7: Create a deletion and retention policy specific to DataLab User accounts and related personal information.</i></p>
--	--	---	--

This PIA articulates the ABS' ongoing commitment to protect privacy and ensure data security when providing access to microdata and handling the personal information of Users. The Cloud DataLab will continue to adapt and evolve to meet user expectations, methodological and technological advancements, and other environmental changes. The ABS is also continuously improving data handling practices and infrastructure, including for the Cloud DataLab, to preserve privacy, ensure data security, and increase data quality and utility.

The ABS will publish a progress report for this PIA on its website within one year of this report being published, to inform on progress of implementing APP recommendations.

1 DESCRIPTION OF THE PROJECT

1.1 Background

The ABS currently provides secure access to de-identified microdata for approved projects and authorised researchers⁴ in the ABS DataLab. The DataLab is currently hosted in the ABS tenancy of a Canberra based Secure Data Centre. The ABS is authorised by legislation to release ABS microdata under certain conditions and carefully manages access through the Five Safes Framework⁵. This is an internationally recognised framework for making effective use of data while controlling for risks using a number of levers – safe people, safe projects, safe settings, safe data and safe outputs.

The ABS DataLab was initially developed to provide secure access to ABS Confidentialised Unit Record Files (CURFs). As the ABS developed more detailed microdata products, including integrated data products such as from MADIP⁶ and the Business Longitudinal Analytical Data Environment (BLADE)⁷, the DataLab became the primary access method for those products. However, the DataLab was not designed for the very large files created through MADIP or BLADE, nor was it designed to manage the number of authorised researchers that the ABS is currently supporting.

While the existing DataLab tools and methods were fit for their original purpose, a move to newer technologies and systems is required to improve performance of the DataLab. The Cloud DataLab project will replace the existing DataLab with a Cloud hosted platform. The Cloud DataLab is being designed to manage the same types of data as the existing DataLab, for the same types of researchers, and using largely the same set of processes and controls currently in place for the DataLab. The move to Cloud also provides additional benefits for ensuring the security of microdata when providing access to authorised Users (as discussed in relation to APP11 below).

A detailed update to the PIA of MADIP, which is one of the main microdata assets accessed in the DataLab, was published in late 2019. It includes information on the ABS DataLab and the ABS' application of the Five Safes Framework in providing Users with access to de-identified microdata.

Appendix A provides a glossary of terms and acronyms used in this PIA.

⁴ ABS applies a rigorous assessment and approval process to applications for access the DataLab. For a project to be approved, the ABS and the Data Custodians (the agencies that collect the data) must agree to the proposed use of the data. The project must be assessed as being in the public interest and be in accordance with the legislation of the relevant agencies. All Users are legally obliged to use data responsibly for approved purposes, comply with the conditions of access, and maintain the confidentiality of data.

⁵ More information on the Five Safes, and managing the risk of disclosure using the Five Safes Framework refer to the ABS Confidentiality Series: <https://www.abs.gov.au/ausstats/abs@.nsf/mf/1160.0>

⁶ <https://www.abs.gov.au/websitedbs/d3310114.nsf/home/statistical+data+integration+-+madip>

⁷ [https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Statistical+Data+Integration+-+Business+Longitudinal+Analysis+Data+Environment+\(BLADE\)](https://www.abs.gov.au/websitedbs/D3310114.nsf/home/Statistical+Data+Integration+-+Business+Longitudinal+Analysis+Data+Environment+(BLADE))

1.2 Does the project involve personal information?

The Privacy Act⁸ defines “personal information” as:

Information or an opinion about an identified individual, or an individual who is reasonably identifiable...

There are two overarching types of data involved in this project: microdata and data about Users of the Cloud DataLab.

Microdata

As described earlier, microdata is the unit record data that provides detailed information about people, households, businesses or other types of entities. It includes data created from ABS surveys and the Census, person centred and business centred administrative data, and integrated data.

The microdata stored in the DataLab is unidentified data. It has had a number of treatments applied to reduce the risk of re-identification of the information, but it may still contain some personal information because of a risk of re-identification of individuals from the unidentified data. Before storage in the DataLab, all microdata is carefully prepared including taking measures to prevent the spontaneous recognition of individuals such as:

- Removal of all directly identifying information such as names and addresses.
- Application of a number of different confidentialisation methods, such as top-coding and grouping.
- Checking for records with remarkable combinations of responses, and possibly altering them slightly to ensure individuals cannot be identified. These treatments are described in more detail in the ABS Confidentiality Series, and the MADIP PIA.

The ABS has a legislative requirement⁹ for the Australian Statistician to release microdata “in a manner that is not likely to enable the identification of an individual”. Further controls on accessing the microdata (through application of the Five Safes Framework) are in place which means that researchers only access de-identified data. The Cloud DataLab will not alter most of the existing Five Safes protections. The main change is to the “safe” settings. As with the current DataLab, the Five Safes will continue to work together to ensure microdata is provided to Users in the Cloud DataLab in a manner that means it is de-identified and does not include personal information. Further detail about the implementation of the Five Safes in the DataLab environment is provided in Appendix B.

Data about Users

Data about Users includes information about researchers/analysts who are applying to access or have been granted access to microdata in the Cloud DataLab as well as ABS staff in Auditor, Administrator, or Systems administrator roles. Cloud DataLab access processes will collect, store and use information about Users including some personal information (such as their name, employer, and contact information).

⁸ <https://www.oaic.gov.au/privacy/the-privacy-act/>

⁹ See section 15 of the *Census and Statistics (Information Release and Access) Determination 2018*



1.3 Scope and purpose of the PIA

The OAIC guidelines¹⁰ recommend a PIA is undertaken for any project which will change how personal information is handled or stored; or for any changes an agency proposes which will:

...substantively change an existing activity or function. This includes a substantive change to the system that delivers an existing function or activity.

While the Cloud DataLab will be used for an existing function (that is, secure access to de-identified microdata by authorised researchers), the move to cloud infrastructure represents a change in how the ABS delivers that function. In addition, the move to cloud technology needs to be examined to ensure the public’s confidence and trust in ABS’s data handling practices, particularly in relation to the security of the data, is maintained. For these reasons, the PIA explores:

- Potential privacy impacts of a Cloud hosted DataLab, including storage of microdata in the Cloud-based system.
- How the ABS collects, manages, and stores information about users of the Cloud DataLab (such as managing Users’ names, addresses, passwords).
- Compliance with the APPs.

1.4 PIA methodology and consultation

The ABS has undertaken this PIA, following the ten-step process recommended by the OAIC. This process included consultation with groups outlined in Table 2.

Table 2: Stakeholders included in the PIA consultation process

Stakeholder	Date of consultation	Feedback
MADIP targeted stakeholders	Mid to late 2019	As part of the MADIP PIA Update consultations, stakeholders were asked for broad perceptions about the Cloud DataLab. Stakeholders were reasonably comfortable with the idea of moving the ABS DataLab to a Cloud-based system with data storage in Australia. While no significant concerns were raised, some stakeholders asked for additional information about security in the Cloud. These points are reflected in the content of this PIA (particularly response to APPs 8 and 11).
MADIP partners (see the MADIP PIA Update for more information)	Mid to late 2019	As part of the stakeholder consultations, the MADIP Board were also consulted about the potential to develop a Cloud DataLab and storage of microdata in the Cloud. They sought further information about the details of the proposal, similar to the questions raised by other stakeholders.

¹⁰ <https://www.oaic.gov.au/privacy/guidance-and-advice/guide-to-undertaking-privacy-impact-assessments/>



Additional stakeholder consultation and communication will be undertaken during the development and implementation of the Cloud DataLab, in particular, to understand Data Custodians' views on hosting their data in the Cloud DataLab and address questions about how their data will be managed.

As an additional measure, ABS engaged external privacy advisers (Maddocks) to independently review the PIA and to provide assurance that all relevant matters were considered. While Maddocks reviewed drafts of the PIA and provided comments to ABS, the analysis, findings and recommendations in the PIA are those of the ABS.

2 TECHNOLOGY AND CLOUD SERVICES

2.1 Service provider

The ABS has engaged Microsoft as the service provider for the Cloud DataLab. Microsoft are a trusted provider of cloud storage and analytics services for the Australian Government.

Microsoft has an overarching agreement with the Australian Government to provide services—the “Microsoft Business and Services Agreement (VSA4) between the Commonwealth of Australia as represented by the Digital Transformation Agency and Microsoft Ireland Operations Limited”.

The ABS has engaged Microsoft to provide services for the Cloud DataLab through the VSA4 agreement and subsidiary bilateral agreements. These collective agreements enforce contractual obligations on Microsoft and the ABS (the Microsoft Agreement).

The Microsoft Agreement specifically acknowledges that in performing the services under the Agreement, Microsoft is a “contracted service provider” under the Privacy Act and impose the obligation that in performing the services under the Microsoft Agreement, Microsoft must comply with the APPs as if it were the ABS.

“Microsoft Azure” is the collective name given to a group of cloud services, developed and managed by Microsoft, and hosted in secure data centres. The Cloud DataLab will make use of a range of these services.

2.2 Cloud DataLab solution design

The Cloud DataLab system uses hosted virtual machines and cloud storage services to provide secure, isolated research spaces for the analysis of microdata.

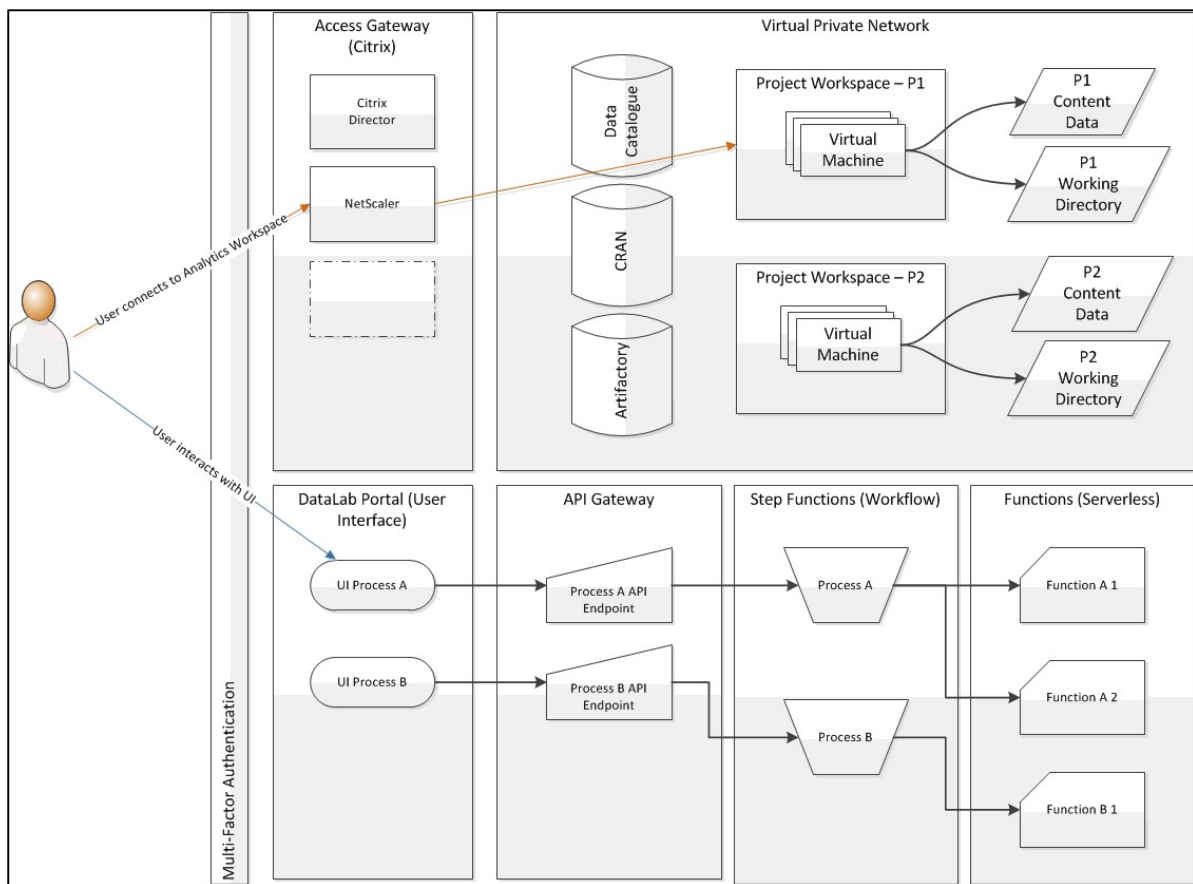
The technology underpinning the Cloud DataLab includes seven components:

1. A front-end web user interface (UI) to manage User and project registrations along with common management functions.
2. Virtual machines (VMs) running Windows along with a range of analytics software (including R and Python).
3. Azure Storage Accounts, to securely hold individual research products and allow querying from authorised researchers.

4. A back-end automation and orchestration system to control the provision of VMs, storage, networks, and User access.
5. A Citrix Cloud platform-as-a-service remote access management system, to allow researchers to securely connect to VMs.
6. A user-directory service in Azure Active Directory, to store and serve User authentication and authorisation information.
7. Ancillary services to provide security, billing, and disaster recovery control and monitoring.

A depiction of the solutions architecture for the Cloud DataLab is included in Figure 1.

Figure 1: Cloud DataLab Solutions Architecture



2.3 Hosting and data centres

For the Cloud DataLab, the Microsoft Agreement ensures the ABS retains effective control of all microdata and User information, and can access, change or remove data and information at any time.

Microdata

All microdata in the Cloud DataLab will be stored in Australian data centres and the ABS will retain full access controls for this data.

All microdata stored in the Cloud DataLab will be encrypted at rest. This provides a strong mitigation against unauthorised access to microdata, by ensuring microdata which is not accessed through the appropriate approved channels is securely encrypted.

Microsoft and its data centre employees will not be permitted access to microdata. This is required by the Microsoft Agreement.

The “Privacy in Microsoft Cloud Services” statement (the Statement)¹¹ provides details on Microsoft’s privacy principles and privacy standards, which guide the collection and use of customer and partner information. The Statement confirms that:

Microsoft contractually commits to not disclose customer data to any third party (including the individual) except upon instruction or permission by the customer, or if required by a lawful demand. Microsoft also regards customer data as confidential.¹²

The Global Privacy Information section of the Statement relating to Australia includes an assessment of Microsoft’s position against each of the APPs.

Personal information about Users

Personal information is collected from Users as part of creating a Cloud DataLab account. This information is required in order to verify the identity of Users, as part of the process to ensure ABS is safely managing access to microdata.

Very few of the Microsoft Azure services will use, manage or store personal information about Users. The Azure Active Directory and three Security services are the few that do. Some of these services are not wholly hosted in Australia. Additional information on these services is included below.

Azure Active Directory (AAD)

The AAD operates from an Australian hosted data centre (as of May 2020). The AAD will be used to store both ABS and external user accounts, passwords and authorisation information. As such, this service uses personal information about Cloud DataLab Users.

The Azure Active Directory does not use, manage, or store microdata.

Various Microsoft Security Services

Three globally hosted security services will access information from Log Analytics stores which are hosted within Australia.

These services scan and monitor the Cloud DataLab logging stores for events, such as a User logging in or out. The services check for unusual or inconsistent events in the logs, but do not retain any data or personal information other than as records of suspected threat activity.

These services also do not use, manage, or store microdata.

¹¹ Privacy in Microsoft Cloud Services <https://aka.ms/MCSPrivacy>

¹² Lawful demand relating to the CLOUD Act is discussed in analysis of APP8.

2.4 Security testing

Penetration testing

Various elements of the Cloud DataLab design have undergone penetration testing throughout the development phase, including the virtual desktop infrastructure. The tests were performed to identify any system vulnerabilities, including the potential for unauthorised parties to gain access to the system, and strengths.

This penetration testing has been independently conducted and the ABS will act on all recommendations from this testing prior to the final production release of the Cloud DataLab.

IRAP assessments

An Independent Security Registered Assessors Program (IRAP) assessment of the entire Cloud DataLab platform will be completed before the final production release of the Cloud DataLab. The IRAP assessment will evaluate the implementation, appropriateness and effectiveness of the Cloud DataLab security controls.

These assessments will be conducted in stages. The first stage relating to the Microsoft Azure services has been completed. The second IRAP will assess the Cloud DataLab design.

The ABS Chief Security Officer will consider the recommendations from these IRAP assessments and approve actions in response to ensure the most appropriate risk management approach for the data and information that ABS is responsible for.

The ABS is aiming to achieve a PROTECTED security rating from these assessments.

2.5 Security certification

Microsoft Azure has been placed on the Certified Cloud Services List (CCSL) by the Australian Signals Directorate (ASD). This means that it has been assessed as meeting a PROTECTED level of security, as defined in the Australian Government Information Security Manual (ISM).

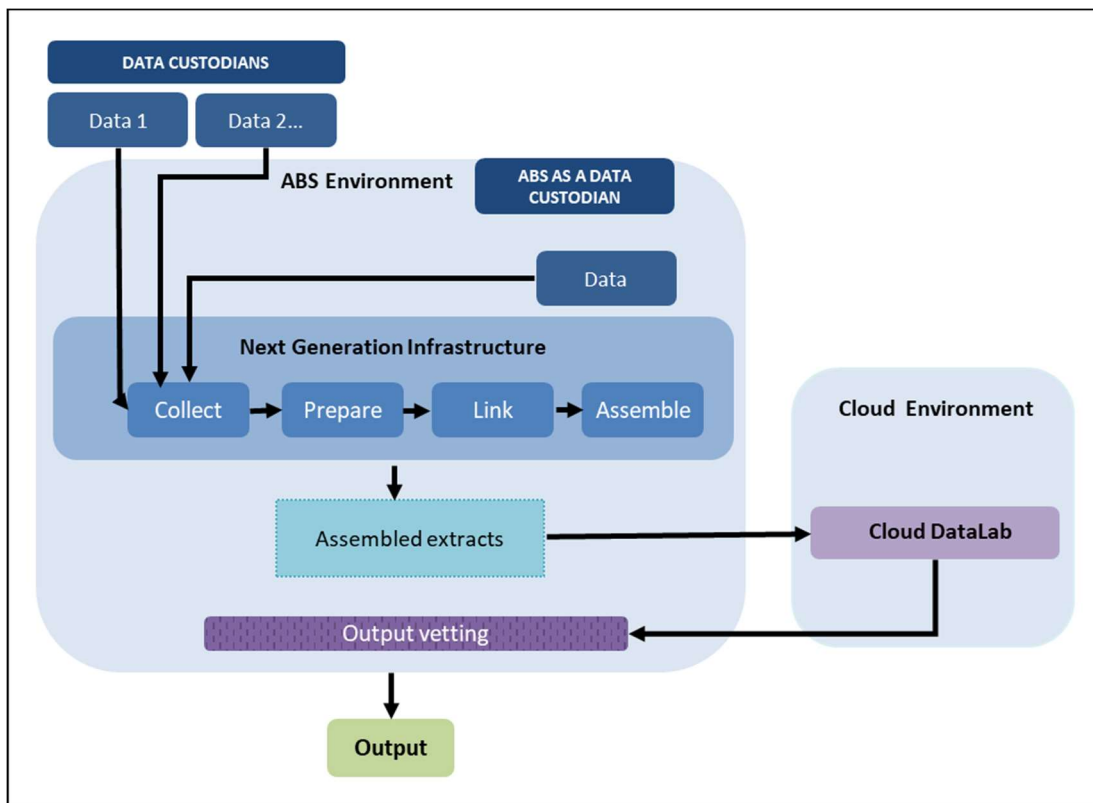
While existing microdata is not PROTECTED and Cloud DataLab authorised researchers do not need to have a Baseline Security Clearance, the ABS is aiming for a PROTECTED status to enable future storage of and access to PROTECTED microdata in the Cloud DataLab. While not required under the APPs, achieving a PROTECTED rating for the Cloud DataLab is consistent with the ABS commitment to carefully manage access to microdata under the Five Safes Framework; in this case, a commitment to ensuring the “safe” setting of the Cloud DataLab exceeds the security rating required for microdata (rated Unclassified DLM).

3 INFORMATION FLOWS

3.1 Information flows – microdata

As discussed earlier, before microdata is made available in the DataLab, direct identifiers are removed and some further confidentialising processes are applied to the data. These “safe data” controls, implemented in applying the Five Safes Framework, help to ensure authorised researchers are prevented from identifying individuals through approved uses of the data. The information flows relating to the creation of statistical microdata products are not in scope of this PIA. However, the high-level information flows for integrated person-centred microdata (based on the description in the MADIP PIA Update¹³) are summarised in the MADIP data sharing model in Figure 2 below and Appendix C.

Figure 2: MADIP data sharing model



While the “safe data” controls make it difficult for authorised researchers to identify individuals from microdata, the ABS recognises that the data may not be truly de-identified in inappropriate settings, such as the public domain. For this reason, the ABS applies additional controls through the other four safes in the Five Safes Framework to ensure an effective balance with providing data that can inform important research while also ensuring privacy is protected and data is secure. This

¹³ The MADIP PIA Update and the ABS response are available on the ABS Privacy Impact Assessments website page: <https://www.abs.gov.au/websitedbs/D3310114.nsf/home/ABS+Privacy+Impact+Assessments>

approach aligns to the mission of the ABS to inform Australia's important decisions while also ensuring it meets its legislative obligations and maintains the trust of the public.

All controls applied under the Five Safes Framework in the existing DataLab will be also be applied in the Cloud DataLab. (Appendix B describes how the Five Safes will be applied in the context of the Cloud DataLab.)

3.2 Information flows – data about Users

The ABS collects personal information from two types of Users – ABS staff and researchers.

ABS staff in Auditor, Administrator, or Systems administrator roles

These staff perform the following functions:

- a. Auditors: vet outputs from the DataLab, ensuring that the results of any research (tables, models, regressions, etc.) are consistent with “safe output” requirements.
- b. Administrators: administer access to the Cloud DataLab for researchers, set up projects and import microdata.
- c. Systems administrators: perform IT security and performance checks and other technology functions.

The personal information collected from these Users is limited to:

- First name and surname
- Work phone number
- Work email address

In line with current processes, ABS staff will create and manage their own accounts (and consequently, their own personal information).

Researchers who are requesting access to microdata in the Cloud DataLab

The personal information collected from these Users is limited to:

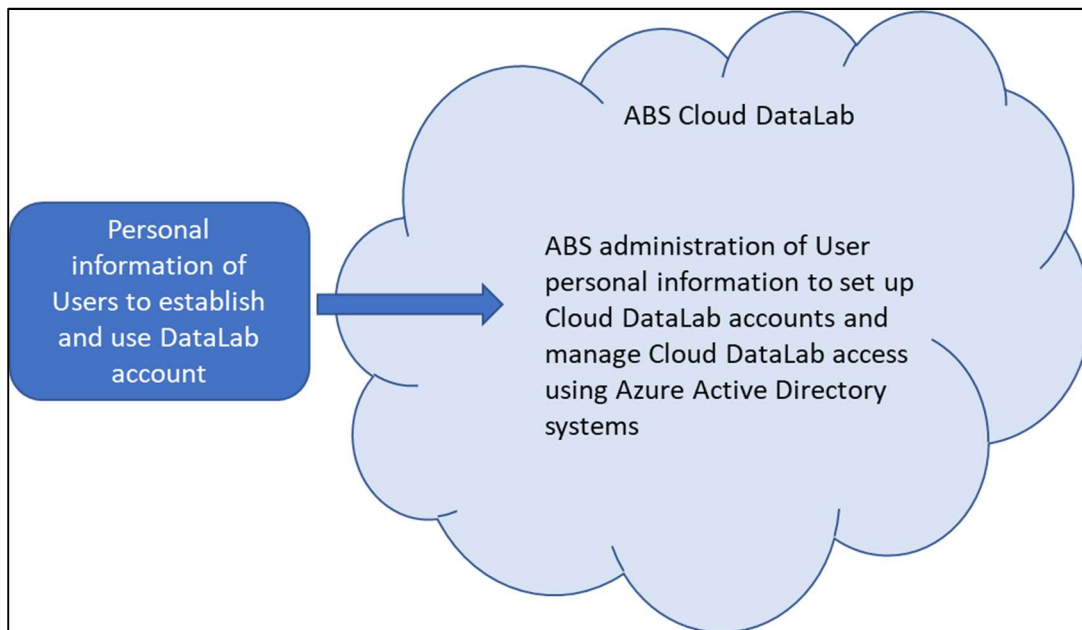
- First name and surname
- Mobile/work phone number
- Work email address
- Employer (organisation)
- Job title or role in the organisation
- Skills and qualifications
- Conflicts of interest related to the microdata

This information is provided by researchers on a standard Project Proposal template. The information in the Project Proposal is used to make assessments about “safe people” and “safe projects”. This template is usually completed by the lead researcher on behalf of the research team and returned to the ABS by email. With any subsequent change to the research team, the lead researcher may provide these details via email, which is then added to the Project Proposal template

by the ABS. Researchers may also change their personal information at any time by notifying the ABS, by phone or email.

Figure 3 provides a high-level summary of key information flows for personal information about Cloud DataLab Users. More detailed information flows for data about Users are provided in Appendix C.

Figure 3: Summary of information flows for personal information about Cloud DataLab Users



Note: Personal information about Users is disclosed to Data Custodians as part of existing microdata access processes. Once the Microdata Access Platform is developed (as discussed immediately below), Data Custodian access approval processes will be managed through that Cloud-based system.

Future development – Microdata Access Platform

To support the Cloud DataLab, the ABS is developing a Microdata Access Platform to assist with management of an increasing number of microdata products available for research and the growing demand for these products. This platform will consist of an expanded User Portal and Customer Relationship Management (CRM) tool. This future development may significantly change how the ABS collects, manages, and stores the personal information of Users. If required, an update to this PIA will be completed to support the development of the Microdata Access Platform.

4 PRIVACY IMPACT ANALYSIS

4.1 Handling personal information

The ABS manages microdata with standards appropriate for personal information even when the microdata may not constitute personal information. The ABS has obligations under the *Census and Statistics Act 1905* to not release information in a manner likely to enable the identification of an

individual. ABS also takes a privacy-by-design approach to ensure privacy impacts of access to microdata are minimised.

To manage the access to microdata, the ABS must create DataLab accounts for all Users. The personal information of DataLab Users required to create these accounts is not bound by the *Census and Statistics Act 1905*, however it is still minimised to reduce privacy impacts. (See Section 3.2 for details of information collected.)

4.2 APP compliance

This section provides an analysis of the Cloud DataLab against the APPs. The analysis covers both storage of microdata in the Cloud DataLab and the collection, storage and use of personal information of Users.

APP1 – open and transparent management of personal information

Compliant, but further action recommended

APP1 requires that an APP entity takes reasonable steps to ensure it complies with the APPs and has a process for dealing with complaints about its compliance with the APPs. APP1 also requires that an APP entity has a clear, up to date, and available privacy policy.

ABS has conducted a number of PIAs about data assets that are accessed using the DataLab, such as MADIP. These PIAs have addressed APP1 requirements for the related microdata.

APP1 requires ABS to be open and transparent about the storage and security of microdata in the cloud environment as part of the Cloud DataLab. ABS plans to update the microdata information on its website to include reference to the use of cloud infrastructure as part of data access systems.

The ABS has policies and practices in place to manage personal information in an open and transparent way. These are outlined in the ABS Privacy Policy¹⁴, which is available on the ABS website. Similarly, Microsoft have significant privacy resources available on their website¹⁵. This includes a document responding to the APPs directly for Azure. (See Section 2.3 and discussion of the Privacy of Microsoft Cloud Services.)

The ABS Privacy Policy provides information about the collection and use of personal information from employees. However, it could be more explicit about the collection, storage and use of personal information about Users of ABS systems such as the DataLab. For example, the privacy policy lists the following reasons for which the ABS might collect personal information. Under the heading “Collection”:

We may collect personal information about you relating to:

- the Census of Population and Housing and surveys conducted under the Census and Statistics Act
- surveys conducted under the ABS Act

¹⁴ <http://abs.gov.au/privacy>

¹⁵ <https://privacy.microsoft.com/en-ca/privacystatement>

- administrative data collections
- other data collection activities (including statistical research)
- your involvement in activities that support statistical coordination and dissemination
- you contacting the ABS
- your employment by or contract to the ABS.

The collection of personal information from researchers in order to manage their access to and use of microdata in the Cloud DataLab is covered by the point “your involvement in activities that support statistical coordination and dissemination”. However, it could be made clearer by reference to the collection of personal information from Users of ABS data products. An updated ABS Privacy Policy is currently in development. The updated version includes simplified language and is designed to be clearer.

Recommendation 1: Update online materials to provide information about the storage and security of microdata in the DataLab, including use of cloud infrastructure.

Recommendation 2: Clearly describe the collection, storage, use and disclosure of personal information about Users in collection notices and/or user agreements.

APP2 – anonymity and pseudonymity

Compliant

APP2 requires that individuals must have the option of not identifying themselves, or of using a pseudonym, when dealing with an APP entity in relation to a particular matter, unless an exception applies.

Consistent with the MADIP PIA Update, APP2 is not relevant to microdata.

The general rule (in APP 2.1) does not apply to the information about Users of the Cloud DataLab if, as per APP 2.2(b), “it is impracticable for the APP entity to deal with individuals who have not identified themselves or used a pseudonym”.

The ABS has a requirement to identify Users accurately in order to facilitate access to data, to ensure data is accessed appropriately, and so Users can be accountable for their access. Criminal penalties apply for non-compliance of use/access.

APP3 – collection of solicited personal information

Compliant

APP3 requires that any personal information collected must be reasonably necessary for one or more of the collecting APP entity’s functions or activities. APP3 imposes an additional requirement for collecting sensitive information, which states that the individual about whom the sensitive information relates must consent to the collection, unless an exception applies.

The Cloud DataLab will not change how the ABS collects personal information as part of the creation of microdata products. As such, this PIA has not considered the impacts of APP3 on microdata for the Cloud DataLab.

The personal information collected from Users is reasonably necessary for the ABS to safely manage access to, and use of, microdata in the Cloud DataLab. All personal information is collected by lawful means. No sensitive personal information is collected.

For all ABS administrative Users, personal information (name, employer, phone number and email) is collected directly from the individual.

Personal information about researchers is collected using a standard template the ABS has developed to facilitate research projects including internal and Data Custodian approvals for microdata access. This template is usually completed by the lead researcher on behalf of the research team and returned to the ABS by email. This personal information is then entered in to the Cloud DataLab by ABS administrative staff in order to create User accounts.

The collection of personal information about an individual (in this case, members of the research team) from a third party (the lead researcher) is permitted by APP3, which allows this third-party collection to occur where it is unreasonable or impractical to collect the personal information directly from the individual (APP 3.6(b)).

A project involving access to the Cloud DataLab could involve a number of researchers, who will only be known by the lead researcher until a Project Proposal is submitted. It would be impractical for the ABS to coordinate obtaining all of the required personal information for all of the researchers for whom access is being sought – this role is more effectively performed by the lead researcher.

In addition, given the only personal information collected from the lead researcher on behalf of another individual is their name and employer, and the individual and lead researcher have a professional working relationship, it is not unreasonable for the ABS to collect the information in this manner. The ABS considers that all authorised researchers would reasonably expect their personal information to be shared with the ABS by the lead researcher for the purposes of administering their Cloud DataLab access. This will be enhanced if the recommendations discussed in APP5 are implemented.

APP4 – dealing with unsolicited personal information

Compliant

APP4 requires that where an APP entity receives unsolicited personal information, it must determine whether it would have been permitted to collect the information under APP3. If so, APPs 5 to 13 will apply to that information. If the information could not have been collected under APP3, and the information is not contained in a Commonwealth record, the APP entity must destroy or de-identify that information as soon as practicable, but only if it is lawful and reasonable to do so.

The Cloud DataLab will not change how the ABS collects personal information as part of the creation of microdata products. As such, this PIA has not considered the impacts of APP4 on microdata for the Cloud DataLab in detail. Through the Five Safes Framework, the ABS makes a considerable effort to ensure personal information, which may have been present in microdata when it was collected, is not disclosed to Users. This applies to solicited and unsolicited personal information.

There are very few opportunities for a User to supply unsolicited personal information to the ABS. While it is possible for a User to include unsolicited personal information in the Project Proposal template, it will be handled in accordance with the ABS' usual policies for handling unsolicited information and destroyed and not stored as part of administering the Cloud DataLab access arrangements.

APP5 – notification of the collection of personal information

Not fully compliant - action required

APP5 requires that where an APP entity collects personal information about an individual, it must take reasonable steps to notify the individual, or otherwise ensure the individual is aware of certain matters.

The Cloud DataLab will not change how the ABS collects personal information as part of the creation of microdata products. As such, this PIA has not considered the impacts of APP5 on microdata for the Cloud DataLab in detail. However, the ABS understands the importance of transparency in informing individuals how their data is managed and is committed to publishing information about the move to cloud infrastructure for microdata access storage and analytics. The ABS will also explore other ways to make individuals aware of the Cloud DataLab project.

Personal information is collected directly from Cloud DataLab Users, or from the lead researcher on behalf of the User, using the Project Proposal template. This template does not specify that personal information will be collected, or how the personal information will be used or stored.

It is reasonable to expect that a researcher will be aware of the purpose of the collection as they are applying for access to ABS microdata. However, they may not be aware that a Microsoft Azure account will be created using their personal information. Further, where the lead researcher has provided information about other individuals, those individuals may not be aware that their personal information will be used to create a Microsoft Azure Cloud DataLab account.

It is not possible for a researcher (or members of the public) to create a Cloud DataLab account. Only ABS DataLab administrators can create an account, using the information provided on the Project Proposal template. However, the User must complete the account set up process by responding to the two-factor authentication email message, logging on, and creating a password.

To increase transparency for Cloud DataLab Users, ABS will create an APP5 notice to make them aware of how their personal information will be used. The notification will be included on the Project Proposal template or clearly linked from the template and require lead researchers to confirm that the collection notice has been provided to all other nominated researchers. It could also be included in the account verification email and accessible from the Cloud DataLab login page.

Recommendation 3: Create an APP5 notice to Cloud DataLab Users to make them aware of how their personal information will be used, including that it will be stored on a Cloud-based service, used by an off-shore service provider, and disclosed to Data Custodians. Take additional steps to ensure that all Users are made aware of the collection notice.

APP6 – use or disclosure of personal information

Not fully compliant – action required

APP6 requires that an APP entity only use or disclose personal information for the particular purpose for which it was collected (known as the “primary purpose”), or for a secondary purpose if the person has consented or if an exception applies, such as where the secondary use or disclosure is required or authorised by or under an Australian law.

As the Cloud DataLab will not change how the ABS collects or uses personal information as part of the creation of microdata products, this PIA has not considered the impacts of APP6 on microdata in detail. The personal information of Users is only used as part of the registration process to access the Cloud DataLab, to manage that access and ensure the effective and efficient functioning of the Cloud DataLab.

Personal information of Users is shared with Data Custodians as part of the existing processes for Data Custodians to approve access for the purposes of an approved project. Once the Microdata Access Platform is developed (as discussed in Section 3.2), Data Custodians will access this information and provide approvals using the Microdata Access Platform. There is no disclosure of personal information of Users for any other purpose.

In order to ensure the process is transparent to Users of the Cloud DataLab, the APP5 notice should specify that personal information is disclosed to Data Custodians as part of the approval process and in the event that the User misuses information held in the DataLab.

Recommendation 4: Ensure the APP5 notice notifies Users that their personal information may be disclosed to Data Custodians in order for them to approve access to microdata.

APP7 – direct marketing

Compliant

APP7 requires that certain classes of APP entities must not use or disclose personal information for the purpose of direct marketing unless an exception applies, such as where the individual has consented.

APP7 is not relevant as it applies to organisations rather than to agencies like the ABS and the ABS has not been otherwise prescribed for the purposes of s7A of the Privacy Act. We also note that direct marketing is not relevant to the Cloud DataLab.

APP8 – Cross border disclosure

Compliant

APP8 requires that before an APP entity discloses personal information to an overseas recipient, the APP entity must take reasonable steps to ensure that the overseas recipient does not breach the APPs (other than APP1) in relation to the information, unless an exception applies, such as the individual has given informed consent.

Microdata

As researchers will not have access to personal information in the Cloud DataLab, APP8 is not relevant to microdata accessed in the Cloud DataLab. However, as best practice, ABS has examined relevant issues relating to storage and use of microdata.

Cloud services by their nature can involve the storage of data outside of Australia and there are many service providers based outside of Australia. The ABS has procured a service where microdata storage will continue to be hosted in Australia. This is ensured through the contract with the service provider (Microsoft Agreement) whereby the ABS retains effective control of microdata and microdata is not disclosed to Microsoft or its staff.

Authorised researchers must be located in Australia to access the Cloud DataLab.

Personal information about Users

Azure Active Directory (AAD) operates from an Australian hosted data centre. This service utilises user names and other details related to the user accounts of Cloud DataLab Users. This data is handled for the exclusive purpose of authenticating and authorising Users of the Cloud DataLab system.

Additionally, first name, surname, email address, and some location information in the form of originating IP address of Users can be captured, stored, and processed by Microsoft Intrusion Detection and Prevention Services, located in the US. This data is handled for the exclusive purpose of security analysis for the Cloud DataLab platform in accordance with the Microsoft Agreement.

Under guidance issued by the OAIC, the majority of handling of the Users' personal information can be regarded as a "use" by ABS rather than a "disclosure" of personal information to Microsoft (for example, because ABS retains effective control of all personal information of Users, and can access, change or remove the data at any time).

For all services provided by Microsoft (including the security analysis for the Cloud DataLab platform), conditions on the provision of services by Microsoft Azure to the ABS are specified in the Microsoft Agreement (as outlined in Section 2.3). For many of the contracted services, those contractual obligations will mean the OAIC's guidance requirements for there being a "use" of personal information by ABS rather than a "disclosure" to Microsoft, have been met.

Importantly, the Microsoft Agreement notes that Microsoft is a contracted service provider, that its use of personal information must be consistent with Australian privacy legislation, and that it cannot access or disclose any personal information except where doing so is required for the delivery of the services it is engaged to provide. In this case, use of personal information by the AAD, and the Intrusion Detection and Prevention Services is necessary to ensure microdata access is securely managed in the Cloud DataLab.

The ABS considers this use of personal information is reasonably necessary for enforcement related activities in connection with the responsible use of microdata, as enabled by the *Census and Statistics Act 1905*.

The ABS considers that it has therefore taken reasonable steps to ensure that Microsoft is contractually bound to ensure that it does not breach the Australian Privacy Principles in relation to any overseas handling of any personal information, in compliance with APP 8.

The ABS acknowledges that having appropriate contractual obligations will only be effective in protecting personal information, if compliance with those obligations is monitored and enforced. The ABS's standard monitoring and compliance regime ensures that contractual requirements are adhered to by providers, such as Microsoft.

CLOUD Act

The *Clarifying Lawful Overseas Use of Data Act* (CLOUD Act) enacted by the USA government allows USA federal law enforcement to compel USA based technology companies via warrant or subpoena to provide requested data stored on servers regardless of whether the data are stored in the USA or overseas. The CLOUD Act therefore applies to the services provided by Microsoft in relation to the Cloud DataLab.

The Australian government is introducing legislation to underpin a future bilateral agreement with the USA under the CLOUD Act. The Telecommunications Legislation Amendment (International Production Orders) Bill¹⁶ will establish a new framework under the *Telecommunications (Interception and Access) Act 1979* to allow for "reciprocal cross-border access to communications data" for law enforcement purposes. It will allow law enforcement and national security agencies to access data directly from communications providers, subject to international agreements being in place.

All microdata stored in the DataLab is encrypted with ABS having sole use of the encryption key. The CLOUD Act does not create any obligation that service providers be capable of decrypting data.

It is unlikely that the CLOUD Act will be used to seek access to any microdata stored in the Cloud DataLab given the unidentified nature of the data would be unlikely to assist enforcement agencies. Information about Users may theoretically be sought under the CLOUD Act but the information again would be of limited use to enforcement agencies. However, the ABS acknowledges that the CLOUD Act is a significant development and in the interests of transparency notes it may impact the Cloud DataLab (subject to the passage of the Australian Bill and entering into the necessary bilateral agreement). The Microsoft Agreement contains protections designed to minimise any impact of the CLOUD Act, including through provisions requiring notification to the ABS.

APP9 – adoption, use or disclosure of government related identifiers

Compliant

APP9 requires that certain classes of APP entities must not adopt, use, or disclose a government related identifier of an individual as its own identifier of the individual unless an exception applies.

The Cloud DataLab will not hold any Government related identifiers. APP9 is not applicable.

¹⁶ https://www.aph.gov.au/Parliamentary_Business/Bills_Legislation/Bills_Search_Results/Result?bld=r6511

APP10 – quality of personal information

Compliant

APP10 requires that an APP entity must take reasonable steps to ensure that the personal information it collects is accurate, up-to-date, and complete.

APP10 requirements for microdata are consistent with the assessment described in the MADIP PIA update that ABS has adequate processes and systems in place that represent reasonable steps to ensure the quality of personal information.

Information is provided to DataLab Users (during training) about how to inform ABS of changes to their personal information as part of the on-boarding process. This is a common practice for the existing DataLab, and the Cloud DataLab project will not alter these processes.

When a request is received, ABS staff responsible for administering the Cloud DataLab will be able to update personal information of Users using the administrator user interface.

APP11 – security of personal information

Microdata: Compliant, but further action recommended

Personal information of Users: Not fully compliant - action required

APP11 requires that an APP entity must take reasonable steps to protect personal information it holds from misuse, interference, and loss, as well as unauthorised access, modification, or disclosure. It must also take reasonable steps to ensure personal information is destroyed or de-identified once it is no longer needed.

The ABS has taken reasonable steps to ensure that microdata stored in the Cloud DataLab is protected from misuse, unauthorised access or disclosure. The MADIP PIA Update outlines a range of protections of microdata covering legislative, protective security, Information and Communication Technology and data governance controls. These findings are all relevant to the security of access to microdata in the Cloud DataLab.

The following section analyses the change to DataLab processes in storing information in a cloud environment and outlines the secure implementation, verification and ongoing support of the Cloud DataLab.

Implementation

The ABS performed a competitive open tender process to select the implementation partner for the Cloud DataLab. The selection criteria included direct assessment of each potential implementation partner on their ability to protect the security of personal information with their proposed solution.

- Microsoft Azure were selected to partner with ABS on the Cloud DataLab and have a demonstrated ability to protect the security of personal information. In particular: Core Azure services were certified as PROTECTED by the Australian Signals Directorate in April, 2018.

- The solution architecture has robust controls to prevent data being copied from the DataLab to the User's PC.

ABS has worked closely with Microsoft to implement the Cloud-based solution in line with the Australian Government Digital Transformation Agency "Secure Cloud Strategy"¹⁷. This strategy notes that "Cloud providers often implement and manage better IT security controls than internal IT teams as it is a core part of their business and reputation." The Cloud DataLab implements multiple Cloud-based IT security controls that are more robust than what could be applied in an on-premise ABS solution.

Verification

Information Security Registered Assessors Program (IRAP) assessments provide an independent assessment of security compliance of projects and highlight information security risks. IRAPs are based on the Australian Signals Directorate's Information Security Manual and the Protective Security Policy Framework.

The ABS is undertaking an IRAP assessment of the Cloud DataLab to verify the security of all aspects of the implemented solution. The IRAP will be completed before the Cloud DataLab is operational. The Microsoft Azure services being used in the Cloud DataLab have also been IRAP assessed. The ABS is committed to act on the outcomes of the Cloud DataLab assessment so that any risk of unapproved access to personal information is effectively managed from an IT security perspective.

Ongoing support

The ABS will re-run IRAP and IT security testing as major functionality is added to the Cloud DataLab. The ABS has ongoing annual funding dedicated to these activities. The ABS also has a dedicated in-house IT security section and an ABS Chief Security Officer dedicated to minimising ongoing security risks.

Deletion and retention of personal information of Users

There is an operational need to retain User information so that, in the event of a disclosure breach, there is a full history of the individuals who have accessed data over time. This approach will be formally documented in a deletion and retention policy.

Recommendation 5: Implement any outcomes arising from security assessments to assure the continued security of microdata in the Cloud DataLab.

Recommendation 6: Implement any outcomes arising from security assessments to assure the continued security of personal information of Cloud DataLab Users.

Recommendation 7: Create a deletion and retention policy specific to DataLab User accounts and related personal information.

¹⁷ <https://www.dta.gov.au/our-projects/secure-cloud-strategy>

APP12 – access to personal information

Compliant

APP12 requires that an APP entity that holds personal information about an individual must give the individual access to that information on request, unless an exception applies.

APP12 requirements for microdata are consistent with the assessment described in the MADIP PIA update.

On request, the ABS can provide a DataLab User with details of the personal information we hold about them. Given this information is very limited, this is not likely to be a frequent request. The ABS will not charge for these requests.

APP13 – correction of personal information

Compliant

APP13 requires that an APP entity must take reasonable steps to correct personal information to ensure that, having regard to the purpose for which it is held, it is accurate, up-to-date, complete, relevant, and not misleading.

APP13 requirements for microdata are consistent with the assessment described in the MADIP PIA update. The ABS has policies and procedures in place for complaints and the correction of inaccurate data.

ABS staff responsible for administering the Cloud DataLab can update their own personal information, and information about Cloud DataLab Users, via the administrator user interface. DataLab Users can request this information be corrected. This is a regular occurrence already, particularly when machinery of government changes require the ABS to update email addresses for government users.

5 ABS RESPONSE AND NEXT STEPS

ABS will implement all recommended actions before the Cloud DataLab becomes fully operational.

In terms of next steps, once operational, the Cloud DataLab will continue to adapt and evolve to meet user expectations, methodological and technological advancements, and other environmental changes. The ABS is also continuously improving data handling practices and infrastructure, such as for the Cloud DataLab, to preserve privacy, ensure data security, and increase data quality and utility. This PIA is a demonstration of the ABS' commitment to managing the privacy impacts of the project.

The ABS will publish a progress report on the ABS website within one year of this PIA being published to inform on progress of implementing APP compliance recommendations. It may also update the PIA if further developments require a privacy impact reassessment.

Appendix A: Glossary and Acronyms

Acronym	Term
ABS	Australian Bureau of Statistics < www.abs.gov.au >
APP	Australian Privacy Principle
IRAP	Information Security Registered Assessors Program < https://www.cyber.gov.au/programs/irap >
ISM	Australian Government Information Security Manual < https://www.cyber.gov.au/ism >
MADIP	Multi-Agency Data Integration Project < www.abs.gov.au/madip >
OAIC	Office of the Australian Information Commissioner < www.oaic.gov.au >
PIA	Privacy Impact Assessment

Term	Description
Accredited Integrating Authority	An agency authorised to undertake high-risk data linkage projects involving Commonwealth data for statistical and research purposes.
administrative data	Data maintained by governments and other entities, including data used for registrations, transactions, and record keeping, usually during the delivery of a service.
Australian Privacy Principles	Principles contained in the <i>Privacy Act 1988</i> that regulate the way we collect, store, provide access to, use, and disclose personal information.
authorised researchers or researchers	As part of the “safe people” element of the Five Safes Framework, access to microdata in the DataLab is only provided to authorised researchers, who are persons who have been authorised by the ABS (and where relevant other Data Custodians) to undertake approved research projects.
cloud	As per the US National Institute of Standards and Technology (NIST) definition of cloud computing .
Data Custodian	The agency that collects or generates data for any purpose, and is accountable and responsible for the governance of that data.
de-identified	Personal information is de-identified “if the information is no longer about an identifiable individual or an individual who is reasonably identifiable” (section 6(1) of the Privacy Act). (De-identified data is different to unidentified data - see the meaning of unidentified data.)
direct identifier	Information which, by itself, is able to identify an individual, organisation, or other entity.
Five Safes Framework	An internationally recognised approach to managing disclosure risk – each “safe” refers to an independent but related aspect of disclosure risk.
microdata	Data in a unit record file that provides detailed information about people, households, businesses or other types of entities.

Microsoft Agreement	The documents making up the contractual agreement between the ABS and Microsoft, in connection with the provision of services required for the provision and operation of the Cloud DataLab.
personal information	As defined in section 6(1) of the Privacy Act 1988 .
Privacy Impact Assessment	A systematic assessment of a project that identifies the impact that it might have on the privacy of individuals, and sets out recommendations for managing, minimising, or eliminating that impact
re-identification	The act of determining the identity of a person or organisation even though directly identifying information has been removed.
sensitive data	Data that would be considered sensitive information under the Privacy Act 1988 (Cth) if the data included personal information.
sensitive information	As defined in section 6(1) of the Privacy Act 1988 .
unidentified	Data is considered “unidentified” when direct identifiers such as name and address are removed or altered into an unidentifiable form. Further confidentialisation or safeguards (such as access controlled through the Five Safes Framework) are often required for the data to be considered de-identified.
Users	Authorised researchers and ABS staff with access to the Cloud DataLab

Appendix B: Application of the Five Safes

Table 3. Application of the Five Safes Framework in the DataLab environments

Safe	DataLab	Cloud DataLab
Safe People	Users must undergo training, complete an authorisation process, sign legally binding confidentiality undertakings and a compliance declaration. Breaches of protocols or disclosure of information may be subject to sanctions and/or legal proceedings.	Unchanged.
Safe Projects	Users must detail the purpose for which they will use the data. This can be compared to what analysis results are actually produced (see Outputs).	Unchanged.
Safe Settings	The DataLab is inside the ABS IT environment (with virtual access available to some Users). It requires secure login and has auditing and monitoring capabilities. No data can be removed from the DataLab without first being checked by ABS staff. The system does not prevent Users from having multiple projects open at the same time.	The Cloud DataLab is hosted within Microsoft Azure tenancies within Australia. It requires secure login and has auditing and monitoring capabilities. The systems prevents Users from concurrently accessing multiple projects. No data can be removed from a project without first being checked by ABS staff.
Safe Data	Direct identifiers are removed and the data are further treated where appropriate. Appropriate control of the data optimises its usefulness for statistical and research purposes.	Unchanged.
Safe Outputs	All statistical outputs are assessed by the ABS for disclosure before being released to the User. The outputs may also be compared for consistency with the original project proposal.	Unchanged.

Appendix C: Information flows

Microdata

As discussed in Section 3.1 of this report, the information flows to generate the assembled microdata extracts are out of scope of this PIA. The main difference in information flows between the current ABS DataLab and the Cloud DataLab is the transfer of the microdata files from the ABS environment to a cloud environment. (As shown in Figure 3, Section 3.1.)

In summary, the process for transferring microdata to the cloud environment is:

- ABS staff create the microdata file in the ABS Next Generation Infrastructure (NGI) environment.
- The file is validated and checked before being authorised for transfer.
- Once authorised, ABS staff place the microdata file in a secure transfer folder in the ABS NGI environment.
- A DataLab Administrator logs into the Azure Admin portal user interface and uses the Azure Storage Explorer system to copy the microdata file from the secure transfer folder in the ABS environment and securely load it to the Cloud DataLab environment.

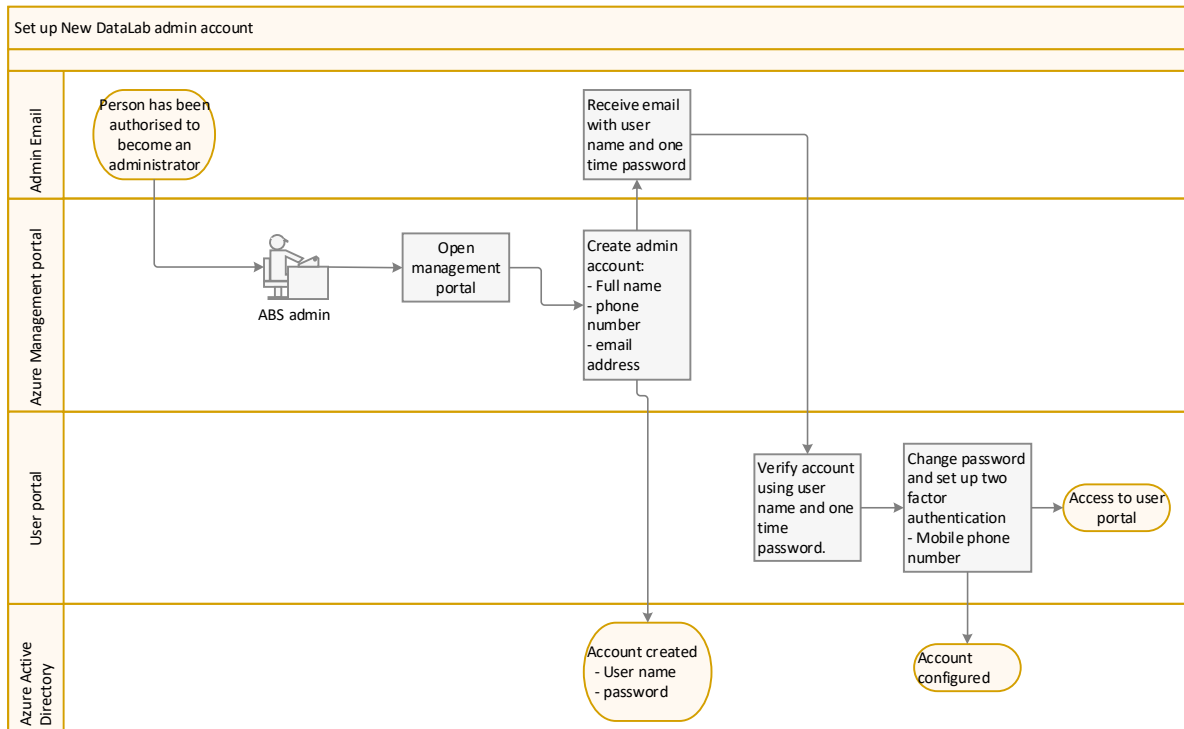
The processes and systems for transferring information from the ABS environment to the cloud environment were covered in the IRAP assessment.

Data about Users

Data flow for creating an administrator account

A small team of ABS staff administer access to the Cloud DataLab. These Users are able to set up their own accounts as described in Figure 4.

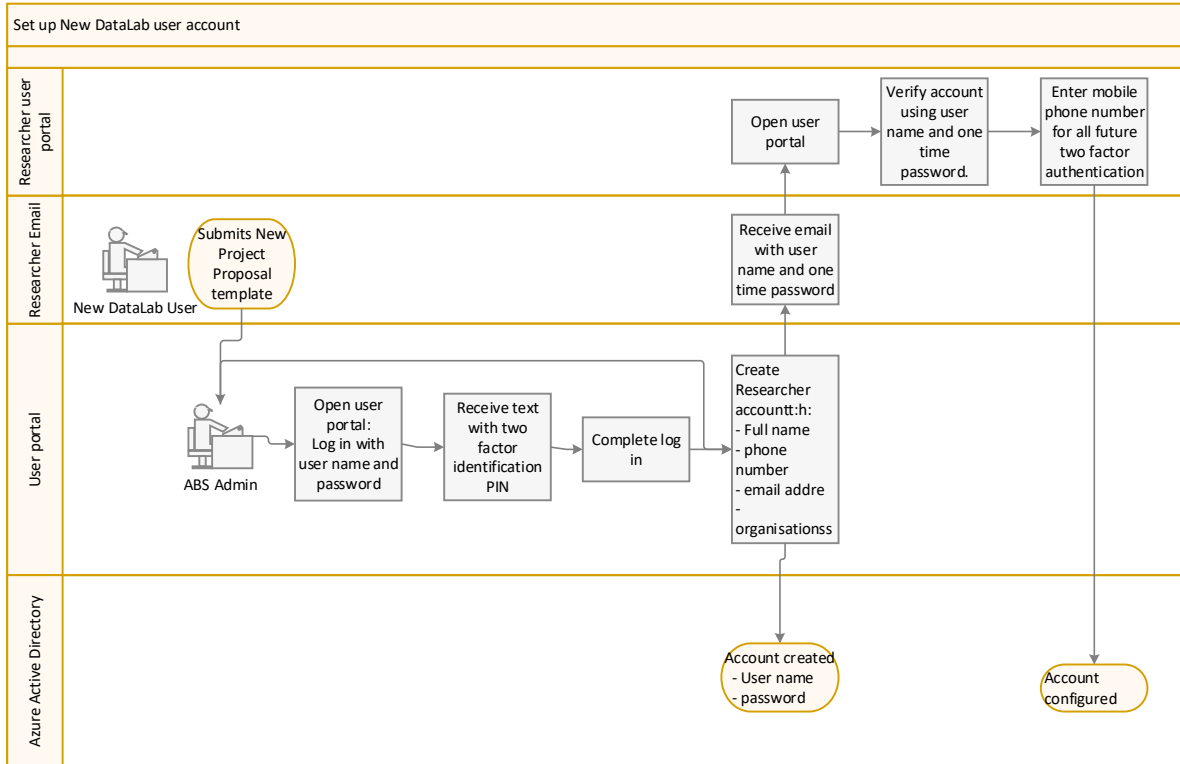
Figure 4: Set up Cloud DataLab Administrator Account



Data flow for creating a researcher account

Once a project has been approved, and the Users have been authorised to access the requested microdata, the Cloud DataLab administrative team will set up a Microsoft Azure account for each approved researcher, using some of the personal information provided in the Project Proposal. See Figure 5.

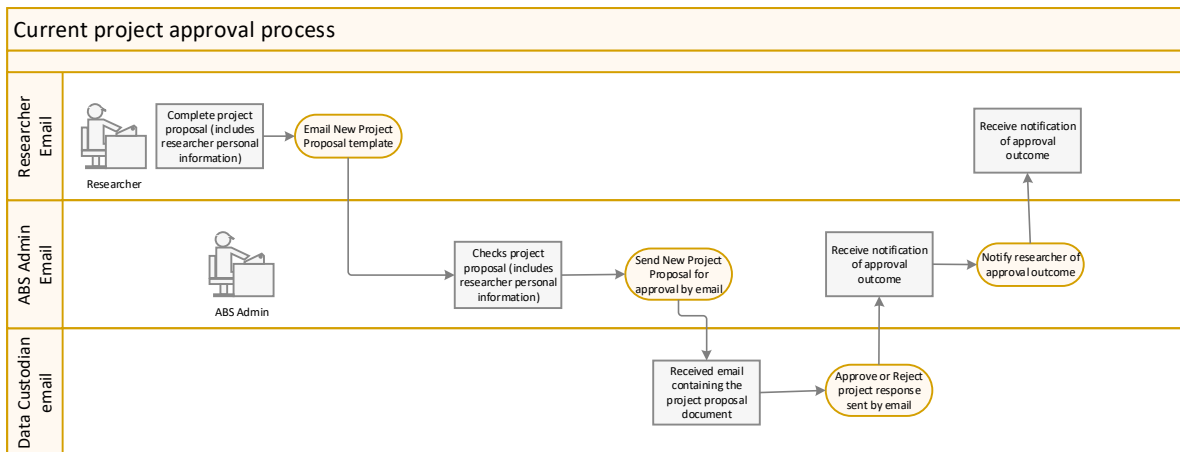
Figure 5: Set up Cloud DataLab Researcher Account



Data flow for project approval process

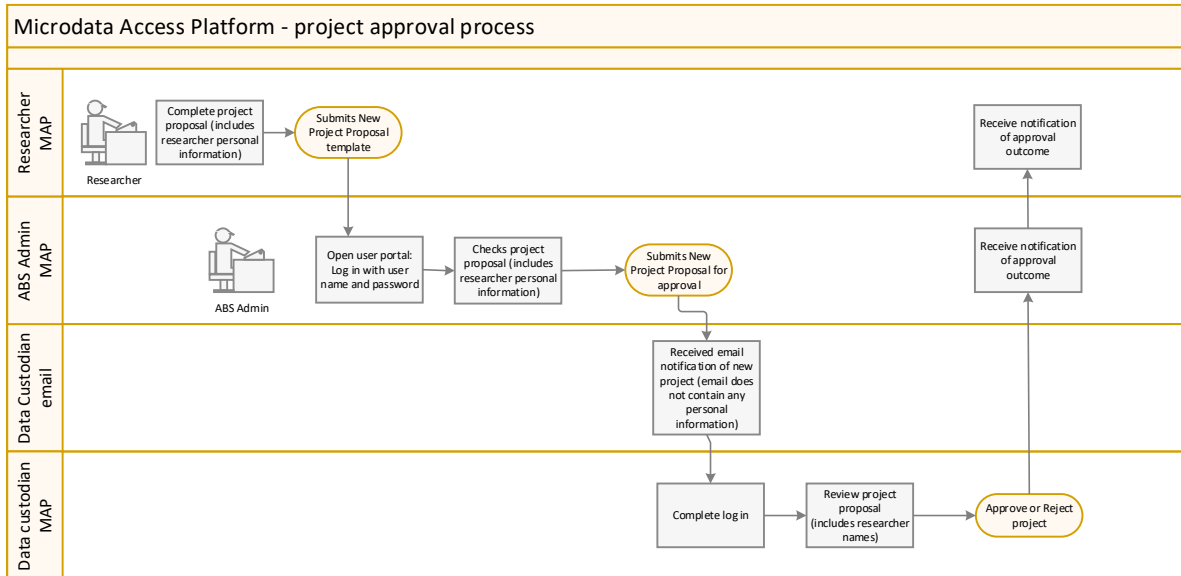
The current project approval process is managed by email as described in Figure 6.

Figure 6: Current project approval process



In future, the project approval process will be managed using the Microdata Access Platform as described in Figure 7.

Figure 7: Future Microdata Access Platform project approval process



Data flow for logging in

When new Users log into their account for the first time they will provide Microsoft with details for their mobile number. Microsoft will use the mobile number to send a 6-digit number as either a text, or an app notification to complete the second login step. See Figure 8.

Figure 8: Logging in to Cloud DataLab

