Cloud DataLab Privacy Impact Assessment update

Implementation of Microdata Access Platform

September 2023

Background

In 2020 the Australian Bureau of Statistics (ABS) conducted the Cloud DataLab Privacy Impact Assessment (PIA). The Cloud DataLab PIA was published in June 2020 accompanied by an assurance report from external privacy advisors Maddocks who were engaged to provide independent advice and assurance for the PIA process and report. At the same time, a Response to 2020 Cloud DataLab PIA was also published, outlining the ABS' response to each of the recommendations and suggestions presented in the PIA.

This PIA update looks to complete work previously identified in the Cloud DataLab PIA report, specifically '**Future development – Microdata Access Platform'** (cloud DataLab PIA, 5 June 2020 report, page 17).

> *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.*

The scope of this update considers the change to the management tools used rather than changes to the existing collection processes for accessing DataLab projects or the nature of the information collected. Hereon in this document we will refer to the 'Microdata Access Platform' as myDATA, to establish connection with the management tool.

Purpose and scope

The Office for Australian Information Commissioner's (OAIC) guidelines recommend a PIA is undertaken when any project or activity impacts on the privacy of individuals. ABS has reviewed the original Cloud DataLab PIA and consider changes made to release myDATA are to the tools used that drives management of DataLab projects and not the process or collection of information.

In this update we will specifically address:

- Design of myDATA Information Communication Technology (ICT) solution – the tool that drives management of DataLab projects
- Data use and information flows within myDATA – how the ABS collects, holds, manages, and discloses information about users of DataLab
- Compliance with the Australian Privacy Principles (APPs)
- Next steps

Design of myDATA ICT solution

The myDATA solution supports the lifecycle of a DataLab project. From application, assessment, approval through to closure of the project. It is new infrastructure comprising of two components:

1. An external facing portal (myDATA user portal) – which is a web application that makes use of the Microsoft Power Platform
2. An internal to ABS management application (myDATA administration portal) – which is a configuration of Microsoft Dynamics 365 (cloud based) inside an ABS managed tenancy

The external myDATA user portal pulls together the project proposal, including details on organisation, portal users and data, which the project is seeking to access. When an organisation has a Responsible Officer Undertaking in place and organisation is added to the project proposal, registered portal users belonging to that organisation can be searched for by name and attached to the project proposal. Subsequently these details are reviewed and approved by the Data custodian of the product.

The primary purpose of myDATA portal is to provide:

- modern infrastructure and process management
- an online project proposal application and approval workflow
- visibility of status relating to, project, portal users and products
- access to services which help manage their projects, including portal users and products
- manage access to future ABS microdata services through one centralised location

Access to the internal myDATA administration portal is managed using Azure Active Directory (AAD) groups synchronised from an ABS on premise role management system. This limits access to ABS employees involved in reviewing and supporting DataLab access. Access is managed by a role owner group, who review access membership quarterly. If reviews are not completed within two weeks, access is automatically revoked. Role owners cannot approve their own access – another role owner will be required to do this.

In addition to the underlying technologies used to host them, to support these two interfaces (myDATA portal and myDATA administration portal), the myDATA system integrates with:

- internal identity systems
- Microsoft Azure components
- internally managed email services

Transfer of personal information to the protected DataLab environment is manual, as indicated in Diagram 1.

Personal information and myDATA

As identified in the original cloud DataLab PIA (2020), personal information is collected from Users for ABS to safely manage access to, and use of, microdata in the DataLab. No sensitive personal information is collected, as defined under Privacy Act 1988.

Personal information is used in the DataLab Project Proposal and subsequently used to make assessments about “Safe People” and “Safe Projects” (Appendix A: Application of the Five Safes Framework with context of myDATA release).

The myDATA registration includes Conditions of use, which reference how personal information is used.

Any personal information collected within myDATA will be retained while there is a business need and will be protected in accordance with the Privacy Act 1988.

The types of information being collected which are defined as personal information under the Privacy Act:

- Full name and preferred name
- Mobile/work phone number
- Work email address
- Employer (organisation)
- Job title or role in the organisation
- Skills and qualifications

Other information collected and could be considered as personal information about an individual include:

- Conflicts of interest
- Completed DataLab application forms including signatures
- What Project and data the individual has had over time

As an added security layer, access to myDATA will use the identity provider Okta Australia. Okta is a cloud hosted software as a service solution that will provide secure 2 factor identity authentication and store access information for myDATA registered users. More information about Okta is available in the section 'Data use and information flows within myDATA.'

Penetration testing

A third party was engaged to design and execute penetration tests across the external facing elements of myDATA.

Tests were performed to identify any system vulnerabilities, including the potential for unauthorised parties to gain access to the system, and strengths. The ABS has addressed changes for all high and medium issues and will implement all remaining recommendations by the end of 2023.

IRAP assessments

An Independent Security Registered Assessors Program (IRAP) assessment of the myDATA system was completed in February 2023. The assessment was against OFFICIAL: SENSITIVE classification criteria.

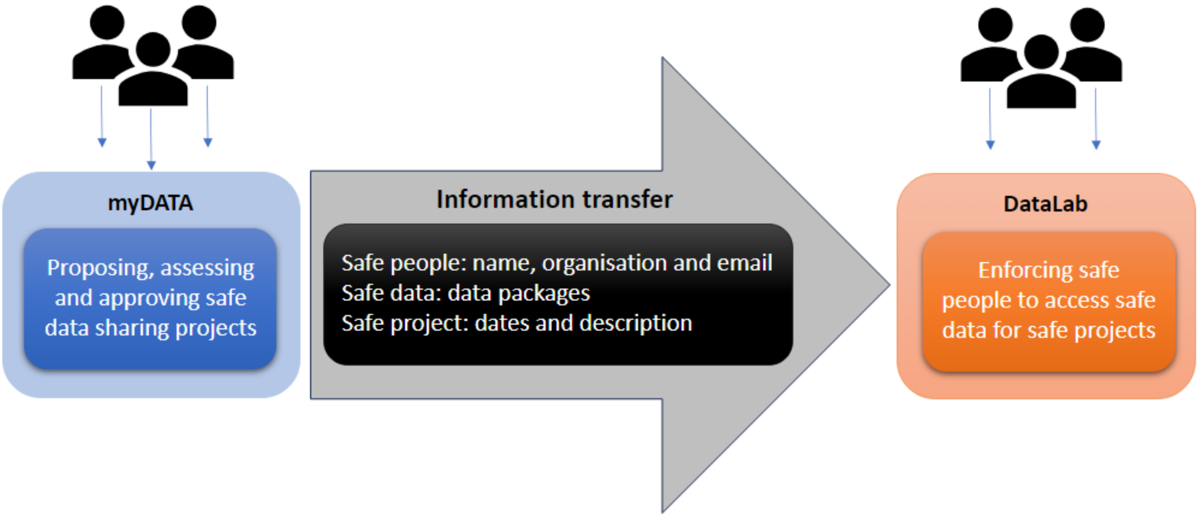
The IRAP assessment evaluated the implementation, appropriateness, and effectiveness of the myDATA security controls. ABS is considering recommendations from the assessment for remedial action, as part of the ABS' Cyber risk management approach.

The Microsoft Dynamics 365, Power Platform and Okta identity solution have all had IRAP assessments with Okta to PROTECTED level.

Data use and information flows within myDATA

MyDATA portal will collect personal information through the registration process, including name, email address, job title and details of the organisation. This personal information is used throughout the DataLab approval process, and passed manually to the protected DataLab environment, which enforces access control. Not all project members gain access to DataLab.

Diagram 1: Information flow from myDATA to DataLab



myDATA

Proposing, assessing and approving safe data sharing projects

Information transfer

Safe people: name, organisation and email
Safe data: data packages
Safe project: dates and description

DataLab

Enforcing safe people to access safe data for safe projects

The myDATA portal will be able to connect registered users to a project, usually initiated by the project lead. Users will be able to view which projects they are participating in and other researchers who are approved to access the project and datasets therein. By providing full preferred name and contact details such as email and phone number users can establish who they are approved to discuss project analyses with and reduce risk of unintentional sharing information with unapproved parties.

MyDATA and DataLab do not directly connect, resulting in approved researcher analyst authenticating to DataLab separately. The ABS will update the PIA if further development changes this interaction.

Portal users maintain their details within the myDATA portal which are passed to the myDATA administration portal, allowing one store of user information across the multiple project applications to which the external user could belong.

Okta stores user identity information within user profiles, for authentication and subsequent access to the myDATA user portal.

The ABS manages the Okta tenancy that hosts these profiles, and the service provides appropriate administrative, physical and technical safeguards to protect the security and integrity of its service.

Okta does not share customer data with other organisations. ABS Okta instances are hosted in Australia; however, disaster recovery centres are based in Singapore as Okta has no option but to use other regions. ABS Security has provided assurance that this is consistent with services other venders have in place.

Minimal administration user profile identity information is also stored in AAD to enable authentication and authorisation of approved ABS employees to the internal component – myDATA administration portal.

Both stores (Okta and AAD) are managed by ABS contracts, which ensures that the ABS retains effective control of all user information, and can access, change, or remove data and information at any time.

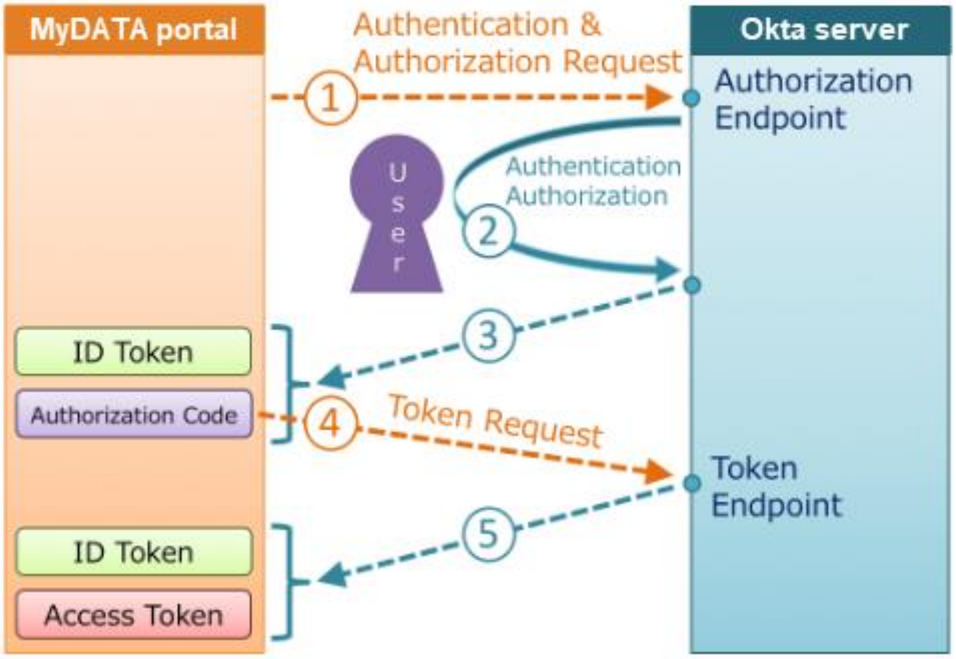
Personal information of Users (full preferred name and Organisation) is shared with Data Custodians, who also register personal information to the myDATA portal, this approval process is part of the existing processes for Data Custodians to approve access for the purposes of an approved project. The myDATA portal will now facilitate the secure review and assessment of approval requests.

The information collected underpins part of what forms the 'Safe People' and 'Safe Project' managerial control within the Five safes framework.

Authentication of identity

Through the registration and login process portal users will authenticate their identity with Okta, this ensures ABS is safely managing access to the system.

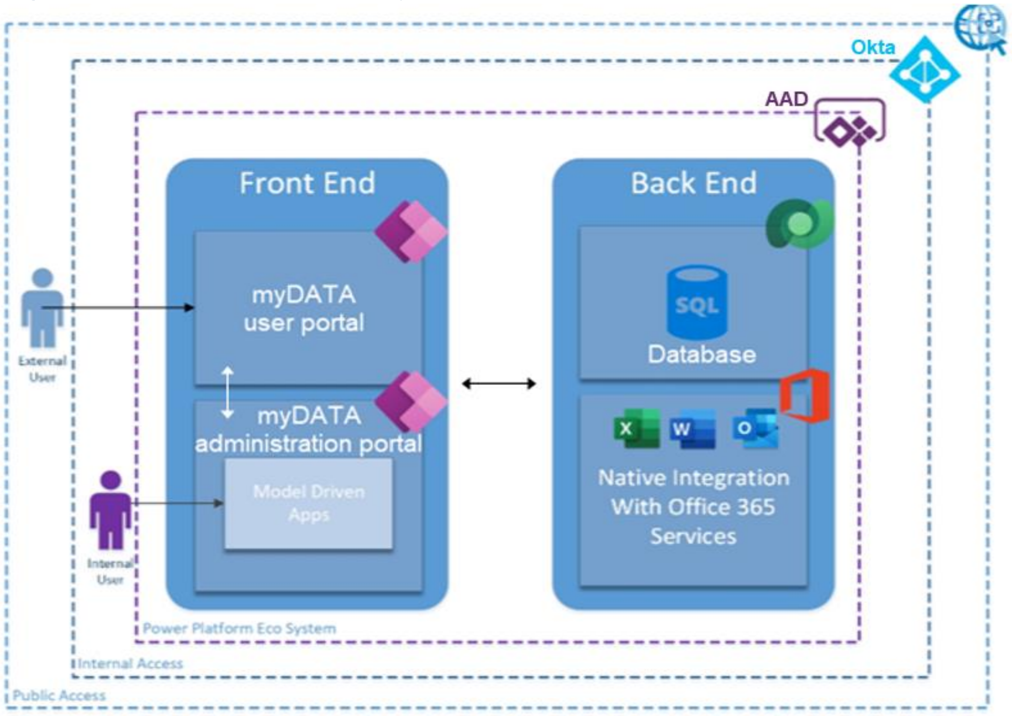
Diagram 2: myDATA portal and Okta server authentication



Information store of myDATA

High level architecture of myDATA is represented in Diagram 3 and described below.

Diagram 3: Information store of myDATA



Dotted lines represent layers of access management. If the external user has not registered or authenticated via Okta, then they would not be able to access the myDATA portal.

Likewise, if the Internal user does not have the appropriate AAD membership, they would not be able to access the myDATA administration portal.

myDATA user portal: front facing interface which collects and displays personal information, for use in:

- Portal user profile
- Organisation details
- Project proposal applications
- Project team views

myDATA administration portal: restricted ABS internal administration portal, displays personal information for staff to perform the following functions:

a. Administrators: apply controls required to ensure ‘Safe People’ and ‘Safe Projects’ are applied under the Five Safes Framework governing access to projects within DataLab
b. System administrators: perform IT security and performance checks among other technology functions

Database: hold the information layers relating to:

- Portal users
- Organisation
- Data Packages
- Projects

Native integration with Office 365 services: Native integration with Office 365 services to support automation, reporting, data migration and other extended function activities for activities like report creation. Data migration refers to the movement of system metadata that drives the input fields used by portal users.

Compliance with the Australian Privacy Principals (APPs)

In reviewing the 2020 cloud DataLab PIA, the following APPs and their recommendations have been updated to reflect the myDATA release.

Australian Privacy Principle (APP)	**2020 Assessment**	**2020 Recommendation**	**June 2023 update**
APP1 – open and transparent management of personal information	Compliant, but further action recommended	2020 recommendation 2: Clearly describe the collection, storage, use and disclosure of personal information about users in collection notices and/or user agreements.	Completed The DataLab Privacy Notice has been updated to inform the users how the ABS is storing and administering the service.
APP2 – Anonymity and Pseudonymity	Compliant	No recommendation	Unchanged
APP3 – Collection of solicited personal information	Compliant	No recommendation	Unchanged
APP4 – Dealing with unsolicited personal information	Compliant	No recommendation	Unchanged
APP5 – Notification of the collection of personal information	Improvements to meet best practice	2020 recommendation 3: Create an APP5 notice to 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.	Completed The DataLab Privacy Notice has been updated to inform the users how personal information is used, stored and disclosed to Data Custodians.

Australian Privacy Principle (APP)	2020 Assessment	2020 Recommendation	June 2023 update
APP6 – Use or disclosure of personal information	Improvements to meet best practice	2020 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.	Completed The DataLab Privacy Notice has been updated to inform the users their personal information may be disclosed to Data Custodians though myDATA for ABS is administering the service.
APP7 – Direct Marketing	Not applicable	No recommendation	Unchanged
APP8 – Cross border disclosure	Not applicable	No recommendation	Unchanged
APP9 – Government related identifiers	Not applicable	No recommendation	Unchanged
APP10 – Quality of Personal Information	Compliant	Compliant in 2020: 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.	Unchanged myDATA portal allows user to provide and update their own personal information.

Australian Privacy Principle (APP)	2020 Assessment	2020 Recommendation	June 2023 update
APP11 – Security	Compliant	2020 recommendation 6: Implement any outcomes arising from security assessments to assure the continued security of personal information of DataLab users.	Ongoing The ABS will continue to maintain its IRAP certification for myDATA until the next security assessment cycle takes place. The ABS is committed to supporting 2-yearly IRAP. Any outcomes arising from security assessments will be reviewed and implemented accordingly to continue assuring the security of personal information of myDATA users.
		2020 recommendation 7: Create a deletion and retention policy specific to DataLab user accounts and related personal information.	Completed. The ABS has developed an internal DataLab deletion and retention policy relating to user accounts and related personal information. The policy governs ABS' approach to deletion and retention of personal information, consistent with requirements of the Records Disposal Authority guidelines. In addition to this recommendation, the ABS routinely review the need to retain personal information relating to users to ensure currency of process.

Australian Privacy Principle (APP)	2020 Assessment	2020 Recommendation	June 2023 update
APP 12 – access to personal information	Compliant	No recommendation	The ABS handles information about myDATA users in accordance with the ABS Privacy Policy for Managing and Operating Our Business. Portal users have control of their personal information through the myDATA portal. Portal users are responsible for maintaining and correcting personal information contained in their profile.
APP 13 – correction of personal information	Compliant	No recommendation	The ABS handles information about myDATA users in accordance with the ABS Privacy Policy for Managing and Operating Our Business. Portal users have control of their personal information through the myDATA portal. Portal users are responsible for maintaining and correcting personal information contained in their profile.

Next steps

The ABS is committed to protecting users' privacy and ensuring data security when handling personal information. As the DataLab continues to adapt and evolve to meet user expectations, the ABS will continue to review and improve the management tools which underpin the application and approval processes to preserve privacy, ensure data security and increase utility.

The ABS will update the PIA as necessary to convey changes to how user information is collected, accessed, and disclosed.

Appendix A: Application of the Five Safes Framework with context of myDATA release

Safe	**DataLab**	**myDATA portals**
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.	Workflow is unchanged, myDATA will facilitate the user's training enrolment and completion of legally binding undertakings and declarations
Safe Projects	Users must detail the purpose for which they will use the data. This can be compared to what analysis results are produced (see Outputs).	Workflow is unchanged, myDATA will modernise the process for project proposal creation, assessment, approval, and closure
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.	Unchanged
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 B: ACRONYMS AND GLOSSARY

Acronyms

Acronym	Term
ABS	Australian Bureau of Statistics
PIA	Privacy Impact Assessment
CRM	Customer Relationship Management (CRM) tool
myDATA	My Data Approvals To Access
OAIC	Office for Australian Information Commissioner's
ICT	Information Communication Technology
APPs	Australian Privacy Principles
IRAP	Independent Security Registered Assessors Program
AAD	Azure Active Directory

Glossary

Term	Description
ABS DataLab	A virtual cloud-based environment created by ABS for enabling scaled analysis of sensitive data and the use of contemporary and cost-effective data science tools.
Microdata	Data in a unit record file that provides detailed information about people, households, businesses or other types of entities.
Personal information	As defined in section 6(1) of the Privacy Act.
Five Safes Framework	A multi-dimensional approach to management disclosure risk which poses specific questions to help assess and describe each risk aspect (or safe) in a qualitative way.
Sensitive information	As defined in section 6(1) of the Privacy Act.

Term	Description
Australian Privacy Principles	Principles contained in the Privacy Act 1988 (Ch) that regulate the way we collect, store, provide access to, use and disclose personal information.
2020 Cloud DataLab PIA	The ABS conducted a Privacy Impact Assessment of Cloud DataLab in 2020. The Cloud DataLab PIA and ABS response are published on the ABS website.
2022 and 2019 MADIP PIA Updates	The ABS conducted a MADIP PIA Update in 2022. The 2022 MADIP PIA Update and MADIP
myDATA	The platform used to collect and manage DataLab project information. Software is run from the Microsoft Dynamics 365 customer relationship management tool
products	A name relating to the data files accessed in DataLab
Portal users	Registered people containing personal information as defined in section 6(1) of the Privacy Act.
Okta	A third-party identity management solution
CAPTCHA	Completely Automated Public Turing test to tell Computers and Humans Apart
Azure Active Directory	Azure Active Directory (AAD), part of Microsoft Entra, is an enterprise identity service that provides single sign-on, multifactor authentication, and conditional access to guard against 99.9 percent of cybersecurity attacks.
Penetration testing	A penetration test, colloquially known as a pentest or ethical hacking, is an authorized simulated cyberattack on a computer system, performed to evaluate the security of the system