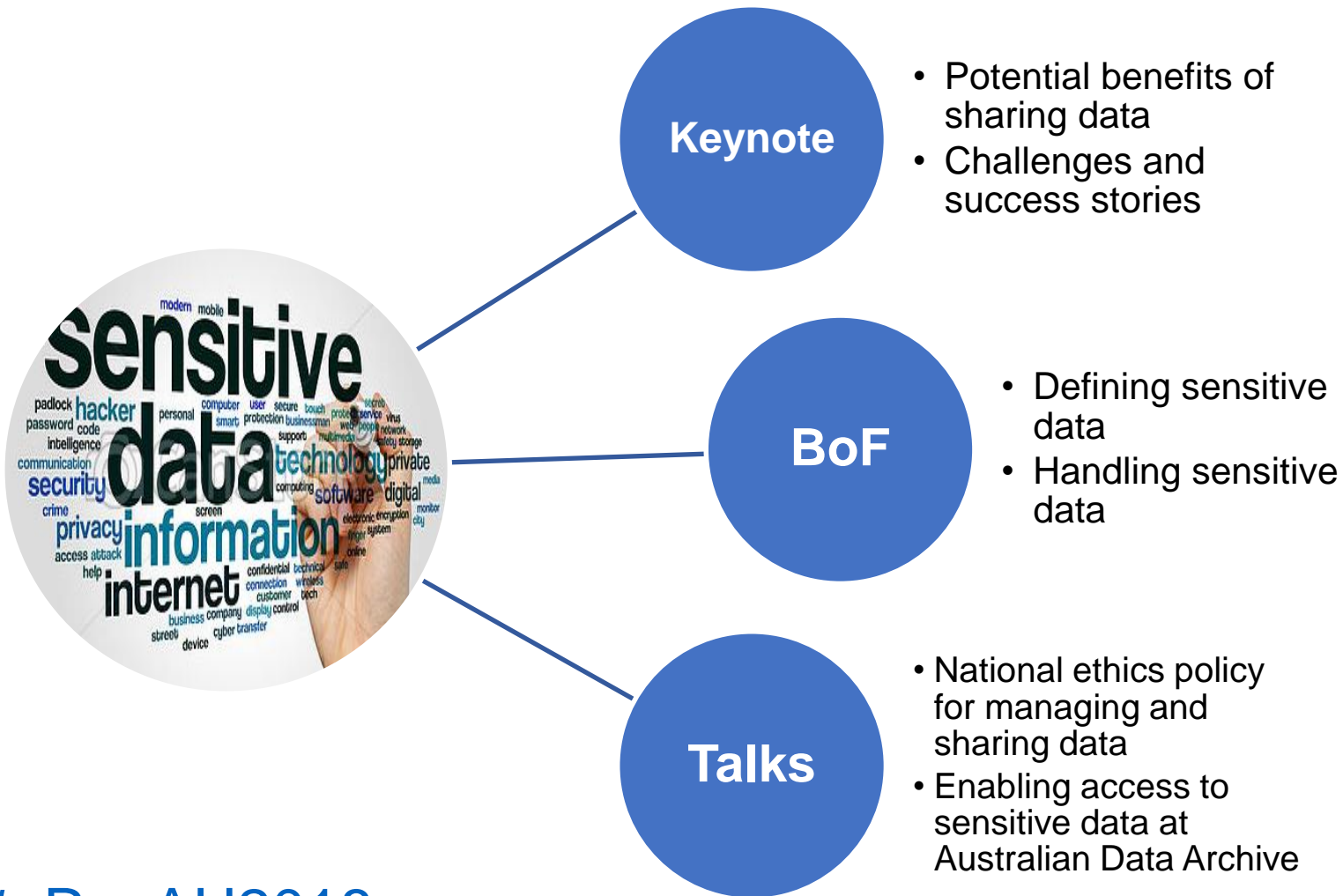


# Enabling data driven outcomes: Safe havens

Anitha Kannan  
Jerico Revote  
Komathy Padmanabhan  
Prof. Paul Bonnington  
Dr. Steve Quenette



# Setting the context..



## #eResAU2018

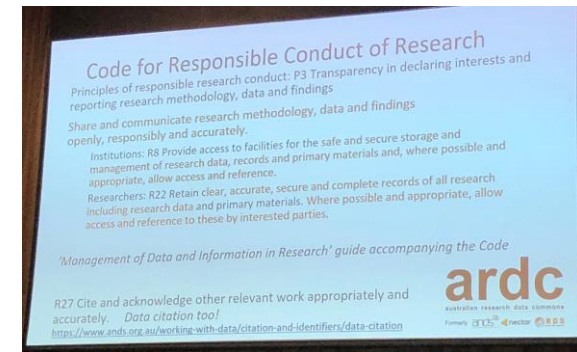
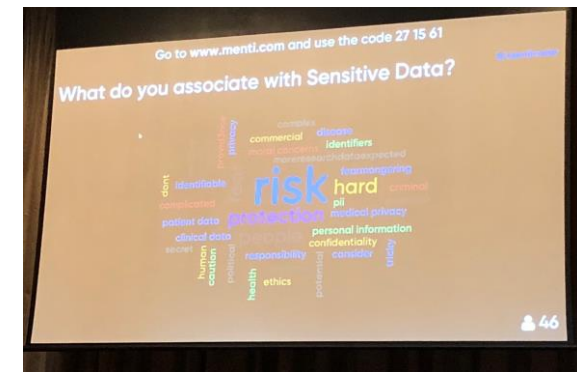
### References:

1. Enabling Data sharing for research, intelligence and public benefit, Prof. David V Ford, Swansea University medical school
2. Sensitive Data – How do you do yours? [BoF, Generating, Collecting and Moving Data](#) Frankie Stevens<sup>1</sup>, Jeff Christiansen<sup>2</sup>, Kate Le May<sup>3</sup>, Steve McEachern<sup>4</sup>, Angela Gackle<sup>5</sup>, Ingrid Mason<sup>1</sup>
3. Changes in national ethics policy for managing and sharing human research data [Managing, Publishing and Finding Data](#) Kate LeMay<sup>1</sup> <sup>1</sup>Australian Research Data Commons, Canberra, Australia, [kate@may@ands.org.au](mailto:kate@may@ands.org.au)
4. Enabling access to sensitive data at the Australian Data Archive [Managing, Publishing and Finding Data](#) Dr Steven McEachern<sup>1</sup>, Ms Janet McDougall<sup>1</sup>, Ms Marina McGale<sup>1</sup>

<sup>1</sup>Australian Data Archive, Acton, Australia

We also know . . .

## #DATA SAVES (& IMPROVES) LIVES

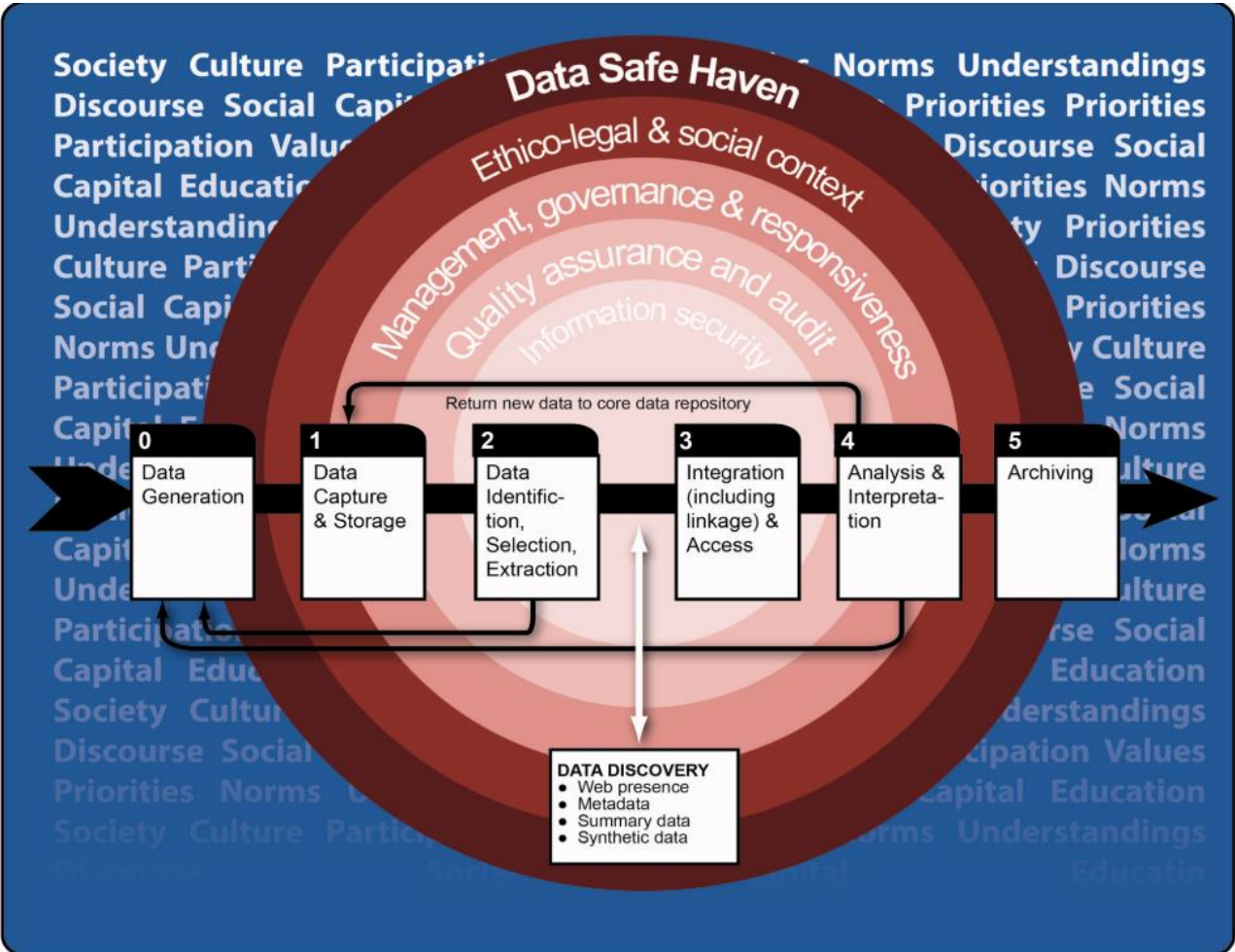


# This talk..

- Quick intro to data safe havens
- Monash eResearch Centre's approach
- Underlying principles
- Our Journey
  - Innovation pathway
  - Collaboration pathway
- Future landscape



# What are data safe havens?



## Proposed criteria for a Data Safe Haven

### Data maintenance and release should be socially acceptable and appropriate

Criterion 1	Consistent with formal ethical and legal requirements
Criterion 2	Responsive to emerging issues
Criterion 3	Discoverable and accessible
Criterion 4	Transparent and accountable

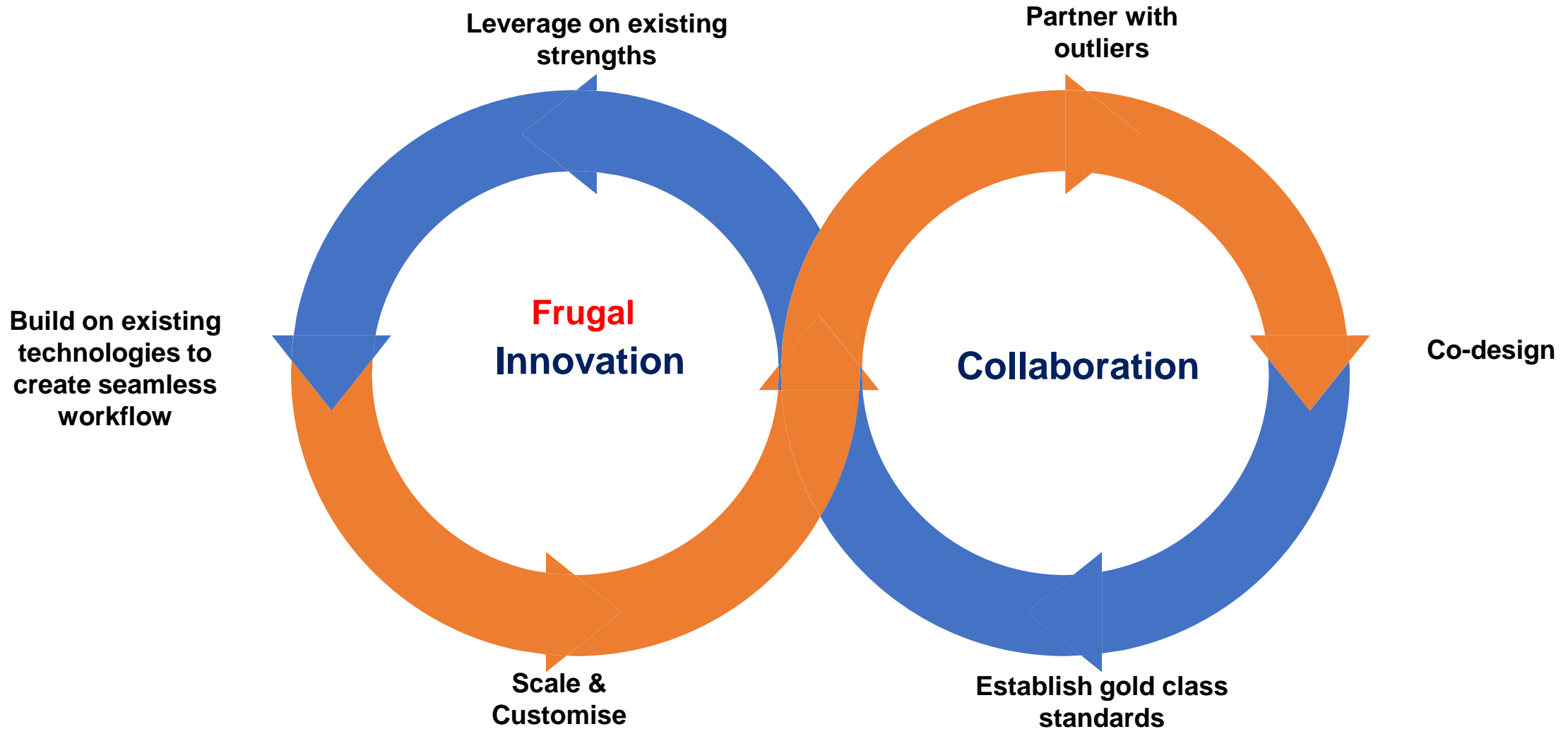
### Data should be veritable

Criterion 5	Data and metadata fidelity
Criterion 6	Quality assurance and control
Criterion 7	Curation and archiving
Criterion 8	Reliable availability including backup
Criterion 9	Effective audit

### Data should be safe and secure

Criterion 10	Preserve confidentiality, integrity and availability of the repository
Criterion 11	Appropriate secure access to individually identifying data
Criterion 12	Appropriate protection of individually identifying data

# Our 2-pronged approach



# Underlying principles



Reference: Desai, Tanvi; Ritchie, Felix; Welpton, Richard (2016). ["Five Safes: designing data access for research"](#) (PDF). Bristol Business School Working Papers in Economics: Footnote 1.

# How we started

## Big data mining market segmentation of ANZ Bank EFTPOS data

[Leave a reply](#)

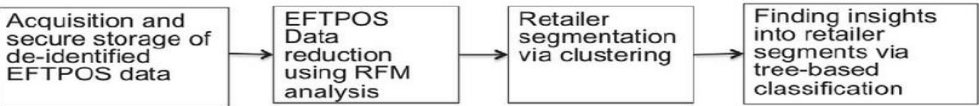
In Australia, the big 4 banks receive large amounts of Electronic Funds Transfer at Point of Sale (EFTPOS) transaction data on a daily basis, but despite this, this information-rich data are not stored nor analysed. The fact that EFTPOS data is both very large and very messy makes it difficult for banks themselves to gain visibility of the characteristics of the stakeholders of the data.

That changed in 2014, when a researcher in Monash's Faculty of IT, Dr. Grace Rumantir, approached us for assistance in accessing/building a secure analysis environment for a data mining project on a collection of commercially sensitive EFTPOS data obtained through an award winning collaboration with the Australia and New Zealand Banking Group (ANZ). To our knowledge this is the first time market segmentation analyses have been applied to such a large amount of EFTPOS data anywhere in the world.



Australian Government  
Department of Industry, Innovation,  
Climate Change, Science, Research  
and Tertiary Education

### TWO-TIERED MARKET SEGMENTATION FRAMEWORK



2014

Secure VLs for mining the EFTPOS data of ANZ Bank for market segmentation analysis

## Sleep Diary Report

Date of report: Thu 20 Apr 2017

### Summary

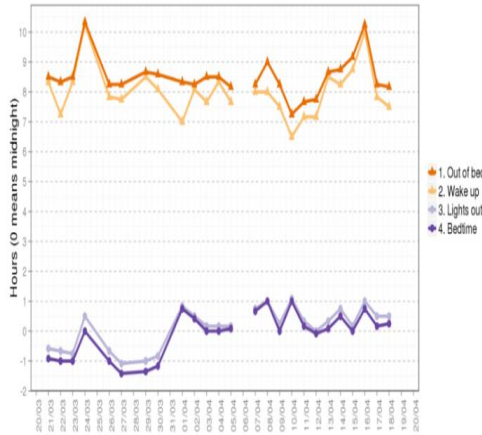
This report is based on information you entered on the Sleep Diary. Your overall sleep over the past week is summarized below, and the figures in this report displays a number of aspects of your sleep on a daily basis.

On average during the past week, you:

- Stayed in bed for 8.6 hours.
- Spent 7.2 hours' sleep per night.
- Took 46 minutes to fall asleep.
- Spent 7 minutes awake at night.
- Had a sleep efficiency of 84.5%.

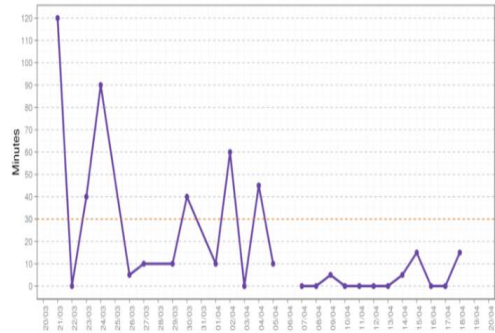
### Sleep Timing

The following figure shows what time you have gone to bed, turn off the light intending to sleep, woke up, and got out of bed.



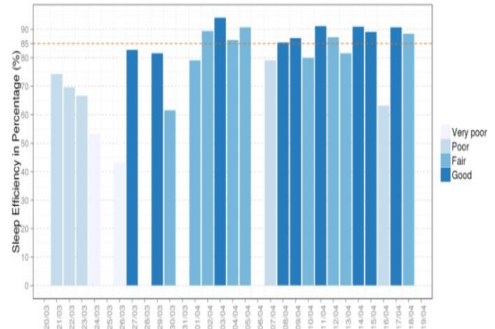
### Nighttime Awakening

The following figure shows the amount of nighttime awakenings you had each day.



### Sleep Efficiency and Quality

The following figure shows your sleep efficiency each day, which is calculated as the percentage of actual sleep you had while in bed. Each bar is coloured based on how you rated your sleep quality to be when you woke up the following day. The darker the bar, the better you rated your sleep to be.



2015

Secure patient reports for patients and treating team from the survey data



# Innovation: Flagship ASPREE Safe Havens

Partnership initiative between Monash eResearch & ASPREE

2016- 2017



## Background

- ASPREE (ASPIrin in Reducing Events in Elderly) is the largest primary prevention aspirin study ever undertaken in healthy older people
- 19,000 Participants – ~ 16500 Australians and 2500 Americans
- 17 sub-studies including Genomics, Hearing, Neuro, Biobank etc.

## Challenges

- ASPREE required a secure, curated platform for analysis of a re-identifiable baseline data set. Applicants included international non-Monash collaborators.
- Applications required approval by several committees before data could be provided.
- Limited resources available to manage import and export of data





# Innovation: Flagship ASPREE Safe Havens

Partnership initiative between Monash eResearch & ASPREE

2016- 2017



## Access Management System

**ASPREE** ASPIrin in Reducing Events in the Elderly - Access Management Site [Logout](#)

[Search](#) [Project Summary](#) [Project Details](#) [Project Administration](#) [Committee](#) [Account](#) [ASPREE Baseline Data Set Handbook](#)

> **Project Summary** You have logged in as jlockery\_ams

ID: P052	ACRONYM: ASPRE - RECRUIT	Group: ASPREE Investigator Group	Project: ASPREE Recruitment Methods Paper	Date initiated: 20 Oct 2017	Awaiting administration review and activation	
<a href="#">New Action</a>						
<a href="#">Outcome of ISC review</a>		16 Sep 2016	8001	Outcome <a href="#">Approved</a>		
<a href="#">Outcome of PPA review</a>		07 Aug 2015	8001	Outcome <a href="#">Approved</a>		
<a href="#">+ Comment...</a>						
<a href="#">Detailed Application</a>		07 Aug 2015	8064	<a href="#">EOI Details</a>	<a href="#">Biosketches</a>	<a href="#">Detailed Proposal</a>
				<a href="#">Submitted</a>	<a href="#">Submitted</a>	<a href="#">Submitted</a>
				<a href="#">Data Request</a>	<a href="#">ADVR</a>	<a href="#">Submit for review</a>
				<a href="#">Submitted</a>	<a href="#">Submitted</a>	<a href="#">Submitted</a>
<a href="#">Outcome of IEC review</a>		07 Aug 2015	8001	Outcome <a href="#">Approved</a>		
<a href="#">+ Comment...</a>						
<a href="#">Outcome of AUS Exec review</a>		07 Aug 2015	8001	Outcome <a href="#">Approved</a>		
<a href="#">+ Comment...</a>						
<a href="#">Expression of Interest</a>		07 Aug 2015	8000	<a href="#">Proposal/ Data/Dates</a>	<a href="#">Investigators/ Staff</a>	<a href="#">Sub-study</a>
				<a href="#">Submitted</a>	<a href="#">Submitted</a>	<a href="#">Submitted</a>
				<a href="#">Biospec Access</a>	<a href="#">Not Entered</a>	<a href="#">N/A</a>
				<a href="#">Submit for review</a>	<a href="#">Submitted</a>	

Copyright © 2015 Monash University ABN 12 377 614 012 - Caution - Privacy  
Maintained by [aspree.ams@monash.edu](mailto:aspree.ams@monash.edu) ASPREE AMS WEB Version 1.0x

- ☐ Developed by ASPREE based on their governance policies
- ☐ Data custodians publish data catalogue
- ☐ Researchers submit expression of interest for the data
- ☐ Application reviewed and approval by local & international steering committees, Ethics committee etc. as per governance policies
- ☐ Applicant notification on approval and instructions to login to safe havens

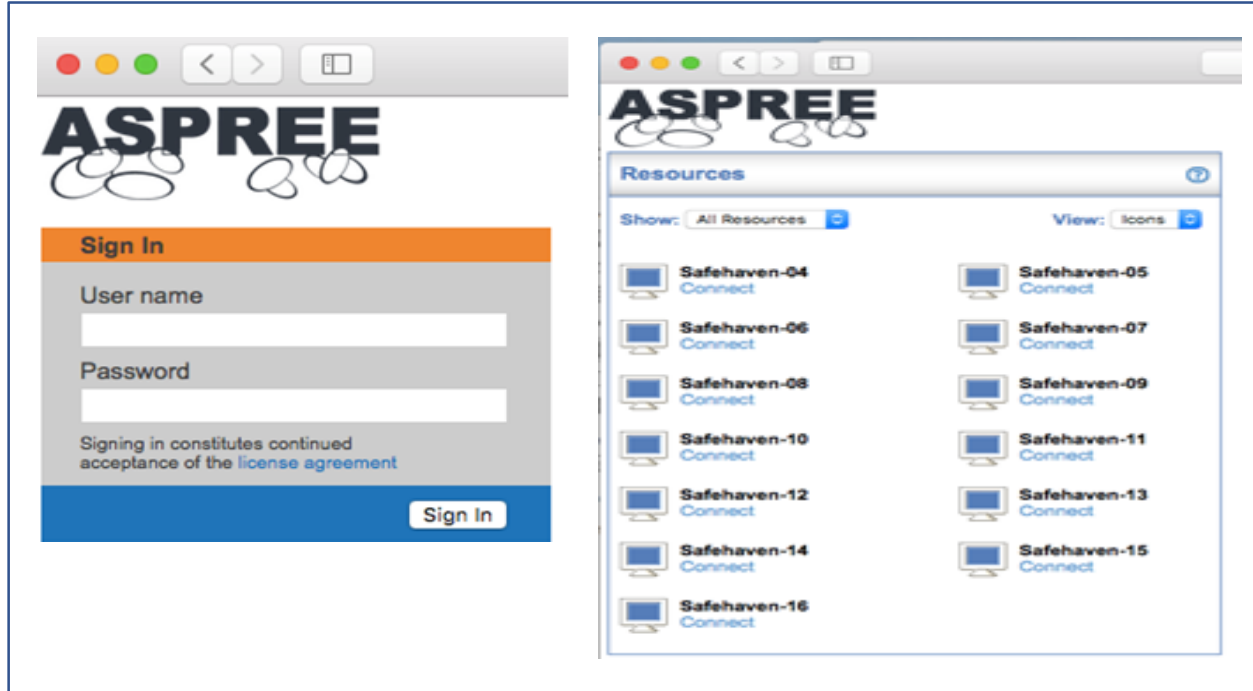
# Innovation: Flagship ASPREE Safe Havens

Partnership initiative between Monash eResearch & ASPREE

2016- 2017

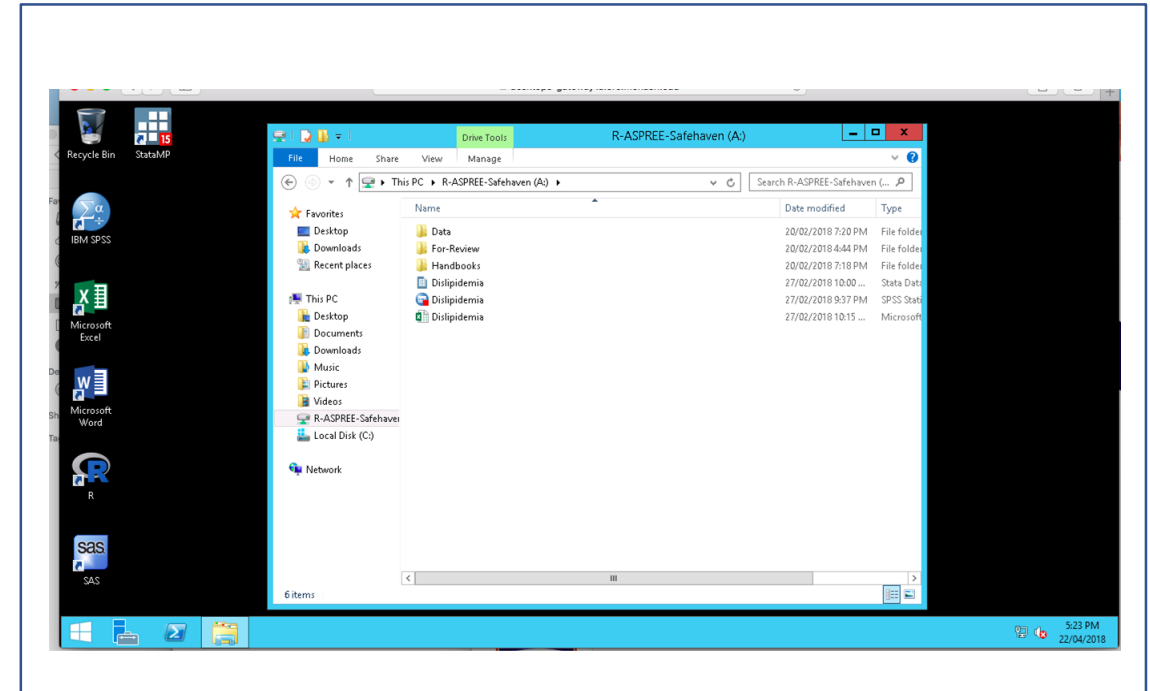


## Safe Havens access gateway



- ✓ Virtual lab of Safe Havens
- ✓ Accessed through 2-factor authentication to Monash VPN
- ✓ Potential to scale it

## Remote Safe Havens



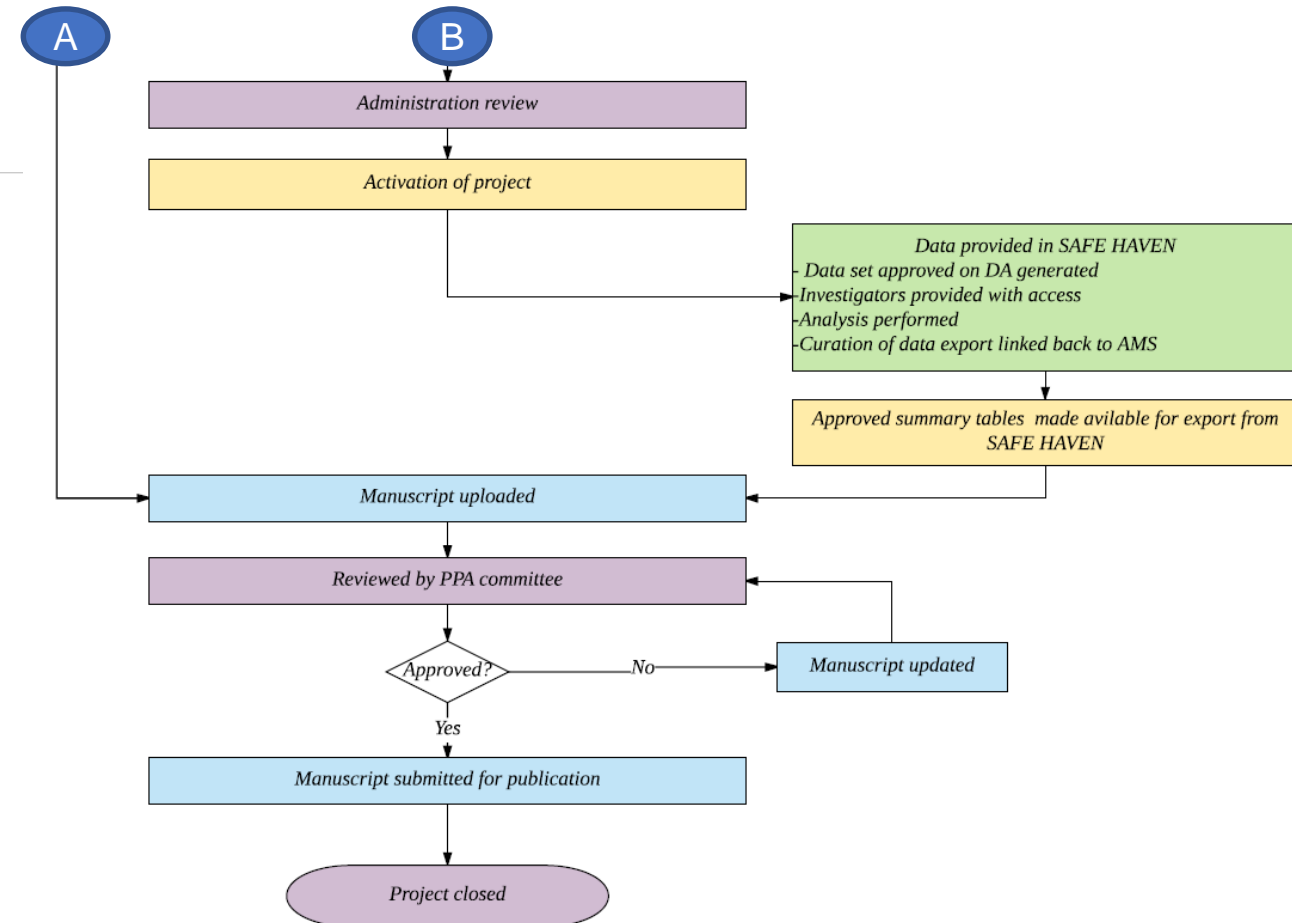
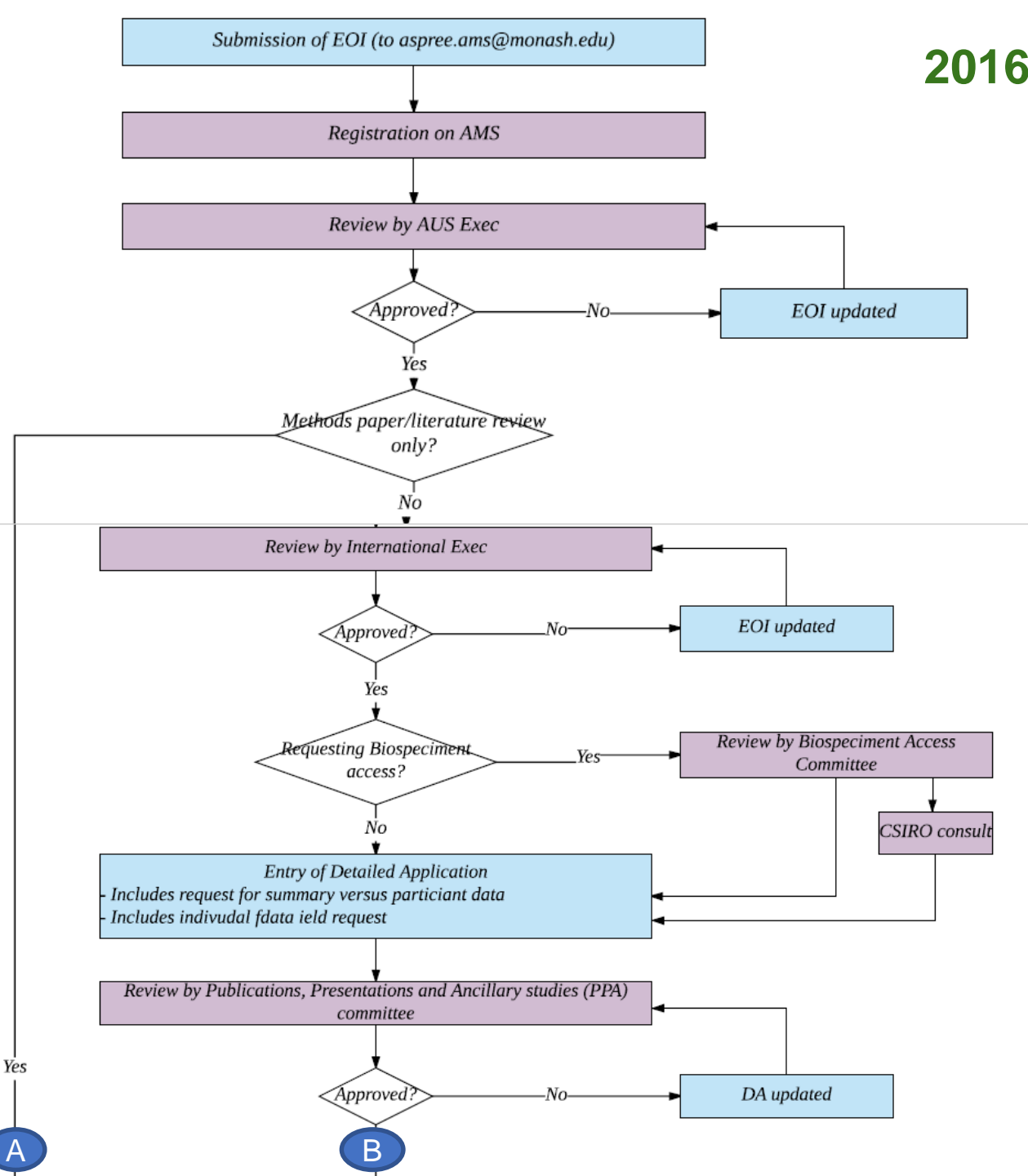
- ✓ Access to approved data
- ✓ Range of analysis tools
- ✓ Options to request export of analysed data

2016- 2017

# Workflow

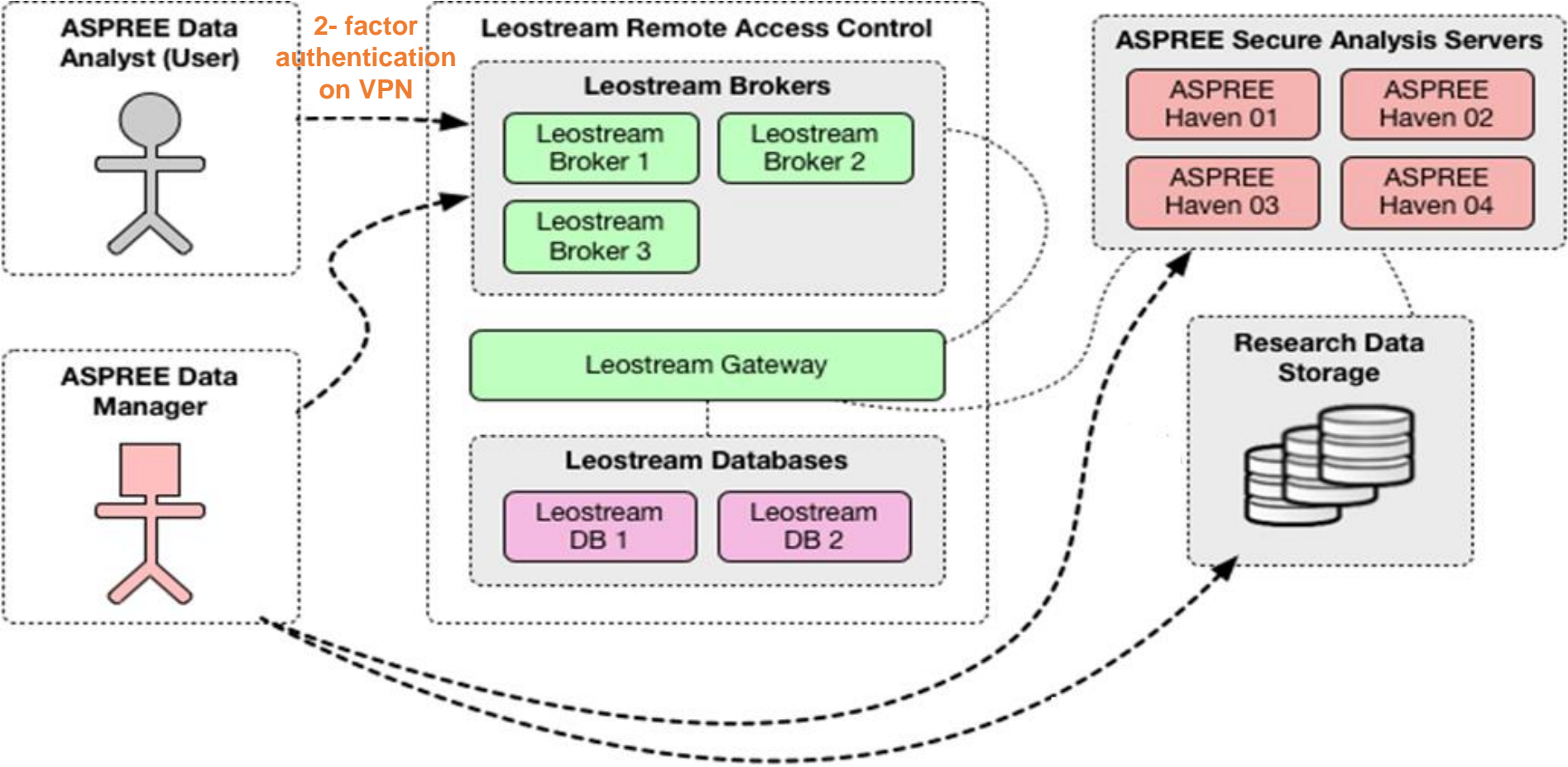
## KEY

- Blue shading indicates an activity completed by the applicant
- Purple shading indicates an activity completed by ASPREE AMS staff
- Yellow shading indicates an activity completed by the ASPREE data custodian
- Green shading indicates an activity completed by eResearch/SAFE HAVEN



# Underlying architecture

2016- 2017





# Extending to other large scale international clinical trials

## TrueNTH – Global Prostate Cancer Registry

2017



Details of 12 countries nominating to participate in TruenTH global registry and distribution of 113 participating sites across the 25 local data centres in 12 countries

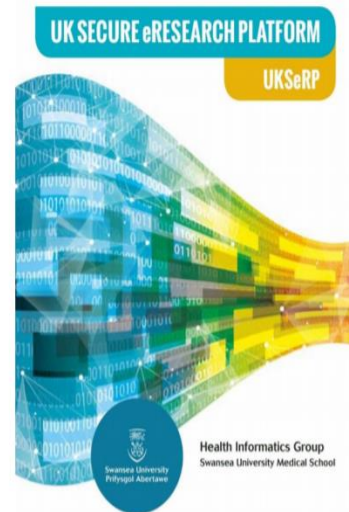
- Reversibly anonymised patient data from all the 12 countries are collected in central data co-ordinating centre at Monash
- Data received through a secure online portal
- Data sharing with central project coordination centre in US and with researchers through safe haven
- Key criteria for being central data coordination centre is having strict security control and capability to securely share back sensitive data with participating sites for analysis

## UK Secure eResearch Platform

The UK Secure Research Platform (UKSeRP) developed by the Health Informatics Group at Swansea

UKSeRP is an ISO27001 approved independent and customisable technology and analysis platform to allow multiple, complex datasets to be managed, analysed and shared safely, subject to legislative and regulatory requirements

Jones KH *et al*, The UK Secure eResearch Platform for public health research: a case study. Lancet (in press August 2016)



Helix, a health data ecosystem operated by expert staff and underpinned by world-class research infrastructure, spearheads a new era in healthcare research that will drive personalised medicine and significantly accelerate progress towards a healthier society. This capability is an initiative of Monash University working collaboratively with key health partners, including Alfred, Eastern, Monash and Peninsula Health.

### HEALTH DATA CAPABILITIES

- Secure data management - REDCap
- Governed environments for data sharing - Monash SeRP
- Secure virtual environments for statistical data analysis
- High performance computing, Research Cloud and Research data storage - eResearch Centre and MASSIVE

### EXPERTISE

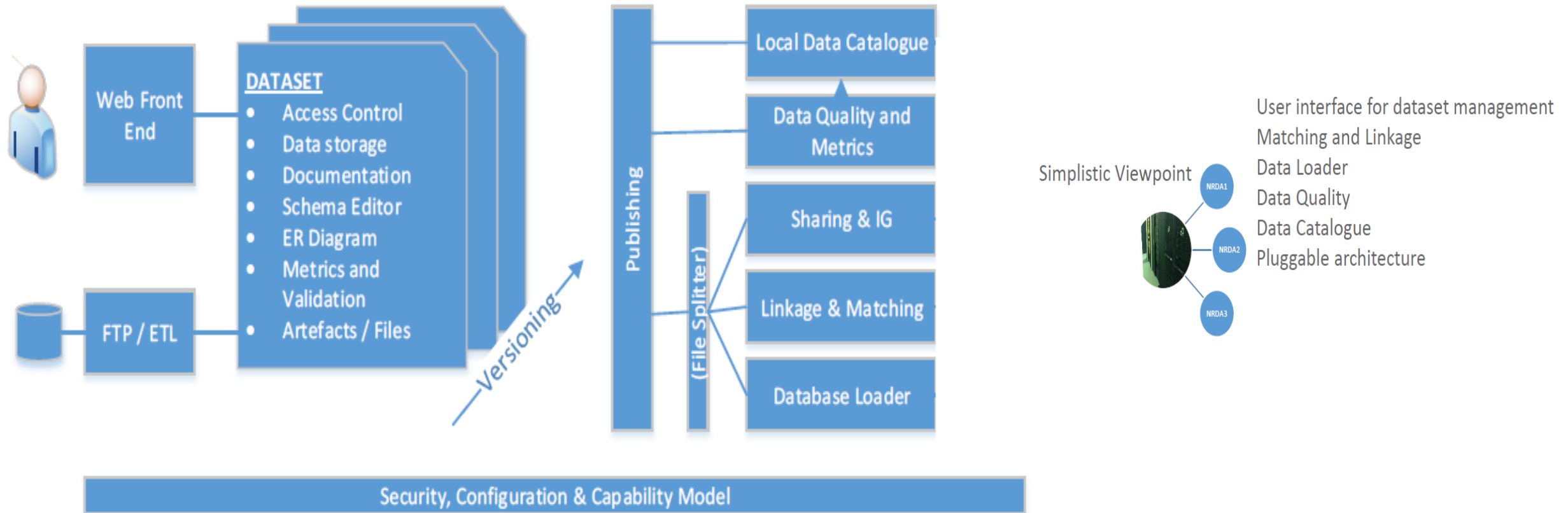
The combination of the sheer scale of Monash University's health datasets - which collate hundreds of thousands of data points within clinical trials, medical and population studies - underpinned by high-performance and specialised IT infrastructure will enable use of the latest developments in technologies like artificial intelligence (AI) offering

[https://saildatabank.com/wp-content/uploads/UKSeRP\\_Brochure\\_v1.5.pdf](https://saildatabank.com/wp-content/uploads/UKSeRP_Brochure_v1.5.pdf)

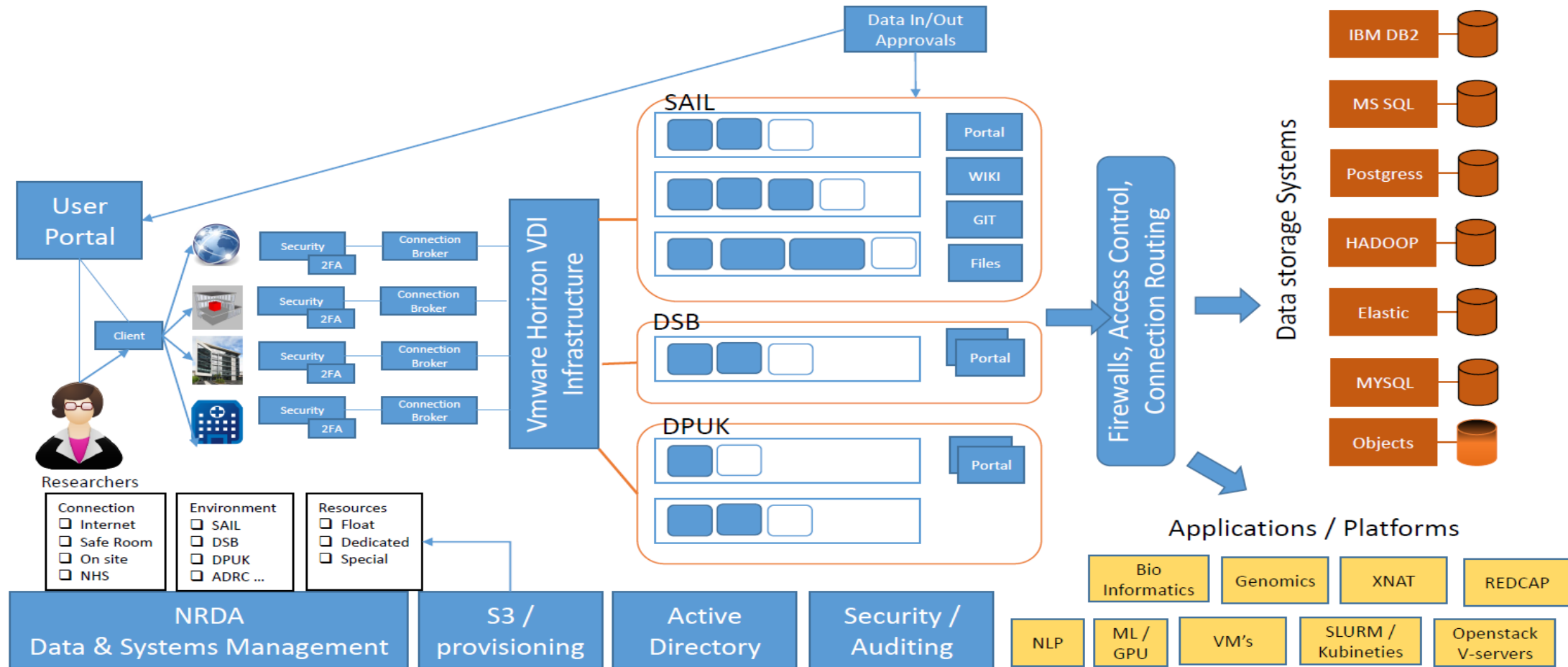
[https://www.monash.edu/data/assets/pdf\\_file/0008/1498337/Monash\\_Uni\\_Helix\\_Flyers.PDF](https://www.monash.edu/data/assets/pdf_file/0008/1498337/Monash_Uni_Helix_Flyers.PDF)

# UKSeRP NRDA

(Data custodian portal)



# UKSeRP - high level representation





# Monash SeRP - Pilot

2018

- Setup on Research Cloud & uses Research Data Storage
- Leverage existing softwares and licenses
- Multi Factor authentication
- Security policies and controls in place
- Includes Data anonymization, data linking, data quality analysis along with simple user interfaces for data custodians and users
- Range of analytical tools and capabilities including NLP and Machine learning



kpap0001 [Sign Out](#)  
To request a new password [click here](#)

## Connect to Research Platform

[Launch Platform](#)[File Out Requests](#)[File In Requests](#)

## Projects

Monash

Project Number	Project Name	Project Start Date	Project End Date
VIBES	Population birth cohort data	8/27/2018	1/1/2025

Monash SeRP

Jerico Revote [Sign Out](#) [Help](#)

[Projects & Datasets](#) [Local Data Catalogue](#) [Security](#)

### Projects & Datasets

Search  [Reset](#)

- Programme Monash
  - ICTACCT-Injury cases and TAC
  - TARN\_VSTR-TARN\_VSTR
  - VIBES-Population birth coh**

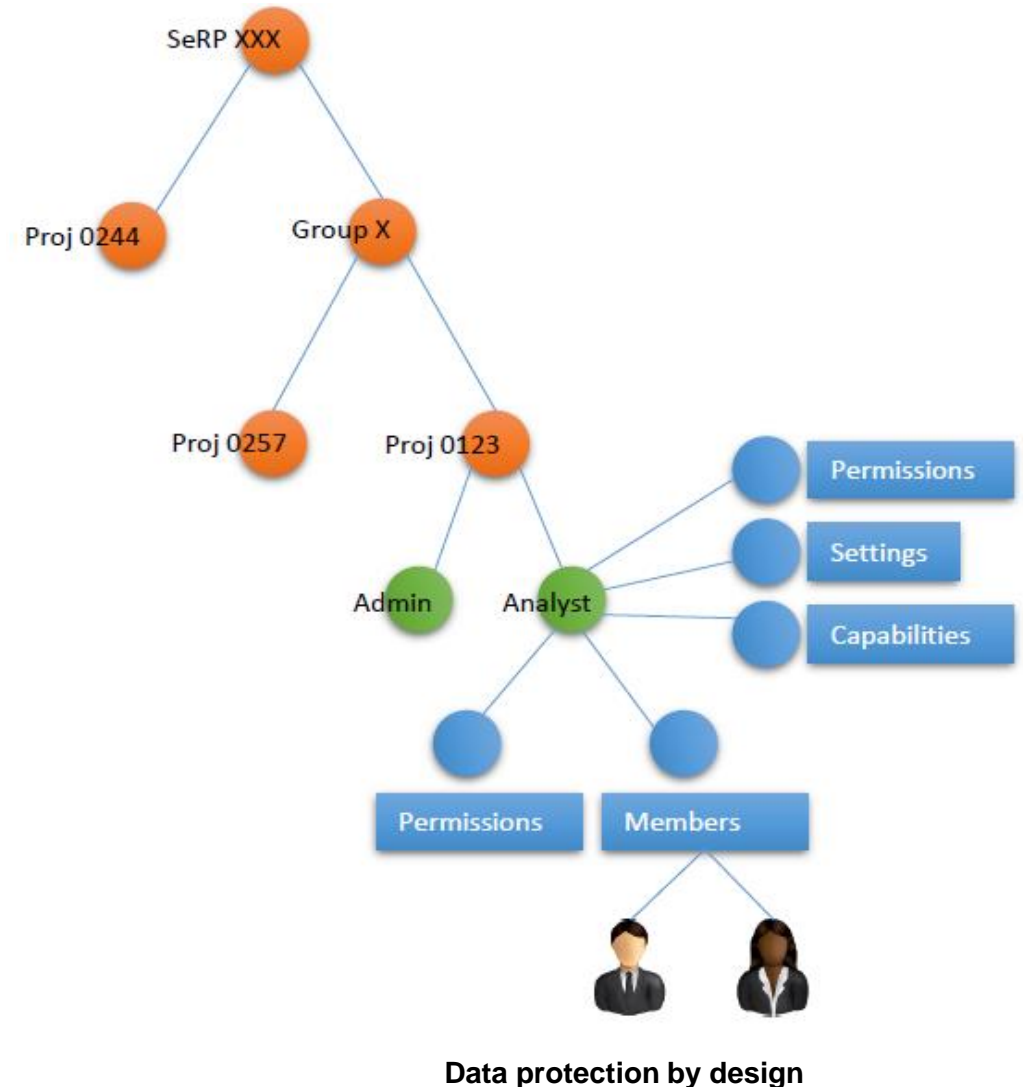
[Create New Dataset](#)

Filter by current version [Please Select...](#)

VSTORM\_data Theme

# Governance

- Researcher led implementation
- Driven by reference group with researchers across different schools
- Determines the technology and services to be made available on Monash SeRP
- Extendible to other disciplines
- Policies defines the access to data (projects), capabilities and administration
- Quality assurance framework (ISO9001) for management and operations



# Future landscape

Innovative collaborations

Information Security framework (ISO 27001)

Policies and Procedures

Extending to different disciplines

- Humanities

- Legal

- Business & Economics

- Health

- Policy & Governance

Data linking at scale

Potential for future federation for National Safe Haven infrastructure

Thank You

If you have any questions, please reach out to us on  
[safehaven-support@monash.edu](mailto:safehaven-support@monash.edu)