saxinstitute | SURE: Secure Unified Research Environment
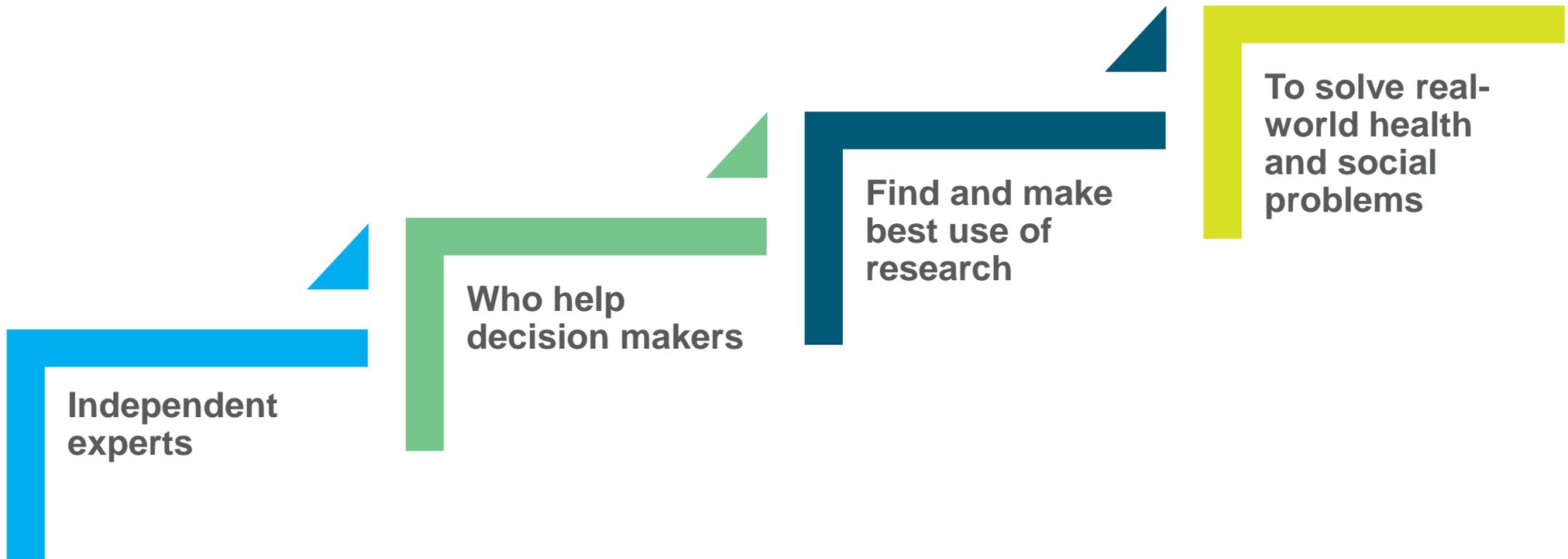
Secure Unified Research Environment

Using SURE to increase access and use of sensitive data

Dr Martin McNamara, Head, Research Assets
Sax Institute, Australia

Who are we?

Independent experts

Who help decision makers

Find and make best use of research

To solve real-world health and social problems

Driving research with real-world impact

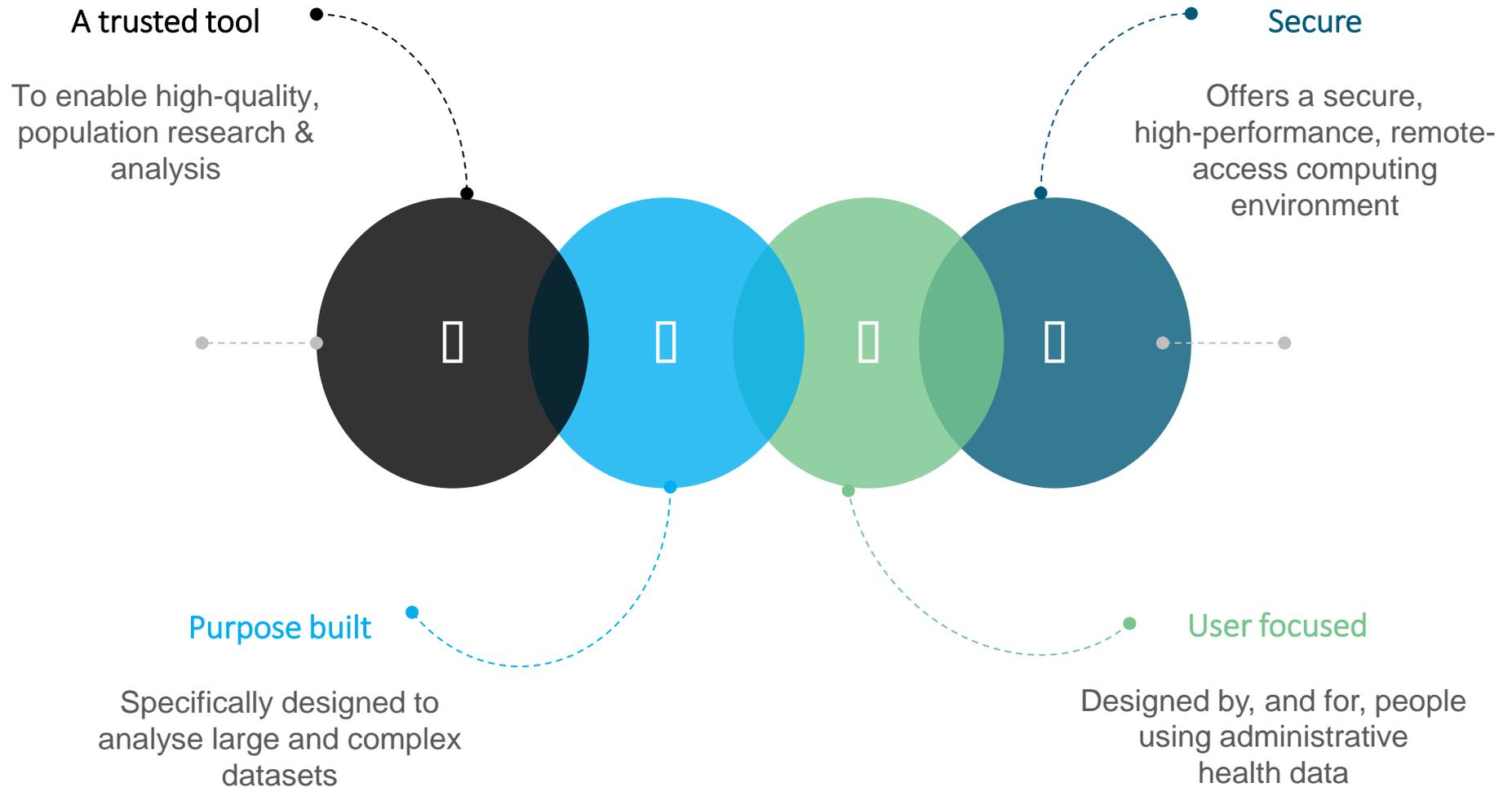we are paving the way for high-impact research

by building research platforms such as SURE

✓ we are a non-profit company limited by guarantee

✓ with expert research and policy staff

✓ and a unique structure – 47 universities and research groups as members

What is SURE?

- A secure, high-performance, remote-access computing environment launched in 2012, purpose-built for the analysis of large and complex datasets, e.g. hospital admission records, cohort studies

- Developed by the Sax Institute with funding from the Australian and NSW governments as part of the National Collaborative Research Infrastructure Strategy (NCRIS)

  - Australian Government Department of Education

  - NSW Ministry of Health and NSW Office of Science and Research

- Part of the Population Health Research Network (PHRN)

What is SURE?

A trusted tool

To enable high-quality, population research & analysis

Secure

Offers a secure, high-performance, remote-access computing environment

Purpose built

Specifically designed to analyse large and complex datasets

User focused

Designed by, and for, people using administrative health data

Development of SURE

Launched

In 2012

Funded

by a combination of Australian government (NCRIS funding) and user access charges

A member

of the Population Health Research Network (PHRN), which has developed national data linkage infrastructure

Why do we need SURE?

Australia has a comprehensive range of administrative data

To make best use of data, different data sources often need to be linked for a more complete picture

But data providers need assurances over data confidentiality and information security

Research projects are increasingly multidisciplinary, so SURE facilitates safe, secure collaboration across diverse organisations in Australia and internationally
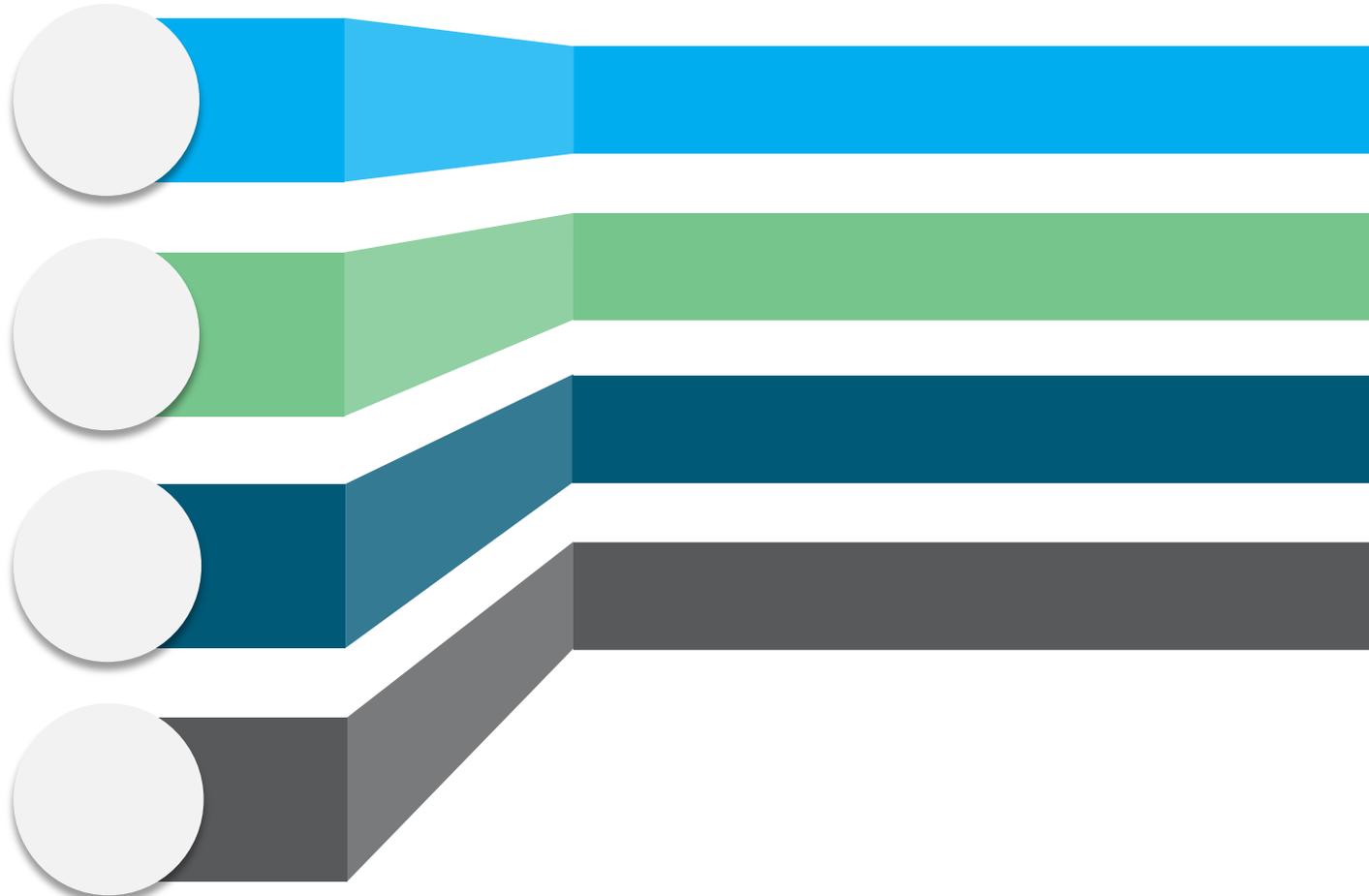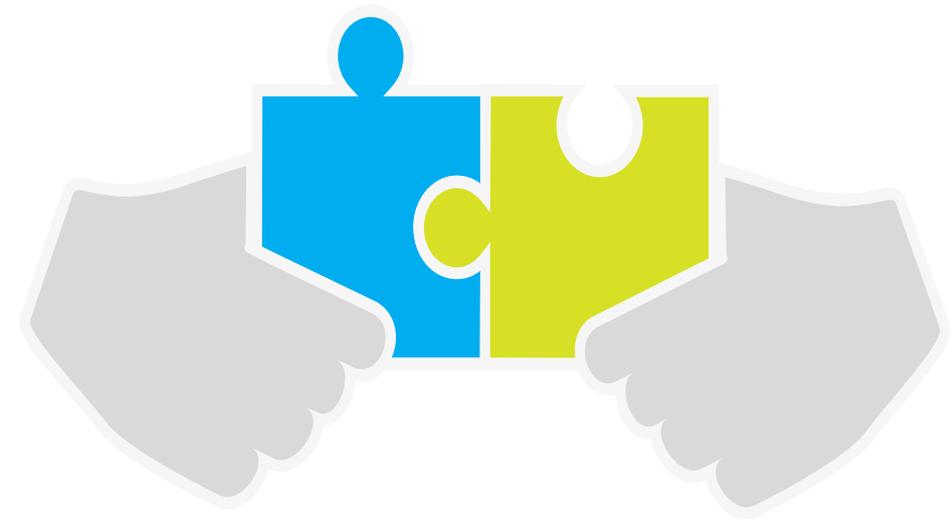
SURE is used for:
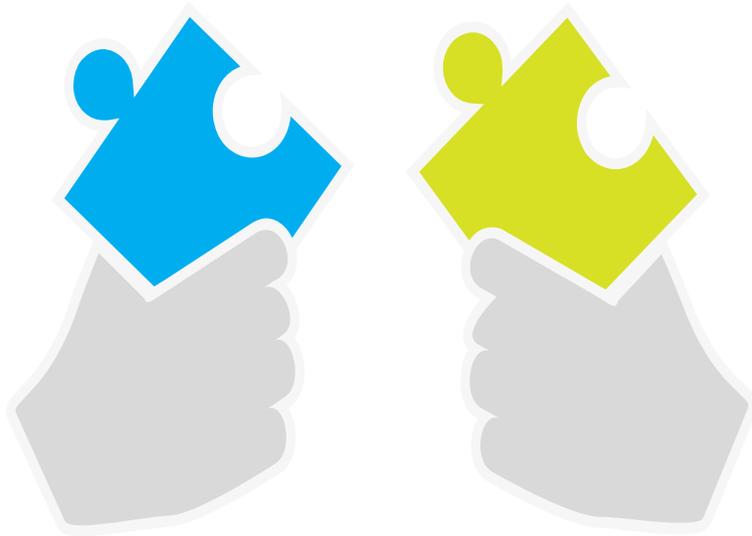
Linked data research

- Researchers use more than 30 different data collections in SURE
- Approvals provided per project
- Administrative data, cohort studies, clinical registries

Analysis

Analysing administrative data extracts:
- Data placed within SURE
- Extracts made available to authorised users
- Large numbers of users accessing the same data source

Features of SURE

- Trusted user model – users gain access to the data
- Controls on data implemented through:
  - Technical infrastructure
  - Governance framework – agreements
  - Training – mandatory for users

Features of SURE

1

2

3

4
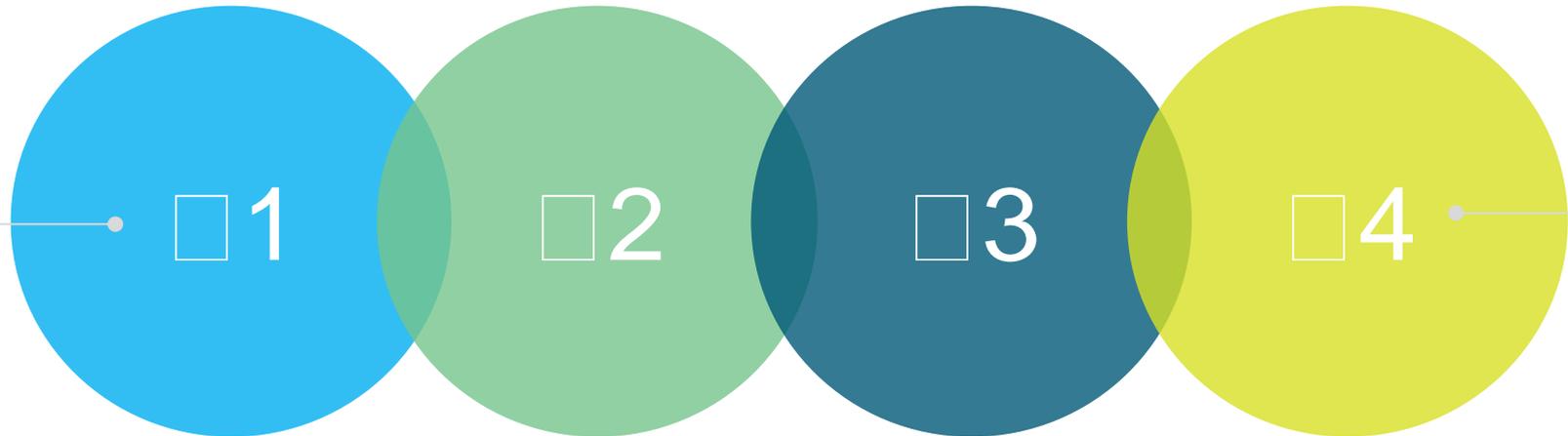
- Movement in and out of SURE tightly controlled

- Environment can be tailored with a range of software and analysis tools

- Strong governance framework

- Responsive technical and user support

SURE infrastructure

- SURE provides project workspaces that are accessed remotely over encrypted internet connections (using Citrix)

- Research team members have shared access to project workspaces

- Each project workspace exists in its own security perimeter within SURE so that data from different projects cannot be combined

- All data storage and analysis happens on dedicated SURE servers housed in a secure commercial data centre

- Movement of data in and out of SURE is carefully controlled via specially designed software called the Curated Gateway

SURE infrastructure
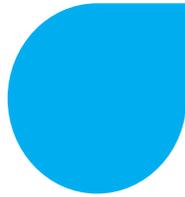
The Curated Gateway

Movement of data in and out of SURE is carefully controlled via specially designed software > files logged and subject to approval and audit

Dedicated Servers

All data storage and analysis happens on dedicated SURE servers housed in a secure commercial data centre

Custodian control

Data custodian can upload data directly to SURE from their own IT system

Curated Gateway

- A purpose-built, secure web-based application

- Provides the only way for files to move in or out of SURE (no Internet, no printing, cannot copy/paste, cannot copy to removable media)

- Files uploaded to the Curated Gateway by data custodian/ researcher for use in SURE are subject to review before being made available within SURE

- Outbound files uploaded to the Curated Gateway for use outside of SURE are reviewed by the data custodian or the study's chief investigator

- All files that pass through the Curated Gateway are logged and are subject to audit

SURE governance
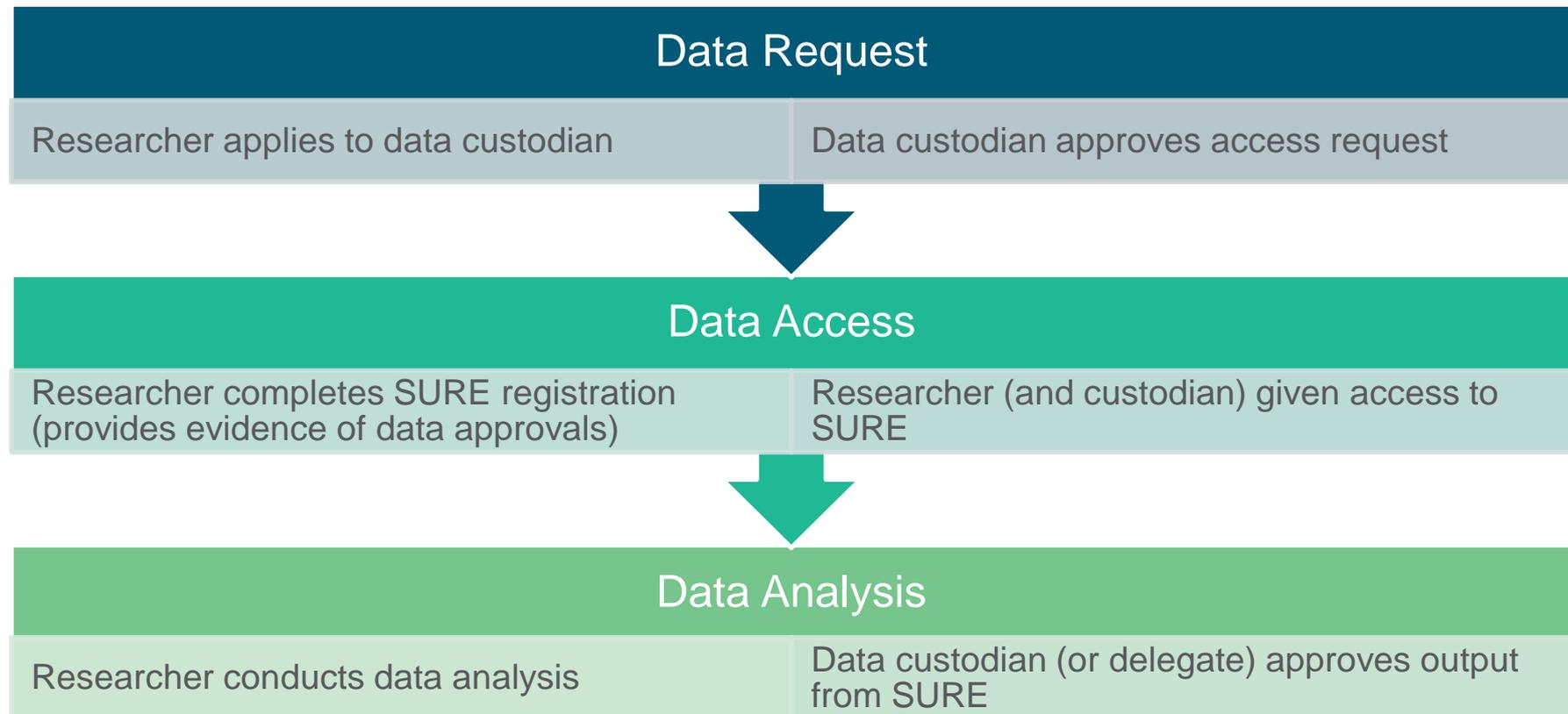
- Projects hosted by SURE must have approval from the relevant data custodian/s and a Human Research Ethics Committee (or equivalent legislative approval)

- Approvals specify the datasets, variables and level of confidentialisation

- Individual researchers AND their institutions must sign deeds of agreement regarding acceptable use, disciplinary arrangements before gaining access to SURE

SURE governance

Individual researchers AND their institutions must sign deeds of agreement regarding acceptable use before gaining access to SURE

SURE training programs are mandatory for new users
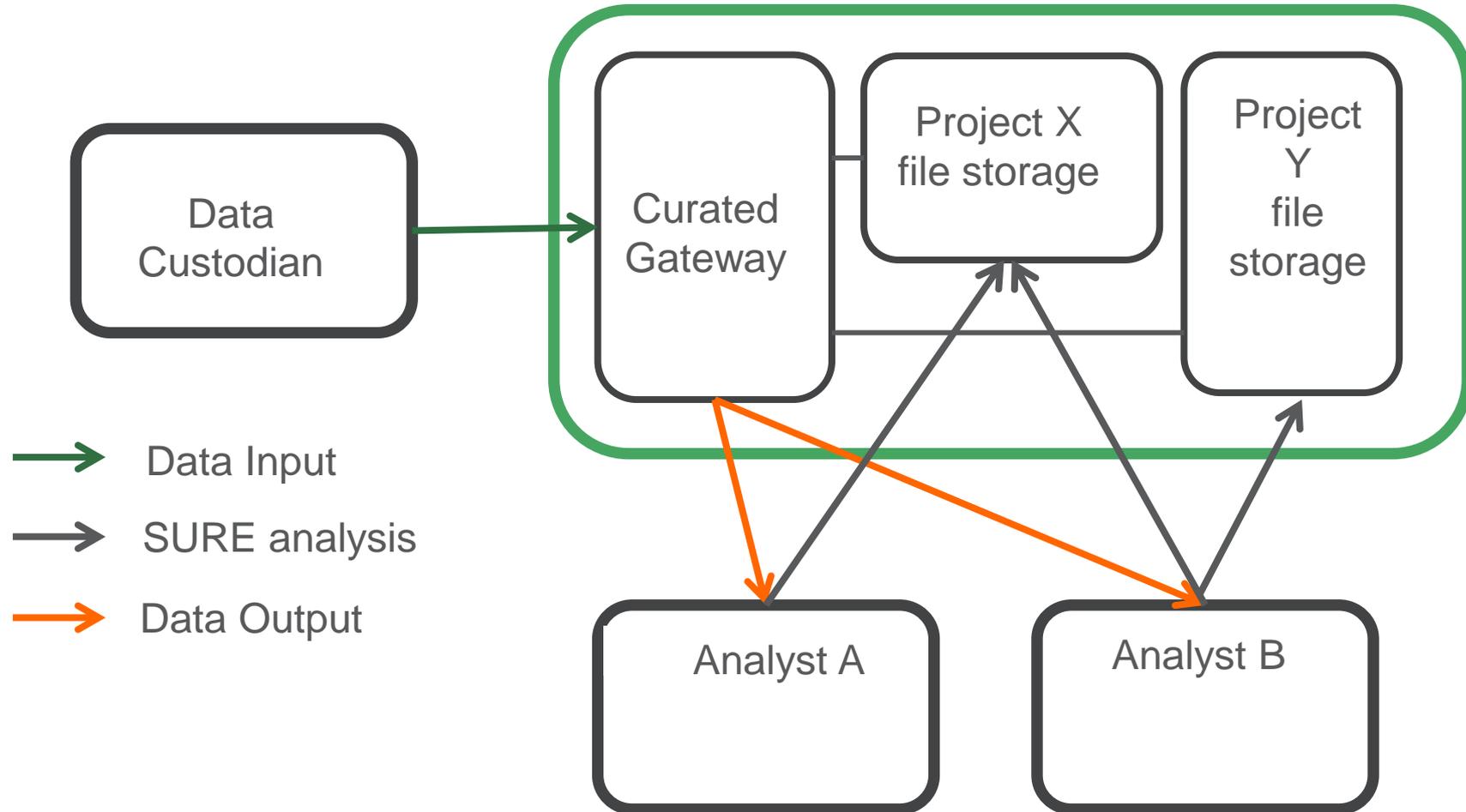
Training covers:

- Responsibilities of a researcher
- Information security
- Statistical disclosure analysis and control

Data access process

Data Request

Researcher applies to data custodian	Data custodian approves access request

Data Access

Researcher completes SURE registration (provides evidence of data approvals)	Researcher (and custodian) given access to SURE

Data Analysis

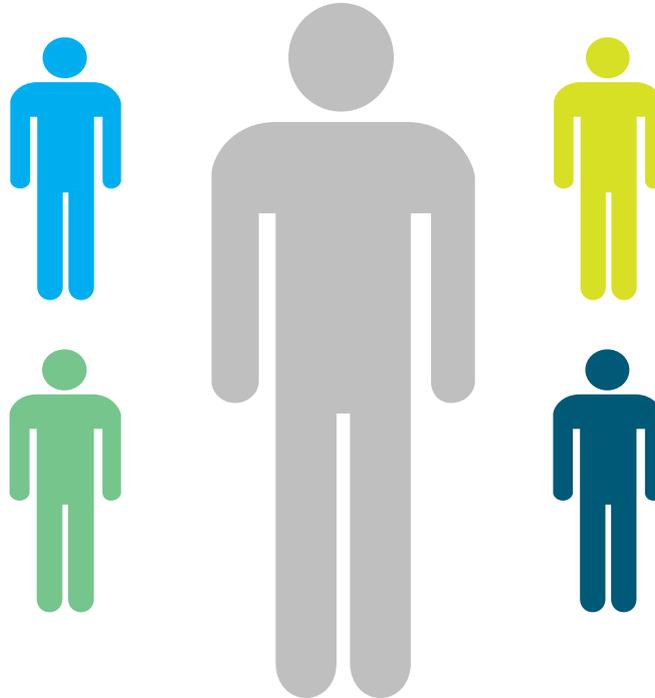| Researcher conducts data analysis	Data custodian (or delegate) approves output from SURE

Using SURE

SURE users

Data custodians

❑ Australian Institute of Health and Welfare

❑ Department of Social Services

❑ State and territory health departments

❑ Cohort studies and clinical registries

72 active study workspaces involving hundres of users

❑ Researchers based in most states in Australia

❑ International users

Government agency case study:
Cancer Institute NSW

Manages data registries and funds cancer research

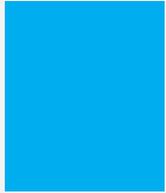Multidisciplinary teams with policy and academic staff

Uses SURE as a **data custodian** and uploads registry data for use by researchers

Uses SURE for **research using linked data** to enhance and better understand CINSW data sources
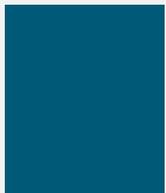
Access arrangements

For individual projects, access charges are paid on a workspace and user basis

- Workspace establishment and annual maintenance
- User costs based on computing resources allocated
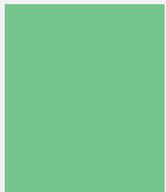
Program and institutional licenses available

Additional services include workspace customisation, data and access management

Levels of access

	Users	Workspaces	Examples of use
Project	1-5	1	Linked data research project with a small number of authorised users
Program	5-10	3	Analysis of multiple administrative data sources for different purposes
Institutional	10+	5	A combination of projects and programs across a range of departments / research teams

Why SURE has been successful

- Strict controls on access but ability to tailor arrangements for different needs
  - For example: Permissions for Curated Gateway approval
- Balance between needs of custodian and research
  - Custodians need to be assured of data security but also have resource limitations on setting up their own secure facilities
  - Researchers (by and large) want to do the right thing but can be let down by their own computing environment or lack of knowledge

Why SURE has been successful – security

- Each user can only access data in SURE that they are authorised to see.

- Once data are loaded into SURE, they cannot be copied, downloaded or transmitted by any means, including the Internet.

- Movement of all files into and out of SURE is curated, recorded and audits are undertaken.

- Data are stored in a high-security facility and not in the user's own computing environment. Rather, users access the data via a virtual workspace.

- Operating since 2012 with no data breaches.

Why SURE has been successful - custodians

- Risk management - help comply with open data policies but at the same time giving them total control over who accesses their data and for what purpose.

- Resourcing - streamlines researcher access to data and removing the need for "bespoke data services".

- Cost - One, high-security location to analyse sensitive data is cheaper and more time efficient than establishing similar access facilities in multiple agencies.

- Public good - increases the value of the data they hold by helping generate insights that will benefit the community.

Why SURE has been successful - researchers

- Access - easier, faster access to large-scale linked datasets previously unavailable to them.

- Big science - data scientists and other researchers to collaborate from anywhere across Australia or the world on large-scale, innovative research projects

- Power - uses the latest analytic software and tools

- Risk management - protects researchers and their institutions from the risks associated with storing sensitive human research data.

THANK YOU

Website: https://www.saxinstitute.org.au/our-work/sure/

martin.mcnamara@saxinstitute.org.au