The Farr Institute of
Health Informatics Research

Farr Institute of Health Informatics Research

Harnessing Data for Health Science and e-Health Innovation

Striking the balance between data privacy and data use

Associate Professor Kerina Jones

Swansea University

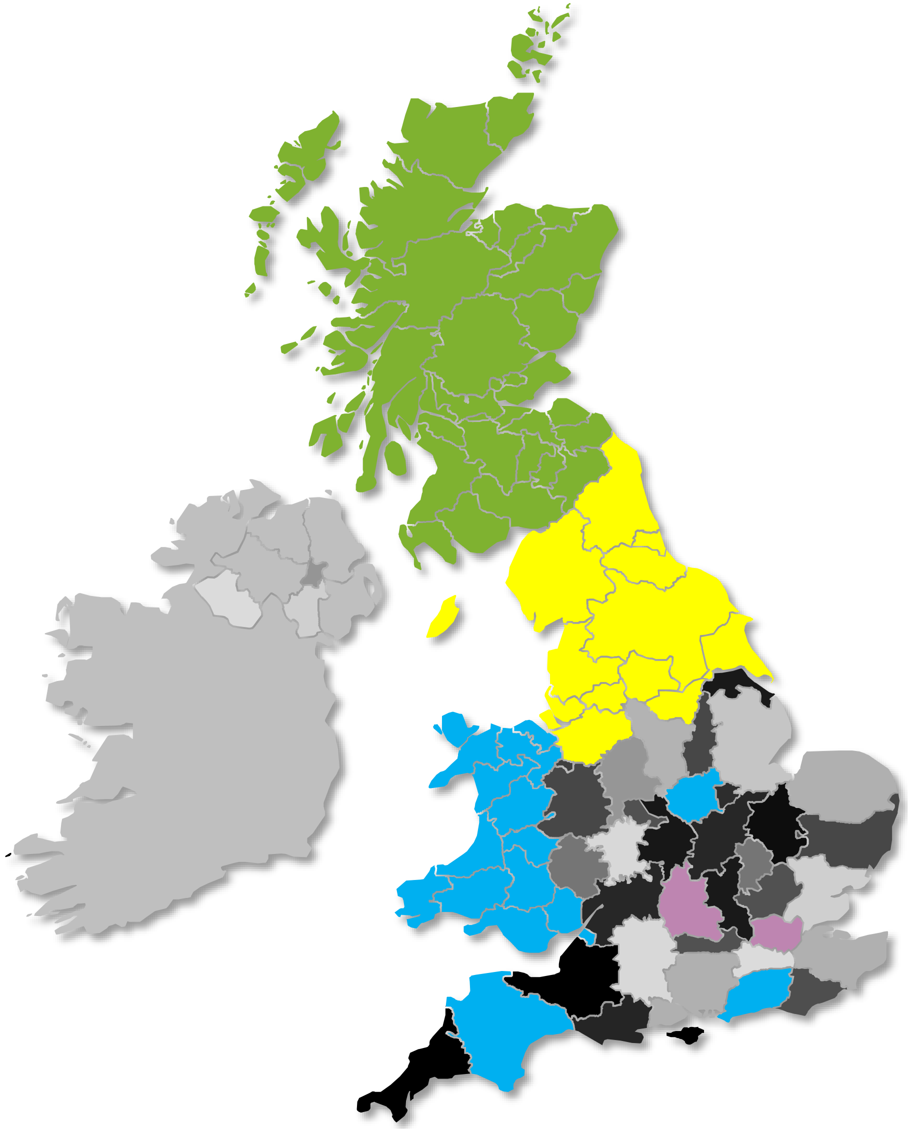
Overview

- Introduction to the Farr Institute for Health Informatics Research
- Data governance models -
  - Secure Anonymised Information Linkage (SAIL) databank
  - UK Secure Research Platform (UKSeRP)
- Health data privacy/utility projects
- Questions/discussion

What is Farr?

- The Farr Institute is a UK-wide research collaboration involving 21 academic institutions and health partners in England, Scotland and Wales
- Publicly funded by a consortium of ten organisations led by the Medical Research Council
- Committed to delivering high-quality, cutting-edge research using Big Data to advance the health and care of patients and the public
- **CIPHER** – Centre for the Improvement of Population Health through eRecords Research

Health Informatics Research Centres



Map Source: www.m62.net

🟢 **Scotland**

Dundee, Glasgow, Edinburgh, St Andrews, Aberdeen, Strathclyde, MRC HGU, NHS NSS

🟡 **HeRC N8**

Manchester, York, Lancaster, Liverpool, Sheffield, AHSNs

🔵 **CIPHER**

Swansea, Bristol, Cardiff, Exeter, Leicester, Sussex, NWIS, Public Health Wales

🟣 **UCL Partners**

UCL, LSHTM, Queen Mary, Public Health England

Farr
The Farr Institute of
Health Informatics Research

Innovative Governance Working Group

Researching and sharing best practice in the trustworthy reuse of health-related data and to promote solutions that meet the relevant regulatory and governance requirements without being unduly burdensome to research

Working with the owners and controllers of data to support the safe use of patient information for medical research across the UK, championing data protection, confidentiality and privacy

A collaborative working group led by Kerina Jones (Swansea) and the IG leads from all Farr nodes: Graeme Laurie (Scotland), Nathan Lea (London) and James Cunningham (Manchester)

The Farr Institute of Health Informatics Research

IGWG Outputs

A Review of Evidence Relating to Harm Resulting from Uses of Health and Biomedical Data

Prepared for the Nuffield Council on Bioethics Working Party on Biological and Health Data and the Wellcome Trust's Expert Advisory Group on Data Access by:

Professor Graeme Laurie,
University of Edinburgh

Dr Kerina H. Jones,
Swansea University

Ms Leslie Stevens,
University of Edinburgh

Dr Christine Dobbs,
Swansea University

Contents lists available at ScienceDirect
International Journal of Medical Informatics
journal homepage: www.ijmijournal.com
ELSEVIER

The other side of the coin: Harm due to the non-use of health-related data

Kerina H. Jones[a,*], Graeme Laurie[b], Leslie Stevens[b], Christine Dobbs[a], David V. Ford[a], Nathan Lea[c]

[a] *Data Science, Swansea School of Medicine, Swansea University, Swansea SA2 8PP, UK*
[b] *School of Law, University of Edinburgh, Old College, South Bridge, Edinburgh EH8 9YL, UK*
[c] *Centre for Health Informatics, University College London, Gower Street, London WC1E 6BT, UK*

- Implicated in the deaths of many thousands of people and potentially £billions in societal financial burdens
- Looking for a large, agile, polymorphic, lethal, black cat that most certainly <u>is</u> there.

http://www.sciencedirect.com/science/article/pii/S1386505616302039

The Farr Institute of Health Informatics Research

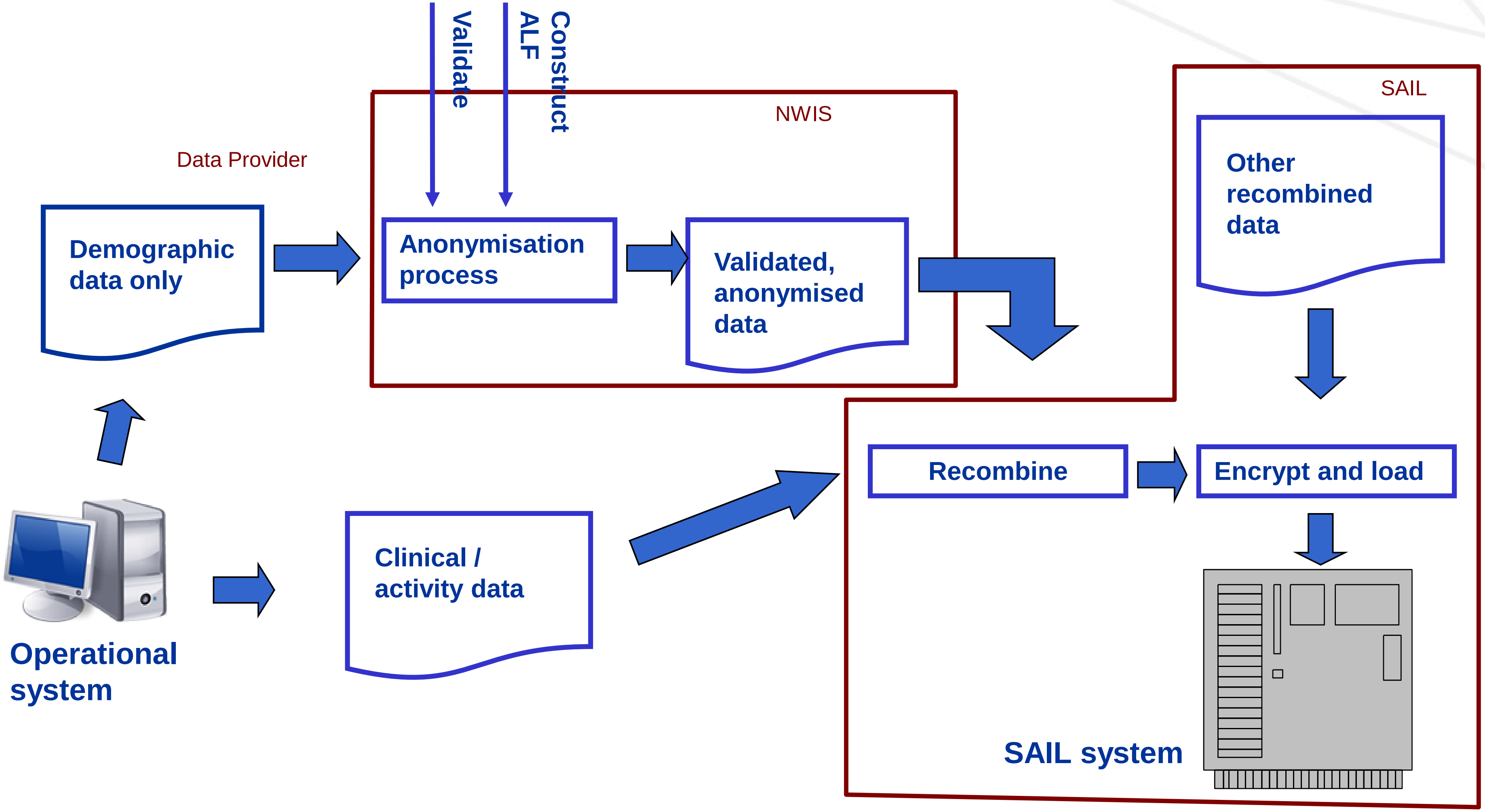
Secure Anonymised Information Linkage

- SAIL is a national Data Safe Haven – anonymised person-based data on the population of Wales
- It underpins CIPHER
- SAIL has anonymised data from: GPs, hospitals, ONS births and deaths, screening services, cancer registry, child health and more
- Linkable at the individual level and available for research
- Privacy by design - robust, proportionate data governance for the safe use of data

SAIL

http://www.saildatabank.com/

Separation principle



Validate

Construct ALF

NWIS

SAIL

Data Provider

Demographic data only

Anonymisation process

Validated, anonymised data

Other recombined data

Recombine

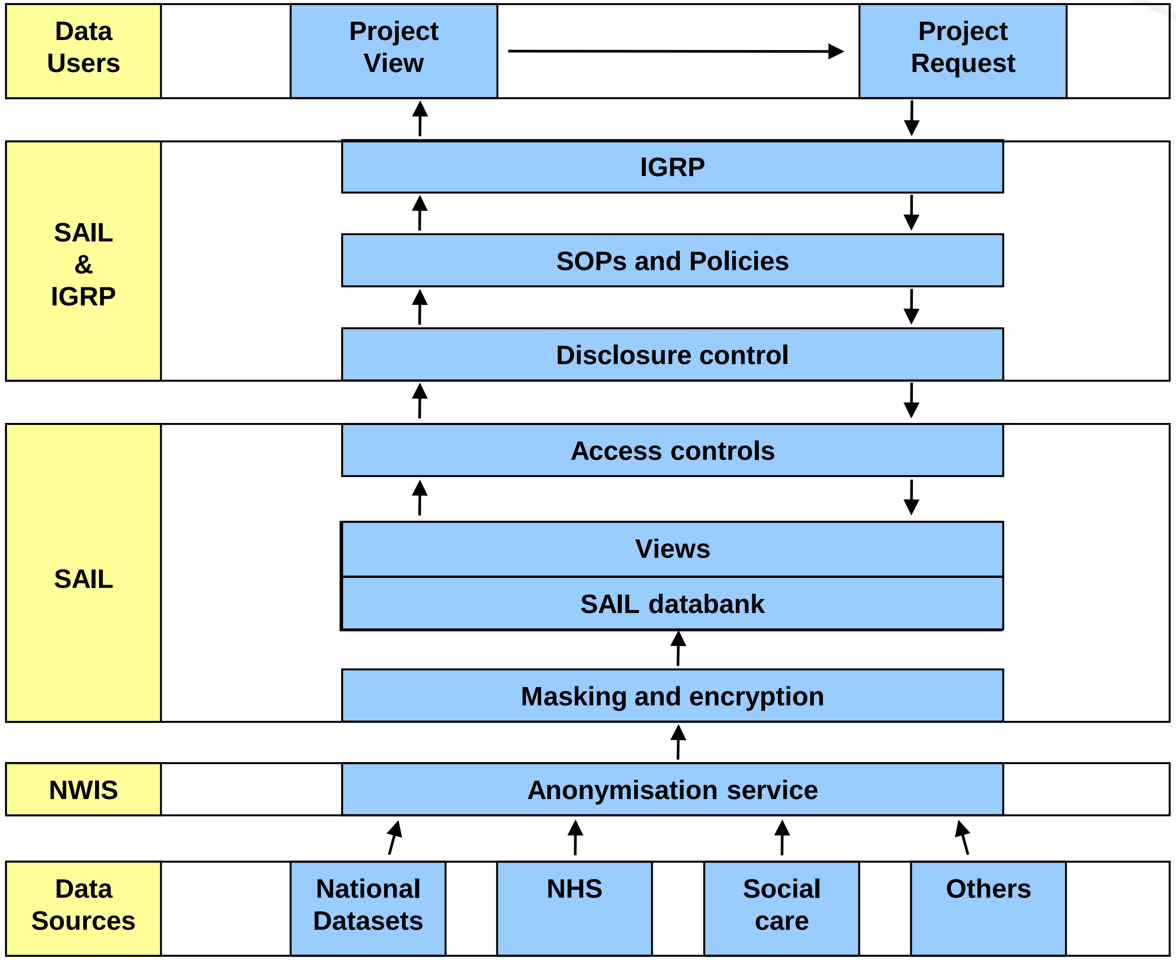
Encrypt and load

Operational system

Clinical / activity data

SAIL system

Enabling safe data access



Data Users

Project View

Project Request

SAIL
&
IGRP

IGRP

SOPs and Policies

Disclosure control

SAIL

Access controls

Views

SAIL databank

Masking and encryption

NWIS

Anonymisation service

Data Sources

National Datasets

NHS

Social care

Others

The Farr Institute of Health Informatics Research

Public Engagement

Consumer panel for data linkage research

- Provide a public perspective on data intensive research
- Guide on how to recruit people to study steering groups
- Provide views on data protection issues
- Discuss proposals for research
- Review information designed for a lay audience
- Advise on how best to engage with the public

The wider public

- Events
- Consultations
- Surveys
- Research projects

Farr The Farr Institute of Health Informatics Research

General Data Protection Regulation

New/significantly changed concepts

- Definitions of personal and sensitive data expanded
- Definition of pseudonymisation added
- Consent to be unambiguous and not assumed from inaction
- Parental consent required for use of online personal data about under 13 year olds
- Right to be forgotten and data portability
- Data protection by design and accountability
- Security breach communication law for data controllers
- Increased monetary penalties for breaches/abuses

- Derogations permitted at member state level

Farr The Farr Institute of Health Informatics Research

Advancing cross-centre research networks - ACoRN

- Data governance challenges in research collaborations across centres
- Scoping the barriers to safe data sharing across centres
- Identifying practical solutions
- Include international relevance
- Exploring the drivers that promote cross-centre research
- Removing barriers will prevent cross-centre research not happening, but drivers are needed to move it forward
- Identify principles for a decision-support framework to guide cross-centre research

The Farr Institute of Health Informatics Research

Genetic Data Integration - GeDI

- MRC-funded project: **Whole Genome Sequencing for Health and Wealth: A bid to establish the Wales Genomic Medicine Centre to contribute to the UK 100,000 Genomes Project and Strengthen Genomics in Wales**
- Variant data to be returned to the Wales Centre
- Sub-project: Genetic Data Integration (GeDI)
  - Types of data - sequences, variants, etc.
  - General and particular challenges
  - Current vs new data governance models
  - Ethical, legal and societal implications (ELSI)
  - Individual and kin privacy, now and future
  - Public engagement on views, expectations and acceptability

Genetic data governance

Particular challenges in working with genetic data:

- Sensitive personal information
- Information persistence – security now and future
- Implications
  - Disease risk/prediction
  - Life/motor insurance
  - Employment prospects
- Impact on kin
- The right to know and not to know – and know what?

- Responsibility to maximise data privacy AND data utility

– 30 - 80 random mutations enough to "individualize"

– Unique ≠ Identification

Mobile Phone Operator-Related Phone and Health Data Governance - MORPHeD

- Smartphone data are increasingly being used for research studies – especially with a geographical component
- Important to look at the data governance – the ethical, legal and societal implications of using these data
- Working with O_2 in connection with one of their platform developments for the health service
- Reviewing the literature on the use of smartphone data in health research
- Developing a method to critically appraise the data governance elements of research studies
- Working with the public to gauge their knowledge of passive and active collection of data via smartphones, the uses the data are put to and their views on acceptability

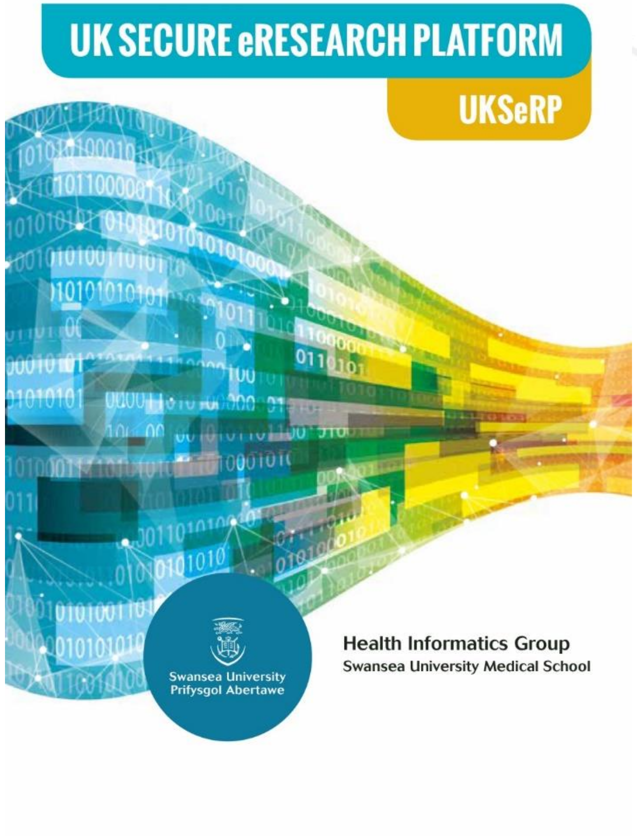
Farr The Farr Institute of Health Informatics Research

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)



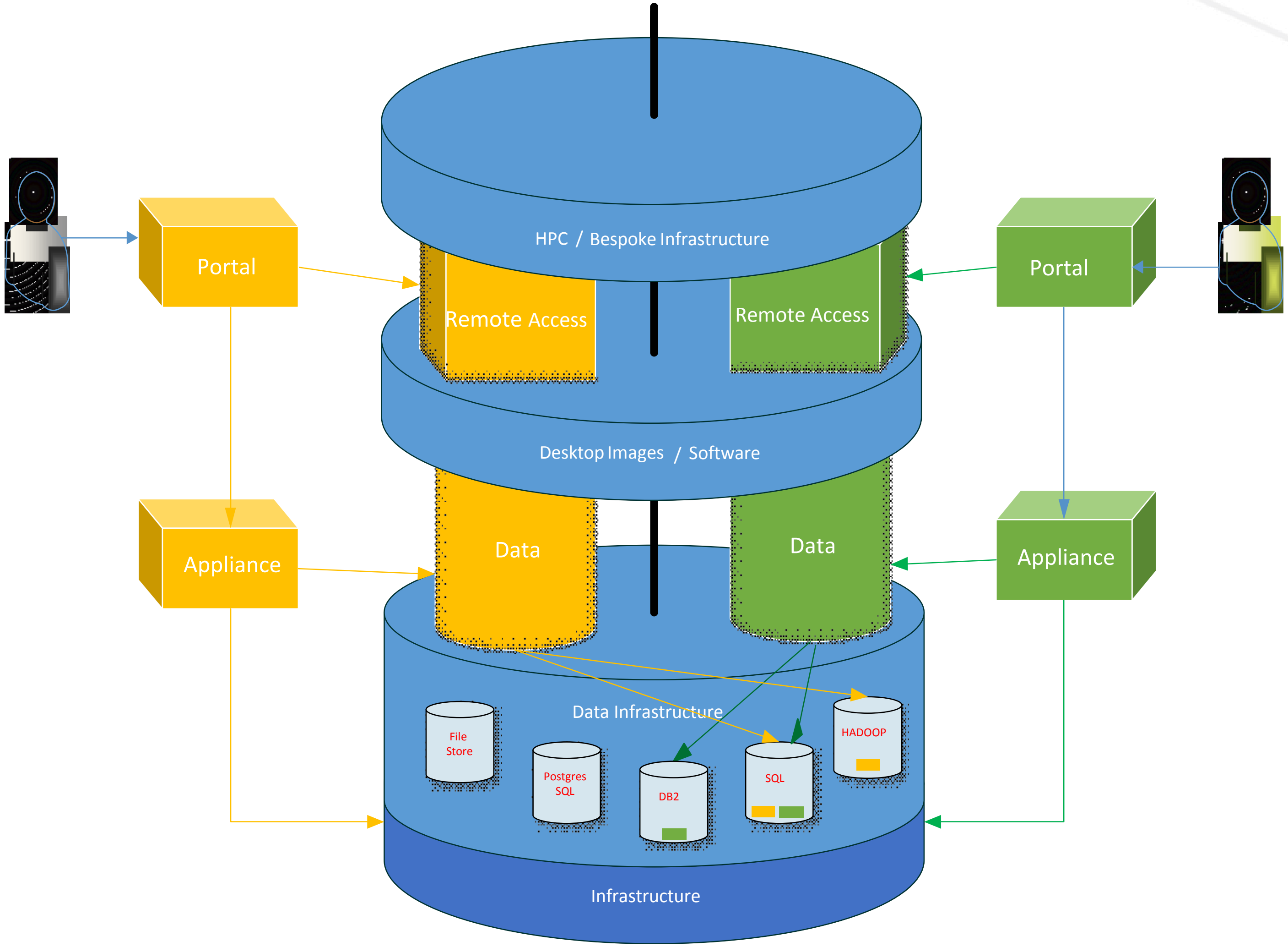
UK SECURE eRESEARCH PLATFORM

UKSeRP

Swansea University
Prifysgol Abertawe

Health Informatics Group
Swansea University Medical School

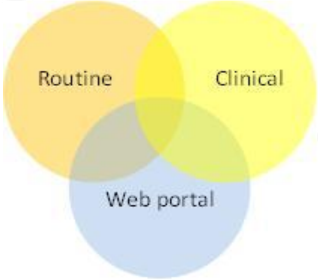
UK SeRP



The Farr Institute of Health Informatics Research

Some UKSeRP projects

- **MS Register**
  - Linking data from neurology clinics, routine sources and directly from people with MS



- **Dementia Platform UK (DPUK)**
  - Linking dementia genotypes to clinical phenotypes

- **Avon Longitudinal Study of Parents and Children (ALSPAC)**
  - Genetic and environmental risk factors of childhood obesity
  - Effect of maternal gestational weight gain on offspring DNA methylation
  - Association of SNPs in hearing loss

THANKS!

Over 100 colleagues in the Data Science Building at Swansea, and many more elsewhere

The Farr Institute of
Health Informatics Research

Any questions?

To find out more, please contact: k.h.jones@swansea.ac.uk
Thank you for listening

The Farr Institute of
Health Informatics Research