## HOW IS THE DATA KEPT SAFE?

All data which is submitted to SAIL is unidentifiable. The commonly-recognised identifying details are removed from the datasets. All work carried out at the Data Science Facility complies with Information Governance and data security. Every centre and project undergoes a programme of voluntary independent audits. SAIL policies have been endorsed by Informing Healthcare as well as a number of Caldicott Guardians and information governance officers in the NHS and Local Government. The National Research Ethics Service has confirmed that no ethical review is required for research projects that use SAIL data, because the data have been anonymised and obtained with the permission of the data provider.

## WHO CAN ACCESS THE DATA?

Only authorised researchers can obtain access to the data. Before any data can be accessed, approval must be given by the independent Information Governance Review Panel (IGRP). The IGRP is comprised of representatives from several organisations including; the British Medical Association, Public Health Wales and National Research Ethics Service. The IGRP gives careful consideration to each project to ensure proper and appropriate use of SAIL data. When access has been granted, the researcher can only view data sets which they have been granted permission for, which are relevant to their research. The viewing of data is done through a privacy-protecting safe haven and remote access system referred to as the SAIL Gateway.  This means that research can be carried out in a secure and protected environment. Data is not able to be taken away from the Gateway (for example saving it to a PC or printing), therefore eliminating risk of attack on an individual's privacy. Following access being granted; all access to SAIL is monitored closely and even though all data is anonymous, SAIL has additional safeguards. Following the end of the project, access to the data expires.

## FOR MORE INFORMATION

www.healthinformatics.swansea.ac.uk

healthinformatics@swansea.ac.uk

Data Science Building | Yr Adeilad Gwyddor Data (Llawr Cyntaf)
Swansea University Medical School | Ysgol Feddygaeth Prifysgol Abertawe
Singleton Park | Parc Singleton
SWANSEA SA2 8PP | ABERTAWE SA2 8PP
Wales, United Kingdom | Cymru, Y Deyrnas Unedig

bsi. ANAB
ACCREDITED
MANAGEMENT SYSTEMS
CERTIFICATION BODY

bsi. ISO/IEC
27001
Information Security
Management

## Health Informatics Group

### Swansea University Medical School

Swansea University
Prifysgol Abertawe

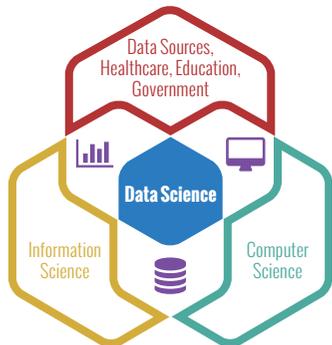# WHERE IS THE HEALTH INFORMATICS GROUP?

The group is based in the Data Science building; a six story purpose-designed building located at Swansea University's Singleton campus. The facility is the hub of health informatics at the University, having a track record of over 20 years on information system research and development, information standards, and public health analyses. The facility houses the Farr Institute at CIPHER, National Centre for Population Health and Wellbeing Research (NCPHWR), the Administrative Data Research Centre Wales (ADRC Wales), and the SAIL Databank.

Our vision is to create a world-class centre in population data science research, training and development through powerful, robust, state-of the-art data linkage and analysis facilities.

# WHAT IS DATA SCIENCE?

Data science is the study of where information comes from, what it represents and how it can be turned into a valuable resource of information. Mining large amounts of structured and unstructured data to identify patterns can improve health treatment and diagnosis , identify risk factors, and increase service efficiency. Big data has the potential to revolutionise our public services.

At the Swansea Data Science Facility certified researchers access anonymised electronic data to carry out specific research with the objective of obtaining a holistic picture across health, social care and housing services . Each  research project must benefit society and the health and wellbeing of the population. The findings of the projects can in turn, support informed decision making  to make changes to policies, health services, and much more.
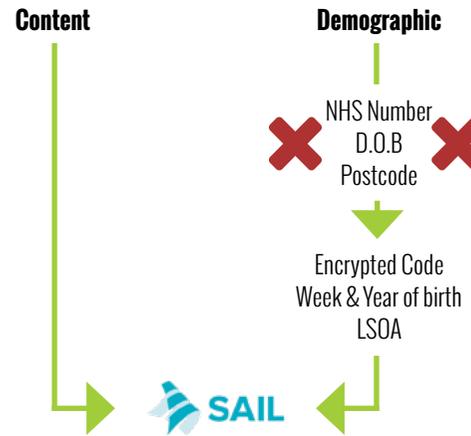


# HOW IS THE DATA COLLECTED AND STORED?

All data which is held at the Data Science facility is collected and stored by SAIL; Secure Anonymised Information Linkage. SAIL is a Wales-wide research resource focused on improving health, well-being and services. Core funding for the data bank is from the Welsh Government's Health and Care Research Wales.



Many organisations collect and hold electronic data about the people for whom they provide services or care. This could be data on hospital admissions, a register of individuals receiving a care package, or a database of blood test results. A provider can then choose to share their data with SAIL, if they do technical support is provided so that everything that arrives at SAIL is in an anonymous form. To anonymise the data, the data provider divides their data into two parts: demographic and content. These go on two different journeys, the content – such as length of hospital stay, medicines prescribed and test results – is sent directly to SAIL. The demographic information is stripped of the commonly-recognised identifying details.



The datasets which are held in SAIL are subject to safeguards and approvals, these can be anonymously linked together to address important research questions. SAIL underpins the work carried out by the Centres housed in the Data Science Facility by safely and securely providing the anonymised data to carry out research.

Some data is not stored by SAIL but researchers may still access this data through safe and secure procedures agreed upon by the data providers such as; Lifelong Learning Wales Record (LLWR) and National Pupil Database (NPD).

# WHAT DOES EACH RESEARCH CENTRE DO?



The Farr Institute at  CIPHER (Centre for the Improvement of Population Health through e-Records Research)

Publically funded by a consortium of organisations led by the Medical Research Council, the Institute is committed to delivering high-quality, cutting-edge research using anonymised data to advance the health and care of patients and the public. The Farr Institute does not own or control data but analyses data to better understand the health of patients and populations. The Institute's research can inform policy in sectors such as health, housing, transport, education,  economic development  and inequality.



National Centre for Population Health and Wellbeing Research (NCPHWR)

An all-Wales research centre with the aim of improving the health and well-being of the population of Wales through applied research. The key themes of the research are;

- children and young people;
- the promotion and maintenance of health through an extended working life.

This contributes to the policy goals of the Welsh Government of giving children a healthy and safe start in life, reducing inequalities and adding more years of high quality life expectancy as well as supporting health and well-being goals for adults.



The Administrative Data Research Centre Wales (ADRC Wales)

Funded by the Economic and Social Research Council (ESRC), the Centre enables access to de-identified administrative data to support economic and social research that has the potential to benefit society.

ADRC facilitates accredited researchers to access de-identified data on themes such as education, health and social mobility data as well as many other long standing records from different government departments. The research findings are used to provide a sound evidence base for policy makers to decide how best to tackle a range of complex social, economic, environmental and health issues.