hic health informatics centre
all the information you need

Alison Bell
Senior Data Analyst
Health Informatics Centre

www.dundee.ac.uk/hic

alison.bell4@nhs.net

University of Dundee

NHS Fife

NHS Tayside

NHS National Services Scotland (Information Services)

- HIC is a research support unit within TASC (TAyside medical Science Centre) University of Dundee, in collaboration with NHS Tayside and NHS Fife.
- Our function is to provide data management support to researchers & others:
  - the collection, preparation and provision of researchable datasets
  - the creation of secure infrastructures for the movement and storage of data and
  - data entry
- Our first priority is to address information governance, data security and confidentiality issues.

TASC Safe Haven @ HIC

health informatics centre
all the information you need

UNIVERSITY OF DUNDEE

NHS Fife

NHS Tayside

NHS National Services Scotland (Information Services)

Release of Research Data

Historically

Anonymised HIC data was released to researchers (emailed, placed on server etc).

Little or no control over where the data went or what it was used for once released

Difficult to do collaborative working

Very difficult to retrieve data back from researchers for archiving at end of study

Our Safe Haven

What it is:

A secure environment isolated from the outside world

Hosting relevant tools/services required for analysis, which can be fine-tuned per study

Its purpose:

To ensure secure containment of data

Ability to audit activity

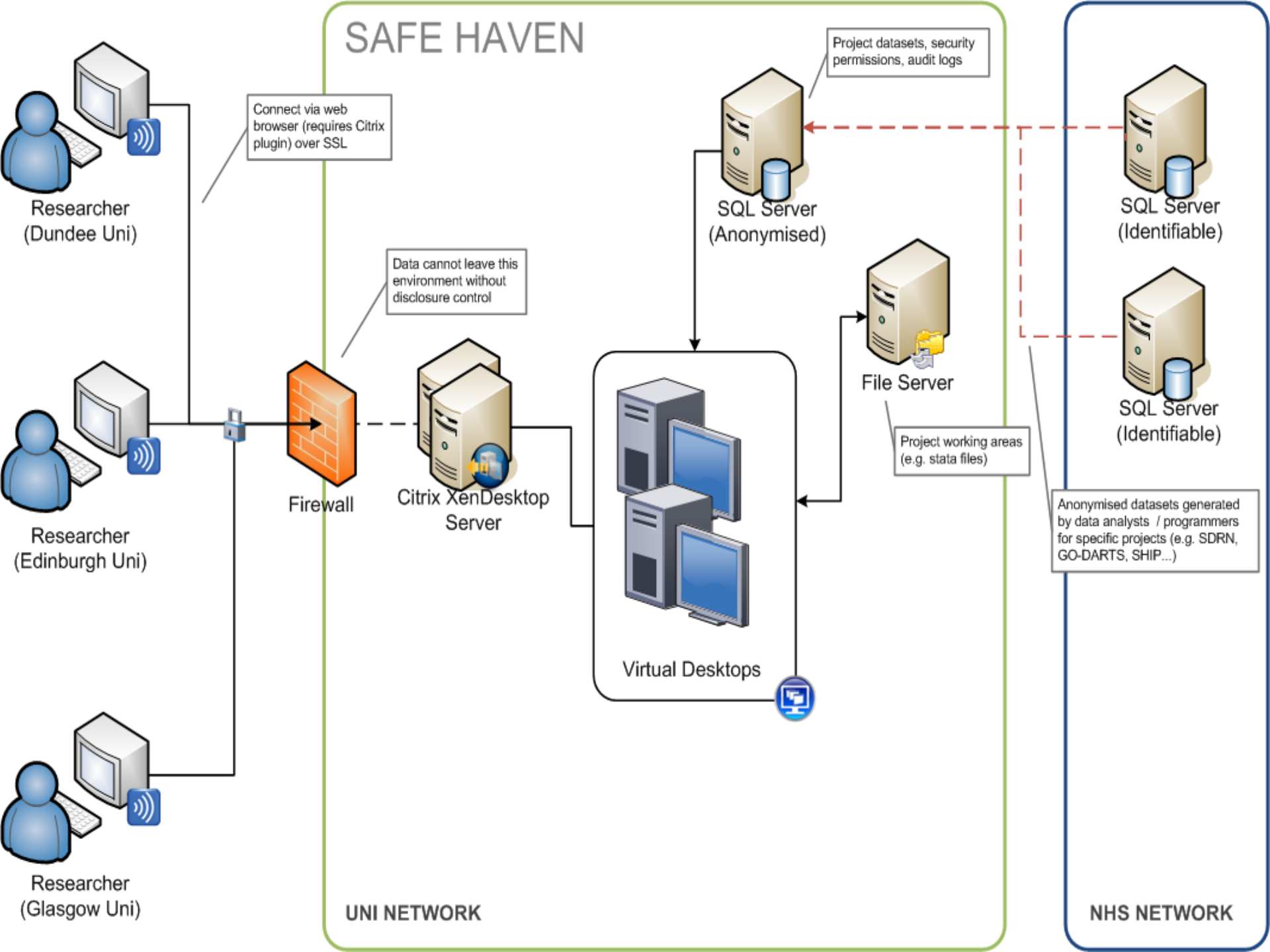
Release of Research Data

Current Process

Access to anonymised patient data is now provided within the TASC safe-haven.

Approved data users log on remotely to a secure server located within HIC.

Data cannot be copied or removed from the secure safe-haven.


SAFE HAVEN
Researcher
(Dundee Uni)
Researcher
(Edinburgh Uni)
Researcher
(Glasgow Uni)
Connect via web browser (requires Citrix plugin) over SSL
Data cannot leave this environment without disclosure control
Firewall
Citrix XenDesktop Server
Virtual Desktops
SQL Server
(Anonymised)
Project datasets, security permissions, audit logs
File Server
Project working areas (e.g. stata files)
SQL Server
(Identifiable)
SQL Server
(Identifiable)
Anonymised datasets generated by data analysts / programmers for specific projects (e.g. SDRN, GO-DARTS, SHIP...)
UNI NETWORK
NHS NETWORK

Output Disclosure Control

No patient-level data is permitted to be removed from the safe haven by the data user.

The data user will move the required file(s) to the output directory within their safe haven personal directory.

An email will automatically be sent to notify the HIC Data Analyst that a file is awaiting release.

Between 9-11am **the next working day** the output file(s) will be reviewed by a HIC Data Analyst and, once verified, emailed to the data user.

Summary of Safe Haven Usage

Pilot

- Ran from April – Nov 2011
- Initially for student projects only
- Gradually rolled out to other users
- 20 research projects in the pilot phase

Production

- From 9th Nov 2011
- 100 users have access
- Over 30 research projects utilising safe haven
- Only exception is legacy projects

health informatics centre
all the information you need

The HIC – OMERO Project

UNIVERSITY OF DUNDEE

NHS Fife

NHS Tayside

NHS National Services Scotland (Information Services)

OMERO

OME (Open Microscopy Environment) develops open-source software and data format standards for the storage and manipulation of biological light microscopy data.

OME is developed as a joint project between research-active teams at labs in Dundee & USA.

The OMERO software is currently used in the overall management of image data (storing, visualising, managing and annotating).

UNIVERSITY OF DUNDEE

NHS Fife

NHS Tayside

NHS National Services Scotland (Information Services)

HIC – OMERO Project

Release of data to researchers as flat files is a historical approach and not scalable to very large datasets.

The HIC-OMERO project objective is to develop a scalable model of managing large population based health datasets, using existing OMERO technology, which will demonstrate good governance and control of data within a safe environment.

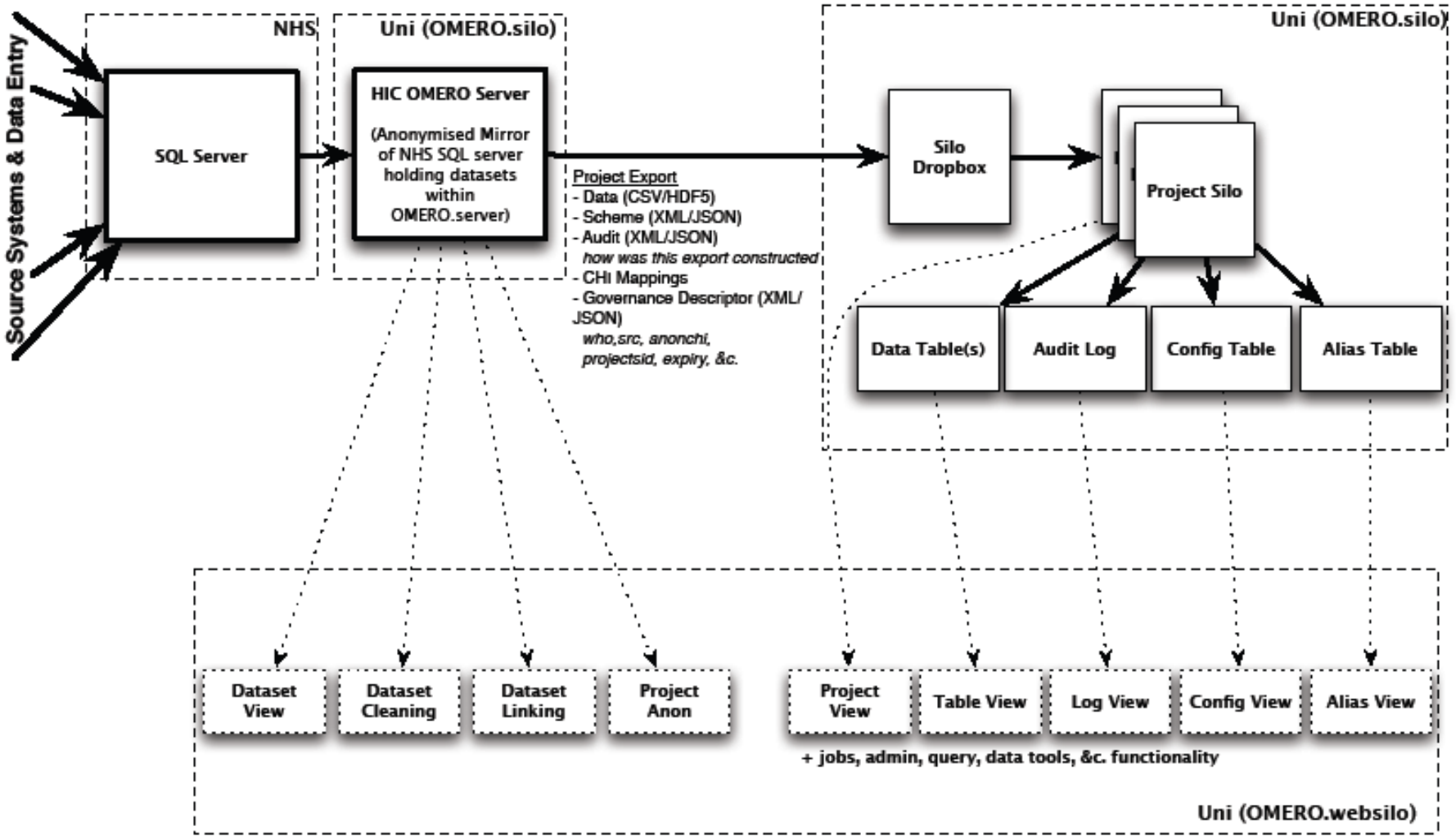
Core functionality to include : import / export of data, audit logs of user actions, basic website to view the logs

UNIVERSITY OF DUNDEE

NHS Fife

NHS Tayside

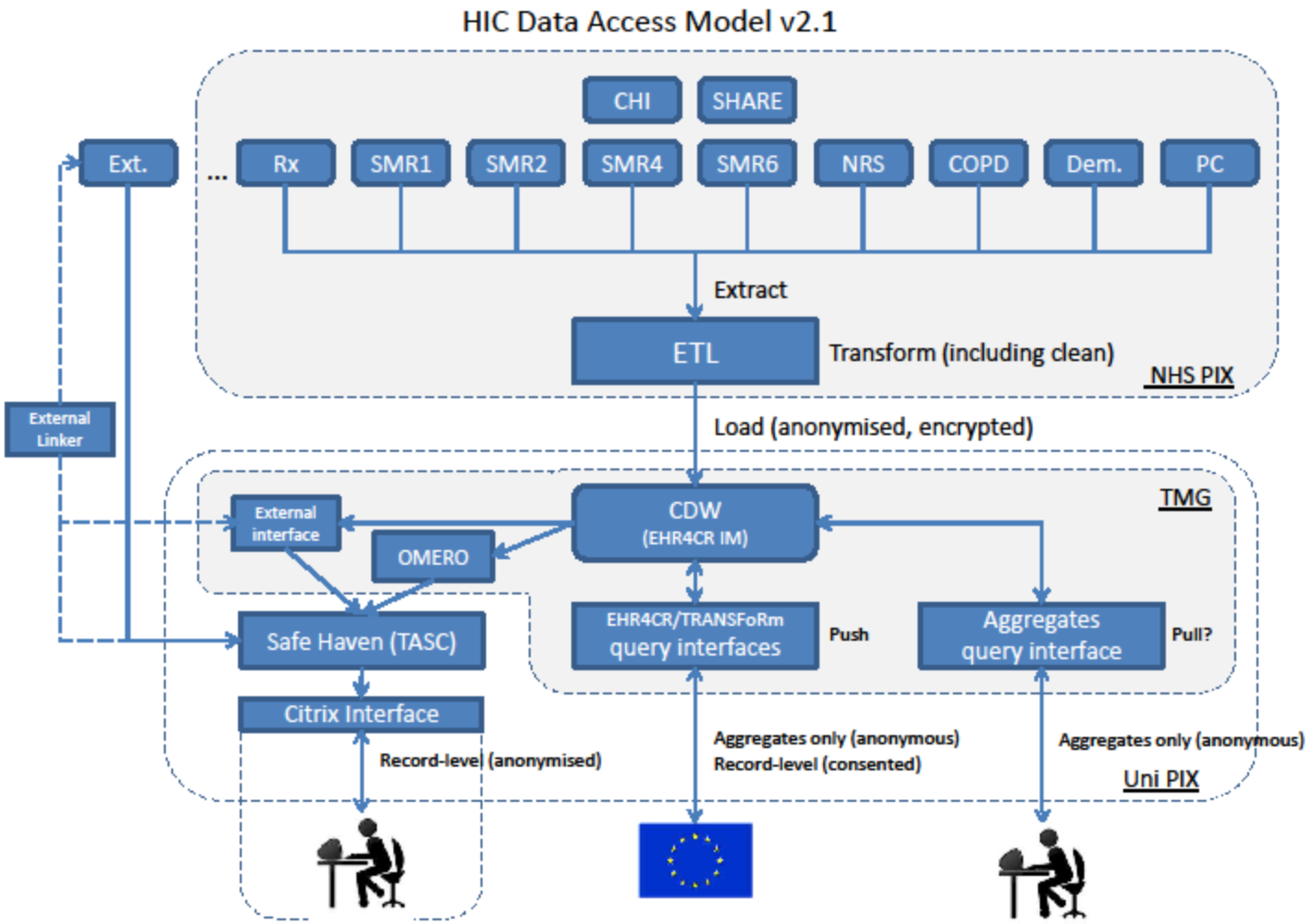
NHS National Services Scotland (Information Services)


Source Systems & Data Entry
NHS
SQL Server
Uni (OMERO.silo)
HIC OMERO Server
(Anonymised Mirror of NHS SQL server holding datasets within OMERO.server)
Project Export
- Data (CSV/HDF5)
- Scheme (XML/JSON)
- Audit (XML/JSON)
how was this export constructed
- CHI Mappings
- Governance Descriptor (XML/JSON)
who,src, anonchi, projectsid, expiry, &c.
Uni (OMERO.silo)
Silo Dropbox
Project Silo
Data Table(s)
Audit Log
Config Table
Alias Table
Dataset View
Dataset Cleaning
Dataset Linking
Project Anon
Project View
Table View
Log View
Config View
Alias View
+ jobs, admin, query, data tools, &c. functionality
Uni (OMERO.websilo)

HIC – OMERO Project

Where we are:

HIC has supplied OMERO with a significant amount of anonymised (scrambled) data to allow them to develop their infrastructure, build their tools and do adequate volume testing

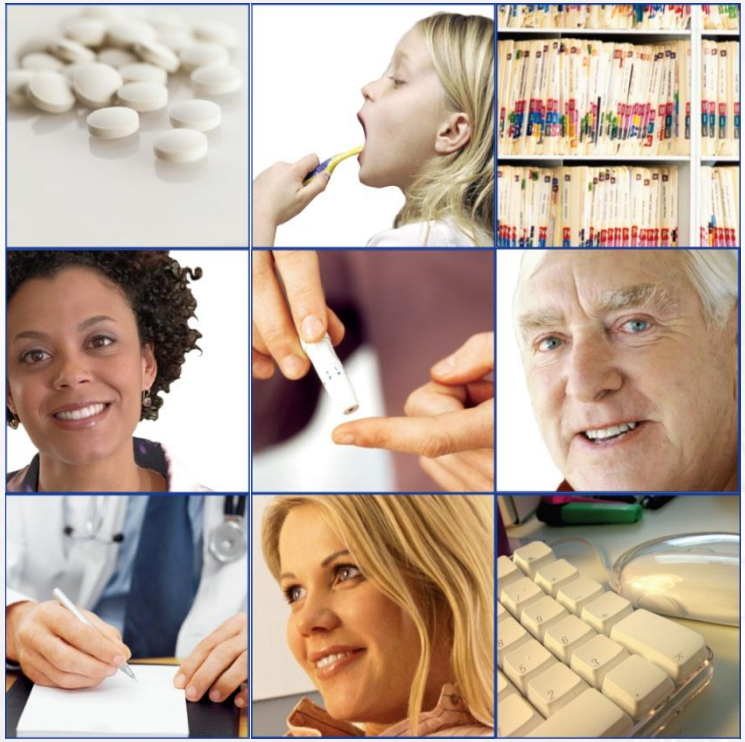


CDW – clinical data warehouse; **ETL** – extract, transform, load; **Ext.** – external data source; **IM** – Information Model

hic health informatics centre

all the information you need

Health in the Information Age



University of Dundee

NHS Fife

NHS Tayside

NHS National Services Scotland (Information Services)