University of Leicester

JISC

NHS National Institute for Health Research

BRISSkit update

Biomedical Research Infrastructure Software Service kit

A vision for cloud-based open source research applications

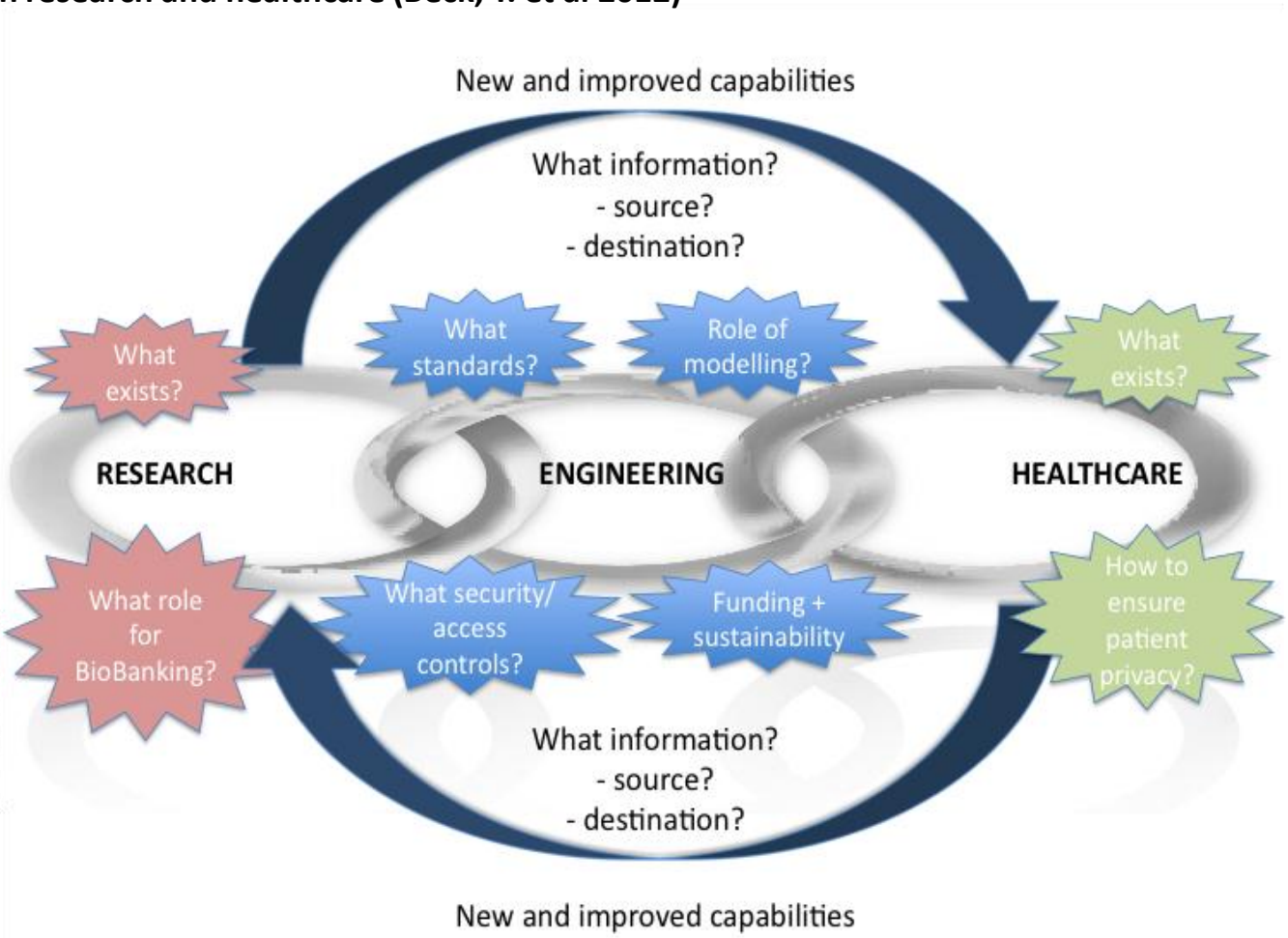
Dr Jonathan Tedds
BRISSkit Project Lead
http://www.brisskit.le.ac.uk

jat26@le.ac.uk
@jtedds

BRISSkit context:

need for data scientists / knowledge engineers!

The I4Health goal of applying knowledge engineering to close the 'ICT gap' between research and healthcare (Beck, T. et al 2012)



BRICCS – *the story so far*

Biomedical Research Informatics Centre for Cardiovascular Science

- End-to-end informatics support for the Leicester Cardiovascular Biomedical Research Unit
- Largely via the use of open source tools
  - Electronic data capture
  - Sample / specimen management
  - Research data management
  - Research project / registration management

BRICCS

Biomedical Research Informatics Centre for Cardiovascular Science

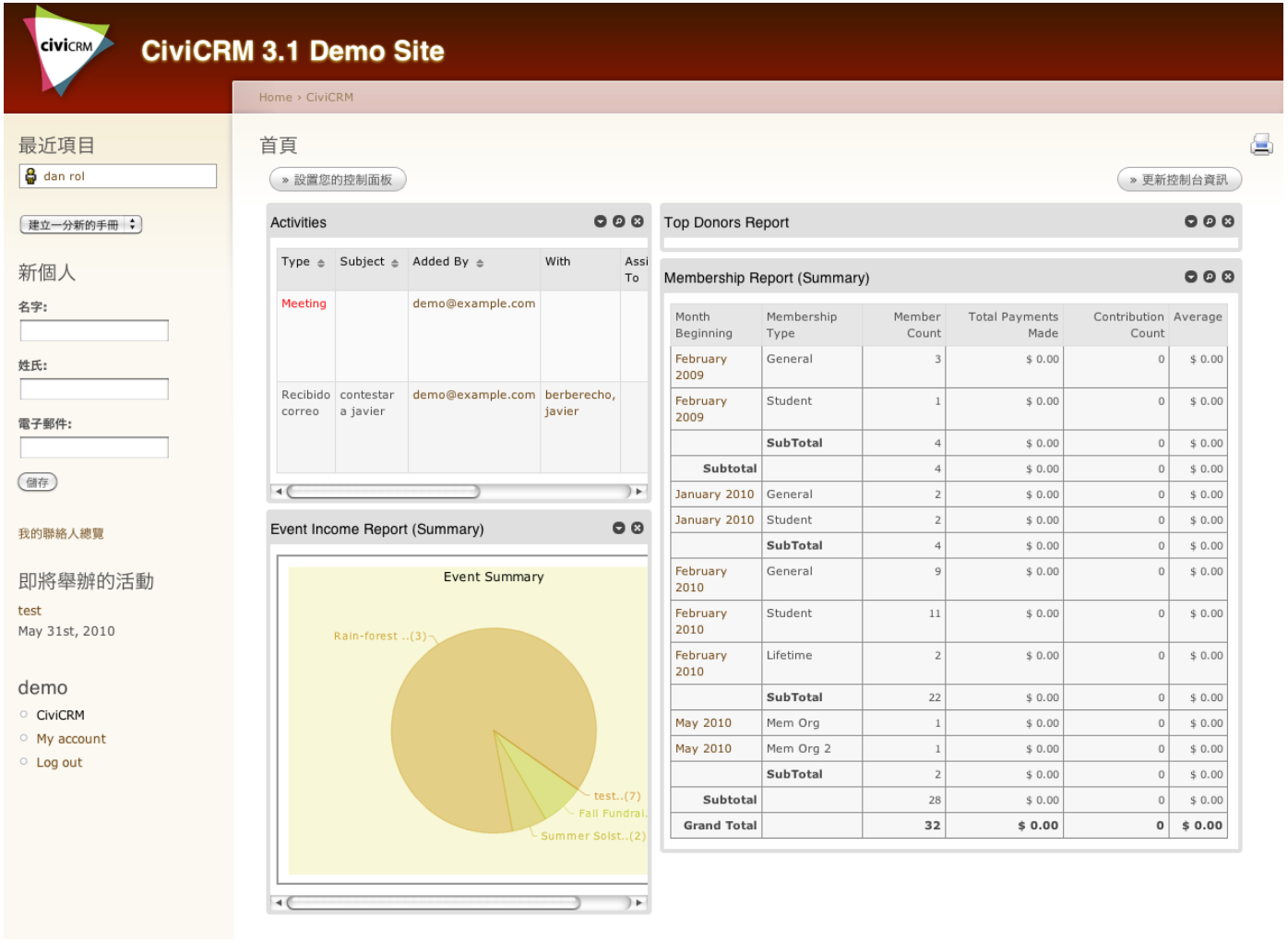
- Collects data only for consented participants.
- Patient identities validated against UHL 'S' number.
- Bridges UHL and university domains.
- Provides sophisticated cohort selection criteria for future research projects (e.g. 'people with existing diagnosis of A, with a family history of B, who have taken C for more than six months, have undergone procedure D and have blood sample E in our lab').

BRISSkit components = web services

CiviCRM

Enables end-to-end contact management for volunteers and research participants, tracking approaches, contact, responses, recruitment, exclusions.

CiviCRM was designed for the 'civic sector' and has an object model that reflects community building and non-profit relationships.

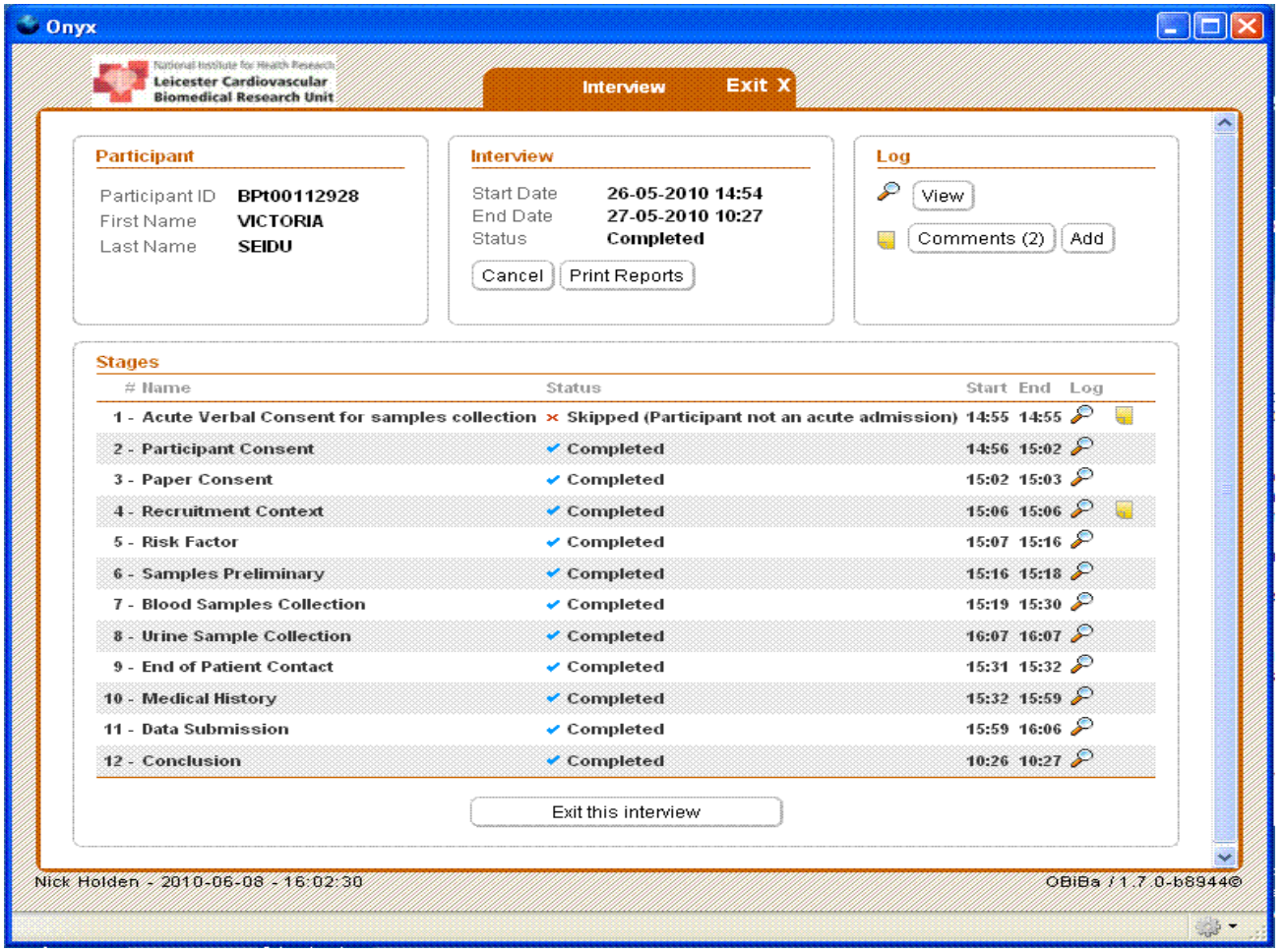


OBiBa Onyx

Records participant consent, questionnaire data and primary specimen IDs.

Web-based, secure data entry by research staff. E.g. used for all patient recruits in LCBRU – mobile computing on wards and outpatient clinic in TMF.

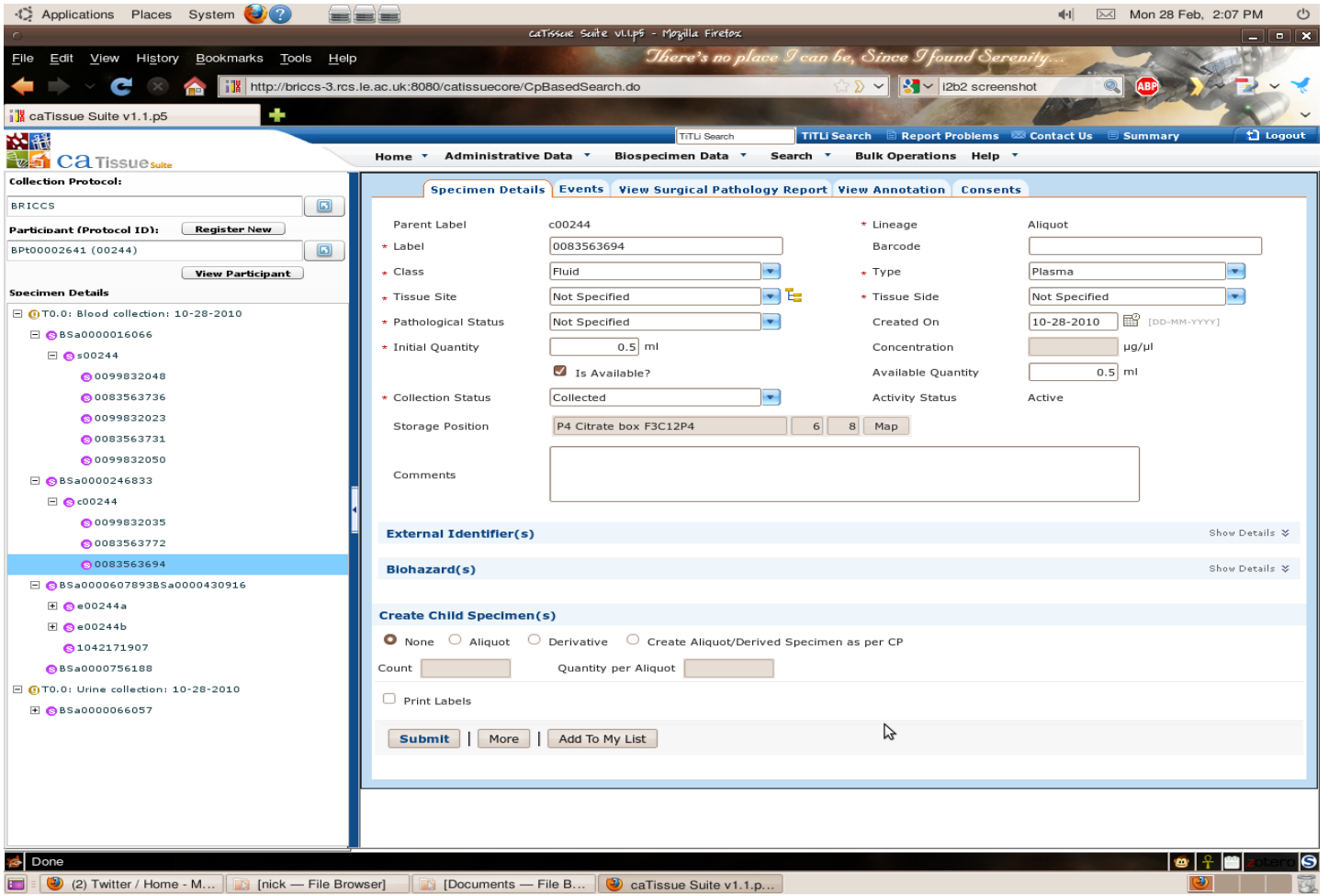
Await significant new release…



caTissue

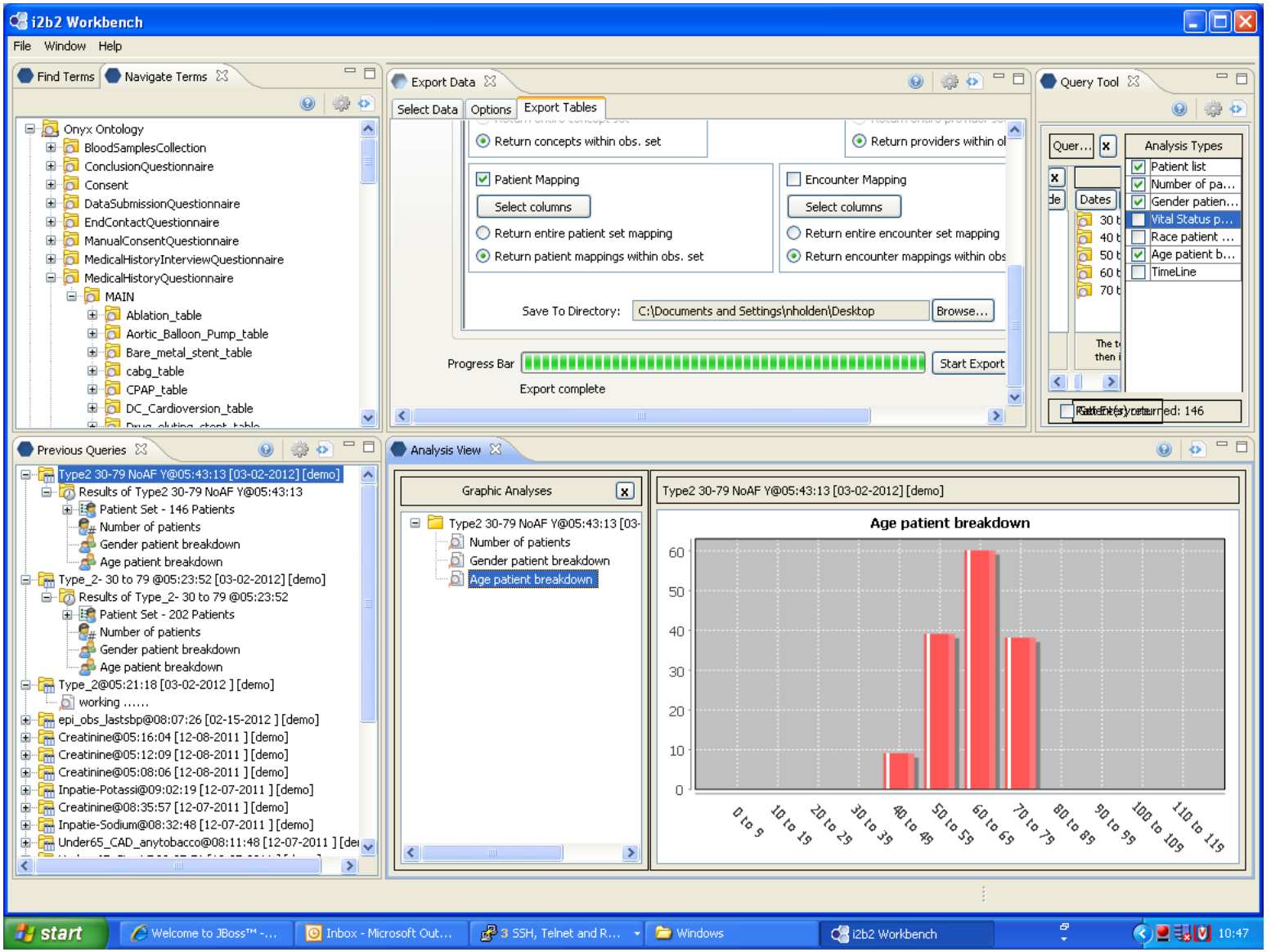
Holds data on primary, derived and aliquot specimen, including linear and 2d barcodes.

Storage inventory, order tracking – currently over 30,000 LCBRU samples stored and recorded.



i2b2

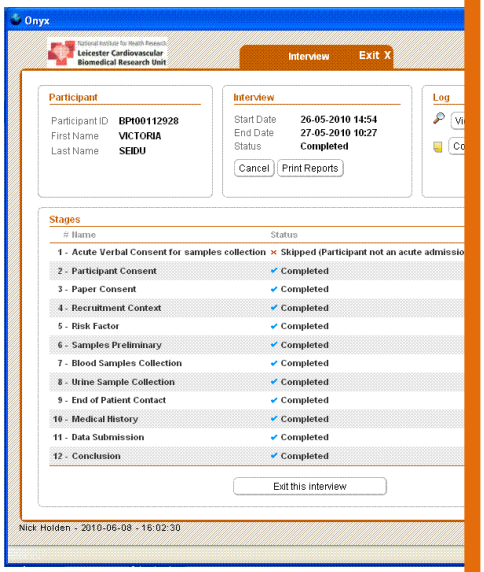
Data from multiple data sources combined into multiple ontologies for flexible and sophisticated searching, cohort discovery and research.



The semantic bridge

OBiBa Ony

Records participa
consent, questio
data and primary
specimen IDs

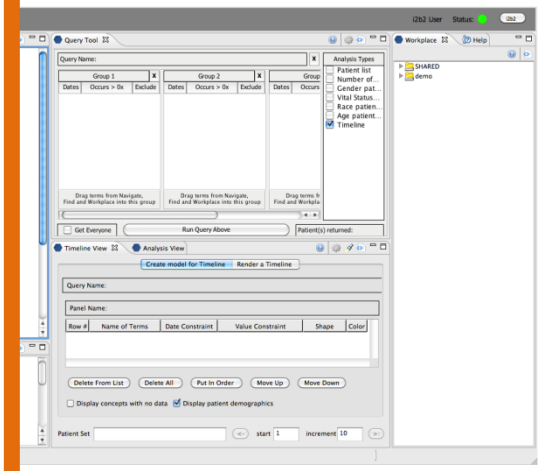


Bio-ontology!

- Classes
  - BrissKitQuestionnaire
    - RiskFactorAlcohol
    - RiskFactorCigars
    - RiskFactorCigs
    - RiskFactorDietExercise
    - RiskFactorFamilyHistory
    - RiskFactorOtherTobacco
    - RiskFactorSocioeconomic
    - RiskFactorTobacco
  - HPO
    - All
      - Phenotypic abnormality
        - Abnormality of the cardiovascular system
        - Abnormality of the hematopoietic system
  - ICD10
    - Diseases of the circulatory system
    - Diseases of the skin and subcutaneous tissue
    - Factors influencing health status and contact with health services
      - Persons encountering health services in other circumstances
  - LOINC
    - LOINCCLASSTYPES
      - Clinical Class
        - Functional status (e.g. Glasgow)
        - Gastrointestinal endoscopy
        - Medical Devices
  - SNOMED-CT
    - Clinical finding
    - Functional finding
    - Hepatorrhaphy
    - Observable entity
    - Organism
    - Social and personal history finding
    - Substance
    - Tobacco smoking behavior - finding
    - Tobacco use and exposure - finding

2

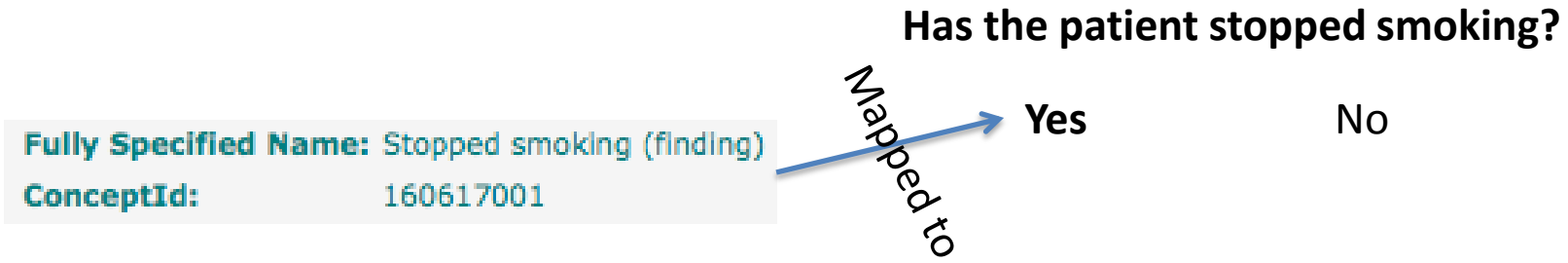
rt selection and
querying



Building Questionnaires with ontologies

- Researcher determines questions and ontologies to use
- While designing questionnaire with Onyx, researcher maps question answers to ontology codes

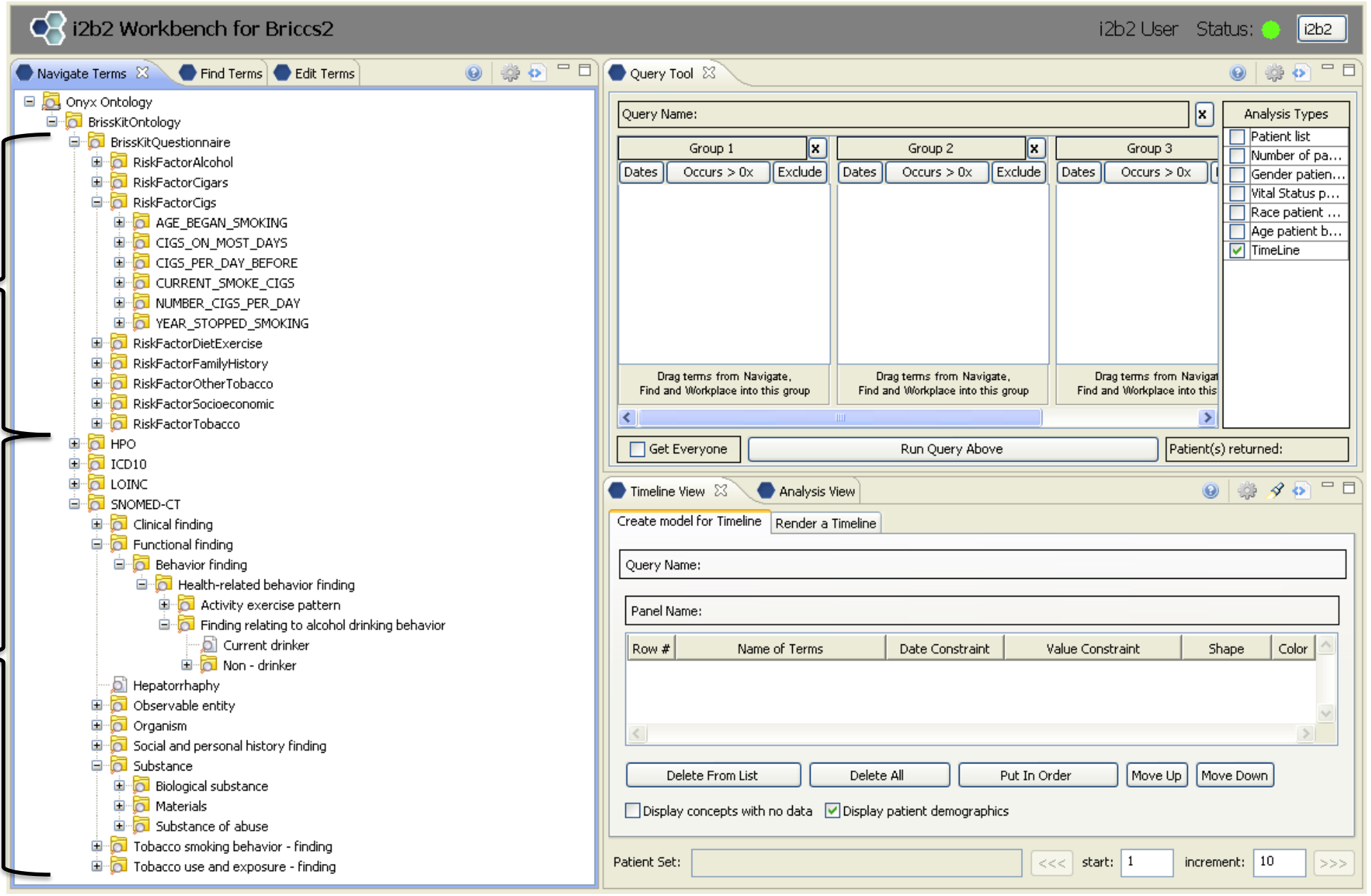
e.g.



BRISSkit ontology tree builder

- Accepts questionnaire design from Onyx
- Allows use any of the 297 ontologies in BioPortal
- Builds an ontology-slim tree:
  - Some ontologies are unmanageable in their full form e.g. SNOMED-CT has 400,000 classes!
  - Only includes terms used in the questionnaire
- Also builds a tree following the questionnaire design
- Outputs in useful formats
  - For use in ontology browsers and of course i2b2!

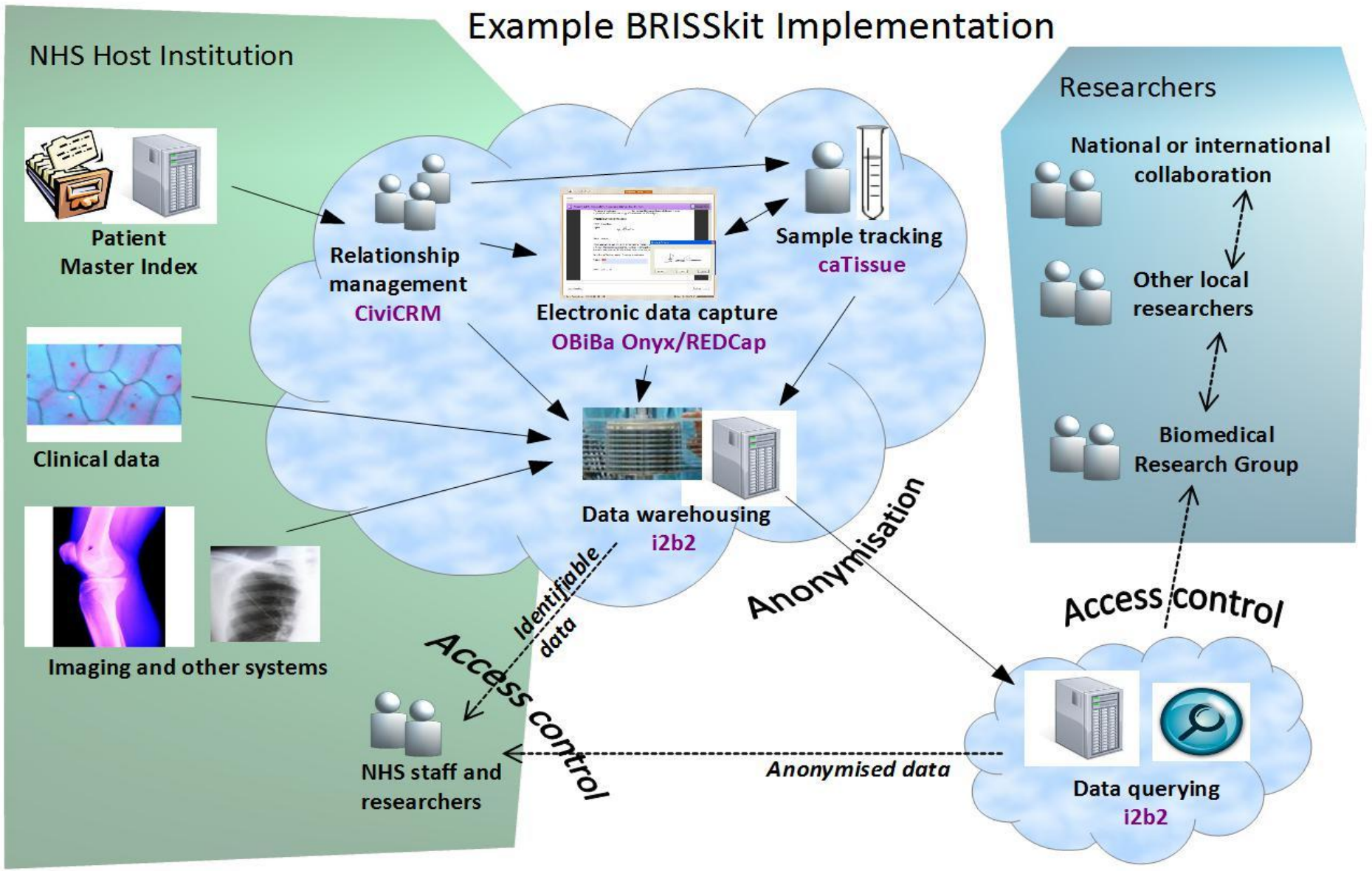
Ontologies displayed in i2b2



Questionnaire ontology

BioPortal ontologies

all components ready for service July 2012

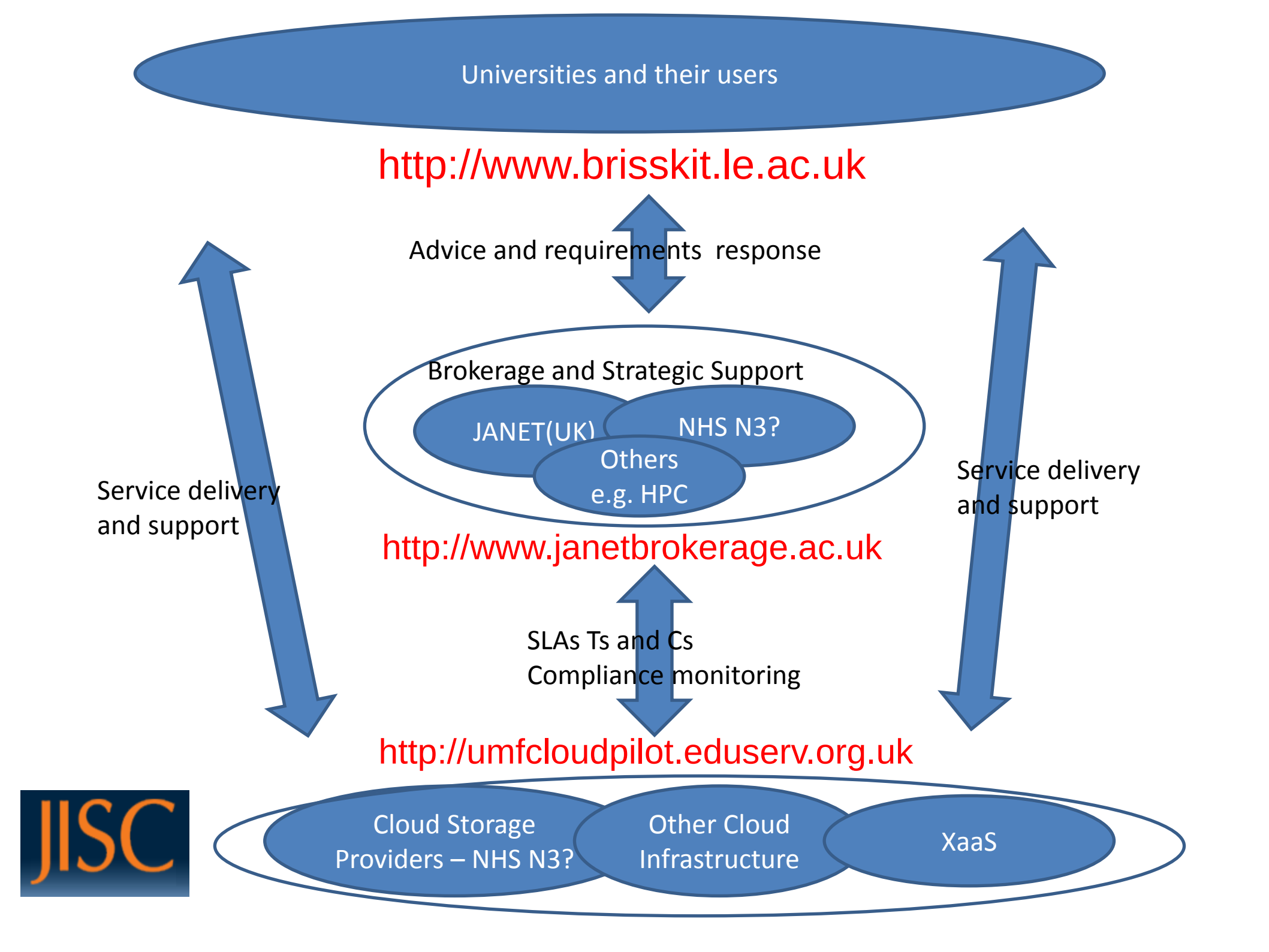


Transition to cloud computing

Virtualisation + Service Computing = Cloud computing

- Infrastructure-as-a-Service: accessing cloud-hosted virtual machines
- Platform-as-a-Service: accessing a cloud-hosted environment
- *Software-as-a-Service: accessing a cloud-hosted application, such as BRISSkit*

- ***NHS-HE JANET connectivity***
  - Web services
  - Rapid flexibility
  - Scalability – location-transparent resource pooling
  - On-demand and self-service
  - Ubiquitous access with high levels of security and compartmentalisation
  - Measured service, with payments geared to usage


Universities and their users
http://www.brisskit.le.ac.uk
Advice and requirements response
Brokerage and Strategic Support
JANET(UK)
NHS N3?
Others
e.g. HPC
Service delivery
and support
Service delivery
and support
http://www.janetbrokerage.ac.uk
SLAs Ts and Cs
Compliance monitoring
http://umfcloudpilot.eduserv.org.uk
Cloud Storage
Providers – NHS N3?
Other Cloud
Infrastructure
XaaS
JISC

BRISSkit Project Phase 1 Status

Project Phase 1 Start Date: 1st July 2011
Project Phase 1 End Date: 30th June 2012

Development is done iteratively under 3 Release Cycles

- ✓Release 1 - 31st July 2011 -31st Jan 2012
  - ✓***Status - Completed***
- ✓Release 2 - 1st Feb 2012 - 31st March 2012
  - ✓***Status - Completed***
- ✓Release 3 - 1st April 2012 - 30th June 2012
  - ✓***Status – Ongoing***

Milestones achieved:

- Implemented cloud infrastructure on Eduserv test platform
- 3 of 4 components operationally ready now, final one by end June 2012
- Onyx-i2b2 integration deployed in genetics
- civiCRM-caTissue-i2b2 integration – under user evaluation
- BRISSkit sustainability plan, community building & partnering support groups

Who is BRISSkit for?

Modular approaches and scalable tools with open source licenses make good investments

- *Individual researchers and associates*
  - enterprise-level tools without the IT overheads
- *Research themes and departments*
  - stand-alone instances of required tools to accelerate research
- *Research units and centres*
  - integrated toolkit with clinical data loading services, or 'jigsaw pieces' to complement existing provision

www.brisskit.le.ac.uk

SEARCH

Email: ***brisskit@le.ac.uk***

BRISSKIT

Biomedical Research Infrastructure Software Service Kit

HOME | ABOUT US | PRODUCTS & SERVICES | CASE STUDIES | EVENTS | BLOGS | WIKI | CONTACT

BRISSkit - Biomedical Research Infrastructure Software Service kit

Overview

BRISSkit will design a national shared service brokered by JANET to host, implement and deploy biomedical research database applications that support the management and integration of tissue samples with clinical data and electronic patient records. We are uniquely positioned to tackle this through our experience in developing the pioneering open source IT infrastructure for the Biomedical Research Informatics Centre for

JONATHAN TEDDS

- My account
- Create content
- Log out

RECENT BLOG POSTS

- CiviCRM
- Advances in Medical Sciences

University of Leicester

JISC

Leicester Cardiovascular Biomedical Research Unit

University Hospitals of Leicester NHS

NHS Trust

Building the BRISSkit community

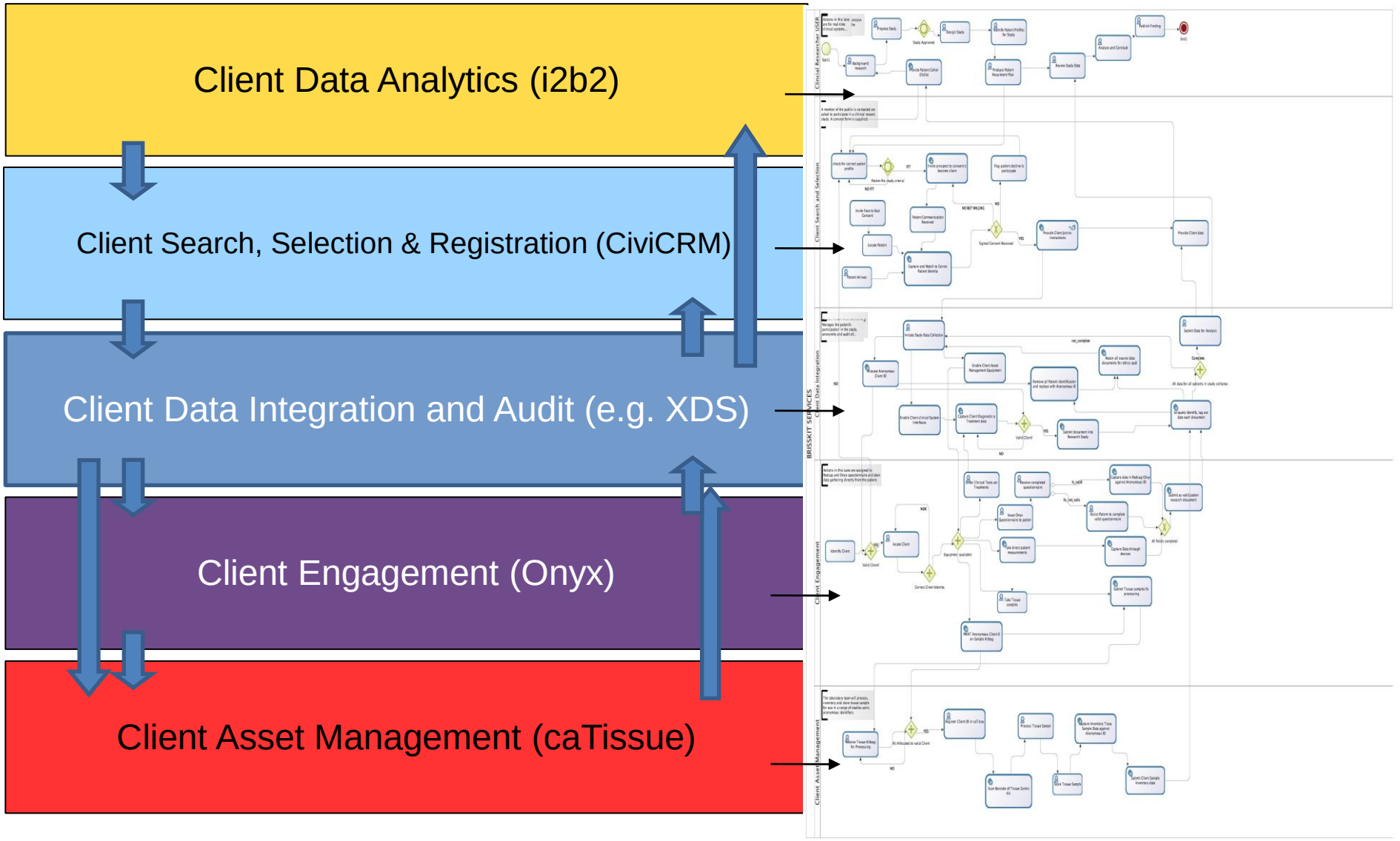
- ***PILOT GROUPS**: new JISC funding for sustainability & uptake, July 2012 - Mar 2013*
  - University Hospitals Leicester Trust (NHS-HE)
    - Cardiovascular BRU & new Respiratory BRU *
    - New Lifestyle BRU (Loughborough-Leicester)
  - Institute for Child Health (UCL) *
  - School of Cancer Studies (Birmingham) *
- *Many **OTHER GROUPS** with firm interest nationally, e.g.*
  - Kings Health Partners BRC, Institute of Cancer Research, Arthritis UK, Cancer Research UK
  - YHMAN Yorks & Humber community cloud (Bradford, Huddersfield, Hull, Leeds, Leeds Met, Sheffield, York) via JANET brokerage?
  - 50+ attendees to Jan 2012 workshop in Leicester (Institute for Cancer Research, Oxford, Birmingham, Sheffield, Loughborough, Manchester, UCL, Leeds & Northumbria)
  - ***Community workshop in Leicester + Hack Day - early Sep 2012***

Sustainability phase outcomes

- Harden service
  - VM management & service management
- Activate new users
  - Integrating distributed data sources
  - Thoroughly test demand through independent evaluation
- Customisation process commoditised
- Investable business plan

→ Transitioning to longer term service & sustainability

BRISSkit - Service Options



Current and Potential Technical Partners

Krishagni

Jefferson

Washington University School of Medicine in St.Louis

Recombinant
HEALTHCARE DATA WAREHOUSING

i2b2
A National Center for Biomedical Computing
Informatics for Integrating Biology & the Bedside

CIVICRM

ngssecure
an ncc group company

OBiBa

National Cancer Institute
caBIG® Cancer Biomedical Informatics Grid®

PERSISTENT

University of Leicester

Total Cost of Ownership of Open Source Software
LSE report for the UK Cabinet Office Sept 2011

Benefits of Open Source approach

- Reliability, stability
- Freedom to alter or adapt the software – no 'tie in'
- Lower costs – no upgrade licenses
- Interoperability – lower barriers to integration
- Capabilities for data sharing enhanced

Phase	Cost Savings	BRISSkit Application
Search	Cost of up-front evaluation study Cost of up-front proof of concept	
Acquisition	Initial Cost of Software Cost of Customisation	Tissue Management Electronic Data Capture Translational Data Management Central Participant Registration
Integration	Cost of Integration	Tissue Management Electronic Data Capture Translational Data Management Central Participant Registration
Use	Annual software costs Software scaling (user or transactions) Data Capture (re-use)	Tissue Management Electronic Data Capture Translational Data Management Central Participant Registration
Retire	Exit costs (hardware and software) Exit costs (changeover, re-training) Exit costs (data migration)	

BRISSkit – Community Sustainability

- **Partnership (indirect service) Model**
  - **BRISSkit partnership agreement**
  - **Partners deliver all direct client services - training, integration, first line support**
  - **Private clouds running BRISSkit software encouraged**
  - **% Revenues sustain central software support, maintenance and training**
- **Direct Service Model**
  - **BRISSkit direct service agreement with clients**
  - **Single central cloud service with preferred partners**
  - **All revenue sustains central support, maintenance and training**
- **Community Model**
  - **Client pay/donate to BRISSkit Community membership fund**
  - **Direct access to central cloud service with preferred partners but no service levels**
  - **Funds sustain central support, maintenance and training**

www.brisskit.le.ac.uk

SEARCH

Email: ***brisskit@le.ac.uk***

BRISSKIT

Biomedical Research Infrastructure Software Service Kit

HOME | ABOUT US | PRODUCTS & SERVICES | CASE STUDIES | EVENTS | BLOGS | WIKI | CONTACT

BRISSkit - Biomedical Research Infrastructure Software Service kit

Overview

BRISSkit will design a national shared service brokered by JANET to host, implement and deploy biomedical research database applications that support the management and integration of tissue samples with clinical data and electronic patient records. We are uniquely positioned to tackle this through our experience in developing the pioneering open source IT infrastructure for the Biomedical Research Informatics Centre for

JONATHAN TEDDS

- My account
- Create content
- Log out

RECENT BLOG POSTS

- CiviCRM
- Advances in Medical Sciences

University of Leicester

JISC

University Hospitals of Leicester NHS

NHS Trust