

Preparing a Research Data Management response: policy, infrastructure and practice

April 2015

Overview

- Developing a RDM response
- Policy principles
- Building infrastructure
- Frameworks and practice
- Challenges and risks

Developing a RDM response

Action	Timeframe
Raise awareness	June 2012 - Dec 2013
Identify best practice and understand capacity of infrastructure in the UK	Jan - June 2013
Audit the data management landscape	June 2013
Analyse costs and benefits of data management infrastructure	March 2013 - Dec 2013
Develop best practice policies and guidelines	Sept - Dec 2013
Establish governance structures	Jan - Feb 2014
Disseminate best practice and importance of RDM	Jan - Dec 2014
Establish infrastructure to enable the preservation, security and public access of research data	March - Dec 2014
Appoint staff	April 2014 - May 2014
Provide skills training in data management	May 2014 - May 2015
Embed RDM training and good practice guidelines	Jan 2015 - May 2015
Offer support help desk on data planning and management	Sept 2014 - May 2015

Key challenges

- Complex and varying funder requirements
- No-one model of best practice
- No common processes, systems or data capture for RDM across the University
- Balancing (often) competing objectives - e.g. sharing vs confidentiality
- Concern / 'fear' over infrastructural solution and cost
- Policy and best practice vs communication and raising awareness
- Different University priorities, not least REF 2014

Understanding the research data management landscape

- Pockets of good practice but generally, lack of
 - Awareness
 - Data sharing
 - Planning for financial sustainability
- Range of datasets contained in a variety of holdings:
 - School servers (with back up)
 - External hard drives
 - PCs
 - Lab note books
 - Laboratories
 - Government offices
 - Cloud services
 - Subject repositories

Understanding the research data management landscape

- Difficult to audit the extent of data holdings
 - MBs - 10s of TB
 - 'Don't know'
 - 'A lot'
 - 'Many files'
 - 'What a silly question'
- Clear signage for needs:
 - Advice and training (funder requirements and legal issues)
 - Research data cataloguer per project / research group
 - Storage /space with back up

Networking and sharing experience

- Learning by doing.....by others
- Digital Curation Centre
- RDM networks and seminars
- Pure User Group
- Individual universities

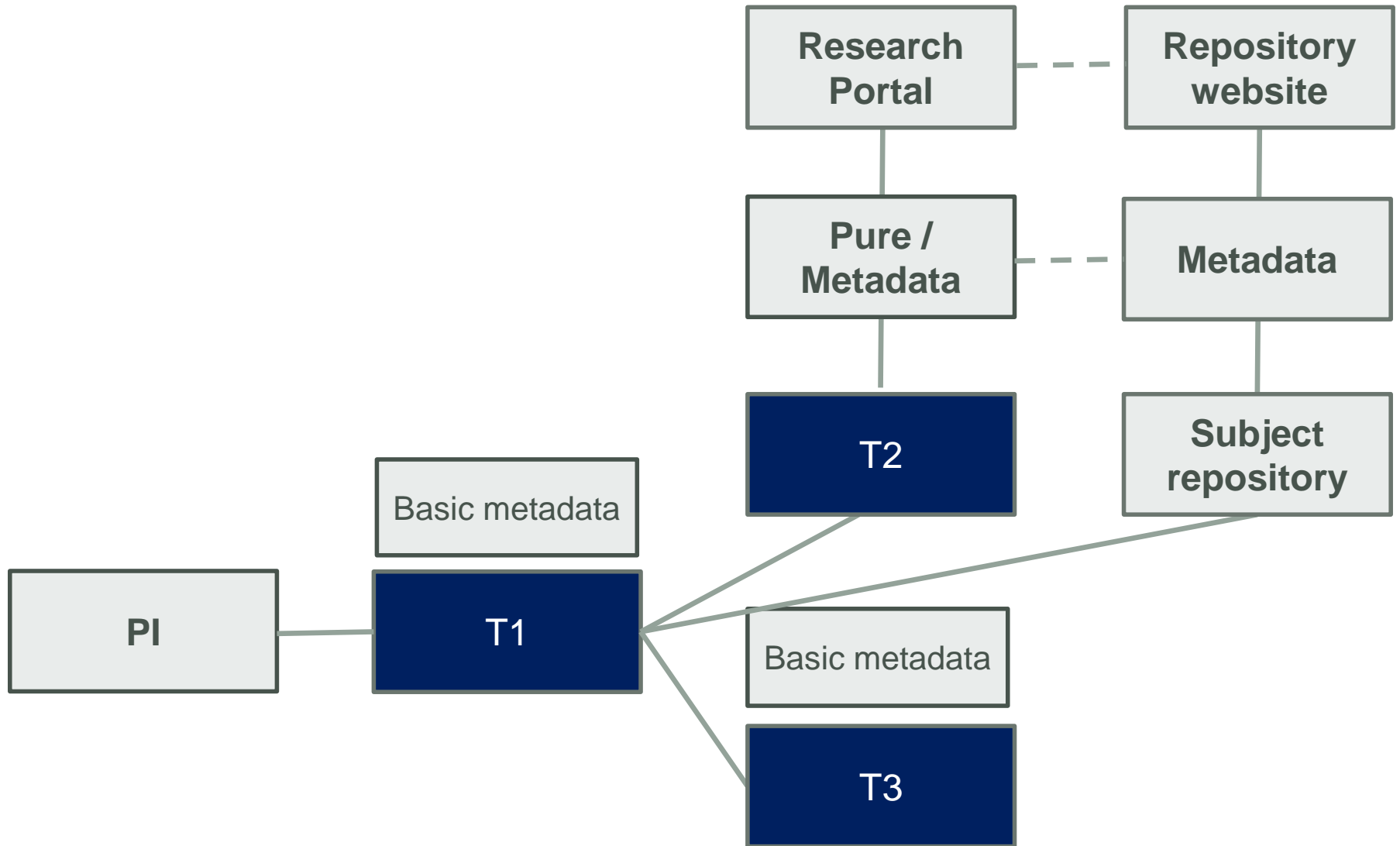
Policy principles

- Restating expected data management practices
 - Institutional
 - Researcher (PI) level
- Sharing data
 - compliance with funder requirements or
 - the University..... 'encourages all researchers to make their publicly funded research accessible and freely available'


Building RDM infrastructure

- Utilising Pure developments
- Future storage requirements unclear
- External storage solutions impressive but costly
- Short / medium term internal 2PB storage solution:
 - Tier 1: active project data
 - Tier 2: data to be made publically available
 - Tier 3: archive data (longer recall period and offsite copy to mitigate against disk failure)
- Researcher access to 2TB (> PIs of larger data driven subjects)

RDM framework



Pure research dataset

 Dataset

Data availability

Publisher *

DOI

-or- If you do not have a DOI already one will be issued during validation where appropriate

Electronic data

Physical data

Links

Date made available

Year * Month Day

Access to the dataset

Access options

▼

QUB Research Portal



[Study at Queen's](#) | [Research](#) | [Alumni and Supporters](#) | [Life in Belfast](#) | [About Queen's](#)

[Research Portal](#) > [Data Sets](#)

Data Sets

[Advanced search »](#)

1 - 3 out of 3

Sort by: Title

Experimental data on time and interventions in children's causal structure learning



[McCormack, T.](#), Creator, UK Data Service, Sep 2014, [10.5255/UKDA-SN-851417](#)
Dataset

Genetics of Nephropathy - an International Effort (GENIE) GWAS of Diabetic Nephropathy in the UK GoKinD and All-Ireland Cohorts



[Maxwell, A.](#), Creator, The National Center for Biotechnology Information, Nov 2011
Dataset

Northern Ireland Life and Times Survey, 2012: Lesbian, Gay, Bisexual and Transgender Issues Teaching Dataset



[Devine, P.](#), Creator, UK Data Service, Aug 2014, [10.5255/UKDA-SN-7546-1](#)
Dataset

Related Content

[Research Portal](#)

[Publications](#)

[Researchers](#)

[Faculties, Schools and Clusters](#)

[Projects](#)

[Impact](#)

Data Sets

[Research Portal Policy](#)

Find Research / Researcher

QUB Research Portal

Genetics of Nephropathy - an International Effort (GENIE) GWAS of Diabetic Nephropathy in the UK GoKinD and All-Ireland Cohorts

Dataset

Alexander Maxwell (Creator)
Amy Jayne McKnight (Creator)

School of Medicine, Dentistry and Biomedical Sciences
Centre for Public Health



[View graph of relations](#)

Description

Diabetic kidney disease, or diabetic nephropathy (DN), is one of the leading causes of end-stage renal disease in the United States and worldwide. DN is a common complication of long-standing type 1 and type 2 diabetes. Type 1 diabetes can have its onset in childhood and affected individuals often develop end-stage renal disease in early adulthood, leading to further loss of quality of life. The genetic basis of the disease is not well understood. The GENIE (GEnetics of Nephropathy an International Effort) consortium was initiated to perform the most comprehensive and well powered DN susceptibility genome wide association study (GWAS) analysis to date, using the largest collection of individuals with type 1 diabetes with and without kidney disease across four study cohorts. The UK-ROI samples were initially GWAS genotyped as part of this project; Dr AJ McKnight was primary analyst. UK-ROI Sample Description The UK-ROI collection consists of samples derived from the Republic of Ireland (Dr. Catherine Godson, PI, at University College, Dublin, Ireland) and the United Kingdom (Warren 3, Diabetes UK, Genetics Of Kidneys IN Diabetes (GOKIND) UK, UK GoKinD, Dr. Alexander P. Maxwell, PI, at Queen's University of Belfast, UK). All study subjects met the inclusion criteria: white individuals with T1D, diagnosed before 31 years of age, whose parents and grandparents were born in the British Isles. Study Type: Case-Control Number of study subjects that have individual level data available through Authorized Access: 1801

Date made available **Nov 2011**
Publisher **The National Center for Biotechnology Information**

Links

http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000389.v1.p1

Projects

US Ireland Research & Development Partnership - Diabetes
Project: Research

Publications

SORBS1 gene, a new candidate for diabetic nephropathy: results from a multi-stage genome-wide association study in patients with type 1 diabetes

Research output: Contribution to journal › Article

Related Content

[Research Portal](#)

[Publications](#)

[Researchers](#)

[Faculties, Schools and Clusters](#)

[Projects](#)

[Impact](#)

[Data Sets](#)

[Research Portal Policy](#)

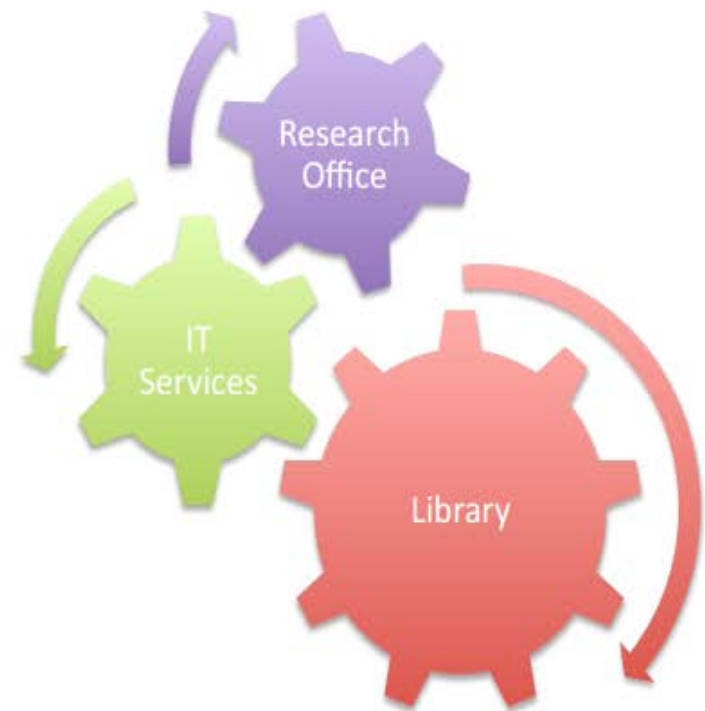
Find Research / Researcher



Challenges and risks

- Not a perfect, comprehensive & everlasting solution
- Commercial products vs in-house solution
- An incremental approach - evidence gathering on future data requirements
 - Extensive storage / infrastructure
 - FTE
 - A broader organisational approach

Organisational collaboration



Future RDM framework

