ReShare &ReCollect
EPrints data repository solutions developed at the UK Data Archive

Veerle Van den Eynden, Alexis Wolton, Tom Ensom, Louise Corti

UK Data Archive

University of Essex

OR2014 conference

Helsinki, 13 June 2014

UK Data Service

Holding & sharing research data collections

- ReCollect = EPrints plugin to install (institutional) data repository
  - pilot for University of Essex
  - > [Research Data @Essex](#) project (2011-2013), Jisc funding
  - multi-disciplinary data collections
  - http://bazaar.eprints.org/280/

- ReShare = subject repository
  - UK Data Archive self-deposit repository for short-term management curation
  - wider social sciences & humanities (ESRC data policy)
  - developed 2013-2014, ESRC funded
  - http://reshare.ukdataservice.ac.uk/

Why EPrints
repository software

- ReCollect:
  - parallel development of University of Essex EPrints publications repository
  - EPrints developer at university

- ReShare:
  - further developed from ReCollect
  - replace older Fedora repository (no longer developer support)
  - good user and admin interface
  - good admin functionality for reviews

ReCollect

- Developed and tested in collaboration with 4 research-data-rich departments, for cross-disciplinary suitability:
    - Biological Sciences
    - Essex Business School
    - Language and Linguistics
    - Computing and Electronic Systems
- Expanded metadata profile for describing research data collections
- Adapted UI and deposit flow

ReCollect metadata profile

- Standards-based approach, producing:
    - [metadata mapping](#) between relevant schemas
    - [metadata profile](#) built from mapped elements

- Standards mapped:
    - **DataCite 3.0** – minimal mandatory metadata (discovery)
    - **UK - Gemini 2.2 –** UK standard for EU INSPIRE Directive implementation (geospatial / environmental data)
    - **DDI 2.1** – social sciences standard, used at UK Data Archive
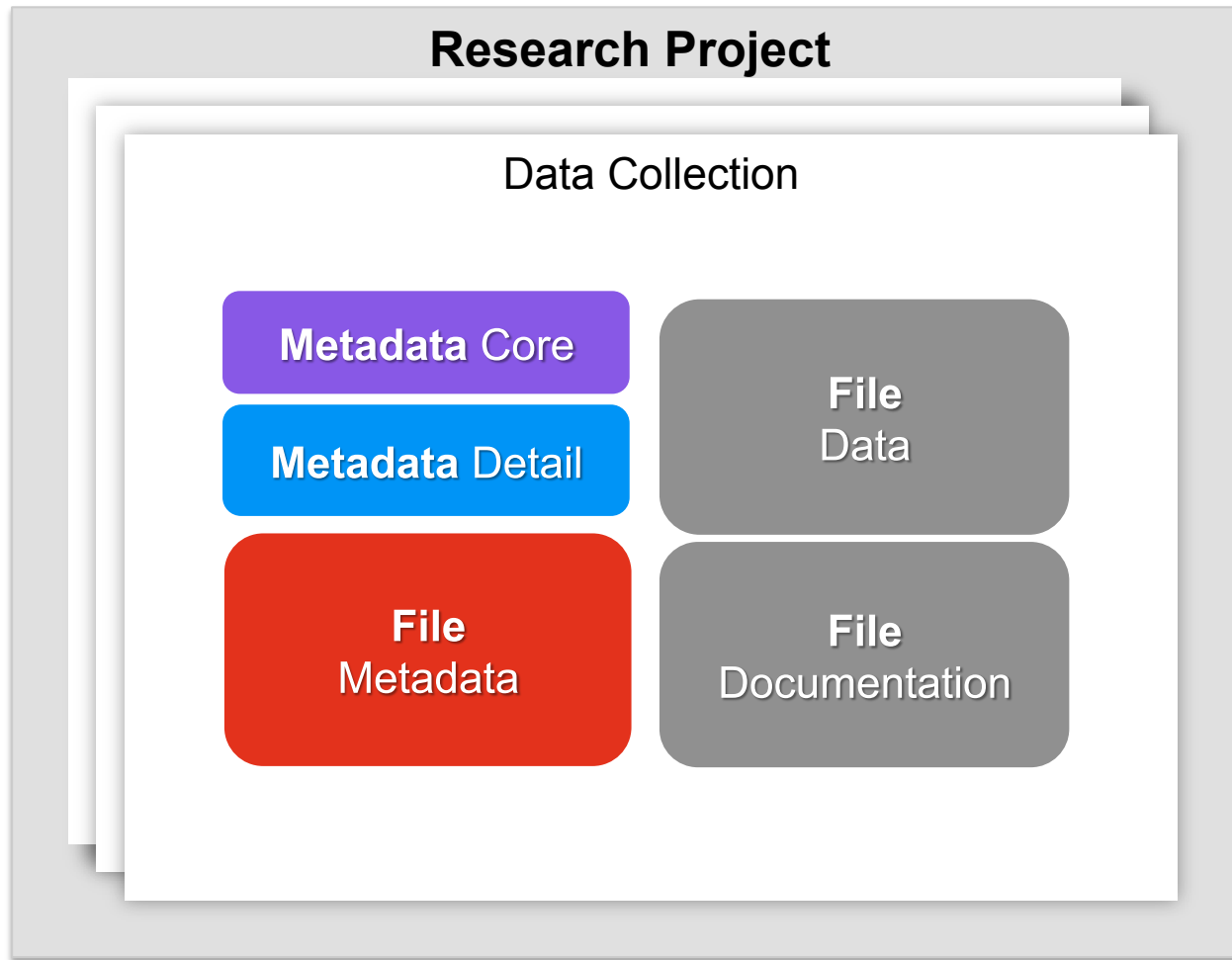
ReCollect metadata

- **Core elements (9) (9M 4A)**
  - citation, identification
- **Detail elements (27) (8M 19O 2A)**
  - descriptive information
- **File-level elements (7) (6M 1O 2A)**

(**M**andatory; **O**ptional; **A**utomatic)

Object model

ReCollect metadata profile

Core metadata	
Title	M
Unique ID	A
Description	M
Keywords	M
Subject	M
Divisions (depts)	M
Depositor	A
Deposit date	A
Version	A

Detail metadata			
Alternative title	O	Provenance	O
Creators	M	Resource language	M
Corporate creators	O	Additional information	O
Data type	M	Related resources	O
Contributors	O	Original data publisher	O
Funder	O	Publication status	M
Grant number	O	Original data publication URL	O
Parent project	O	Original data publication date	O
Collection dates	O	Copyright	M
Temporal extent	O	Publication date	A
Geographic coverage	O	Publisher	A
Geographic bounding box	O	Contact	M
Data collection method	O	Administrative note	O
Legal / ethical issues	O		

File-level metadata	
MIME type	A
Description	M
Content	M
Embargo date	O
Metadata version	A
Access	M
Licence	M

UK Data Service

ReCollect user interface

- Adapted for deposit and display of complex collections with multiple files & metadata
- Summary page
  - display 2nd set of extended metadata at eprint level
    - javascript box show/hide
  - data box + documents box
  - objects (files) sorted ~ content type
    - data
    - documentation
    - metadata
    - readme
    - full archive
- Help text & guidance

UK Data Service

University of Essex

Home   **About**   **Browse**   **Help**

Logged in as RD Admin | Manage deposits | New Deposit | Manage records | Profile | Saved searches | Review | Admin | Logout

Search

Sample Data: Bioimaging - Surface Receptors

Laissue, Philippe (2012) *Sample Data: Bioimaging - Surface Receptors.* [Data Collection]. Research Data Essex, University of Essex.
http://researchdata.essex.ac.uk/id/eprint/77 (Unpublished)

Collection description

PLEASE NOTE THAT THIS IS A SAMPLE DATA COLLECTION MADE VISIBLE FOR DEMONSTRATION PURPOSES ONLY. Confocal microscopy derived images of cells, used to demonstrate the development of a novel method for the quantified, superresolved, three-dimensional colocalisation of histocompatibility class I receptors.

Keywords:	Colocalisation, quantification, confocal microscopy, surface receptors, bioimage informatics
Subjects:	Biological Sciences > Biology > Cell Biology
Divisions:	Faculty of Science and Engineering > Biological Sciences, School of
Depositing User:	RD Admin
Date Deposited:	22 Aug 2012 10:58
Revision:	78
URI:	http://researchdata.essex.ac.uk/id/eprint/77

Available Files

Data

➕ **data_tiff.zip**

➕ **data_nd2.zip**

Documentation

➕ **doc_thumbnails.zip**

➕ **Additional details**

Essex Research Data is powered by *EPrints 3* which is developed by the School of Electronics and Computer Science at the University of Southampton. More information and software credits.

eprints

Logged in as RD Admin | Manage deposits | New Deposit | Manage records | Profile | Saved searches | Review | Admin | Logout

Search

Sample Data: Bioimaging - Surface Receptors

Laissue, Philippe (2012) *Sample Data: Bioimaging - Surface Receptors*. [Data Collection]. Research Data Essex, University of Essex.
http://researchdata.essex.ac.uk/id/eprint/77 (Unpublished)

Collection description

PLEASE NOTE THAT THIS IS A SAMPLE DATA COLLECTION MADE VISIBLE FOR DEMONSTRATION PURPOSES ONLY. Confocal microscopy derived images of cells, used to demonstrate the development of a novel method for the quantified, superresolved, three-dimensional colocalisation of histocompatibility class I receptors.

Keywords:	Colocalisation, quantification, confocal microscopy, surface receptors, bioimage informatics
Subjects:	Biological Sciences > Biology > Cell Biology
Divisions:	Faculty of Science and Engineering > Biological Sciences, School of
Depositing User:	RD Admin
Date Deposited:	22 Aug 2012 10:58
Revision:	78
URI:	http://researchdata.essex.ac.uk/id/eprint/77

➕ Additional details

Available Files

Data

➖ **data_tiff.zip**

Visible to:	Repository staff only
Content:	Data
Document description:	Images (TIFF format)
Metadata Revision:	5
Mime-Type:	application/zip
File size:	2175Mb

➕ **data_nd2.zip**

Documentation

➕ **doc_thumbnails.zip**

eprints

Search

Sample Data: Bioimaging - Surface Receptors

Laissue, Philippe (2012) *Sample Data: Bioimaging - Surface Receptors*. [Data Collection]. Research Data Essex, University of Essex. http://researchdata.essex.ac.uk/id/eprint/77 (Unpublished)

Collection description

PLEASE NOTE THAT THIS IS A SAMPLE DATA COLLECTION MADE VISIBLE FOR DEMONSTRATION PURPOSES ONLY. Confocal microscopy derived images of cells, used to demonstrate the development of a novel method for the quantified, superresolved, three-dimensional colocalisation of histocompatibility class I receptors.

Keywords:	Colocalisation, quantification, confocal microscopy, surface receptors, bioimage informatics
Subjects:	Biological Sciences > Biology > Cell Biology
Divisions:	Faculty of Science and Engineering > Biological Sciences, School of
Depositing User:	RD Admin
Date Deposited:	22 Aug 2012 10:58
Revision:	78
URI:	http://researchdata.essex.ac.uk/id/eprint/77

Available Files

Data

⊞ data_tiff.zip

⊞ data_nd2.zip

Documentation

⊞ doc_thumbnails.zip

⊟ Additional details

	Creators	Email
Creators:	Laissue, Philippe	plaissue@essex.ac.uk

Type of data:	Image	

	Contribution	Name	Email
	Contributor	Obara, Boguslaw	UNSPECIFIED
Contributors:	Contributor	Asma, Jabeen	UNSPECIFIED

Edit item: Sample Data: Bioimaging - Surface Receptors

[Upload] → [Details] → [Subjects]

[Save and Return]   [Cancel]   [Next >]

Add a new document

⚠ If this is your first time depositing, or if you require more information on any aspect of the deposit process, you may wish to view our Help page.

Click the Browse button below to select a file and the Upload button to upload it to the repository. If you are uploading a large number of files, you may find uploading and describing each file separately cumbersome. In these situations we suggest you upload the files in logically grouped zip bundles instead.

Please note that your progress will automatically be saved when you click 'Next' or 'Save and Return' above. Clicking 'Cancel' will result in the loss of any changes since last save.

The orange star symbol (⭐) indicates a mandatory field.

File	From URL

[Choose File] No file chosen

⊞ **Documentation (Thumbnails)**

13Mb

Hide options ➖

⭐ **Visible to:**	Repository staff only ▾	❓
⭐ **Content:**	Documentation ▾	❓
⭐ **Document description:**	Thumbnails	❓
Embargo date:	Year: [____] Month: Unspecified ▾ Day: ? ▾	❓
⭐ **License:**	Creative Commons: Attribution-No Derivative Works 3.0 ▾	❓

ReShare

ReShare metadata

- Further fine-tuning of metadata profile:
  - DDI 2.5 compliant
  - include DataCite DOIs, ORCID
  - customise for UK Data Service, e.g. CVs

- **File-level elements (6) (2M 3O 1A)**
  - describe individual data files or zip packages and their access conditions
- **Data collection level elements (31) (20M 11O 7A)**
  - descriptive information, people, coverage, methodology
- **Project-level elements (7) (2M 5O 7A)**
  - can be harvested from [RCUK Gateway to Research](#) (API)

(**M**andatory; **O**ptional; **A**utomatic)

UK Data Service

Controlled vocabularies

- elements commonly used by the Archive as facets in its search and browse platform [Discover](#)

    - subject categories (list)

    - HASSET keywords (webservice; Humanities and Social Science Electronic Thesaurus)

    - data type (list)

    - kind of data (list)

    - country (list)

    - spatial units (hierarchical CV)

    - observation units (list

Apache **Solr**™

UK Data Service

ReShare metadata profile

Project-level metadata	
GtR ID	A
Grant reference	O
Project title	O
Project description	O
Research funders	M/(A)
Grant holders	O
Project dates	O

File-level metadata	
Accessible to	M
File/bundle content	O
File/bundle description	O
Embargo date	O
Licence	M
Metadata version	A

Data collection metadata			
T&C agree	M	Geographic area	O
Date of deposit	A	Geographic location	O
Publication date	A	Country	M
Publisher	A	Spatial unit	O
Collection ID	A	Data collection methods	O
Data collection title	M	Observation unit	M/CV
Alternative title	O	Kind of data	M/CV
Data collection description	M	Provenance/ lineage	O
Keywords	M/CV	Type of data	M/CV
Subjects	M/CV	Language	M
Collection period	M/(A)	Relates resources	O/(A)
Temporal coverage	O	Notes administrator	O
		DOI	A

People	
Creators	M/(A)
Copyright holders	M/(A)
Contact	M
Contributors	O

Discover

Search and browse our data collections, support guides, case studies, and related publications.

Do you mean: biodiversity ?

🔍 biodiversity [GO]

Reset filters | Clear search | ☑ Auto-complete | ☑ Map search to HASSET thesaurus? | Help

👥 Case study Ⓒ Data collection Ⓢ Series record 📄 ESRC output 👤 Support guide

SEARCH RESULTS

Sorted by: [Relevance ▾]

Displaying 1-10 of 31 results

1 **2** **3** **4** ▶ ▶▶

- **Discover**
- **Variable and question bank**
- **QualiBank**

Type +
Subject +
Date +

Data type −

- ☐ UK Survey data (18)
- ☐ Qualitative and mixed methods data (11)
- ☐ Time series (2)
- ☐ Geospatial data (1)
- ☐ Historical data (1)

Show more... Refine

Key data +
Country +
Kind of data +
Spatial unit +

Analysis unit −

- ☐ Families and households (17)
- ☐ Groups (17)
- ☐ Individuals (13)
- ☐ Organisations (4)
- ☐ Geographic Units (2)

Show more... Refine

Access +
Depositor +

Ⓒ SN 6159 Eating Biodiversity: An Investigation of the Links between Quality Food Production and Biodiversity Protection, 2005-2007
Buller, H., University of Exeter. Department of Geography

+ Full record...

🛒 Download/Order | Get full DDI XML | Similar data collections

Ⓒ SN 7100 Linking Evidence and Policy for Managing Biodiversity in the Agricultural Landscape, 2011
Sutherland, W., University of Cambridge. School of the Biological Sciences. Department of Zoology

+ Full record...

🛒 Download/Order | Get full DDI XML | Similar data collections

Ⓒ SN 5023 Eurostat New Cronos, 1950-2060
Eurostat

+ Full record...

Explore Online | Get full DDI XML | Similar data collections

Ⓒ SN 3475 Survey of English Housing, 1993-1994
Office of Population Censuses and Surveys. Social Survey Division

+ Full record...

🛒 Download/Order | Get full DDI XML | Similar data collections

Ⓒ SN 3561 Survey of English Housing, 1994-1995
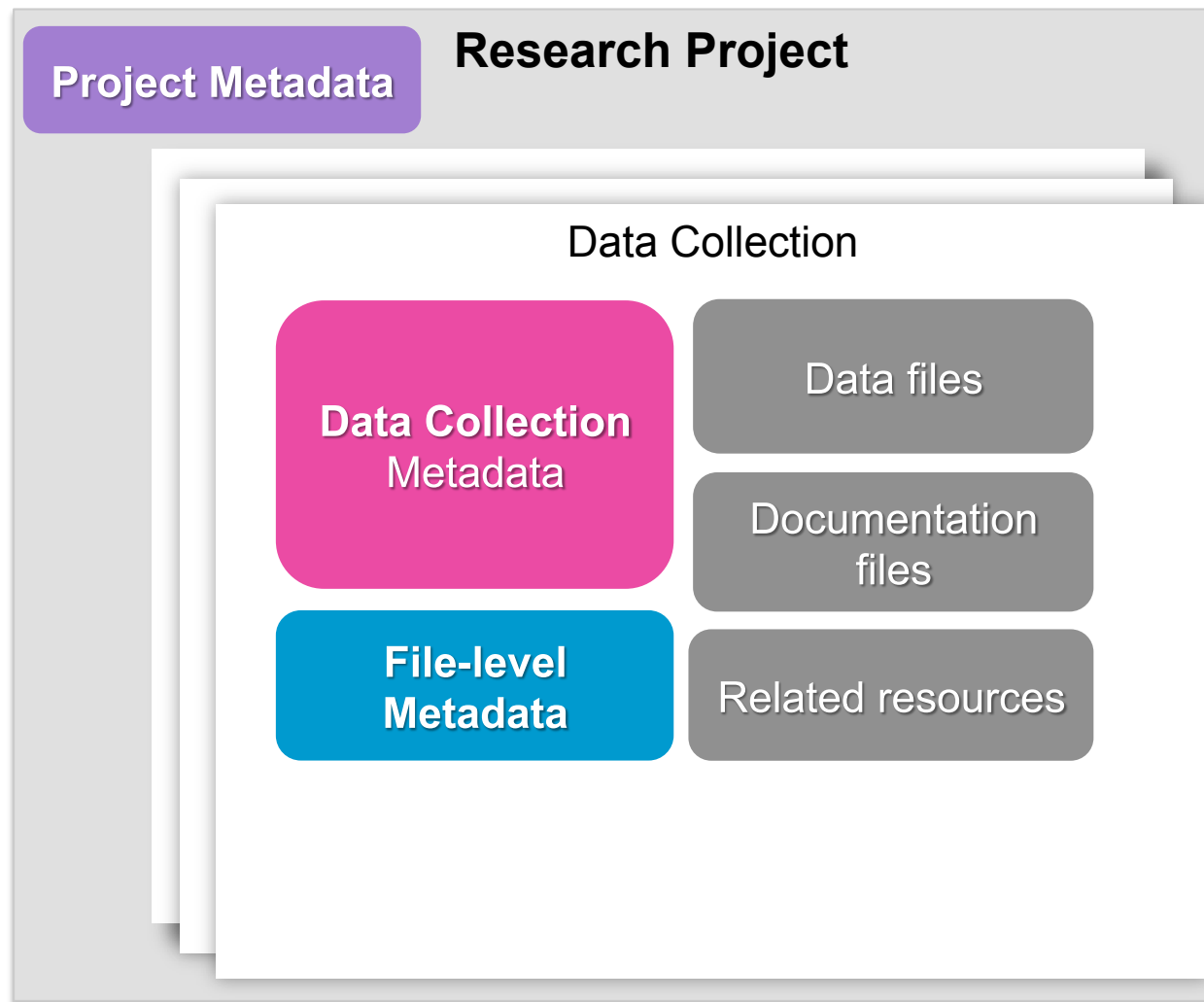Office for National Statistics. Social Survey Division

+ Full record...

🛒 Download/Order | Get full DDI XML | Similar data collections

Data Service

ReShare object model

Research Project

Project Metadata

Data Collection

Data Collection Metadata

Data files

Documentation files

File-level Metadata

Related resources

UK Data Service

ReShare user interface

- Revised deposit workflow
    - agree terms & conditions
    - project metadata
    - people
    - data collection metadata
    - upload files
    - file-level metadata, access & licence
    - deposit
- Extended help text & guidance, linked to our RDM guidance
- Customised UK Data Service styling, brand, legal info,…

UK Data Service

ReShare access options

- Access options for individual files or zip bundles:
    - open access
    - safeguarded access for users registered with the UK Data Service (controlled via Shibboleth authentication)
    - permission access, with access requests enabled by the depositor
    - embargo (up to 12 months)

- New user access control plugin to be developed (with EPrints Services and the University of Leeds)
    - depositors can authorise and execute permission access for confidential data, is still being developed in collaboration with

ReShare features

- UK Federation authentication of data depositors and data users
- Pull related project information from [RCUK Gateway to Research](#) (auto-fill 11 elements: project-level, people, dates)
- Publish metadata records to [Discover](#) portal
- DOI minted via customised DataCite DOI EPrints plugin
- Fedora Commons import plugin developed to transfer metadata records and data objects from legacy Fedora Commons system (ESRC Data Store), merging parent and child XML and transforming into EP3 XML
- UKDS staff review data collections before publishing:
  - confidentiality
  - copyright
  - documentation sufficient
  - long-term file formats

Edit collection: Endometriosis: improving the well being of couples

| Terms and conditions | → | Award details | → | People | → | Data collection | → | Upload |

Please note:

The pink asterisk symbol * indicates a mandatory field.

Your progress will be automatically saved when you click 'Next' or 'Save'. 'Cancel' will result in the loss of any changes since last save.

---

*

If your data collection results from a funded research project, provide project information here (else progress to the next step). This is different information from the information you will submit for the data collection.

Grant reference:
Submit your RCUK grant number (e.g. ES/K000000/1) to retrieve project information from the RCUK Gateway to Research. If grant number is unknown, you can find it by searching the Gateway to Research.

ES/J003662/1

Project title:
Project title; pre-filled if project information is retrieved from the RCUK Gateway to Research.

Endometriosis: improving the well being of couples

Project description:
The project description or abstract explains the general aims, purpose and background to the research project. This field will be pre-filled if project information is retrieved from the RCUK Gateway to Research.

three phases.

Phase One is a contextual phase consisting of interviews with key informants including healthcare providers, patients, and support group representatives.

Phase Two includes in-depth interviews with heterosexual women and their male partners.

Phase Three involves a Stakeholder Workshop which will collaboratively explore the findings and discuss how insights from this work contribute to improved support for couples and also inform theories of chronic illness more generally.

Edit collection: [Endometriosis: improving the well being of couples](#)

[Terms and conditions] → [Award details] → [People] → [Data collection] → [Upload]

✱ Creators   ?

	Family name	Given name / Initials	Email	Affiliation	ORCID	
1.	Culley	Lorraine				⌄
2.						⌄ ⌃
3.						⌄ ⌃

[Add row]

✱ Copyright holders   ?

	Family name	Given name / Initials	Email	Affiliation	ORCID	
1.	Culley	Lorraine				⌄
2.						⌄ ⌃
3.						⌄ ⌃

[Add row]

✱ Contact   ?

	Family name	Given name / Initials	Email	Affiliation	ORCID	
1.	Culley	Lorraine	lac@dmu.ac.uk	De Montfort Univer		⌄
2.						⌄ ⌃
3.						⌄ ⌃

[Add row]

Contributors   ?

	Family name	Given name / Initials	Email	Affiliation	ORCID	
1.	Hudson	Nicky				⌄
2.	Mitchell	Helene				⌄ ⌃

Edit collection: [Endometriosis: improving the well being of couples](#)

Terms and conditions → Award details → People → Data collection → Upload

* Data collection title   ?

Endometriosis: improving the well being of couples

+ Alternative title

* Data collection description   ?

Endometriosis is a common, disabling gynaecological condition affecting approximately 2 million women in the UK. Common symptoms are chronic pelvic pain, fatigue, heavy periods, and pain experienced during sexual intercourse. It may also affect fertility. There is no consensus on what causes endometriosis and there is no definitive cure. Despite the chronic and potentially disabling nature of endometriosis, along with its impact across a wide range of life domains, there is relatively little work on what it means to live with this condition. The overall aim of this study is to explore the impact of endometriosis on couples in particular, and to enhance the well being of people living with endometriosis by providing an evidence base for improving couple support. This is a qualitative study in three phases. Phase One is a contextual phase consisting of interviews with key informants including healthcare providers, patients, and support group representatives. Phase Two includes in-

* Keywords   ?

health **x**
couples **x**

[   ] Add

* Subjects   ?

Remove   Health

Add Economics

Add Education

Add Employment and labour

Add Environment, conservation and land use

☑ Health

Add History

Add Housing

Edit collection: Endometriosis: improving the well being of couples

Terms and conditions → Award details → People → Data collection → Upload

Add a new file or zip bundle

Upload data files and documentation for your data collection. Click the Browse button below to select a file; and the Upload button to upload it to the repository. Large files (>2GB) make take a long time to upload, depending on your browser speed. When uploading many files (>10), please upload them in logically grouped zip bundles instead, e.g. grouping by content (data, documentation), by file format (text, images, databases) or by folder structure.

Make sure data files contain no disclosive personal information.

Include documentation on data collection methods and context to inform future reuse of the data.

File	From URL

Choose File no file selected

+
Endopart_study_summary_report_and_recommendations.pdf

904Kb

Show options +

+ Endopart interview schedule men.pdf

250Kb

Hide options —

✱ Accessible to:	Registered users only (safeguarded data)	?
File or bundle content:	Documentation	?
File or bundle description:		?
Embargo date:		?
✱ License:	UK Data Service End User Licence	?

Save

UK Data Service

Conclusion

- Significant developments for EPrints repository software holding collections of research data
    - expanded metadata profile
    - optimised interface and workflow
    - connections with external systems
- Expanding data infrastructure
    - ReCollect: institutions can install data repository with relative ease
    - ReShare augments UK Data Archive data infrastructure for storing, discovery and dissemination of social sciences and humanities data collections
- Ability to hold restricted confidential / sensitive

UK Data Service

Ways forward

- ReCollect was first EPrints plugin to implement standards-based metadata for data
- Release ReCollect2 ([metadata profile fine-tuned](#) with University of Leeds & University of Glasgow)
- ReCollect uptake by institutions:
    - University of Glasgow
    - University of Leeds
    - University of Southampton
    - Goldsmiths, University of London
    - London School of Hygiene and Tropical Medicine
    - University of East London
    - UK Data Service ReShare
    - Western Balkan countries (shared infrastructure under development)
- Master ReCollect could be used as 'Repository as a Service' for small institutions to buy a space on federated server

UK Data Service

Ways forward

- Release more plugins to EPrints bazaar
  - Gateway to Research harvest
  - Fedora import
  - user access module

UK Data Service

Questions

Veerle Van den Eynden

veerle@essex.ac.uk