



STUDY-LEVEL METADATA FOR DATA COLLECTIONS IN A REPOSITORY

VEERLE VAN DEN EYNDEN
UK DATA ARCHIVE
UNIVERSITY OF ESSEX

OR2014 Conference
Practical data curation and management for institutional repositories
Helsinki, 9 June 2014



OVERVIEW

How do repositories 'describe' data collections ?

- data creator perspective
- data user perspective

- data should be easily discoverable, accessible, assessable, intelligible, useable, and wherever possible interoperable to specific quality standards

[G8 Science Ministers Statement, 13 June 2013](#)

As examples:

- ReCollect metadata
- ReShare metadata

OVERVIEW

- ReCollect = EPrints plugin for institutional data repository



- pilot for University of Essex
- multi-disciplinary data collections
- <http://bazaar.eprints.org/280/>

- ReShare = subject repository



- UK Data Archive repository for short-term curation
- wider social sciences & humanities
- <http://reshare.ukdataservice.ac.uk/>



WHY METADATA

- Describe data collection
 - Title, abstract, keywords, coverage, collection method,
- Find a data collection (discovery)
 - Keyword, subject, controlled vocabularies, standardisation of metadata across repositories & disciplines (e.g. data registry)
- Assess use potential of data
 - Collection methods, sample size, data content,
- Know how to access data
 - Access conditions, licence
- Citation of data use
 - Author/creator, publisher, dates



METADATA IN PRACTICE

- Data creator provides metadata; wants minimal effort
- Data user needs optimal / maximum detail to judge usefulness of data
- Standardised terminology
- Typically combination of mandatory and optional elements



METADATA STANDARDS

- Dublin Core
- DataCite
- DDI (Data Documentation Initiative)
- Gemini
- INSPIRE directive

RECOLLECT METADATA PROFILE



Core

- Metadata for **findability**
- e.g. author, discipline, project

Detail

- **Generic dataset description**
- e.g. collection methodology

Discipline

- As additional **file**
- Structured e.g. XML

Based on JISC IDMB project outputs, Southampton University

RECOLLECT METADATA PROFILE



- DataCite core as minimal metadata standard for IRs in UK
- Minimising barriers for researchers to deposit data
 - ...while satisfying requirements for re-use

I have no time, depositing data needs to be quick and easy!

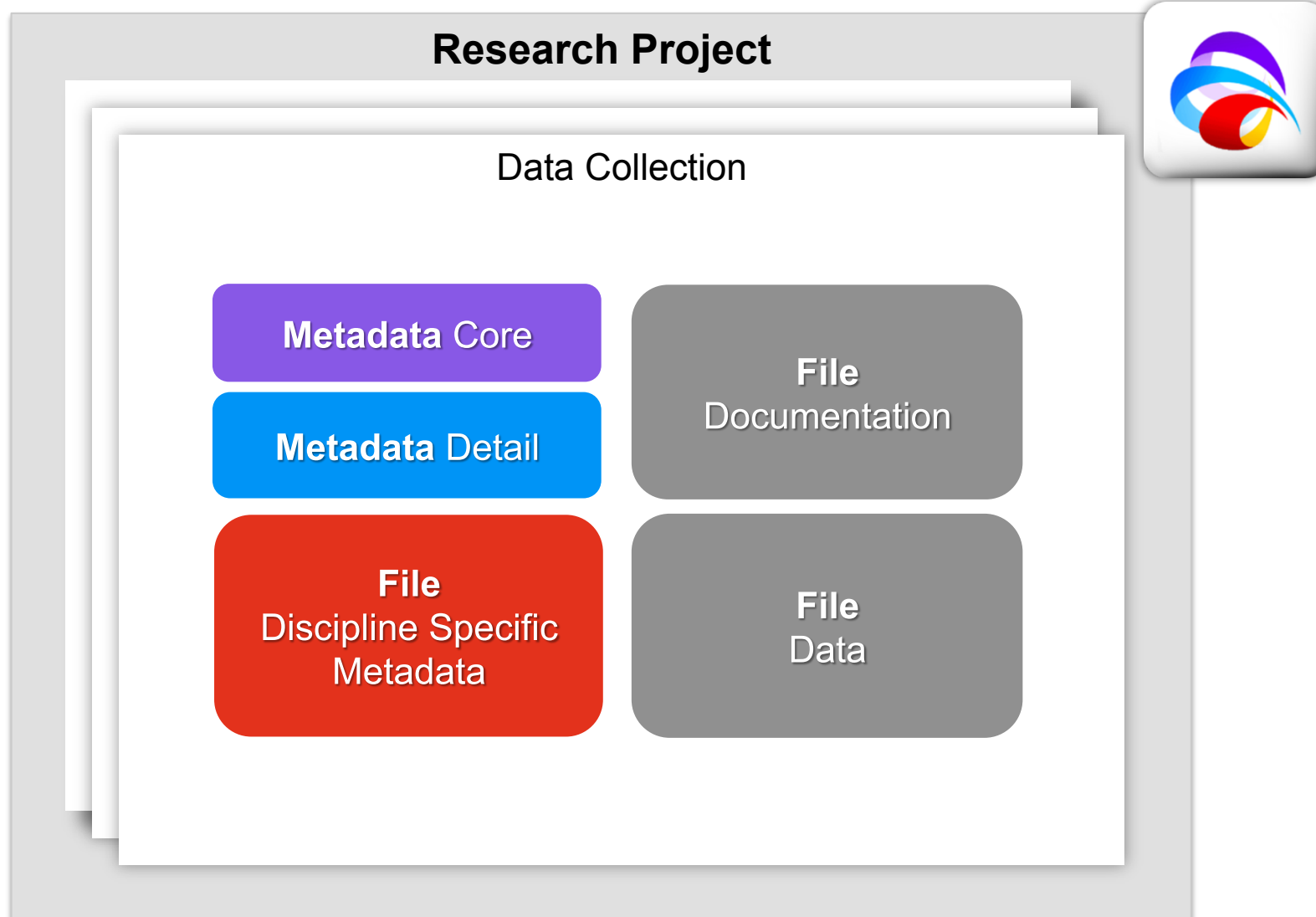
I want to use this dataset, but I can't understand it!

METADATA PROFILE

- We took standards-based approach, producing:
 1. **Metadata mapping** between relevant schemas
 2. **Metadata profile** built from mapped elements
- Standards that we mapped:
 - **DataCite** – minimal mandatory metadata (discovery)
 - **INSPIRE** – basic contextual information (EU directive for geospatial / environmental data) (= Gemini 2.1 in UK)
 - **DDI 2.1** – social sciences standard



OBJECT MODEL



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](#)

[+ data_nd2.zip](#)

Documentation

[+ doc_thumbnails.zip](#)

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: 5


Revision:

Mime-Type: application/zip

File size: 2175Mb

 [data_nd2.zip](#)

Documentation

 [doc_thumbnails.zip](#)

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	plaisue@essex.ac.uk

Type of data: Image

	Contribution	Name	Email
Contributors:	Contributor	Obara, Boguslaw	UNSPECIFIED
	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

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

"Archiving and
sharing research
data"



ReShare is the UK Data Service's online data repository for the archiving and sharing of research data.

ReShare offers access to open and safeguarded data collections for research and learning. Data can be found via Discover, the discovery portal of the [UK Data Service](#).

Researchers can easily deposit collections of data and accompanying documentation, after registering with the UK Data Service. ESRC grant holders deposit their data with ReShare, as a contractual requirement under the [ESRC research data policy](#). When creating a metadata record for a data collection, researchers can retrieve related project information from the [RCUK Gateway to Research](#).

The UK Data Service sustainably manages and archives the data collections and reviews data before their release, to ensure they conform with ethical and legal requirements.

[LOGIN](#) / [REGISTER](#)

[DISCOVER UK DATA SERVICE](#)



[GO](#)

Data

Website

QUICK ACCESS TO

[Deposit data collection](#)

RESHARE METADATA



- **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 1A)**
 - harvested from [RCUK Gateway to Research](#)

(Mandatory; Optional; Automatic)

CONTROLLED VOCABULARIES



- elements commonly used by the Archive as facets in its search and browse platform [Discover](#)
 - subject categories
 - HASSET keywords
 - data type
 - kind of data
 - country
 - spatial units
 - observation units

RESHARE METADATA PROFILE

1.1	Terms and conditions confirmation	Mandatory
2.1	GtR ID	Automatic
2.2	Award or project grant number	Optional
2.3	Award or project title	Optional
2.4	Award or project description	Optional
2.5	Award or project funders	Mandatory
2.6	Award or project PI / Co-PI	Optional
2.7	Award or project period	Optional
3.1	Creators	Mandatory
3.2	Copyright holders	Mandatory
3.3	Contact	Mandatory
3.4	Contributors	Optional
4.1	Depositor	Automatic
4.2	Date of deposit	Automatic
4.3	Version	Automatic
4.4	Date of publication	Automatic
4.5	Publisher	Automatic
4.6	Collection identifier	Automatic
4.7	Title	Mandatory
4.8	Alternative title	Optional
4.9	Description	Mandatory
4.10	Keywords	Mandatory

4.11	Subjects	Mandatory
4.12	Dates of collection	Optional
4.13	Temporal coverage	Optional
4.14	Geographic coverage	Optional
4.15	Geographic bounding box	Optional
4.16	Country	Mandatory
4.17	Spatial unit	Optional
4.18	Method of data collection	Optional
4.19	Observation unit	Mandatory
4.20	Kind of data	Mandatory
4.21	Provenance / lineage	Optional
4.22	Datatype	Mandatory
4.23	Resource language	Mandatory
4.24	Related resources	Optional
4.25	Access note	Optional
4.26	Persistent identifier (DOI)	Automatic
5.1	Document metadata version	Automatic
5.2	Document access level	Mandatory
5.3	Document type	Optional
5.4	Document description	Optional
5.5	Document embargo	Optional
5.6	Document licence	Mandatory





Discover

- 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 +

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

Do you mean: **biodiversity** ?

[Reset filters](#) | [Clear search](#) | [Auto-complete](#) | [Map search to HASSET thesaurus?](#) | [Help](#)

- Case study
- Data collection
- Series record
- ESRC output
- Support guide

Sorted by:

SEARCH RESULTS

Displaying 1-10 of 31 results 1 2 3 4 >>>

- 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, 1950-2080**
 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](#)

ARCHIVE

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.

Project title:

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

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.						▼ ^

*** Copyright holders** ?

	Family name	Given name / Initials	Email	Affiliation	ORCID	
1.	Culley	Lorraine				▼
2.						▼ ^
3.						▼ ^

*** Contact** ?

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

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 ?

healthx
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) may 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





CONTACT

UK DATA ARCHIVE
UNIVERSITY OF ESSEX
WIVENHOE PARK
COLCHESTER
ESSEX CO4 3SQ

T: +44 (0)1206 872001

E: datasharing@ukdataservice.ac.uk

W: www.ukdataservice.ac.uk
