



# Research Data @ Essex: implementing EPrints for research data

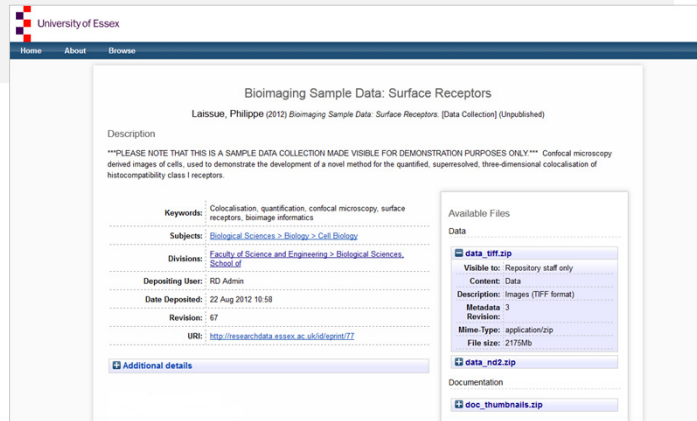


Tom Ensom, Alexis Wolton, Veerle Van den Eynden and Louise Corti  
UK Data Archive, University of Essex, United Kingdom



This project is piloting a research data management and sharing infrastructure for the University of Essex. The project team, lead by the UK Data Archive, includes staff from Information Systems & Services (ISS), the Research and Enterprise Office (REO) and Records Management.

- Core**
  - Metadata for **findability**
  - e.g. author, discipline, project
- Detail**
  - **Generic dataset detail**
  - e.g. collection methodology
- Discipline**
  - As additional **file**
  - Structured e.g. XML



## TECHNICAL / REPOSITORY

- **EPrints data instance** populated with test data
- **User interface & workflow customisations** for complex data collections
- **Metadata profile** to describe research data at discovery and 'generic description' levels of detail (see diagram - based on IDMB project outputs).
- Metadata compliant with **standards** relevant to UK higher education and research data (DataCite, GEMINI/INSPIRE) and drawing on others e.g. DDI.
- To be published as an EPrints **plugin** for re-use by wider community
- Currently exploring persistent identifier options, including implementation of UK Data Archive's **DataCite DOI** methodology



## SUPPORT & TRAINING

- **Data management planning guidance** produced for research office website
- Providing data management and sharing **training sessions** for university staff & students
- **User guide** for repository users including guidance on best practise preparation
- **Moodle training modules** on data management issues
- Will examine **business case** using UK Data Archive Costing Tool and Institutional Data Management Blueprint business model.

## POLICY & ADVOCACY

- **11 principle** research data management policy
- **EPSRC roadmap** drafted, identifying steps to policy compliance
- Approved by strategic groups

DEPARTMENT	RESEARCH AREA	SAMPLE DATA
Biological Sciences	Proteomics	mass spectrometry data from tumour tissue samples
	Bioimaging	high resolution image data collected to examine cellular structure
Essex Business School	Management	tabular data on football managers, using publicly influenced performance metrics to examine managerial performance and succession
Language & Linguistics	Second language acquisition	audio and transcripts of classroom second language learners
	Sociolinguistics	audio and transcripts of interviews with multiple generations of Indian English speakers
Computing and Electronic Systems	Artificial intelligence	crowd sourced AI scripts with results of a competition between these AIs

## PILOT DEPARTMENTS

- Comprehensive **data inventory template** published
- Plus minimal *Understanding a researchers' data in 10 questions* strategy
- Report published internally based on **data asset assessment & interviews**
- **Sample data collections** ingested and described (see table above)
- Positive **feedback from researchers** on repository pilot and training opportunities

