



**British
Geological Survey**

NATURAL ENVIRONMENT RESEARCH COUNCIL

Geoscience for our changing Earth

ESAA and Data Citation for UKCCS

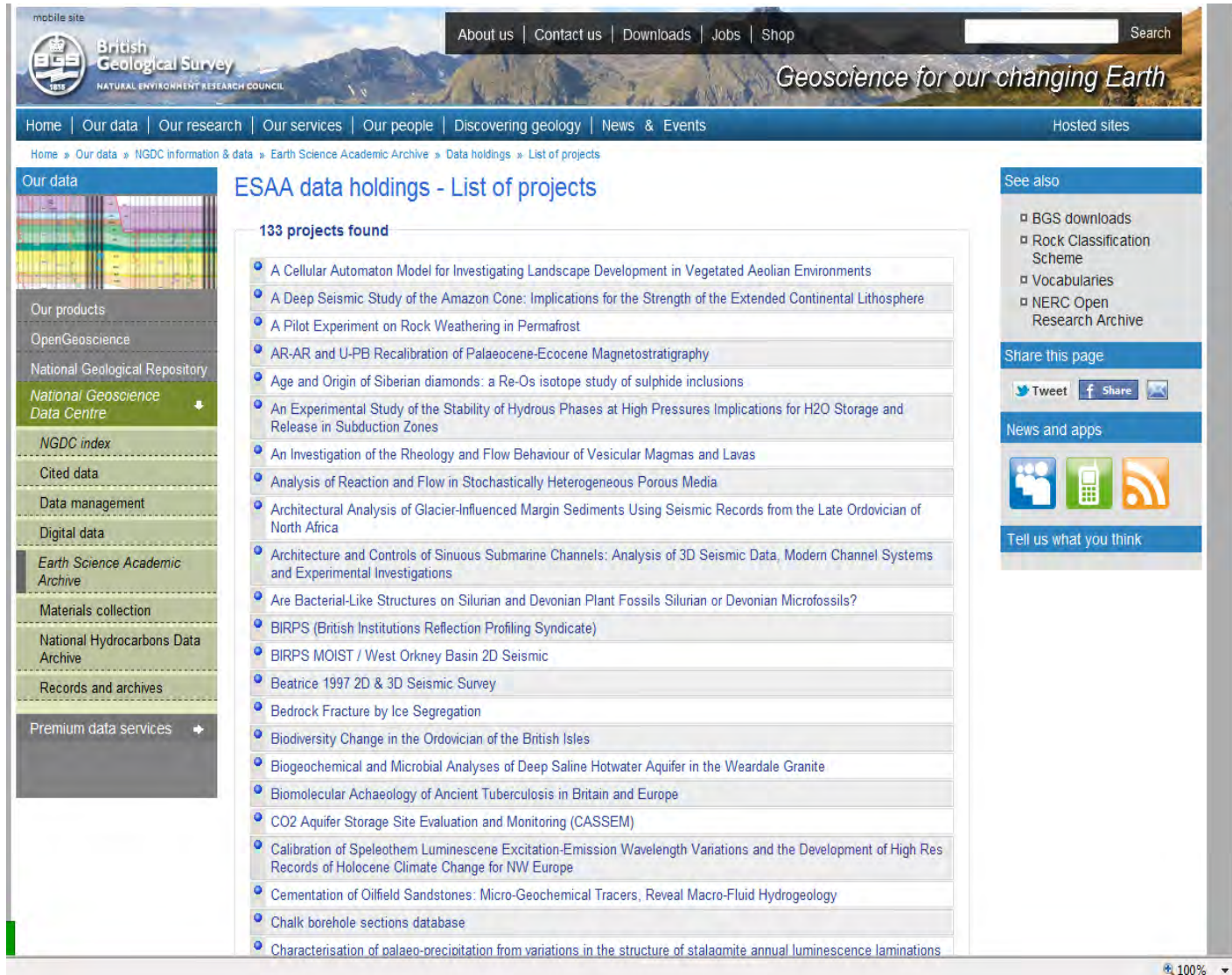
Rod Bowie September 2013



What is ESAA

- <http://www.bgs.ac.uk/services/ngdc/esaa/home.cfm>
- Set up to manage, within the NGDC, the data received from NERC Earth Science grants
- Expanded to include all similar data archived for academic research projects
- Discovery metadata to ISO 19115:2003
- Currently 133 projects on-line
- Confidentiality of data assured within embargo period
- Knowledge of availability of data means potential collaboration and re-use
- This well proven system forms the basis for the management of data in the new UKCCS Data Hub

Projects in ESAA



The screenshot shows the British Geological Survey website. The header includes the BGS logo, navigation links (About us, Contact us, Downloads, Jobs, Shop), a search bar, and the slogan "Geoscience for our changing Earth". The main navigation bar lists "Home", "Our data", "Our research", "Our services", "Our people", "Discovering geology", "News & Events", and "Hosted sites". The breadcrumb trail reads: "Home » Our data » NGDC information & data » Earth Science Academic Archive » Data holdings » List of projects".

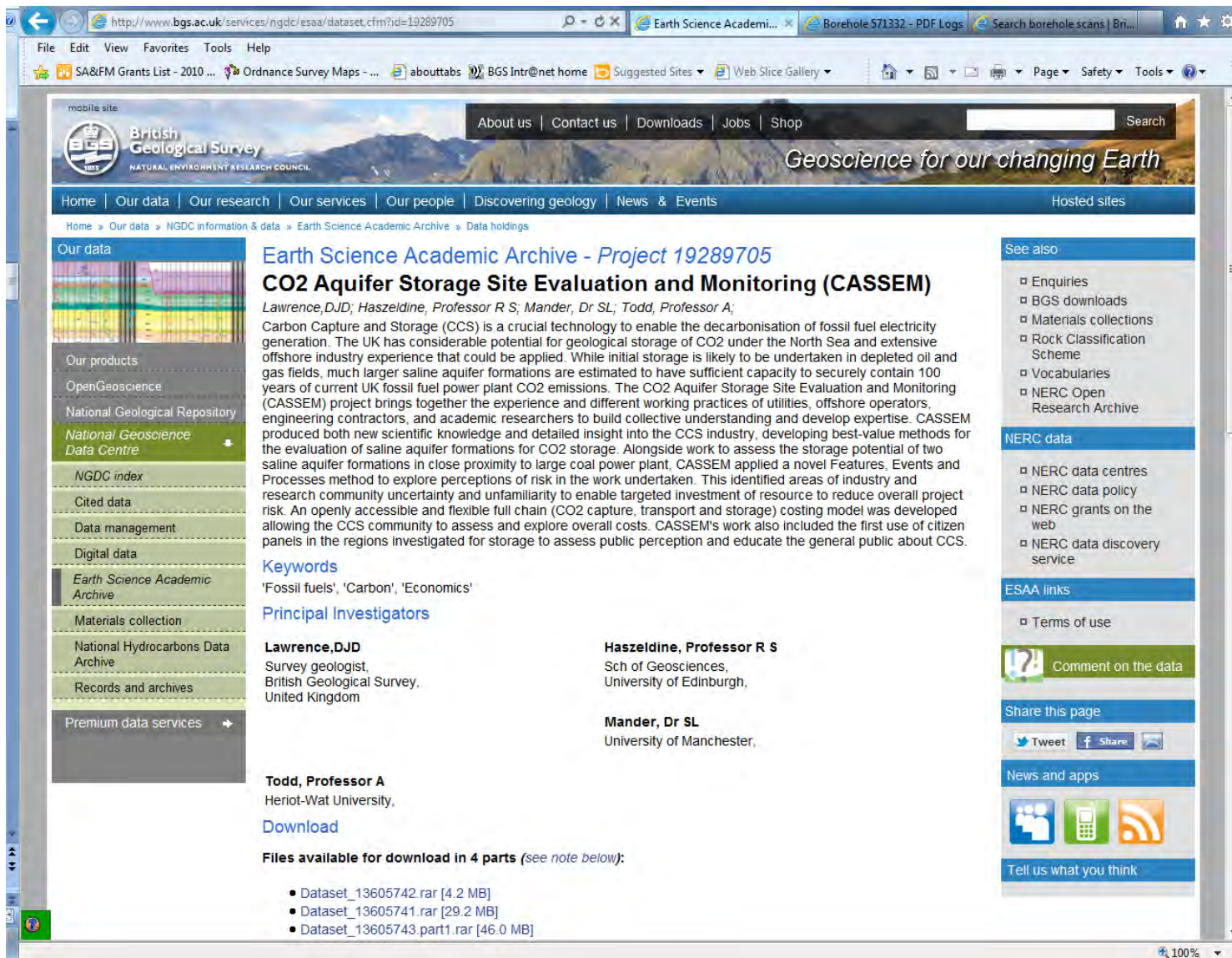
The main content area is titled "ESAA data holdings - List of projects" and displays "133 projects found". A list of project titles is shown, including:

- A Cellular Automaton Model for Investigating Landscape Development in Vegetated Aeolian Environments
- A Deep Seismic Study of the Amazon Cone: Implications for the Strength of the Extended Continental Lithosphere
- A Pilot Experiment on Rock Weathering in Permafrost
- AR-AR and U-PB Recalibration of Palaeocene-Eocene Magnetostratigraphy
- Age and Origin of Siberian diamonds: a Re-Os isotope study of sulphide inclusions
- An Experimental Study of the Stability of Hydrous Phases at High Pressures Implications for H2O Storage and Release in Subduction Zones
- An Investigation of the Rheology and Flow Behaviour of Vesicular Magmas and Lavas
- Analysis of Reaction and Flow in Stochastically Heterogeneous Porous Media
- Architectural Analysis of Glacier-Influenced Margin Sediments Using Seismic Records from the Late Ordovician of North Africa
- Architecture and Controls of Sinuous Submarine Channels: Analysis of 3D Seismic Data, Modern Channel Systems and Experimental Investigations
- Are Bacterial-Like Structures on Silurian and Devonian Plant Fossils Silurian or Devonian Microfossils?
- BIRPS (British Institutions Reflection Profiling Syndicate)
- BIRPS MOIST / West Orkney Basin 2D Seismic
- Beatrice 1997 2D & 3D Seismic Survey
- Bedrock Fracture by Ice Segregation
- Biodiversity Change in the Ordovician of the British Isles
- Biogeochemical and Microbial Analyses of Deep Saline Hotwater Aquifer in the Weardale Granite
- Biomolecular Archaeology of Ancient Tuberculosis in Britain and Europe
- CO2 Aquifer Storage Site Evaluation and Monitoring (CASSEM)
- Calibration of Speleothem Luminescence Excitation-Emission Wavelength Variations and the Development of High Res Records of Holocene Climate Change for NW Europe
- Cementation of Oilfield Sandstones: Micro-Geochemical Tracers, Reveal Macro-Fluid Hydrogeology
- Chalk borehole sections database
- Characterisation of palaeo-precipitation from variations in the structure of stalaomite annual luminescence laminations

The left sidebar contains a "Our data" menu with options like "Our products", "OpenGeoscience", "National Geological Repository", "National Geoscience Data Centre", "NGDC index", "Cited data", "Data management", "Digital data", "Earth Science Academic Archive", "Materials collection", "National Hydrocarbons Data Archive", "Records and archives", and "Premium data services".

The right sidebar includes a "See also" section with links to "BGS downloads", "Rock Classification Scheme", "Vocabularies", and "NERC Open Research Archive". Below this is a "Share this page" section with "Tweet" and "Share" buttons, and a "News and apps" section with icons for people, a mobile phone, and RSS. At the bottom of the sidebar is a "Tell us what you think" section.

Data



mobile site

British Geological Survey
NATURAL ENVIRONMENT RESEARCH COUNCIL

About us | Contact us | Downloads | Jobs | Shop

Geoscience for our changing Earth

Home | Our data | Our research | Our services | Our people | Discovering geology | News & Events

Home » Our data » NGDC information & data » Earth Science Academic Archive » Data holdings

Our data

Our products

OpenGeoscience

National Geological Repository

National Geoscience Data Centre

NGDC index

Cited data

Data management

Digital data

Earth Science Academic Archive

Materials collection

National Hydrocarbons Data Archive

Records and archives

Premium data services

Earth Science Academic Archive - Project 19289705

CO2 Aquifer Storage Site Evaluation and Monitoring (CASSEM)

Lawrence, DJD; Haszeldine, Professor R S; Mander, Dr SL; Todd, Professor A;

Carbon Capture and Storage (CCS) is a crucial technology to enable the decarbonisation of fossil fuel electricity generation. The UK has considerable potential for geological storage of CO2 under the North Sea and extensive offshore industry experience that could be applied. While initial storage is likely to be undertaken in depleted oil and gas fields, much larger saline aquifer formations are estimated to have sufficient capacity to securely contain 100 years of current UK fossil fuel power plant CO2 emissions. The CO2 Aquifer Storage Site Evaluation and Monitoring (CASSEM) project brings together the experience and different working practices of utilities, offshore operators, engineering contractors, and academic researchers to build collective understanding and develop expertise. CASSEM produced both new scientific knowledge and detailed insight into the CCS industry, developing best-value methods for the evaluation of saline aquifer formations for CO2 storage. Alongside work to assess the storage potential of two saline aquifer formations in close proximity to large coal power plant, CASSEM applied a novel Features, Events and Processes method to explore perceptions of risk in the work undertaken. This identified areas of industry and research community uncertainty and unfamiliarity to enable targeted investment of resource to reduce overall project risk. An openly accessible and flexible full chain (CO2 capture, transport and storage) costing model was developed allowing the CCS community to assess and explore overall costs. CASSEM's work also included the first use of citizen panels in the regions investigated for storage to assess public perception and educate the general public about CCS.

Keywords

'Fossil fuels', 'Carbon', 'Economics'

Principal Investigators

Lawrence, DJD Survey geologist, British Geological Survey, United Kingdom	Haszeldine, Professor R S Sch of Geosciences, University of Edinburgh,
Todd, Professor A Heriot-Wat University,	Mander, Dr SL University of Manchester,

Download

Files available for download in 4 parts (see note below):

- Dataset_13605742.rar [4.2 MB]
- Dataset_13605741.rar [29.2 MB]
- Dataset_13605743.part1.rar [46.0 MB]

See also

- Enquiries
- BGS downloads
- Materials collections
- Rock Classification Scheme
- Vocabularies
- NERC Open Research Archive

NERC data

- NERC data centres
- NERC data policy
- NERC grants on the web
- NERC data discovery service

ESAA links

- Terms of use

Comment on the data

Share this page

Tweet Share

News and apps

Tell us what you think

ESAA provides

- Secure storage for all formats of data and materials
- Documentation guides and advice
- Support
- Help with Data Management Planning
- Advice about IPR and third party ownership
- Maintains and manages confidentiality and release of data
- Data Citation
- The same services will be provided through the UKCCS Data Hub

Receipt of Data

- All data is accessioned in the our Accession system. This high level system records basic details about the data; who it comes from and records any terms and conditions
- All data is transferred to the original data folders in the UKCCS space on the SAN.
- This is backed up off site and also archived
- File transfers from ftp or disc take time and may involve conversion or manipulating the data in some way

Checking and confirming metadata

- Basis data entered in ISO compliant Metadata system
- A script checks all mandatory fields have been completed
- However checking for the quality of the abstract, lineage etc. has to be carried out manually.
- Discussions between Data Centre staff and data supplier may involve several conversations






*SIMPLY EXPLAINED:
METADATA*



geek & poke

Data quality checks

- The NGDC/NGR deals with a wide variety of datasets
- We may have no specific experience of your data

Name	Date modified	Type	Size
 Electron Microprobe Data-Entrada sandstone	15/11/2012 15:11	Adobe Acrobat D...	73 KB
 Green River Bulk rock chemistry data 1	21/05/2012 17:14	Microsoft Office E...	10,321 KB
 internal rsd's trace element data1	21/05/2012 17:15	Microsoft Office E...	163 KB
 Sequential leaching Entrada sandstone	15/11/2012 15:21	Microsoft Office E...	274 KB
 Sequential leaching method for file - sequential leaching Entrada sandstone	15/11/2012 15:21	Adobe Acrobat D...	32 KB

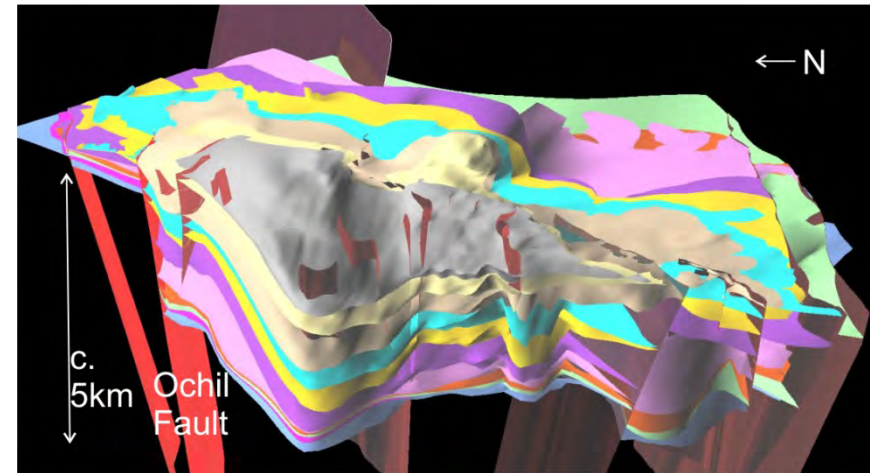
- We have to accept a lot on faith and work with you
- Many projects will create multiple files in multiple formats
- Details for the abstract may have to be checked with and updated from other sources and papers

Data Quality

51101
Bram Moor Borehole

Sample	Bed / Member	Depth M	K+C	M	Sum	OP	Porosity % of grain	Grain size	Grain shape	K+C	M	Sum
BM 42	Conistone Bed	82.00	3	71	26							
43	Conistone Bed	83.50	4	79	17							
40	Outfall Clay	85.00	11	57	32		4.6	NA	(-)			
41	Conistone Bed	86.50	16	39	45		7.6	NA	(-)			
37		88.00	17	46	37		6.1	NA	(-)			
38		89.00	18	49	32		6.5		(-)			
39		91.00	24	56	20		6.6		(-)			
36	Kellaways Rock	95.69	31	45	23		1.3		(-)			
35		98.30			6.2				(?)			
34	Scalby Fm.	100.88	76	15	9							
33		102.44	60	40	0							
31	Scarborough Fm.	104.22							(-)			
30		105.94	21	44	34		6.1		(-)			
29		107.42					6.2		(-)			
28	Clayton Fm.	108.41	52	18	21		6.3		(?)			
27		110.69	60	33	7				(-)			
26		111.82	50	23	27				(-)			
24		114.22	68	18	14				(-)			
25	Lebanon Mem.	114.80							(-)			
23		115.72					6.3		(?)			
22		117.70	61	27	11		6.2		(-)			
21		119.40	55	34	11		6.0		(-)			
22		121.25	53	32	16		6.2		(?)			
20		122.07	39	57	23		6.0		(?)			
19		125.41	34	41	25		6.6		(-)			
18		127.51	37	43	20		7.3		(-)			
17	Wheatstone Mem.	129.20	60	33	7				(-)			
16		131.03	22	46	22				(-)			
15		132.87	24	25	21		5.0		(-)			
14		134.94					6.0		(-)			
13	Eller Bank Fm.	136.31					6.2		(-)			

VS



■ This gives an indication of the variety of data held

Archiving / creation of deposited data

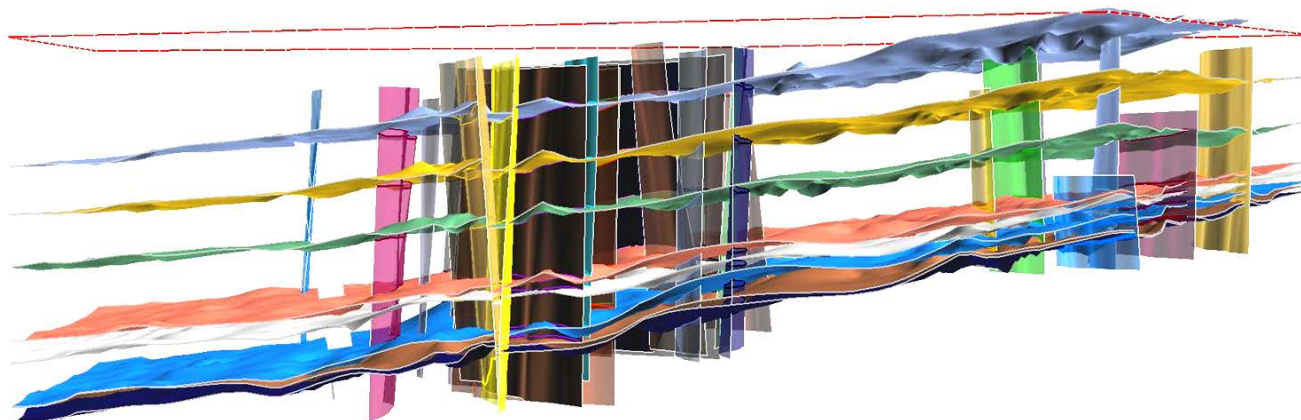


- Committing the data to archive is relatively easy; we have established routines and processes
- The factors we have to consider are:
 - ▶ Is the descriptive and technical metadata good enough to describe this dataset uniquely?
 - ▶ How big is the dataset? 1TB can take 10 hours to consign to tape and how will we make this information available?



Additional Data Centre specific steps

- We will link the datasets into the UKCCS/ESAA
- This means creating a copy of the data
- The UKCCS/ESAA website provides additional metadata and direct download facilities



Data Citation - Why

- A DOI enables scientists to cite datasets in the same manner as a scientific journal article
- Enabling credit to be assigned to the dataset creators, and ensuring the discoverability, permanence and stability of the dataset.
- Issuing DOIs and the publication of data recognises the value of the data and the effort that has gone into its creation, capture and effective management.
- DOIs allow formal publication of the dataset in data journals.

What sort of data can we/will we assign a DOI to?

Dataset has to be:

- Stable (i.e. not going to be modified)
- Complete (i.e. not going to be updated)
- Permanent – by assigning a DOI we're committing to make the dataset available for posterity
- Good quality – by assigning a DOI we're giving it our data centre stamp of approval, saying that it's complete and all the metadata is available

Landing Page

- When a dataset is cited that means:
 - There will be bitwise fixity
 - With no additions or deletions of files
 - No changes to the directory structure in the dataset “bundle”
- A DOI should point to a *html representation of some record which describes a data object – i.e. a landing page.*
- Upgrades to versions of data formats will result in new editions

Will all data get DOI

- No – If it doesn't meet the requirements or a DOI is not requested.
- The dataset will need to be permanent, by being deposited in UKCCS. A copy will be placed in a separate area on the Data Centre SAN.
- The ownership of the dataset has to be clearly defined, along with the licensing terms to allow us to distribute it.
- The metadata describing the dataset must be complete.

Issuing of DOI to depositor

- Creation of landing page is currently a manual process taking up to a maximum 0.5 days per DOI
- At the current rate of ingestion this is OK, but when rates increase this will be automated by linking it to our metadata systems
- Issuing a DOI can be as quick as 15 minutes if all the correct information is to hand
- DOI issued through DataCite's agent in the UK , the British Library

Questions?

