METADATA MARVELS AT THE UK DATA ARCHIVE

CONTROLLED VOCABULARIES
The UK Data Archive is guardian of the Humanities and Social Science Electronic Thesaurus (HASSET) and of its multilingual sister, the European Language Social Science Thesaurus (ELSTST):
- ELSTST has been translated into nine languages, with three more on the way
- HASSET has been in use and developed by the UK Data Archive over more than 50 years
- useful for indexing data, in study level and variable level
- allows retrieval of data and related documentation using hierarchies of keywords
- recently converted to SKOS, using the Pudgy tool
- between HASSET and ELSTST represent 1,769 concepts and 4,032 synonyms
- in SKOS, that’s 10,808 relationships as triples

QUDEXT
Schema that enables discovery, locating, retrieving and citing complex qualitative data collections: inc:
- highly structured and consistently marked-up data
- rich descriptive metadata for file entry, interview characteristics, interview setting, type of object
- logical links between data objects; text to related audio, images, and other research outputs
- preserves references to annotations performed on data
- common metadata elements that enable federated catalogue across providers and borders
- maintained by the UK Data Archive, University of Essex

US DATA ARCHIVE
With over 40 years of use and promotion of metadata standards for collection description and data description in the social science domain, the UK Data Archive:
- is an active promotor and guardian of standards and is engaged in:
  - Discovery open metadata principles
  - Jisc-CASRAI
  - DDI and DDI Controlled Vocabularies Group
  - DDI Qualitative metadata working group
- undertakes R&D metadata projects:
  - ESRC, Digital Futures, UK Quali Bank
  - Jisc Research Data Registry
  - Welcome Data Discoverability
  - JiscSKOS-HASSET
  - Jisc LiGeo

CITATION
Digital Object Identifiers (DOI) ensure persistent web address for digital collections
- British Library is the registration agency that provides our DOIs
- not-for-profit organisation managed by the German National Library of Science and Technology with other international members
- Our DOIs:
  - DataCite DOIs are assigned to each data collection, our DOIs support changes, so that if the location of the data changes, the DOI will always link to the data cited
- APA:
  - standardised referencing format for various materials
  - innovation: we have developed a paragraph-level reference list citation format
- e.g., Meredian, D., (Interview with Cloud) in Mathers Alonso; Poverty and the Fabian Family, 05/01454, UK Data Service (distributor), February 2003.
  - from: http://datahub.is/datastore/uk/ukdatasource/Documents/了一会儿

ENTITY IDENTITY
- We support ORCID in our self-deposit repository
- We are part of Jisc-CASRAI Authoritative Lists Working Group for a recommended organisational identifier standard
- We will support Premis the agreed standardised way to report funding sources for published scholarly research

RESEARCHERS DIRECTORY
- We support Open Archival Information System (OAIS)
  - for organisational structure and systems for archives with responsibility for preserving information and making it available to a designated community
- We will support Premis to provide granular structured machine-actionable preservation event statements

DDI CODEBOOK 2.5
The Data Documentation Initiative (DDI) is a specification for capturing metadata about social science data:
- originally created to capture the information found in survey codebooks, which remains the focus of earlier versions
- used for:
  - basic study-level catalogue metadata
  - rich variable description for survey files
- maintained by the Data Documentation Initiative Alliance, a non-profit service consortium including universities, data archives, and international organisations
- made available via OAI/PMH and as DDI XML

www.data-archive.ac.uk