CESSDA Data Catalogue (CDC)

Will launch in June 2018

Does not hold any data itself, or provide direct access to data

Automatically, frequently and periodically harvests metadata from Service Providers and other sources

Fully automated harvesting is crucial to ensure that the CESSDA Data Catalogue content is up to date

Researchers are directed to the source/service where the data may be obtained or analysed

See also https://www.cessda.eu/Research-Infrastructure/Products-Services
CESSDA Data Catalogue

Circa 13,000 records to date:
  FSD
  GESIS
  Progedo
  UKDS

Actively working with other Service Providers
Basic Features

13260 results found in 124ms

Topic
Collection years
Country
Search countries
Publisher
Search publishers
Language of data files
Search languages

Results per page 30
Sort by Title (ascending)

Perspective Germany 2005/06
McKinsey & Company, Düsseldorf; Stern, Hamburg; Zweites Deutsches Fernsehen (ZDF), Mainz; (1 more)

Applicant Countries Eurobarometer 00 (pilot study)
European Commission, Brussels DG Education and Culture (EAC-D2, Public Opinion Analysis)
Attitude towards the EU and EU enlargement. Topics: life satisfaction; subjective rating of the development of the general life situation, the economic situation, the financial situation of the household, personal job situation, national labor market situation and the personal professional outlook in the coming year; native language; knowledge of foreign languages; frequency of political discussions with friends; self-rated opinion leadership; frequency of news consumption (television, newsp...
Help - Tooltips

Perspective Germany 2005/06
McKinsey & Company, Düsseldorf; Stern, Hamburg; Zweites Deutsches Fernsehen (ZDF), Mainz; (1 more)

Applicant Countries Eurobarometer 00 (pilot study)
European Commission, Brussels DG Education and Culture (EAC-D2, Public Opinion Analysis)
Attitude towards the EU and EU enlargement. Topics: life satisfaction; subjective rating of the development of the general life situation, the economic situation, the financial situation of the household, personal job situation, national labor market situation and the personal professional outlook in the coming year; native language; knowledge of foreign languages; frequency of political discussions with friends; self-rated opinion leadership; frequency of news consumption (television, newsp...
Filter summary

The following filters have been applied to your search.

Topic: medical behavior
Collection years: 1942 - 2018
Country: England Finland
Language of data files: fi

Select a filter to remove it from this search.

Close
Uses Deep Links

Results - Pagination

Consortium of European Social Science Data Archives

Find Social and Economic Research Data

13260 results found in 124ms

Results per page: 30
Sort by: Title (ascending)

Topic
Search topics

Collection years

Perspective Germany 2005/06
McKinsey & Company, Düsseldorf; Stern, Hamburg; Zweites Deutsches Fernsehen (ZDF), Mainz; (1 more)
Results - Sort Order

Consortium of European Social Science Data Archives

Find Social and Economic Research Data

13,260 results found in 124ms

- Topic
  - Search topics
- Collection years

Results per page: 30
Sort by:
- Relevance
- Title (ascending)
- Title (descending)
- Date of collection (ascending)
- Date of collection (descending)

Perspective Germany 2005/06
McKinsey & Company, Düsseldorf; Stern, Hamburg; Zweites Deutsches Fernsehen (ZDF), Mainz; (1 more)
Select UI Language
Study title
Effects of Television News in British General Elections, 1997

Creator
Sanders, D., University of Essex. Department of Government
Norris, P., Harvard University. John F. Kennedy School of Government

Study Persistent Identifier
4167 (UKDA)
10.5255/UKDA-SN-4167-1 (datacite)

Abstract
Abstract copyright UK Data Service and data collection copyright owner.

The research used an experimental design to assess the extent to which UK voters' political perceptions are shaped by the content of television news coverage. The research involved:
1. administering a pre-test questionnaire to a randomly selected sample of 1125 respondents;
2. randomly assigning respondents to a series of groups;
3. exposing each group of respondents to a distinctive selection of video news items; and
4. administering a post-test questionnaire to each respondent.

The purpose of the research was to establish the extent to which changes between pre-test and post-test responses varied according to the type of video footage that respondents had seen.
Three core sets of hypotheses were tested:
Similar Results

Study title
Gallup Ecclesiastica 2007

Creator
Church Research Institute

Study Persistent Identifier
FSD2710 (FSD)
urn:nbn:fi:fsd:T-FSD2710 (The National Library of Finland)
Abstract copyright UK Data Service and data collection copyright owner. The NOP National Political Surveys are conducted by NOP on a regular basis. These surveys are designed principally to ascertain public opinion on political parties, leaders and government, and to record voting intention. Data of this nature are given for each survey. In addition, the majority of surveys present data of topical interest and of social importance. Variables covered in the past include: Voting intention; Satisfaction with the Government, the Prime Minister the Leader of the Opposition, the political system and institutions; Unopinions on problems facing Government, the political parties, voting and electoral Behaviour, foreign affairs and defence, economic and industrial affairs, social issues (e.g. welfare, education, health, housing, environment), law, crime and justice, religious beliefs, role of women, race relations, attitudes to pornography, birth control, family planning, marriage, abortion, drugs, gambling, leisure, special issues and events. The Archive also holds several merged files of NOP data and some ad hoc surveys. These are listed separately. Should readers desire to pursue the information presented here in greater depth, they are referred to the <i>Political, social and economic review</i>. The Archive does not, however, necessarily hold the data referred to in the Review. It will endeavour to obtain such data as required, but cannot guarantee success. For more information on the actual data held by the Archive, and the subjects covered, write to the User Enquiry Section, Data Archive, "classifications": [{"vocab": ""}, "vocabUri": ""}, "id": ""}, "term": "Regular opinion polls"}], "keywords": [{"vocab": ""}, "vocabUri": ""}, "id": ""}, "term": "Repeated cross-sectional study"}], "studyAreaCountries": [{"abbr": "not available", "country": "Great Britain"}], "unitTypes": [{"vocab": ""}, "vocabUri": ""}, "id": ""}, "term": "Individuals"}], "publisher": {"abbr": "not available", "publisher": "UK Data Service"}, "publicationYear": "1981-01-01T00:00:00+02:00", "pidStudies": [{"agency": "UKDA", "pid": "409"}, {"agency": "datcite", "pid": "10.5255/UKDA-EN-409-1"}], "fileLanguages": [{"creator": "[NOP Market Research Limited]"}, {"samplingProcedureFreeTexts": "{"Two-stage stratified, area cluster sample. For details see the National Opinion Polls publication <i>Political, social and economic review</i>, 30, April 1981."}, {"nodeIDCollectionStarts": [{"vocab": ""}, "vocabUri": ""}, "id": ""}, {"term": "Opinion polls"}], "nodeIDCollectionEnds": [{"vocab": ""}, "vocabUri": ""}, "id": ""}, {"term": "Opinion polls"}], "dateCollectionPeriodStartDate": "1974-07-02T00:00:02Z", "dateCollectionPeriodEndDate": "1974-07-02T00:00:00+02:00", "dateAccessFreeTexts": "{"The depositor has specified that registration is required and standard conditions of use apply. The depositor may be informed about usage. See <a href="http://ukdataservice.ac.uk/get-data/how-to-access/conditions.aspx#terms and conditions of access</a> for further information.",\"lastModified\": "2018-04-22T14:13:18Z","isActive\": true,\"id\": "UK-Data-Service__409",\"studyUri\": "http://doi.org/10.5255/UKDA-EN-409-1"}
Metadata Issues

Many and various, including:

- Leading and trailing spaces or non-printing characters
- Capitalisation
- Missing/incorrect language tags
- Date and time formats
- Alternative accepted spellings (e.g. Surinam, Suriname*)

*State in South America, formally referred to as Republic of Suriname
Technical Details

- Java-based
- Spring Boot framework
- Elastic search
- UI - Javascript & Searchkit
- Docker containers running in Kubernetes clusters on Google Cloud Platform
Technical Details

- Uses CESSDA’s Open Source Metadata Harvester
- Supports CESSDA Metadata Model v1
- Repository Handlers
  - adapter for OAI-PMH (DDI XML 2.5)
  - adapter for NESSTAR (DDI XML 1.2)
Version 1 (December 2017)

Fixed and upgraded Open Source Metadata Harvester

- Crawls endpoints for XML metadata and ingests transformed data into Elasticsearch

Leveraging a Microservice architecture with Spring Boot, Java 8 Stream for a declarative sets of business instructions:

- Harvester Microservice: Sole responsibility is to harvest and store XML metadata
Indexer Microservice: An ETL tool responsible for extraction of harvested metadata, transformation with named-entity recognition (NER) enrichment and ingestion into Elasticsearch.

Spring Boot Admin App: To monitor and manage the above Microservice’s logs, configuration, JMX exported beans, fail fast feature toggles, etc.

Search UI: Worked with a Frontend developer to customise SearchKit for FullText Search with advanced search query and result filtering.
Version 2 (June 2018)

Similar functionality to Version 1 - total rebuild to handle currently and future XML Schemas.

- Adopted a pluggable Microservice architecture style with multiple Repository Handlers, exposed as Restful APIs.
- Harvester responsible for discovering handlers and handing over requests to a given handler to process its XML schema specifics.
- Consumer Service responsible for orchestrating harvesting of any XML Metadata endpoint and ingestion of output into Elasticsearch.
Version 3 (Date TBC)

- More metadata cleaning, transformation, disambiguation
- Support for CESSDA Metadata Model v2
- Harvest from other sources
- Repository Handler adapter for OAI-PMH (DDI XML 3.2)
- Make metadata harvestable by other aggregators

- Identify relationships among all research entities
  publications, data, funding, researchers, organisations and data sources
Thanks for listening

Any Questions?