



**Metadata en  
interoperabiliteit:  
Hoe dan?**

# Programma

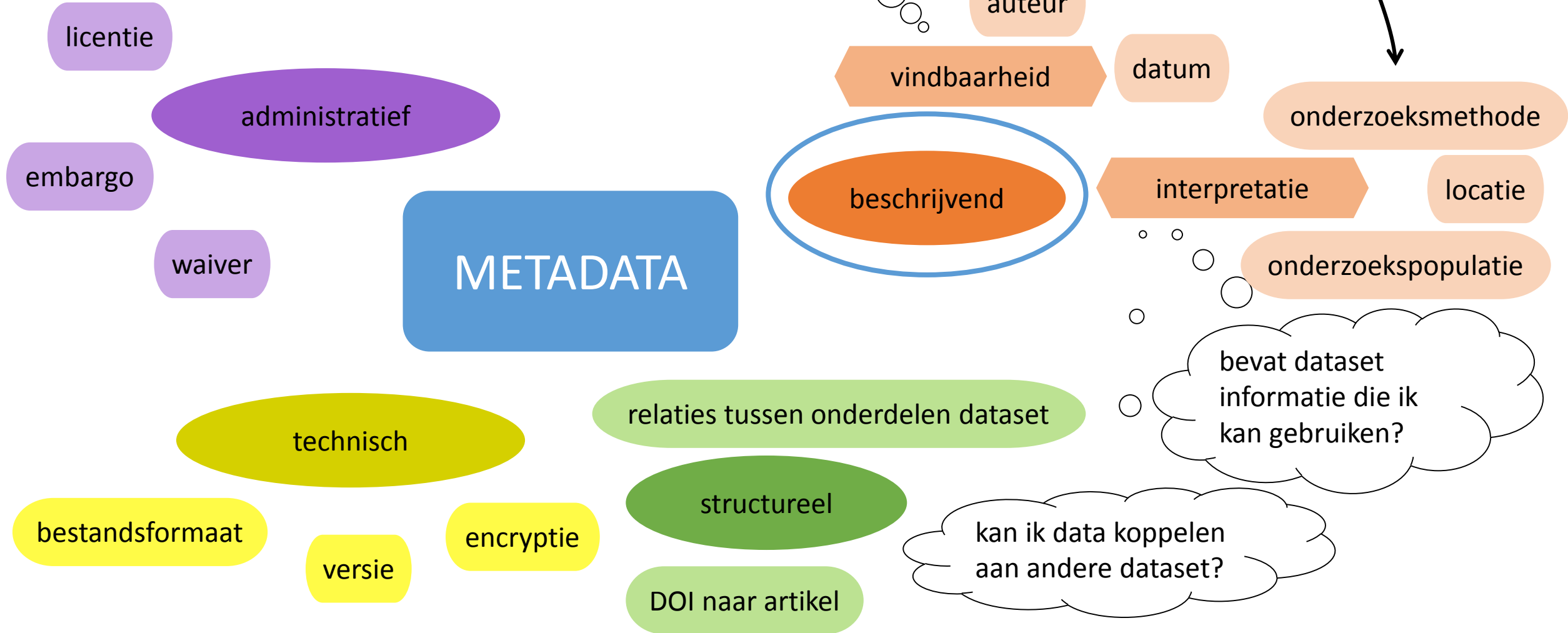
Tijd	Onderwerp	Sprekers
13:00 – 13:30	Opening - Metadata-advies en support vanuit de UBs	Jolien Scholten, Erik Jansen
13:30 – 14:00	Algemene problematiek rond FAIR data	Patrick Aerts (DANS)
14:00 – 14:45:	Metadataschema's in de onderzoekspraktijk	Maurice Vanderfeesten, Jolien Scholten
14:45 – 15:00:	Pauze	
15:00 – 15:30:	Metadata en interoperabiliteit in het GODAN-project en hun toepassingen	Ben Schaap ( <a href="#">GODAN-project</a> )
15:30 – 16:00:	Praktische implementatie van de I in FAIR	Erik Schultes ( <a href="#">GO FAIR</a> )
16:00 – 16.30:	Afsluiting - Discussie: Ontwikkeling, verantwoordelijkheid en ondersteuning	Jolien Scholten, Erik Jansen

# Support UBs (Post-it sessie)

- Vermeld op elke Post-it de naam van je **universiteit of organisatie** (afkorting);
- Noteer op de Post-its de verschillende **manieren** waarop jullie als UB **adviseren en/of ondersteuning** bieden met betrekking tot metadatatieren en interoperabel maken van gegevens;
- Plak de post-its onder de juiste categorie.
  - Informeren via statische communicatiekanalen (websites, flyers)
  - Groepsonderwijs
  - 1-op-1 ondersteuning
  - ?



# Typen metadata

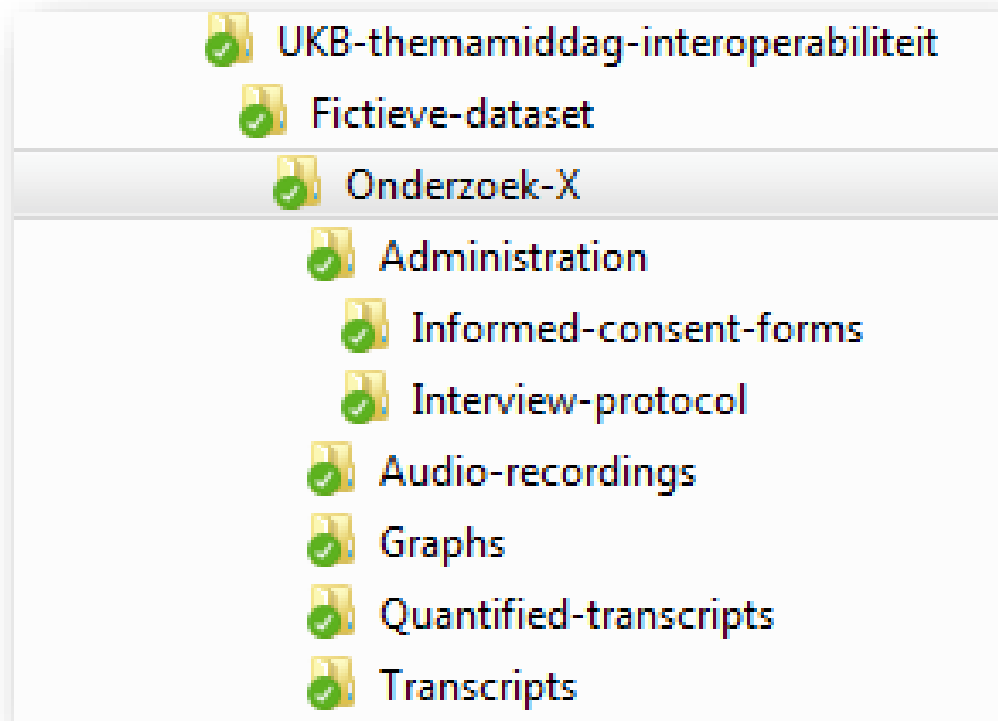


(gebaseerd op cursus [Essentials 4 Data Support](#))





Data assets :  
folder structure based on data assets





# Data assets

Description	Type	Form at	Sensitivity	Size estimate	Storage (during)	Archive (after)
audio-recordings	Audio files	mp3				
transcripts	Text files	txt				
quantified-transcripts	spreadsheet	xlsx				
graphs	illustrations	png				
interview-protocol	Text files	docx				
informed-consent-forms	scans	pdf				

# Data assets

Description	Type	Form at	Sensitivity	Size estimate	Storage (during)	Archive (after)
audio-recordings	Audio files	mp3	Privacy	5,5GB*		
transcripts	Text files	txt	Privacy	10MB		
quantified-transcripts	spreadsheet	xlsx	None	1MB		
graphs	illustrations	png	None	10MB		
interview-protocol	Text files	docx	None	1MB		
informed-consent-forms	scans	pdf	Privacy	500MB		

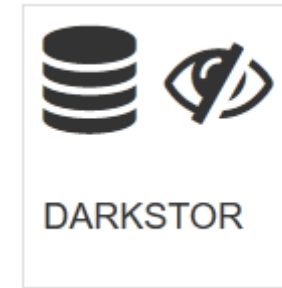
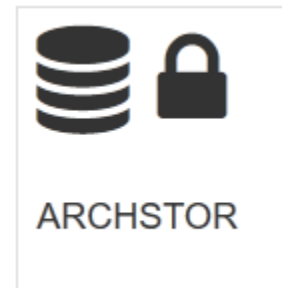
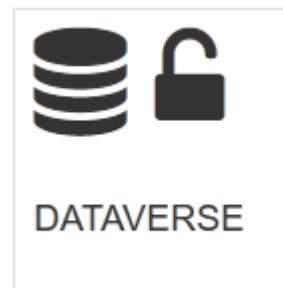
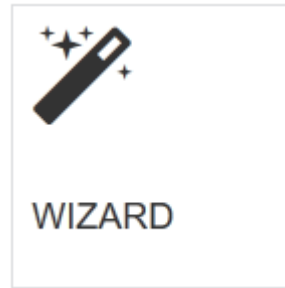
\*) 100 recording of approx . 1 hour, in mp3; compression: 128mbps, 16bit 44000Hz : 55MB pr recording

# Data assets



Description	Type	Form at	Sensitivity	Size estimate	Storage (during)	Archive (after)
audio-recordings	Audio files	mp3	Privacy	5,5GB	SURFdrive + cryptomator	
transcripts	Text files	txt	Privacy	10MB	SURFdrive + cryptomator	
quantified-transcripts	spreadsheet	xslx	None	1MB	SURFdrive	
graphs	illustrations	png	None	10MB	SURFdrive	
interview-protocol	Text files	docx	None	1MB	SURFdrive	
informed-consent-forms	scans	pdf	Privacy	500MB	SURFdrive + cryptomator	

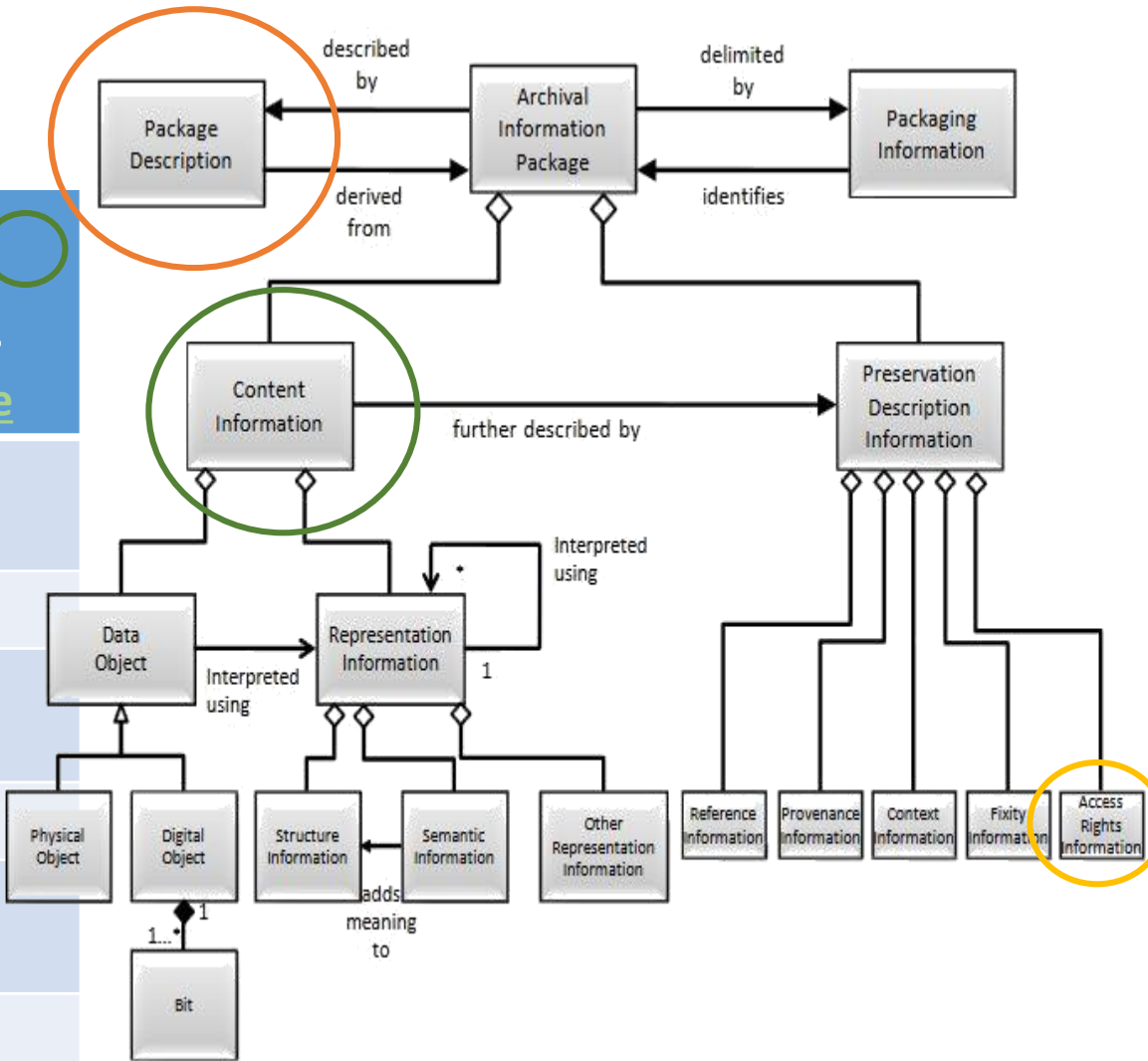
# Data assets



Description	Type	Form at	Sensitivity	Size estimate	Storage (during)	Archive (after)
audio-recordings	Audio files	mp3	Privacy	5,5GB	SURFdrive + cryptomator	DarkStor
transcripts	Text files	txt	Privacy	10MB	SURFdrive + cryptomator	DarkStor
quantified-transcripts	spreadsheet	xslx	None	1MB	SURFdrive	DarkStor + ICPSR
graphs	illustrations	png	None	10MB	SURFdrive	DarkStor + ICPSR
interview-protocol	Text files	docx	None	1MB	SURFdrive	DarkStor + ICPSR
informed-consent-forms	scans	pdf	Privacy	500MB	SURFdrive + cryptomator	DarkStor

# Data assets

Description	Archive (after)	Metadata: Package description > <u>Findable</u>	Metadata: Content information > <u>Interoperable</u>
audio-recordings	DarkStor	DC	Codebook*
transcripts	DarkStor		Codebook*
quantified-transcripts	DarkStor + ICPSR		DDI
graphs	DarkStor + ICPSR		Codebook*
interview-protocol	DarkStor + ICPSR		DDI
informed-consent-forms	DarkStor		Codebook*



\*) includes DC elements

# Metadata Archiving > Generic Repo> DarkStor

## DarkStor (< 50GB)

Vul uw datasetgegevens in. Deze gegevens worden voor VU-onderzoekers opgenomen in Metis, tezamen met uw publicaties. Als uw gegevens zijn ingevoerd klikt u op "Ga door naar gebruiksvoorwaarden".

Uw naam

VUnetID

Faculteit / instelling \*

Afdeling

Promotor / coördinator

Titel van de dataset \*

Extra opmerking

€2.00 per GB per jaar

## package-description.txt - Notepad

```
File Edit Format View Help
dc:title      (dataset belonging to project x)
dc:description (description)
dc:creator    (uploader name, vunet-id)
dc:date
dc:contributor (promotor / coordinator, department)
dc:publisher  (University, Faculty)
```

File Explorer window showing the directory structure of a dataset. The path is: Shared > UKB-themamiddag-interoperabiliteit > Fictieve-dataset > Onderzoek-X.

The left pane shows the directory tree with "Onderzoek-X" selected.

Name	Date modified	Type	Size
Administration	27-6-2018 14:49	File folder	
Audio-recordings	27-6-2018 14:26	File folder	
Graphs	27-6-2018 14:48	File folder	
Quantified-transcripts	27-6-2018 14:27	File folder	
Transcripts	27-6-2018 14:27	File folder	
package-description.txt	27-6-2018 16:01	Text Document	1 KB

# Metadata Archiving > Subject Specific Repo > ICPSR

## ICPSR Find & Analyze Data

FIND DATA SEARCH/COMPARE VARIABLES DATA-RELATED PUBLICATIONS RESOURCES FOR STUDENTS

### Log into ICPSR

Use MyData


Email:

Password:

[Request a new password](#)

You must set your browser to accept cookies in order for login to succeed. Your login session will last for nine hours (or until you log out).

### New User?



[Create Account](#)

By creating a MyData account, you can:

- Download data and use online analysis tools
- Download data available only to ICPSR members, if you are a member
- Generate utilization reports and download statistics
- Update your account profile and settings
- Access your past searches, downloads, and deposits (forthcoming)

### Type of data

Quantitative  
 Geospatial  
 Qualitative  
 Mixed Method

Please describe any methods used to de-identify these data.

### Mode of data collection (select all that apply)

Audio computer-assisted self interview (ACASI)  
 Coded video observation  
 Coded on-site observation  
 Cognitive assessment test  
 Computer-assisted personal interview (CAPI)  
 Computer-assisted self interview (CASI)  
 Computer-assisted telephone interview (CATI)  
 Face-to-face interview  
 Mail questionnaire  
 Mixed mode  
 On-site questionnaire  
 Paper and pencil interview (PAPI)  
 Record abstracts  
 Self-enumerated questionnaire  
 Telephone interview  
 Web-based survey  
 Other

### Universe

The group of persons or other elements that are the object of the study

### Geographic coverage area(s)

Geographic area(s) to which data are relevant (e.g., New York City, Singapore, United States, Springfield, Ohio).

### Geographic Unit

The smallest geographic unit represented in the data that can be analyzed.

Census tract  
 State  
 Precinct  
 County  
 Other

### Is the data collection one of a series?

Yes  
 No

### Type of data collection (select all that apply)

records data  
 location data  
 interview data  
 audio data  
 video data  
 text data  
 image data  
 code data  
 other data

by the data collection (include range(s) of dates).

are collected (provide specific dates and ranges, e.g., month/day/year-month/day/year).

# Example DDI Codebook: survey metadata

```
▼<ddi:codeBook xmlns:ddi="ddi:codebook:2_5" version="2.5" ID="ANES_1960_TS" elementVersion="1.0.0">
  ▼<ddi:docDscr>
    ▼<ddi:citation>
      ▼<ddi:titlStmt>
        <ddi:titl>NES1960.SAS</ddi:titl>
      </ddi:titlStmt>
      ▼<ddi:prodStmt>
        <ddi:prodDate date="2013-07-23 00:35:39">2013-07-23 00:35:39</ddi:prodDate>
        <ddi:software version="1.0.0">DataForge:SledgeHammer</ddi:software>
      </ddi:prodStmt>
    </ddi:citation>
  </ddi:docDscr>
  ▼<ddi:stdyDscr source="producer">
    ▼<ddi:citation source="producer">
      ▼<ddi:titlStmt source="producer">
        <ddi:titl source="producer">ANES 1960 Time Series Study</ddi:titl>
        <ddi:altTitl source="archive">American National Election Study, 1960</ddi:altTitl>
        <ddi:IDNo source="archive" agency="ICPSR">7216</ddi:IDNo>
        <ddi:IDNo source="archive" agency="CrossRef">10.3886/ICPSR07216.v3</ddi:IDNo>
      </ddi:titlStmt>
      ▼<ddi:rspStmt source="producer">
        <ddi:AuthEnty source="archive" affiliation="Unknown">Campbell, Angus</ddi:AuthEnty>
        <ddi:AuthEnty source="archive" affiliation="Unknown">Converse, Philip</ddi:AuthEnty>
      </ddi:rspStmt>
      ▼<ddi:othId source="producer">
        THE SURVEY RESEARCH CENTER 1960 AMERICAN NATIONAL ELEC- TION STUDY WAS DIRECTED BY ANGUS CAMPBELL, PHILI
        DONALD STOKES. CONDUCTED IN SEPTEMBER THROUGH DECEMBER OF 1960, THE STUDY WAS DESIGNED TO PROVIDE INFORM
        AND BEHAVIOR OF A REPRESENTATIVE CROSS-SECTION OF UNITED STATES VOTERS. THE STUDY IS THE LAST WAVE OF A |
```



Q. 1, 6. WE'RE INTERESTED IN THIS INTERVIEW IN FINDING OUT WHETHER PEOPLE PAID MUCH ATTENTION TO THE ELECTION CAMPAIGN THIS YEAR. TAKE NEWSPAPERS FOR INSTANCE -- DID YOU READ ABOUT THE CAMPAIGN IN ANY NEWSPAPER. (IF YES) Q. 6. HOW MUCH DID YOU READ NEWSPAPER ARTICLES ABOUT THE ELECTION -- REGULARLY, OFTEN, FROM TIME TO TIME, OR JUST ONCE IN A GREAT WHILE.

```
</ddi:qstnLit>
</ddi:qstn>
<ddi:sumStat wgt="not-wgt" type="max">5</ddi:sumStat>
<ddi:sumStat wgt="not-wgt" type="min">1</ddi:sumStat>
<ddi:sumStat wgt="not-wgt" type="mean">2.454</ddi:sumStat>
<ddi:sumStat wgt="not-wgt" type="mode">1.0</ddi:sumStat>
<ddi:sumStat wgt="not-wgt" type="stdev">1.575</ddi:sumStat>
<ddi:sumStat wgt="not-wgt" type="invd">100</ddi:sumStat>
<ddi:sumStat wgt="not-wgt" type="vald">1081</ddi:sumStat>
<ddi:catgry>
  <ddi:catValu>1</ddi:catValu>
  <ddi:labl>YES, REGULARLY</ddi:labl>
  <ddi:catStat type="freq" wgt="not-wgt">481</ddi:catStat>
</ddi:catgry>
<ddi:catgry>
  <ddi:catValu>2</ddi:catValu>
  <ddi:labl>YES, OFTEN</ddi:labl>
  <ddi:catStat type="freq" wgt="not-wgt">136</ddi:catStat>
</ddi:catgry>
<ddi:catgry>
  <ddi:catValu>3</ddi:catValu>
  <ddi:labl>YES, FROM TIME TO TIME</ddi:labl>
  <ddi:catStat type="freq" wgt="not-wgt">173</ddi:catStat>
```

# DDI Tool: Colectica

- Free plug-in for Excel
- Exports data to XML, in DDI standard

The screenshot shows the Microsoft Excel interface with the Colectica ribbon active. The ribbon includes options for Document Workbook, Data Documentation, From SPSS, From Stata, Save as DDI, and Create Documentation. The 'Save as DDI' option is highlighted with a red box. Below the ribbon, a data table is visible with columns labeled 'foreign' and 'displace'. The table contains 20 rows of data for a 1978 Automobile dataset.

	E	F	G	H	I	J	K	L
1	headroom	trunk	weight	length	turn	displace	gear_ratio	foreign
2	2.5	11	2930	186	40	121	3.58	0
3	3	11	3350	173	40	258	2.53	0
4	3	12	2640	168	35	121	3.08	0
5	4.5	16	3250	196	40	196	2.93	0
6	4	20	4080	222	43	350	2.41	0
7	4	21	3670	218	43	231	2.73	0
8	3	10	2230	170	34	304	2.87	0
9	2	16	3280	200	42	196	2.93	0
10	3.5	17	3880	207	43	231	2.93	0
11	3.5	13	3400	200	42	231	3.08	0
12	4	20	4330	221	44	425	2.28	0
13	3.5	16	3900	204	43	350	2.19	0
14	3	13	4290	204	45	350	2.24	0
15	2.5	9	2110	163	34	231	2.93	0
16	4	20	3690	212	43	250	2.56	0
17	3.5	17	3180	193	31	200	2.73	0
18	2	16	3220	200	41	200	2.73	0
19	2	7	2750	179	40	151	2.73	0
20	3.5	13	3430	197	43	250	2.56	0

This close-up view of the Colectica ribbon highlights the 'Save as DDI' option, which is circled in red. The ribbon also shows other options like 'Document Workbook', 'Data Documentation', 'From SPSS', 'From Stata', and 'Create Documentation'.

The screenshot shows a file explorer window displaying the output of the DDI export process. The file 'content-information-quantified-transcripts.ddi25.xml' is highlighted, indicating it has been successfully generated. Other files in the folder include 'quantified-transcripts.xlsx' and various system folders like 'Presentaties', 'Pictures', and 'Recycle Bin'.

# Websites

- Colectica: <https://www.colectica.com/software/colecticaforexcel/>
- Inter-university Consortium for Political and Social Research (ICPSR): <https://www.icpsr.umich.edu/icpsrweb/>
- Nestar Publisher: <http://www.nesstar.com/software/download.html>
- Open Archival Information System, zie bijvoorbeeld: [https://www.kb.nl/sites/default/files/docs/sierman\\_oiasmodelned.pdf](https://www.kb.nl/sites/default/files/docs/sierman_oiasmodelned.pdf)
- RightField: <http://www.rightfield.org.uk/>
- XML beautifier: <http://xmlbeautifier.com/>
- Data repositories: <https://www.re3data.org/>
- Cloud Encryption: <https://cryptomator.org/>

# Discussion

