

DDI, Dataverse and Colectica: our data management combo

Alina Danciu, Geneviève Michaud

▶ **To cite this version:**

Alina Danciu, Geneviève Michaud. DDI, Dataverse and Colectica: our data management combo. 12th Annual European DDI User Conference (EDDI20)„ Nov 2020, En ligne, Paris, France. 10.5281/zenodo.4299083 . hal-03956404

HAL Id: hal-03956404
https://sciencespo.hal.science/hal-03956404

Submitted on 25 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

SciencesPo
CENTRE DE DONNÉES SOCIO-POLITIQUES

DDI, Dataverse and Colectica: our data management combo

Alina DANCIU, Geneviève MICHAUD
Center for Socio-Political Data

EDDI 2020, 12th Annual European DDI Users Conference
Virtual event
December, 1, 2020

Creative Commons Attribution 4.0 International (CC BY 4.0)

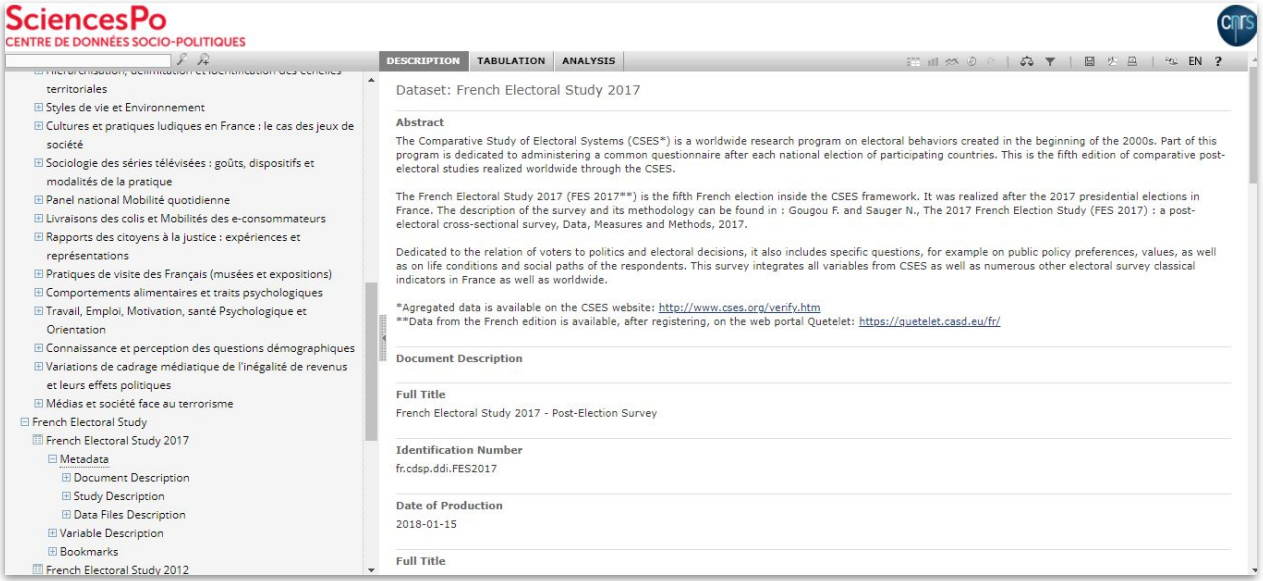
A quick look on...

Going from Nesstar to Dataverse (study-level metadata)

Migrating DDI 1.2 to DDI 2.5

Using Colectica for variable management

Previous metadata & data management platforms



The reason why



cessda [Metadata Model](#) (CMM)

The Dataverse® Project

colectica

DDI xml files migration

1st step - a R script to transform Nesstar DDI 1.2 files to DDI 2.5 files

The R script is available on CDSP's Github.

2nd step - a Python custom program to:

- Import DDI 2.5 files into Dataverse (study level). During this step, a PID (DOI) was given to each study.
- The corresponding DOI was then inserted back in DDI 2.5 file.
- The DDI 2.5 file is uploaded into Dataverse.

Using the CMM controlled vocabularies

Dataverse property	Controlled vocabulary
Keywords	ELSST thesaurus
Kind_of_Data	DDI Alliance Controlled Vocabulary for General Data Format
Time_Method	DDI Alliance Controlled Vocabulary for Time Method
Analysis_Unit	DDI Alliance Controlled Vocabulary for Analysis Unit
Mode-Collection	DDI Alliance Controlled Vocabulary for Mode Of Collection
Sampling-Procedure	DDI Alliance Controlled Vocabulary for Sampling Procedure
Topic_Classification	CESSDA Topic Classification
Type_Instrument	DDI Alliance Controlled Vocabulary for Type of Instrument

1 tailor-made VC for institutions(using ROR and ISNI)

Benefits of using Dataverse

Dataverse is a good solution for networked institutions:

- Single Sign On using institutional credentials
- Can harvest and be harvested
- General and Dataverse-level branding capabilities

Dataverse supports discoverability and persistence:

- Powerful search engine and filters
- Good Google referencing
- DOI

Pending issues with Dataverse



DDI import and export definitively are a hard route to follow.

Adding a language creates additional difficulties.

(see : Customizing a DDI compliant repository: a detailed course with Dataverse, session Managing DDI)

Icon source

We are getting there



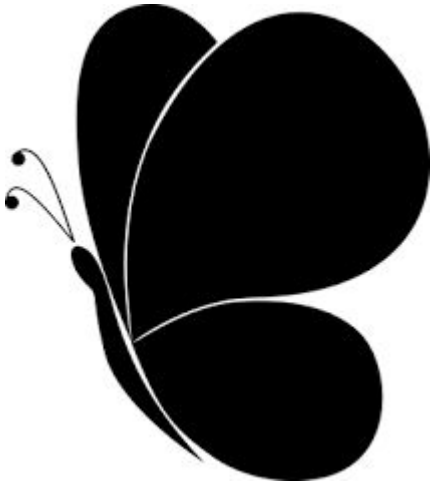
6. Integrating Colectica in our toolbox



- For **variable-level metadata**, we continue to use Nesstar Publisher and will test R scripts
- We have also been using **Colectica** for a year now, for **longitudinal studies**
- We use Colectica Designer, Colectica Portal and Colectica Repository
- The pending question : “single” studies

Photo source

7. What's next



Improving data.sciencespo.fr

Testing **Dataverse variable visualisation tools**

Setting up n**ew data and metadata management and publication workflows**

cc @Pixabay

Get in touch!

alina.danciu@sciencespo.fr

genevieve.michaud@sciencespo.fr



@alina_danciu_
@GenevievMichaud
@CDSP_ScPo

Thank you

DOI: 10.5281/zenodo.4299083