

# DDI @ CESSDA

---

*Carsten Thiel | CESSDA ERIC*

*Dagstuhl, Sep 20-24, 2021*

 [cessda.eu](https://cessda.eu)

 [@CESSDA\\_Data](https://twitter.com/CESSDA_Data)



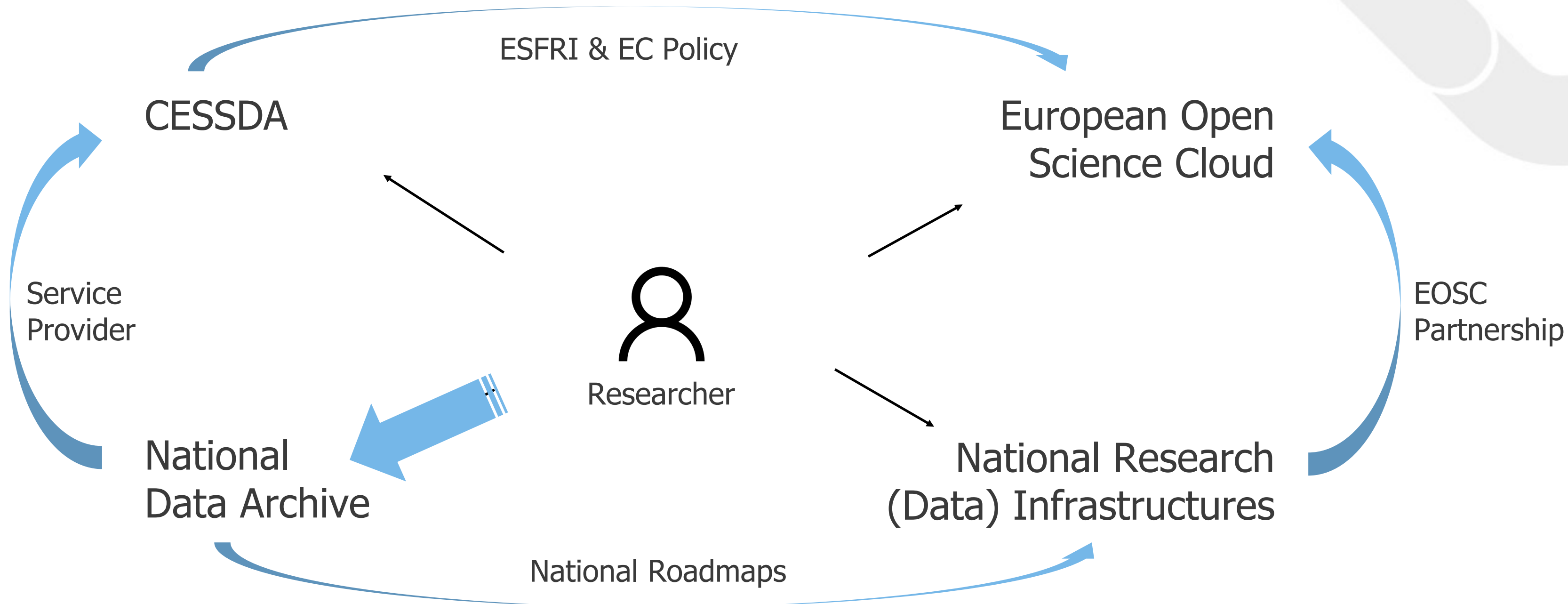
Licence: CC-BY 4.0



# CESSDA

- ◇ Consortium of European Social Science Data Archives  
– European Research Infrastructure Consortium
- ◇ 22 members and 1 observer
- ◇ ⇒ Many European languages
- ◇ ⇒ Some with decades of history, some just starting up
  
- ◇ The Social Sciences' gateway to the European Open Science Cloud

# CESSDA in EOSC



# Metadata Exchange

# Internal Challenges

- ◇ Metadata has been collected since the 1970s (at some SPs)
- ◇ Inconsistent interpretation of DDI 2.5 elements
- ◇ Different technologies and solutions
  - ◇ Nesstar
  - ◇ Dataverse
  - ◇ Colectica
  - ◇ Custom or self-developed solutions
- ◇ ⇒ Very hard to have a common baseline

# DDI Profiles

- ◇ Machine-readable specification
- ◇ Available for CDC (DDI 1.2.2, DDI 2.5 & DDI 3.2) and EQB (DDI 2.5 & DDI 3.2)
- ◇ CESSDA Metadata Validator to check compliance
- ◇ <https://cmv.CESSDA.eu/>
- ◇ <https://cmv.CESSDA.eu/profiles/cdc/ddi-3.2/0.1.1/profile.xml>

- ◇ From 2022: CDC harvestable by EOSC

Consortium of European Social Science Data Archives

CESSDA  
MV Metadata Validator

Documentation API

Configuration

Validation Gate: BASIC

Profile:  BY\_PREDEFINED  BY\_URL  BY\_UPLOAD

Select resource

Documents:  BY\_PREDEFINED  BY\_URL  BY\_UPLOAD

Browse... No files selected.

Validate

# Continuing Challenges

- ◇ Metadata licenses unclear
- ◇ Metadata clean-up requires manual work
- ◇ Identification of studies across SP catalogue records
- ◇ Where do the researchers look for data?
- ◇ What else to present and how? Variables: across studies and countries?
- ◇ *And I won't even mention actual access!*

# New Requirements

- ◇ EOSC compatibility
  - ◇ References (publications, authors/ORCIDIDs) and funding
  - ◇ FAIR
  - ◇ CTS status
- ◇ User experience improvements
  - ◇ Curation level: data steward and self-deposit
  - ◇ License CVs
  - ◇ Access condition CVs
  - ◇ Vocabularies and thesauri by reference
- ◇ Semantic interoperability

⇒ Codebook 2.5 is not enough



What to do?

# Considerations

- ◇ SPs have Codebook, Lifecycle not commonly used
- ◇ Codebook is much more suited to what we have at CESSDA level
- ◇ EOSC-interoperability is limited with both Codebook and Lifecycle
  - ◇ OpenAIRE or DCAT more viable options there
- ◇ Investment at all SPs necessary, independent of decision

# Needs and requirements from DDI

- ◇ New and improved standards – Codebook 2.6?
- ◇ Concrete implementation examples, not only abstract models
  - ◇ E.g. full schema.org JSON-LD markup

# Thank you

---

*carsten.thiel@cessda.eu*

 [cessda.eu](https://cessda.eu)

 @CESSDA\_Data



Licence: CC-BY 4.0

