Facilitating FAIRness with a Stockholm University machine-actionable DMP Online template

Joakim Philipson - Research Data Analyst

(Stockholm University, RDM team)

DMP Online Dropin 2021-10-26

CC BY

National requirement

- As from 2019, all grantees from Swedish Research Council must have a **data management plan** if the research generates research data.
- *Swedish Research Council* (**VR**) & Association of Swedish Higher Education Institutions (**SUHF**) collaborated on partially reworked version of **Science Europe**'s Core Requirements for Data Management Plans – now a template in **DMP Online**, *but only with free text fields*

- **Stockholm University** offers DMP Online tool as default to all our researchers at dmp.su.se via SSO local login

- Other DMP tools: ARGOS, DSW, DMPTool , easyDMP, RDMO - … all more or less converging to RDA DMP Common Standard to produce more *machine-actionable, maDMPs*

Stockholm University

Multiple purpose of SU maDMP template

1. *Easing the administrative burden on researchers* to fill in the DMP by means of preset multiple-choice answer options, drop-down menus, radio buttons etc., including extensive local *Guidance*

2. *Facilitating* semi-automated *review* and *evaluation of FAIRness,* while retrieving needed information for RDM administration

3. *Conforming to the RDA DMP Common Standard* for machine-actionable DMPs by validation against *maDMP-schema-1.1*

This work is also part of **EOSC Nordic wp. 5.3.2** on maDMPs (with other members from Uninett - Norway, SND - Sweden, GoFAIR – Netherlands)

Stockholm University

Features of SU-VR maDMP template

- *Preserves structure* (sections I-VI) *and original questions from* Science Europe – *Swedish Research Council* (**VR**) in headings*,* but with more specific questions and answers (checkboxes, drop-down menus, radio buttons: yes/no/unknown)
- Using DMP Online (Roadmap) API (v0) to get output (as JSON) of *entire* DMPs (questions and answers) for further processing

- Workflow:
  1. APIv0 curl GET *https://dmp.su.se/plans/72835* ->

```
$ curl -G -k "https://dmponline.dcc.ac.uk/api/v0/plans?plan=72835" -H
"Authorization:Token token=XYZ123456789sTuV012345" > DMP72835.json
```

  2. Oxygen: *dmp72385.json**2**dmp72385.xml* conversion ->
  3. *dmp72385.xml**2**dmp72385**transform**.xml* with XSLT [allowing FAIR-evaluation] ->
  4. *dmp72385transform.**xml**2dmp72385transform.**json*** ->
  5. *dmp72385transform.**json*** validated with RDA maDMP

Stockholm University

Challenges creating a maDMP template in DMP Online

1. *Sharing template with others* (FAI***Re-usable***) only through PDF !-<
   [Solution: Dataverse ”private link”:
   https://dataverse.harvard.edu/privateurl.xhtml?token=345aaf2e-4648-4ba3-b93a-7ce3547f9354
   to download HTML of preview to see a ”blank”, editable version of the template]

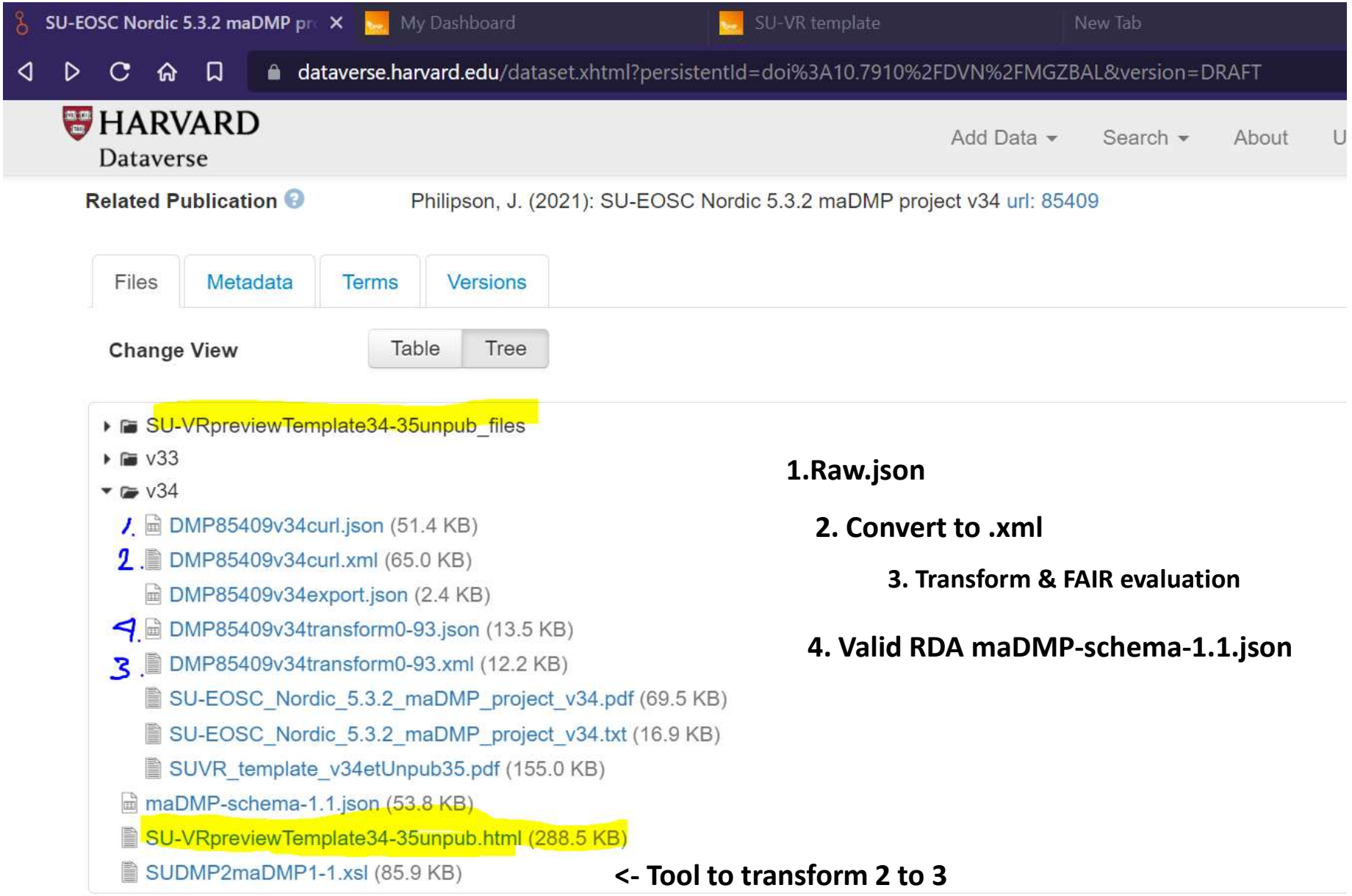
2. *Cover page details (e.g. Funder, Grant ID) not part of API v0 output* (from v34 added separately)*, while API v1 does not give full content of other fields*

3. *Compliance with RDA DMP Common Standard (RDCS) required JSON output conversion to XML, transformed (by XSLT) and reconverted to JSON, to validate against the RDCS madmp-schema-1-1.json*

4. *No pre-ingest validation of datatypes or pattern matching of input strings in template webform text fields -> little control of input type / output received, other than through very clear guidance - not always getting attention necessary to avoid less meaningful answers.*

5. *No repetition of input text fields possible, as required e.g. for dataset identifiers*
   [DCAT model inherent in RDCS].

Stockholm University



1.Raw.json

2. Convert to .xml

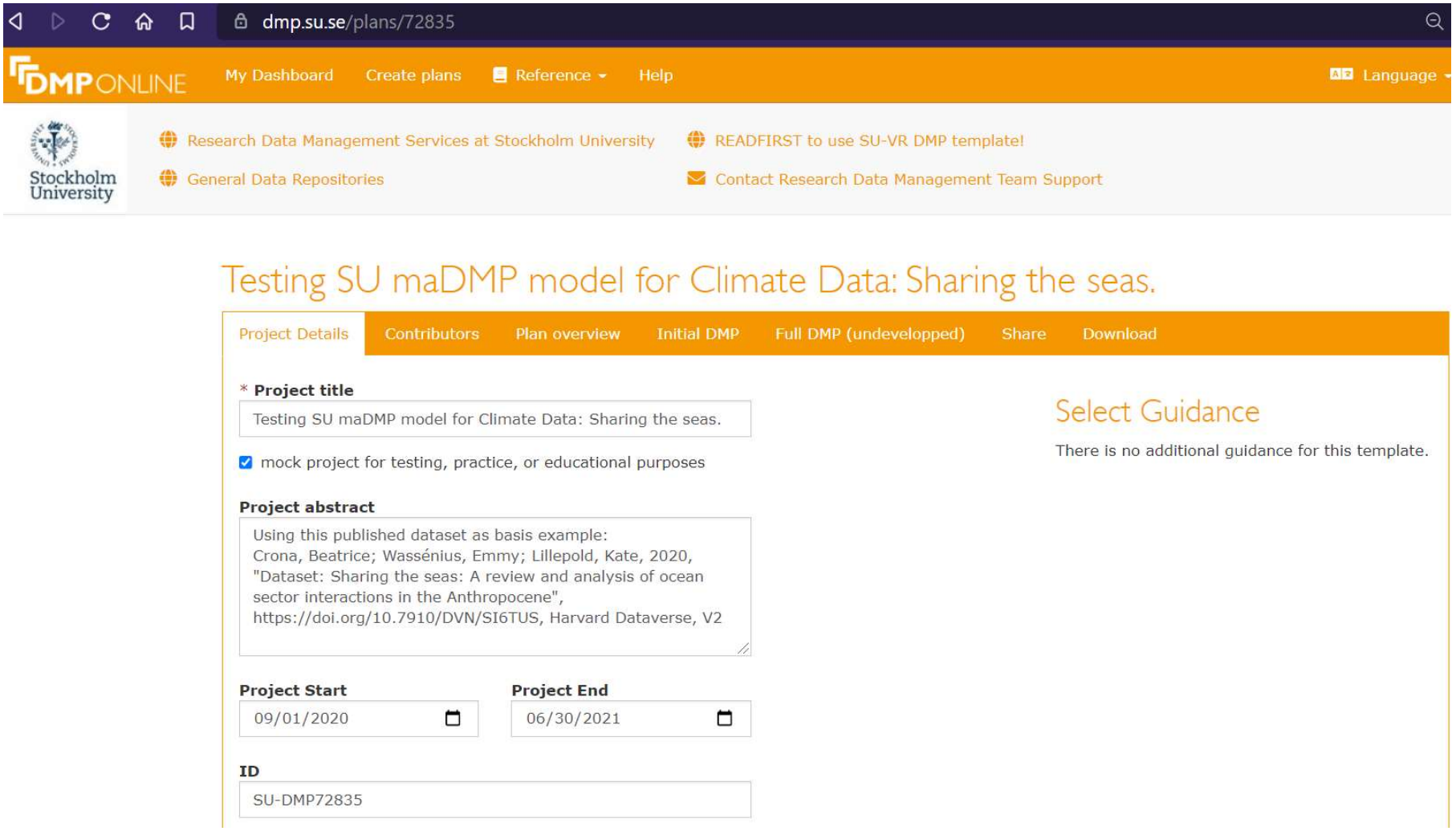
3. Transform & FAIR evaluation

4. Valid RDA maDMP-schema-1.1.json

<- Tool to transform 2 to 3

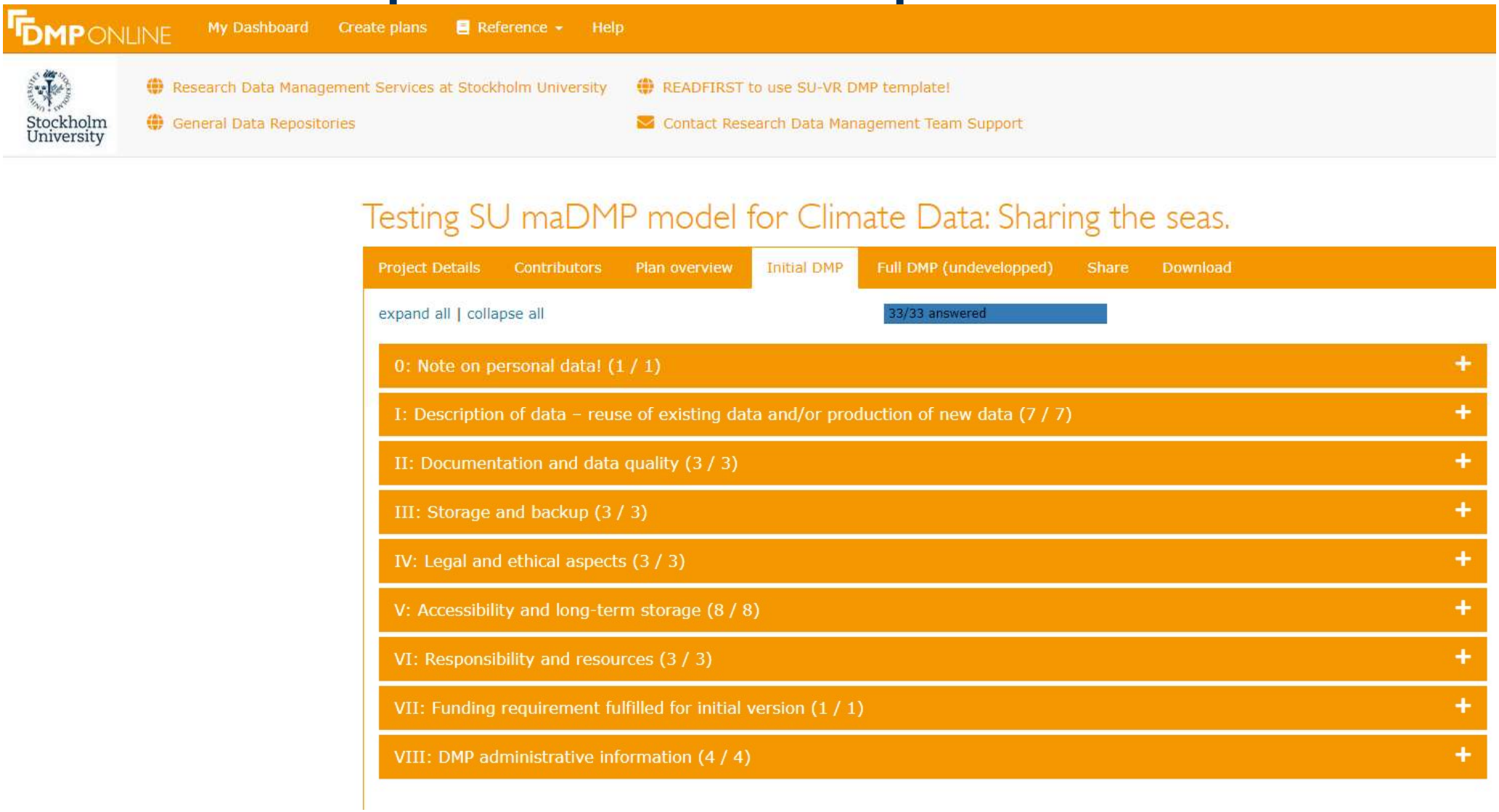
Stockholm University

Test example DMP of SU-VR template

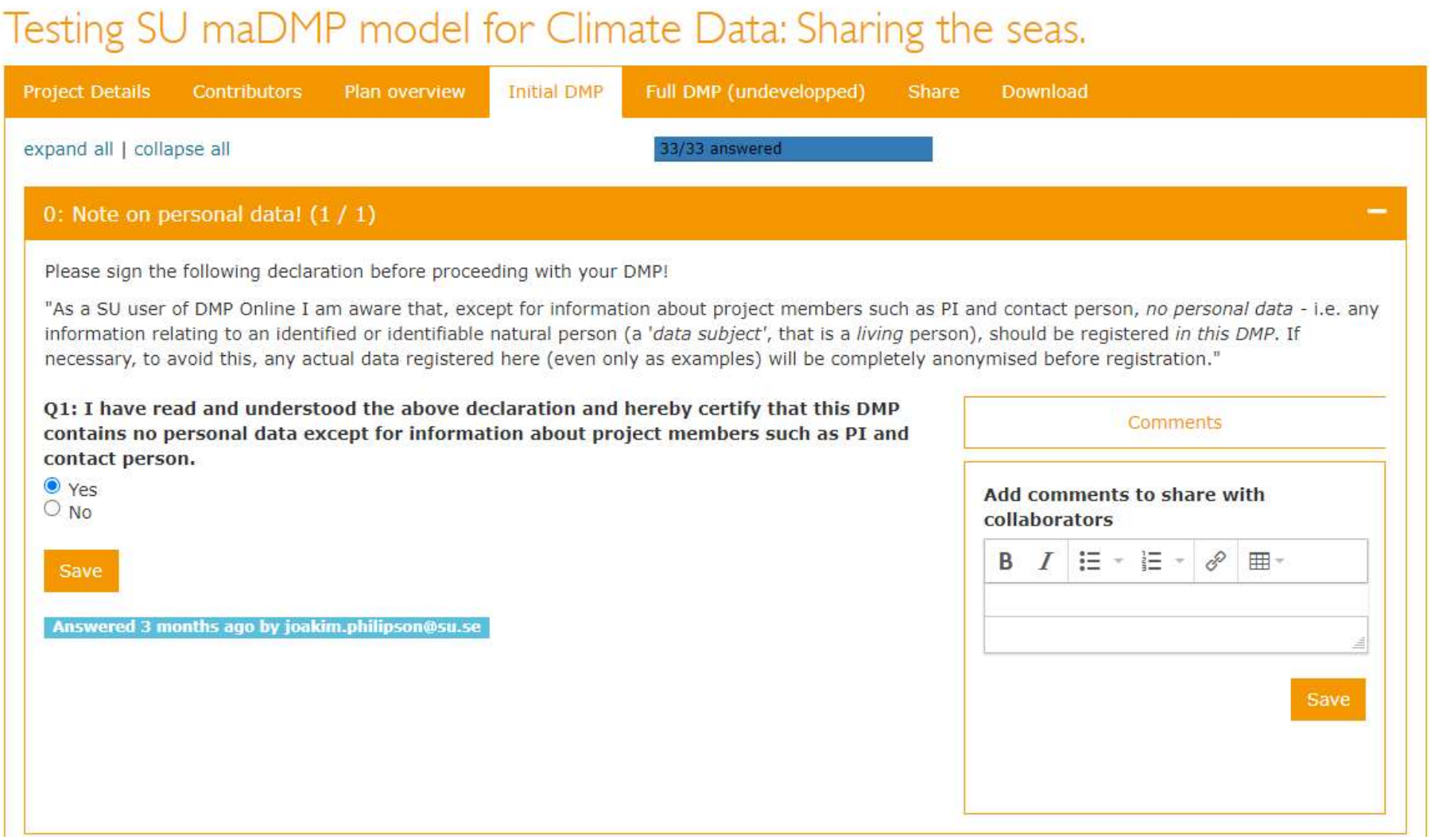


Stockholm University

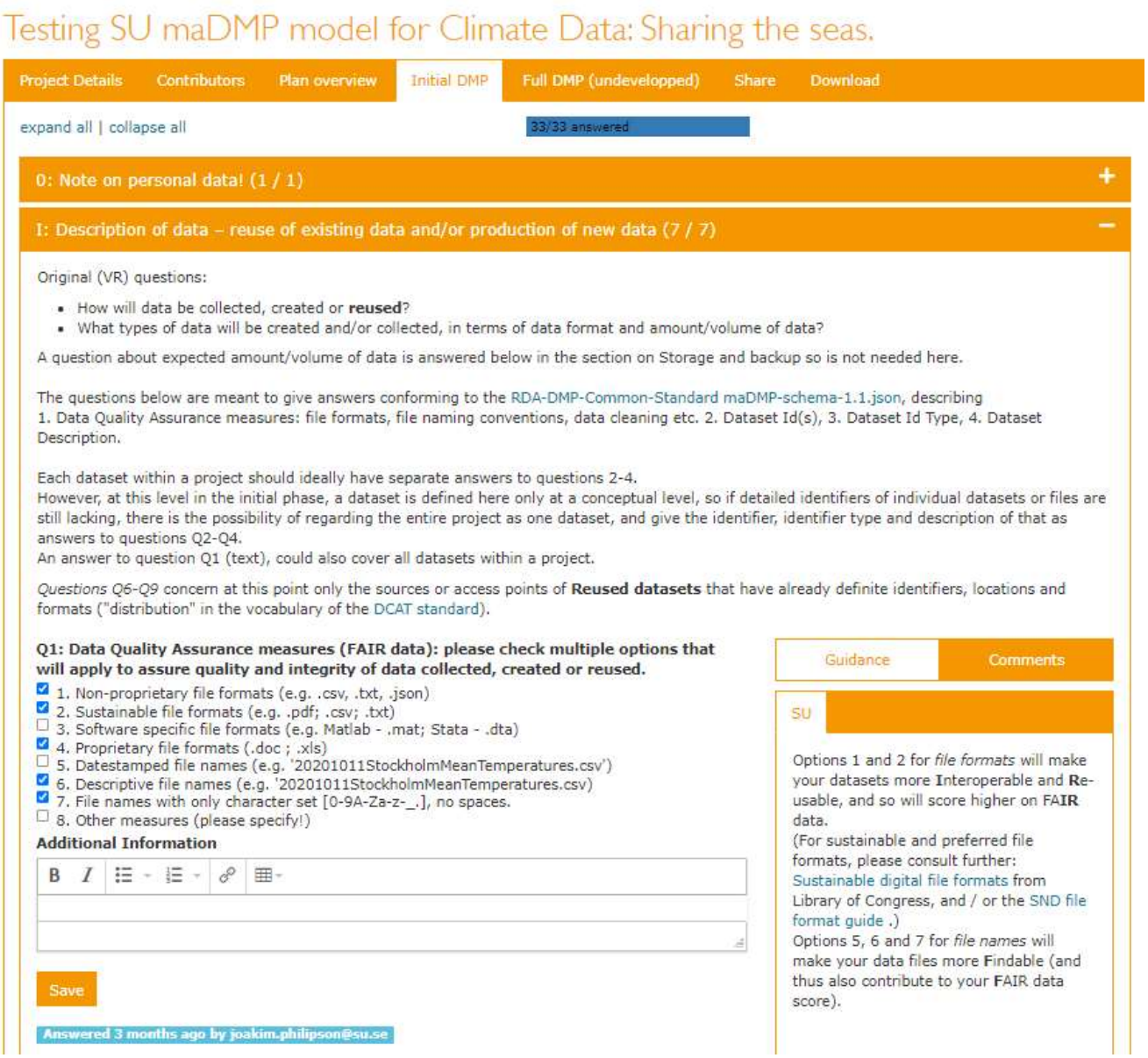
Test example DMP of SU-VR template



Demo DMP example of SU-VR template: section 0

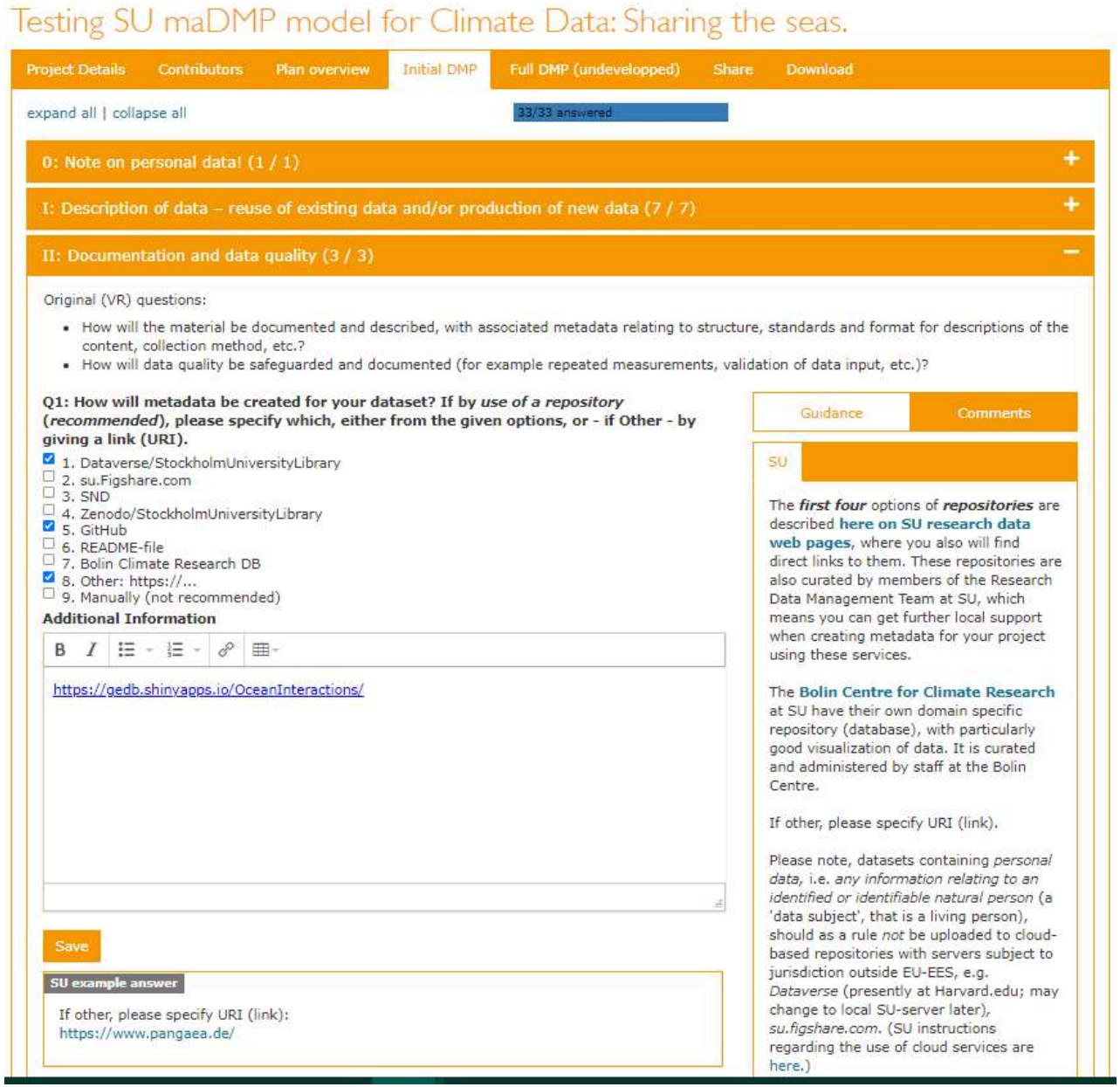


Stockholm University



Demo DMP example of SU-VR template: I-Q1

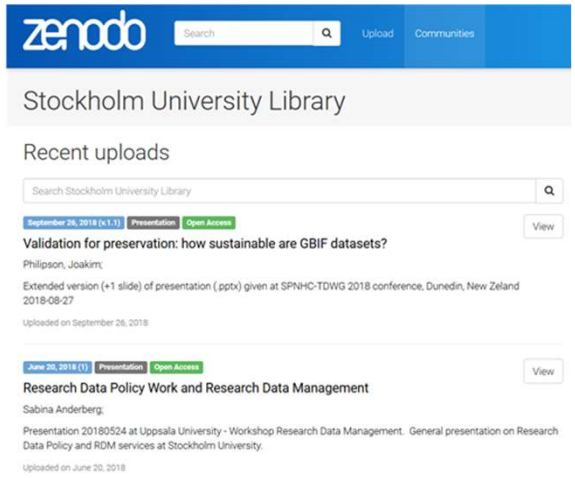
Stockholm University



Demo DMP example of SU-VR template: II-Q1

Stockholm University

Data repositories curated by SU RDM-team



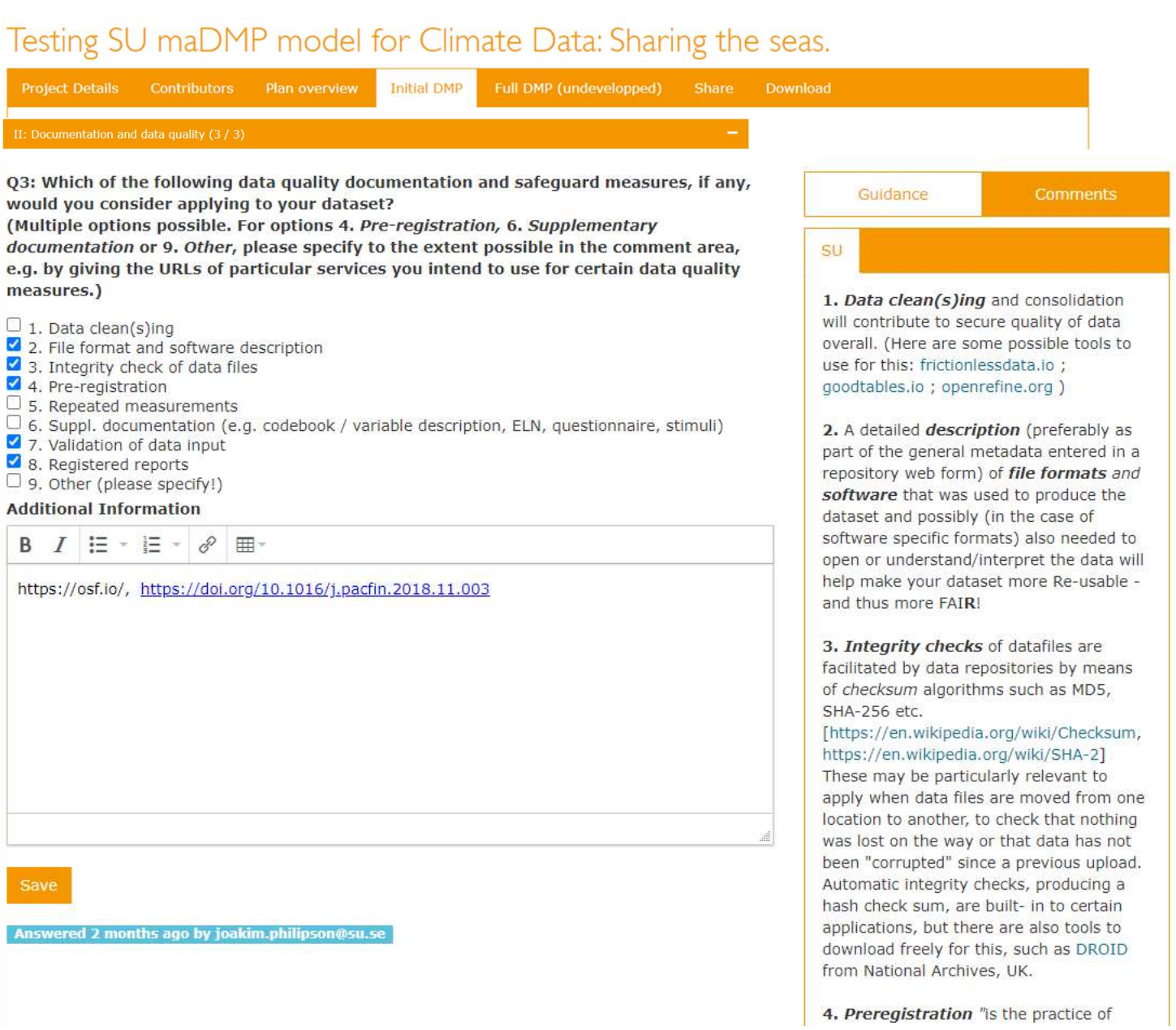
https://zenodo.org/communities/stockholmuniversitylibrary

- **Bolin Centre DB**
- **Dataverse**
- **SND - https://snd.gu.se**

- ✓ **Create standard metadata** (for DMP q & a)
- ✓ Pre-order **DOI** for your dataset (or private link)
- ✓ **Curation** and counseling from SU
- ✓ *Harvest and transform for **archival** (su.*figshare, so far*)*
- ✓ *One **storage** / backup option during project (*embargo*)*?

Stockholm University

Demo DMP example of SU-VR template: II-Q3



Stockholm University

Demo DMP example of SU-VR template IV-Q1

Stockholm University

Demo DMP example of SU-VR template IV-Q3

IV: Legal and ethical aspects (4 / 4)

Original (VR) questions:

- How is data handling according to legal requirements safeguarded, e.g. in terms of handling of personal data, confidentiality and intellectual property rights?
- How is correct data handling according to ethical aspects safeguarded?

Q3: Will the creation, collection or reuse of dataset(s) in your project entail any of the following:

- **(a) processing of *special categories of personal data* according to the General Data Protection Regulation (EU 2016/679), i.e. personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation**
- **(b) processing of personal data regarding violations of law that include crimes, judgments in criminal cases, penal law sanctions, or administrative deprivation of liberty**
- **(c) physical interventions on research subjects or deceased persons**
- **(d) methods with the purpose of affecting a research person physically or mentally, or which includes an apparent risk of injuring the research subject either physically or mentally**
- **(e) studies of biological material that has been taken from a living or deceased person, and can be traced to that person**

Or, further, will the creation, collection or reuse of dataset(s) in your project include:

- **(F) Data from animal research**
- **(G) Data on genetic resources and/or traditional knowledge associated with genetic resources**
- **(H) Data that can be used for military purposes or concerning products that can be used for military purposes**
- **(I) Data that are sensitive in some other respect**

(Possible ethical review documentation applying to any of points (a-e) will be asked for separately.) If you answer yes to any of the points (F-I) and there is already relevant documentation or applications, please provide (a) reference(s) to any application(s)/ approval(s)/decision(s)/document(s), if possible by URL(s) / PID(s) such as DOI(s) giving direct access, or registration no. (Swe. *diarienr.*) in the text field below.

- ○ yes
- ◉ no
- ○ unknown
- ○ yes, but not (a-e)

Additional Information

Guidance | Comments

SU

First list (a.-e.): Answering *yes* to any of the **points a.- e.** above means that you may need to apply for ethical review for research taking place *in Sweden*. See information here. For research taking place in other countries, other rules may apply.

For support concerning *ethical review of research involving humans*, please contact etik@fs.su.se.

Second list (F-I):
F. Animal research, and the use of data from animal research, may require ethical review or other approvals. See information concerning animal research here (information in Swedish). Some information in English is provided here.

For guidance concerning what applies to animal research, please contact etik@fs.su.se.

G. The use of **genetic resources and/or traditional knowledge associated with genetic resources** is regulated in the **Nagoya Protocol**, implemented through the **EU Access and Benefit-Sharing (ABS) Regulation**. Information from the *Swedish Environmental Protection Agency* is found here. For guidance concerning matters related to the Nagoya Protocol and The ABS Regulation, please contact info@fs.su.se.

Stockholm University

DMP example of SU-VR template: API output & transform

DMP72835.json

1.Raw.json

SUDMP2maDMP1-1.xsl | **DMP72835.json** | DMP72835.xml | DMP72835transform.xml | DMP72835transform.json

```
[
  {
    "id": 72835,
    "title": "Testing SU maDMP model for Climate Data: Sharing the seas.",
    "grant_number": null,
    "last_updated": "2021-06-16 14:05:01 UTC",
    "creation_date": "2021-03-04 10:42:54 UTC",
    "test_plan": true,
    "template": {
      "title": "SU-VR template",
      "id": 1456260114
    },
```

2. Convert to .xml

SUDMP2maDMP1-1.xsl | DMP72835.json | **DMP72835.xml** | DMP72835transform.xml | DMP72835transform.json

array

```
<?xml version="1.0" encoding="UTF-8"?>
<array>
    <id>72835</id>
    <title>Testing SU maDMP model for Climate Data: Sharing the seas.</title>
    <grant_number>null</grant_number>
    <last_updated>2021-06-16 14:05:01 UTC</last_updated>
    <creation_date>2021-03-04 10:42:54 UTC</creation_date>
    <test_plan>true</test_plan>
    <template>
        <title>SU-VR template</title>
        <id>1456260114</id>
```

3. Transform & FAIR eval.

SUDMP2maDMP1-1.xsl | DMP72835.json | DMP72835.xml | **DMP72835transform.xml** | DMP72835transform.jsor

dmp description

```
<?xml version="1.0" encoding="UTF-8"?>
<dmp xmlns:xlink="http://www.w3.org/1999/xlink"
     xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
     xmlns:mets="http://www.loc.gov/METS/"
     xmlns:local="local-functions"
     OBJID="72835"
     LABEL="Testing SU maDMP model for Climate Data: Sharing the seas.">
  <title>Testing SU maDMP model for Climate Data: Sharing the seas.</title>
```

SUDMP2maDMP1-1.xsl | DMP72835.json | DMP72835.xml | **DMP72835transform.xml** | DMP72835transform.jsor

dmp description

```
<FAIRscore>
  <sectionI>
    <DataDescription>
      <FileFormatsUNamesQ1>
        <Option>7. File names with only character set [0-9A-Za-z-_.], no
                  spaces.</Option>
        <F-value>1</F-value>
        <Option>6. Descriptive file names (e.g.
                  '20201011StockholmMeanTemperatures.csv) </Option>
        <F-value>1</F-value>
        <Option>4. Proprietary file formats (.doc ; .xls)</Option>
        <Option>1. Non-proprietary file formats (e.g. .csv, .txt, .json) </Opti
        <I-value>1</I-value>
        <R-value>1</R-value>
        <Option>2. Sustainable file formats (e.g. .pdf; .csv; .txt)</Option>
        <I-value>1</I-value>
        <R-value>1</R-value>
      </FileFormatsUNamesQ1>
```

UDMP2maDMP1-1.xsl | DMP72835.json | DMP72835.xml | DMP72835transform.xml | **DMP72835transform.json**

```
{
    "dmp": {
        "xmlns:xlink": "http://www.w3.org/1999/xlink",
        "xmlns:xsi": "http://www.w3.org/2001/XMLSchema-instance",
        "xmlns:mets": "http://www.loc.gov/METS/",
        "xmlns:local": "local-functions",
        "OBJID": 72835,
        "LABEL": "Testing SU maDMP model for Climate Data: Sharing the seas.",
        "title": "Testing SU maDMP model for Climate Data: Sharing the seas.",
        "description": "Using this published dataset as basis example:\r Crona, Beatrice
        "created": "2021-03-04T10:42:54Z",
        "modified": "2021-06-16T14:05:01Z",
        "contact": {
            "mbox": "kl@su.se",
            "name": "Lillepold, Kate",
            "contact_id": {
```

Validate with

URL: /C:/Users/joph9849/Desktop/DMPetPolicy/RDA-DMPvalidation/maDMP-schema-1.1.json

Schema type: JSON

Embedded Schematron rules

Public ID:

Schematron phase:

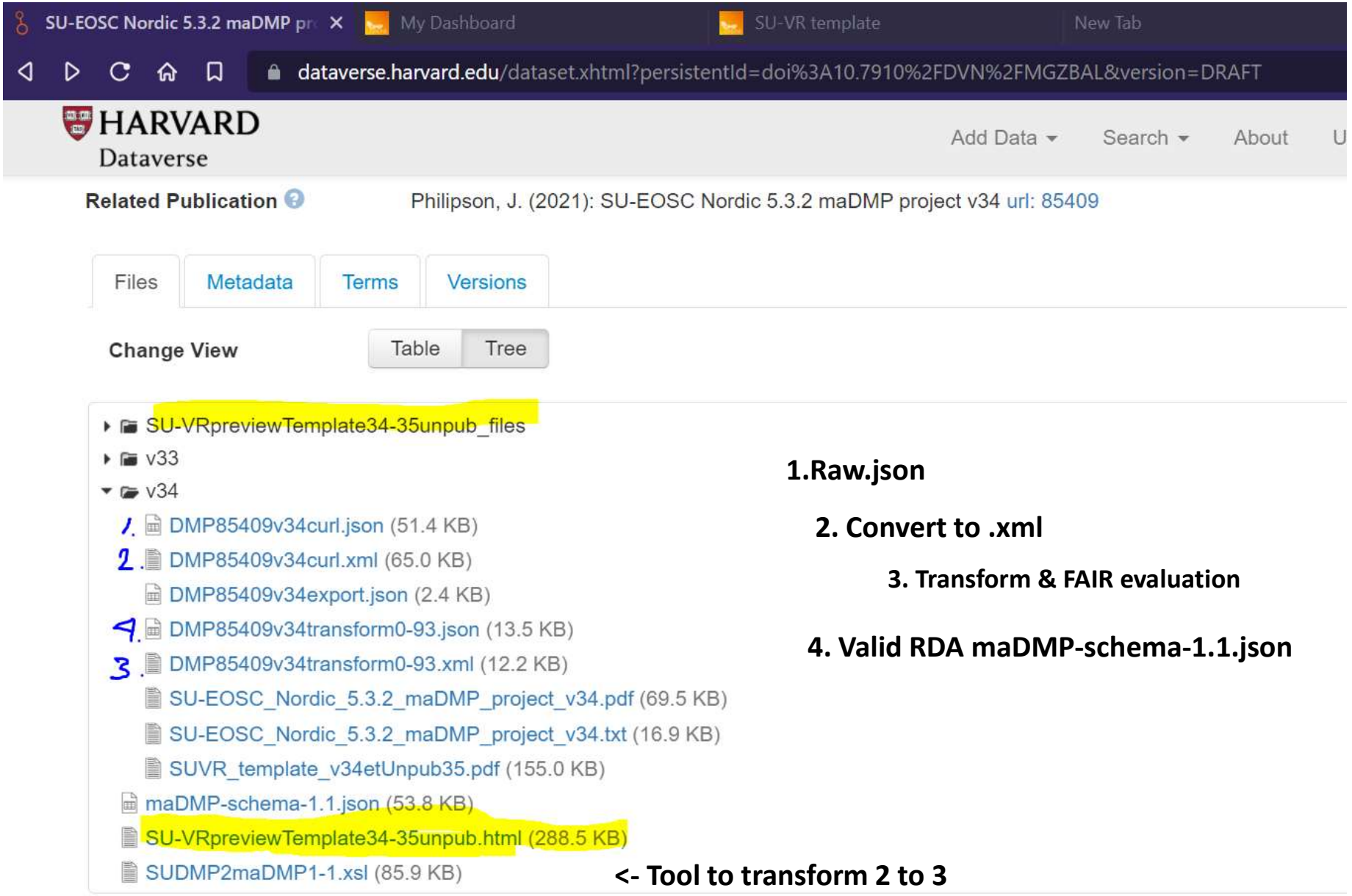
OK | Cancel

Text | Grid

UCL-SUpresent\DMP72835transform.j... Validation successful

4. Valid RDA maDMP-schema-1.1.json

Stockholm University



1.Raw.json

2. Convert to .xml

3. Transform & FAIR evaluation

4. Valid RDA maDMP-schema-1.1.json

<- Tool to transform 2 to 3

Stockholm University

DMP example of SU-VR template API output & transform

2. Convert to .xml

1.Raw.json

DMP85409v34curl.json | DMP85409v34curl.xml | DMP85409v34transfori

```
[
    {
        "id": 85409,
        "title": "SU-EOSC Nordic 5.3.2 maDMP project v34"
        "grant_number": "EOSC-Nordic-857652",
        "last_updated": "2021-10-05 16:05:41 UTC",
        "creation_date": "2021-10-05 13:38:51 UTC",
        "test_plan": false,
        "template": {
            "title": "SU-VR template",
            "id": 1456260114
        },
        "funder": {"name": null},
        "principal_investigator": {
            "name": "Joakim Philipson",
            "email": "joakim.philipson@su.se",
            "phone": ""
        },
        "data_contact": {
            "name": "Joakim Philipson",
            "email": "joakim.philipson@su.se",
            "phone": ""
        },
        "users": [
            {"email":
            {"email":
```

#: expected type: JSONObject, found:

Text Grid

Validation failed.

Validate with

URL: /C:/[redacted]RDA-DMPvalidation/maDMP-schema-1.1.json

Schema type: JSON

Embedded Schematron rules

DMP85409v34curl.json | DMP85409v34curl.xml | DMP85409v34transform0-93.xm

array

```
<?xml version="1.0" encoding="UTF-8"?>
<array>
    <id>85409</id>
    <title>SU-EOSC Nordic 5.3.2 maDMP project v34</title>
    <grant_number>EOSC-Nordic-857652</grant_number>
    <last_updated>2021-10-05 16:05:41 UTC</last_updated>
    <creation_date>2021-10-05 13:38:51 UTC</creation_date>
    <test_plan>false</test_plan>
    <template>
        <title>SU-VR template</title>
        <id>1456260114</id>
    </template>
    <funder>
        <name>null</name>
    </funder>
    <principal_investigator>
        <name>Joakim Philipson</name>
        <email>joakim.philipson@su.se</email>
        <phone/>
    </principal_investigator>
    <data_contact>
        <name>Joakim Philipson</name>
        <email>joakim.philipson@su.se</email>
        <phone/>
    </data_contact>
    <users>
```

Text Grid Author

Stockholm University

DMP example of SU-VR template API output & transform

2. Convert to .xml

DMP85409v34curl.json | **DMP85409v34curl.xml** | DMP85409v34transform0-93.xml | DMP85409v3

array

```
<number>2</number>
<questions>
    <text>Q1: Data Quality Assurance measures (FAIR data): please check mu
        that will apply to assure quality and integrity of data collected,
        reused.</text>
    <number>1</number>
    <format>Check box</format>
    <option_based>true</option_based>
    <answered>true</answered>
    <answer>
        <text/>
        <options>
            <text>7. File names with only character set [0-9A-Za-z-_.], no
                spaces.</text>
        </options>
        <options>
            <text>6. Descriptive file names (e.g.
                &apos;20200910SanFranciscoDaylight1pm.mp4&apos;) </text>
        </options>
        <options>
            <text>5. Datestamped file names (e.g.
                &apos;20201011StockholmMeanTemperatures.csv&apos;)</text>
        </options>
        <options>
            <text>2. Sustainable file formats (e.g. .pdf; .csv; .txt)</tex
        </options>
```

Text | Grid | Author

Stockholm University

DMP example of SU-VR template API output & transform: 3. Transform & FAIR evaluation

DMP85409v34curl.json | DMP85409v34curl.xml | **DMP85409v34transform0-93.xml** | DMP85409v34transform

dmp array project

```
23  <array>
24     <project>
25        <title>SU-EOSC Nordic 5.3.2 maDMP project v34</title>
26        <array>
27           <funding>
28              <funder_id>
29                 <identifier>Other</identifier>
30              </funder_id>
31              <funder_id>
32                 <identifier>501100009244</identifier>
33                 <type>fundref</type>
34              </funder_id>
35              <funder_id>
36                 <identifier>EOSC Nordic</identifier>
37                 <type>other</type>
38              </funder_id>
39              <grant_id>
40                 <identifier>EOSC-Nordic-857652</identifier>
41                 <type>other</type>
42              </grant_id>
43              <funding_status>granted</funding_status>
44           </funding>
45        </array>
46     </project>
47  </array>
48  <ethical_issues_exist>no</ethical_issues_exist>
```

DMP85409v34curl.json | DMP85409v34curl.xml | **DMP85409v34transform0-93.xml** | DMP85409v34transform0-93.json

dmp extension FAIRscore sectionI DataDescription FileFormatsUNamesQ1 R-value

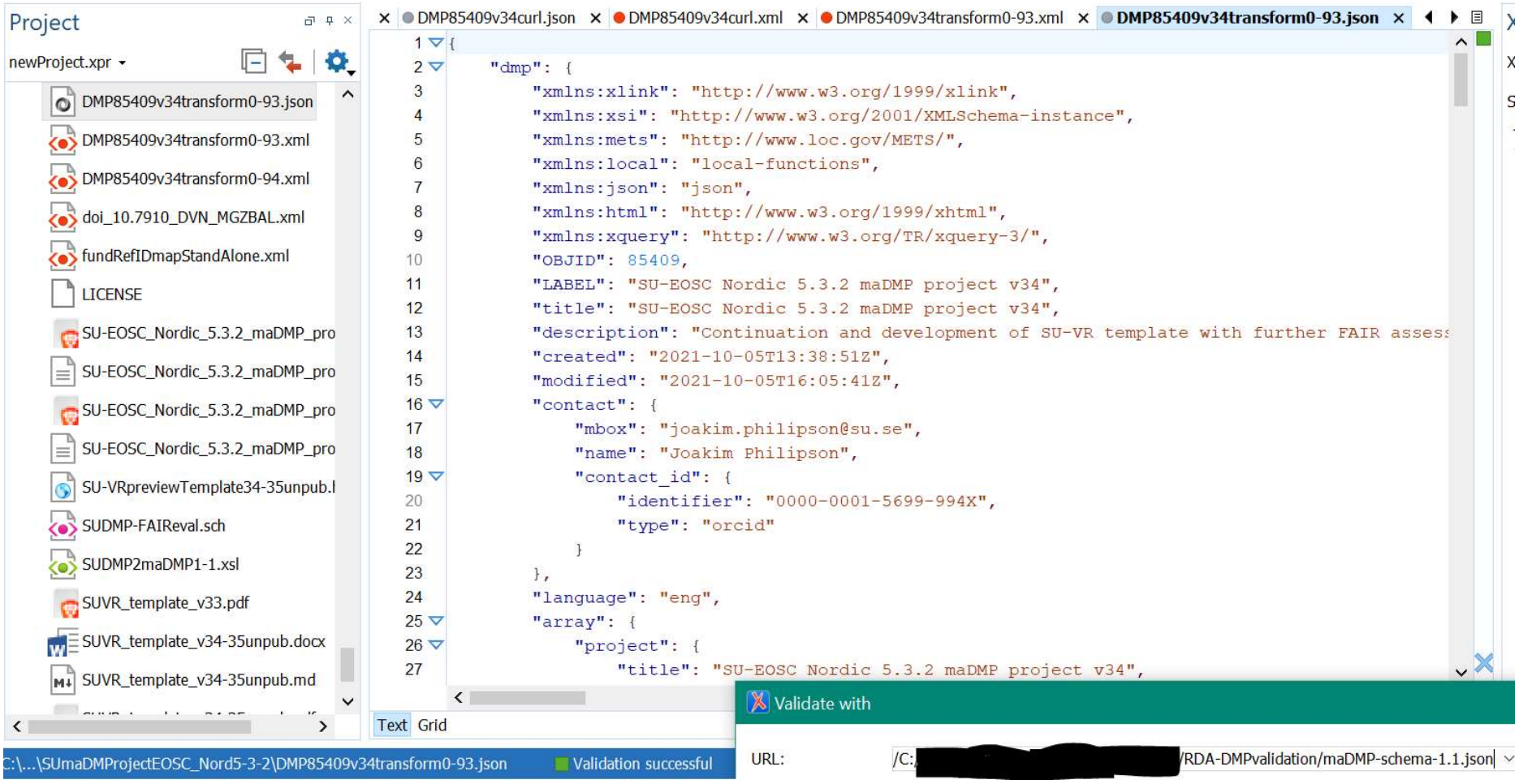
```
115  <extension>
116     <FAIRscore>
117        <sectionI>
118           <DataDescription>
119              <FileFormatsUNamesQ1>
120                 <Option>7. File names with only character set [0-9A-Za-z-_.], no spaces.</
121                 <F-value>1</F-value>
122                 <Option>6. Descriptive file names (e.g. '20200910SanFranciscoDaylight1pm.mp
123                 <F-value>1</F-value>
124                 <Option>5. Datestamped file names (e.g.
125                    '20201011StockholmMeanTemperatures.csv')</Option>
126                 <F-value>1</F-value>
127                 <Option>2. Sustainable file formats (e.g. .pdf; .csv; .txt)</Option>
128                 <I-value>1</I-value>
129                 <R-value>1</R-value>
130                 <Option>1. Non-proprietary file formats (e.g. .csv, .txt, .json, netCDF) </
131                 <I-value>1</I-value>
132                 <R-value>1</R-value>
133              </FileFormatsUNamesQ1>
134              <URL-idQ2>
135                 <Text>https://doi.org/10.7910/DVN/MGZBAL</Text>
136                 <A-value>1</A-value>
137              </URL-idQ2>
138              <idTypeQ3>
139                 <Option>doi</Option>
```

Text | Grid | Author

Stockholm University

DMP example of SU-VR template API output & transform

4. Valid RDA maDMP-schema-1.1.json



Stockholm University

Stockholm University

Thanks for your attention!

Contact

opendata@su.se

CC BY