Horizon 2020

1. Data Summary

What is the purpose of the data collection/generation and its relation to the objectives of the project?

Dataset Salzburg middle-aged marriages: Land Salzburg registers all marriages in its area

Dataset US wheat yield: US Department of Agriculture, collects this data by counting and summing the farmers data nationwide.

Dataset Combined: processed combination of "Salzburg middle-aged marriages" dataset and "US wheat yield" dataset

What types and formats of data will the project generate/collect?

Dataset Salzburg middle-aged marriages: structured relational data in CSV format

Dataset US wheat yield: structured relational data in XLS format

Dataset Combined: structured relational data in CSV format

Will you re-use any existing data and how?

Dataset Salzburg middle-aged marriages: existing data

Dataset US wheat yield: re-used data

Dataset Combined: non existing data, combination of two datasets

What is the origin of the data?

Dataset Salzburg middle-aged marriages: Land Salzburg

Dataset US wheat yield: US Department of Agriculture

Dataset Combined: Helmuth Breitfenfellner

What is the expected size of the data?

Dataset Salzburg middle-aged marriages:

Dataset US wheat yield:

Dataset Combined:

To whom might it be useful ('data utility')?

Dataset Salzburg middle-aged marriages: Everyone interested in marriage patterns of Salzburg citizens

Dataset US wheat yield: Everyone interested in US wheat output

Dataset Combined: Everyone interested in correlation between Salzburg middle-age marriages and US wheat yield

2. FAIR data

2.1 Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

Dataset Salzburg middle-aged marriages: Nein

Dataset US wheat yield: Nein

Dataset Combined: Ja https://doi.org/10.5281/zenodo.2648330

What naming conventions do you follow?

Dataset Salzburg middle-aged marriages: none

Dataset US wheat yield: none

Dataset Combined: none

Will search keywords be provided that optimize possibilities for re-use?

-
-

Do you provide clear version numbers?

Dataset Salzburg middle-aged marriages: none

Dataset US wheat yield: none

Dataset Combined: none

What metadata will be created? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Dataset Salzburg middle-aged marriages:

- automatic: Es wird kein festgelegtes System zur Beschreibung genutzt
- semi-automatic: Es wird kein festgelegtes System zur Beschreibung genutzt
- manual: Es wird kein festgelegtes System zur Beschreibung genutzt

Dataset US wheat yield:

- automatic: Es wird kein festgelegtes System zur Beschreibung genutzt
- semi-automatic: Es wird kein festgelegtes System zur Beschreibung genutzt
- manual: Es wird kein festgelegtes System zur Beschreibung genutzt

Dataset Combined:

- automatic: Es wird kein festgelegtes System zur Beschreibung genutzt
- semi-automatic: Es wird kein festgelegtes System zur Beschreibung genutzt
- manual: Es wird kein festgelegtes System zur Beschreibung genutzt

2.2. Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default? If certain data sets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

Dataset Salzburg middle-aged marriages: Ja

Dataset US wheat yield: Ja

Dataset Combined: Ja

How will the data be made accessible (e.g. by deposition in a repository)?

Dataset Salzburg middle-aged marriages: www.data.gv.at

Dataset US wheat yield: www.ers.usda.gov

Dataset Combined: Zenodo

What methods or software tools are needed to access the data?

Dataset Salzburg middle-aged marriages: Browser

Dataset US wheat yield: Browser

Dataset Combined: Browser

Is documentation about the software needed to access the data included?

Dataset Salzburg middle-aged marriages:

Dataset US wheat yield:

Dataset Combined:

Is it possible to include the relevant software (e.g. in open source code)?

Dataset Salzburg middle-aged marriages: Ja

Dataset US wheat yield: Ja

Dataset Combined: Ja

Where will the data and associated metadata, documentation and code be deposited? Preference should be given to certified repositories which support open access where possible.

Dataset Salzburg middle-aged marriages: www.data.gv.at (not certified)

Dataset US wheat yield: www.ers.usda.gov (not certified)

Dataset Combined: Zenodo (not certified)

Have you explored appropriate arrangements with the identified repository?

Dataset Salzburg middle-aged marriages: none

Dataset US wheat yield: no

Dataset Combined: no

Is there a need for a data access committee?

Nein

Are there well described conditions for access (i.e. a machine-readable license)?

Dataset Salzburg middle-aged marriages: None, open access

Dataset US wheat yield: None, open access

Dataset Combined: None, open access

How will the identity of the person accessing the data be ascertained?

Dataset Salzburg middle-aged marriages: open access

Dataset US wheat yield: open access

Dataset Combined: open access

2.3. Making data interoperable

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards

for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

Dataset Salzburg middle-aged marriages:

- Ja

Dataset US wheat yield:

- Ja

Dataset Combined:

- Ja

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

Dataset Salzburg middle-aged marriages:

- Dublin Core

Dataset US wheat yield:

- Dublin Core

Dataset Combined:

- Dublin Core

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Dataset Salzburg middle-aged marriages: Nein

Dataset US wheat yield: Nein

Dataset Combined: Nein

2.4. Increase data re-use (through clarifying licences)

How will the data be licensed to permit the widest re-use possible?

Dataset Salzburg middle-aged marriages: CC-BY-3

Dataset US wheat yield: CC-0

Dataset Combined: CC-BY-3

When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Dataset Salzburg middle-aged marriages: 1. Juli 2019

Dataset US wheat yield: 1. Juli 2019

Dataset Combined: 1. Juli 2019

Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

Dataset Salzburg middle-aged marriages: Everyone interested in marriage patterns of Salzburg citizens

Dataset US wheat yield: Everyone interested in US wheat output

Dataset Combined: Everyone interested in correlation between Salzburg middle-age marriages and US wheat yield

How long is it intended that the data remains re-usable?

Dataset Salzburg middle-aged marriages: infinite

Dataset US wheat yield: infinite

Dataset Combined: infinite

3. Allocation of resources

What are the costs for making data FAIR in your project?

Personnel: 0

Other: 0

How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).

Who will be responsible for data management in your project?

Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

Preservation decisions

- Zenodo

Costs

Personnel: 0

Other: 0

4. Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

Is the data safely stored in certified repositories for long term preservation and curation?

Dataset Salzburg middle-aged marriages: www.data.gv.at, (certified)

Dataset US wheat yield: www.ers.usda.gov, (not certified)

Dataset Combined: Zenodo, (certified)

5. Ethical aspects

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

Dataset Salzburg middle-aged marriages: Nein

Dataset US wheat yield: Nein

Dataset Combined: Nein

Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

Dataset Salzburg middle-aged marriages:

- Personal data:
- Informed consent: Nein

Dataset US wheat yield:

- Personal data:
- Informed consent: Nein

Dataset Combined:

- Personal data:
- Informed consent: Nein

6. Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

Nein