The I-ADOPT Variable Modelling Challenge

RDA **I**nteroper**A**ble **D**escriptions of **O**bservable **P**roperty **T**erminology WG (I-ADOPT WG)

Core members:
Barbara Magagna, GO FAIR Foundation, NL
Gwenaëlle Moncoiffé, BODC, UK
Anusuryia Devaraju, CSIRO, AU
Maria Stoica, University of Colorado, US
Sirko Schindler, German Aerospace Center, DE
Alison Pamment, Centre for Environmental Data Analysis, UK

Created by

In collaboration with

RESEARCH DATA ALLIANCE
I-ADOPT WG

The I-ADOPT Variable Modelling Challenge

Participation, submission and scoring rules

RDA **I**nteroper**A**ble **D**escriptions of **O**bservable **P**roperty **T**erminology WG (I-ADOPT WG)

Core members:
Barbara Magagna, GO FAIR Foundation, NL
Gwenaëlle Moncoiffé, BODC, UK
Anusuryia Devaraju, CSIRO, AU
Maria Stoica, University of Colorado, US
Sirko Schindler, German Aerospace Center, DE
Alison Pamment, Centre for Environmental Data Analysis, UK

Created by

In collaboration with

RESEARCH DATA ALLIANCE
I-ADOPT WG

Participation Rules

- Subscribe to the Challenge before September 22
- As individual or as a group
- Reviewers are excluded from participation (I-ADOPT core group & external reviewers)
- Don’t share your modelling results with other competitors
- Per participant we accept only one modelling solution per variable
- Participants can submit as many variables as they wish from the predefined list

Submission rules

- The list of variables will be published on 16 September
- All submission should be made as a zip file
- Label your zip file with the names of the participant(s)
- Via email: iadopt.variable@gmail.com
- Include the questions and your answers (next slide) in the email
- Submission period during the Challenge Week:

16 - 22 September, 2024 – Anywhere on Earth (AoE)

Questionnaire

1. Are you submitting as a team or as an individual?
2. Please provide your name(s), ORCID and affiliations. Provide one contact email.
3. What is your professional background?
   (e.g. data steward, researcher, semantic expert, database manager, …)
4. What research domains are you currently involved in?
   (e.g. environmental science, social science, health science, ...)
5. What is your level of knowledge about semantic technologies?
   (limited knowledge, basic understanding, in-depth knowledge)
6. Have you worked with the I-ADOPT Framework before, or is this your first time?
7. Have you found the instructional videos helpful?
8. What were the main challenges in modelling variables using I-ADOPT?
9. Why is the I-ADOPT Framework useful in your case? What are your objectives?
10. We are planning to develop an I-ADOPT service for providing FAIR variable descriptions. How likely will you consider using such a service in near future? *(e.g. unlikely, likely, very likely)*

Different options how to submit your variable

You can choose to submit your variable descriptions in one of the formats below (either human-readable or RDF), depending on your skills. Note, however, that machine-readable output as well as the use of semantic concepts rather than pure terms score higher.

Creating human readable output:

A. Excel template

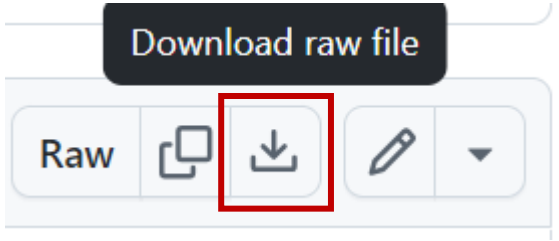
B. Text template

Creating machine-readable (RDF) output:

C. Online form

D. Turtle file

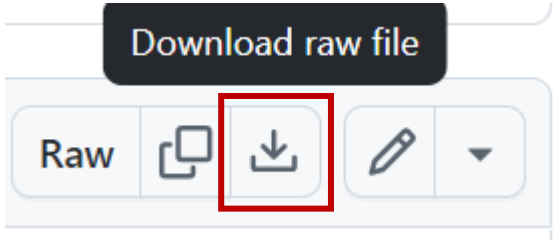
Option A: Excel template



- Go to this template and download it
- Use one tab per variable, copying the template to each tab
- Name the tab with the provided variable name
- Use new rows for multiple Constraints or ContextObjects as shown in the provided example
- Green columns are mandatory for a variable to be valid
- White columns are optional, but add to the overall score
- Purple columns are for the links (URIs) to semantic concepts, lead to a higher score
- Yellow columns can be useful in the modelling process but don't score

Template: https://github.com/i-adopt/variables/blob/main/templates/VariableModellingTemplate.xlsx

Option B: Text template



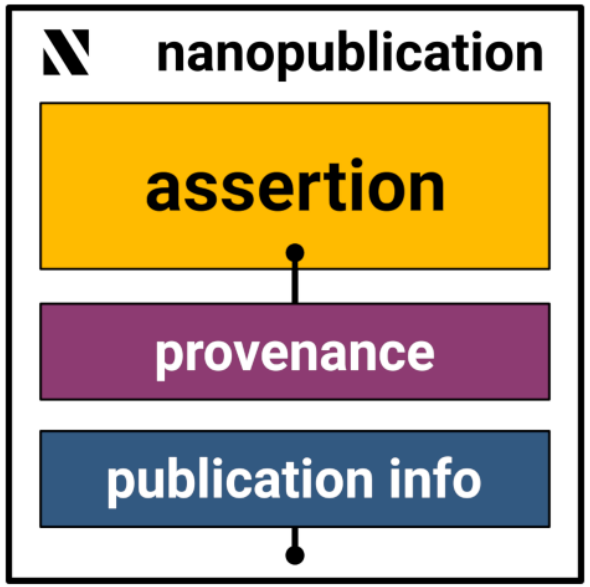
- Go to this template and download it
- Use one text file per variable, see example
- Name the text file with the provided variable name
- List additional information you used for the modelling process, if you like (doesn't score)

Template: https://github.com/i-adopt/variables/blob/main/templates/VariableModellingTemplate.txt

Example: https://github.com/i-adopt/variables/blob/main/templates/VariableModellingExample.txt

Option C: Online form

- The online form is a nanopublication template.
- A nanopublication is a small RDF knowledge graph which allows to create machine readable RDF statements with associated provenance information.
- To use the template you need to have an ORCID. Please follow the instructions online. Find here an example.
- Provide the full identifiers of the published variable nanopublications in a zipped txt file.



Template:
https://nanodash.knowledgepixels.com/publish?16&template=https://w3id.org/np/RAGUWnXKhfKYwmMoDK-LVXlEnnGdAuzFZKR9FZsXHJsxQ
Example:
https://w3id.org/np/RANsVBnIjFjay8xxt_Iw3io3zq4IOei23TkTUPx4WsLFc

Option D: Turtle file

- Create one turtle file per variable
- See this RDF turtle file as an example
- Use a mocked, local identifier for your variable (=name of the turtle file)
- If you can not find a semantic concept for a component, use a local identifier and assign a proper label

https://github.com/i-adopt/variables/blob/main/EnvThes/30351.ttl

Scoring Rules

- Submissions are evaluated according to the following criteria

Identify components and roles	up to 5 point(s) per variable
Annotate components with concepts	up to 1 point(s) per variable component
Provide descriptions of the variable	up to 3 point(s) per variable
Provide RDF output	up to 2 point(s) per variable
Adhere to formal cardinality requirements	up to 5 point(s) per variable

- Depending on the difficulty of the variable a modifier is applied:

Simple variables	multiply overall score by 1
Advanced variables	multiply overall score by 2

Winners

- Final ranking is based on both quality and quantity.
- Per participant, we select the 10-highest scoring variables.
  The sum of their scores will be the score of the participant.

We have a **prize fund** to distribute among **successful participants**!

Valid contributions will be published

- We will publish valid contributions on this website:

https://i-adopt.github.io/variables/index.html

- All winners will be listed on the I-ADOPT website
- The challenge outcome will be presented at the RDA 23rd Plenary Meeting in Costa Rica (12-14 November 2024)

Acknowledgements

Created by: GO FAIR Foundation

In collaboration with: RDA RESEARCH DATA ALLIANCE I-ADOPT WG

GO FAIR Foundation BY NC ND