# COSMOS 2024


# I-ADOPT:
# A systematic Way to Represent Variables

Barbara Magagna, GO FAIR Foundation
George Alter, University of Michigan, Institute for Social Research
Fran Cotton, INSEE
Flavio Rizzolo, Statistics Canada
Steven McEachern, Australian National University

COSCMOS Conference 2024, April 11, 2024, I-ADOPT, cite as: https://osf.io/ce9gy

# I-ADOPT Framework – A Semantic Broker

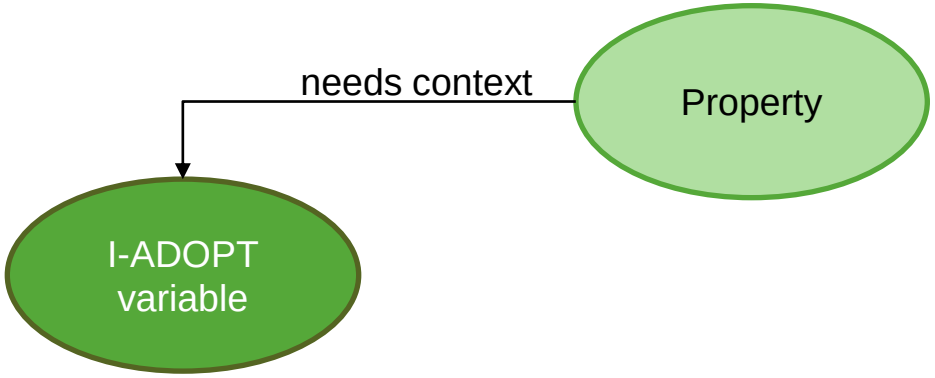
I-ADOPT provides a **standardized descriptions for variables** and supports **interoperability** between existing terminologies by

- ❏ Enabling **mappings** between variable descriptions **across terminologies**
- ❏ Requiring **no change to existing structures**
- ❏ Adding rich (**human-readable and machine-actionable**) descriptions with **qualified references**

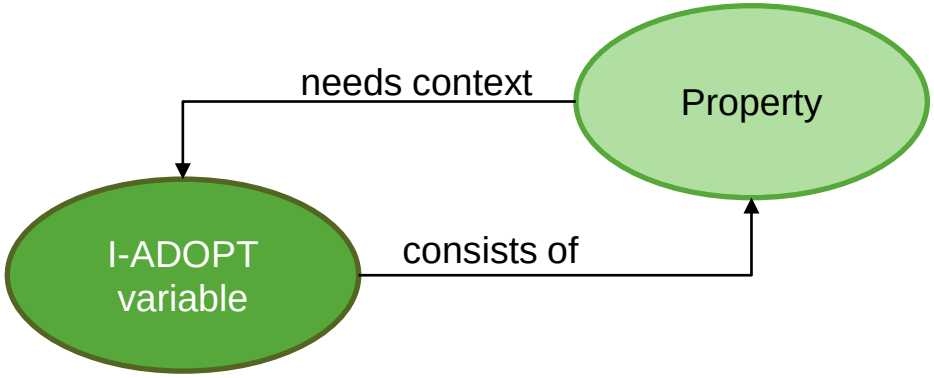
COSCMOS Conference 2024, April 11, 2024, I-ADOPT, cite as: https://osf.io/ce9gv

# Descriptions of Properties

# Contextualised properties -> observable properties -> I-ADOPT variables

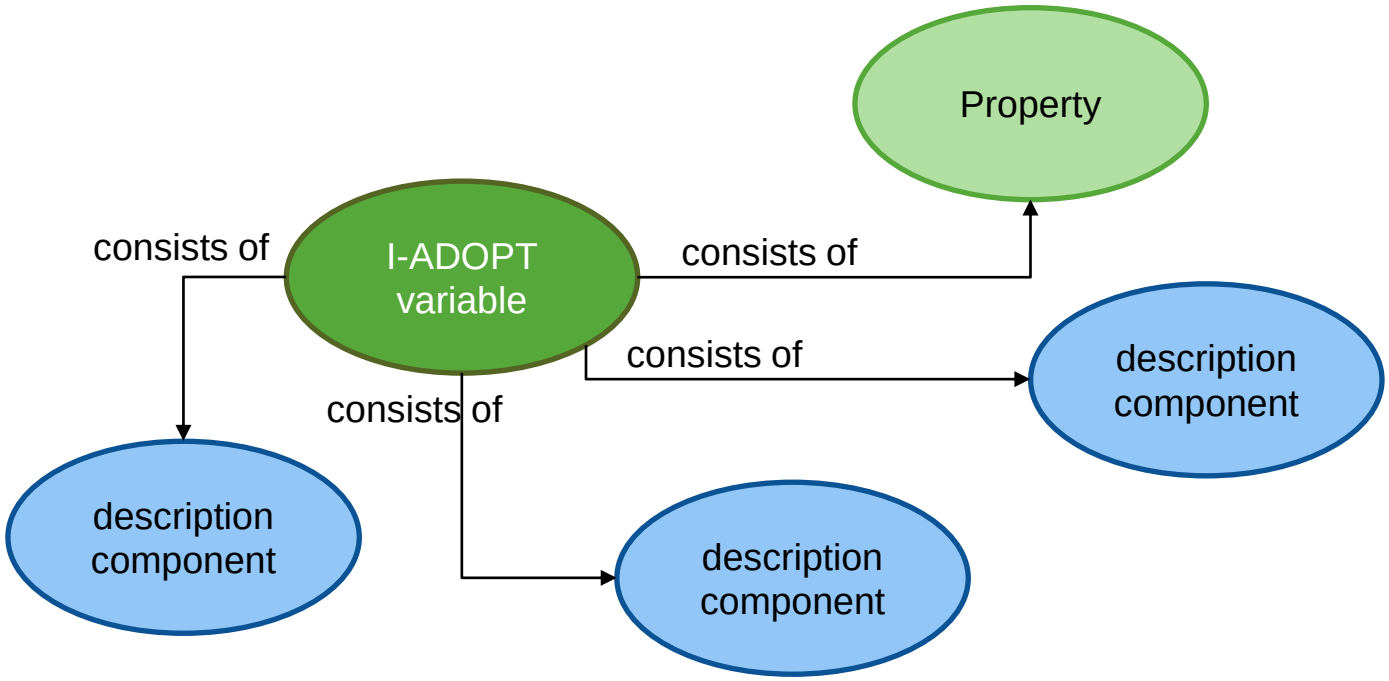


# Contextualised properties -> observable properties -> I-ADOPT variables



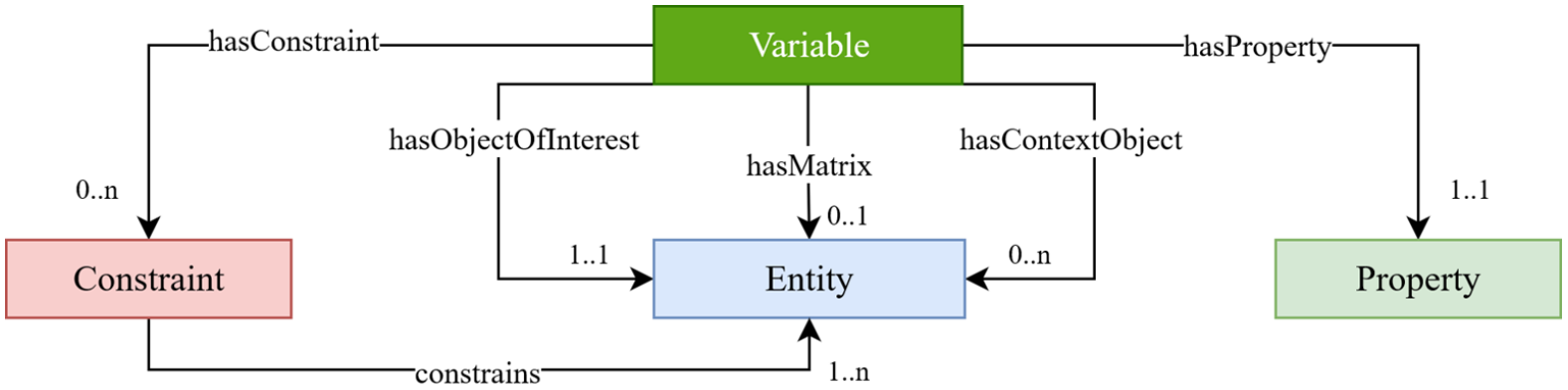

COSCMOS Conference 2024, April 11, 2024, I-ADOPT, cite as: https://osf.io/ce9gy

# I-ADOPT variable consists of various description components



# The I-ADOPT Ontology

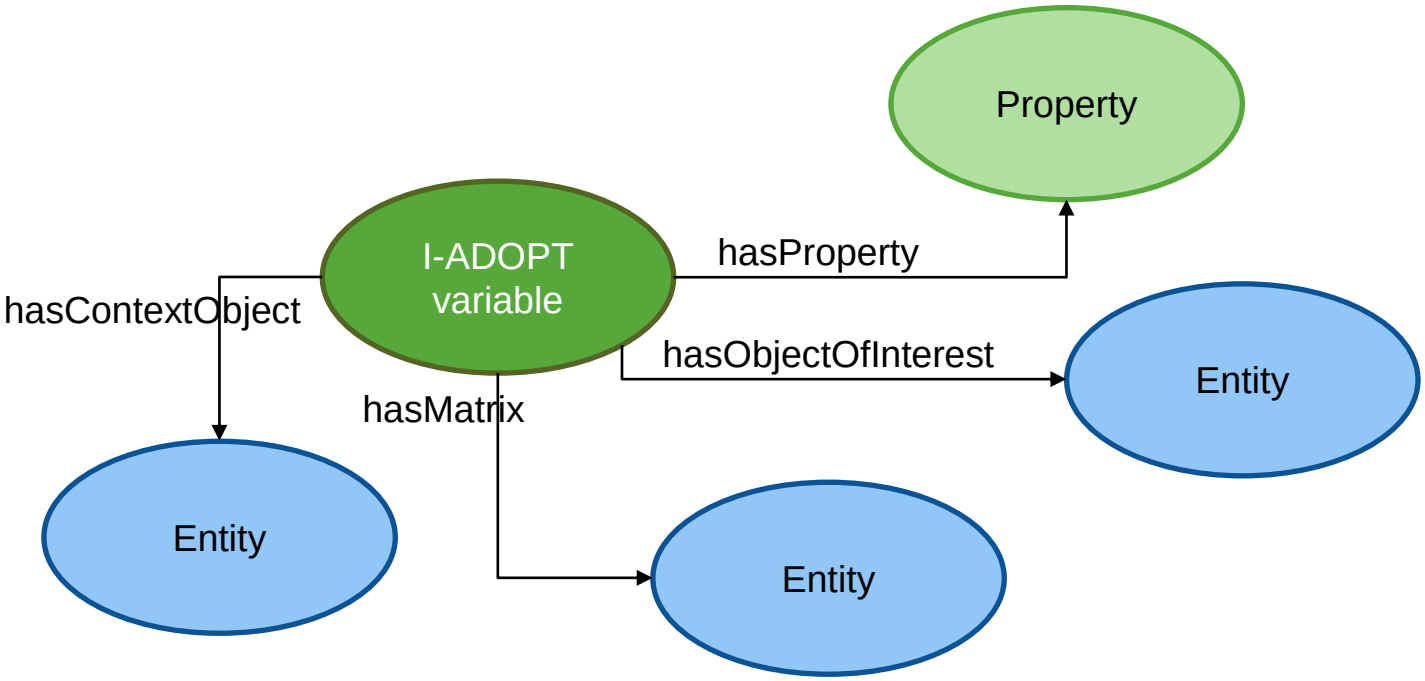
https://w3id.org/iadopt



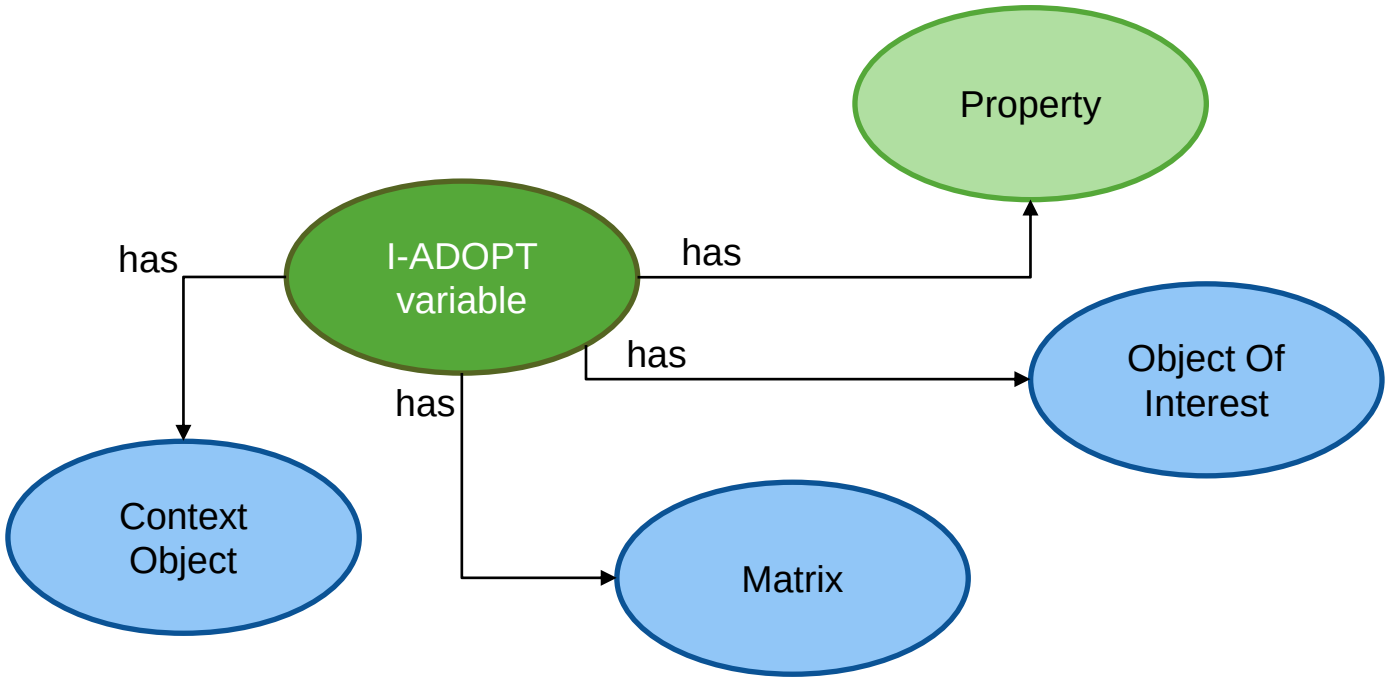

- https://github.com/i-adopt
- DOI: 10.15497/RDA00071

RESEARCH DATA ALLIANCE
I-ADOPT WG

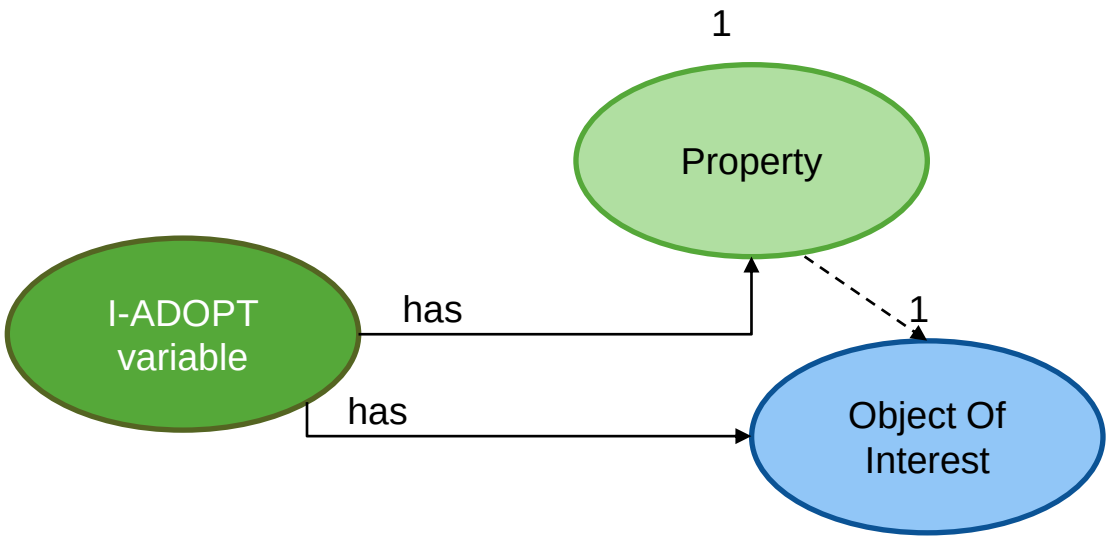
# I-ADOPT ontology explained



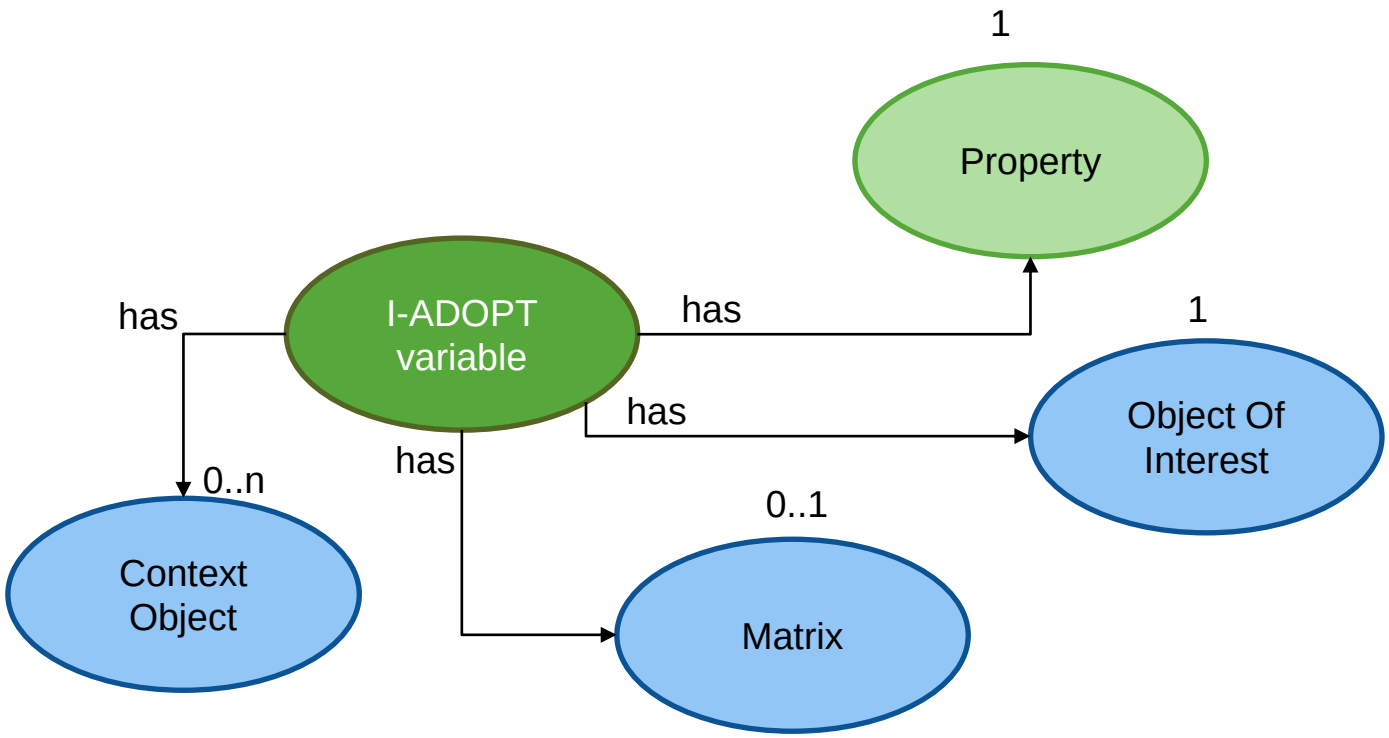
# I-ADOPT ontology simplified explanation



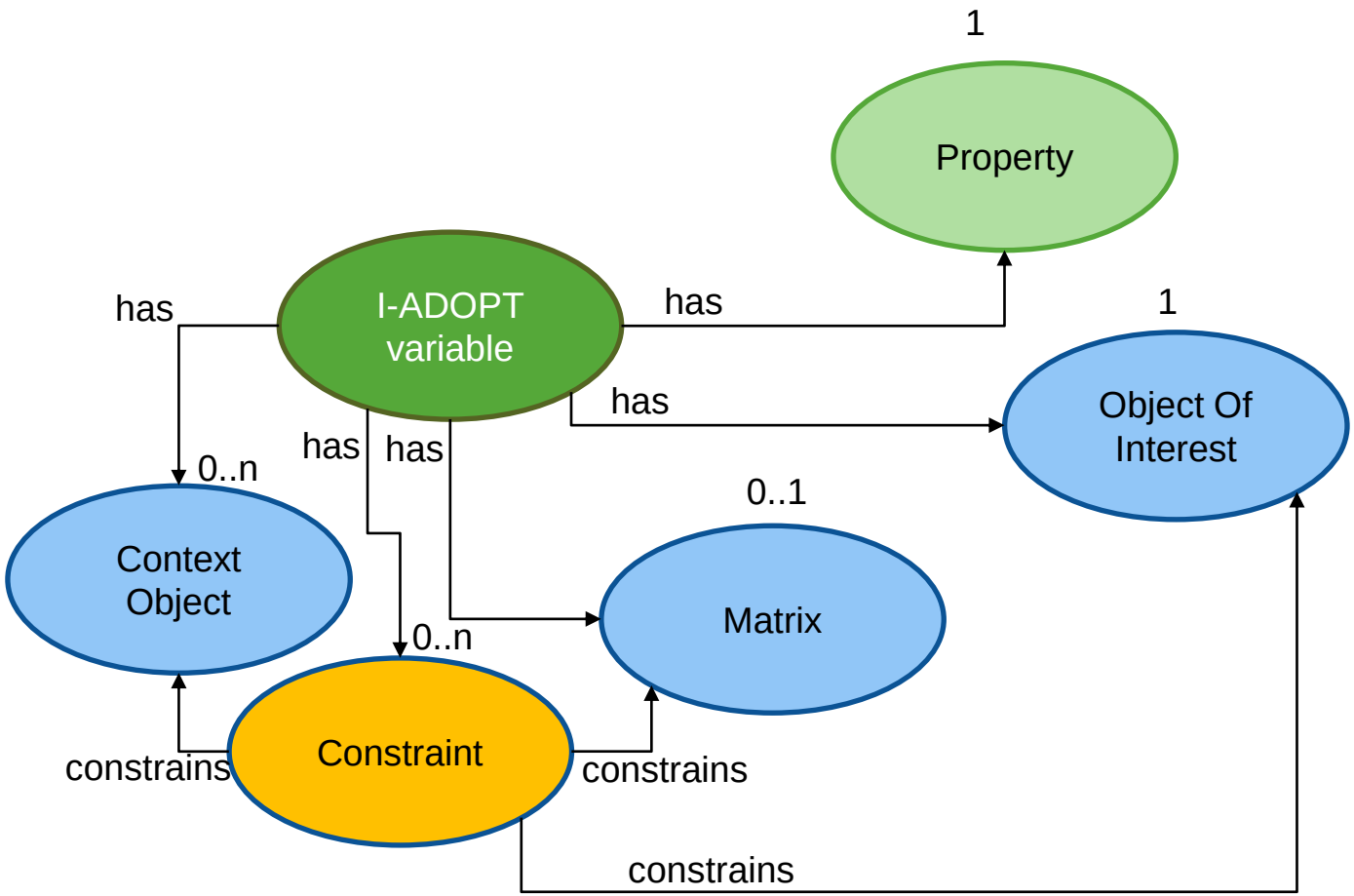
# I-ADOPT ontology simplified explanation – minimal description



# I-ADOPT ontology simplified explanation – extended description



# I-ADOPT (simplified) in an OWL ontology



# RDA I-ADOPT recommendations (DOI: 10.15497/RDA00071)

1. *Descriptions should be human and machine-readable*

Data creators and data curators or data publishers should describe the variable(s) of their datasets in a human- and machine-readable format.

2. *Descriptions should be explicit and sufficient*

The variable description should contain sufficient information so that the data can be re-used with minimum reliance on free-text documentation.

3. *Use of semantic artefacts*

The description should use FAIR semantic artefacts (e.g., controlled vocabularies or ontological relationships) and be compatible with Linked Data.

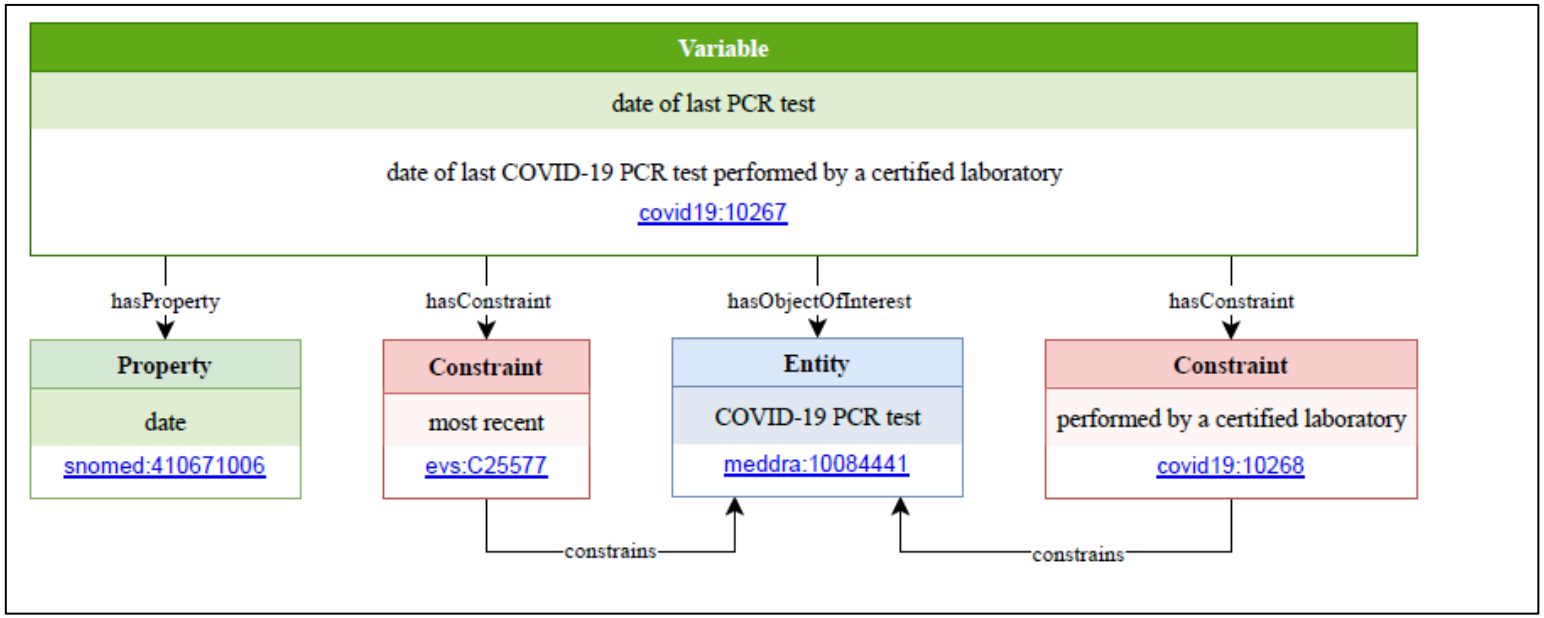
4. *Use of I-ADOPT ontology*

The description should follow a decomposition approach consistent with the classes and relations defined in the I-ADOPT ontology

5. *Reuse of I-ADOPT aligned terminology*

Reuse existing FAIR terminologies that are aligned with the I-ADOPT Framework. If no such terminology is available, you may either extend the existing variable description or create a new variable following the I-ADOPT framework.

# I-ADOPT example: Date of last PCR Test



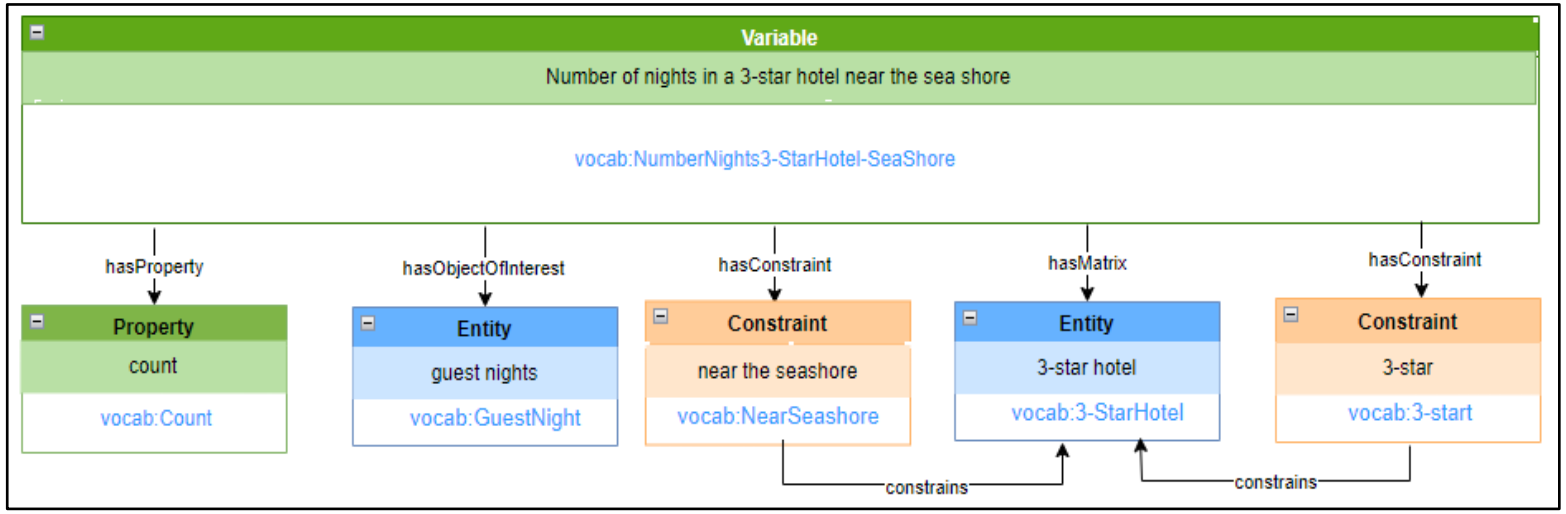

**Object of interest**: *COVID-19 PCR test*

**Property**: *date*

**Constraint**: *most recent* (constrains *COVID-10 PRC test*)

**Constraint**: *performed by a certified laboratory* (constrains *COVID-10 PRC test*)

# I-ADOPT example:
# Number of nights in a 3-star hotel near the seashore




**Object of interest**: *guest nights*

**Property**: *count*

**Matrix**: *hotel*

**Constraint**: *3-star* (constrains *hotel*)

**Constraint**: *near the seashore* (constrains *hotel*)

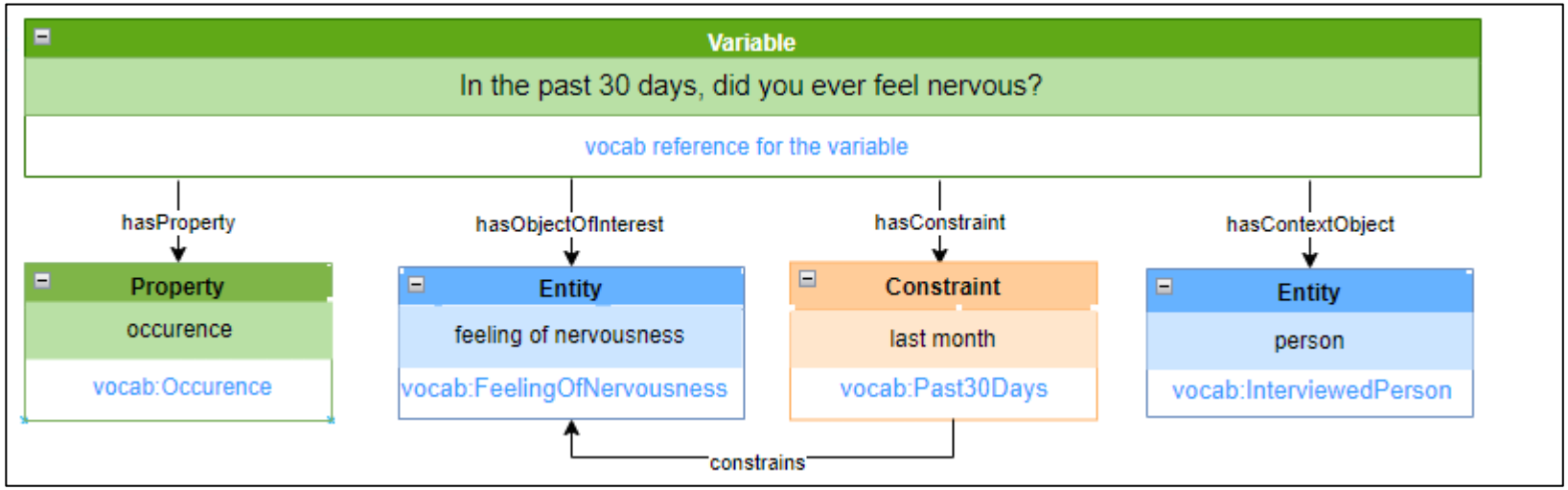
# Disambiguate variable descriptions using I-ADOPT
# 3 questions about feeling nervous

Consider these three questions:

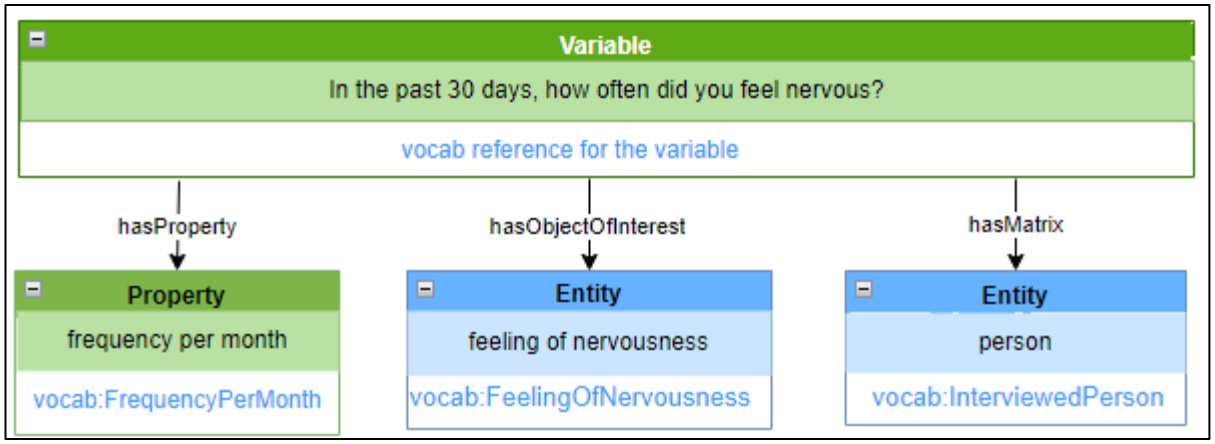
Q1: In the past 30 days, did you ever feel nervous?

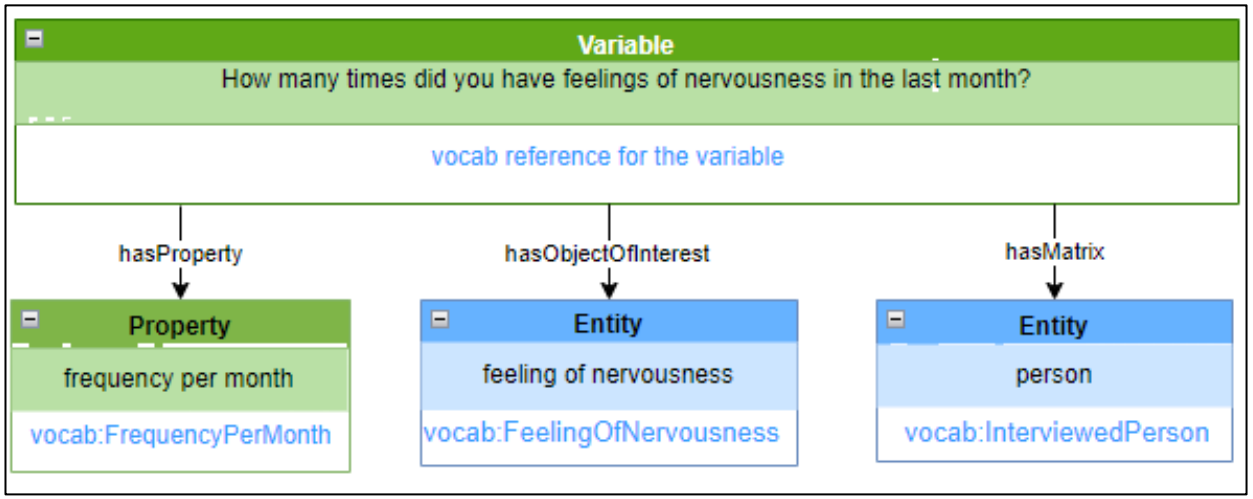
Q2: In the past 30 days, how often did you feel nervous?

Q3: How many times did you have feelings of nervousness in the last month?



# Disambiguate variable descriptions using I-ADOPT
# 3 questions about feeling nervous



# Outlook: Developing I-ADOPT services for reuse by researchers

- Based on:
  - User input (free text description) and specify FSRs (vocabularies) chosen by the research community
  - Named Entity Recognition (NER) and/or
  - Large Language Models (LLM)

  both require large sets of pre-modelled I-ADOPT variables (training set)

  Results into:
  - Decomposition in atomic parts based on NER
  - Arrangement of parts in I-ADOPT roles based on Variable Design Patterns