The FAIR Principles and the Role of Ontological Conceptual Modelling



Robert Pergl robert.pergl@fit.cvut.cz



KEG, University of Economics, May 19, 2022

Contents

- 1. Motivation
- 2. The FAIR Principles Explained
- 3. The Challenge of "I": Semantic Interoperability
- 4. 3 Case Studies

Motivation



The FAIR Principles Explained



The Challenge of "I": Semantic Interoperability

Stamper's Semiotic Ladder









Giancarlo Guizzardi; Ontology, Ontologies and the "I" of FAIR. Data Intelligence 2020; 2 (1-2): 181–191. doi: https://doi.org/10.1162/dint_a_00040

- 1). I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation;
- 2). I2. (meta)data use vocabularies that follow FAIR principles;
- 3). 13. (meta)data include qualified references to other (meta)data.

"In order to address these three interoperability requirements, we need a language that supports us in: (a) **systematically making ontologically consistent representation choices**; (b) **making explicit the ontological nature of the elements represented**, i.e., the ontological commitment that is being made; and (c) **identifying and characterizing the nature of the relations between real-world entities represented in these different data artifacts**. In order words, we need **a language that is truly ontological in nature**, i.e., a language that explicitly commits to a foundational ontology".

- Just logic-based languages do not suffice (FOL, OWL)
- Foundational ontologies suggested, namely Unified Foundational Ontology (UFO)



March 12, 2020	Project deliverable Open Access		
D2.2 FAIR Semantics: First recommendations © Le Franc, Yann; © Parland-von Essen, Jessica; © Bonino, Luiz; Lehväslaiho, Heikki; © Coen, Gerard; © Staiger, Christine		3,406 © views	2,648 ≛ downloads
		See more details	
December 21, 2020	Project deliverable Open Access		
D2.5 FAIR Semantics Recommendations Second Iteration		1,484	1,211
		@ views	📥 downloads
		See more details	
😰 Hugo, Wim; 😰 Le Franc, Yann; 💿 Coen, Gerard; 💿 Parland-von Essen, J	essica; 💿 Bonino, Luiz		

- P-Rec. 10: Foundational Ontologies may be used to align semantic artefacts.
- P-Rec. 11: A standardized knowledge representation language should be used for describing semantic artefacts

Foundational Ontology

• Also known as *upper ontology* or *top-level ontology*.

"Foundational ontologies are **axiomatic accounts of high-level domain-independent categories about the real world**. They constitute toolboxes of reusable information modeling primitives for **building application ontologies in specific domains**. As such, they enhance **semantic interoperability** between agents by specifying descriptively adequate **shared conceptualisation**." [1]

- 18 listed at https://en.wikipedia.org/wiki/Upper_ontology
- We use the Unified Foundational Ontology (UFO) + OntoUML

[1] Schneider, Luc. "How to build a foundational ontology." Annual Conference on Artificial Intelligence. Springer, Berlin, Heidelberg, 2003.

Unified Foundational Ontology

- Developed by prof. Giancarlo Guizzardi and associates.
- Incorporating developments from GFO, DOLCE and the Ontology of Universals underlying OntoClean in a single coherent foundational ontology.
- Has been used to analyze, redesign and integrate reference conceptual models in a number of complex domains.
- 3 fundamental parts:
 - UFO-A: Structural aspects
 - UFO-B: Behavioral aspects
 - UFO-C: Social aspects
- + several others (UFO-S, UFO-L, ...)

OntoUML

- Ontologically well-founded language for ontology-driven conceptual modelling based on UFO.
- Built as a UML profile: Looks familiar X has a significantly different semantics





Guizzardi, G. et al. (2015) Towards ontological foundations for conceptual modeling: The unified foundational ontology (UFO) story. Applied Ontology. 10, 3-4





OntoUML community portal

Learn more about UFO, OntoUML and ontology-driven conceptual modelling with us. Share your experience and knowledge and join our community to make it better!

https://ontouml.org

Case Studies





Reference Ontology

FAIR Convergence Matrix



Currently known as <u>FAIR Implementation Profile</u> (FIP)

Sustkova, H.P. et al. 2020. FAIR Convergence Matrix: Optimizing the Reuse of Existing FAIR-Related Resources. Data Intelligence. 2, 1–2

GO FAIR Reference Ontology



https://github.com/go-fair-ins/GO-FAIR-Ontology

Sustkova, H.P. et al. 2020. FAIR Convergence Matrix: Optimizing the Reuse of Existing FAIR-Related Resources. Data Intelligence. 2, 1–2





Marrow Donor Registry Simulator









https://www.elixir-czech.cz htt

https://ccmi.fit.cvut.cz/en

https://www.ntis.zcu.cz/en

Marrow Donor Registry Domain

For the successful operation of the donors' registry, it is important:

- to coordinate individual executive registry units (e.g., donor centers, laboratories)
- to collaborate with other donors' registries, organizations specialized in immunology, hematology, oncology, and transplantation (e.g., WMDA, BMDW)

Interoperability

ORGANIZA-

TIONAL

level executive units

responsibility

TECHNICAL

level

database structure

for exchanged data

levels

EXPERT

level

search

donor-patient match

examinations

Interoperability is a key factor for registries' national operation and international cooperation.

- OntoUML + UFO serve to ensure the precise definition of terms, relationships and rules between them.
- A simulator application was built to make the ontology operational.
 - Visualisation
 - Validation
 - Learning, knowledge exchange
 - Plans to apply it also for information systems interoperability

Structural model (UFO-A) visualisation + editing



Behaviour - process (UFO-B)



Simulation





Data Platform for Marketing Communications









EVROPSKÁ UNIE Evropské strukturální a investiční fondy OP Praha – pól růstu ČR



Data Platform for Marketing Communications: Challenge

- Integration of heterogeneous closed & open data sources
 - Different formats XLS, CSV, custom formats
 - Syntactic heterogeneity different encoding of entities
 - Semantic heterogeneity different conceptualisations, e.g. term "Person"
 - Synonyms
- Business Intelligence practices, such as ETL scripts just partially applicable:
 - "Infinite extendability" 100s and 1000s of datasets
 - Updates and changes of original datasets
 - Various usage scenarios direct analysis, downloading for analysis, API for apps

Data Platform for Marketing Communications: Solution

- 1. Ontological analysis of the domains (OntoUML+UFO) \rightarrow solid "backbone"
- 2. Mapping of ontological entities to data elements
- 3. Data transformations using the models
- 4. Using the models in the developed Data Platform (WIP)

1. Ontological analysis of the domains



Currently >6 800 distinct entities (>13 000 total)

2. Mapping of ontological entities to data elements



3. Data transformations using the models

```
"identifiers": [
        "eu-ceny-plynu-pro-domacnosti-2roky"
     1,
      "stereotype": "Data"
    },
      "name": "Spotřeba plvnu Data",
     "rules": "\"consom\" = 4141901 -> \"Spotřeba plynu pod 20GJ\";\n\"consom\" = 4141902 -> \"Spotřeba plynu mezi 20 -
200 GJ\";\n\"consom\" = 4141903 -> \"Spotřeba plynu nad 200 GJ\";",
     "identifiers": [
        "eu-cenv-plvnu-pro-domacnosti-2rokv"
      ],
      "stereotype": "Data"
    },
      "name": "Jednotky spotřeby plynu Data",
     "rules": "\"unit\" -> \"Jednotky spotřeby plynu\";",
      "identifiers": [
       "eu-ceny-plynu-pro-domacnosti-2roky"
      1,
      "stereotype": "Data"
    },
      "name": "Zdanění cen dodávek plvnu Data",
     "rules": "\"tax\" = \"I TAX\" -> \"Cena dodávek plynu včetně daně\";\n\"tax\" = \"X TAX\" -> \"Cena dodávek plynu
bez daně\";",
      "identifiers": [
       "eu-ceny-plynu-pro-domacnosti-2roky"
      1,
      "stereotype": "Data"
    },
      "name": "Měna ceny dodávek plynu Data",
     "rules": "\"currency\" = \"EUR\" -> \"MEA EUR\";\n\"currency\" = \"NAC\" AND \"geo\" = \"AT\" -> \"MEA EUR
\";\n\"currency\" = \"NAC\" AND \"geo\" = \"BA\" -> \"MEA BAM\";\n\"currency\" = \"NAC\" AND \"geo\" = \"BE\" ->
\"MEA EUR\";\n\"currencv\" = \"NAC\" AND \"geo\" = \"BG\" -> \"MEA BGN\";\n\"currencv\" = \"NAC\" AND \"geo\" = \"CZ\" ->
```

4. Using the models in the developed Data Platform (WIP)



4. Using the models in the developed Data Platform (WIP)



Thank you for your attention!