

Ethics of Vocabulary Development & Use

Robert J. Rovetto

ontologos @ protonmail.com (Independent)

r.rovetto@unsw.edu.au (PhD student)

<http://purl.org/rrovetto/ethics-of-ontology>

Brief Bio

- Born USA
- Masters via Philosophy Dept., USA.
- The Reluctant / Unintended Ontologist
- PhD student, UNSW, AU

- Web:
 - <http://purl.org/rrovetto/ethics-of-ontology>
 - <https://github.com/rrovetto>
 - <http://orcid.org/0000-0003-3835-7817>

- Open to opportunities

Note

- Has been independent & unfunded project to date. Need formal support, collaborations & mentors to sustainably pursue this promising line of research.

- *Wearing 2 hats*: independent available for hire (ontologos@protonmail.com), and PhD student available for collaborations (r.rovetto@unsw.edu.au)

- University PPT template in what follows.

In a nutshell

- **WHAT :**
 - Exploring ethical and moral aspects of vocabulary (broadly construed) development & use.
 - Ethics of ontology, of metadata, of knowledge representation, ...
 - Developing guidelines
- **WHY you should care :**
 - mutual interest in vocabulary, intellectual exercise, determination of actual or potential ethical implications, informing policy, overlap with ethics of AI, FAIR guidelines call for ethical considerations, ...
 - *Impact statement:* can implementing FAIR guidelines potentially (if inadvertently) do harm?

A note on terminology

Vocabulary = systems or artifacts that declare* (typically in natural language) terminology for use in (non)technological contexts and systems.

- a.k.a. a *knowledge organization system (KOS)* or *semantic artifact /system*
 - Term lists, Controlled vocabularies
 - Taxonomies & Classification systems
 - Metadata schemas
 - Ontologies & Knowledge graphs
 - ...

*declare = explicitly state, assert, create or encode in writing within documents, or in computable formalisms/languages within computational systems.

Linking the Motivating Case to Broader Applications & Relevance

- *Motivation: A more complex system—computational ontology dev. & use*
 - Specifically, *highly abstract* ontology development and use (upper/foundational ontologies)
 - Observations, participation, critical reflection in/of specific ontology circles since 2009.
 - Both *ethics in* and *of* these systems is significant → conduct and content
 - Paper:
 - *The Ethics of Conceptual, Ontological, Semantic and Knowledge Modeling*, in *AI & Society*.
[Contact me to read full paper online.] [Contact me to collaborate on follow-up case-study paper]
- Commonality with other (similar) systems: meaning & semantics
 - Ontologies (developed in a certain way) are an example system aligning with FAIR concept
- Relevant to Artificial Intelligence (AI)
 - Knowledge Representation & Reasoning (KRR) → computational ontologies, symbolic approach
 - Topics in Ethics of AI

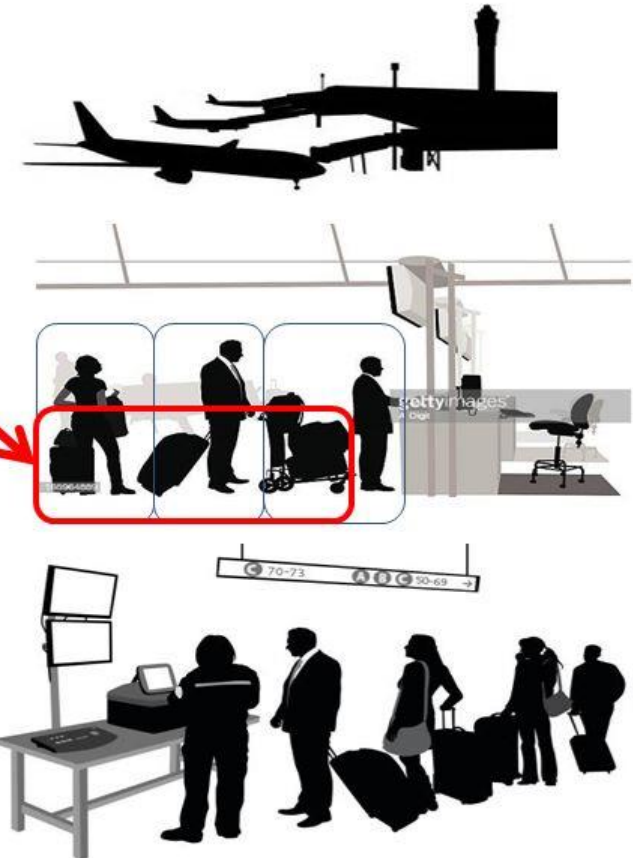
The core point: Meaning Matters

Check the semantic & ontological baggage (The Airport Analogy)

Meaning does matter

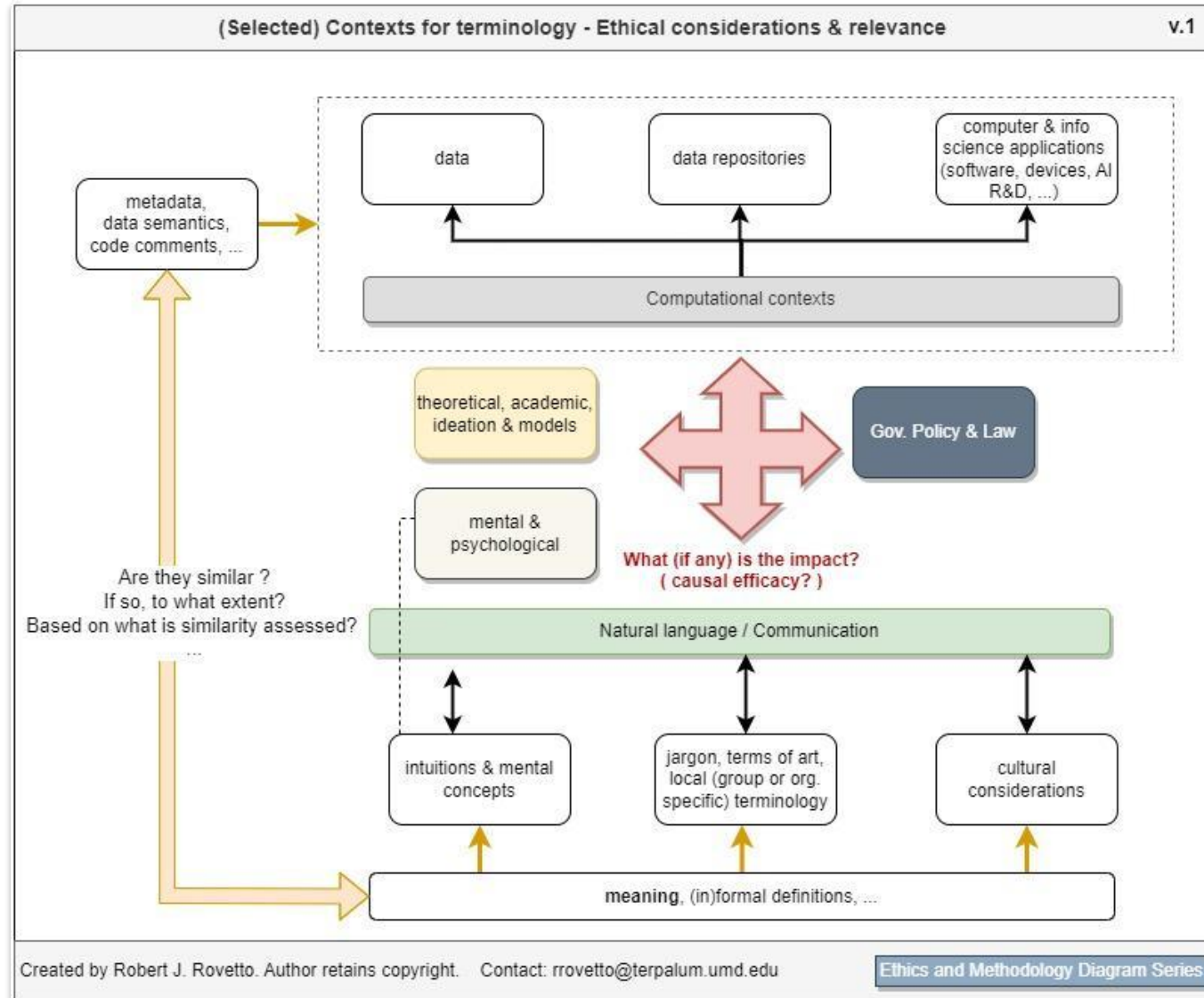
- (in)formal semantics → personal or group assumptions, worldviews, ontological (philosophical or computational) commitments, supposedly encoded by the vocab. or via its implementation language.

- This is you—your airplane
 - may take on passengers* and cargo**
 - * users, clients
 - ** data, content, other vocabs. / ontologies / ...
- These are some (not necessarily all) vocabularies
 - each with baggage (semantic, logical, philosophical)
- Just as airport security needs to know about baggage, so should you
 - Inspect the semantic baggage!



Slide by Robert Rovetto.
Images from internet search and
<https://www.istockphoto.com/nl/search/2/image?excludenudity=false&mediatype=illustration&phrase=airport%20security%20line>

Why it matters? (1)



Both are relevant:

- ethics in (*conduct*) and
- *ethics of* (content, use)

If a vocab or semantics is applied to personal or societal data/content, these considerations & concerns matter.

Why it matters? (2)

- Using a vocab. that labels or describes your data or the real-world things you're interested in, does so *according to* the definitions, characterizations, worldviews of the vocab (or its developers & owners) *and* any other vocabs it in turn uses
- Potential harm (technical & conceptual):
 - Incorrect or misleading labeling, definitions, semantic descriptions;
 - Non-consensual & (in)direct commitment to: a meaning, definition, philosophical worldview about... the given natural language terms, their supposed referents, or questions that are otherwise profound, controversial, unanswered, ...
 - person, life, mind, disease, disorder, love, ...
 - Ethical, legal and cultural implications?
 - Claims or attempts to universally force a single definition, or re-define dictionary definitions?
 - Ontology monopoly

What is the impact of a particular abstract characterization upon the ethical, legal, psychological and cultural values, doctrine and systems of both a given society and internationally?

Why it matters (post presentation addition)

(Voiced during the closing session)

- Another presentation at this event echoed the concerns I present in this talk. We read in “Data Science Transforming Maintenance” slides by Jens Klump:
 - “The Standardisation of high-level ontologies for industrial applications has the potential for excluding alternative models.”
 - “The risk of these standardization efforts at this stage is that they could act as barriers to alternative solutions.”
 - “Incompatible structures can lead to exclusionary structure.”
- Correct. And the motivating case reflects that, i.e., upper ontologies (at various degrees of abstraction)
 - *Potential to Pidgeon-hole* your data, content, know-how, into a particular model or conceptualization → Technical or architecturally locked-in?
 - Socio-psycho-political pressures to use a given vocab / ontology / model ...
 - Monopolization by one or more.
 - Observations in relevant communities at times shows hostile, arguably unethical, competitive tactics & environments [Contact me for specifics]

Example observations (& consequences) & confusions

Observation	Consequence (ethical implication?) – The harm in semantic harmonization
Limitations, errors, bias in actual vocabulary harmonization efforts	<ul style="list-style-type: none">• incorrect, inaccurate, superficial, misleading mappings;• ...• (un)intended unfair advantage for particular vocab (or its owners)

Terminological & Semantic harmonization should not do harm!

Confusions

Linguistic aspect (words) vs. Models or Theories (no matter how abstract) thereof (i.e., of linguistic terms and/or supposed referents) vs. the Computational construct vs. ...

Meta- language vs. Domain language

*Myth of not reinventing wheel → Misleading & can lead to harms mentioned (exclusion, monopoly, mislabeling, ...). The wheel analogy does *not necessarily or univocally* apply to vocabs / semantic artifacts (at least ontologies).*

Why it matters (3):

Topics in Ethics of AI applied to Ethics of Vocabulary

- **Bias** → To what degree does a vocabulary* exhibit bias? What sort?
 - Motivating case: concerns over “metaphysical bias” (bias of metaphysics) & its impact
 - Technical & Non-technical
- **Privacy** → To what degree does a vocab. (as a policy- and data-impacting system) affect privacy of persons and organizations.
 - Concerns over computational ontology violating privacy in virtue of their design & technical affordances such as data-sharing, linking data/content (“Non-consensual data sharing”)
- **Machine ethics** → to the extent that ethics applies to machines with autonomy (mimicking human intelligence, replacing some human tasks), how does use of a vocab. in those machines affect its autonomy, i.e., how does it influence actions by machines?
 - Question of the causal role played by vocabs. in AI.
 - Good and bad actions, and AI-recommended actions, ...

* = semantic system/artifact , or KOS

Ethics of the FAIR guidelines

FAIR guideline	Concerns & Considerations
Findable	What about for internal systems/project? What about sensitive or security information? What about use-cases and preferences not intended to be public-facing? → Consider the developers/owners business model & context.
Accessible	Too accessible? → Privacy concerns
Interoperable	In what sense? How is interoperability measured? Is it desired or part of use-case? If involving semantics, then concerns previously mentioned apply. Two semantic models are not necessarily interoperable at superficial levels.
Reusable	If involving semantics, then previously mentioned concerns apply. The meaning-focused aspect of semantic artifacts mean that reuse is not and should not be a mandatory requirement.

*How much is trend, hype? How much is demonstrably useful?
Serve as devil's advocate for greater understanding and risk management.
Applies also to 'open... x'*

Other interests – Invitation to Collaborate

- Interest in multi-language systems
- Interest in indigenous languages & concepts
 - at least for highly abstract terms
 - on topics of interest
- Relevant / Precedent Work:
 - I made formal recommendations to collect & research local & indigenous knowledge → made in (as a member of) national standards development in water safety [contact me for details]

Thank you.

Questions / Constructive comments?

- Direct Hire & Scheduling meetings: <https://tinyurl.com/yas7trzy>
- Contact for collaborations via/with univ.: r.rovetto@unsw.edu.au
- Contact for other (e.g., indep consultant): ontologos@protonmail.com
- Project Description: <http://purl.org/rrovetto/ethics-of-ontology>