Ontology 101: An Introduction

Lyle D. Burgoon, Ph.D.

Leader, Bioinformatics and Computational Toxicology

Environmental Laboratory

The view and opinions expressed are those of the author and not those of the US Army or any other federal agency.

AOPXplorer

U.S.ARMY

US Army Corps of Engineers

ERDC
Engineer Research and Development Center

If You Remember Nothing Else, Remember This:

- Ontologies are a way to represent our knowledge on a specific topic

If You Remember 2 Things, Remember:

- Ontologies are a way to represent our knowledge on a specific topic

- Ontologies allow us to share information using a common language

If You Remember 3 Things, Remember:

- Ontologies are a way to represent our knowledge on a specific topic
- Ontologies allow us to share information using a common language
- Ontologies help computers “understand” a subject and apply logic

Today's Goal

- To give you background on ontologies so that you can understand why you care about them, what they are, and how they are built

Are We Talking Philosophy or Computer Science?

- Strictly speaking, when we speak of ontologies here, I mean in the computer science sense
- Ontology is a core and critical area of philosophy
  - Specifically metaphysics (describing what exists and categories of existence)
- Computer science borrowed the concept of ontology from philosophy, but put its own spin on it
- In computer science, ontologies originated in the artificial intelligence community
  - Computers needed to understand human logic and decision-making

Let's Set The Stage

- Some of us have heard of the term "ontology"
- Most biologists who have heard of an ontology heard of the "Gene Ontology"
- Forget anything you know about ontologies – you are now a blank slate

So, What Is An Ontology?

- A representation of knowledge
  - A model of knowledge
- A means to describe concepts and their relationships in a way that a computer can use that information

How About Something More Concrete?

- I want to create a computer program that can order diet-appropriate pizzas for me
- To do that, the computer needs to know what a pizza is
- Think of the computer as a young child – how would you explain to them what a pizza is?

Pizza Defined

- Pizza
  - Has a base (we can argue about whether or not yeast-risen dough is a requirement another day)
  - May have sauce (sauce is optional)
  - Has at least one topping
    - Toppings may be cheese, fruits, vegetables, meats
  - Baked in an oven

A Pizza Ontology?

- We've defined a pizza
- But wait – there are lots of terms we didn't define
  - Base
  - Sauce
  - Toppings
  - Baked
  - Oven
- We go through the same process, defining each of these terms, and any other new terms

Wait, Wait!

- This sure looks like a rabbit hole…
- When/where do I stop defining and describing concepts? This could go on forever
- Toppings: how do I define spinach?
  - Is it enough to say it's a vegetable?
  - Do I need to specify that it's a flowering plant, it's an Amaranthaceae, or that it's native to Asia?
  - Do I need to include it's high in iron and calcium (although both may be difficult to absorb)?
  - What about the fact that Popeye seems to love it only when he or someone is in trouble (but can't be bothered to eat it otherwise)?

Fit for Purpose

- Question: When do you stop adding details?
- Answer: Only add in those details that are necessary for you to meet your goals

Example

- My program needs to understand pizza dietary restrictions
  - Vegan
  - Vegetarian
  - No dairy
  - No fish
  - No pork
  - No vegetables

- Fit For Purpose Ontology
  - Does knowing the anthropological history of spinach help the computer make informed decisions about dietary restrictions?

Let’s Talk Types of Pizzas

- We’ve defined a pizza:
  - Has a base
  - Optional sauce
  - Has at least one topping
  - Baked in an oven

- Types of pizzas
  - Vegetarian
  - Supreme
  - Meat lovers
  - Fungus delight
  - Margarita pizza

Let's Explore This Type/Subclass Thing Some More

- Vegetarian pizza
  - All qualities of a pizza
  - Toppings are of type vegetable, cheese is optional
  - Sauce is optional

- Sweet, this is our vegetarian pizza…

Is This A Vegetarian Pizza?

- The Supreme (sauce + cheese, too)

Topping	Type
Onion	Vegetable
Green Bell Pepper	Vegetable
Olive	Vegetable
Sausage	Meat
Pepperoni	Meat

Is This A Vegetarian Pizza?

- The Supreme (sauce + cheese)

Topping	Type
Onion	Vegetable
Green Bell Pepper	Vegetable
Sausage	Meat
Pepperoni	Meat

- Vegetarian Pizza Criteria

Criterion	Yes/No/Optional/ Silent
Vegetable	Yes
Cheese	Optional
Sauce	Optional
Meat	Silent

The Open World Assumption

- The Supreme (sauce + cheese)

Topping	Type
Onion	Vegetable
Green Bell Pepper	Vegetable
Sausage	Meat
Pepperoni	Meat

- Vegetarian Pizza Criteria

Criterion	Yes/No/Optional/ Silent
Vegetable	Yes
Cheese	Optional
Sauce	Optional
Meat	Silent

Closing the Loophole

- Vegetarian pizza
  - All qualities of a pizza
  - Toppings of type vegetable, cheese is optional
  - Toppings cannot be meat
  - Sauce is optional
  - Sauce cannot be a meat sauce

Is This A Vegetarian Pizza?

- The Supreme (sauce + cheese)

Topping	Type
Onion	Vegetable
Green Bell Pepper	Vegetable
Sausage	Meat
Pepperoni	Meat

- Vegetarian Pizza Criteria

Criterion	Yes/No/Optional/ Silent
Vegetable	Yes
Cheese	Optional
Sauce	Optional
Meat	**No**

Pizzas Are Great, But…

- Let's move to something a little more relevant to our topic at hand
- What this will be:
  - A means to explore the process I use when designing an ontology
- What this won't be:
  - A prescription for how to design an ontology for zebrafish, toxicology, developmental toxicology, etc…

Before We Begin

- I want you to think of the ontology we're going to start hashing out in the next several slides as a blueprint
- You are the architect!
  - That's actually what we call people who design high level blueprints like ontologies for large systems
- In computer speak, what we are doing is putting together the "classes" – or the blueprints – that model what things we need to understand, and how different parts relate to each other
  - Kind of like how a blueprint for a house shows you where the windows are in relationship to the kitchen, and where the sink is in relation to the shower, tub, and toilet

Design Step 1

- Ask what the purpose or goal of the ontology is
  - Is this ontology going to help computers perform an isolated, specific type of task? If so, what is the task?
  - Is this ontology going to be used by other ontologies as a source of expert information?

Design Step 2

- Start thinking about, and listing, all of the “nouns” in the field
- Don’t worry if you don’t get everything
- The next step will help you build out

Design Step 3

- One noun at a time, break down the important parts, and identify what makes that noun what it is, identify relationships between nouns

Step 4

- Repeat Steps 2 and 3

Exercise

- Step 1: Purpose – integrate behavioral data from zebrafish assays

Exercise

- Step 2: Think about and list the “nouns”
  - I’m looking at zebrafish behavioral assays, in males and females, following exposure, for some time, to some chemical (ignoring mixtures for now to keep it simple)

Exercise

- Step 2: Think about and list the “nouns”
  - I’m looking at zebrafish behavioral assays, in males and females, following exposure, for some time, to some chemical (ignoring mixtures for now to keep it simple)
- Some nouns
  - Zebrafish, tanks/chambers, chemical, sex, time, concentration, acclimation time, study site, IACUC approval number, optokinetic reflex, brain morphology, potentiated startle, impaired habituation

Exercise

- Step 3: Break down the important parts of each noun, identify what makes the noun what it is, identify relationships between nouns

- Zebrafish
  - Has_Sex {male, female, intersex}
  - Has_Age_At_Exposure {number greater than 0 in days}
  - Has_Exposure_Duration {number greater than 0 in hours}
  - Has_Pathology some pathologies {0 or more pathologies}

Exercise

- Pathology
  - Defined: some adverse event
  - Subclasses
    - Behavioral
      - Impaired habituation
      - Potentiated startle
      - Reduced locomotion
      - Memory deficit
      - Lack of optokinetic reflex
    - Morphological
      - Brain (has_organ {brain}, disjoint with all other organs)
        - Adverse morphology of amygdala
        - Adverse morphology of habenula

Exercise

- Sex
  - Male
    - Has_gonad {testes}, disjoint with has_gonad {ovary}
  - Female
    - Has_gonad {ovary}, disjoint with has_gonad {testes}
  - Intersex
    - Has_gonad {testes} and has_gonad {ovary}

Once We Have Our Ontology And All The Parts…

- We Test!
- Real fish with real data (or fake fish with fake data) are used to test out this ontology to see what we forgot, or what we might want to model a different way
- Our real/fake fish with real/fake data are called “individuals”

What If We Forgot Something?

- It's fairly common that I forget about an "-icity", some toxicity or pathology that I didn't think of
- That's okay – I just extend my ontology.
  - Add it and move on
- I've never built a perfect ontology in my life
  - It's not uncommon to go back to the drawing board and start from scratch
  - It's also not uncommon for this to take much longer than you ever imagined

So This Was All Abstract and Cool, But…

- You want to do this for real? So a computer can actually use it?
- I use Protégé (http://protege.stanford.edu/) to put together my ontologies
- I make my ontologies in a language called OWL (Web Ontology Language)

If You're Fired Up and Want To Participate…

- There are lots of ontology projects out there
  - ► And lots of philosophies on how to build an ontology

- AOP-Ontology project (https://github.com/DataSciBurgoon/aop-ontology)

- Make sure you talk with the ontology community coordinators for an ontology you would like to contribute to to find out their rules for engagement

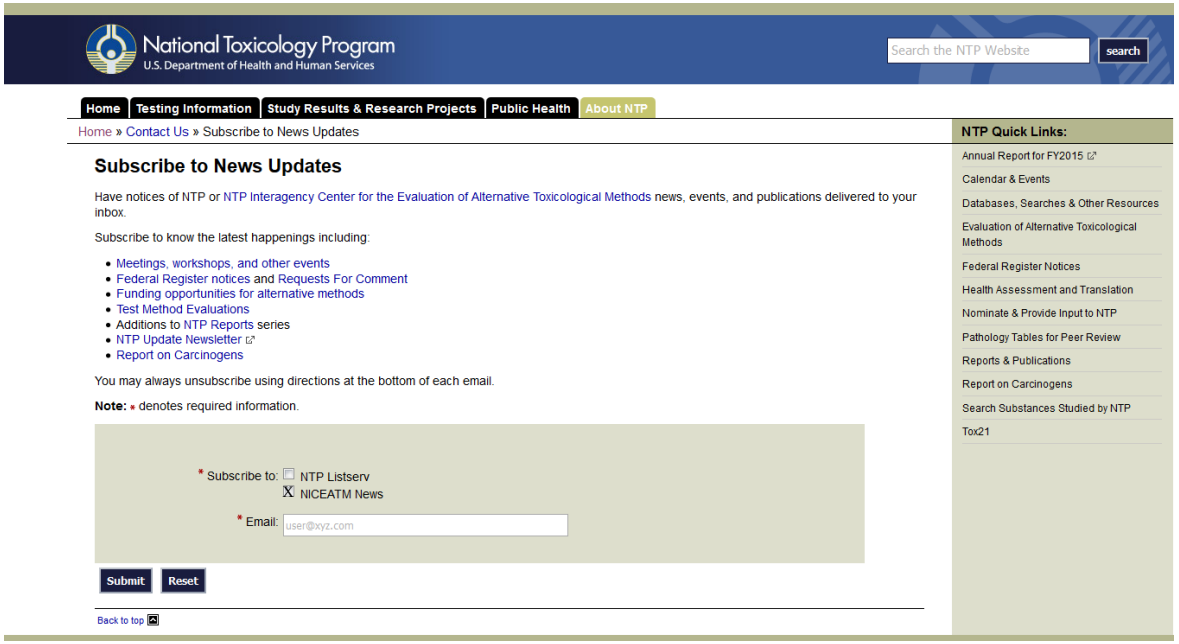
Reasoning

- We did not discuss reasoning. That's coming up in a subsequent webinar
- That's where ontologies become really cool and useful and neat

NICEATM News

For updates on the SEAZIT project and other activities related to *in vitro* alternatives, subscribe to the NICEATM News email list.

- To subscribe to the NICEATM News email list, go to: https://tools.niehs.nih.gov/webforms/index.cfm/main/formViewer/form_id/361
- Check the NICEATM News box and click submit



Thanks!

- Email: lyle.d.burgoon@usace.army.mil
- Twitter: @DataSciBurgoon
- Github: https://github.com/DataSciBurgoon/