A biomedical ontology of wetlab procedures

Akihila Moturu[1], Julia Daniel[1], Gaby Steiner[2], Eric Cramer[2]

[1]Stanford University School of Engineering, Department of Computer Science; [2]Stanford University School of Medicine, Department of Biomedical Informatics

Background and Motivation

- One of the most serious issues currently facing the scientific community is faulty experimental design.
- There are many different types of biomedical wet-lab and clinical experimental methods, each of which yields different data with a specific scope of soundly derivable conclusions.
- There have been a few efforts to standardize protocols and build ontologies that can generate workflows for various experimental needs.
- **We sought to create an educational tool to help researchers and students learn how to design biomedical wet lab workflows that are best suited to their research questions.**
- The ultimate goal is producing *sound* research and facilitating systematic methodology selection.

Methods

- Develop an ontology of wet lab techniques and tools commonly found in a university setting that are most relevant to research in molecular/cell biology.
- Include names of standard experimental methods, what type of species is required for each method (protein, DNA, RNA, etc.), what types of data the methods produce, and what types of data can be used to support specific conclusions.
- Utilizes top-bottom ontology design, classes are not disjoint

PSM

- A Python program builds the ontology as a data structure.
- User completes a series of questions about their hypothesis and intended analysis.
- Uses forward chaining to reason over research question requirements.

Evaluation

- Method 1: Ontology and tool presented to students who were asked to:
  - Run the tool using a hypothesis of their choice, and check the tool's output against their domain knowledge given a standardized input.
    - "Good" result if the tool and the expert are in agreement.
    - "Bad" result if non-analogous deviations are proposed by the tool.
  - Run the tool using a hypothesis we gave them, and verify that they land on the expected output.
- Method 2: Gather a collection of published studies that are marked "good" or "bad" (eg. sound or retracted)
  - Hypothesis and methods extracted from the paper.
  - Based on extracted information, steps in paper are compared to steps proposed by tool.
  - Tool "classifies" paper as either good or bad, number of correct classifications measured.

Figures:
1. Branch of ontology containing Protein Experiments
2. Experiment Navigator interface screenshot

3-4. Two students testing the Experiment Navigator (photos used with consent)

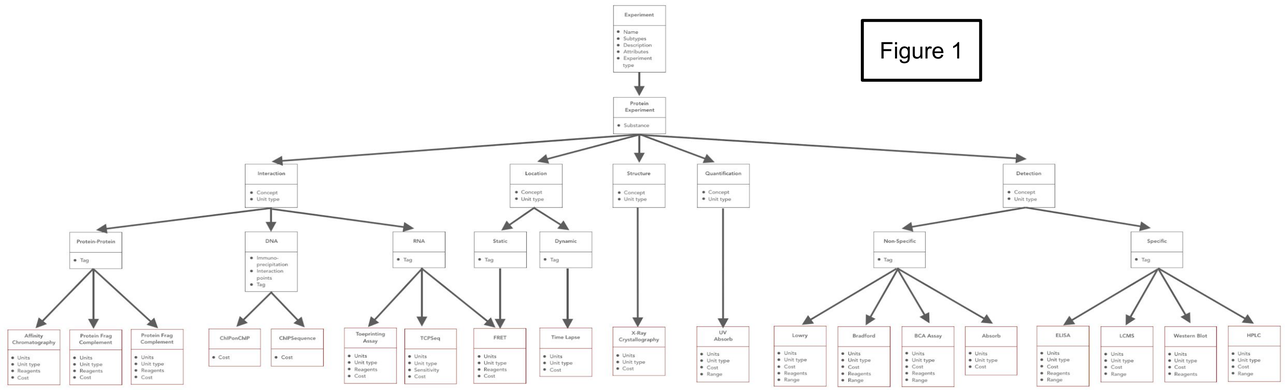
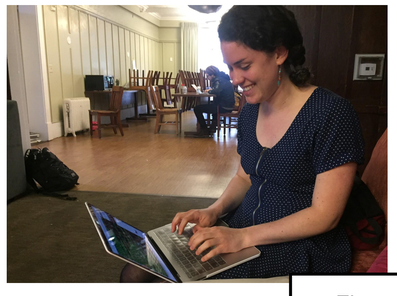
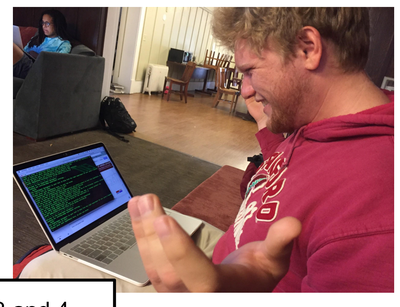


Figure 1

```
* ~ * ~ * ~ * ~ * ~ * ~ * ~ * ~ * ~ * ~ * ~ *
Welcome to the Experiment Navigator!
Loading our cutting-edge experiment ontology...
...done!
The Experiment Navigator is your tool to help find the right experiment
for you, given a hypothesis you'd like to test.
Please enter your hypothesis here, in your own words:
 > RNA polymerase interacts with RNA in the cytoplasm
RNA polymerase interacts with RNA in the cytoplasm
Great! In order to determine which experiment(s) you should conduct to t
est this hypothesis, we need to know more about the substances and infor
mation involved.
Is the molecule you're working with a type of protein? yes
What do you want to learn about the protein? Please type one of the foll
owing: location, structure, interaction, quantification, detection. inte
raction
What type of molecule is the protein interacting with? DNA, RNA, or anot
her protein? RNA
Do you know what reagants you will be using for the experiment? no
Do you need high sensitivity? no
The type of experiment that you want to perform is FRET!
```

Figure 2





Figures 3 and 4

Results

- We developed a tool that correctly navigated our ontology to identify the most appropriate experimental technique for a particular hypothesis.
- We tested the tool on 6 users, all students with varying levels of background knowledge in biology.
- 5/6 students correctly navigated to the right experiment type output when given a specific hypothesis to input.
- Additionally, 3 students tested their own hypotheses, and all three received the expected output or agreed with the output the Experiment Navigator showed.

Discussion

- This tool, we ultimately think, can be an effective pedagogical tool when integrated with an experimental design curriculum.
- Our ontology is relatively extensive, which made it challenging to implement. We learned a great deal by attempting multiple different knowledge representation strategies, particularly Protege (using OWL) and Python classes with inheritance. Each strategy had pros and cons in implementation.
- Our problem-solving method is rule-based, and manually curated for now. We would have liked to further automate the PSM, but our approach shows a successful proof of concept.

Future Directions

- This tool was initially conceived as a way for researchers to design studies or to verify that their work conforms to standard or sound practices.
- Pivoted to become an educational tool for informing researchers how to design their projects.
- Can expand to include more wet lab techniques or species of interest.
- Can expand to include more research domains (eg. clinical or psychological research).

References

1. Russell & Norvig 2009, p. 337, https://www.ics.uci.edu/~rickl/courses/cs-171/aima-resources/Artificial%20Intelligence%20A%20Modern%20Approach%20(3rd%20Edition).pdf
2. Atlantic article publicizing John Ioannidis's research in PLOS One. https://www.theatlantic.com/magazine/archive/2010/11/lies-damned-lies-and-medical-science/308269/
3. FAIR data principles: https://www.force11.org/group/fairgroup/fairprinciples
4. Hayes-Roth, Frederick; Donald Waterman; Douglas Lenat (1983). Building Expert Systems. Addison-Wesley. ISBN 0-201-10686-8
5. Marin D. Simina et al. "Opportunistic Reasoning: A Design Perspective" in Proceedings of the Seventeenth Annual Conference of Cognitive Science edited by Johanna D. Moore, 1995 ISBN 0-8058-2159-7, page 78