Towards an Ontology for Propaganda Detection in News Articles

Kyle Hamilton Early stage PhD Technological University Dublin ML-Labs—Science Foundation Ireland kyle.i.hamilton@mytudublin.ie

ESWC 2021 PhD Symposium

Problem – The Propaganda Supply Chain

- Related work and State of the Art
- Approach
- Methodology
- Evaluation Plan



DEMAND

"Americans gorge themselves daily on empty informational calories, indulging their sugar fixes of self-affirming half-truths and even outright lies."

> Chris Stirewalt (former FOX NEWS reporter) January 2021

CONSEQUENCES



• Problem

• Related work and State of the Art

- Approach
- Methodology
- Evaluation Plan

Related work and State of the Art

Propaganda technique identification

Span identification

Rank	Team	F1	Precision	Recall
1	Hitachi	0.51551	0.56544	0.47368
2	ApplicaAl	0.49153	0.59954	0.41650
3	aschern	0.49100	0.53232	0.45564

Technique identification

Rank	Team	F1	F1 Appeal_to_Authority	F1 Appeal_to_fea
1	ApplicaAl	0.62067	0.48148	0.45490
2	aschern	0.62011	0.35294	0.41778
3	Hitachi	0.61732	0.40000	0.38938

https://propaganda.qcri.org/semeval2020-task11/ leaderboard.php

False information ("Fake News") detection

- Style based
- Knowledge based
- Network based
- Human in the loop

Туре	F1	Dataset
Network	0.98	Modified Twitter15
Network & Style	0.93	PolitiFact
Network	0.88	PolitiFact

Argument mining (AM)



There is a lot of uncertainty involved in argumentation. How to close this gap is an open research challenge!

- Problem
- Related work and State of the Art

• Approach

- Methodology
- Evaluation Plan

Propaganda:

Any information which is intended to influence beliefs or modify behaviors in order to further an agenda.

Propaganda techniques

1. Presenting Irrelevant Data (Red Herring)	10. Exaggeration or Minimisation
2. Misrepresentation of Someone's Position (Straw Man)	11. Flag-waving
3. Whataboutism	12. Doubt
4. Causal Oversimplification	13. Appeal to fear/prejudice
5. Obfuscation, Intentional vagueness, Confusion	14. Slogans
6. Appeal to authority	15. Thought-terminating cliché
7. Black-and-white Fallacy, Dictatorship	16. Bandwagon
8. Name calling or labeling	17. Reductio ad hitlerum
9. Loaded Language	18. Repetition

Propaganda techniques – example 1



In a glaring sign of just how ⁴⁰⁰ stupid and petty⁴¹⁶ things have become in Washington these days, Manchin was invited on Fox News Tuesday morning to discuss how he was one of the only Democrats in the chamber for the State of the Union speech ⁶⁰⁷ not looking as though Trump ⁶³⁵killed his grandma⁶⁵³.

9. Loaded Language

10. Exaggeration and Minimization

Dependency Parse Tree (via SpaCy python library)



Propaganda techniques – example 2

The event, called the Maunder Minimum, is credited with plunging temperatures to the point where the Thames River and Baltic Sea froze over. However, other parts of the planet heated up, including Alaska and Greenland, to far above their normal highs. This temporary shakeup of temperatures could happen again, though the ²²³⁵researchers say it will have very little effect on the overall global warming trend that mankind has created for itself.²³⁵⁴

10. Appeal to authority

Constituency Parse Tree

(via Berkeley Neural Parser https://github.com/nikitakit/self-attentive-parser)



Possible approaches – stylistic

Model each propaganda technique by its logical structure.

Appeal to Authority/Testimonial X is an expert X makesClaim Y ∴ Y



Model each technique as a Class



Harris, R. A., Marco, C. D., Mehlenbacher, A. R., Clapperton, R., Li, I., Ruan, S., & O'Reilly, C. (n.d.). *A Cognitive Ontology of Rhetorical Figures*. 8. 1

Possible approaches – multidisciplinary

An ontology that models the current socio-political context in relation to propaganda.

Domains:

- 1. Rhetoric
 - Propaganda Techniques
- 2. Psychology
 - Cognitive Biases
- 3. Social/Political Science
- 4. Behavioral Economics

General Steps:

Term excerption -

- From literature ex) via textRank
- Prioritize*

Taxonomy -

- Identify categories*
- Identify hierarchical structure*

Ontology -

- Identify relationships*
- Identify inference rules*
- * With the assistance of subject matter experts

Research Question

To what extent can reasoning improve the accuracy and explainability of propaganda detection in news articles?

Hypothesis

By utilizing techniques from Semantic Web Neuro-Symbolic Reasoning and Argument Mining, and a theoretical foundation from the social sciences, the accuracy, and interpretability of propaganda detection can be improved over existing deep learning methods such as the SEMEVAL-2020 top ranked results.

- Problem
- Related work and State of the Art
- Approach
- Methodology
 - Evaluation Plan

Objectives for ontology development

- 1. Define domain of ontology: *Propaganda techniques in news text*.
- 2. Define the purpose of the ontology: The identification of propaganda in news text with description logic.
- 3. Write competency questions what types of questions should the information in the ontology provide answers for? ex) Does this news article contain propaganda techniques?
- 4. Review/reuse existing ontologies:
 - a. The Rhetorical Figure Ontology Project;
 - b. Computational Rhetoric Project at the University of Waterloo, organized around a comprehensive OWL based ontology of rhetorical figures, RhetFig;
 - c. the work of Mitrovic et al on Ontological representations of rhetorical figures
- 5. Enumerate important terms in the ontology start with the 18 propaganda techniques in PTC. Expand with terms from the literature concerning propaganda techniques.
- 6. Define the classes and the class hierarchy *data-driven*, *bottom up approach*.
- 7. Define the properties of classes—slots *data-driven, bottom up approach*.
- 8. Define the facets (restrictions) of the slots (cardinality, data-types, domain/range).
- 9. Create instances use the PTC dataset as a first iteration.
- 10. Evaluate and debug see next slide.

- Problem
- Related work and State of the Art
- Approach
- Methodology

• Evaluation Plan

Evaluation Plan for the Ontology

Frame of reference (baseline):

Since this is a new ontology, what constitutes a gold-standard in this case is unclear, but again, this is early stages, and such a gold-standard may yet emerge from a more thorough understanding of ontology engineering. In the meantime, the following frames of reference can be utilized:

Data-driven - This is a measure of coverage based on the corpus. Currently, this is the PTC corpus as described earlier.

Assessment-by-humans - Experts in behavioral/social sciences will be recruited to augment the data-driven approach in determining domain coverage. Experts in ontology engineering will be recruited to help assess semantic quality and reasoning capability.

Evaluation goals:

Domain coverage - This is an iterative process where the ontology will be compared with the proposed social context mapping and vocabulary.

Quality of modeling - Reasoner software to be determined. Human assessments of syntactic, structural, and semantic quality may also be employed.

Suitability to application/task - This will be determined by the quality of the results from the overall software pipeline, based on precision/recall of the identified propaganda techniques, as well as the ability to answer the competency questions.

Evaluation Plan for the Prediction Task

Using metrics from SEMEVAL-2020 compare prediction results to leaderboard. See figure on next slide.



Evaluate the full pipeline for the prediction task using F1 as per SEMEVAL.



To evaluate the preprocessing and knowledge representation stage, implement best semeval BERT based solution to compare results using F1 as per SEMEVAL.

Application flow diagram



Limitations

Public data only

- No news articles behind paywalls
- No private social media

English text only

- No images (i.e. Instagram)
- No audio/visual media (i.e. Youtube).

Proposed **Contributions**

- 1. An ontology for propaganda detection
- 2. An application for the identification of propaganda techniques in unstructured text.

The ability to identify propaganda and disinformation can be utilized in the development of media literacy projects such as the BBC's *Evidence Toolkit*.

The addition of a suitable ontology and social context mapping for detecting propaganda techniques can help further research in this area, and more broadly in the identification of rhetorical figures in unstructured text for Argument Mining for example.