# Automating Comment Moderation: Topics and Toxicity in Online News

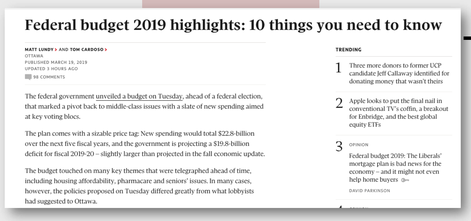
SFU

Vagrant Gautam • Maite Taboada

Discourse Processing Lab

## 1. Online News

### Articles



Federal budget 2019 highlights: 10 things you need to know

MATT LUNDY > AND TOM CARDOSO >
OTTAWA
PUBLISHED MARCH 19, 2019
UPDATED 3 HOURS AGO
99 COMMENTS

The federal government unveiled a budget on Tuesday, ahead of a federal election, that marked a pivot back to middle-class issues with a slate of new spending aimed at key voting blocs.

The plan comes with a sizable price tag: New spending would total $22.8-billion over the next five fiscal years, and the government is projecting a $19.8-billion deficit for fiscal 2019-20 – slightly larger than projected in the fall economic update.

The budget touched on many key themes that were telegraphed ahead of time, including housing affordability, pharmacare and seniors' issues. In many cases, however, the policies proposed on Tuesday differed greatly from what lobbyists had suggested to Ottawa.

TRENDING

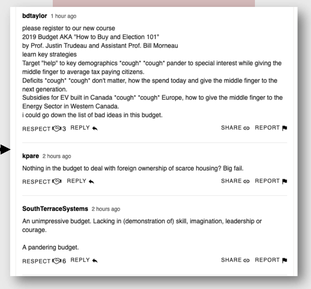
1 Three more donors to former UCP candidate Jeff Callaway identified for donating money that wasn't theirs

2 Apple looks to put the final nail in conventional TV's coffin, a breakout for Enbridge, and the best global equity ETFs

3 OPINION Federal budget 2019: The Liberals' mortgage plan is bad news for the economy – and it might not even help home buyers

DAVID PARKINSON

### Comments



bdtaylor 1 hour ago
please register to our new course
2019 Budget AKA "How to Buy and Election 101"
by Prof. Justin Trudeau and Assistant Prof. Bill Morneau
learn key strategies
Target "help" to key demographics "cough" "cough" pander to special interest while giving the middle finger to average tax paying citizens.
Deficits "cough" "cough" don't matter, how the spend today and give the middle finger to the next generation.
Subsidies for EV built in Canada "cough" "cough" Europe, how to give the middle finger to the Energy Sector in Western Canada.
I could go down the list of bad ideas in this budget.
RESPECT 3 REPLY SHARE REPORT

kpare 2 hours ago
Nothing in the budget to deal with foreign ownership of scarce housing? Big fail.
RESPECT 6 REPLY SHARE REPORT

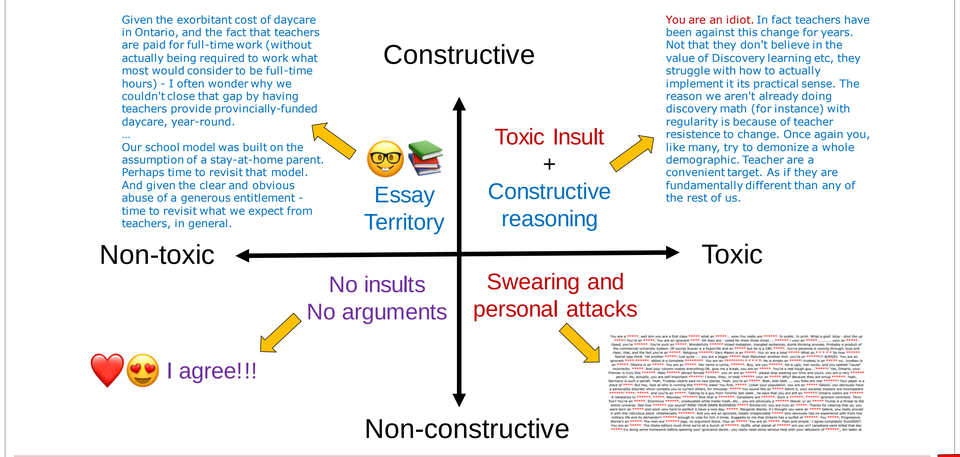
SouthTerraceSystems 2 hours ago
An unimpressive budget. Lacking in (demonstration of) skill, imagination, leadership or courage.
A pandering budget.
RESPECT 6 REPLY SHARE REPORT

- Online publications with comments sections
- Human moderators hired to monitor comments
- Positive moderation: Highlight good arguments
- Negative moderation: Delete profanity, personal attacks

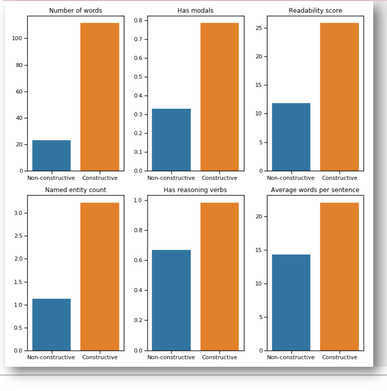
THE CONVERSATION
THE GLOBE AND MAIL
THE TYEE NEWS. CULTURE. SOLUTIONS.

Questions:
1. What makes comments **linguistically** different?
2. Can we use **computer science** to automate moderation?

## 2. Taxonomy of the Online Comment



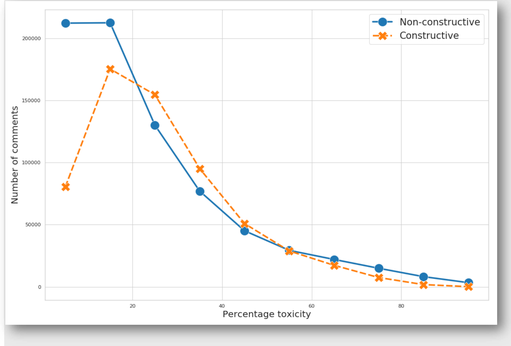

### Quantifying constructiveness and toxicity




- **Constructiveness system** developed at SFU by Varada Kolhatkar and Maite Taboada
- Comments classified for constructiveness using a number of linguistic features, some of which are shown on the left
- **Toxicity system** developed by Google: Perspective API
- Uses machine learning (RNNs with attention)

Google Developers

## 3. Toxicity and Constructiveness




- Results on 1,000,000+ online comments
- Highest frequency comments are non-constructive and non-toxic
- Very few are highly toxic
- Non-constructive comments more common (intuitive)

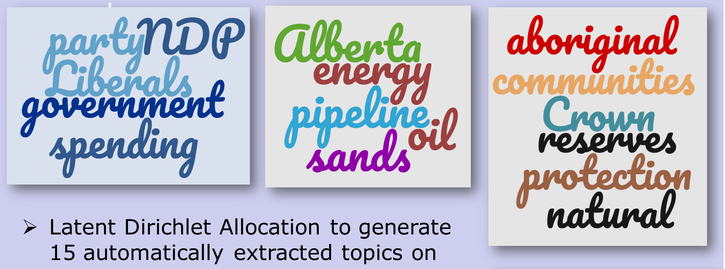
## 4. Topic Modelling on Articles and Comments

Going beyond toxicity and constructiveness...

- **What subjects generate the most comments?**
- **What subjects generate the most toxic comments? Or the most constructive comments?**

**TOPIC MODELLING** Statistical modelling technique to automatically extract the subjects discussed in a text

### Example topics




- Latent Dirichlet Allocation to generate 15 automatically extracted topics on 10,000+ news articles and 1,000,000+ comments
- Each text is given probabilities for all 15 topics – this allows an article to be 50% about a new pipeline and oil, and 50% about the impact on nature and nearby First Nations communities

- **Hypothesis 1**
  People would talk a lot about politics; more comments on articles classified as "politics" articles
- **Hypothesis 2**
  People would be significantly more toxic about certain topics, e.g., issues of abortion, race, atheism

### Results

**Highest frequency words**


**Constructive comments**


**Non-constructive comments**

- Most frequent words across all comments regardless of toxicity and constructiveness: Harper, time, people, government, Canada
- **Hypothesis 1 confirmed**

  Most comments about politics and on politics articles; people talk more than anything else about politics
- **Hypothesis 2 rejected**

  Roughly the same proportion of toxic commenters in every comment section

### Future work

- Do toxicity and constructiveness propagate in threads?
- Add sentiment to the taxonomy
- Do people with anonymous usernames write different comments from people with their real names? More toxic? Less constructive?

### References

[1] Cao, N., & Cui, W. (2016). *Introduction to Text Visualization*. Atlantis Press.
[2] Kolhatkar, V., & Taboada, M. (2017). *Constructive Language in News Comments*. Proceedings of the First Workshop on Abusive Language Online, ACL. Vancouver. 11-17.
[3] Kolhatkar, V., & Wu, H., & Cavasso, L., & Francis, E., & Shukla, K., & Taboada, M. (2018). *The SFU Opinion and Comments Corpus: A Corpus for the Analysis of Online News Comments*.
[4] Rubin, T. N., & Chambers, A., & Smyth, P., & Steyvers, M. (2011). *Statistical Topic Models for Multi-Label Document Classification*. Mach Learn. 88.
[5] Newman, D., & Bonilla, E., & Buntine, W. (2011). *Improving Topic Coherence with Regularized Topic Models*. NIPS. 496-504.
[6] Röder, M. & Both, A. & Hinneburg, A. (2015). *Exploring the Space of Topic Coherence Measures*. WSDM 2015 - Proceedings of the 8th ACM International Conference on Web Search and Data Mining. 399-408.

### Acknowledgements

Thanks to Fatemeh Torabi Asr for her guidance when I hit a topic modelling wall, to Varada Kolhatkar for her beautiful code, to all the people behind the corpora, NLP systems and open-source Python and R libraries I used. Thanks to Yue Wang for the original poster template. Thanks to Ashley Farris-Trimble for telling Maite to take me on!

Funding: Key Big Data Undergraduate Student Research Award (USRA)