

FIRST ACL WORKSHOP ON ETHICS IN NLP  
4 APRIL 2017, VALENCIA SPAIN

# GENDER AS A VARIABLE IN NLP

© 2017 Brian N. Larson, JD PhD – [www.Rhetoricked.com](http://www.Rhetoricked.com) – @Rhetoricked

My name is Brian Larson.

- I'm presently assistant professor at Georgia Institute of Technology in Atlanta.
  - I'm moving this summer to the Texas A&M's law school in Fort Worth.
  - Before academia, I was a lawyer for a dozen years, practicing in information technology and intellectual property law.
  - I am not an ethicist by training, but rather a rhetorician (though those two fields have had close relations for some 2500 years).
  - My interests now include digital media law, communication ethics, and research methodology.
  - My slides and supporting notes are available on my website, for which the URL [rhetoricked.com](http://rhetoricked.com) appears on each slide.
- I'm going to start by briefly summarizing my paper. But then I'd like to move on to think in broader terms how to put the work of this workshop into a map of research questions for ethics in NLP.

## MY PAPER PRESENTED A SIMPLE OUTLINE

- ▶ Identify some theories of gender
- ▶ Identify some theories of ethics
- ▶ Explain that some studies using gender as a variable didn't do a terrific job of it
- ▶ Offer some guidelines for similar studies in the future

[Just read quickly.]

## MY RECOMMENDATIONS FOR USING GENDER AS A VARIABLE WERE ALSO SIMPLE

[www.Rhetoricked.com](http://www.Rhetoricked.com) -@Rhetoricked

[Just read quickly.]

## MAKE YOUR THEORY OF GENDER EXPLICIT

- ▶ Read and cite someone else's theory, or
- ▶ Offer your own definition



[www.Rhetoricked.com](http://www.Rhetoricked.com) -@Rhetoricked

If you are going to use a variable in research, you should probably be able to say what it measures...  
[Read the rest.] This is Judith Butler, author of an influential perspective on gender.

---

Image: University of California, Berkeley - The photo was submitted personally by Judith Butler, CC0, <https://commons.wikimedia.org/w/index.php?curid=31967265>



## DON'T USE GENDER AS A VARIABLE UNLESS NECESSARY

- ▶ Don't make gender a "throw-away crosstab"

[www.Rhetoricked.com](http://www.Rhetoricked.com) -@Rhetoricked

There is a temptation when gathering data from human participants to ask demographic questions that are not theorized. Gender is a good example. Without thinking about what they are really collecting, researchers ask participants to share their genders; then, when presenting findings, researchers will use this gender value as a way of cross-tabulating their data.

---

Image: © 2011 brownpau "Necessary Room" [flic.kr/p/a8PAkG](https://flic.kr/p/a8PAkG). CC BY 2.0 license.

## MAKE YOUR METHOD OF ASCRIBING GENDER CATEGORIES EXPLICIT

- ▶ Readers should be able reproduce it
- ▶ Critique it



[www.Rhetoricked.com](http://www.Rhetoricked.com) -@Rhetoricked

If you are assigning a social category as a variable to participants in your study—in other words, if you put people into category-boxes—you should tell us how you are doing it. If you employ some sort of automated or qualitative process, you should show us how you did it so that we can reproduce and interrogate your work.

The gold standard should be to allow participants to indicate their own genders. That, too, is fraught with issues.

---

Image: © 2006 Ben Ward “Boxes” <http://flic.kr/p/h2RxE>. CC BY-NC-ND 2.0 license



## RESPECT STUDY PARTICIPANTS

- ▶ If possible, allow participants to indicate their own genders
- ▶ Avoid asking them to identify as “other” or “none of the above”

[www.Rhetoricked.com](http://www.Rhetoricked.com) –@Rhetoricked

Natural language is human language. We have an ethical duty to treat other humans with respect. So, if you ask a transgender or intersex person to self-identify for gender, you should consider how your question exhibits respect or not. Offering two options—male and female—leaves no place for persons who do not identify with one of these genders. Offering three by adding an “other” or “none of the above” still slots transgender and intersex folks into an outsider group.

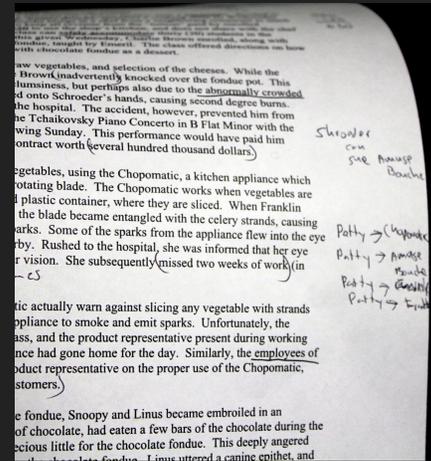
So, those are my substantive guidelines. To encourage researchers to follow them, however, we need to take one further step...

---

Image: © 2008 Sally Payne “Faith 7th AMTG-Saturday Night” <http://flic.kr/p/5ughZn>. CC BY-NC-ND 2.0 license

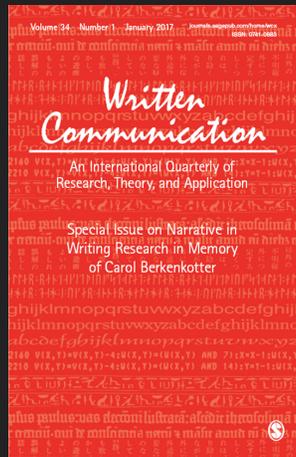
## PEER REVIEWERS: EXPECT THESE BEHAVIORS FROM PAPER/ARTICLE AUTHORS

- ▶ Announce that you are adopting these standards
- ▶ Expect authors to meet them or explain why they do not



[www.Rhetoricked.com](http://www.Rhetoricked.com) -@Rhetoricked

Image: © 2012 Quin Dombrowski "Absurdist prose" <http://flic.kr/p/btvv6W>. CC BY-SA 2.0 license



## LARSON (2016) OBSERVES THESE GUIDELINES

- ▶ Writing by law students (N=197)
- ▶ Theory of gender adapted (after discussion) from Butler
- ▶ Participants self-identified for gender in a free-form field
- ▶ Gender binary observed: Gender F (n=104) and Gender M (n=89). Four did not answer.
- ▶ No meaningful difference observed on “informational-involved” dimension of Biber (1988)

[www.Rhetoricked.com](http://www.Rhetoricked.com) –@Rhetoricked

Go through bullets.

I explained my findings by adapting Wilson and Sperber’s (1995) relevance theory. The students had received a year or so of training in a professional discourse and were asked to solve a problem within that discourse. The men and women performed indistinguishably on Biber’s “informational-involved” dimension, which had been used in previous NLP studies (e.g. Argamon et al. 2003).

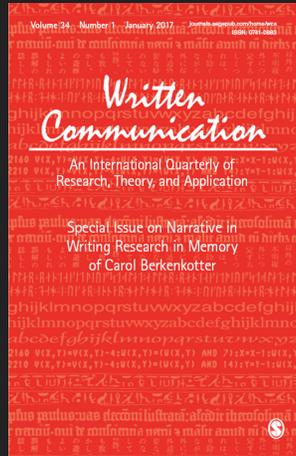
---

Argamon, S., Koppel, M., Fine, J., & Shimoni, A. R. (2003). Gender, genre, and writing style in formal written texts. *Text*, 23(3), 321–346.

Biber, D. (1988). *Variation across speech and writing*. Cambridge U.K.: Cambridge University Press.

Larson, B. N. (2016). Gender/Genre: The Lack of Gendered Register in Texts Requiring Genre Knowledge. *Written Communication*, 33(4), 360–384.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition* (2nd ed.). Wiley-Blackwell.



## LARSON (2016) ALSO MAKES DATA/CODE AVAILABLE

- ▶ Journal is (unfortunately) not open.
- ▶ Pre-publication version of article available on my website
- ▶ Corpus available from LDC
- ▶ Code available on Github

[www.Rhetoricked.com](http://www.Rhetoricked.com) –@Rhetoricked

- Pre-pub version of article: <https://rhetoricked.com/2016/09/24/my-written-communication-article-has-dropped/>
- Corpus at LDC (I can share if you want to collaborate :-): <https://catalog ldc.upenn.edu/LDC2017T03>
- Code on Github: <https://github.com/rhetoricked/WrittenCommunication2016>

## PEER REVIEWS AND OTHER PAPERS RAISED INTERESTING ISSUES

- ▶ Reviewers wanted answers to questions
- ▶ Reviewers distinguished ethics from questions of good vs. bad science

[www.Rhetoricked.com](http://www.Rhetoricked.com) –@Rhetoricked

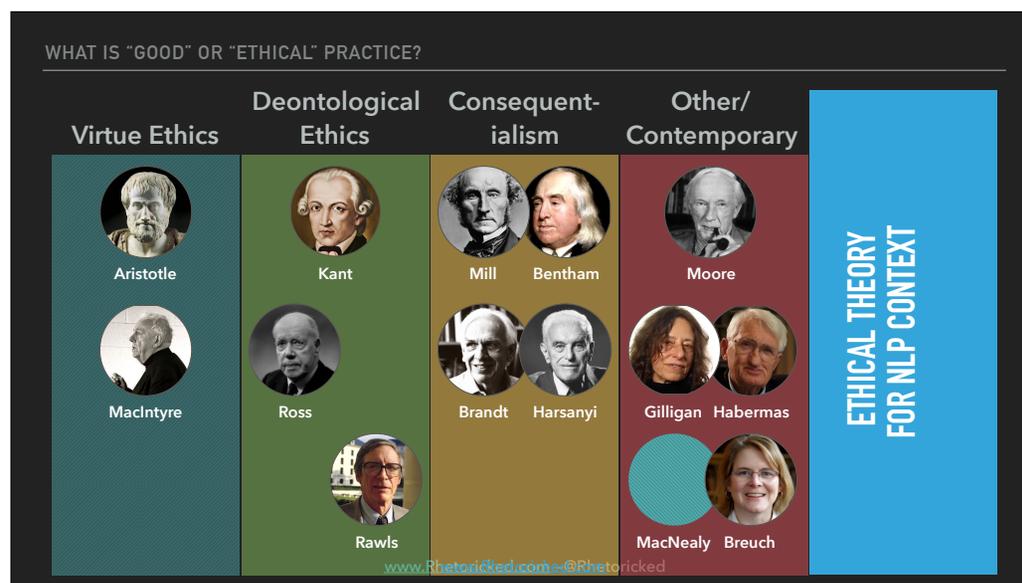
Answers to questions:

- “What *\*is\** a good definition of gender and *\*how should\** one ideally assign gender to participants?” “What theory or framework of gender would you recommend?”
- “What is your proposed solution to [Facebook e.g. of multiplicity of labels]? Should NLP practitioners use all labels available or group them?”
- “...regarding how or if to gender-balance datasets. For instance, Wikipedia written overwhelmingly by male authors ([https://en.wikipedia.org/wiki/Wikipedia:Wikipedians#Number\\_of\\_editors](https://en.wikipedia.org/wiki/Wikipedia:Wikipedians#Number_of_editors)). What effect does that have on NLP research? Should data samples be random, therefore having this same bias, or sampled evenly from different genders in order to avoid models even unwillingly becoming gender biased? What if gender information about authors is not available?”

Ethical vs. bad science:

- “The conceptual sloppiness of defining gender variables in previous NLP-and-gender studies not only raises ethical questions (as addressed here), but also is just bad science.”

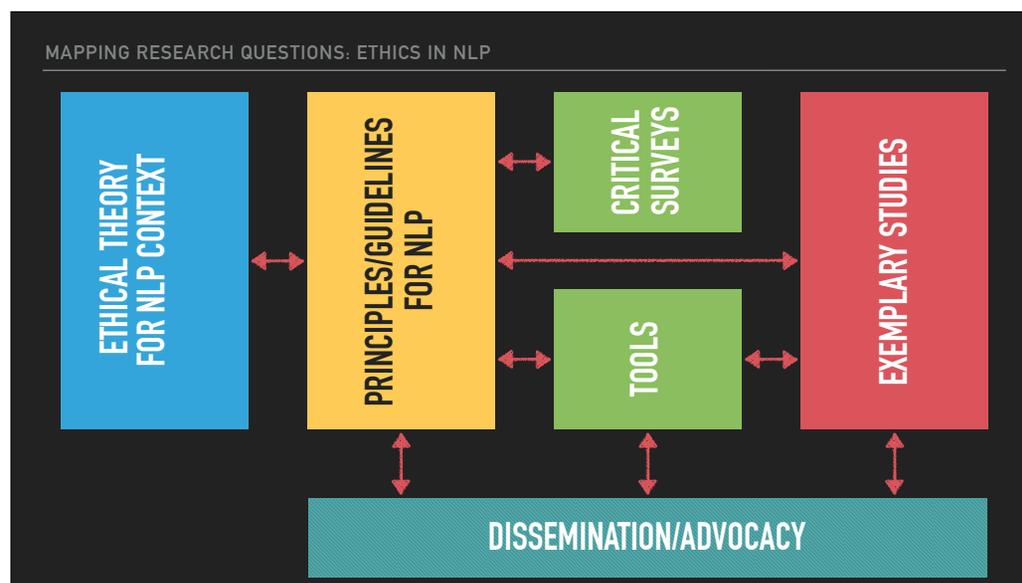
Reading those questions and the other papers for this workshop, it occurred to me that it would be helpful to map the research questions we are talking about onto a research agenda for the field. For me, that begins with a review of the underlying ethical principles we might want to observe...



Ethical thinkers are often divided into ‘camps.’

- Virtue Ethics: What kind of character one should have. Classic is Aristotle’s *Nicomachean Ethics*. Alasdair MacIntyre provides contemporary gloss.
- Deontological ethics: What are one’s duties? What rules must one follow? Classic is Kant. Contemporary glosses by W(illiam) D(avid) Ross (prima facie duties), and John Rawls (veil of ignorance).
- Consequentialism: What are the effects of an action or rule? Mill and Bentham were early advocates of utilitarianism. Richard Brandt and John Harsanyi provided contemporary focus on “rule utilitarianism”.
- My “other” bucket includes non-foundationalists like G(eorge) E(dward) Moore (who challenged the notion that we can define what is “good”); contemporary thinkers like Carol Gilligan (ethics of care in *In a Different Voice*) and Jürgen Habermas (discourse ethics). I’ve also identified a couple contemporary scholars in research methods in written communication, MacNealy and Breuch, who argue that validity and reliability in research are actually ethical concerns. (That is, what makes “good” research?)
- I HAVE noted that the vast majority of these thinkers are white, European males. There is certainly room for more women and for perspectives outside the West.
- Taken together with principles from research ethics provided by the US Common Rule and the Helsinki Declaration, these thinkers provide a basis for developing an ethical theory for the NLP context. But that’s just a first step. If we want to map potential research question for ethics in NLP, we need more...

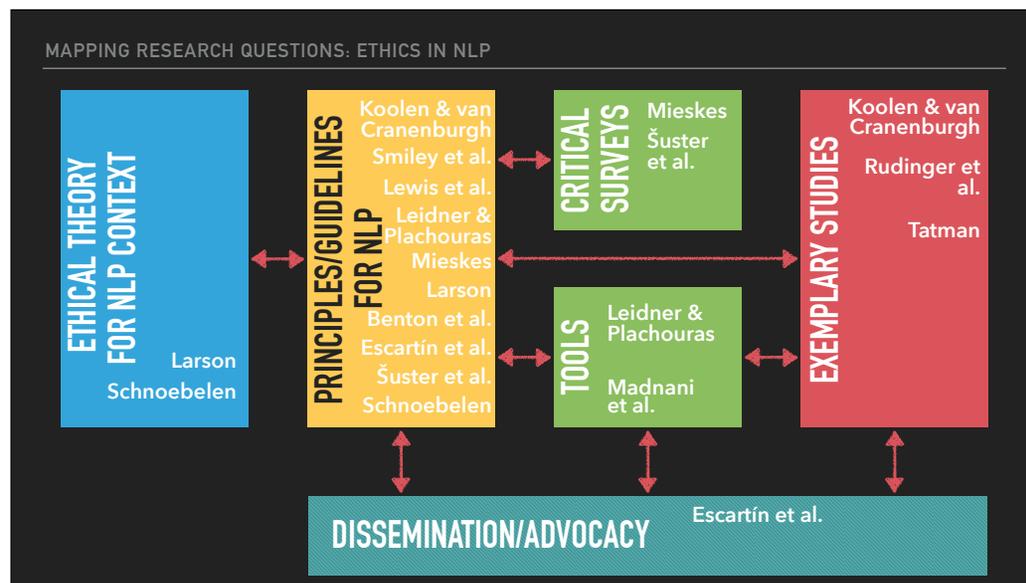
Fair use asserted with regard to images used here.



If we think of the kinds of research questions we can ask regarding ethics in NLP, we can consider the questions and what the objects and products of their study might be:

- **Ethical theory for NLP context.** “What theories are best equipped to provide ethical principles for NLP?” The objects of study here are existing ethical theories and the products are position papers. They enable a move to [CLICK]
- **Principles/guidelines for NLP.** The question: “What ethical principles and guidelines should we apply to NLP contexts?” These questions should result in dialog with ethical theory...thus the bidirectional arrows. The objects of study are the ethical theories and the academic and practitioner contexts of NLP. The products are also position papers. [CLICK]
- **Critical surveys.** We then ask: “Are researchers and practitioners in various NLP specialties employing these guidelines and principles?” The objects of study here are studies conducted by NLP researchers. The products are critical surveys, meta-studies that look systematically at other studies, identifying the “state of the art.” [CLICK]
- **Tools.** “What tools and technology can encourage ethical practices in NLP?” Here, the products are tools grounded in the ethical principles. [CLICK]
- **Exemplary studies.** If we ask, “What does an ethical NLP study look like?”, the responses to this question are reports of studies that are conducted ethically and are reported with attention to ethical concerns. [CLICK]
- **Dissemination/advocacy.** A very important question is “How to we encourage ethical practices in NLP?” The objects of study here might other NLP researchers, studied for example by means of surveys and tests of ethical principles and guidelines. The products of this question may thus include papers discussing findings of such studies, but they should also take the form of advocacy within the NLP community.

Today’s workshop has contributed significantly to answering some of these questions.



I did this plot very quickly, so my apologies to authors whose work I've miscategorized. In principle, for NLP ethics to be a viable subfield, it probably needs to host research on all these topics. Thoughts for the next workshop perhaps.

# THANK YOU!

Slides available at [www.Rhetoricked.com](http://www.Rhetoricked.com)  
Contact: [brian@TendalLarson.com](mailto:brian@TendalLarson.com) – @Rhetoricked