

Ethics in NLP: Standardization, Impact

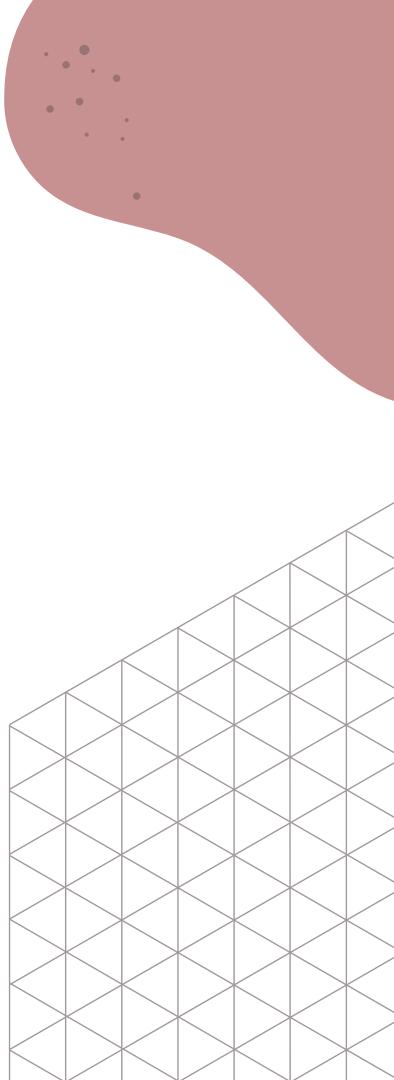
Hana Galijasevic
October 28, 2021



01.

Roadmap

Overview of presentation



Roadmap for today

01

Ethics Crash Course

Brief introduction to basic ethics concepts and ideas

02

NLP & Ethics

How do the two interact?
Historical examples,
overview of peer review
processes in NLP

03

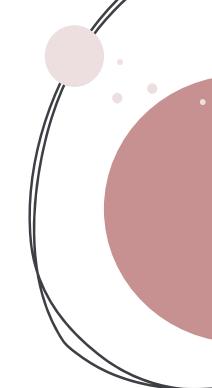
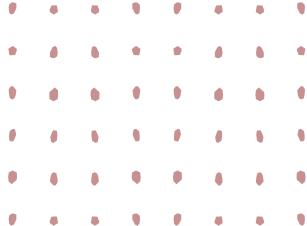
Current Research on Ethical NLP

Proposed solutions to issues, state of the field

04

Required Papers & Discussion

Analysis of the publications, discussion on ideas therein, further reading suggestions



“The L in NLP means language,
and language means people”



-Schnoebelen, 2017

Important Concepts not covered in this presentation

(Harmful) Bias

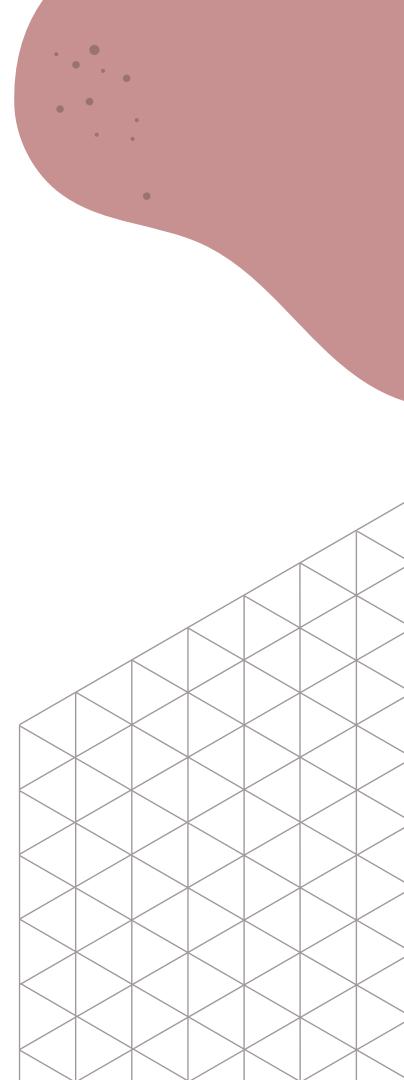
Privacy

etc.

Language
Models as
Ground Truth

Dual Use

Crash Course in Ethical Practices in Research



Why Research Ethically?

Learn from mistakes

Most if not all current practices in human-centered research stem from events and harms that occurred because no such regulations were in place. Examples include: Tuskegee experiments, the Nuremberg trial experiments, Stanford Prison, etc.

Remember the Big Picture

Ultimately, everything we do affects people. Research is only as valuable as the extent to which it benefits humanity (whether through knowledge, opportunity, or immediate physical impact).

Who is benefitting? Who is suffering?

What is ethical research?

Since all research can be tied back to human impact, all research is then human-subject experimentation.

If we assume we are all participating in a large experiment, we need to make sure it meets certain criteria of responsible experimentation:

- **Beneficence** (no harm to subjects, maximize benefits, minimize risk)
- **Respect** for subjects' autonomy (informed consent)
- **Justice** (benefits vs. harms, protection of vulnerable subjects)

These concepts are well-established principles from the Belmont Report. Research is usually conducted after approval from the Institutional Review Board (IRB) that takes these principles and other important factors into account.

What about non-research technology?

We are building technology that affects people in the world, and we therefore have a responsibility towards those people:

- People whose data we collect
- People who will use the technology
- People who will be affected by others' use of the technology

Common Principles of Research Ethics, cont.

- respect for persons (autonomy, privacy, informed consent)
- balancing of risk to individuals with benefit to society
- careful selection of participants
- independent review of research proposals
- self-regulating communities of professionals
- funding dependent on adherence to ethical standards

(Metcalf 2014, p.2.)

Why have a Code of Ethics?

- Provides a framework for ethical judgments within a profession
- Expresses the commitment to shared minimum standards for acceptable behavior
- Provides support and guidance for those seeking to act ethically
- Formal basis for investigating unethical conduct, as such, may serve both as a deterrence to and discipline for unethical behavior

(from UMD Electrical Engineering slides; accessed through nlpers.blogspot.com)



**And so
on...**

Questions about ethical practices in research?



03.

NLP & Ethics

How is NLP impacted by these practices? And vice versa.





**So what's this got
to do with NLP?**



Well... everything.

A few misconceptions of many

Machine learning is often seen as “playing with data” and “solving abstract problems”

*But: “The L in NLP means language, and language means people”
(Schnoebelen 2017)*

Linguistics is not uniformly better: ethics board approval for child language work, psycholinguistic studies, and some documentary work

So we need to do IRB review for NLP research?

- **Ethics boards tend to focus on preventing abuses of research subjects, not further downstream harms**

Do better.

- Engagement in discussions of ethics/societal impact that go **beyond** “this work is **approved/unacceptable**”
 - *(Though some systems should not be built.)*
- Working together to **understand** what **potentials for harm** there are, how to mitigate them, and how to educate the public about them
- Working together to **develop best practices** that help us do this work
- Move away from leaderboardism and towards work that is **situated** and **interdisciplinary**
 - *(This requires a different, but healthier, time-scale.)*

Best Practices

- Treat subjects fairly (informed consent, fair compensation)
- Abide by licenses and terms of use
- Data and model documentation (data statements, datasheets, model cards)
- Connecting model development work to specific problems in the world (with specific use cases and stakeholders)
 - Especially important in benchmark development
- Identifying stakeholders and, if possible, seeking their input
- Writing ethics statements which are proactive, rather than defensive
- Own your point of view ... and learn from others

Questions about NLP ethical impact?



04.

(Ethical) Review in NLP

And how it needs upgrading



It's inadequate, includes only IRB approval (if anything at all), mostly considers only data collection; has not been adjusted for emerging and disruptive technologies like ML/AI; no analysis of downstream risk or dual use consequences

This will be addressed in more detail in a later section.

Questions about review standards in NLP?



05.

Warning Examples

How have things gone wrong in the past? How do we prevent it from happening again?

Facebook and Myanmar.

<https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/>

GPT-3, Timnit Gebru, and the culture around ethics.

<https://www.wired.com/story/google-timnit-gebru-ai-what-really-happened/>

A whole corpus of examples of bias in Language Models, Machine Translation, etc.

- Gender
- Demographics
- Class
- Race
- etc.

False sense of security in ‘anonymized’ data
Identifiability does not rely just on tokens, for example
Data privatization, protection

<https://www.theguardian.com/technology/2019/jul/23/anonymised-data-never-be-anonymous-enough-study-finds>

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=998565 (“I’ve Got Nothing to Hide” and Other Misunderstandings of Privacy, Solove, 2007, 2014)



etc.



06.

Current research on ethical NLP

... and proposed solutions/harm
mitigation methods



start simple.

RECOMMENDATIONS FOR MORE ETHICAL NLP

Data Statements for NLP: Transparent documentation

(Bender & Friedman 2018)

- Foreground characteristics of our datasets (see also: AI Now Institute 2018, Gebru et al 2018, Mitchell et al 2019)
- Make it clear which populations & linguistic styles are and are not represented
- Support reasoning about what the possible effects of mismatches may be
- Recognize limitations of both training and test data:
- Training data: effects on how systems can be appropriately deployed
- Test data: effects on what we can measure & claim about system performance

think big picture.

Situating ML tasks in the world

- Make time to consider, early & often, the following questions:
 - What are the use cases of the technology being developed?
 - How does the specific ML task (inputs, outputs) relate to the intended use case?
 - What are the failure modes and who might be harmed?
 - What kinds of bias are likely to be included in the training data?
- Broaden our notion of ‘scaling up’: It’s not just about large numbers but also about diverse communities & experiences with the software

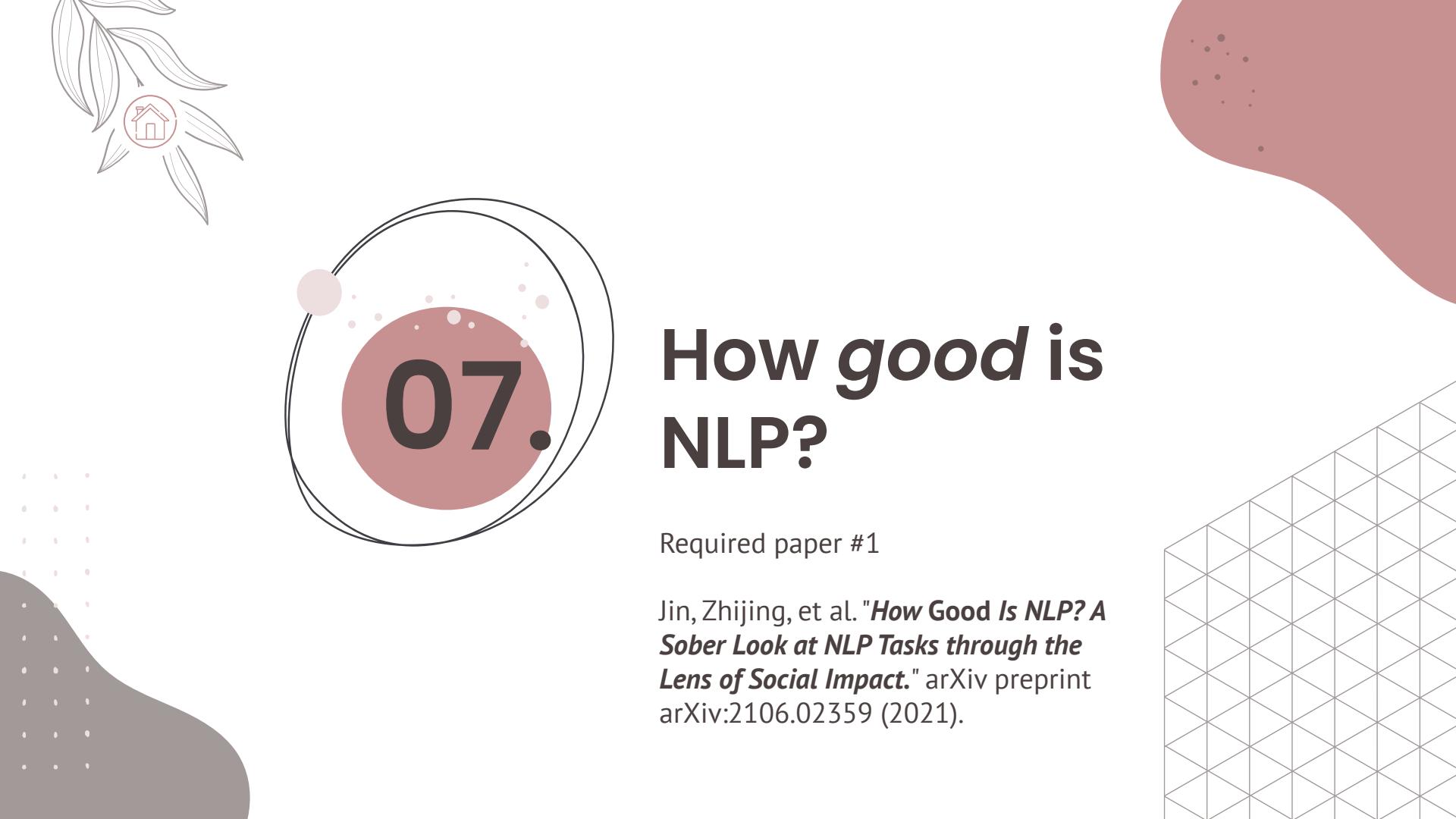
(go on a (reading)) Bender.

A great way to familiarize yourself with the current state of all of the many aspects of NLP research is to look at a curated program of papers and discussions on them.

Professor Emily Bender at University of Washington has done extensive work in developing frameworks for researchers and teachers to talk about ethical NLP. She has created multiple courses on Ethics in NLP, and is one of the biggest names in this area.

See the website for one of her courses; you'll find *dozens* of relevant and interesting papers, modularized by topic/aspect. It's NLP ethics *for days*.

https://faculty.washington.edu/ebender/2019_575/



07.

How good is NLP?

Required paper #1

Jin, Zhijing, et al. "***How Good Is NLP? A Sober Look at NLP Tasks through the Lens of Social Impact.***" arXiv preprint arXiv:2106.02359 (2021).

How good is NLP? – recap

“Anticipate rising importance of developing NLP technologies for social good.”

Aim to promote a guideline for social good in the context of NLP

Propose a framework to evaluate the direct and indirect real-world impact of NLP tasks, and adopt the methodology of global priorities research to identify priority causes for NLP research

Want to provide recommendations for ethical development of algorithms



Unintended negative consequences that early theoretical researchers did not anticipate have also emerged:

- o Microsoft's Twitter bot Tay
- o Leak of privacy of Amazon Alexa
- o GPT-3 with risks and harms of encoding gender and racist biases

Existing Guidelines (or, lack thereof)

These conversations are not uncommon in NLP spaces such as conferences like **ACL** and **NAACL**, but they're still emerging and only gaining relevance.

An example of a common outcome is **requiring all research papers to submit broader impact statements** in non-NLP conferences; NLP then followed suit and in 2021 ACL instituted *optional* ethical and impact statements.

We need comprehensive guidelines for following ethical standards to result in positive impact and prevent unnecessary societal harm.

Right now, there is no scientific evaluation framework for researchers to work within – they're relying on intuition and word of mouth recommendations to base their research agenda on.

Paper's main question

Given a specific researcher or team with skills s , and the set of NLP technologies T they can work on, what is the best technology $t \in T$ for them to optimize the social good impact I ?

What is 'social good'?

“[it] is context-dependent, relevant to *people, times, and states of nature*”
(Broome, 2017)

Moral Philosophy Theories:

- **Intuitionism** – relative certainty on judging impact as positive or negative (i.e. easy, given cases)
- **Deontology** – emphasizes duties or rules
- **Consequentialism** – emphasizes consequences of acts
- **Virtue ethics** – emphasizes virtues and moral character

Using the philosophies

Decomposition of impact judgment/analysis into these philosophy dimensions;

common in business analysis/decisions

Principles of governance for future AI / NLP

- Humanity
- Privacy
- Security
- Fairness
- Safety
- Accountability
- Transparency
- Collaboration
- Share/Equality
- AGI
(superintelligence)

How to evaluate indirect impact in NLP?

It's not easy – linguistic theories, for example, don't seem to have direct human impact, though biases that arise from those theories might.

(Jin et al., 2021) classifies NLP on a **dimension of theory → application**, for easier differentiation of expectations, implicit impact, etc.

4 dimensions / stages in developing a new NLP technology:

1. Fundamental theories
2. Building block tools
3. Applicable tools
4. Deployed applications/products

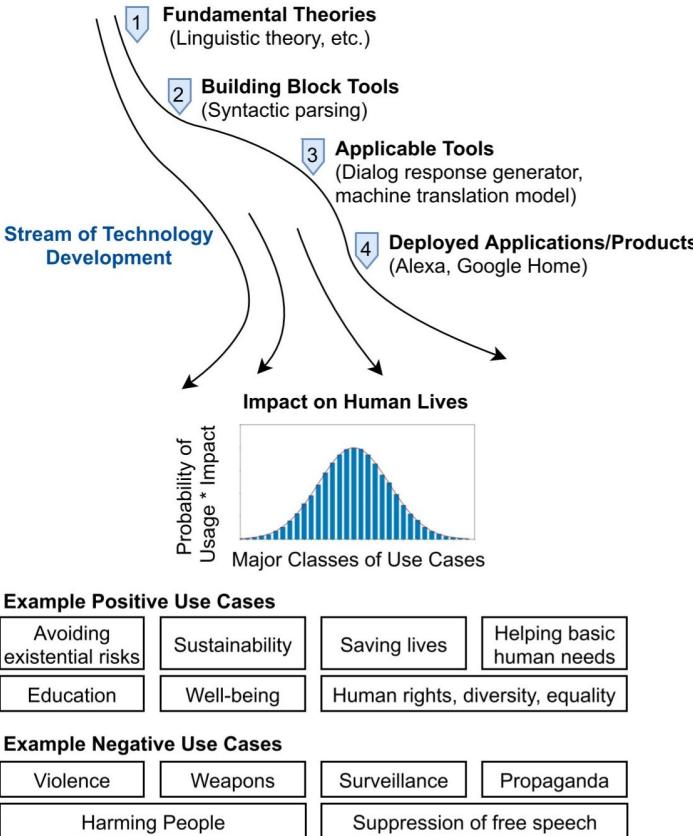


Figure 1: Stream of technology development from theory to application with end impacts. The end impacts are a distribution of use cases and their corresponding weighted impacts.

(Jin et al., 2021)

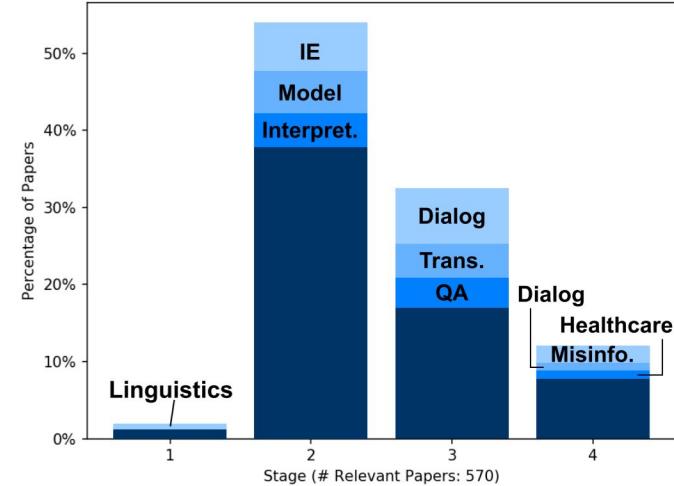


Figure 2: Distribution of ACL 2020 papers by the four stages. For each stage, we highlight the top several topics of the papers. We only list the top one topic for Stage 1 due to visual space limit. Abbreviations of technologies include Information Extraction (IE), Interpretability (Interpret.), machine translation (Trans.), question answering (QA), and misinformation (Misinfo.).

ESTIMATING IMPACT



Direct Impacts of Stage-4 technologies:

- Allowing users more free time (e.g., machine translation, NLP for healthcare, fake news detection, etc.)
- However, even the free time factor is varied, and can be both positive and negative

$$I(t) = \sum_{as \in AS} \text{scale}_{as}(t) \cdot \text{impact}_{as}(t), \quad (1)$$

summation of scale*impact of each consequence/risk

(Jin et al., 2021)

Indirect Impacts of Early-Stage Technologies:

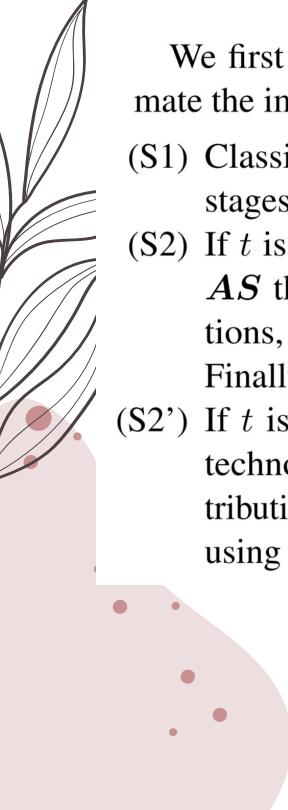
Different approach because it's harder to predict impact when tech is in early-stage

$$I(t) = \sum_{x \in \text{Stage-4 DE}(t)} p(x) \cdot c_x(t) \cdot I(x) , \quad (2)$$

p is probability that descendent (from t) tech x can be successfully developed, $c_x(t)$ is contribution of t to x ; I is formula from direct stage-4 impact slide

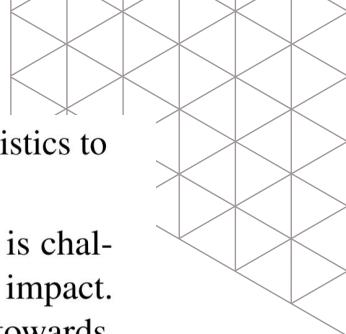
nb: this is a meta framework, not a developed one.

(Jin et al., 2021)

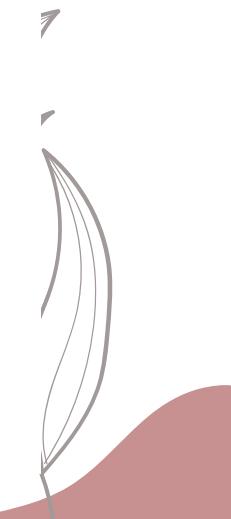


We first introduce some thinking steps to estimate the impact of research on an NLP task t :

- (S1) Classify the NLP task t into one of the four stages (§3.1)
- (S2) If t is in Stage 4, think of the set of aspects AS that t will impact, the scale of applications, and aspect-specific impact magnitude. Finally, estimate impact using Eq. (1).
- (S2') If t is in Stage 1-3, think of its descendant technologies, their success rate, and the contribution of t to them. Finally, estimate impact using Eq. (1) and (2).



Next, we introduce some high-level heuristics to facilitate fast decisions:

- (H1) For earlier stages (i.e., Stage 1-2), it is challenging to quantify the exact social impact. Their overall impact tends to lean towards positive as they create more knowledge that benefits future technology development.
 - (H2) Developers of Stage-4 technologies should be the most careful about ethical concerns. Enumerate the use cases, and estimate the scale of each usage by thinking of the stakeholders, economic impact, and users in the market. Finally, evaluate the final impact before proceeding. (E.g., if the final impact is very negative, then abandon or do it with restrictions).
 - (H3) For Stage-3 technologies, if their Stage-4 descendants are tractable to enumerate and estimate for their impacts, then aggregate the descendants' impacts by Eq. 2. Otherwise, treat them like (H1).
- 

DECIDING RESEARCH PRIORITY:

Authors propose a research priority list, given that NLP is such a vast and unexplored field of possibilities. They do this under the impression of UN's global priorities.

The high-level intuitions are drawn from the **Important/Neglected/Tractable (INT) framework** (MacAskill, 2015), a commonly adopted frame-work in global priorities research on social good.

Authors propose explaining this framework in mathematical terms, but do not suggest using it as a one-and-done for quantification of impact (because none of it is directly calculable; money and profit gain is only but a fraction of overall impact, for example)

- Needs holistic review of **expected social impact** because most research doesn't directly produce it – but sometimes it is worth it to achieve it incrementally.
- **Marginal impact** – focusing on important but neglected fields, because extensive research in a highly saturated field may have marginal impact
- **Opportunity cost** – largest expected improvement of social impact
- can also **borrow some criteria** from effective altruism, and also statistical calculations of social good.

Evaluating NLP Tasks: Case Study of ACL 2020

Authors evaluate the distribution of NLP papers for social good in ACL 2020 (there were 89)

Notice geographical difference in priorities (e.g., USA focuses almost only on interpretability; China almost not at all)

Compared with UN's SDGs, current NLP research lacks attention to other important cause areas:

- Tackling global hunger
- Extreme poverty
- Clean water and sanitation
- Clean energy
- Education (very poorly represented)

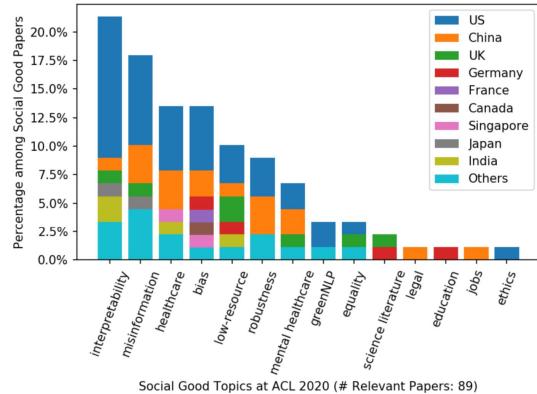


Figure 4: Social good topics at ACL 2020 by countries.

(Jin et al., 2021)

Cause? Value misalignment

- Authors give a proposed list of NLP priority research aligned with UN's SDGs

Priority	Example NLP research topics
Poverty	<ul style="list-style-type: none">• Predicting poverty by geo-located Wikipedia articles (Sheehan et al., 2019)• Parsing fund applicant profiles (proposed)
Hunger	<ul style="list-style-type: none">• NLP for agriculture (Yunpeng et al., 2019)• NLP for food allocation (proposed)
Health & Well-being	<ul style="list-style-type: none">• NLP to analyze clinical notes (Dernoncourt et al., 2017a,b; Luo et al., 2018; Gopinath et al., 2020; Leiter et al., 2020a,b)• NLP for psychotherapy and counseling (Biester et al., 2020; Xu et al., 2020; Pérez-Rosas et al., 2019)• NLP for happiness (Asai et al., 2018; Evensen et al., 2019)• Assistive speech generation (proposed)
Education	<ul style="list-style-type: none">• NLP for educational question answering (Atapattu et al., 2015; Lende and Raghwanshi, 2016)• Improving textbooks (Agrawal et al., 2010)• Automated grading (Madnani and Cahill, 2018; Taghipour and Ng, 2016)• Plagiarism detection (Chong et al., 2010)• Tools for learners with disabilities (proposed)
Equality	<ul style="list-style-type: none">• Interpretability (Köhne, 2015; Belinkov et al., 2017; Nie et al., 2020)• Ethics of NLP (Hovy and Spruit, 2016; Stanovsky et al., 2019; Sap et al., 2019)• NLP for low-resource languages (Zoph et al., 2016; Kim et al., 2017)• NLP on resource-limited devices (Sun et al., 2020)• NLP tools that signal bias in human language and speech (proposed)
Clean water	<ul style="list-style-type: none">• Raising public awareness of water sanitation (proposed)
Clean energy	<ul style="list-style-type: none">• Green NLP (Strubell et al., 2019; Schwartz et al., 2020)• NLP to analyze cultural values regarding climate change (Jiang et al., 2017; Koenecke and Feliu-Fabà, 2019)• Cross-cultural models of climate change perceptions (proposed)

Table 3: Top priorities and some NLP research related to each of them. This list may not be exhaustive. We also propose a high-impact research problem in each of the areas which has received less attention so far.

Heuristics

Checklist heuristics for social goodness of an NLP research direction:

- (Q1) What kind of people/process will **benefit** from the technology?
- (Q2) Does it **reinforce** the traditional structure of beneficiaries? I.e., what groups of underprivileged people can be benefited? (e.g., by gender, demographics, socio-economic status, country, native languages, disability type)
- (Q3) Does it **contribute** to SDG priority goals such as poverty, hunger, health, education, equality, clean water, and clean energy?
- (Q4) Can it **directly improve** quality of lives? E.g., how many QALYs might it result in?
- (Q5) Does it count as
 - ◆ (a) mitigating problems brought by NLP, or
 - ◆ (b) proactively helping out-of-NLP social problems?

What does it look like in practice?

Paper has some case studies on examples of NLP impact in these areas, with tasks such as:

- Low-resource NLP and machine translation,
- transparency, interpretability, algorithmic fairness and bias,
- green NLP,
- QA & dialog,
- information extraction,
- NLP-powered search engine & summarization,
- NLP for social media.

Notice how these cover almost all the topics considered in this course so far.

Conclusion: How good *is* NLP?

Ethical and societal implications:

Could benefit a focused, organized, and accountable development of NLP for social good

Recommended direction on how to achieve social good in context of NLP research.



Discussion



How do you feel about this typology of social good research goals?

What should NLP as a field/community focus on most (in terms of resources allocated, efforts invested)?

What's your take on the ACL by-country analysis of topics covered?

Do you agree/disagree with the presented principles of ethics and governance? Why?

Can you think of examples for the 4 stages of NLP technology development? Do you agree that they should be assessed relative to it?

What other ways can we estimate/evaluate social impact? (for those only convinced by numbers and metrics)



08.

Use of formal ethical reviews in NLP

Required paper #2

Santy, Sebastin, Anku Rani, and Monojit Choudhury. "*Use of Formal Ethical Reviews in NLP Literature: Historical Trends and Current Practices*." arXiv preprint arXiv:2106.01105 (2021).

General overview

Standard practice to get a study involving human subjects reviewed and approved by a professional ethics committee/board of the institution.

This is not the case in NLP research (most of the time).

Look at ACL Anthology with specific questions in mind. (see next slide)

Compare to other, related fields.

Suggest pathways for improvement.

Questions paper poses

How commonly do we see mention of ethical approvals in NLP research?

- What types of research or aspects of studies are usually subject to such reviews?
- With the rising concerns and discourse around ethics of NLP, do we also observe a rise in formal ethical reviews of NLP studies?
 - If so, would this imply that there is a heightened awareness of ethical issues that was previously lacking?

Introduction

Most of the ethical issues in NLP are rooted in the data being used for research.

Other works propose best practices for dealing with ethical implications of NLP research and deployment.

A number of reforms at NLP conferences have arisen from this heightened awareness around the topic of NLP ethics.

- (New tracks focusing on Ethics in NLP; some recommend the inclusion of broader impact statement in their papers...)

INSTITUTIONAL ACCOUNT OF ETHICS IN NLP?

The Where's Waldo of the research
review / ethics-oriented community.

Institutional Review

There exists **no specialized IRB** (Institutional Review Board) for NLP.

NLP falls under computing research. (This means it mostly considers data collection and annotation processes, as well as user studies.)

However, it is not prevalent to have IRB approval in NLP studies. Almost nonexistent in 2006, but a steady increase since 2016.

The most notable finding of the paper was that:

“IRB permission was mostly sought for either data collection or annotation studies, but hardly ever for data re-purposing or system design/deployment - a void that [we] think the NLP community should be conscious about.”

Methods and Findings

Analyze (210) papers from major NLP conferences, journals, and workshops from ACL Anthology. For comparative analysis, do the same for other related conferences (CogSci, InterSpeech, NeurIPS, CVPR, ICWSM, CHI, COMPASS)

Keyword search for IRB-related concepts. Also wildcard string *ethic* (used Allen AI SPIKE interface, and pdfgrep)

IRB approval was taken for the resulting papers; or it was mentioned as a recommendation for the field.

3.1 How many papers seek IRB approvals?

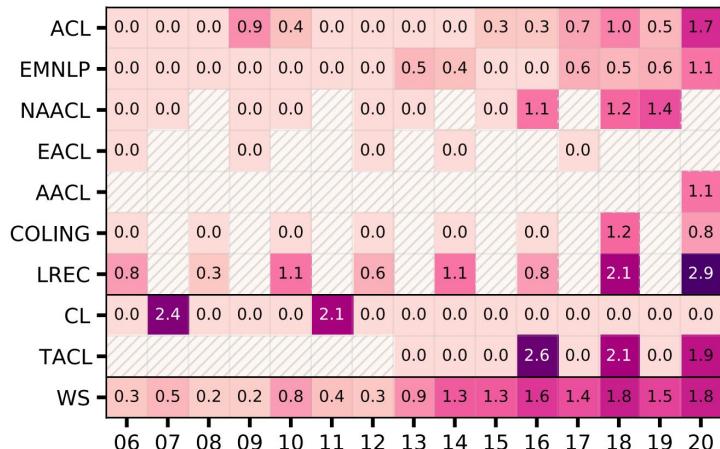


Figure 2: The percentage (%) of papers in NLP conferences, journals and workshops over the past 15 years that mention IRB-related terms. The intensity of color is proportional to the % values. The boxes with a gray hatch reflect the years when that particular conference was not held.

Aspects as metric

Results what you expect – those who take approval, take it for data collection, annotation of data.

However, only 12% took IRB approvals for data-scraping.

Only ONE paper seeks the approval for re-purposing and further annotation of data.

(Note GDPR makes this type of research subject to IRB approval due to the nature of the personal data collected by media platforms)

Only ONE paper has taken IRB approval for the whole system (owing to its sensitive nature).

Topics as metric

Nearly half of the papers that mention IRB have sensitive topics (e.g., mental health hate speech, clinical/medical NLP)

20.3% are about collection of eye movement, EEG, and audio/video recordings of human subjects.

The rest are generic data collection/user study approval.

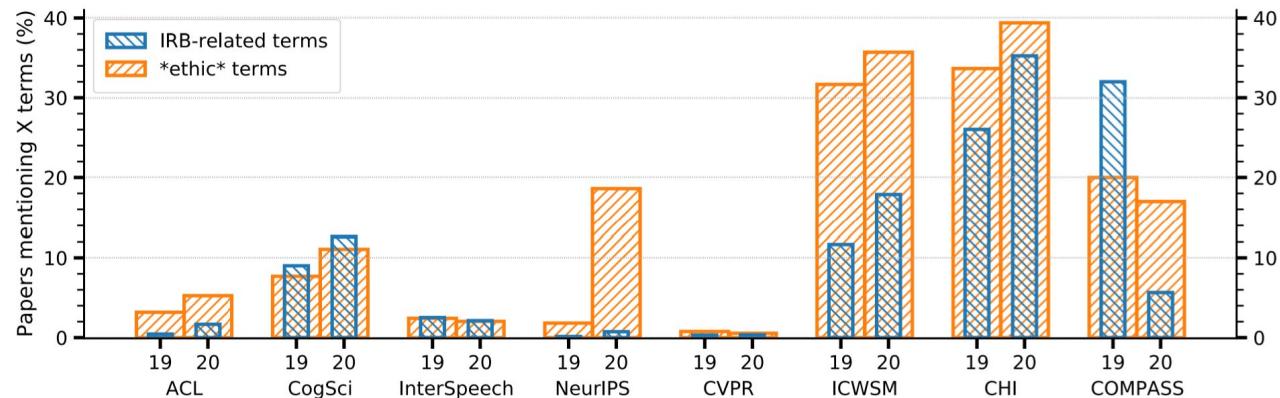


Figure 4: Percentage(%) of papers mentioning IRB-related and *ethic* terms in related conferences.

nb.

Distribution of IRB approvals **by country and industry vs. academia?**

*(note only ONE industry paper of 841 took IRB approval)

**highlight my own

3.3 What is the distribution of IRB approvals by country and industry/academia?

	Total Papers	Percent Papers
Countries		
USA	47/5368	0.88%
Canada	5/358	1.40%
Germany	5/850	0.59%
UK	5/1088	0.46%
Netherlands	3/226	1.33%
Sweden	2/100	2.00%
South Korea	2/151	1.32%
China	2/2350	0.09%
Affiliation Types		
University	52/7730	0.67%
Industry	1/841	0.12%
National Lab	1/182	0.55%
Joint/Collaboration	11/2651	0.41%

Table 1: Distribution of % IRB-related term mentions among countries and different types of affiliations for NLP conferences (excluding **LREC** and **WS**) from 2012 to 2020.⁶

(Santy et al., 2021)

trends.

“InterSpeech and CVPR have significantly fewer papers with IRB mentions (< 0.35% and < 2.5%, respectively) and the trends have hardly changed over the years, despite the fact that they conduct research with speech, multimodal, and vision data that may have been collected from human subjects.”

“NeurIPS, on the other hand, has seen a meteoric rise in their *ethic* mentions, which, on manual inspection reveals, is due to their mandatory inclusion of broader impact statements.”

hope.

“This quantitative testimony from NeurIPS shows that ACL and other *CL conferences **are moving in the right direction** with their inclusion of stringent ethics reviews for their papers.”

takeaways.

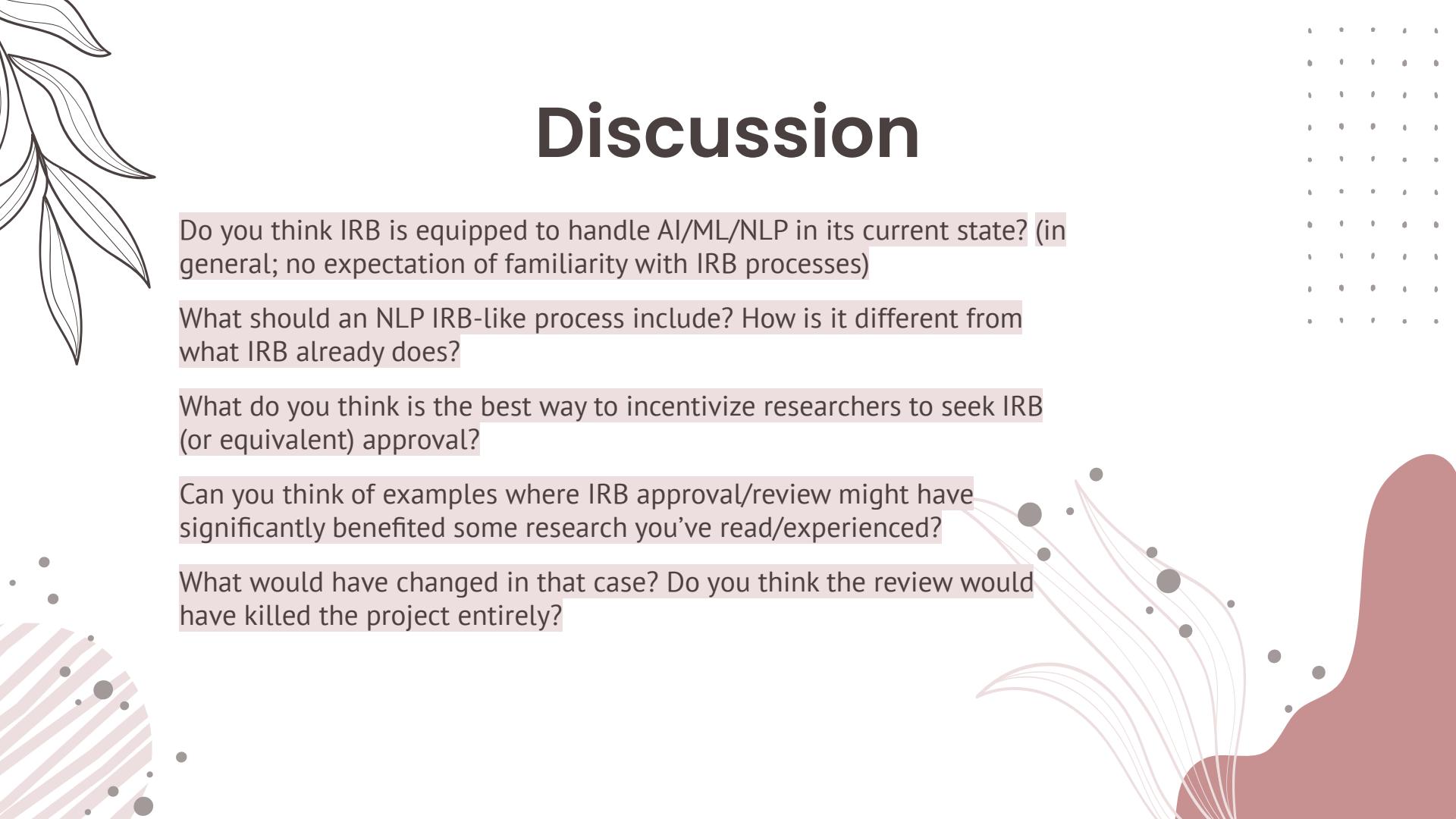
Very very few papers in NLP/CL in recent research seek IRB approval or other forms of ethics review process.

This is **DANGEROUS**.

“As such, repurposing data collected from human subjects without their explicit consent on how the data will be re-used is potentially dangerous and may even have legal repercussions.”

*“Furthermore, with the exception of a couple of papers, to date, **there is no practice or trend of taking IRB approval for designing, developing, and deploying systems**”*

... and so much of the **harm caused by a system** could actually come from **its design or type of training or deployment**; rather than underlying datasets.



Discussion

Do you think IRB is equipped to handle AI/ML/NLP in its current state? (in general; no expectation of familiarity with IRB processes)

What should an NLP IRB-like process include? How is it different from what IRB already does?

What do you think is the best way to incentivize researchers to seek IRB (or equivalent) approval?

Can you think of examples where IRB approval/review might have significantly benefited some research you've read/experienced?

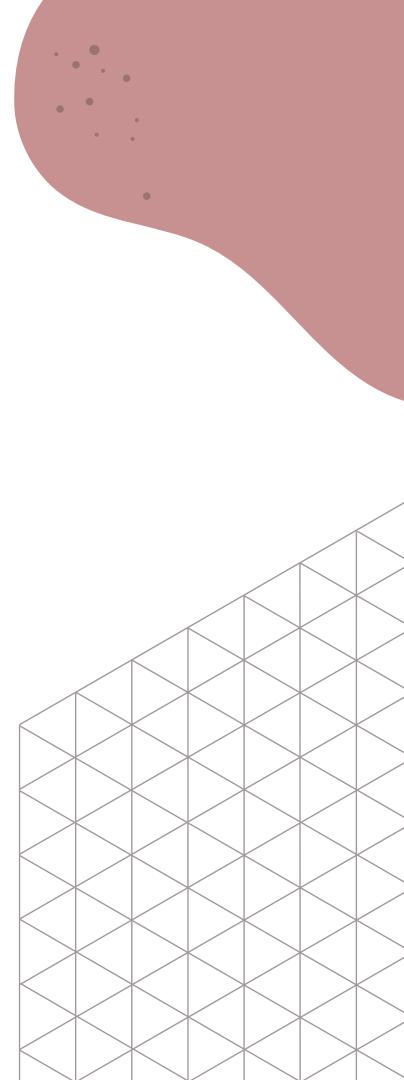
What would have changed in that case? Do you think the review would have killed the project entirely?



09.

Further reading

And more topics about NLP ethics



<https://arxiv.org/abs/2103.07929> <- A Systematic Review of Reproducibility Research in Natural Language Processing (Belz et al., 2021)
<https://dl.acm.org/doi/pdf/10.1145/3442188.3445875> <- Differential Tweetment: Mitigating Racial Dialect Bias in Harmful Tweet Detection (Ari Ball-Burack et al., 2021)
<http://arxiv.org/abs/1803.09010> <- Datasheets for Datasets (Gebru et al., 2018, 2020)
<http://arxiv.org/abs/2104.10097> <- Beyond Fair Pay: Ethical Implications of NLP Crowdsourcing (Shmueli et al., 2021)
<https://doi.org/10.1145/2073276.2073282> <- Toward a Formal Model of Accountability (Feigenbaum et al., 2011)
<https://www.reuters.com/investigates/special-report/myanmar-facebook-hate/> <- Facebook in Myanmar crisis
<https://arxiv.org/abs/1906.06425> <- Principled Frameworks for Evaluating Ethics in NLP Systems (Prabhumoye et al., 2019)
<https://aclanthology.org/O18-1041> <- “Data Statements for Natural Language Processing” (Bender and Friedman, 2018)
<https://aclanthology.org/W17-1612/> <- “Ethical Research Protocols for Social Media Health Research” (Benton et al., 2017)
<https://doi.org/10.1145/1713066.1713069> <- “Internet Research Ethics and the Institutional Review Board” (Buchanan and Ess, 2009)

<https://github.com/acgrissom/2016-ml-course/blob/master/assignments/writing.md>
https://faculty.washington.edu/ebender/2019_575/
https://faculty.washington.edu/ebender/2019_575/scicomm.html
[WWW 2019 SRNLP Tutorial Final](#)
[NAACL 2018 SRNLP Tutorial](#)
[NAACL 2018 SRNLP Tutorial](#)
http://demo.clab.cs.cmu.edu/ethical_nlp/
<http://web.cs.ucla.edu/~kwchang/talks/emnlp19-fairnlp/>
<https://web.stanford.edu/class/cs384/>
<https://uclanlp.github.io/CS269-Winter2020/>
<https://2021.naacl.org/calls/papers/#ethics-policy-updated-in-final-call>
<https://2021.naacl.org/ethics/faq/>
<https://2021.naacl.org/calls/reproducibility-checklist/>

(or, email me/Drago for even more; especially on topics not covered today;
getting official ethical review requirements and standards is all the rage these days)