

Ethics Sheets for AI Tasks (ACL 2022)

Saif M. Mohammad

National Research Council Canada

http://saifmohammad.com

saif.mohammad@nrc-cnrc.gc.ca

@SaifMMohammad

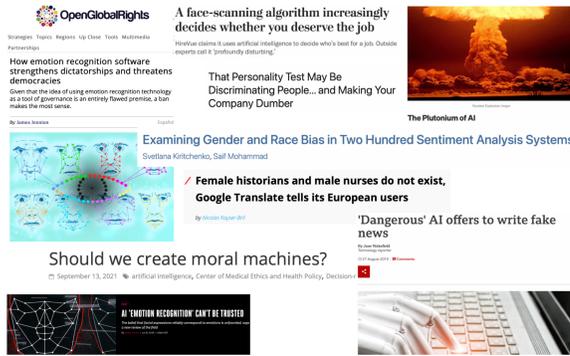


1. TECHNOLOGY OFTEN AT ODDS WITH PEOPLE

More adverse outcomes for those that are already marginalized

AI Tasks

- Face recognition
- Emotion recognition
- Personality trait identification
- Machine translation
- Image generation
- Text generation
- Deception detection
- Information retrieval
- ...



Recent Innovations to Bolster Ethics in AI

- Datasets: Datasheets (Gebru et al., 2018; Bender and Friedman, 2018)
 - Systems: Model Cards (Mitchell et al., 2019)
 - Papers, Funding Applications: Impact Statements
- Limitations:**
- Conflict of interest: incentives to show work in positive light
 - Tendency to produce boiler-plate text
 - Scope is limited to individual pieces of work

Criticisms of AI Systems and Published Research

- Physiognomy, racism, bias, perpetuating stereotypes, discrimination, ignoring indigenous world views, and more
- Thoughtlessness in machine learning
 - e.g., is automating this task really going to help people?
- Disregard for the variability and complexity of human behavior

Fletcher-Watson et al. 2018; McQuillan 2018; Birhane 2021

Motivation

- All AI tasks have their own unique ethical considerations: with various degrees of societal impact
- Need engagement with ethics at a level beyond individual projects
- **Important ethical considerations apply at the level of AI Tasks**

2. PROPOSAL

A Call to Document Ethics Considerations at the Level of AI *Tasks*

Ethics Sheet: A survey-style article that substantively engages with the ethical issues relevant to a task

- draws on knowledge from a body of relevant past work (from AI Ethics, Psychology, Linguistics, Social Science, etc.)
- goes beyond individual systems and datasets
- centers those most affected

Paper



Poster



What Does an Ethics Sheet Do?

- Fleshes out assumptions
 - in how the task is commonly framed
 - in the choices often made regarding the data, method, and evaluation
- Presents ethical considerations unique / especially relevant to the task
- Communicates societal implications
 - to researchers, engineers, the broader public
- Lists common harm mitigation strategies

Not so much telling one what is right and wrong. More about helping one determine what may be appropriate for their context.

Components of an Ethics Sheet

- **Preface**
 - why and how the sheet came to be written, the process, challenges faced,...
- **Introduction**
 - task definition & terminology, scope, ways in which the task can manifest
- **Motivations and Benefits**
 - list of benefits, research interests, commercial motivations
- **Ethical Considerations** (this is the star of the show)
 - Associated trade-offs, choices, societal implications, harm-mitigations strategies,...

Notable Benefits of an Ethics Sheet

Great introductory document for a topic

Acts as a record of what we, as a community, value

A written document allows stakeholders to

challenge our assumptions and conclusions

...and that is a good thing!

No One Sheet to Rule them All

A single sheet does not speak for the whole community

Multiple ethics sheets (by different teams and approaches) for the same or overlapping tasks can reflect multiple perspectives, viewpoints, and what is important to different groups of people at different times.

Be wary of the world with single authoritative ethics sheets per task and no dissenting voices.



Template of 50 Ethical Considerations

TASK DESIGN

A. Theoretical Foundations

1. Task Design and Framing
2. Theoretical Models and their Implications
3. Meaning and Extra-Linguistic Information
4. Wellness and Health Implications
5. Aggregate Level vs. Individual Level Prediction

B. Implications of Automation

6. Why Automate
7. Embracing Diversity
8. Participatory/Emancipatory Design
9. Applications, Dual Use, Misuse
10. Disclosure of Automation

DATA

C. Why This Data

11. Types of Data
12. Dimensions of Data

D. Human Variability–Machine Normativeness

13. Variability of Expression, Conceptualization
14. Norms of Emotions Expression

...

50. Attributing People to Social Groups

3. EXAMPLE



Ethics Sheet for Automatic Emotion Recognition and Sentiment Analysis
CL Journal, June 2022

What are the ethical considerations for your task?

