



What is AI Ethics? A Network Analysis of Scholarship

ONUR BAKINER

POLITICAL SCIENCE DEPARTMENT

Motivation

- ▶ Explosion of interest in AI ethics
- ▶ Unclear if concepts & their meanings are shared
- ▶ Unclear if an epistemic/academic community of AI ethics scholarship is emerging

Research Questions

- ▶ Do scholars use shared concepts to examine AI ethics?
- ▶ What kinds of social, political, economic, and environmental issues are mentioned?
- ▶ Do AI ethics scholars cite one another?
- ▶ Is there an AI ethics canon (i.e., texts cited across the board)?
- ▶ Does the field of specialization shape the choice of concepts and citations?

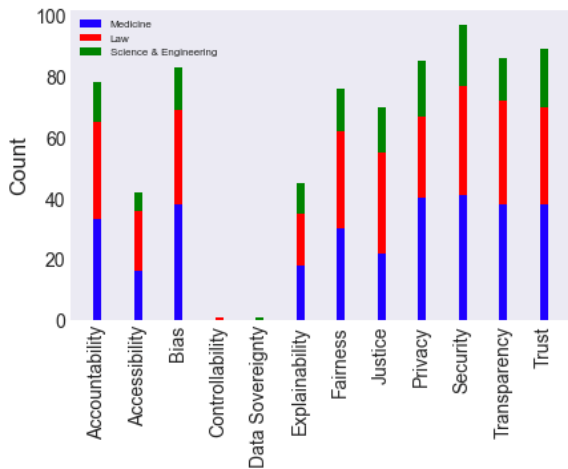
Data & Methods

- ▶ Data collection
 - ▶ 107 peer-reviewed articles (2017-present) that have “AI ethics” as keyword or part of abstract
 - ▶ Three fields of specialization: medical science (44), law (38), science & engineering (25)
 - ▶ Author(s), title, publication, year, bibliography, main text collected
- ▶ Data analysis
 - ▶ Exploratory/descriptive analysis (so far)
 - ▶ Network analysis of citations

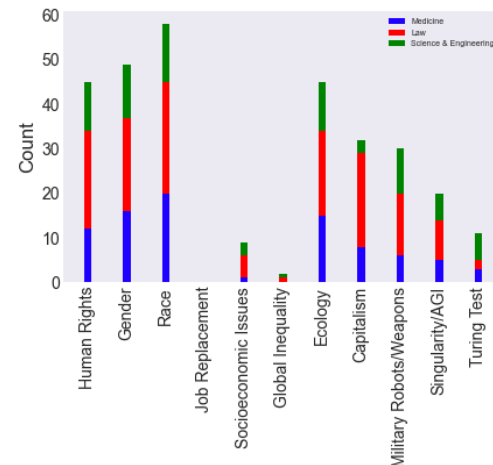
Findings I (tentative)

- ▶ Shared concepts exist
- ▶ Some social issues (rights, gender, race/racism, ecology) are shared, whereas others (especially economic ones) are not
- ▶ Only 28% of article mention philosophical paradigms

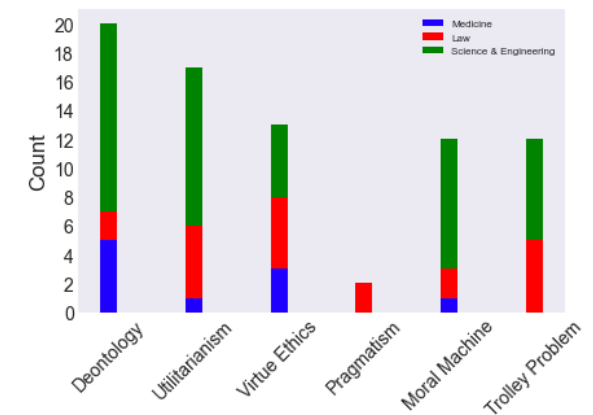
Article Count of Key AI Concepts, Disaggregated by Specialization



Article Count of Social Issues, Disaggregated by Specialization



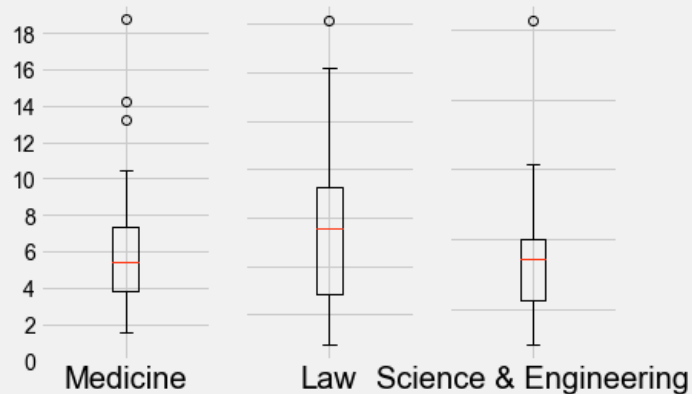
Article Count of Philosophical Paradigms and Concepts Mentioned, Disaggregated by Specialization



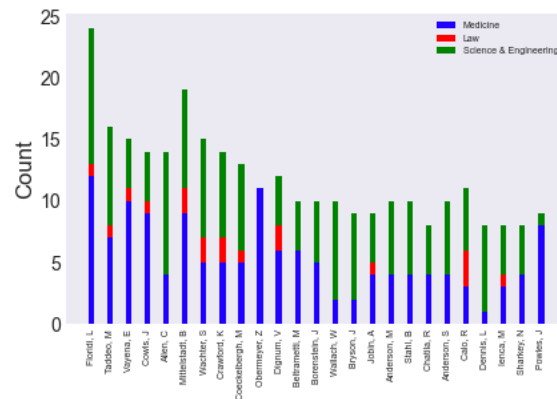
Findings II (tentative)

- ▶ Only recent literature is cited (except in law)
- ▶ Medicine and science & engineering have some shared literature, law has none
- ▶ Law and science & engineering cite some of the popular literature

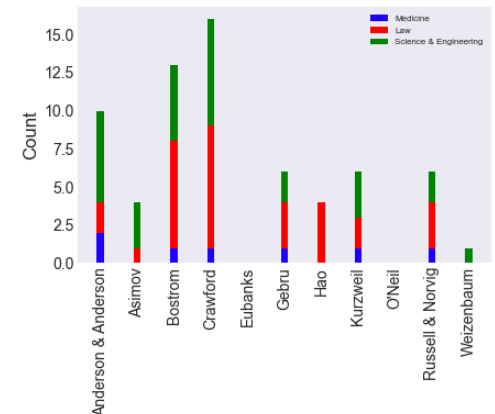
Age of Cited Sources per Discipline



Breakdown of Top 25 Authors Cited in the Literature

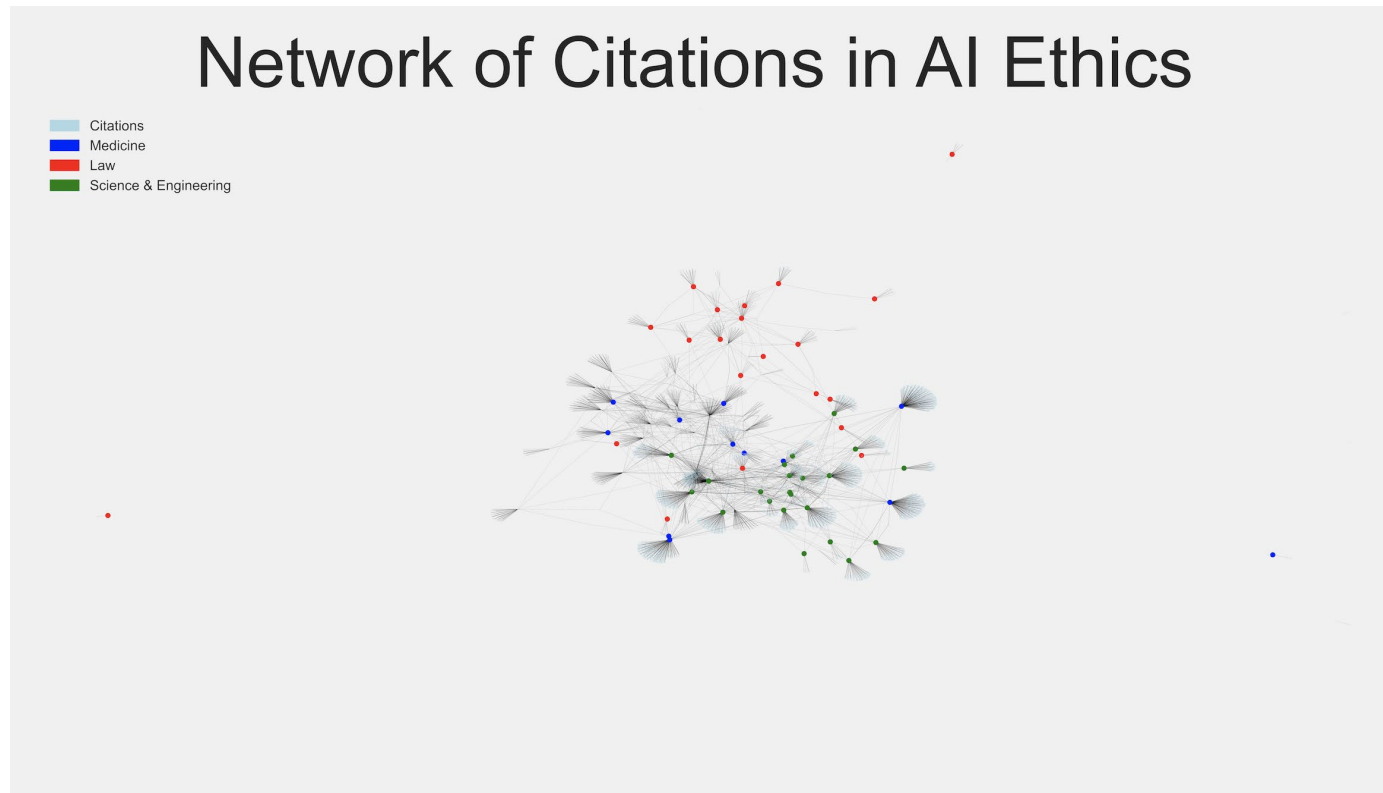


Breakdown of Popular AI Authors Cited in the Literature



Findings III (tentative)

- ▶ Citation network shows some degree of specialization



Conclusion

- ▶ AI ethics has emerged as a novel field with a short intellectual history
- ▶ Scholarship cannot be said to constitute an academic/epistemic community (yet)
- ▶ There is an interesting lack of socioeconomic writing on AI ethics
- ▶ Future work
 - ▶ Replicate with more articles
 - ▶ Compare/contrast average citation age with non-AI articles
 - ▶ Analyze meaning & sentiment in the texts