

COLM 2024: AI Ethics panel

1. My talk

1.1. *About me / Framing*

- A legal scholar focused on:
 - privacy
 - surveillance
 - CS & law
 - AI & law
- A repeat offender government regulator and law enforcer:
 - US DOJ
 - US Federal Trade Commission
 - Office of the Attorney General of Colorado

1.2. *The end of the 20-year hiatus of Cyberlaw*

- Early in the life of the commercial internet, a fizzy, lively, bustling new interdisciplinary community: Cyberlaw
- At some point, we retreated to our subdisciplinary corners:
 - Biggest Players
 - Intellectual Property
 - Privacy
 - Smaller Groups
 - Telecom
 - Antitrust
 - Robotics
 - Groups focused elsewhere but with a heavy dose of tech engagement
 - National Security
 - Antidiscrimination
 - First Amendment
 - Consumer Protection

- Labor
- We resisted successive waves of **Technological Developments**
 - Web 2
 - Smartphones
 - Social Media
 - Platforms
 - Cloud Computing
 - Blockchain & Crypto
- Along the way, we welcomed newer subgroups spurred by some of these:
 - Platform Law (mostly around section 230)
 - Fairness Accountability and Transparency
- Since November 2022, we've started to have an awkward family reunion.
 - The "one of each" conference.

1.3. The Copyright Problem

- **Thesis:** In the early days of large language models and attention on the ethical and legal problems they present, there has been an overly narrow focus on copyright law.
- **The Why** is not hard to see:
 - 17 USC 504:
 - \$750 - \$30,000 per work!
 - Raised to \$150,000 per work for willful infringement!
 - NYT's complaint against OpenAI: "billions of dollars in damages"
 - Also: other remedies:
 - Especially as relates to model pretraining, an argument that liability means an injunction ordering the company to delete the model and start over!
 - Existential stakes
 - But also unimaginable, politically if not legally.
 - To be clear: fundamentally interesting questions
 - Law:
 - Substantial similarity

- Fair Use
 - Secondary Liability
- CS:
 - These rules are all superficially quantitative and thus reduce themselves to metrics and evaluations!
- The result: a generation of computer scientists is being taught that the law, policy, and **ethics** problem that matters, is the copyright question.
 - Small data point to support: ICML Gen Law 1 and 2 papers

1.4. SO What's the Problem?

- **Key claim:** Each subdiscipline of law brings along a lot of underlying baggage:
 - Values
 - Methods
 - Interdisciplinary interlocutors
 - Overarching frameworks and worldviews.
- A world built around Copyright will be shaped like the values, methods, disciplines, and worldviews of Copyright.
 - Focused primarily on property and incentives
 - Close cousins to Economics and friend to measurement
 - Dominated by libertarians
 - To be clear, there are brilliant copyright scholars who focus outside these lanes. But they are on the outside looking in.
 - For example: those who focus on:
 - moral rights
 - labor and workers' rights
 - the public domain
- The focus on Copyright also brings along a large number of practical policy considerations:
 - Agencies:
 - Copyright Office

1.5. What is needed, Part I

- The main thing that is needed is a "big tent"
- Welcome other points of view, methods, values, and issues

1.6. But what is probably needed, Part II

- I think we're best off if we identify a counterweight
 - Otherwise it will be Copyright and The Rest
- We need another field that is big and substantial and associated with a different set of values, methods, disciplines, and parts of the political firmament.
- So we should eliminate:
 - Antitrust, Copyright's close cousin
 - Telecom, oddballs
 - Even Labor: just the other side of the Copyright Debate
- My proposal: Privacy

Field	Values	Method	Neighbors	Agencies
IP	Property and Labor	Economic	Economics	Copr Office
Privacy	Autonomy, Dignity, Choice	Sociotechnical	Philosophy, Sociology, STS	FTC / CFPB
NatSec	Security and Secrecy	Technocratic	PoliSci / Int'l Relations	IC

- Key point: NOT the narrow privacy that the computer scientists think we mean
 - Not "memorization" and not "training on personal data"
- Better: focused on harms including
 - deepfakes
 - child safety
 - misinfo and disinfo
 - bias
- My definition: "Any information harm or problem that isn't covered by information privacy"

Author: Paul Ohm

Created: 2024-10-07 Mon 11:33