Regulation and the Liability Problem for Increasingly Autonomous Systems

Peter M. Asaro, PhD
asarop@newschool.edu

Assistant Professor
School of Media Studies, The New School

Affiliate Scholar
Center for Internet and Society, Stanford Law School

Co-Founder & Vice-Chair
International Committee for Robot Arms Control
Overview

• Regulating Robots & AI

• Promote Safety & Innovation
  – Markets
  – Public Opinion

• Expectations & Predictability

• Managing Failures
The Liability Problem

• As increasingly autonomous systems act in the world, in increasingly complex and unpredictable ways, how do we manage the liability for the harms they may cause?

• Separation of:
  – Causal Agency
  – Legal Agency
  – Moral Agency
Liability & Accountability

• Compensation for Harms

• Punishment
  – Retributive Justice
  – Feedback Signal (Reform/Learning)
  – Deterrence (Impact on Future Decisions)

• Intention & Human-Centric

• Accountability & Transparency
Legal Approaches

• Agents & Diminished Agents
  – Children, Slaves, Animals
  – Agency Law
  – Employees

• Product Liability & Negligence (Corporations)
  – Joint & Several Liability
  – Strict Liability
  – Insurance or State/Society
Responsibility

- **Retroactive**
  - Someone to Blame & Punish
  - Target of Reform (Feedback)
  - Source of Retribution

- **Proactive**
  - Active Taking of Responsibility
  - Making Moral & Legal Judgements
Human Responsibility

• Meaningful Human Control

• Kill Switch
  – Recognizing Misaligned Values

• Policy Lever
  – Laws Act on Humans/Institutions

• Inappropriate Delegation
  – Lethal Decisions
  – Deprivation of Rights (Due Process)
Accountability Gap in AWS

• Who is responsible for the deaths?
  – Programmers
  – Commanders
  – Operators
  – The AWS
  – The State
  – Nobody? (de facto)

• Minimal Liability in War

• War Crimes Require Intent
Policing & Lethal Robots

• Higher Standards for Use of Force
  – “In order to prevent an immanent threat of
death or grave bodily harm.”

• Threat ID Requires:
  – Physical Modeling Capability
  – Psychological Model of Intent

• Could Disrupt Threat w/o Lethal Force

• Most Cases are Self-Defense of Officer

• Answer: No Autonomous Use of Force
Future Work

• Regulatory Mechanisms
  – On Humans/Manufacturers
    • Law
    • Ethics Boards
    • Training Engineers
    • Ethics in Design Process (IEEE P7000 Standard)
  – Internal to Autonomous Systems
    • Technical Safety Mechanisms
    • AI/Machine Ethics
  – Learned vs. Imposed by Design
Papers


Thank You!

- Email: asarop@newschool.edu
- Twitter: @peterasaro
- WWW: www.peterasaro.org