European University Institute

DEPARTMENT OF LAW

Models of Law and Regulation for AI - and what to learn (or not) from Sci-Fi?

Prof. Nicolas Petit

Nicolas.petit@eui.eu

Twitter: @CompetitionProf

Outline

- How to make sense of the disordered legislative and regulatory interventions towards AI?
- What does science fiction tell us about efforts to regulate emerging technology?

1. Quick reminder on AI
2. Models of law and regulation of AI
3. Four fallacies
4. Law v public policy for AI
5. A fifth model?
6. A quick look at science fiction

European University Institute
DEPARTMENT OF LAW

1. AI in a nutshell

Core ideas

- Any physical process including the mind process can be modelized as a computable algorithm (Church Turing thesis)
- Machines can learn: “Learning is any process by which a system improves performance from experience” (Herbert Simon)

Today

- Brute force computational power now available (due to Moore’s law)
- Zettabytes (or Yottabytes) of data now available, and distributed, and pre labelled (cloud)
- « End of theory »
    - Deep learning, neural networks, etc.
- Use cases: autonomous vehicles, predictive justice, automated law enforcement

European University Institute
DEPARTMENT OF LAW

2. Models of law and regulation for AI

- Black letter law model
- Emergent phenomena model
- Ethical model
  - Virtue ethics, deontological ethics and consequentialism
- Risk regulation model

European University Institute
DEPARTMENT OF LAW

3. Four fallacies

- The paradox of irrelevant law
- The problem or redundant law
- The failure of good intentions
- Knee jerk regulation

4. Law v Public Policy for AI

- Against public policy for AI?
- Against existing law for AI?

5. Alternative model: externalities with a moral twist

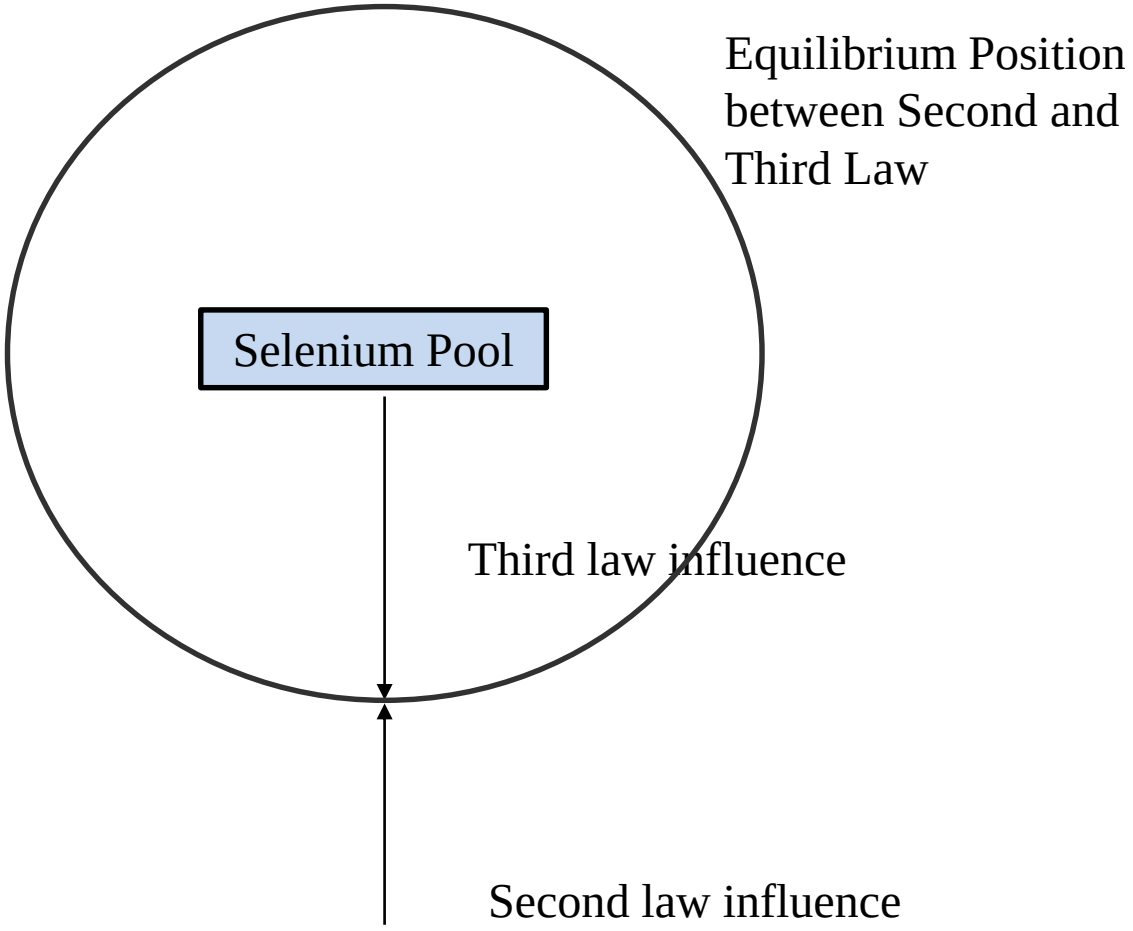
- Discrete externality (micro; personal)
  - For harms, ex post resolution, abstract standards, impact assessment
- Systemic externality (macro; societal)
  - Negative
    - Substitution effect in jobs
    - Privacy
  - Positive
    - Complementarity effect, Generative or General Purpose technologies
  - For harms, more *ex ante*, prescriptive, impact assessment
- « Existernality » (global; existential)
  - Negative: existential risk (Terminator)
  - Positive: pure human enhancement
  - For harms, more *ex ante*, proscriptive, deontological (no impact assessment or fat tail risks)

European University Institute
DEPARTMENT OF LAW

6. What Sci-Fi says?

- Why sci-fi might be relevant (and in particular Isaac Asimov's work)?
- Technological neutrality, not determinism
  - The 3 laws of robotics, *Runaround* (1942) +1
- Inevitability, but fallibility of man-made law
- Human intervention correcting errors in man-made law
- Emergent behavior – *Evitable Conflict*

Runaround (1942)



European University Institute

DEPARTMENT OF LAW

Thank you