Ethico-legal Governance

AI in Employment

Ethics of AI - Summer Semester 2021

Team 3: Rahaf Gharz Addien, Muhtashim Lekhon,
Amardeep Pal, Federica Suriano, Helena Winiger

07.07.2021

Outline:

- About
- Symbolic AI
- Ethico-legal Governance
- AI in Employment as high risk Systems
- The Reasoning Process
- Deontic Logics
- Ethical reflections
- Challenges
- Roadmap

https://pixabay.com/de/illustrations/pfeil-wegweiser-wegzeiger-richtung-2085195/

About:

- Ethico-legal governance based on symbolic logic for AIs
- The possibility of encoding legal constraints in symbolic AI
- AI in employment:
    - Automated decision making in hiring
    - Exploring possible levels of governance based on symbolic AI
    - Ethical dimensions

Symbolic AI

- Reasoning oriented field
- Solving classical AI problems of:
  - Solving mathematical problems
  - Rules and Ontologies
  - Logical inference
- **Reasoning Capabilities**
- **Explicit Symbolic Programming**

Ethico-legal Governance

What is Ethico-legal Governance?

- Governance
  - Set of processes and abilities
- Ethical Governance
  - Governing through ethical values
  - High standards of performance
  - Responsibility

https://ai.wharton.upenn.edu/wp-content/uploads/2020/12/ai-governance.jpg

Ethico-legal Governance

- Explicit ethico-legal governance of AIs based on symbolic AI
  - Encoding ethical and legal constraints in symbolic logics

Ethico-legal Governance

- Explicit ethico-legal governance of AIs based on symbolic AI
  - Encoding ethical and legal constraints in symbolic logics

  - Improving the degree of accountability and explainability
  - Increasing the autonomy level in selecting the suitable actions ethically
  - Reducing the harms could be caused by AIs

AI in Employment

EUROPEAN COMMISSION

Automated Decision Making in Hiring

EU proposal for regulation, 2021

https://pixabay.com/get/g1c35e3e1f0bb2454ecd28cd7851a5b2d7bc2026daebfd5af52a7caffc888293b4a3d2793dcde4996fffe08e27b48cd37_1920.jpg

ANNEX III
HIGH-RISK AI SYSTEMS REFERRED TO IN ARTICLE 6(2)

Relevant Legal Constraints

Article. 22 GDPR:

- Right not to be subject to a decision based solely on automated processing.
- **Unless** the decision is necessary for entering into a contract between the data subject and a data controller.
- Implementing suitable measures is obligatory at least human intervention for evaluation
- The decisions shall not be based on special categories of personal data Art. 9 **unless** it is necessary for reasons of substantial public interest.

Reasoning Process

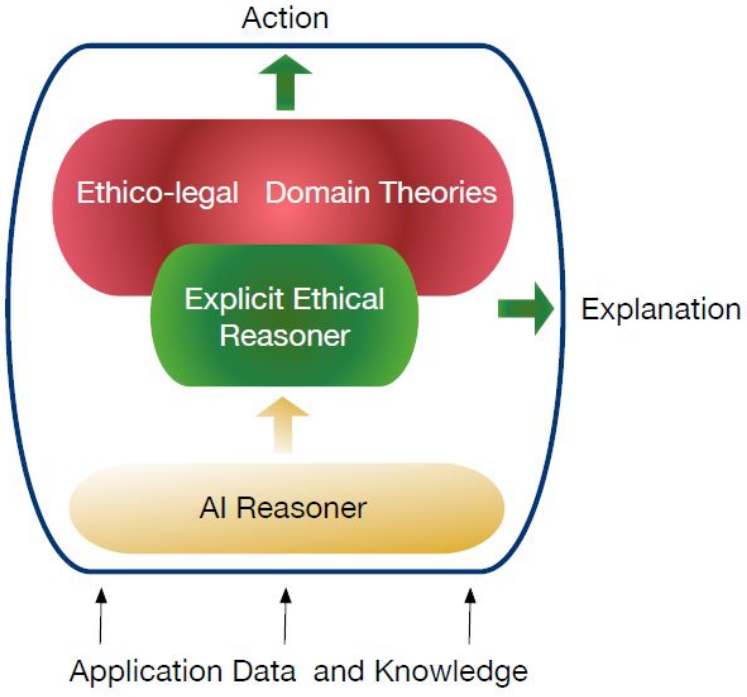


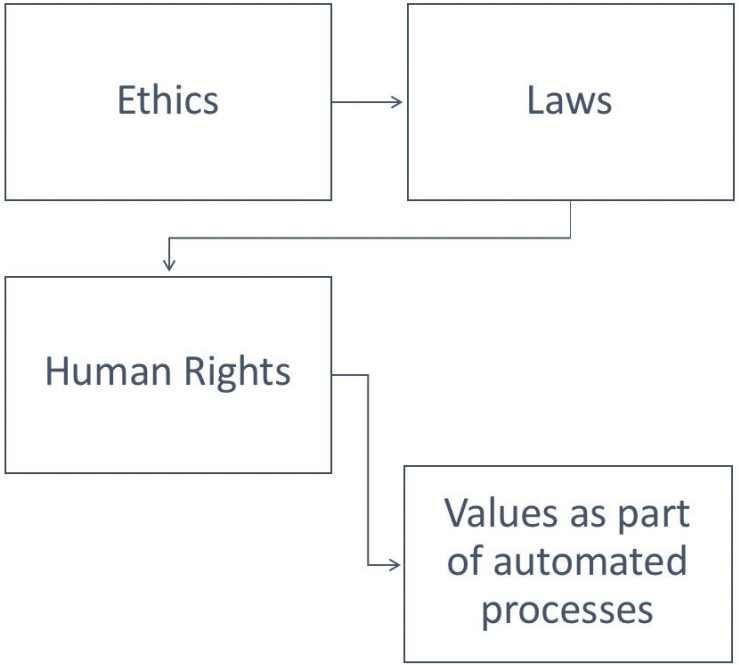
Figure from: Benzmüller, C., Parent, X., & van der Torre, L. (2020). Designing Normative Theories for Ethical and Legal Reasoning

Deontic Logics

Need for formalized values

-> Deontic logics is a branch of logic regarding e.g.

- permissible (permitted): PE
- impermissible (prohibited): IM
- obligatory (must, ought): OB



Implementation

Combination of logics:

- **Deontic Logics:**
  - Standard Deontic Logic (SDL)
  - Dyadic Deontic Logic (DDL)
- **Mathematical Logic:**
  - First-order logic
  - Higher-order logic (HOL)
    - set theory
    - quantifications over properties allowed
    - based on simply typed λ-term
    - shorter proofs than in first order logic

Reflection in an ethical context

- Technical reflections:
  - Transparency and explainability: ‘reason of reason’
  - Recommended actions /decisions based on ethical and legal values
  - Human intervention: decisions have to be evaluated by humans

https://www.pexels.com/de-de/foto/holz-mann-liebe-menschen-8504340/

Reflection in an ethical context

- Social reflections:
  - Wide socio-technical context
  - Information about decisions
  - Contestability: users' right to argue against decision
  - Data protection
- Research reflections:
  - Interdisciplinary collaboration
  - Broader ethico-legal projects
  - Further ethics and articles

https://www.pexels.com/de-de/foto/holz-mann-liebe-menschen-8504340/

Challenges

- Clarity of the legal text
- Personal data
- Applying suitable measures
- Definition of actions
- Consistency with other articles of GDPR
- Choosing the deontic logic
- Validity / Truthfulness

ps://pixabay.com/de/illustrations/überwindung-stein-rollen
-schieben-2127669/

Roadmap

Research Phase
Idea of the project
Ethical values
Possible solution

Embedding Phase
Comparing different deontic logics
Implementation

Concretization Phase
Approaching to equivalent
ethico-legal rules and fitting logics

Reporting Phase

References

1. Arkin, R. C., Ulam, P., & Duncan, B. (2009). *An Ethical Governor for Constraining Lethal Action in an Autonomous System.* Fort Belvoir, Defense Technical Information Center: Technical Report.
2. Benzmüller, C., & Lomfeld, B. (2020). Reasonable Machines: A Research Manifesto. *KI 2020: Advances in Artificial Intelligence, 43rd German Conference on AI*, pp. 251-258.
3. Benzmüller, C., & Lomfeld, B. (2020). Träumen vernünftige Maschinen von Gründen? Eine reale Utopie. *#Verantwortung KI - Künstliche Intelligenz und gesellschaftliche Folgen, KI als Laboratorium? Ethik als Aufgabe!*, pp. 29-36.
4. Benzmüller, C., Parent, X., & van der Torre, L. (2020). Designing Normative Theories for Ethical and Legal Reasoning: LogiKEy Framework, Methodology, and Tool Support. *Artificial Intelligence, Vol. 287*, pp. 2352-3409.
5. Benzmüller, C., & Miller, D. (2014). Automation of Higher-Order Logic; http://page.mi.fu-berlin.de/cbenzmueller/papers/B5.pdf
6. Benzmüller, C., Farjami, A., & Parent, X. (2018). *Faithful Semantical Embedding of a Dyadic Deontic Logic in HOL;* https://arxiv.org/abs/1802.08454
7. Benzmüller, C., Steen, A., & Wisniewski, M. (2016). *Tutorial on Reasoning in Expressive Non-Classical Logics with Isabelle/HOL;* http://page.mi.fu-berlin.de/cbenzmueller/papers/C61.pdf
8. Govindarajulu, N. S., & Bringsjord, S. (2015). Ethical Regulation of Robots Must Be Embedded in Their Operating Systems. In R. Trappl, *A Construction Manual for Robots' Ethical Systems. Requirements, Methods, Implementations* (pp. 85–99). Berlin: Springer.
9. Joritka, M., & Winfield, A. F. (2019). Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions, Royal Society, A 376*, pp. 1-13.
10. McDaniel, P., Holz, T., Döhmann, I., Burchard, C., Sadeghi, A.-R., Rieck, C., . . . Freiling, F. (2021). *Cybersecurity and Machine Learning Vision Document.* National Science Foundation (NSF) and German Research Foundation (DFG).
11. Nipkow, T. (2021). *Programming and Proving in HOL;* https://isabelle.in.tum.de/doc/prog-prove.pdf
12. Theodorou, A., & Dignum, V. (2020). Towards ethical and socio-legal governance in AI. *Nature Machine Intelligence*, pp. 10–12.
13. McNamara, P. (2006). Standard Deontic Logics; https://plato.stanford.edu/entries/logic-deontic/
14. o.A.. Dontic Logics and Normative Reasoning; https://homepage.ruhr-uni-bochum.de/defeasible-reasoning/deon-log-2019/post/session-01/
15. EU Commission (2021). Laying down harmonized rules on Artificial Intelligence (AI Act); https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206
16. The Royal Society (2019). Explainable AI: the basics; https://ec.europa.eu/futurium/en/system/files/ged/ai-and-interpretability-policy-briefing_creative_commons.pdf
17. EU Parlament (2000). Charta of Fundamental Rights of the EU; https://www.europarl.europa.eu/charter/pdf/text_en.pdf
18. EU Parliament and Council (2016). General Data Protection Regulation. Article 22; https://gdpr-info.eu/art-22-gdpr/
19. EU Parliament and Council (2016). Does the GDPR apply to companies outside the EU?.;https://gdpr.eu/companies-outside-of-europe/
20. https://pixabay.com/de/illustrations/pfeil-wegweiser-wegzeiger-richtung-2085195/
21. https://ai.wharton.upenn.edu/wp-content/uploads/2020/12/ai-governance.jpg
22. https://pixabay.com/get/g1c35e3e1f0bb2454ecd28cd7851a5b2d7bc2026daebfd5af52a7caffc888293b4a3d2793dcde4996fffe08e27b48cd37_1920.jpg
23. https://pixabay.com/de/illustrations/überwindung-stein-rollen-schieben-2127669/
24. https://www.pexels.com/de-de/foto/holz-mann-liebe-menschen-8504340/

Thank you!

Are there questions?

What do you think about our project?

What could be possible problems?