Can we teach AI Ethics?
A field feedback

Pr. François Goulette, ENSTA Paris
Deputy director U2IS (Robotics & AI Lab)

Hi! PARIS Meet Up on "AI & Ethics"
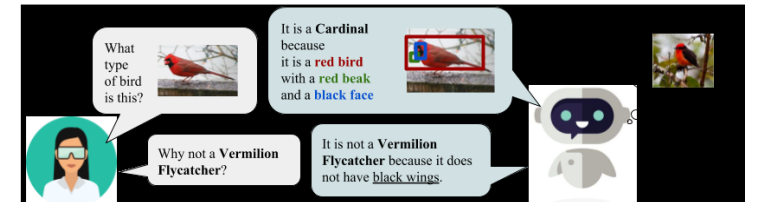
April 22nd, 2024

How to teach AI ethics?

- In charge of setting up a 30 hour-course for M2 students
  - Master 2 level, mainly for computer scientists
    - open to other backgrounds (PSL Week – PSL University)
  - ~20-35 students, started in 2019-2020
  - Web page:
    - https://www.caor.minesparis.psl.eu/presentation/cours-eia/

- Field feedback on how to teach AI ethics
  - Having students **interested & involved**
  - Teach **knowledges**
  - Teach / develop **know-how** of individual / collective ethical thinking

Content - knowledges

- Moral philosophy
  - Ethics (virtues, deontology, utilitarism, consequentialism...)
- Law
  - Digital regulation (GDPR, AI Act in Europe ; USA...)
- Computer science and specific issues for AI, deep learning, digital
  - Biases, explainability...
- Professional / applied fields
  - Autonomous systems, medicine, letal weapons...

- ALL these fields are moving FAST
- → It MUST BE taught by pluridisciplinary experts

Content – know-how

- Content - know-how
  - Develop an ethical thought on products and services using AI
  - Practical, operational

- Field experience :
  - hands-on group studies

→ It MUST be

- individual AND collective
- hands-on and operational

Kargu (kamikaze drone)

Mixed Reality: Apple Vision Pro

Content Generation: HeyGen

Examples of course projects 2023

Take-away field experience

- Scope of AI Ethics
  - Moral philosophy is large (various approaches), and not enough
  - Law is necessary, but not enough
  - Computer science and applied concepts are necessary
  - All are evolving fast

- Request from the students
  - Expectations on meaning in their professional lives (values / alignment)
  - Need for operational methods
    - Cannot wait years to develop a critical thinking on product design / use / deployment
  - Need for collective methods
    - Includes complexity, search for wisdom
    - Avoid biases, moral relativism
    - Express personal opinion, respect different viewpoints, elaborate a collective thought

To go further – how to avoid AI Ethics washing ?

- Development of AI Ethics is a REQUIREMENT today
  - Consciousness of global need

- BUT
  - It is hard, pluridisciplinary, evolving
  - It concerns multiple levels of product design, use, deployment, communities…

- How to avoid ETHICS WASHING ?
  - Develop further EDUCATION on AI ethics
  - Develop CONTINUOUS LEARNING in companies in AI ethics
  - Develop new TOOLS / METHODS for advanced collective thought
  - Set-up advisory AI Ethics Committees in organizations
    - for projects and product development…

An insight of complete 1 year - training

(30 hours modules)

- Introductory course (existing)
- Advanced moral philosophy
  - Morals and ethics, dilemmas, global ethics, ethical guidelines, auditability
- Epistemology
  - Man's conception of himself, upheavals in representations, intelligence, consciousness, transhumanism
- Law and regulations
- Sociology
  - "Parliament of objects" (Bruno Latour)
- Product design, ethics-guided design
  - CK method ; ethics adapted to TRL scale
- Advanced computer science, Generative AI, Natural Language Processing…
- Ethics and business development (start-up, unicorns, disruption)
- Professional modules : medicine (decision aided ; robotics) ; transportation ; autonomous systems ; defense
- Seminars
- Internship

Thank you!