

The Ethical Implications of AI Technologies in Society

Jean-Gabriel Ganascia

15 November 2024

Sorbonne University (SU), LIP6 (computer science lab of SU)

Ex-Chairman of the COMETS (CNRS Ethics Committee)

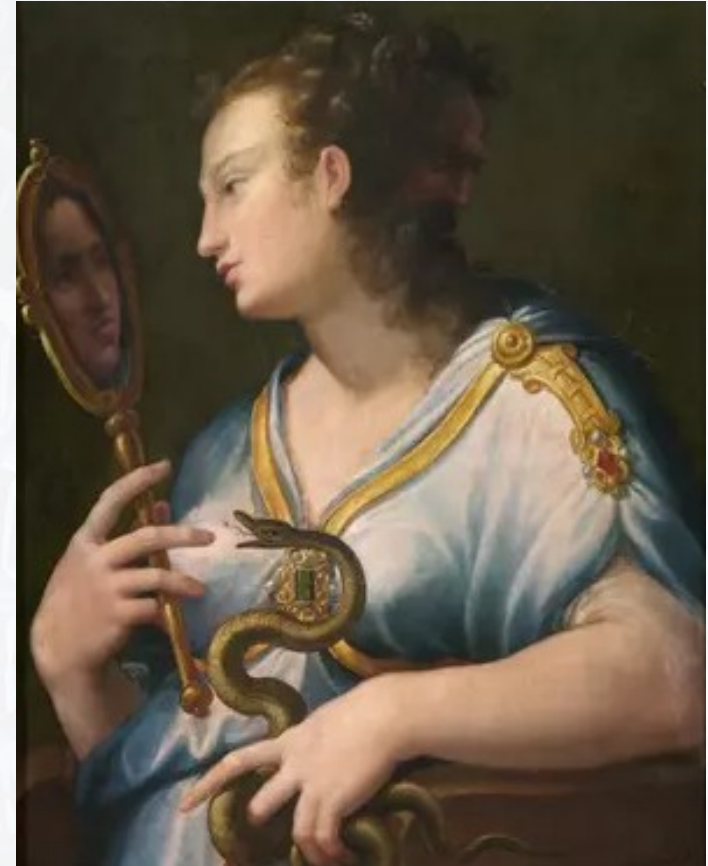
Chairman of France Travail Ethics Committee

Jean-Gabriel.Ganascia@lip6.fr



Synoptic

1. Two Characters from Mythology
2. Why is There a Problem with AI Technologies ?
3. What is the Problem?
4. Regulation is not the Solution!
5. How to Manage to Make Machines Virtuous and Trustworthy?



1

TWO CHARACTERS FROM MYTHOLOGY

Moral and Ethics

Ethics: Latin *ethica*; Greek *êthikos*, *êthikê*, from *êthos*, 'custom', 'mores'

Originally, in Greek, *êthos* meant a **place familiar to animals**, e.g. a stable.

With Aristotle, means the **rational deliberation** necessary to act well.



Moral : Latin *moralis* from *mores* → Mores

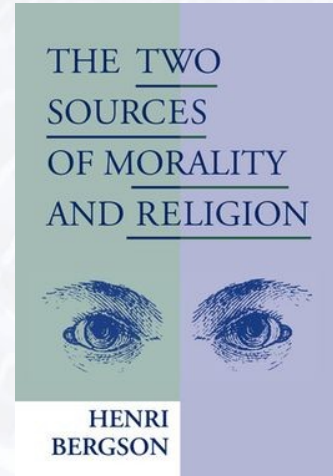
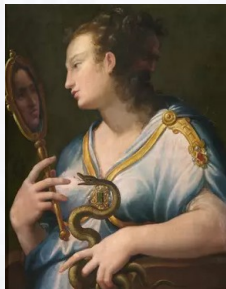
The **art (or the science?) of directing one's conduct**

Two sources:

- Experience, traditions
- Rationality, laws and rules

Approaches:

- *Traditions, precepts*
- *Principles:*
 - **Deontism:** Themis
 - **Utilitarianism:** prudence



Two Characters from Mythology: Prudence & Themis

Themis

Allegory of the **Justice**

Two symbols:

- she Carries a **Balance**, Sign of Equity
- she wears a **Blindfold** as a Token of Impartiality

Blind Application of Laws



Prudence

Personification of the **Virtue of Prudence**

Two symbols:

- she looks both the past in a **mirror** and the future ahead
- the **Snake** represent the knowledge

Anticipation and wisdom are drawn from lessons of experience



2

WHY IS THERE A PROBLEM WITH AI TECHNOLOGIES?

Examples of Laudable Applications of AI

Example: Health Sciences (agriculture, environment, etc.)

Processing Huge Masses of Medical Data

- Extracting medical knowledge from patient data (X-rays, clinical signs, etc.)
- Extracting biological information (e.g. genetic factors explaining the evolution of the disease, etc.)

Bioinformatics

- Modeling biological processes (e.g. mechanism of introduction of the virus into cells, genetic factors explaining the evolution of the disease, protein folding — Nobel Prize — etc.)

Extraction of Knowledge from the Scientific Literature

- More than 87,000 papers on CoViD-19 were produced between March and October 2020!

Robotics

- Robotics for the Elderlies
- Surgical Robots and Prothesis

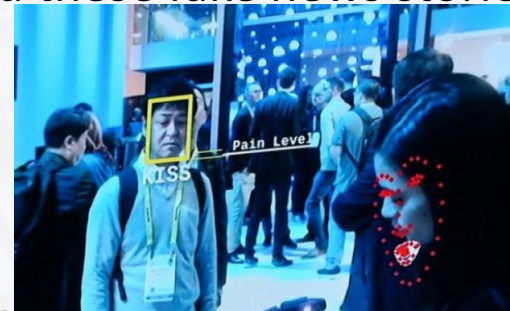


Misuses of AI and ML that affect society

Irresponsible, Unjust and Unfair uses of AI & ML

Uses of AI that could violate human dignity and autonomy

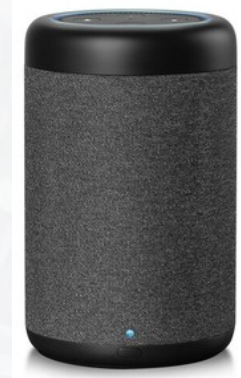
- Surveillance systems that would track every move — social credit in China...
... and Border Surveillance in US or EU
- Biased AI systems that are discriminatory — facial recognition
- Public shaming of those who break the rules
- Video, speech and/or image synthesis that could produce fake news
- AI-based targeting techniques to spread these fake news stories.
- ...



Misuses of Language Generation

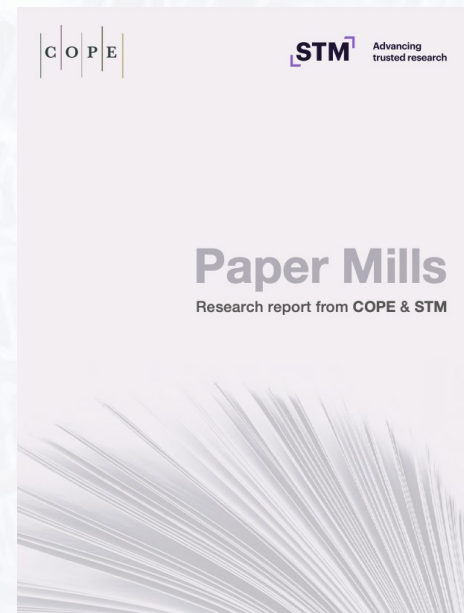
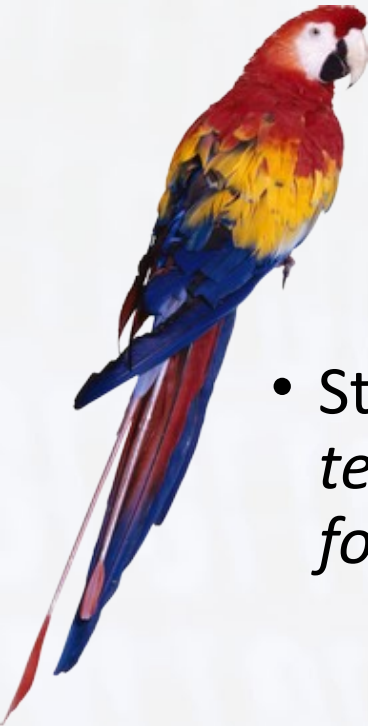
- “Chatbots” (dialogues)

- Domestic spies
- Misuses in education
- Industrial disruption
- ...



- Stochastic Parrots (*AI Generation techniques based on the use of foundation models*)

- Paper Mills
- Bias
- Environmental cost of learning
- “Deep Fakes”



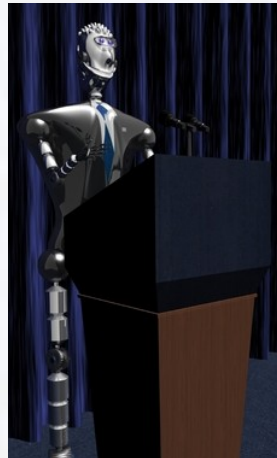
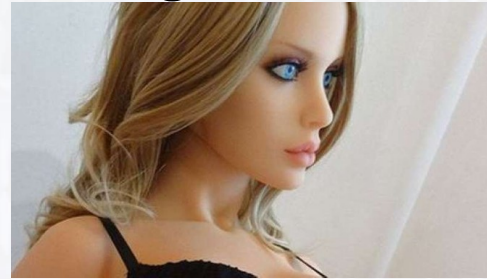
Current Application of AI in Robotics

Robots (i.e. “artificial workers”), bots (i.e. virtual robots) and automata are extensively used in a great multiplicities of activities, where they are useful.



But some uses of these technologies are more debatable as

- Autonomous weapons
- Sexbots
- Chatbots
- Fight Bots
- Influencers



3

WHAT IS THE PROBLEM?

AAAI Spring Symposium 2006 — W⁵

W⁵ – What Went Wrong and Why Workshop: *Lessons from AI Research and Applications*

Location: Stanford

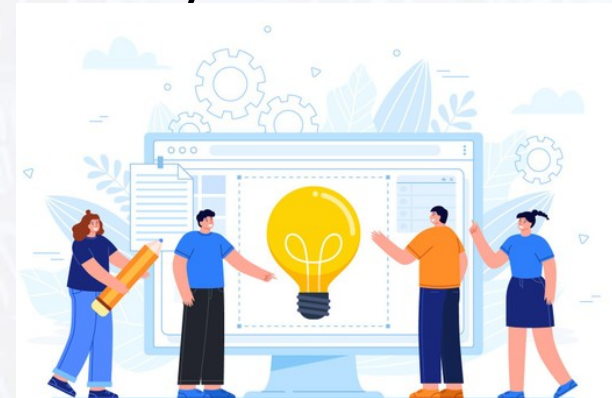
Participants: John McCarthy, Ed Feigenbaum, Carl Hewitt, Doug Lenat, Manuella Veloso, Mike Pazzani, Craig Knoblock ...

Example: “electric elves” are useful and efficient intelligent agents

But, it happened that they awake their owner at 3am to advise him that his 8am plane should have to be 2 hours delayed...

Lesson:

AI system are not just **technical systems**,
that also **socio-technical systems**.



Technical and Sociotechnical Dimensions

Technical devices are there to facilitate and help us.

They cannot be considered in isolation, as such...



And yet, this is not enough to inspire confidence!

Why?

The machines make errors

- How to make them sound?
- How to prove their soundness?

The machines contribute to transform the society

- How to restrict their use?
- How to anticipate and to preclude their misuses?



4

REGULATION IS NOT THE SOLUTION!

“Ethics” Committees (>84 in 2019* — 350 in 2022**) More than 67* (510**) Principles & recommendations

- The **Asilomar AI Principles**, developed under the auspices of the Future of Life Institute, 2017
- The **Montreal Declaration for Responsible** developed under the auspices of the University of Montreal, 2017
- The **General Principles of Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems**, IEEE, 2017
- The **Ethical Principles offered in the Statement on Artificial Intelligence, Robotics and ‘Autonomous’ Systems**, European Commission’s European Group on Ethics in Science and New Technologies, 2018

These principles are not only numerous, but also contradictory!

- **Privacy vs. Transparency**
- **Privacy vs. Security**
- **Lack of discrimination vs. Inclusive policy**
- ...

of academics, researchers, civil society organizations, companies building and utilizing AI technology, 2018

- **20 recommendations** of the AI4People, an Atomium-EMSD initiative designed to lay the foundations for a “Good AI Society”, 2018
- **Ethical Guidance for a Trustworthy AI**, High Level Expert Group on Artificial Intelligence, European Commission, 2019

• **AI Act: European Union**

*Brent Mittelstadt, “Principles alone cannot guarantee ethical AI”, *Nature Machine Intelligence*, vol. 1, November 2019, pp. 501–507, <https://doi.org/10.1038/s42256-019-0114-4>

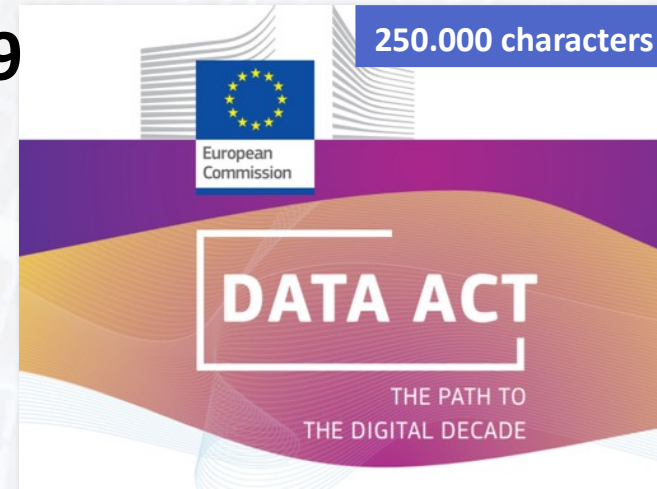
**Lionel Tidjon and Foutse Khomh, The Different Faces of AI Ethics Across the World: A Principle-Implementation Gap Analysis, 1922, <https://hal.science/hal-03689128/document>



European Regulations of the Digital!

Frantic digital regulation!

Acceleration since December 2019



AI Act – European Commission



EU regulation based on risk

Risk level	Examples of AI systems	Allowed in the EU?
Unacceptable risk	Social scoring used by governments; toys using voice assistance which encourages dangerous behaviour	No
High risk	Scoring of exams; AI application in robot-assisted surgery; verification of authenticity of travel documents	Yes, subject to mandatory requirements, ex-ante and ex-post enforcement
Limited risk	Chatbots; "deep fake" videos	Yes, subject to transparency requirements
Minimal risk	AI-enabled video games; spam filters	Yes

Three AI unacceptable risk

- Manipulation of persons through **subliminal techniques (!)**
- AI based social scoring done by public authorities
 - exemptions: **creditworthiness** assessment, credit scoring when put into service by **small-scale providers**
- prohibition of driving license penalty point system
- Real time biometric identification



James Vicary admitted making false claims about subliminal advertising in 1957 forcing many people to think subliminal messages didn't work!

5

HOW TO MANAGE TO MAKE MACHINES
VIRTUOUS AND TRUSTWORTHY?

How to Manage?

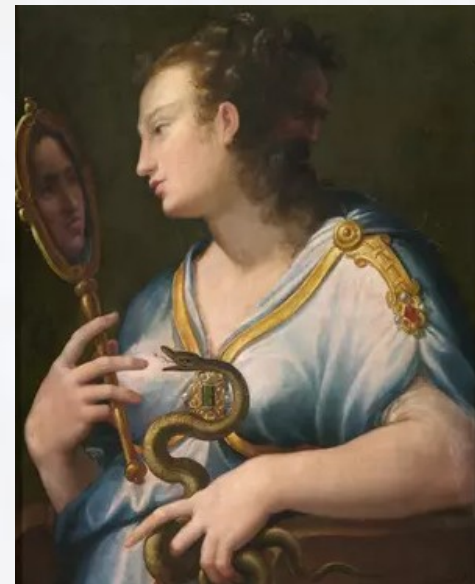
- **Make them more sound, robust and “trustworthy”**
- **Prove or test** their soundness on past examples

Add prescriptions to machines in order they:

- Respect justice
- Protect Privacy
- Preserve Security
- Be Transparent
- Preserve Human Autonomy



Ethics by design
Computational Ethics



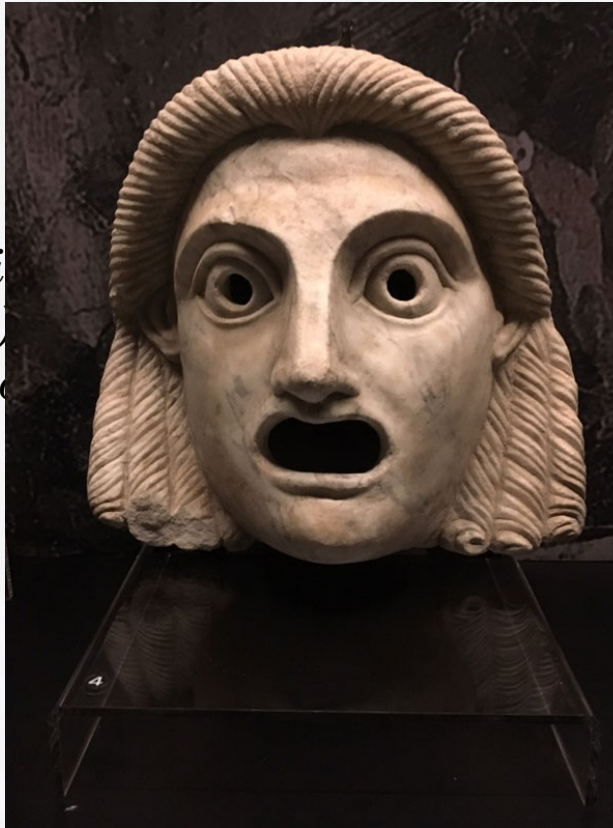
Ethics is not only *prudence* and *virtue*, it is also openness to “what (or who) is coming”...

Chaos:

Mathematical Theory → Unpredictable Dynamic Systems

Jacques Derrida: “*The chaos, it is the form of all future as such, of all that comes*”

“*The open abyss of Khaos, it is also the open and gaping form of my mouth (Khainô), when I do not know what to say*”



Think What Happens
Jacques Derrida

Derrida
pour les temps
à venir



sous la direction de René Major

l'autre pensée Stock

Thank You!