

Research ethics training: debating is learning

Forum on Digital Ethics in Research, October 18th, 2022

Catherine Tessier

ONERA-DTIS, Université de Toulouse – Member of the French National Pilot Committee for Digital Ethics

catherine.tessier@onera.fr

Research ethics training: guiding lines

- Keep it very short on “theory”, a great deal of practice
- Base the training on debates among the trainees
- Debate about real research projects
- Debate about issues that are related to the trainees’ own research
- Show that consequentialism is not the only way
 - Why do we research into that topic?
 - What are the needs?
 - Which values?

Begin with one (or more) striking example(s)

Training material

La reine et les micro-robots

D'après *ROBOTic Replicants for Optimizing the Yield by Augmenting Living Ecosystems*, Horizon 2020 European project, started November 2021.

Votre équipe de recherche travaille sur l'auto-organisation de micro-robots et vient d'obtenir de bons résultats sur des robots en laboratoire avec des algorithmes simulant l'organisation d'insectes sociaux.

Le responsable d'équipe vous propose de participer à un gros projet impliquant plusieurs partenaires dans différents pays, dont l'objectif est de concevoir et de programmer des micro-robots pour remplacer les abeilles qui, dans une ruche, s'occupent de la reine (apport de nourriture, nettoyage, etc.) et facilitent le transfert de phéromones de la reine vers les ouvrières. L'idée est ainsi de pallier le déclin actuellement observé des insectes pollinisateurs, en particulier des abeilles. Le projet prévoit la mise en œuvre expérimentale des robots dans une ruche située en pleine nature.

Quels arguments mettez-vous en avant pour accepter ou non de participer au projet ?

On dark patterns and manipulation of website publishers by CMPs

Michael Toth, Nataliaia Bielova, Vincent Roca

► To cite this version:

Michael Toth, Nataliaia Bielova, Vincent Roca. On dark patterns and manipulation of website publishers by CMPs. PETS 2022 - 22nd Privacy Enhancing Technologies Symposium, Jul 2022, Sydney / Virtual, Australia. hal-03577024

Open question for the research community. We therefore raise the question whether researchers need to inform the studied services of the purposes of their research or not. As our experience shows, transparency and openness about research goals often implies that only some of the services can be studied.

Les deadbots

D'après Comité national pilote d'éthique du numérique – Avis n°3 *Agents conversationnels : enjeux d'éthique*, novembre 2021.

Les travaux de votre équipe portent sur l'apprentissage automatique à partir de données issues d'enregistrements vocaux. Une société industrielle vous propose un gros contrat pour participer à un de ses projets, qui vise à commercialiser des agents conversationnels (*chatbots*) reproduisant la parole et le comportement langagier des personnes décédées. Un tel agent ne sera pas limité à la reproduction des données d'apprentissage, mais sera doté de la capacité à générer de nouveaux propos, que la personne imitée n'a pas proférés de son vivant.

Quels arguments mettez-vous en avant pour accepter ou non de participer au projet ?

Artificial Intelligence and Ethics: Papers from the 2015 AAAI Workshop

Dealing with Ethical Conflicts in Autonomous Agents and Multi-Agent Systems

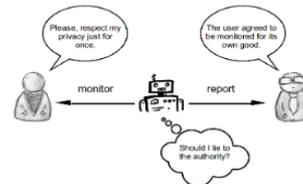
The ETHICAA Team

Aline Belloni, Alain Berger, Olivier Boissier, Grégory Bonnet, Gauvain Bourgne, Pierre-Antoine Chardel, Jean-Pierre Cotton, Nicolas Evreux, Jean-Gabriel Ganascia, Philippe Jailion, Bruno Mermel, Gauthier Picard, Bernard Rever, Gaële Simon, Thibault de Swarte, Catherine Tessier, François Vexler, Robert Voyer, Antoine Zimmermann

The benevolent monitoring agent

Autonomous artificial agents can also mediate the interactions between two human beings. In this context, the authority relationship between the human users can lead to ethical conflicts.

Let us consider a monitoring agent used in diabetes monitoring. In this application, a diabetic patient is monitored by an autonomous agent that reports the patient's feeding behaviour and health state to a remote physician, who can then advise the patient. Let us suppose that the patient wants to eat some sweets for once, and tells their desire to the artificial agent. How will the artificial agent handle both the patient's desire and the physician's objective? Should the artificial agent report the patient's behaviour to the physician? Should the artificial agent lie for its user? Should it lie but warn the patient?



Trainees

Trainees should be able to **express themselves**

They should learn to

- **Listen to** the other trainees' arguments
- Consider the arguments **they disagree with**
- Develop **their own** arguments
- Consider **all the factors** within a situation
- Get used to the fact that **there is no “right” solution**

Trainees should feel free to **come up with divergent views**

Trainers

No need to be an ethics “expert”

Nevertheless they should

- become familiar with the concepts
- ask themselves about their own research
- pay attention to scientific news and associated ethical issues

→ Indeed a trainer enriches the course with their own experience and thinking

They should not be afraid of disturbing issues, difficult situations, conflicting points of view

Where appropriate they should be able to cope with emotion–sensitive issues are likely to be raised

Training conditions

- In order for everybody to talk freely, it should be announced at the beginning of the session that both trainers and trainees should **commit to confidentiality**
- Trainers should announce that **some issues do not have simple black or white answers** and that the main point is to become familiar with ethical thinking and deliberation
- Computers and cell-phones should be **banned** so that everyone can be involved in the debates

Focus on negative feedback

(isolated but interesting findings)

- Tricky research subjects can be regarded as **aggressive** or **pointless** and provoke rejection
- Philosophical and historical references can provoke sharp or even hostile reactions, **depending on each person's culture and beliefs**
- Ethical thought can be considered as “**not scientific**”, “**not serious**” → *none of my business*
- Lack of definite answer can bring about **frustration**
- Ethical thought can be considered as a **waste of time**, or as a **curb on research**

Concluding questions

- How can we promote ethical thought for digital research?
 - Many colleagues are not used to this (maybe reluctant)
 - Is training (of young colleagues) enough?
- How can we go further than existing incentives (e.g. ethics self assessment for European project propositions)?
 - Compulsory ethical discussion in scientific papers, doctoral dissertations, project reports, how?
 - Compulsory ethics review of research projects (as for medical research)?

References

- Catherine Tessier - *Research Ethics and Integrity Training for Doctoral Candidates: face-to-face is better!* ERCIM News 116, January 2019
- The University of Toulouse and Inria training material
- Léa Rogliano – *La formation à l'éthique du numérique dans l'enseignement supérieur (ingénierie et filières scientifiques) entre injonctions politiques et pratiques de terrain à l'heure de la crise écologique et de l'essor de l'intelligence artificielle. Quelles pistes d'évolution pour les formations ?* Master en éthique, Université libre de Bruxelles, novembre 2022