

# Teaching Ethics in Natural Language Processing: DIY

Karën Fort

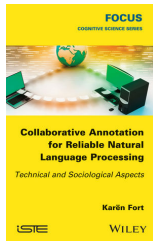
karen.fort@loria.fr / <https://members.loria.fr/KFort/>

ERCIM, Oct. 18th, 2022

# Where I'm talking from

See <https://members.loria.fr/KFort/>

- ▶ Language resources creation for NLP, esp. using crowdsourcing

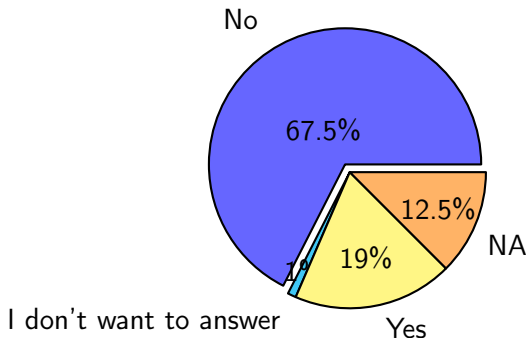


- ▶ Ethics and NLP



## Survey in the community in 2015: ethics in your curricula

*Is there any class on ethics in the curricula in which you participate? [approx. 200 responses]*



Karën Fort, Alain Couillault. Yes, We Care! Results of the Ethics and Natural Language Processing Surveys. LREC 2016.

# DIY Ethics: teaching experience

- ▶ from 2018
- ▶ in Grenoble, Tours, Nancy, Panthéon-Assas and Malta
- ▶ different contexts:
  - ▶ an introduction to a real class on the subject ("pédagogie du faire")
  - ▶ a one-shot seminar

## DIY Ethics: organization

- ▶ 2.5 to 3h
- ▶ groups of 3 to 4
- ▶ one newspaper article per group (**carefully chosen** subjects) in their language (FR/EN)
- ▶ 40 min. of reading+discussion+preparation
- ▶ 5 min. of presentation to everyone, without slides (everybody participates) + 5 min. questions
- ▶ **no** evaluation

**They** think (autonomy), **they** explain (peer-learning), **they** debate

Newsletters

*The Atlantic*

---

NEXT ECONOMY

## The Internet Is Enabling a New Kind of Poorly Paid Hell

For some Americans, sub-minimum-wage online tasks are the only work available.

By Alana Semuels

<https://www.theatlantic.com/business/archive/2018/01/amazon-mechanical-turk/551192/>

## Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter

**Attempt to engage millennials with artificial intelligence backfires hours after launch, with TayTweets account citing Hitler and supporting Donald Trump**

[https://www.theguardian.com/technology/2016/mar/24/  
tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter](https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter)

# Data minimization and informed consent

**MOTHERBOARD**  
TECH BY VICE

## Data Broker Is Selling Location Data of People Who Visit Abortion Clinics

It costs just over \$160 to get a week's worth of data on where people who visited Planned Parenthood came from, and where they went afterwards.



By [Joseph Cox](#)

<https://www.vice.com/en/article/m7vzjb/location-data-abortion-clinics-safegraph-planned-parenthood>



# Carbon footprint

ARTIFICIAL INTELLIGENCE

## Training a single AI model can emit as much carbon as five cars in their lifetimes

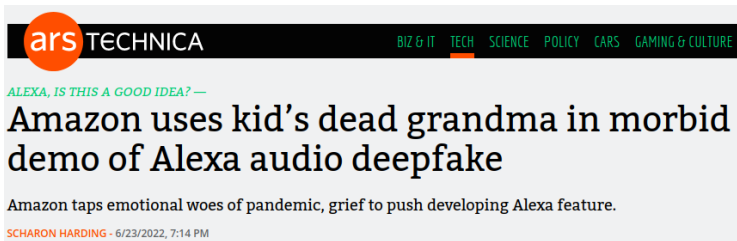
Deep learning has a terrible carbon footprint.

By Karen Hao

June 6, 2019

<https://www.technologyreview.com/2019/06/06/239031/training-a-single-ai-model-can-emit-as-much-carbon-as-five-cars-in-their-lifetimes/>

# Long term consequences



The image shows the top portion of a web article. At the top left is the Ars Technica logo, consisting of an orange circle with the word 'ars' in white lowercase letters, followed by the word 'TECHNICA' in white uppercase letters on a black background. To the right of the logo is a navigation bar with the following categories: 'BIZ & IT', 'TECH' (which is highlighted with an orange underline), 'SCIENCE', 'POLICY', 'CARS', and 'GAMING & CULTURE'. Below the navigation bar is a sub-header in green text: 'ALEXA, IS THIS A GOOD IDEA? —'. The main title of the article is 'Amazon uses kid's dead grandma in morbid demo of Alexa audio deepfake', written in a large, bold, black serif font. Below the title is a sub-headline in a smaller, black sans-serif font: 'Amazon taps emotional woes of pandemic, grief to push developing Alexa feature.' At the bottom left of the article header is the author's name and date: 'SCHARON HARDING - 6/23/2022, 7:14 PM'.

[https://arstechnica.com/gadgets/2022/06/  
amazon-uses-kids-dead-grandma-in-morbid-demo-of-alexa-audio-deepfake/](https://arstechnica.com/gadgets/2022/06/amazon-uses-kids-dead-grandma-in-morbid-demo-of-alexa-audio-deepfake/)

# Confirmation bias (and so much more)

WIRED

BACKCHANNEL BUSINESS CULTURE GEAR IDEAS SCIENCE SECURITY

SIGN IN

SUBSCRIBE

KHARI JOHNSON

BUSINESS MAR 7, 2022 7:00 AM

## How Wrongful Arrests Based on AI Derailed 3 Men's Lives

Robert Williams, Michael Oliver, and Nijeer Parks were misidentified by facial recognition software. The impact cast a long shadow.

<https://www.wired.com/story/wrongful-arrests-ai-derailed-3-mens-lives/>

# Red line?

**WIRED**

BACKCHANNEL BUSINESS CULTURE GEAR IDEAS SCIENCE SECURITY

SIGN IN

SUBSCRIBE

CARISSA VÉLIZ

IDEAS DEC 27, 2021 7:00 AM

## If AI Is Predicting Your Future, Are You Still Free?

Part of being human is being able to defy the odds. Algorithmic prophecies undermine that.

<https://www.wired.com/story/algorithmic-prophecies-undermine-free-will/>

# Opening the ethical chakras: not just consequentialism



Ethics in Bricks

@EthicsInBricks



What is good?

Kant: Act in line with duties

Aristotle: Strive for excellence

Bentham: Do the math

Arendt: Think

[Traduire le Tweet](#)



7:54 AM · 8 juin 2022 · Twitter for iPhone

# A question of time

Shows that:

- ▶ ethics is about time
- ▶ in **one hour** we can start the reflection
- ▶ as specialists, we can go further than the journalist

## The teacher's role: if necessary

- ▶ widening their perspective:
  - ▶ the world is not only about white (male) human beings living in rich countries/regions, with Internet access

# The teacher's role: if necessary

- ▶ widening their perspective:
  - ▶ the world is not only about white (male) human beings living in rich countries/regions, with Internet access
- ▶ deepening their analysis:
  - ▶ what happened technically?
  - ▶ can we avoid this? How?
  - ▶ who made the decisions?



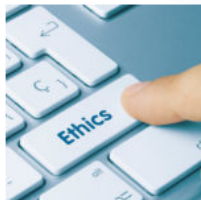
# The teacher's role: if necessary

- ▶ widening their perspective:
  - ▶ the world is not only about white (male) human beings living in rich countries/regions, with Internet access
- ▶ deepening their analysis:
  - ▶ what happened technically?
  - ▶ can we avoid this? How?
  - ▶ who made the decisions?
- ▶ making them feel how **important** (ie responsible) they are

## Next Step: doctoral training at LORIA (Nov. 3rd & 4th)

### Explore the ethical issues of computer science research through dystopia.

The workshop "Ethics in computer science: write your own dystopia" is intended for PhD students, students and researchers in computer science, natural language processing and philosophy. It aims to create a space for reflection and exchange to explore together the issues of research ethics through science fiction. It will take place on November 3 and 4, 2022 at the Château du Montet, in Vandoeuvre-lès-Nancy.



[https:](https://www.loria.fr/en/explore-the-ethical-issues-of-computer-science-research-through-dystopia/)

[//www.loria.fr/en/explore-the-ethical-issues-of-computer-science-research-through-dystopia/](https://www.loria.fr/en/explore-the-ethical-issues-of-computer-science-research-through-dystopia/)

# Next Step: doctoral training at LORIA (Nov. 3rd & 4th)

## Day 1

9h – 12h : Introduction covering the philosophical bases (Karën Fort), a case study on smart office setting (Sarah Carter) and the role of dystopia and science fiction to help thinking about ethical issues (Ilaria Tiddi).

Practical activities : creating the teams, brainstorming on research topics and their potential ethical issues

12h – 13h30: Lunch

13h30 – 17h: Practical activities: reflecting on ethical issues, preparing a backbone scenario from a chosen situation/story

17h30 – 18h: Pitching

[https:](https://www.loria.fr/en/explore-the-ethical-issues-of-computer-science-research-through-dystopia/)

[//www.loria.fr/en/explore-the-ethical-issues-of-computer-science-research-through-dystopia/](https://www.loria.fr/en/explore-the-ethical-issues-of-computer-science-research-through-dystopia/)

# Next Step: doctoral training at LORIA (Nov. 3rd & 4th)

## Day 2 Writing a dystopia

A professional writer will guide your team to create your own dystopian short story. Starting from your theme, you will identify your story's moral debate, imagine your characters and explore dramatic situations. Then you will learn to structure your storyline in order to build momentum and meaning, before actually writing the final text and ultimately sharing it with the rest of the group.

9h - 12h:

- Introduction to creative writing & dystopia
- Characters' moral debate
- Dystopian spiral

12h - 13h30: Lunch

13h30 - 17h:

- Unfolding the narrative
- Writing the story

17h - 18h30:

Sharing the stories (12mn/group including 5mn questions)

<https://www.loria.fr/en/explore-the-ethical-issues-of-computer-science-research-through-dystopia/>

## Expected outcome

- ▶ 20-25 PhD students trained, more to come
- ▶ 6 or 7 productions, hopefully publishable (leaflets, short stories, comics?)
- ▶ training to be "exported" to other labs (Inria?)

Thank you for your attention and use our Bingo!

## NLP Ethics Excuse Bingo

If I don't, someone else will	Who are you to decide?	Ethics is relative to culture	There are positive uses too
Ethics review is censorship	Science is neutral	Well, you flew to this conference	People are biased too
Negative outcomes are not predictable	Workers there are happy for \$0.05	There are no alternatives	Don't slow down progress
You want to go back to candles?	Ethics review is US imperialism	The data was publicly accessible	Stop being political

CC-BY-SA

Emily M. Bender & Karën Fort 2022

<https://twitter.com/KarnFort1/status/1532699418067193857>



Fort, K. and Couillault, A. (2016).

Yes, We Care! Results of the Ethics and Natural Language Processing Surveys.

*In international Language Resources and Evaluation Conference (LREC) 2016, Proceedings of the international Language Resources and Evaluation Conference (LREC) 2016, Portoroz, Slovenia.*