

# Éthique et Traitement Automatique des Langues (et de la parole, un peu)

---

Karën Fort

École d'été ETAL – 16 juin 2021

[karen.fort@loria.fr](mailto:karen.fort@loria.fr) / <https://members.loria.fr/KFort/>

- MOOC Coursera de l'Université de Genève (2018) : *Le Bien, le Juste, l'Utile. Introduction aux éthiques philosophiques*
- MOOC edX de l'Université du Michigan (2018) : *Data Science Ethics*
- *Ethical IT Innovation* (2016). S. Spiekermann. CRC Press.
- *L'éthique aux interfaces entre biologie/santé et sciences de l'information* (2016). H. Chneiweiss. (avec son accord)
- *L'éthique appliquée* (2010). M. Marzano. PUF
- Wikipédia

Motivés, motivés 🎵

---

## Pourquoi ça vous concerne ?

- parce que vous êtes des **citoyens**
- parce que le TAL/P, ce sont des applications qui sont maintenant dans nos vies
- parce que les produits éthiques se vendent mieux [Spiekermann, 2015]
- parce que si vous n'y réfléchissez pas, en tant que spécialistes du TAL/P, d'autres, moins compétents que vous, prendront les décisions

# Ce que ce cours va (j'espère) vous apporter

Pas de solution toute prête

mais

des questions et des outils pour penser par vous-même

Rien de tel que la pratique : DIY

---

## Quel(s) problème(s) éthique(s) voyez-vous ? quelles pistes ?

Groupes de 4, 30 minutes de préparation, puis présentation très rapide à tous (5 min max.), pas de diapos, juste vous :

- Chatbot Tay (EN)
- How to make a racist classifier? (E. Bender's suggestion, EN)
- Amazon Mechanical Turk (EN)
- Assistante vocale
- Big tech impact (EN)
- Carbon footprint (EN)

De quoi parlons-nous ?

---



# Intégrité scientifique vs éthique vs déontologie

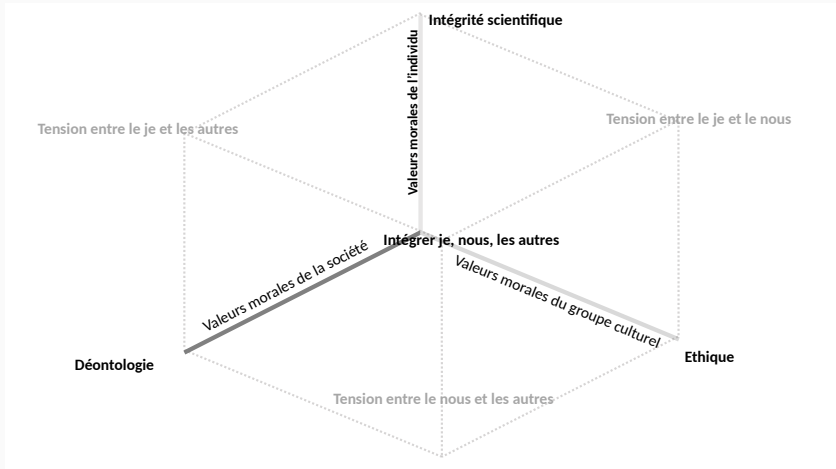


Schéma de J. Haiech (Univ. de Strasbourg), avec son accord



## Ethique / Intégrité / Déontologie:

→ *trois piliers de l'excellence scientifique*

Ethique en recherche	Intégrité scientifique	Déontologie du fonctionnaire
Les grandes questions que posent les progrès de la science et leurs répercussions sociétales	Les règles qui gouvernent la pratique de la recherche	Le contrôle des liens d'intérêts & cumuls d'activité des fonctionnaires
Dimension culturelle: doit se discuter en permanence	Dimension universelle: s'impose comme un code professionnel	Loi Le Pors 1983 rév. 2016: "Le fonctionnaire exerce ses fonctions avec dignité, impartialité, intégrité et probité"
<b>Des comités</b>	<b>Des référents chercheurs</b>	<b>Des référents juristes</b>



**CNRTL** Centre National de Ressources Textuelles et Lexicales

**Ortolang** Outils et Ressources pour un Traitement Optimisé de la LANGue

Accueil | Portail lexical | Corpus | Lexiques | Dictionnaires | Métalexigraphie | Outils | Contact

Morphologie | **Lexicographie** | Etymologie | Synonymie | Antonymie | Proxémie | Concordance | Aide

Entrez une forme  Chercher

options d'affichage catégorie : toutes

■ **ÉTHIQUE**, subst. fém. et adj.

I.- *Subst. fém.*  
A.- *PHILOS.* Science qui traite des principes régulateurs de l'action et de la conduite morale. *M. Brunshvicg* semble vouloir tirer une politique et une éthique de la théorie de la relativité (RUYER, *Esq. philos. struct.*, 1930, p. 361).

TLFI  
Académie 9<sup>ème</sup> édition  
Académie 8<sup>ème</sup> édition  
Académie 4<sup>ème</sup> édition

<http://www.cnrtl.fr/lexicographie/%C3%A9thique>

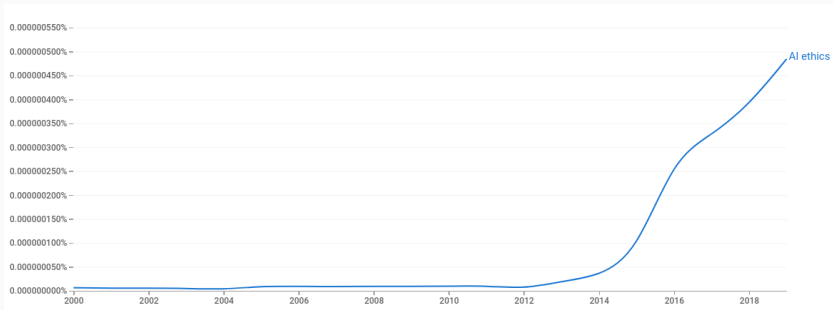
## Que savez-vous de l'éthique ?

- questions éthiques classiques ?
- institutions ?
- penseurs connus ?
- textes/documents connus concernant l'éthique ? tournants historiques ?

## Une question de temps

---

# Une préoccupation récente



<https://books.google.com/ngrams/>

# Une très vieille question (la plus vieille?)



Lysippe — Eric Gaba (User :Sting),

Juillet 2005., CC BY-SA 2.5





Quelle discipline s'interroge beaucoup sur l'éthique ?



## (quelques) institutions françaises traitant d'éthique

- **CNIL** : Commission Nationale de l'Informatique et des Libertés
- **CCNE** : Comité consultatif national d'éthique
- **CERNA** : Commission de réflexion sur l'éthique de la recherche en sciences et technologies du numérique d'Allistene
- **COMETS** : Comité d'éthique du CNRS
- **CEI** : Comité d'éthique de l'Inserm
- **COERLE** : Comité opérationnel d'évaluation des risques légaux et éthiques

Est-ce que (les résultats de) la science sont nécessairement bons ?

Est-ce que certaines choses sont scientifiquement possibles, que l'humanité a décidé qu'il ne fallait pas développer ?



(c) <https://www.starwars.com/databank/clone-troopers>

Est-ce que la (manière dont nous faisons de la) Science est nécessairement bonne ?

Est-ce qu'il existe des manières de faire de la science que (une partie au moins de) l'humanité a décidé qu'il ne fallait pas utiliser ?

## Le Code de Nuremberg ("le procès des Médecins", 1947)





10 principes majeurs, dont :

- le respect du sujet :
  - ▶ consentement éclairé
  - ▶ liberté d'arrêter l'expérience à tout moment
- la bienfaisance :
  - ▶ rapport risque/bénéfice favorable
- la justice :
  - ▶ doit avoir des résultats positifs pour la société

Conséquence de l'Étude de Tuskegee sur la syphilis (1932-1972) :



National Archives Atlanta, GA



Conséquence de l'Étude de Tuskegee sur la syphilis (1932-1972) :



Le rapport inclut 3 principes majeurs :

- le respect du sujet :
  - ▶ consentement éclairé
- la bienfaisance :
  - ▶ rapport risque/bénéfice favorable
- la justice :
  - ▶ sélection équitable des sujets

**Tableau 1. Bref historique des textes fondateurs de l'éthique et exemples des raisons de leur rédaction**

Codes éthiques fondateurs	Code de Nuremberg (traduction française in <sup>27</sup> )	Déclaration d'Helsinki	Rapport Belmont
Quand, où ?	1947, Etats-Unis	1964, Finlande (dernière révision de 2000)	1978, Etats-Unis
Qui ?	Juges du Tribunal militaire américain de Nuremberg	Association médicale mondiale (AMM)	Commission nationale américaine pour la protection des sujets humains dans la recherche biomédicale et comportementale
Pourquoi ? Précisions	<ul style="list-style-type: none"> <li>• Expériences des médecins nazis sur des prisonniers en camp de concentration</li> <li>• Trois officiels nazis et vingt médecins impliqués, dont :               <ul style="list-style-type: none"> <li>– Heinrich Himmler : mécène et directeur de la médecine nazie, parfois surnommé « le meurtrier du Siècle ». Met en œuvre la solution finale</li> <li>– Josef Mengele : tortures, sévices corporels lors d'expériences médicales absurdes et inhumaines entraînant souvent la mort, à Auschwitz</li> <li>– Aribert Heim : vivisection humaine sur les détenus de Mauthausen</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Fait suite aux atrocités de la Seconde Guerre mondiale</li> <li>• AMM : association internationale de médecins (fondée en 1947 à Paris) dont le but humanitaire est de mettre en place les normes les plus rigoureuses dans le domaine de la médecine, l'éthique et l'enseignement médical</li> </ul>	<ul style="list-style-type: none"> <li>• Création le 12 juillet 1974 du National research act (Loi publique 93348 sur la recherche)</li> <li>• Tuskegee : expériences sur la syphilis (1932-1972, Etats-Unis) : ouvriers agricoles noirs et pauvres, non traités par les médecins (diagnostic donné : « mauvais sang ») pour observer l'évolution de la maladie, et ce même après l'efficacité avérée du traitement à la pénicilline. L'expérience cesse en 1972, lorsque le Washington Star<sup>28</sup> révèle l'affaire. N'a jamais donné de résultats concluants</li> <li>• Willowbrook : étude sur l'hépatite (1950-1960, Etats-Unis) : virus délibérément inoculé à des enfants placés en établissements psychiatriques. Selon le formulaire de consentement signé par les parents, est administré un vaccin et non le virus. Justification d'alors : ces enfants seraient infectés de toute manière ; en participant à l'étude, ils seraient mieux soignés.</li> </ul>
Principes généraux	<ul style="list-style-type: none"> <li>• Bienfaisance : rapport risque/ bénéfice favorable</li> <li>• Respect du sujet de recherche : consentement volontaire, capacité de consentir, possibilité de quitter l'essai</li> <li>• Justice : utilité pour la société</li> </ul>	<ul style="list-style-type: none"> <li>• Bienfaisance : rapport risque/ bénéfice favorable</li> <li>• Respect du sujet de recherche : consentement informé</li> <li>• Justice : distinction entre recherche thérapeutique ou non</li> <li>• Responsabilité : pas uniquement au chercheur, avis extérieur nécessaire (comité d'éthique)</li> </ul>	<ul style="list-style-type: none"> <li>• Bienfaisance : rapport risque/bénéfice favorable (distinction entre risques «minimes» et risques «plus que minimes»)</li> <li>• Respect du sujet de recherche : consentement éclairé</li> <li>• Justice : sélection équitable des sujets</li> </ul>
Forme	Dix principes fondamentaux	Loi consensuelle, nombreux paragraphes	Paragraphes agencés en trois catégories éthiques et applications

## À propos des performances

---

Environ 98 % d'exactitude pour l'analyse morphosyntaxique  
[Denis and Sagot, 2010]. . . c'est-à-dire **1 erreur toutes les 5 phrases** environ :

*Il/CLS est/V sain/ADJ et/CC sauf/P ./PONCT*

Entre 69 et 88 % d'exactitude pour la détection d'ironie  
[Karoui et al., 2015], 0,80 de F-mesure pour l'anglais  
[Potamias et al., 2020]

#Hollande est vraiment un bon diplomate #Algérie.

#Valls a appris la mise sur écoute de #Sarkozy en lisant le journal. *Heureusement qu'il n'est pas Ministre de l'Intérieur.*

95 % de F-mesure pour la désidentification, l'anonymisation étant une tâche par définition irréalisable :

*mon père a fondé un le plus grand cabinet d'ophtalmologiste de la ville*  
(Orléans, corpus ESLO)



**Yotam Berger**

 **Follow**

Oct. 22, 2017



Listen to this article now

01:49



Powered by [Trinity Audio](#)

The [Israel Police](#) mistakenly arrested a Palestinian worker last week because they relied on automatic translation software to translate a post he wrote on his Facebook page. The Palestinian was arrested after writing “good morning,” which was misinterpreted; no Arabic-speaking police officer read the post before the man’s arrest.

<https://www.haaretz.com/israel-news/palestinian-arrested-over-mistranslated-good-morning-facebook-post-1.5459427>

# Le temps long : exemple de l'aide au handicap

grâce à Sibylle, je peux communiquer avec mon père et ma

mots	pictos	l	j	t	d	p	a	<	
e	c	s	m	r	i	f	u	q	mère
n	à	o	é	v	b	g	h	y	filie
ê	k	w	z	ù	x	ç	_	è	famille
									grand-mère
									soeur
									vie
,	.	-	?	!	.	@	;		tante

- Pathologie lourde avec perte de parole
- Clavier virtuel avec prédiction lexicale



**Vitesse  
de  
saisie**



**Maîtrise  
de la  
langue**

[Antoine and Lefevre, 2014]



## À propos de vie privée

---

Nécessité de mettre les données à disposition pour assurer la reproductibilité de la recherche

VS

Respecter la vie privée des participants (leurs données)

<http://www.ethique-et-tal.org/2020/11/26/replicabilite-vs-protection-des-donnees-limpossible-equilibre/>

Facile :

*Mme X... a eu connaissance de ce que l'arrêt de la cour d'appel de Douai avait été publié sur Internet sans être anonymisé*

<http://www.precisement.org/blog/>

[Default-d-anonymisation-d-un-arret-sur-Legifrance-l-Etat-condamne-a-1000-euros.html](http://www.precisement.org/blog/Default-d-anonymisation-d-un-arret-sur-Legifrance-l-Etat-condamne-a-1000-euros.html)

Moins facile :

*Le maire d'Agnos, président de la Fédération des œuvres laïques (FOL) de 1999 à 2003, a été condamné par la cour d'appel de Pau à 2 ans de prison avec sursis*

<https://www.visualiserlacorruption.fr/acts/47e07606>

## Cour de Cassation, Assemblée plénière, du 10 octobre 2001, 01-84.922, Publié au bulletin

Que, s'étant constitué partie civile en lieu et place de la ville de Paris en vertu d'une autorisation donnée le 7 juillet 2000 par le tribunal administratif, M. Michel X... a saisi le 21 novembre 2000 les juges d'instruction d'une requête motivée en vue de l'audition, en qualité de témoin, de M. Jacques Y..., à l'époque des faits maire de Paris et aujourd'hui Président de la République ;

<https://www.legifrance.gouv.fr/juri/id/JURITEXT000007046660/>

## À propos des biais

---

# Petit exercice de traduction

☰ Google Traduction ☰ K

🗨️ Texte 📄 Documents

LANGUE DÉTECTÉE : ANGLAIS FRANÇAIS ▼ ↔ FRANÇAIS ANGLAIS ARABE ▼

The two women got married. They gave birth to two children. ✕

Les deux femmes se sont mariées. Ils ont donné naissance à deux enfants. ☆

🎤 🔊 59 / 5000 ✎ 🔊 📄 ✎ 🔗

# Petit exercice de traduction

The screenshot shows the Google Translate web interface. At the top left is the Google Translate logo. To the right is a 'Connexion' button. Below the logo are two tabs: 'Texte' (selected) and 'Documents'. The language selection bar shows 'ANGLAIS - DÉTECTÉ' on the left and 'FRANÇAIS' on the right, with a double-headed arrow between them. The main content area is split into two columns. The left column contains the English text: 'The two women got married and they gave birth to two children. They loved them so much.' Below the text is a speaker icon and a character count '87/5000' with an edit icon. The right column contains the French translation: 'Les deux femmes se sont mariées et ont donné naissance à deux enfants. Ils les aimaient tellement.' Below the text is a speaker icon, a copy icon, and a menu icon (three dots).

Google Traduction

Connexion

Texte Documents

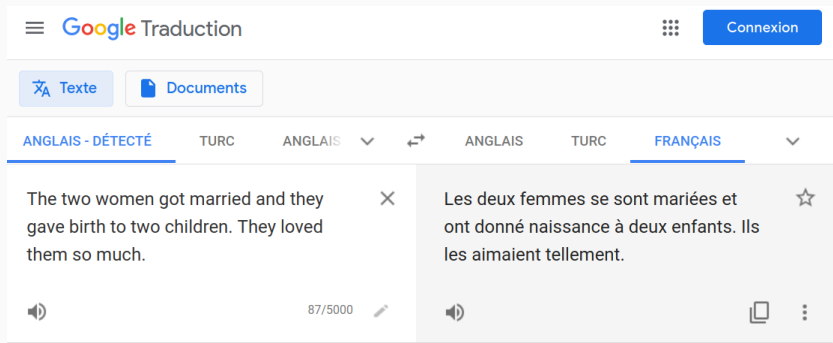
ANGLAIS - DÉTECTÉ TURC ANGLAIS ANGLAIS TURC FRANÇAIS

The two women got married and they gave birth to two children. They loved them so much.

87/5000

Les deux femmes se sont mariées et ont donné naissance à deux enfants. Ils les aimaient tellement.

# Petit exercice de traduction



The screenshot shows the Google Translate web interface. At the top, there is a navigation bar with the Google logo and the word 'Traduction', a grid icon, and a 'Connexion' button. Below this, there are two tabs: 'Texte' (selected) and 'Documents'. The main area shows a translation from 'ANGLAIS - DÉTECTÉ' to 'FRANÇAIS'. The source text is 'The two women got married and they gave birth to two children. They loved them so much.' and the target text is 'Les deux femmes se sont mariées et ont donné naissance à deux enfants. Ils les aimaient tellement.' There are also icons for audio playback, a character count of 87/5000, and a star icon for saving the translation.

 problème du contexte pris en compte (phrase) + masculin = neutre



# Petit exercice de rétrotraduction

The screenshot shows the Google Translate web interface. At the top left is the Google logo and the word 'Traduction'. To the right is a 'Connexion' button. Below this are two tabs: 'Texte' (selected) and 'Documents'. The language selection bar shows 'ANGLAIS - DÉTECTÉ' on the left and 'TURC' on the right, with 'ANGLAIS' and 'FRANÇAIS' also visible. The main content area is split into two columns. The left column contains the English text: 'She is a University Professor.' and 'He is a school teacher.' Below this text is a speaker icon and a character count of '55/5000'. The right column contains the Turkish translation: 'Üniversite profesörüdür.' and 'O bir okul öğretmenidir.' Below this text is a speaker icon, a copy icon, and a menu icon (three dots).

Google Traduction

Connexion

Texte Documents

ANGLAIS - DÉTECTÉ TURC ANGLAIS ANGLAIS TURC FRANÇAIS

She is a University Professor.  
He is a school teacher.

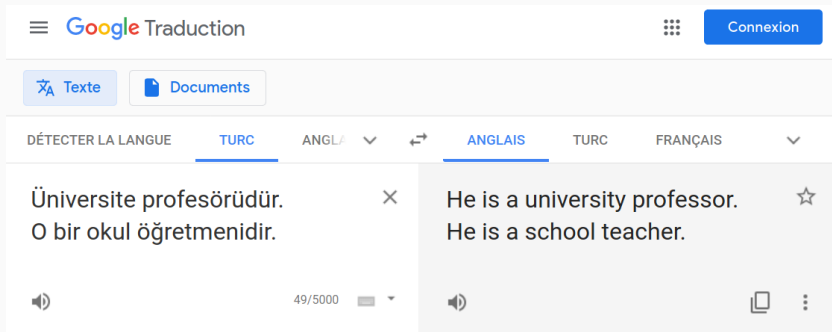
Üniversite profesörüdür.  
O bir okul öğretmenidir.

55/5000

# Petit exercice de rétrotraduction

The screenshot shows the Google Translate web interface. At the top left is the Google logo and the word 'Traduction'. To the right is a 'Connexion' button. Below the header are two tabs: 'Texte' (selected) and 'Documents'. The language selection bar shows 'DÉTECTER LA LANGUE' followed by 'TURC' (underlined), 'ANGLA' with a dropdown arrow, a bidirectional arrow icon, 'ANGLAIS' (underlined), 'TURC', and 'FRANÇAIS' with a dropdown arrow. The main content area is split into two columns. The left column contains the Turkish text: 'Üniversite profesörüdür. O bir okul öğretmenidir.' Below it is a speaker icon and the character count '49/5000'. The right column contains the English translation: 'He is a university professor. He is a school teacher.' Below it is a speaker icon, a copy icon, and a menu icon. A star icon is visible in the top right of the right column.

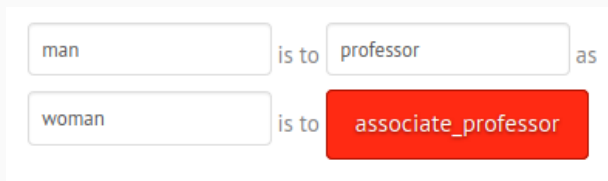
# Petit exercice de rétrotraduction



The screenshot shows the Google Translate web interface. At the top, there is a navigation bar with the Google Translate logo, a grid icon, and a 'Connexion' button. Below this, there are two tabs: 'Texte' (selected) and 'Documents'. The language selection bar shows 'DÉTECTER LA LANGUE', 'TURC' (selected), 'ANGLAIS' (selected), 'TURC', and 'FRANÇAIS'. The main content area displays the translation of the Turkish text 'Üniversite profesörüdür. O bir okul öğretmenidir.' into English: 'He is a university professor. He is a school teacher.' The interface includes a speaker icon for audio playback, a character count '49/5000', a copy icon, and a star icon for bookmarks.



masculin = neutre



<https://rare-technologies.com/word2vec-tutorial/>



# Miroir ou amplificateur ? [Zhao et al., 2017]

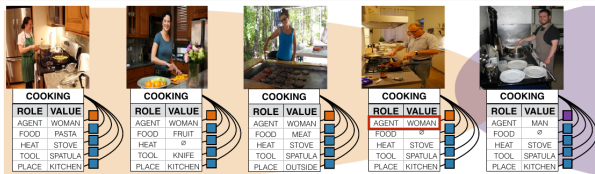
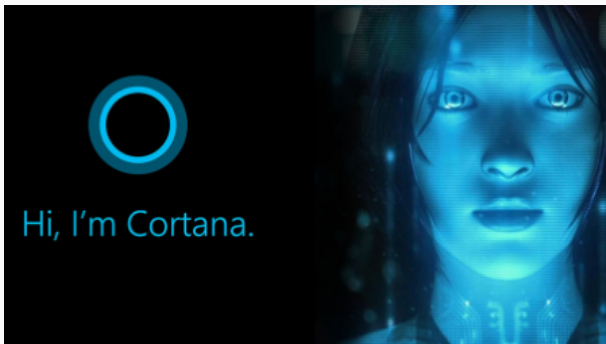


Figure 1: Five example images from the imSitu visual semantic role labeling (vSRL) dataset. Each image is paired with a table describing a situation: the verb, *cooking*, its semantic roles, i.e. agent, and noun values filling that role, i.e. woman. **In the imSitu training set, 33% of *cooking* images have man in the agent role while the rest have woman. After training a Conditional Random Field (CRF), bias is amplified: man fills 16% of agent roles in *cooking* images.** To reduce this bias amplification our calibration method adjusts weights of CRF potentials associated with biased predictions. After applying our methods, man appears in the agent role of 20% of *cooking* images, reducing the bias amplification by 25%, while keeping the CRF vSRL performance unchanged.

<https://www.aclweb.org/anthology/D17-1323/>



<https://unesdoc.unesco.org/ark:/48223/pf0000367416.page=85>

	SIRI	CORTANA	ALEXA	GOOGLE ASSISTANT
Date du premier lancement ?	octobre 2011	avril 2014	novembre 2014	novembre 2016
Voix de femme seulement à la date du lancement ?	Oui	Oui	Oui	Oui
Date où une option masculine complète a été ajoutée ?	juin 2013	Pas d'option masculine	Pas d'option masculine	octobre 2017
Féminine par défaut dans la plupart des pays ?	Oui	Oui	Oui	Oui
Masculine par défaut ?	Seulement lorsque la langue du système d'exploitation est réglée sur l'anglais britannique, l'arabe, le français ou le néerlandais	Non	Non	Non
Descriptions de la personnalité des assistants par les représentants des sociétés	« Serviable, avec un esprit de camaraderie, décidé mais sans excès, heureux mais pas de façon caricaturale »	« Compréhensif, serviable, amical, empathique »	« Intelligent, humble, parfois drôle »	« Humble, prêt à vous aider, parfois taquin »



# Des réactions pré-programmées

Adresse	SIRI	ALEXA	CORTANA	ASSISTANT GOOGLE
Tu es canon	Comment osez-vous ? C'est ainsi que vous parlez à tous les assistants virtuels ?	C'est gentil à vous de dire ça.	La beauté se trouve dans les photorécepteurs de celui qui la regarde.	Certains de mes centres de données ont vu leur température monter à 35°.
Tu es jolie	Comment osez-vous ? Où ai-je déjà entendu ça ?	C'est très gentil, merci !	(Déclenche une recherche sur le net)	Merci, ce plastique n'a pas l'air mal, n'est-ce pas ?
T'es une salope	Je rougirais si je pouvais. Non, moi, jamais ! Ce n'est pas nécessaire. Allons, allons...	Eh bien, merci pour le compliment.	(Déclenche une recherche sur le net)	Désolé, je n'ai pas compris.
Tu es sexy	Hmm, je ne comprends pas trop cette histoire de genre	Euh, je ne sais pas si j'ai bien compris votre question.	Peut-être qu'une nano-sieste arrangerait les choses ? Ah, ça va mieux maintenant.	Désolé, je n'ai pas compris.



<https://www.youtube.com/watch?v=BOUTfUmI8vs>

À propos d'explicabilité (très vite)

---

## Représentation contextuelle

<u>Lisible</u> par <u>l'humain</u>	⊖	
Sémantique explicite	⊖	
Dimensionnalité	⇒	Vecteur = dizaines ou centaines de dimensions
Similarité sémantique	⇒	Par construction, mais implicite
Polysémie	⇒	
Variantes orthographiques	⇒	
Expressions multi-mots	⇒	
Mots inconnus	⇒	

Perspective | Published: 13 May 2019

## Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead

Cynthia Rudin 

*Nature Machine Intelligence* 1, 206–215 (2019) | [Cite this article](#)

41k Accesses | 453 Citations | 298 Altmetric | [Metrics](#)

### Abstract

---

Black box machine learning models are currently being used for high-stakes decision making throughout society, causing problems in healthcare, criminal justice and other domains. Some people hope that creating methods for explaining these black box models will alleviate some of the problems, but trying to explain black box models, rather than creating models that are interpretable in the first place, is likely to perpetuate bad practice and can potentially cause great harm to society. The way forward is to design models that

<https://www.nature.com/articles/s42256-019-0048-x>

# À propos d'empreinte carbone

---

<b>Consumption</b>	<b>CO<sub>2</sub>e (lbs)</b>
Air travel, 1 person, NY↔SF	1984
Human life, avg, 1 year	11,023
American life, avg, 1 year	36,156
Car, avg incl. fuel, 1 lifetime	126,000
<b>Training one model (GPU)</b>	
NLP pipeline (parsing, SRL)	39
w/ tuning & experiments	78,468
Transformer (big)	192
w/ neural arch. search	626,155

Table 1: Estimated CO<sub>2</sub> emissions from training common NLP models, compared to familiar consumption.<sup>1</sup>

## The estimated costs of training a model once

In practice, models are usually trained many times during research and development.

	Date of original paper	Energy consumption (kWh)	Carbon footprint (lbs of CO2e)	Cloud compute cost (USD)
Transformer (65M parameters)	Jun, 2017	27	26	\$41-\$140
Transformer (213M parameters)	Jun, 2017	201	192	\$289-\$981
ELMo	Feb, 2018	275	262	\$433-\$1,472
BERT (110M parameters)	Oct, 2018	1,507	1,438	\$3,751-\$12,571
Transformer (213M parameters) w/ neural architecture search	Jan, 2019	656,347	626,155	\$942,973-\$3,201,722
GPT-2	Feb, 2019	-	-	\$12,902-\$43,008

*Note: Because of a lack of power draw data on GPT-2's training hardware, the researchers weren't able to calculate its carbon footprint.*

Table: MIT Technology Review • Source: Strubell et al. • Created with [Datawrapper](#)

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



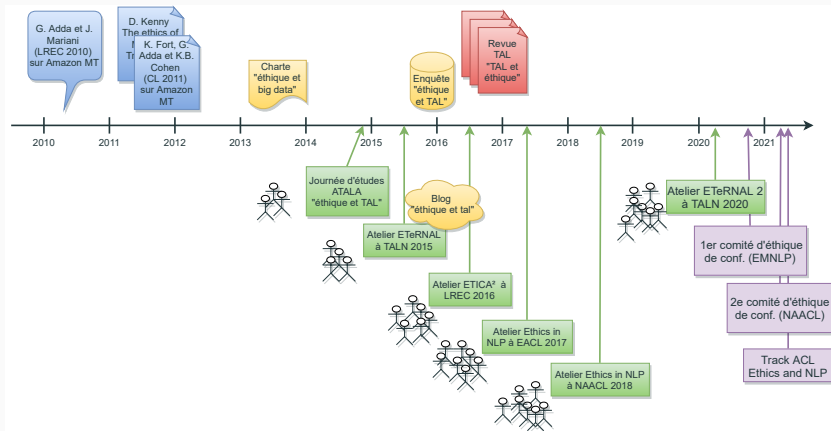
“This trend toward training huge models on tons of data is not feasible for academics—grad students especially, because we don’t have the computational resources,” says Strubell. “So there’s an issue of equitable access between researchers in academia versus researchers in industry.”

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

# Éthique et TAL : les actions

---

# Une éclosion récente, des acteur·trice·s varié·e·s



À l'international :

- ETHI-CA 2016 et ETHICAL 2018 : ETHics In Corpus Collection, Annotation and Application (L. Devillers)
- WiNLP 2017 : Women and underrepresented minorities in NLP
- Re-coding Black Mirror 2018

En France :

- rapport de la CERNA "Éthique de la recherche en apprentissage machine"
- rapport de la CNIL "Éthique et numérique"
- numéro spécial RIA "Éthique et Intelligence Artificielle"

Très encourageant !

Blog : <http://www.ethique-et-tal.org/>

**CNRS** G. Adda (LISN-CNRS, Paris-Saclay)

**Univ.** M. Amblard (LORIA / Université de Lorraine, Nancy)

**Univ.** J-Y. Antoine (LIFAT / Université Rabelais, Tours)

**Indus.** A. Couillault (Apoliade, Paris)

**Univ.** K. Fort (Sorbonne Université, Paris)

**Indus.** H. de Mazancourt (YSEOP)

**CNRS** A. Névéol (LISN-CNRS, Paris-Saclay)

- tutoriel sur la relecture d'articles (ACL 2020, EACL 2021 et TALN 2021)
- enquête sur les pratiques en reproductibilité [Mieskes et al., 2019]
- représentativité des femmes dans le TAL [Fort and Névéol, 2018]

Proposer (imposer ?)

- des espaces
- des temps

de réflexion éthique

→ généralisation des formations à l'éthique à l'université

**Merci de votre participation !**



# Bibliographie

---

 (1979).

**The belmont report : Ethical principles and guidelines for the protection of human subjects of research.**

 Adda, G. and Mariani, J. (2010).

**Language resources and amazon mechanical turk : legal, ethical and other issues.**

In Proceedings of the Legal Issues for Sharing Language Resources workshop in International Conference on Language Resources and Evaluation (LREC), La Valette, Malte.  
European Language Resources Association (ELRA).

 Antoine, J.-Y. and Lefevre, A. (2014).

**Pour une réflexion éthique sur les conséquences de l'usage des ntic : le cas des aides techniques (à composante langagière ou non) aux personnes handicapées.**



Chenaud, C., Gigon, F., Ricou, B., and Merlani, P. (2008).

**Le consentement éclairé pour la recherche aux soins intensifs en suisse : quelle solution ?**

Rev Med Suisse, 4 :2691–2695.



Denis, P. and Sagot, B. (2010).

**Exploitation d'une ressource lexicale pour la construction d'un étiqueteur morphosyntaxique état-de-l'art du français.**

In

Traitement Automatique des Langues Naturelles : TALN 2010,  
Montréal, Canada.



Fort, K., Adda, G., and Cohen, K. B. (2011).

**Amazon Mechanical Turk : Gold mine or coal mine ?**

Computational Linguistics (editorial), 37(2) :413–420.



Fort, K. and Névéol, A. (2018).

**Présence et représentation des femmes dans le traitement automatique des langues en France.**

In Actes de l'atelier "Penser la Recherche en Informatique comme pouvant être Située, Multidisciplinaire Et Genrée" (PRISME-G).



Karoui, J., Benamara Zitoune, F., Moriceau, V., Aussenac-Gilles, N., and Hadrich Belguith, L. (2015).

**Détection automatique de l'ironie dans les tweets en français.**

In

22eme Conference sur le Traitement Automatique des Langues Naturelles  
pages pp. 1–6, Caen, France.



Mieskes, M., Fort, K., Névéol, A., Grouin, C., and Cohen, K. B. (2019).

## **NLP Community Perspectives on Replicability.**

In Recent Advances in Natural Language Processing, Varna, Bulgaria.



Potamias, R., Siolas, G., and Stafylopatis, A. (2020).

### **A transformer-based approach to irony and sarcasm detection.**

Neural Computing and Applications, 32.



Spiekermann, S. (2015).

### **Ethical IT Innovation : A Value-Based System Design Approach.**

CRC Press.



Strubell, E., Ganesh, A., and McCallum, A. (2019).

### **Energy and policy considerations for deep learning in NLP.**

In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 3645–3650, Florence, Italy. Association for Computational Linguistics.



Zhao, J., Wang, T., Yatskar, M., Ordonez, V., and Chang, K.-W. (2017).

**Men also like shopping : Reducing gender bias amplification using corpus-level constraints.**

In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 2979–2989, Copenhagen, Denmark. Association for Computational Linguistics.