

*(L'homme) sait souvent ce qu'il fait,
mais il ne sait pas ce que fait ce qu'il fait.* [Paul Valéry]

Éthique et Traitement Automatique des Langues

Karën Fort

23 janvier 2018



Éthique et Traitement Automatique des Langues (TAL)

La langue : un matériau particulier

Des applications qui impactent nos vies

Des chercheurs employeurs

Enjeux éthiques

Les actions du groupe Éthique et TAL

Conclusion : et maintenant ?

Éthique et Traitement Automatique des Langues (TAL)

La langue : un matériau particulier

Des applications qui impactent nos vies

Des chercheurs employeurs

Enjeux éthiques

Les actions du groupe Éthique et TAL

Conclusion : et maintenant ?

Ambiguïtés à tous les étages (1/2)

B. Habert (ENS de Lyon)

- ▶ Délimiter les unités : une même suite de sons ou de mots peut être découpée de plusieurs manières :

mon beau-frère et ma soeur ; mon beau-frère est masseur ; mon beau-frère hait ma sœur

(il a cassé) (sa pipe) ; (il) (a cassé sa pipe)

- ▶ Délimiter les groupes :

(Nadine couvre) (la corbeille de fleurs) ;

(Nadine couvre la corbeille) (de fleurs)

- ▶ Déterminer la fonction des groupes :

Il attend la nuit

[que la nuit vienne | pendant la nuit]

Ambiguïtés à tous les étages (2/2)

B. Habert (ENS de Lyon)

- ▶ Déterminer les relations grammaticales entre groupes :

Georges admire Marie autant que Jean [Georges admire Jean | Jean admire Marie]

- ▶ Déterminer les relations sémantiques entre groupes :

Toutes les victimes n'avaient pas été vaccinées [il existe des victimes non vaccinées | l'ensemble des victimes n'a pas été vacciné];

Chacun cherche son chat [tous le même | chacun le sien]

- ▶ Déterminer le but pratique visé :

Cela ne se dit pas [constat | obligation]

Du jeu dans la langue

B. Habert (ENS de Lyon)

L'ambiguïté n'est pas forcément un défaut des langues, mais aussi la possibilité de dire « plusieurs choses » à la fois :

*Dans le milieu (c'est le cas de le dire) du football...
[entourage | réseau basé sur la corruption]*

On joue aussi avec la langue, en poésie, dans les jeux de mots :

*Un membre de l'Assemblée nationale perd son calme :
La moitié de ce parlement est composée d'imbéciles !
On exige des excuses.
Les voici : Je retire ce que j'ai dit. La moitié de ce
Parlement n'est pas composée d'imbéciles !*

La langue vit

B. Habert (ENS de Lyon)

Le sens d'un mot, d'une construction, d'un texte n'est pas univoque :

Exemple : le rôle de l'impératif dans

Empruntez le passage souterrain / Buvez Coca-Cola

Une langue « bouge » : arrivée de nouveaux mots, de nouvelles constructions. Les règles d'une langue évoluent nécessairement, pas celles d'un langage artificiel

- *Monsieur, n'est-ce pas vous qui vous appelez*

Sganarelle ?

- *Eh ! qui ?*

- *Je vous demande si ce n'est pas vous qui se nomme*

Sganarelle [Molière, 17ème siècle]

Conséquences : des analyses plus ou moins maîtrisées

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]

#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 (anonymisation) et une tâche par définition irréalisable :

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

Éthique et Traitement Automatique des Langues (TAL)

La langue : un matériau particulier

Des applications qui impactent nos vies

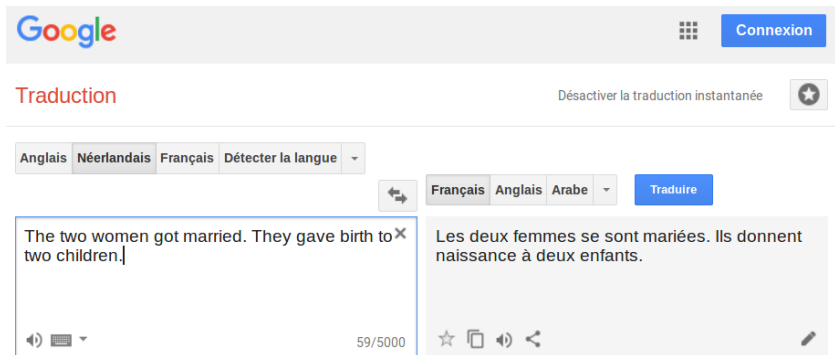
Des chercheurs employeurs

Enjeux éthiques

Les actions du groupe Éthique et TAL

Conclusion : et maintenant ?

Traduction automatique



The screenshot shows the Google Translate web interface. At the top left is the Google logo. To its right is a grid icon and a blue button labeled "Connexion". Below the logo, the word "Traduction" is displayed in red. To the right of "Traduction" is the text "Désactiver la traduction instantanée" and a star icon in a square. Below this is a horizontal line. Underneath the line, there are language selection buttons: "Anglais", "Néerlandais", "Français", and "Détecter la langue" with a dropdown arrow. To the right of these is a double-headed arrow icon. Further right are buttons for "Français", "Anglais", and "Arabe" with a dropdown arrow, followed by a blue "Traduire" button. Below the language buttons is a text input area containing the English text: "The two women got married. They gave birth to two children." with a close icon (X) at the end. Below the input area are a speaker icon, a text icon with a dropdown arrow, and the character count "59/5000". To the right of the input area is the translated text in French: "Les deux femmes se sont mariées. Ils donnent naissance à deux enfants." Below the translated text are icons for a star, a copy icon, a speaker icon, and a share icon, followed by an edit icon (pencil).

<https://translate.google.com/>

Aide au handicap

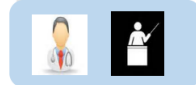
grâce à Sibylle, je peux communiquer avec mon père et ma

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]

Désidentification (anonymisation)

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/>

[Defaut-d-anonymisation-d-un-arret-sur-Legifrance-l-Etat-condamne-a-1000-euros.html](#)

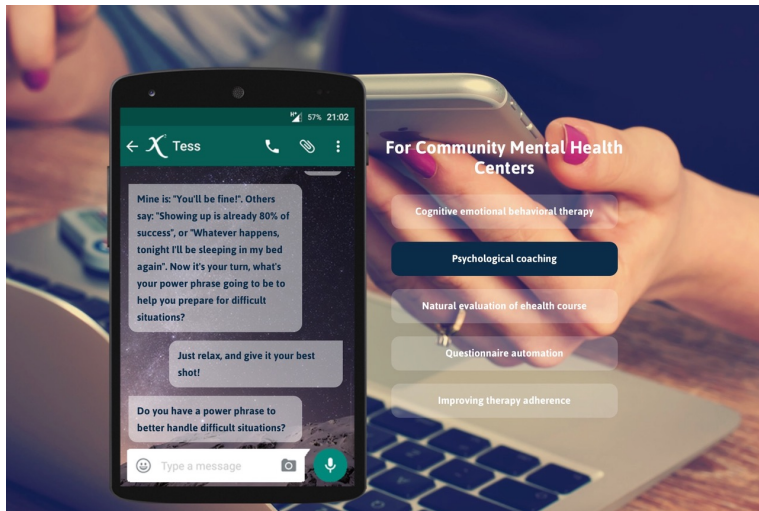
Moins facile :

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>

Analyse de tweets

ici, pour le monitoring de la santé mentale



<https://x2.ai/>

Éthique et Traitement Automatique des Langues (TAL)

La langue : un matériau particulier

Des applications qui impactent nos vies

Des chercheurs employeurs

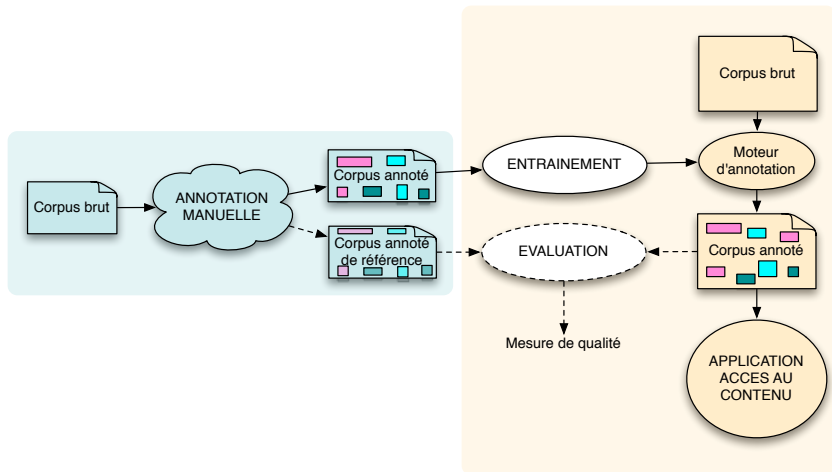
Enjeux éthiques

Les actions du groupe Éthique et TAL

Conclusion : et maintenant ?

Sous le capot du TAL

Les outils sont majoritairement à base d'apprentissage supervisé (dans la recherche)



L'ogre a faim !

pour le nourrir, nous devenons des employeurs

Nécessité de grandes masses de données **annotées** pour

entraîner

et

évaluer les systèmes

Exemple :

→ 100 000 mots annotés pour entraîner un tagger :

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

⇒ besoin d'**annotateurs humains**, ce qui coûte cher (600 000 \$ pour le Prague Treebank)

Éthique et Traitement Automatique des Langues (TAL)

La langue : un matériau particulier

Des applications qui impactent nos vies

Des chercheurs employeurs

Enjeux éthiques

Les actions du groupe Éthique et TAL

Conclusion : et maintenant ?

Des questions éthiques variées

D'un point de vue conséquentialiste

- ▶ applications qui représentent un risque concernant la vie privée des citoyens :
 - ▶ extraction automatique de « sentiments »,
 - ▶ identification de « leaders d'opinion », etc.
- ▶ applications dont les effets négatifs ne sont pas toujours évalués sur le temps long :
 - ▶ aide au handicap : qui peut faire perdre de l'autonomie
 - ▶ systèmes par apprentissage : qui peuvent être biaisés par la manière dont ils sont entraînés et généraliser des stéréotypes

Des questions éthiques variées

D'un point de vue déontologique

D'un point de vue déontologique :

- ▶ des outils performants, mais pas autant qu'on pourrait le croire :
 - ▶ désidentification (anonymisation) forcément imparfaite,
 - ▶ ironie difficile à détecter, etc.
- ▶ une maîtrise imparfaite (impossible ?) de l'apprentissage automatique :
 - ▶ biais à l'entraînement
- ▶ utilisation de données du Web, plus ou moins légalement :
 - ▶ *Twitter*, *Doctissimo*, etc.
- ▶ « uberisation » de la production de données :
 - ▶ plate-formes à la *Amazon Mechanical Turk*

Des enjeux largement sous-estimés

Des chercheurs

- ▶ employeurs
- ▶ qui ne sont pas soumis au secret (« médical ») et ne peuvent pas s'en réclamer
- ▶ qui ignorent les règles de droit, quand elles sont claires (pas de *fair-use* en France)
- ▶ qui n'ont pas de vrais moyens pour assurer leur rôle de lanceur d'alertes

Des questions politiques et organisationnelles

Évolutions récentes :

- ▶ temps de la recherche raccourci (projets sur 3 ans, voire moins)
- ▶ évaluation quantitative de la recherche (h-index)
- ▶ recherche de plus en plus appliquée et en lien avec les entreprises

Conséquences éthiques ?

Éthique et Traitement Automatique des Langues (TAL)

Les actions du groupe Éthique et TAL

- Créer des espaces

- Prendre le pouls de la communauté

- Proposer des leviers

Conclusion : et maintenant ?

Groupe éthique et TAL

Des chercheur-se-s et des industriels

CNRS G. Adda (LIMSI-CNRS, Paris-Saclay)

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

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

Indus. A. Couillault (Apoliade, Paris)

Univ. K. Fort (U. Paris-Sorbonne, Paris)

Indus. H. de Mazancourt (Mazancourt Conseil)

CNRS A. Névéol (LIMSI-CNRS, Paris-Saclay)

Éthique et Traitement Automatique des Langues (TAL)

Les actions du groupe Éthique et TAL

- Créer des espaces

- Prendre le pouls de la communauté

- Proposer des leviers

Conclusion : et maintenant ?

Se rencontrer, échanger, produire

- ▶ Journées d'études et ateliers dans la communauté :
 - ▶ journée d'études ATALA « éthique et TAL » (22 nov. 2014)
 - ▶ atelier ETeRNAL (conférence TALN, 22 juin 2015)
 - ▶ numéro spécial de la revue TAL (fin 2016) : « TAL et éthique » (57-2)
- ▶ Journées inter-disciplinaires, notamment « éthique et données » (28 janvier 2016)
 - ▶ échanger de bonnes pratiques
 - ▶ mettre au jour des problèmes pas toujours vus dans les autres disciplines
- ▶ Blog <http://www.ethique-et-tal.org/>, notamment sur :
 - ▶ la relecture par les pairs [Névéol et al., 2017]
 - ▶ les femmes dans le TAL [Fort and Névéol, 2018]
 - ▶ évaluation et classements dans les *shared tasks*
 - ▶ réflexions et annonces (journées CERNA) diverses

Éthique et Traitement Automatique des Langues (TAL)

Les actions du groupe Éthique et TAL

- Créer des espaces

- Prendre le pouls de la communauté

- Proposer des leviers

Conclusion : et maintenant ?

Donner la parole aux chercheur·se·s en TAL

K. Fort et A. Couillault (U. de la Rochelle)

2 questionnaires adressés à la communauté francophone (FR) puis internationale (INT), en 2015

- ▶ publicité sur les listes de diffusion du domaine
- ▶ 100 personnes pour le questionnaire FR (180 à 200 participants à TALN)
- ▶ 200 personnes pour le questionnaire INT (1 000 participants à ACL)

En pratique :

- ▶ résultats disponibles en ligne
<http://www.schplaf.org/kf/pdf/EthicsAndNLPResults.zip>
- ▶ analyse dans une publication [Couillault et al., 2014]

Des résultats relativement homogènes, sauf

Avez-vous déjà refusé ou limité un projet pour des raisons éthiques ?

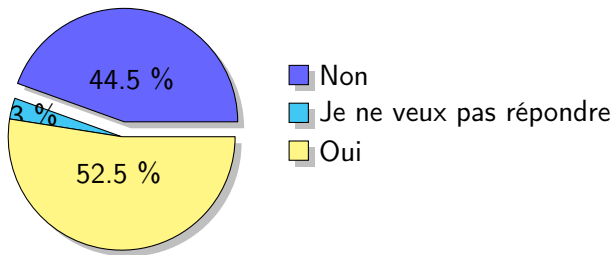
Oui :

- ▶ 40,20 % dans l'étude FR
- ▶ 23,50 % dans l'étude INT

→ lié à l'obligation de prendre l'éthique en compte lors du dépôt de projets européens (?)

Responsabilité des chercheur·se·s (*moral buffer*)

Vous considérez-vous responsable des utilisations faites des outils que vous développez ?



Questionnaire INT

Oui, c'est tout à fait mon rôle	40,20 %
C'est un rôle partagé par l'ensemble de l'équipe	34,31 %
C'est le rôle d'un des membres de l'équipe	0,98 %
Mon sujet de recherche n'est pas pertinent pour cette question	10,78 %
Non, ce n'est pas la responsabilité du chercheur	23,53 %

Questionnaire FR

Responsabilité des chercheur·se·s : commentaires (-) FR

« C'est l'utilisation des savoirs qui est soumise au questionnement éthique, non pas la science même. Occupons-nous des questions scientifiques et laissons les questions éthiques aux utilisations. »

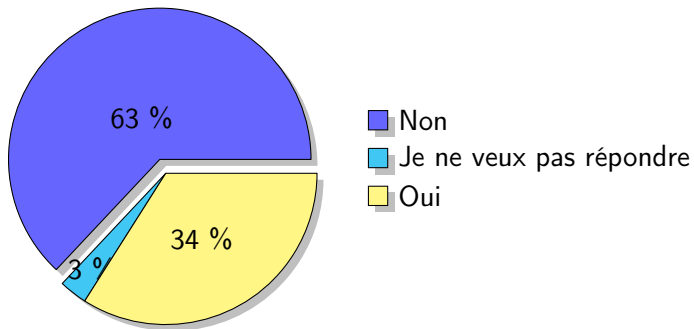
« je crois que cela concerne beaucoup de domaines, mais qu'il n'est pas nécessairement de la responsabilité du chercheur de se poser ces questions »

« Le principe de précaution et l'exception culturelle sont les mamelles du déclin. »

Aucun commentaire de ce type dans la version internationale, mais des réflexions et suggestions intéressantes (voir annexe)

Lancer une alerte ?

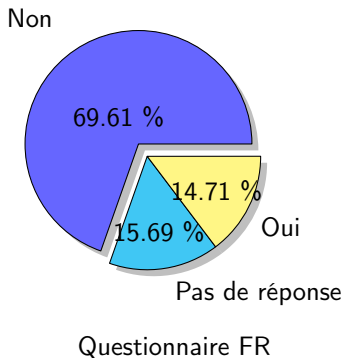
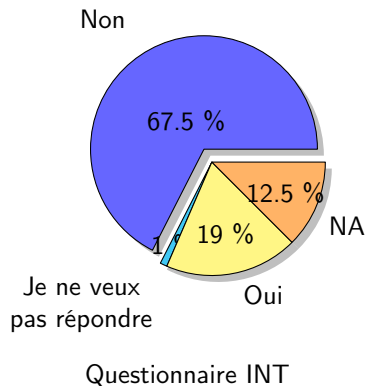
Do you know of any way to blow the whistle (issue an alert) in case of a major ethical infringement in your domain?



Questionnaire INT uniquement

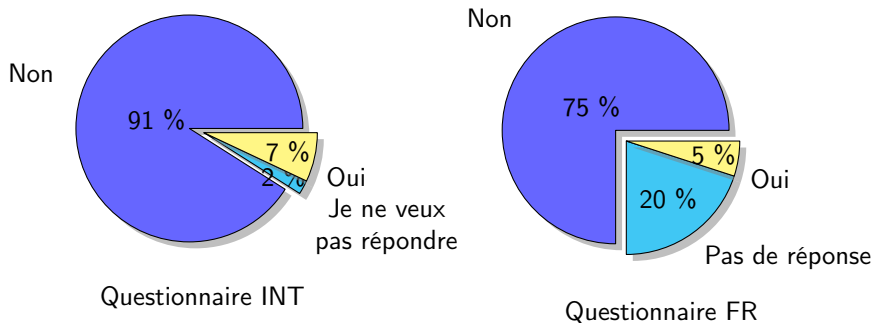
Formation à l'éthique

Existe-t-il une sensibilisation à l'éthique dans les formations dans lesquelles vous intervenez ?



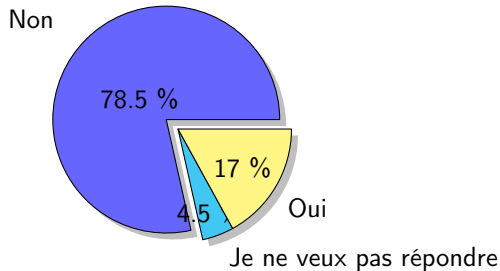
Rapports avec le citoyen

Pensez-vous que le grand public est conscient des limites des capacités des outils de TAL?

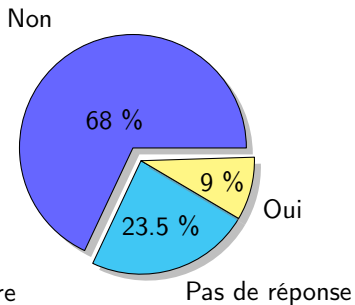


Rapports avec les pouvoirs publics

Pensez-vous que les pouvoirs publics sont conscients des limites des capacités des outils de TAL?



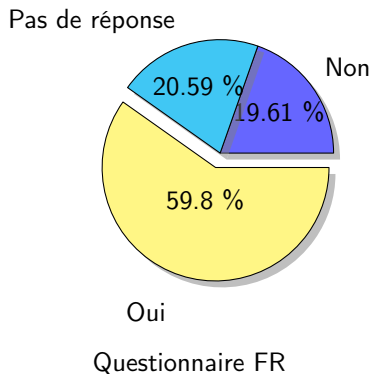
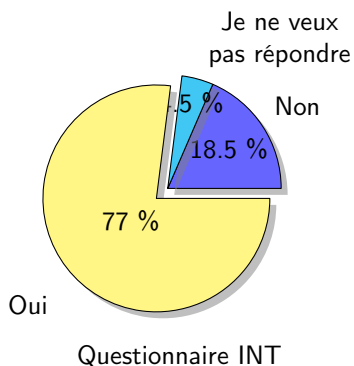
Questionnaire INT



Questionnaire FR

L'éthique dans les appels à publications

Pensez-vous que l'éthique doit faire partie des sujets de l'appel général des conférences du domaine ?



Volonté de participer à un groupe de travail

Êtes-vous d'accord pour participer à un groupe de travail sur l'éthique dans le TAL?

- ▶ FR : 28 volontaires ont laissé leur email
- ▶ INT : 73 volontaires (31 ont laissé leur email)

→ une communauté à **faire vivre**!

Éthique et Traitement Automatique des Langues (TAL)

Les actions du groupe Éthique et TAL

- Créer des espaces

- Prendre le pouls de la communauté

- Proposer des leviers

Conclusion : et maintenant ?

Charte éthique (et Big Data) : création

Des rédacteurs et des contributeurs (animateur : A. Couillault)



Processus

Version 1 (fin 2012)

- ▶ 1 réunion par mois, de juin à décembre 2012
- ▶ une validation par chaque organisme

Version 2 (fin 2015)

- ▶ élargie à de nouveaux intervenants : GFII, INSERM, CIGREF, ESG, Capital Games
- ▶ mise en cohérence avec : Ethical Guidelines H2020, CNIL, données de santé

But

Faire adopter la charte par les agences de moyens, afin que celles-ci puissent définir une **politique de sélection**.

Charte éthique (et Big Data) : forme

Charte auto-déclarée, remplie par le producteur de données, disponible en ligne en français et en anglais :

<http://wiki.ethique-big-data.org>

Utilisable

- ▶ *a priori* (avant la création de ladite donnée)
 - sélection des projets
- ▶ *a posteriori*
 - documentation des données produites

Charte éthique (et Big Data) : contenu

Documenter des données en termes de :

1. traçabilité

- ▶ qui a fait quoi ?
- ▶ statut, rémunération ?
- ▶ processus qualité ?
- ▶ consentement, anonymisation ?

2. propriété intellectuelle

- ▶ licences
- ▶ effet « matriochka » (imbrication)

3. législation(s) spécifique(s)

- ▶ déclaration à la CNIL, *fair use*, etc

Éthique et Traitement Automatique des Langues (TAL)

Les actions du groupe Éthique et TAL

Conclusion : et maintenant ?

Autres actions pour le TAL

À l'international :

- ▶ ETHI-CA 2016 et ETHICAI 2018 : ETHics In Corpus Collection, Annotation and Application (L. Devillers)
- ▶ Ethics in NLP 2017 et 2018
- ▶ 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 !

Donner les moyens des choix politiques

- ▶ valoriser l'implication dans l'éthique du domaine (y compris dans l'évaluation des chercheurs)
- ▶ proposer|imposer la Charte dans toutes les agences de moyens (déjà adoptée par *Cap Digital*)
- ▶ imposer une analyse éthique pour toute réponse à AAP
- ▶ communiquer de manière plus nuancée
 - ▶ aux politiques
 - ▶ au grand public
- ▶ créer une véritable **protection** du chercheur et du lanceur d'alertes

Surtout...

Proposer (imposer ?)

- ▶ des espaces
- ▶ des temps

de réflexion éthique

→ journée CERNA éthique et TAL

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

Merci



<http://www.moddb.com/games/bloop-boop/news/ethics-of-monetisation>

Annexes

Commentaires de l'enquête

Commentaires à noter

Propositions, suggestions

Licence spécifique :

I would greatly appreciate a kind of open source license that explicitly excludes the use of my code for military purposes or by military organizations

Plate-forme éthique de travail parcellisé :

Remunerating workers should be (I think) part of a larger goal of getting most value out of crowdworkers' time (as well as ensuring fairness within platforms for organizing crowdwork). Part of this may be an initiative to build a (non-AMT) platform that also allows you to do many things that AMT doesn't allow you to do (recruiting people who are native speakers, paying with some EU-compatible way instead of a US credit card)

Remarques

Analyser les erreurs :

While there was some tremendous progress in nearly all fields of NLP through the statistical/machine learning approach in the last decade or so, often consequences of the errors the systems make are not discussed. By focusing on quantitative evaluation we often omit a qualitative grid for error classification (although for a good reason : comparability).

Des ressources pour se protéger :

The breaches of individual freedom brought about by NLP methods and tools are insufficiently matched by similar resources given to individuals to defend themselves against the intrusions.

À noter

Ethics apply to lower members in practice, neither to leaders nor to institutions

I have organized a seminar on Ethics for PhD students at UPMC (University Pierre and Marie Curie) a few years ago. After two or three years, I had to stop, because of the lack of institutional support.



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.

In Actes de la journée ATALA Éthique et TAL.



Couillault, A., Fort, K., Adda, G., and De Mazancourt, H. (2014).

Evaluating Corpora Documentation with regards to the Ethics and Big Data Charter.

In

International Conference on Language Resources and Evaluation (LREC) 2014, Reykjavik, Islande.



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. 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.



Névéol, A., Fort, K., and Hwa, R. (2017).

Report on EMNLP Reviewer Survey.

Technical report, Association for computational linguistics.