



Éthique et Traitement Automatique des Langues DIY

Karën Fort

karen.fort@sorbonne-universite.fr

27 novembre 2018



Sources d'inspiration

- ▶ 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

Éthique et Traitement Automatique des Langues (TAL)

La langue : un matériau particulier

Des applications qui impactent nos vies

Motivation et but

Rien de tel que la pratique : DIY

Une question de temps

Éthique et TAL

Éthique et Traitement Automatique des Langues (TAL)

La langue : un matériau particulier

Des applications qui impactent nos vies

Motivation et but

Rien de tel que la pratique : DIY

Une question de temps

Éthique et TAL

Ambiguïtés à tous les étages (1/2)

B. Habert (ex. 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 (ex. 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 (ex. 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 (ex. 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



Motivation et but


Rien de tel que la pratique : DIY

Une question de temps


Éthique et TAL

Traduction automatique



[Connexion](#)

Traduction Désactiver la traduction instantanée 






Anglais Néerlandais Français Détecter la langue ▼

 Français Anglais Arabe ▼ [Traduire](#)

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

  ▼ 59/5000

Les deux femmes se sont mariées. Ils donnent naissance à deux enfants.



















    

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

Aide au handicap

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

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

mère			
filie			
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/>

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

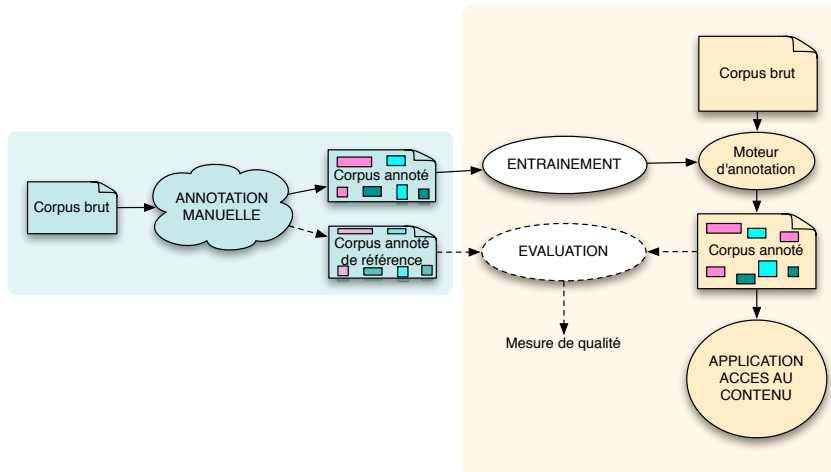
Moins facile :

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>

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

Motivation et but

Rien de tel que la pratique : DIY

Une question de temps

Éthique et TAL

Éthique ?

soft law

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 Métalexicographie 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. Brunswick semble vouloir tirer une politique et une éthique de la théorie de la relativité (RUYER, Esq. philos. struct., 1930, p. 361):

<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 ?

Pourquoi ça vous concerne ?

- ▶ parce que vous êtes des citoyens
- ▶ parce que les produits éthiques se vendent mieux [Spiekermann, 2015]
- ▶ parce que si vous n'y réfléchissez pas, en tant que spécialistes du TAL, 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 outils pour penser par vous-même

Éthique et Traitement Automatique des Langues (TAL)

Rien de tel que la pratique : DIY

Une question de temps

Éthique et TAL

C'est à votre tour de réfléchir (et de partager vos réflexions)
quel(s) problème(s) éthique(s) voyez-vous en tant que professionnels ? quelles solutions ?

Groupes d'étudiants générés automatiquement, 20 minutes de préparation, puis présentation très rapide à tous (5 min max.), pas de diapos, juste vous et le tableau blanc :

- ▶ Mise à disposition des décisions de justice
- ▶ Chatbot Tay
- ▶ Amazon Mechanical Turk
- ▶ Amazon Alexa et Google Home

Éthique et Traitement Automatique des Langues (TAL)

Rien de tel que la pratique : DIY

Une question de temps

Qui s'en occupe ?

Est-ce que la Science est nécessairement bonne ?
(certains) tournants du XXe siècle

Éthique et TAL

Éthique et Traitement Automatique des Langues (TAL)

Rien de tel que la pratique : DIY

Une question de temps

Qui s'en occupe ?

Est-ce que la Science est nécessairement bonne ?
(certains) tournants du XXe siècle

Éthique et TAL

Une très vieille question (la plus vieille?)



Lysippe — Eric Gaba (User :Sting),
Juillet 2005., CC BY-SA 2.5

L'éthique est au cœur de quelle discipline ?



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

Éthique et Traitement Automatique des Langues (TAL)

Rien de tel que la pratique : DIY

Une question de temps

Qui s'en occupe ?

Est-ce que la Science est nécessairement bonne ?

(certains) tournants du XXe siècle

Éthique et TAL

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 ?

Éthique et Traitement Automatique des Langues (TAL)

Rien de tel que la pratique : DIY

Une question de temps

Qui s'en occupe ?

Est-ce que la Science est nécessairement bonne ?

(certains) tournants du XXe siècle

Éthique et TAL

Le Code de Nuremberg

date d'avant, mais finalisé pendant "le procès des Médecins" (1947)



Le Code de Nuremberg

date d'avant, mais finalisé pendant "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

Le rapport Belmont [Bel, 1979]

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



National Archives Atlanta, GA

Le rapport Belmont [Bel, 1979]

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 ²⁷)	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"> • Expériences des médecins nazis sur des prisonniers en camp de concentration • Trois officiels nazis et vingt médecins impliqués, dont : <ul style="list-style-type: none"> – Heinrich Himmler : mécène et directeur de la médecine nazie, parfois surnommé « le meurtrier du Siècle ». Met en œuvre la solution finale – Josef Mengele : tortures, sévices corporels lors d'expériences médicales absurdes et inhumaines entraînant souvent la mort, à Auschwitz – Aribert Heim : vivisection humaine sur les détenus de Mauthausen 	<ul style="list-style-type: none"> • Fait suite aux atrocités de la Seconde Guerre mondiale • 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 	<ul style="list-style-type: none"> • Création le 12 juillet 1974 du National research act (Loi publique 93348 sur la recherche) • 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 Star²⁸ révèle l'affaire. N'a jamais donné de résultats concluants • 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.
Principes généraux	<ul style="list-style-type: none"> • Bienfaisance : rapport risque/ bénéfice favorable • Respect du sujet de recherche : consentement volontaire, capacité de consentir, possibilité de quitter l'essai • Justice : utilité pour la société 	<ul style="list-style-type: none"> • Bienfaisance : rapport risque/ bénéfice favorable • Respect du sujet de recherche : consentement informé • Justice : distinction entre recherche thérapeutique ou non • Responsabilité : pas uniquement au chercheur, avis extérieur nécessaire (comité d'éthique) 	<ul style="list-style-type: none"> • Bienfaisance : rapport risque/bénéfice favorable (distinction entre risques « minimes » et risques « plus que minimes ») • Respect du sujet de recherche : consentement éclairé • Justice : sélection équitable des sujets
Forme	Dix principes fondamentaux	Loi consensuelle, nombreux paragraphes	Paragraphes agencés en trois catégories éthiques et applications

Éthique et Traitement Automatique des Langues (TAL)

Rien de tel que la pratique : DIY

Une question de temps

Éthique et TAL

Les actions du groupe Éthique et TAL

Aller plus loin ?

Éthique et Traitement Automatique des Langues (TAL)

Rien de tel que la pratique : DIY

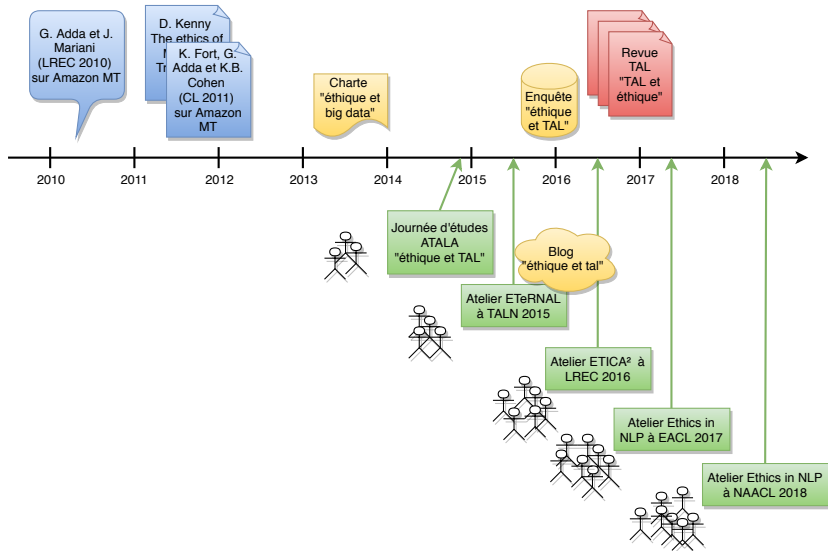
Une question de temps

Éthique et TAL

Les actions du groupe Éthique et TAL

Aller plus loin ?

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



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 (LIFAT / Université Rabelais, Tours)

Indus. A. Couillault (Apoliade, Paris)

Univ. K. Fort (Sorbonne Université, Paris)

Indus. H. de Mazancourt (YSEOP)

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

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)

Rien de tel que la pratique : DIY

Une question de temps

Éthique et TAL

Les actions du groupe Éthique et TAL

Aller plus loin ?

Autres actions pour le TAL

À l'international :

- ▶ ETHI-CA 2016 et ETHICAL 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)
- ▶ 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

pour votre temps

et votre attention

Annexes

Prendre le pouls de la communauté

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

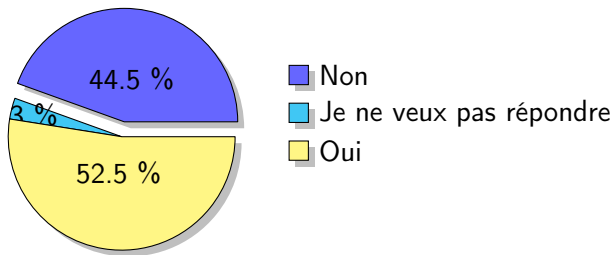
- ▶ publicité sur les listes de diffusion du domaine
- ▶ 100 personnes pour le questionnaire FR (180 à 200 participants à TALN)
- ▶ 200 personnes pour le questionnaire INT (1 000 participants à ACL)

En pratique :

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

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. »

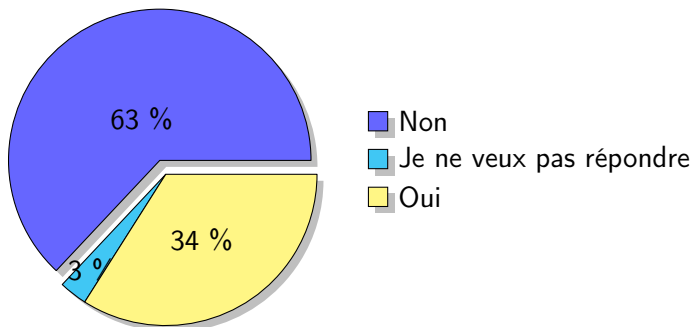
« je crois que cela concerne beaucoup de domaines, mais qu'il n'est pas nécessairement de la responsabilité du chercheur de se poser ces questions »

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

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

Lancer une alerte ?

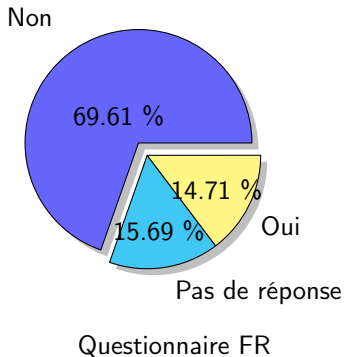
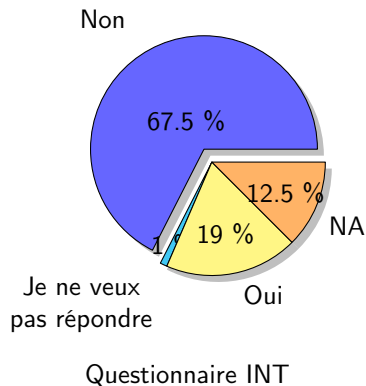
Do you know of any way to blow the whistle (issue an alert) in case of a major ethical infringement in your domain?



Questionnaire INT uniquement

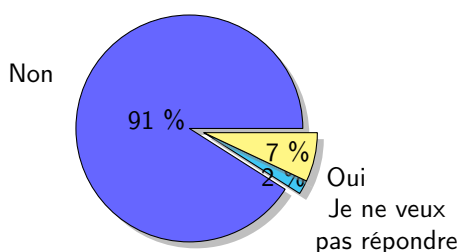
Formation à l'éthique

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

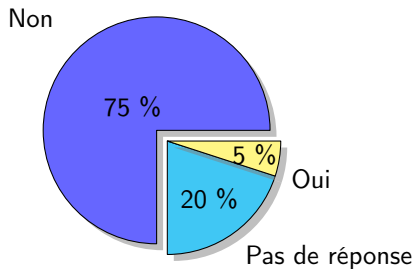


Rapports avec le citoyen

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



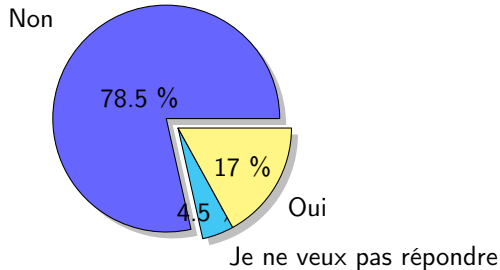
Questionnaire INT



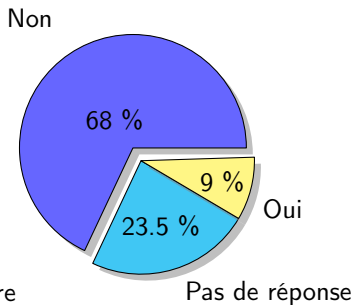
Questionnaire FR

Rapports avec les pouvoirs publics

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



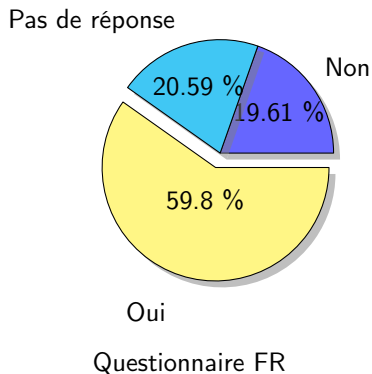
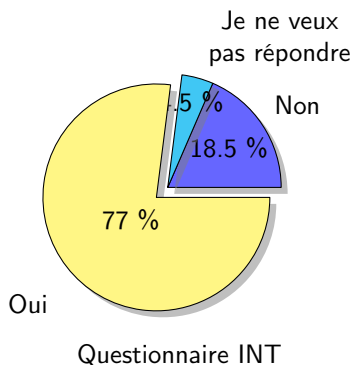
Questionnaire INT



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 ?





(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.

In Actes de la journée ATALA Éthique et TAL.



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.



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

Amazon Mechanical Turk : Gold mine or coal mine ?



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.



Spiekermann, S. (2015).

Ethical IT Innovation : A Value-Based System Design
Approach.
CRC Press.