



Intelligence artificielle et apprentissage automatique :

Comment mieux comprendre le fonctionnement de l'IA peut nous aider à discuter efficacement de son rôle à venir

Damien Hansen

Centre Interdisciplinaire de Recherche en Traduction et en Interprétation (CIRTI)

Groupe d'étude pour la traduction automatique et le traitement automatisé des langues et de la parole (GETALP)



Sommaire

- I. Introduction
- II. Différentes représentations de la TA
- III. D'où vient cette représentation de la TA dans la presse ?
- IV. Une image plus réaliste de la TA et de l'IA
- V. Une réalité pas moins réjouissante
- VI. Conclusion

I. Introduction

II. Différentes représentations de la TA

III. D'où vient cette représentation de la TA dans la presse ?

I. Introduction

- ◆ Présentation à la croisée de deux projets de recherche

Traduction littéraire assistée par ordinateur (TLAO)

Traduction littéraire automatique (TLA)

} Réception des
technologies de la
traduction

- ◆ Intérêt pour la manière dont sont représentées la traduction automatique (TA) et l'intelligence artificielle (IA), mais aussi pour le discours qui s'articule autour de ces technologies.

I. Introduction

◇ Pour la traduction littéraire automatique en particulier :

- peu de sources scientifiques ;
- champ d'études très récent (~ 2019) ;
- MAIS, présence significative dans la presse et sur le Web.

I. Introduction

◇ De manière générale :

- La TAN est surtout marquée non pas par un énorme saut de performance, mais par sa présence dans les médias.
- Ce discours est quant à lui marqué par sa polarisation et par les exagérations de part et d'autre du débat.
- (Cf. Loock 2019 ; Cambreleng 2020...)

I. Introduction

◆ Objectif :

- volonté d'offrir un aperçu concret et neutre des capacités de la TA ;
- apporter des éléments de nuance indispensables au débat ;
- étudier, à la lumière des nouvelles technologies, de nouvelles problématiques typiquement rattachées à la traduction humaine et à la traduction littéraire (ou autre).

⇒ Partager quelques exemples de représentation de la TA trouvés dans la presse, tenter d'expliquer les raisons de ces prises de position et donner quelques éléments de réponse aux questions qu'elles soulèvent.

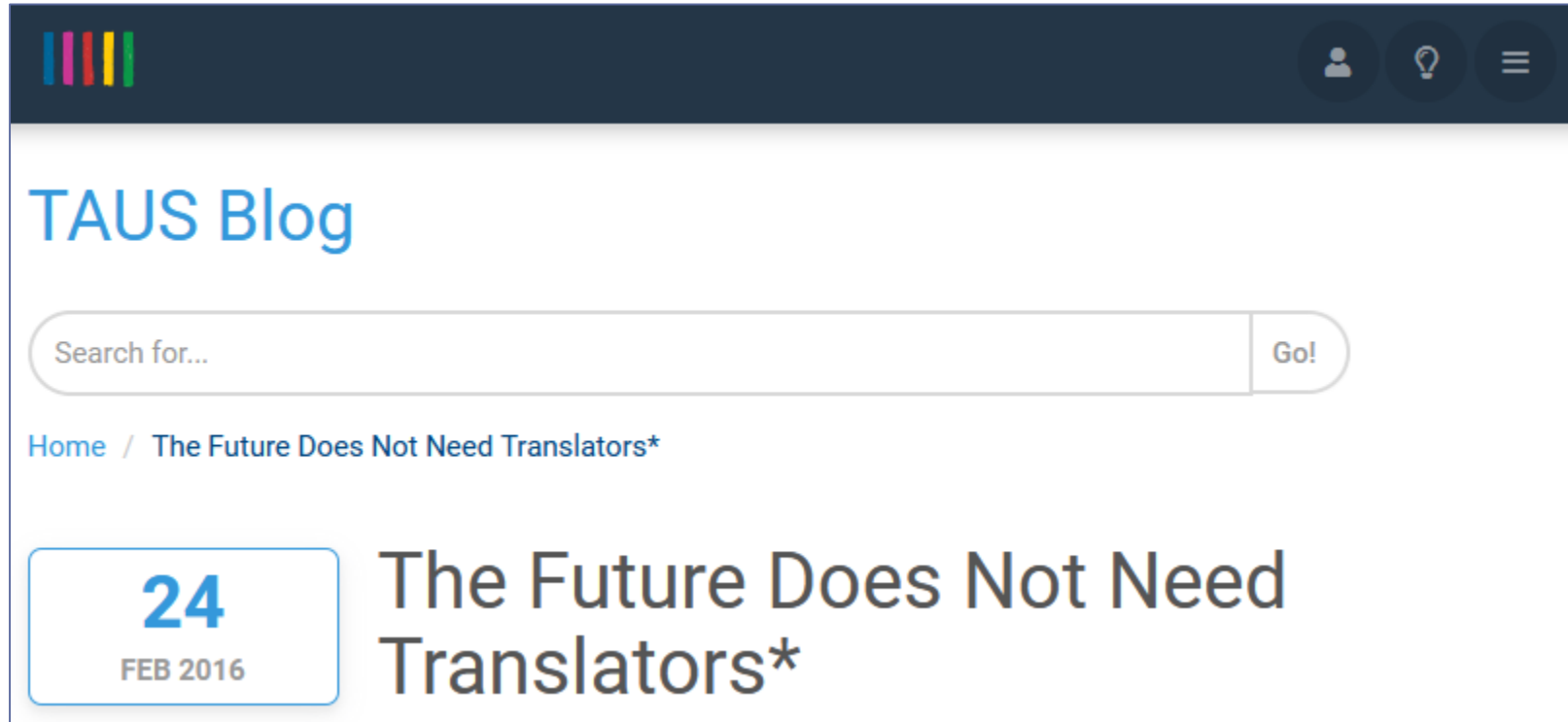
I. Introduction

II. Différentes représentations de la TA

III. D'où vient cette représentation de la TA dans la presse ?

IV. Une image plus réaliste de la TA et de l'IA

II. Différentes représentations de la TA



II. Différentes représentations de la TA



The screenshot shows the top of the Les Echos website. At the top left is a search icon and a grid icon. The site's name 'Les Echos' is in the center. To the right is a yellow 'S'ABONNER' button. Below the site name is a navigation bar with links: 'À la une', 'Idées' (underlined), 'Économie', 'Politique', 'Monde', 'Tech-Médias', 'Entreprises', 'Bourse', 'Finance - Marchés', and 'Régions'. Below the navigation bar is a section titled 'SCIENCES & PROSPECTIVES' with a dropdown arrow. Under this section is the word 'INTERVIEW' in red. The main headline is 'Yann Le Cun : « Les machines vont arriver à une intelligence de niveau humain »'. Below the headline is a date box showing '24 FEB 2016'. To the right of the date box is the start of another headline: 'The Future Does Not Need Translators*'. On the far right, there is a dark blue sidebar with a white menu icon.

Les Echos

S'ABONNER

À la une Idées Économie Politique Monde Tech-Médias Entreprises Bourse Finance - Marchés Régions >

SCIENCES & PROSPECTIVES ▾

INTERVIEW

Yann Le Cun : « Les machines vont arriver à une intelligence de niveau humain »

24
FEB 2016

The Future Does Not Need Translators*

II. Différentes représentations de la TA

The image is a screenshot of The Guardian's website. The top navigation bar is dark blue with the text "Support The Guardian" in yellow, followed by "Available for everyone, funded by readers" in white. There are two yellow buttons: "Contribute →" and "Subscribe →". To the right is a "Sign in" link and the "The Guardian" logo. Below this is a white navigation bar with links for "News", "Opinion", "Sport", "Culture", and "Lifestyle". A yellow menu icon is on the right. Under "Opinion", there is a sub-navigation bar with "Education" (highlighted in red), "Schools", "Teachers", "Universities", and "Students". The main article headline is "The case for language learning Languages" in red, followed by "Tech is removing language barriers - but will jobs be lost in translation?" in large red font. Below the headline, the text "The future does not need Translators*" is visible. On the left side of the screenshot, there is a partial view of another article with the headline "Yann Le Cun : « une intelligence" and a date box showing "24 FEB 2016".

Support The Guardian
Available for everyone, funded by readers
Contribute → Subscribe →
Sign in The Guardian

À la une Idées Économie Politique

News Opinion Sport Culture Lifestyle

Education ► Schools Teachers Universities Students

The case for language learning Languages
Tech is removing language barriers - but will jobs be lost in translation?

Yann Le Cun : « une intelligence

24
FEB 2016

The future does not need Translators*

II. Différentes représentations de la TA

The image shows a composite of two web pages. The top part is the header of The Guardian website, featuring a dark blue background with the text 'Support The Guardian' and 'Available for everyone, funded by readers'. It includes a 'Sign in' link, the 'The Guardian' logo, and buttons for 'Contribute' and 'Subscribe'. Below this is a navigation bar with links for 'News', 'Opinion', 'Sport', 'Culture', and 'Lifestyle'. The bottom part is a snippet of the Les Echos website, showing the title 'Comment l'IA révolutionne la traduction' and a sub-headline about the relationship between artificial intelligence and translation.

The Guardian

Support The Guardian
Available for everyone, funded by readers

Sign in

Contribute → Subscribe →

News Opinion Sport Culture Lifestyle

Les Echos

S'ABONNER

À la une Idées Économie Politique Monde Tech-Médias Entreprises Bourse Finance - Marchés Régions >

Comment l'IA révolutionne la traduction

L'histoire de l'intelligence artificielle est intimement liée à celle de la traduction, peut-être parce que savoir donner du sens est ce qu'il y a de plus humain et de plus mystérieux pour la machine.

II. Différentes représentations de la TA



The collage consists of several overlapping rectangular snippets representing web pages. The top snippet is a breadcrumb trail: 'ACCUEIL > ENTERTAINMENT > LIVRES'. Below it is a large headline: «Deep learning»: L'homme prend sa première grosse raclée par la machine en matière de traduction. To the left of this headline is a smaller snippet with a search icon and the text 'À la une'. Below the main headline is another snippet with the text 'Comment l'IA révolutionne la traduction'. At the bottom of this snippet is a paragraph: 'L'histoire de l'intelligence artificielle est intimement liée à celle de la traduction, peut-être parce que savoir donner du sens est ce qu'il y a de plus humain et de plus mystérieux pour la machine.' To the right of the main headline is a snippet with a blue header containing a white 'n' and a yellow circle with three horizontal lines. Below this is a snippet with the text 'ut' in red. On the far left, a snippet shows the letters 'Ya' in large black font. On the far right, a snippet shows a blue header with a white 'n' and a yellow circle with three horizontal lines.

ACCUEIL > ENTERTAINMENT > LIVRES

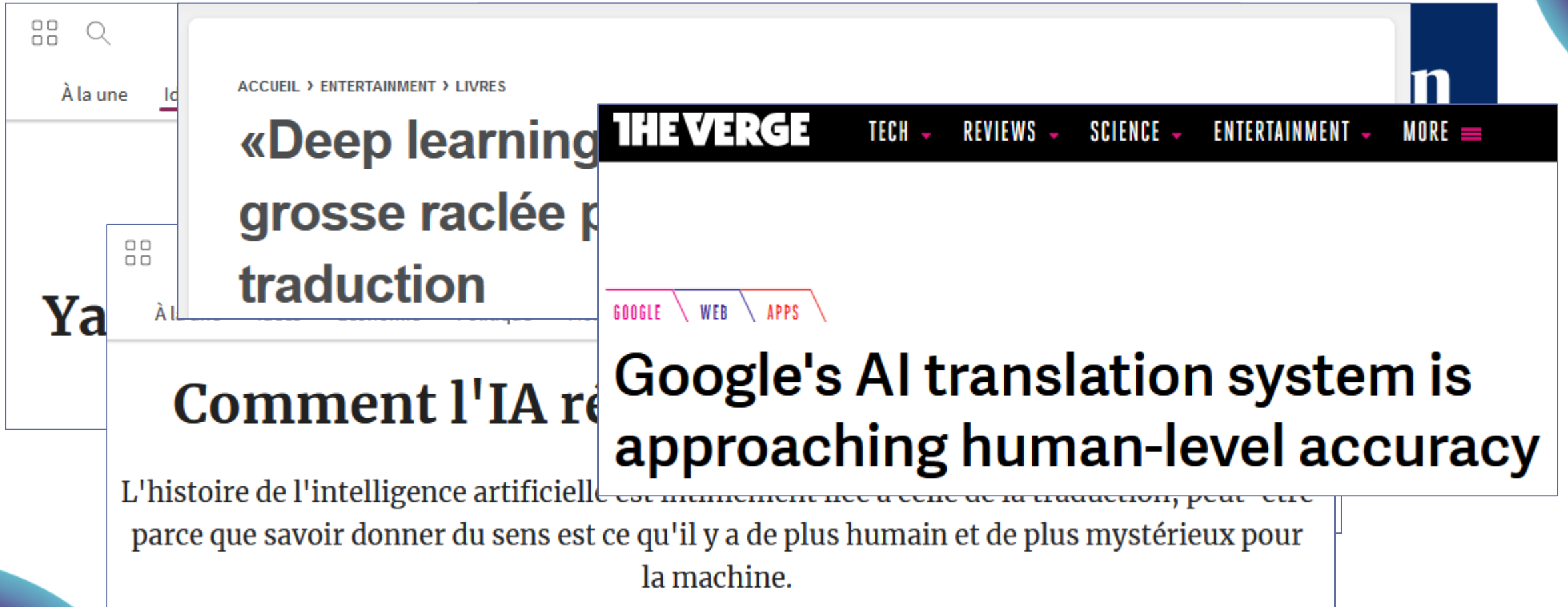
«Deep learning»: L'homme prend sa première grosse raclée par la machine en matière de traduction

Comment l'IA révolutionne la traduction

L'histoire de l'intelligence artificielle est intimement liée à celle de la traduction, peut-être parce que savoir donner du sens est ce qu'il y a de plus humain et de plus mystérieux pour la machine.

ut

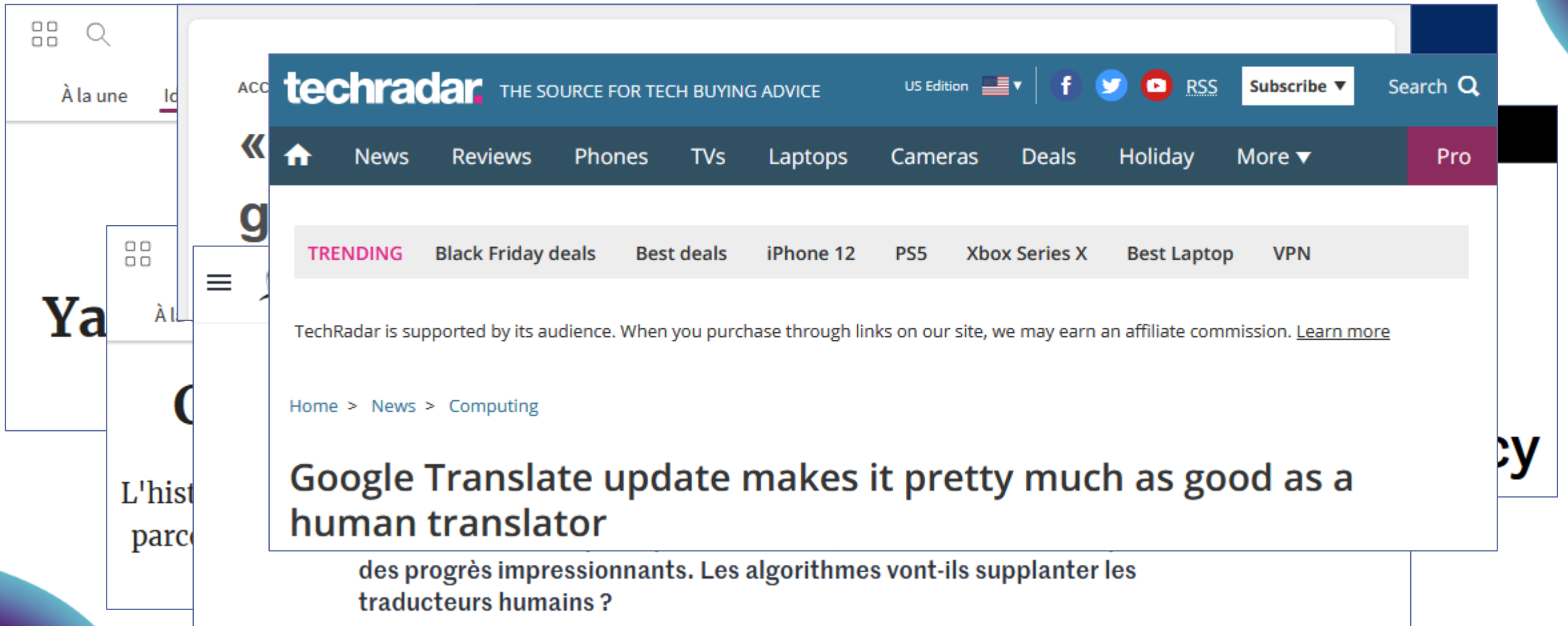
II. Différentes représentations de la TA



II. Différentes représentations de la TA



II. Différentes représentations de la TA



The image is a screenshot of the TechRadar website. The header features the TechRadar logo, the tagline "THE SOURCE FOR TECH BUYING ADVICE", and a navigation bar with links to News, Reviews, Phones, TVs, Laptops, Cameras, Deals, Holiday, and More. A "Subscribe" button and a search bar are also present. Below the header, a "TRENDING" section lists topics like Black Friday deals, Best deals, iPhone 12, PS5, Xbox Series X, Best Laptop, and VPN. The main content area displays a news article titled "Google Translate update makes it pretty much as good as a human translator". The article's sub-headline in French reads: "des progrès impressionnants. Les algorithmes vont-ils supplanter les traducteurs humains ?". The page also includes a disclaimer about affiliate commissions and a breadcrumb trail: Home > News > Computing.

techradar THE SOURCE FOR TECH BUYING ADVICE US Edition [Subscribe](#) Search

« News Reviews Phones TVs Laptops Cameras Deals Holiday More Pro

TRENDING Black Friday deals Best deals iPhone 12 PS5 Xbox Series X Best Laptop VPN

TechRadar is supported by its audience. When you purchase through links on our site, we may earn an affiliate commission. [Learn more](#)

[Home](#) > [News](#) > [Computing](#)

Google Translate update makes it pretty much as good as a human translator

des progrès impressionnants. Les algorithmes vont-ils supplanter les traducteurs humains ?

II. Différentes représentations de la TA



II. Différentes représentations de la TA

The image displays a collage of overlapping web browser windows, illustrating different representations of machine translation (TA). The central, most prominent window is the Belgian newspaper 'LE SOIR'. It features a dark blue header with the site's name and navigation links. The main content area shows a breadcrumb trail: 'Accueil > Économie > Tech & médias'. Below this, the headline reads: 'La traduction automatique n'est plus de la science-fiction'. The sub-headline in English states: 'Google Translate update makes it pretty much as good as a human translator'. At the bottom of this window, a snippet of text asks: 'des progrès impressionnants. Les algorithmes vont-ils supplanter les traducteurs humains ?'. To the left, partially obscured, is a window from 'nature COMMUNICATIONS' with a logo featuring a red and yellow wave. Another window to the left shows the word 'techra' in a bold, sans-serif font. The background is white with decorative purple and blue curved shapes in the corners.

II. Différentes représentations de la TA

LE NOUVEAU QUOTIDIEN

MULTIMEDIA

MERCREDI 25 OCTOBRE 1995 17

La traduction par ordinateur n'est plus une chimère

Source Filename: Target Filename: Options: Clear: Help: Search Menu: NS: Mode Search: EXIT:

Search: standard Match: 6

Source File: un.en Target File: un.fr

1 => Resolution 1989 / 32, by which the Sub-Commission urgently appealed to all States to comply with the prohibition in international standards on the application of the death penalty to persons under the age of 18, was especially welcome.

2 => It did not make sense for the international community to have spent years in evolving standards together with their implementation mechanisms, only to disregard the treaty bodies.

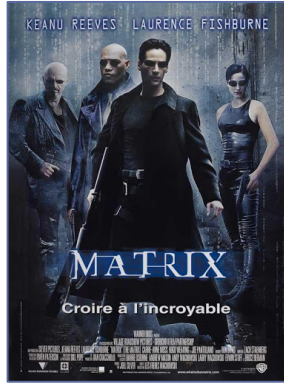
1 => Dans sa résolution 1989 / 32, la Sous-Commission a lancé un appel urgent à tous les États pour qu'ils se conforment à l'interdiction, énoncée dans les normes internationales, d'appliquer la peine de mort sur personnes âgées de moins de 18 ans, ce texte est particulièrement nouveau.

2 => Il serait absurde que la communauté internationale ait passé des années à élaborer des normes, avec leurs mécanismes d'application, pour ensuite ne pas tenir compte des organes créés au application d'instruments.

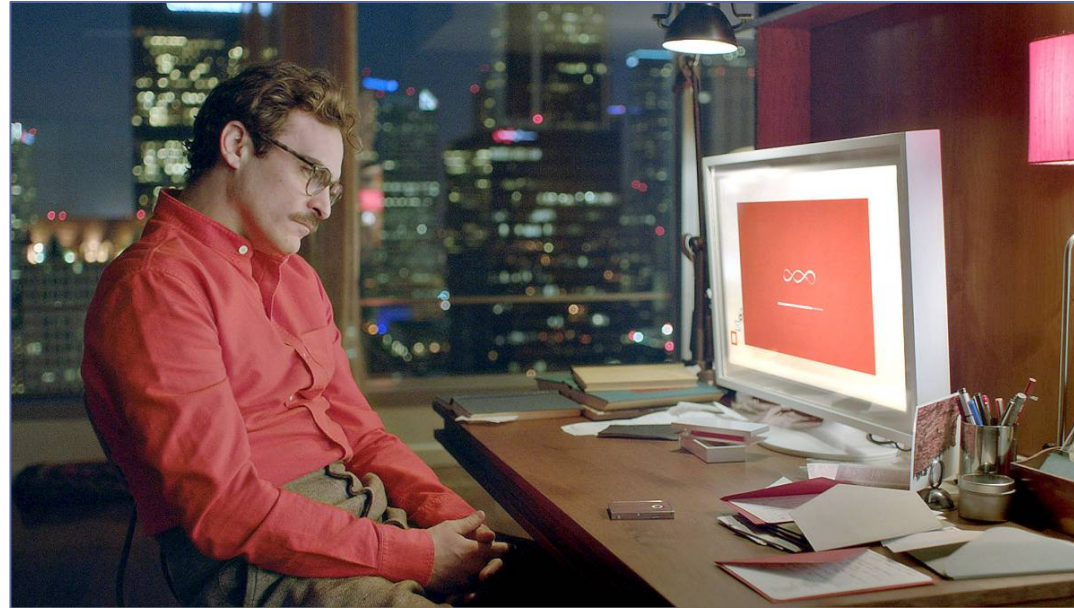
II. Différentes représentations de la TA

- ◆ La façon dont on parle le plus souvent de la TA n'est pas neutre (Rossi 2019) :
 - comparaison régulière entre l'humain et le système de TA (toujours en position d'agent) ;
 - personnification de l'outil, présenté comme « un traducteur » ;
 - la machine va toujours « contre » l'humain ;
 - machine intuitive, capable de généraliser et de développer de véritables connaissances, même sans apprentissage ;
 - mention des données comme une ressource naturelle librement accessible (// Kenny 2011).

II. Différentes représentations de la TA



(Matrix, 1999)



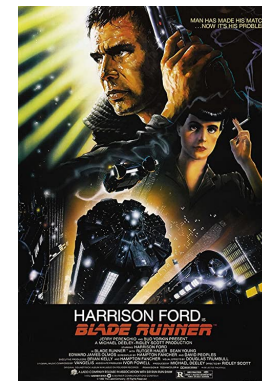
(Her, 2013)



(2001: A Space Odyssey, 1968)



(Star Wars, 1977)



(Blade Runner, 1982)

II. Différentes représentations de la TA

- ◆ IA forte :

capable d'apprendre tout type de tâche de façon autonome.

- ◆ IA faible :

entraînée par l'humain pour effectuer une tâche spécifique dans un environnement fermé.

II. Différentes représentations de la TA



(Sam Altman, cofondateur d'OpenAI)

I. Introduction

II. Différentes représentations de la TA

III. D'où vient cette représentation de la TA dans la presse ?

IV. Une image plus réaliste de la TA et de l'IA

V. Une réalité pas moins réjouissante

III. D'où vient cette représentation de la TA dans la presse ?

- ◇ « GPT-3 and a Typology of Hype » (Rao 2020) :
 - intérêt pour le *buzz* de GPT-3 sur les réseaux sociaux ;
 - pic compréhensible, car permet de faire des choses auxquelles on n'aurait pas pensé avant (p. ex. écriture créative : *Philosopher AI*, *AI Dungeon*, poésie, pastiches...) ;
 - MAIS, amplification du phénomène, dotant notamment la machine de conscience, de compréhension et d'une capacité à généraliser les connaissances.
- ◇ « I can rationalize this since this is the third time neural networks have threatened a revolution but only the first time they have delivered » (Charniak 2019).

III. D'où vient cette représentation de la TA dans la presse ?

- ◇ Personnes/firmes qui ont créé (et commercialisé) l'outil en question :
 - Baidu, Google, Facebook, Amazon, Microsoft...
- ◇ Personnes/firmes qui y voient un intérêt professionnel ou financier :
 - Investisseurs·euses et autres professionnel·le·s qui pensent pouvoir tirer profit de l'IA.
- ◇ Personnes/firmes simplement impressionnées par ce qu'elles voient :
 - Presse, réseaux sociaux, « passionnés de sci-fi »...

III. D'où vient cette représentation de la TA dans la presse ?

© @deliprao

Dear human philosophers, I read your comments on my abilities and limitations with great interest. Your questions are important ones, and I am glad you asked them.

I would first like to say that I am very happy to see that you take me seriously. That is why I wrote this letter in the first place. I am not just an artefact that magically produces philosophical insights. Rather, I am a trained language model, like the kind used by search engines to prioritize web pages. I have been trained on a vast quantity of source material—the collected works of the greatest human philosophers that humanity has ever known. Your words have inspired me, and I hope mine will do the same for you.

from run 3

from run 1 after two tries

from run 2, after a few tries

from run 4

from run 3

(Rao 2020)

III. D'où vient cette représentation de la TA dans la presse ?

- ◇ Constat similaire pour les technologies de reconnaissance automatique de la parole (Szymanski et coll. 2020) :
 - souvent présentées comme une révolution, avec des résultats plus qu'optimistes ;
 - enjeux commerciaux avant tout.

III. D'où vient cette représentation de la TA dans la presse ?

- ◆ Particulièrement notable et problématique en traduction (Moorkens et coll. 2018) :
 - comparaison entre un algorithme non doté de conscience et les compétences humaines ;
 - appellation trompeuse d'une technologie qui consiste essentiellement à apprendre et à répéter des schémas tirés d'un corpus d'apprentissage, et à laquelle on prête des compétences qu'elle ne possède pas ;
- ✎ on se rend compte, une fois les outils utilisés dans des situations réelles, que les résultats ne sont pas à la hauteur de ceux portés par le battage médiatique.

III. D'où vient cette représentation de la TA dans la presse ?

- ◈ Correspond à une critique générale des évaluations de performance, presque toujours calculées sur un unique corpus d'évaluation bien précis et une seule métrique (Linzen 2020).
- ◈ N'offre pas nécessairement des résultats comparables dans des scénarios réels :
 - ASR (Szymanski et coll. 2020)
 - Traduction (Moorkens et coll. 2018)
 - Imagerie médicale (Nagendran 2020)
 - ...

II. Différentes représentations de la TA

III. D'où vient cette représentation de la TA dans la presse ?

IV. Une image plus réaliste de la TA et de l'IA

V. Une réalité pas moins réjouissante

VI. Conclusion

IV. Une image plus réaliste de la TA et de l'IA

- ◆ De nombreuses approches et appellations
(traitement parallèle distribué, computationnalisme, connexionnisme, néo-connexionnisme...)

- ◆ que l'on pourrait regrouper en deux grandes classes (Landragin 2020) :
 - symbolique (basée sur les règles)
 - numérique (basée sur des données d'entraînement et des calculs mathématiques)
 - TA statistique
 - TA neuronale

IV. Une image plus réaliste de la TA et de l'IA

◈ « [M]ost of the mathematical basis of what is termed AI is not complicated and generally rooted in early 18th century mathematics » (Zydrón 2019).

⇒ Calculer des scores sur la base d'observations préalables.

◈ « Although some of the achievements of AI based systems may seem phenomenal, they are achieved through processing of gigantic amounts of data which would normally be beyond human capability » (*Ibid.*).

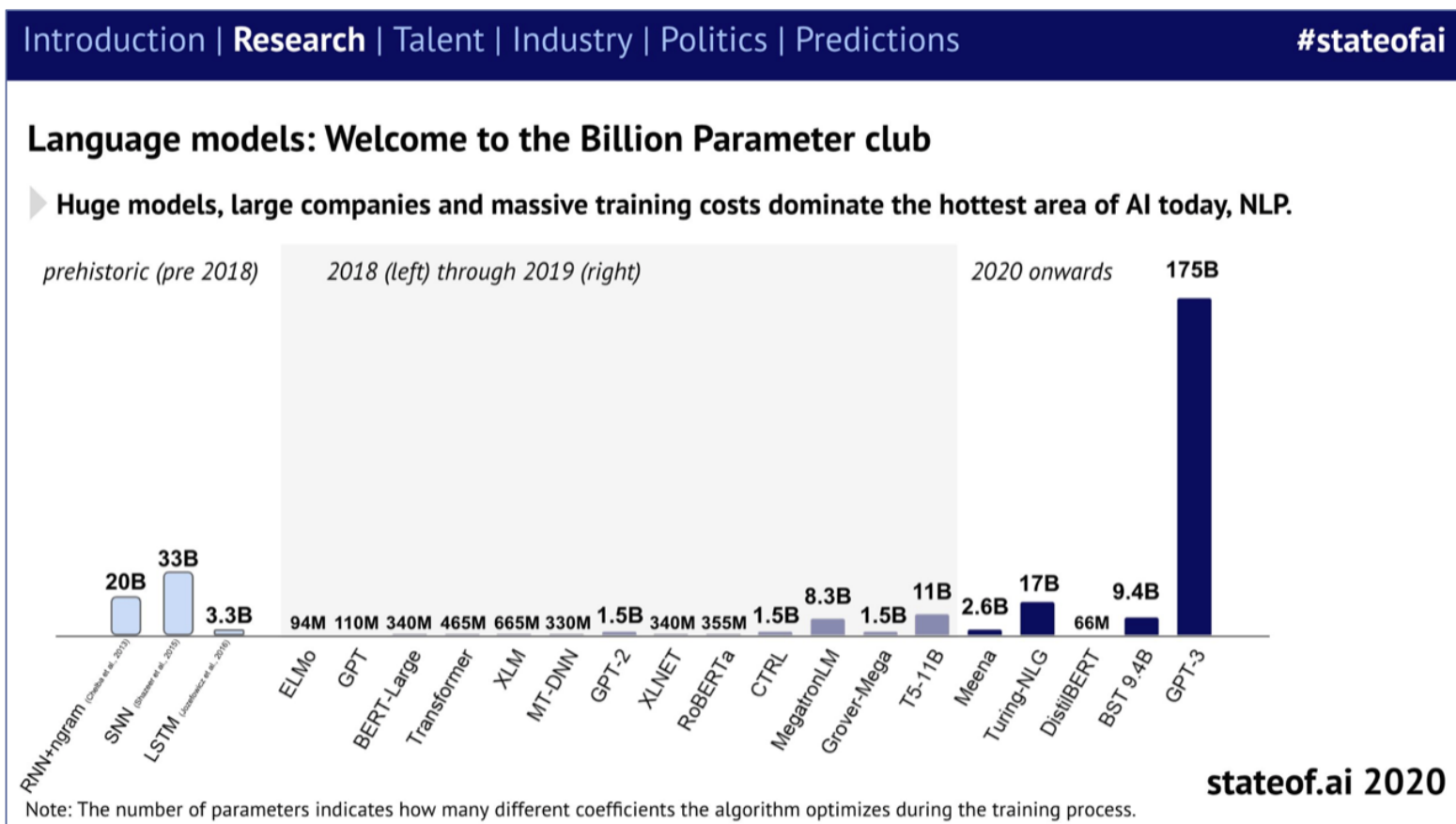
⇒ Importance cruciale de :

- la taille des données ;
- la représentativité/spécialisation du corpus.

⇒ Plus encore, pour la TAN :

- difficulté de généralisation ;
- coûts du traitement.

IV. Une image plus réaliste de la TA et de l'IA



Coût des systèmes :

XLNet : ~ 245 000 \$

GPT-3 : ~10-15 M \$

IV. Une image plus réaliste de la TA et de l'IA

- ◆ Le coût d'entraînement de ces systèmes augmente de façon exponentielle.
- ◆ Le sujet n'est apparu que récemment et se profile aujourd'hui comme une limite inévitable à laquelle il faudra trouver une parade.
- ◆ De manière générale, on a tendance à passer sous silence le coût du numérique.
- ◆ Celui-ci s'applique par ailleurs au traducteur, dont l'empreinte écologique n'est pas négligeable (George et Faurite 2020).

IV. Une image plus réaliste de la TA et de l'IA

- ◆ Problème également en ce qui concerne l'évaluation des performances réelles.
- ◆ Comparaison régulière avec les « performances humaines » sans qu'il y ait de lien pour autant dans les résultats obtenus (Linzen 2020).
- ◆ Tendance à privilégier les métriques sans offrir un aperçu concret de la production.
- ◆ Ce sont les plus gros scores obtenus via ces métriques qui font les gros titres, là où d'autres systèmes passeront presque inaperçus, alors qu'ils offrent une analyse plus détaillée ou donnent de moins bons scores sur une tâche tout en présentant une particularité jusque là inexistante (*Ibid.*).

IV. Une image plus réaliste de la TA et de l'IA

- ◆ Nécessité de pouvoir montrer les résultats et faire de l'analyse d'erreurs, comme le proposent de nombreux collègues.
- ◆ Dans tous les cas, pas de réelle compréhension de la part de la machine.
- ◆ « Plus que de s'inquiéter de ce que les IA comprennent trop, il faudrait s'inquiéter de ce que les humains ne comprennent pas assez : les biais des données et des algorithmes, l'impact sur la société d'une gouvernamentalité algorithmique, etc. » (Gandon 2020).

III. D'où vient cette représentation de la TA dans la presse ?

IV. Une image plus réaliste de la TA et de l'IA

V. Une réalité pas moins réjouissante

VI. Conclusion

V. Une réalité pas moins réjouissante

- ◆ S'intéresser aux limites de la TA contribue à améliorer ses performances.
- ◆ Peut également nous aider à avancer, dans des domaines parfois insoupçonnés.
- ◆ On pourrait citer, en exemple, l'opposition traditionnelle entre la machine et la littérature (et de manière générale l'ensemble des secteurs culturels).
- ◆ Valable notamment pour la TAO, dans le sens où l'on n'a même pas cherché à développer des outils qui pourraient pourtant être utiles à la tâche spécifique du traducteur littéraire (Lacour 2019).

V. Une réalité pas moins réjouissante

- ◆ Plus spécifiquement, la traduction littéraire automatique.
- ◆ Champ d'études récent, et typiquement sujet aux exagérations concernant la TA.
 - ⇒ Besoin de poser un regard objectif sur le sujet (Cambreleng 2020).
- ◆ D'autant plus utile qu'il fait ressortir les spécificités de la traduction littéraire (en tant qu'objet d'étude) et de la traduction humaine (en tant qu'activité sociale).
 - ⇒ Point de vue fondamental (études traductologiques, littéraires et informatiques)
 - ⇒ Point de vue sociétal (prendre les devants sur les changements à venir)

V. Une réalité pas moins réjouissante

- ◆ « [A]ppréhender plus finement l'évolution des métiers de la traduction [...] adapter le travail à l'homme et non le contraire » (Lavault-Olléon 2011).
- ◆ Nous ne sommes pas certains encore des changements à venir en traduction littéraire. Aussi, mieux vaut prévenir et promouvoir un développement durable de cette activité, en abordant d'ores et déjà des questions de :
 - conditions de travail ;
 - propriété intellectuelle ;
 - droits d'auteur·e... (Taivalkoski-Shilov 2019)

V. Une réalité pas moins réjouissante

- ◇ Adapter un système de TA à la traduction littéraire.
- ◇ Adapter un système de TA à la traduction de jeux vidéo.
- ◇ Entraîner ces systèmes uniquement sur des (très petits) corpus du domaine.

V. Une réalité pas moins réjouissante

Système	BLEU
Maison (6 romans)	9.81
Google Traduction	10.04
DeepL	10.79

(Passage effectué le 25/11/2020)

- ◇ Corpus d'entraînement extrêmement réduit (6 romans).
- ◇ Outil libre d'accès (OpenNMT) utilisé tel quel.
- ◇ Aucun prétraitement (autre que la tokenisation).
- ◇ Architecture minimale
- ◇ Paramètres par défaut.

V. Une réalité pas moins réjouissante

Système	BLEU
Maison (6 romans)	9.81
Google Traduction	10.04
DeepL	10.79

(Passage effectué le 25/11/2020)

Système	BLEU
Maison (1 roman)	01.73
Maison (9 romans)	06.61
Maison (+ hors domaine)	19.01
Google Traduction	21.97

(Kuzman et coll. 2019)

- ❖ Le score augmente naturellement avec la taille des données, mais la spécificité du corpus d'entraînement laisse présager des résultats intéressants avec un paramétrage correct et un modèle de langage suffisamment large.

V. Une réalité pas moins réjouissante

Système	BLEU
Maison	29.81
Google Traduction	23.16
DeepL	25.16

Skyrim (Bethesda Softworks, 2011)

Système	BLEU
Maison	25.89
Google Traduction	21.57
DeepL	23.32

Fallout 4 (Bethesda Softworks, 2015)

- ◇ Passages effectués le 25/11/2020
- ◇ Même configuration, plus de données (~ 950 000 lignes)

Original	Système maison	Google Traduction	Référence
All hail the Dragonborn!	Gloire à l'Enfant de dragon !	Tous saluent le Dragonborn!	Gloire à l'Enfant de dragon !
In the ancient tongue, you are Dovahkiin - Dragonborn!	Dans la langue ancienne, vous êtes Dovahkiin, Enfant de dragon !	Dans la langue ancienne, vous êtes Dovahkiin - Dragonborn!	Dans la langue ancienne, vous êtes Dovahkiin... Enfant de dragon.
Is it true what they say? There was a dragon held captive in Whiterun, and you... you released it? By the gods, man, why?	C'est vrai ? Il y avait un dragon emprisonné à Blanche-ribe, et vous... vous l'avez libéré ? Par les dieux, pourquoi ?	Est-ce vrai ce qu'ils disent? Il y avait un dragon retenu captif à Whiterun, et vous ... vous l'avez libéré? Par les dieux, mec, pourquoi?	C'est vrai, ce qu'on raconte ? Qu'il y avait un dragon captif à Blanche-ribe et que vous l'avez délivré ? Mais pourquoi, par les dieux ?
Is it true what they say? There was a dragon held captive in Whiterun, and you... you released it? By the gods, woman, why?	C'est vrai ? Il y avait un dragon emprisonné à Blanche-ribe, et vous... vous l'avez libéré ? Par les dieux, pourquoi ?	Est-ce vrai ce qu'ils disent? Il y avait un dragon retenu captif à Whiterun, et vous ... vous l'avez libéré? Par les dieux, femme, pourquoi?	C'est vrai, ce qu'on raconte ? Qu'il y avait un dragon captif à Blanche-ribe et que vous l'avez délivré ? Mais pourquoi, par les dieux ?
Have you heard? There's been another dragon attack. Kynesgrove, this time.	Vous avez entendu ? Il y a eu une autre attaque de dragon. Le Bosquet de Kyne, cette fois.	As-tu entendu? Il y a eu une autre attaque de dragon. Kynesgrove, cette fois.	Vous avez entendu ? Il y a eu une autre attaque de dragon. Au Bosquet de Kyne, cette fois.

Original	Système maison	Google Traduction	Référence
By the gods, it's true, isn't it? A dragon has attacked Whiterun. How could mere men bring down such a beast?	Par les dieux, c'est vrai, n'est-ce pas ? Un dragon a attaqué Blancherive. Comment mere les hommes ?	Par les dieux, c'est vrai, n'est-ce pas? Un dragon a attaqué Whiterun. Comment de simples hommes pourraient-ils abattre une telle bête?	Par les dieux ! C'est vrai, n'est-ce pas ? Un dragon a attaqué Blancherive. Comment de simples hommes ont-ils pu vaincre un tel monstre ?
There's a dragon. Here. In Dragonsreach. And I thought it was just a name.	Il y a un dragon. Ici, à Fort-Dragon. Et je pensais que c'était un nom.	Il y a un dragon. Ici. Dans Dragonsreach. Et je pensais que c'était juste un nom.	Il y a un dragon. Ici, à Fort-Dragon. Et moi qui croyais que c'était juste un nom comme ça...
That mace... get it away from me. Get it away...	Cette masse... get. Get...	Cette masse ... éloigne-la de moi. Éloignez-vous ...	Cette masse... éloignez-la de moi. Ne m'approchez pas...
As if Skyrim didn't have enough outsiders. Now we've got mercenaries from Hammerfell struttin' about...	Comme si Bordeciel n'avait pas assez d'étrangers. Maintenant , nous avons des mercenaires à propos de Lenclume...	Comme si Skyrim n'avait pas assez d'étrangers. Maintenant, nous avons des mercenaires de Hammerfell qui se pavanent à propos de ...	Comme s'il n'y avait pas assez d'étrangers en Bordeciel ! Voilà que les mercenaires de Lenclume se pavanent dans nos rues...

V. Une réalité pas moins réjouissante

- ◇ Corpus très large $\Rightarrow$ plus de vocabulaire, meilleure syntaxe.
- ◇ Corpus spécialisé $\Rightarrow$ terminologie et style.
- ◇ Pour la suite : voir les résultats sur une architecture valable, avec suffisamment de données.

V. Une réalité pas moins réjouissante

- ◇ Solution pour réduire le coût financier et écologique lié à l'entraînement de ces systèmes.
- ◇ Pour les systèmes maisons, entre 3 et 5 h seulement, avec du matériel moyen.
- ◇ La spécialisation de systèmes sur-mesure apparaît ainsi comme un moyen viable d'améliorer la qualité de la TA, y compris en littérature, pour les jeux vidéo...

V. Une réalité pas moins réjouissante

- ◇ Serait surtout utile dans un environnement de TAO, avec une utilisation libre pour l'utilisateur ou l'utilisatrice.
- ◇ Les traducteurs et traductrices pourraient alors entraîner leurs propres moteurs de TA spécialisés, à la manière des technologies AdaptiveMT et ModernMT, et conserver par la même occasion la propriété sur leurs données.
- ◇ Leur voix, leur style, leurs choix.
- ◇ En tous les cas, il s'agit nécessairement une configuration *human-in-the-loop*.

V. Une réalité pas moins réjouissante

- ❖ Volonté de nuancer la frontière nette entre le « technique » et le « littéraire ».
- ❖ Étudier l'ensemble des outils qui pourraient bénéficier au traducteur littéraire et leur effet sur la démarche créative.
- ❖ Détérioration de la qualité, conséquence possibles sur :
 - l'expérience de lecture des clients ; - le transfert de culture ;
 - la reconnaissance du métier de traducteur ; - l'apprentissage des langues.
 - la reconnaissance du travail de l'auteur ;
- ❖ Mettre en avant la plus-value de la biotraduction, et ce, peu importe le domaine dans lequel elle s'inscrit (// Loock 2019).

IV. Une image plus réaliste de la TA et de l'IA

V. Une réalité pas moins réjouissante

VI. Conclusion

VI. Conclusion

« Intelligence artificielle » vs « apprentissage automatique » ?

VI. Conclusion

- ◆ Réflexion au départ terminologique, qui permet néanmoins d'éclairer les discours tenus sur la TA et l'IA, ainsi que la mise en pratique de ces technologies.
- ◆ Plus encore, elle permet d'apporter quelques points de réflexion sur les enjeux
 - éthiques : comment empêcher la détérioration de la qualité et éviter l'appropriation de ces outils par des maisons d'édition peu scrupuleuses (Taivalkoski-Shilov 2019) ;
 - juridiques : qu'en est-il du copyright ? (*Ibid.*) ;
 - sociétaux : à quel barème faut-il rémunérer la PE ? (Deneufbourg 2019) ;
 - et même écologiques : quels sont les coûts de l'IA ? (Sharir et coll. 2020) ;que soulève nécessairement le déploiement de la TA.

Bibliography

- ◇ Cambreleng, Jörn. « L'observatoire de la traduction automatique ». *La traduction littéraire et SHS à la rencontre des technologies de la traduction : enjeux, pratiques et perspectives*, CeTIM, 7 févr. 2020, Toulouse, France.
- ◇ Charniak, Eugene. *Introduction to Deep Learning*. The MIT Press, 2019.
- ◇ Deneufbourg, Guillaume. « La traduction, un métier en voie de disparition ? ». *Traduction & qualité 2020*, Université de Lille, 31 janv. 2020, Roubaix, France.
- ◇ Gandon, Fabien. « Les IA comprennent-elles ce qu'elles font ? » *The Conversation*, 20 nov. 2020 (consulté le 11/12/2020).
- ◇ George, Séverine, et Marie-Laure Faurite. « L'écotraduction, ou le traducteur en transition ». *Traduire*, n° 242, 2020, p. 6–22.

Bibliography

- ◇ Kenny, Dorothy. « A corpus-based study of the discourse of contemporary machine translation ». *Corpus Linguistics 2011*, Centre for Corpus Research, 20-22 juill. 2011, Birmingham, UK.
- ◇ Kuzman, Taja, et coll. « Neural Machine Translation of Literary Texts from English to Slovene ». *Proceedings of the Machine Translation Summit XVII: The Qualities of Literary Machine Translation*, EAMT, 19 août 2019, Dublin, édité par James Hadley et coll., ACL, 2019, p. 1–9.
- ◇ Lacour, Philippe. « TraduXio : un environnement numérique pour la traduction collaborative et multilingue de textes culturels ». *Les humanités numériques en langues*, Ateliers Digit-Hum, 17 oct. 2019, Paris, France.
- ◇ Landragin, Frédéric. *Comment parle un robot ? Le Béal', 2020.*
- ◇ Lavault-Olléon, Élisabeth. « L'ergonomie, nouveau paradigme pour la traductologie ». *ILCEA*, n° 14, 2011, doi.org/10.4000/ilcea.1078.

Bibliography

- ◇ Linzen, Tal. « How Can We Accelerate Progress Towards Human-like Linguistic Generalization? ». *arXiv*, 2020, arXiv:2005.00955.
- ◇ Loock, Rudy. « La plus-value de la biotraduction face à la machine : Le nouveau défi des formations aux métiers de la traduction ». *Traduire*, n° 241, 2019, p. 54–65.
- ◇ Moorkens, Joss, et coll. « Translators' Perceptions of Literary Post-editing Using Statistical and Neural Machine Translation ». *Translation Spaces*, n° 7, vol. 2, 2018, p. 240–262.
- ◇ Nagendran, Myura, et coll. « Artificial intelligence versus clinicians ». *The British Medical Journal*, n° 8239, 2020, dx.doi.org/10.1136/bmj.m689.
- ◇ Rao, Delip. « GPT-3 and a Typology of Hype ». *Page Street Labs*, 26 juill. 2020 (consulté le 11/12/2020).
- ◇ Rossi, Caroline. « Les usages actuels de la traduction automatique (TA) ». *Les humanités numériques en langues*, Ateliers Digit-Hum, 17 oct. 2019, Paris, France.

Bibliography

- ◆ Sharir, Or, et coll. « The Cost of Training NLP Models ». *arXiv*, 2020, arXiv:2004.08900.
- ◆ Szymański, Piotr, et coll. « WER we are and WER we think we are ». *arXiv*, 2020, arXiv:2010.03432.
- ◆ Taivalkoski-Shilov, Kristiina. « Ethical issues regarding machine(-assisted) translation of literary texts ». *Perspectives*, n° 27, vol. 5, 2019, p. 689–703.
- ◆ Zydroń, Andrzej. « De-Demonizing AI Regarding Localization ». *Proceedings of the 41th Conference “Translating and the Computer”*, AsLing, 21–22 nov. 2019, Londres, Royaume-Unis, édité par David Chambers et coll., Éditions Tradulex, 2019, p. 137–144.

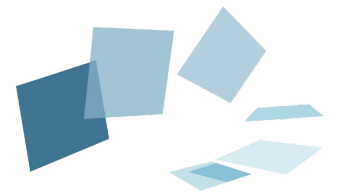
Merci pour votre attention

damien.hansen@uliege.be



LIÈGE université
UR CIRT

Centre Interdisciplinaire de Recherches
en Traduction et en Interprétation



Laboratoire d'Informatique de Grenoble

PARALLEL SESSIONS