

Un rapide état des lieux

François YVON
LISN, CNRS et Univ. Paris Saclay
francois.yvon@limsi.fr

for the European Language Resource Consortium

www.lr-coordination.eu



Une ambition: traiter la langue dans tous ses états

- Ecrits y-c. gestion de document / fouille de texte
- Paroles et Voix
- Signes (langue des signes)
- Documents (écriture manuscrite et OCR)
- Images et vidéos (inscrustation)

+ Variation, Multimodalité, Multimédia, Multilinguisme

Traitement automatique de la parole

- Transcription automatique: indexation, sous-titrage, dictée vocale
- Synthèse vocale
- Traduction de parole
- Caractérisation du locuteur: identification, reconnaissance, émotions

Analyse sémantique, « compréhension »

- Systèmes de dialogues
- Résumé automatique
- Réponse à des questions précises (QA)
- Génération: des données aux texte

Fouille de textes

- Analyse d'opinion et de sentiments
- Stylométrie et attribution d'auteurs
- Extraction d'informations structurées (terminologie, ontologies, bases de connaissances)
- Monitoring des médias (discours haineux, fallacies, etc)

Technologies de la traduction

- Traduction automatique (générale, spécialisée)
- Traduction assistée et post-édition (TAO)
- Gestion de mémoires de traductions
- Alignement automatique (phrases, mots)
- Aides à la rédaction, à la localisation



Traitement de réunions multilingues

- Repérage des tours de parole
- Génération automatique des verbatims
- Traduction automatique du texte et/ou de la parole
- Reconnaissance des notes manuscrites
- Recherche (multilingue) dans les CR des réunions passées





J' ai commandé une glace au serveur



J' ai commandé une glace au serveur



J' ai commandé une glace au serveur



J' ai commandé une glace au serveur

Une recette universelle : exploiter des données **annotées**

➤ Pour **la traduction**

textes source et cible

(the cat is on the mat, le chat est sur le matelas)

➤ Pour **l'analyse de sentiments**

texte et polarité

*(Ce restaurant thai était délicieux, ***)*

➤ Pour **la transcription vocale**

enregistrement et transcriptions

(, *les technos du langage humain*)

➤ Pour **les réponses aux questions**

(texte + question, réponse correcte)

En général: des exemples (entrées – sorties)

Input
(textes)

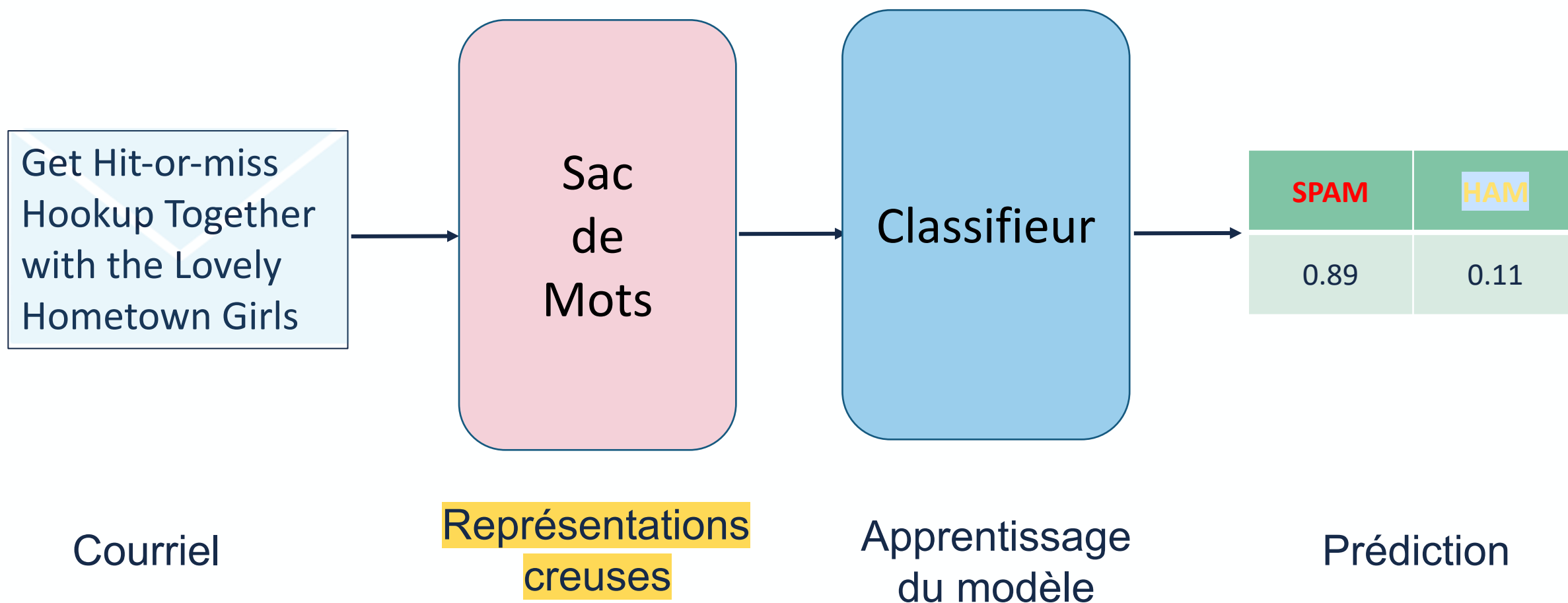
Get Hit-or-miss Hookup
Together with the
Lovely Hometown Girls

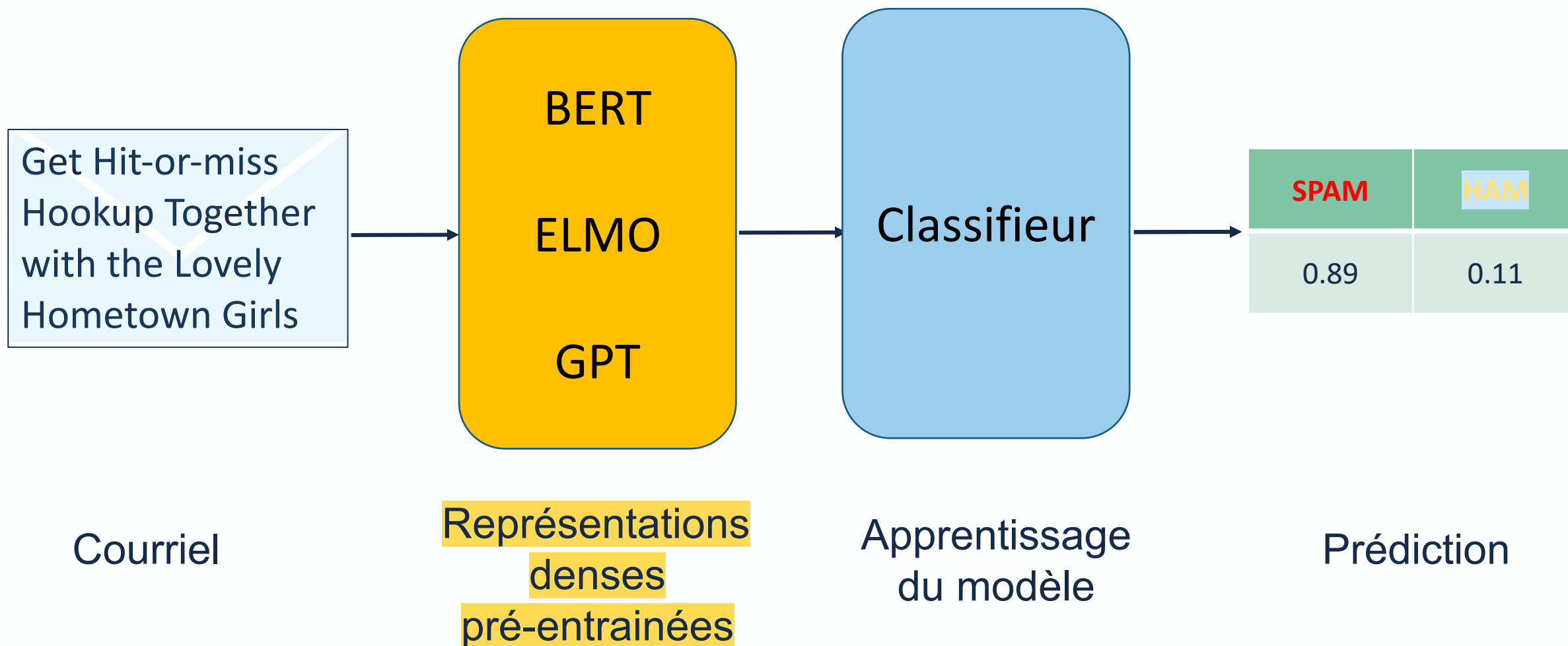
Classifieur

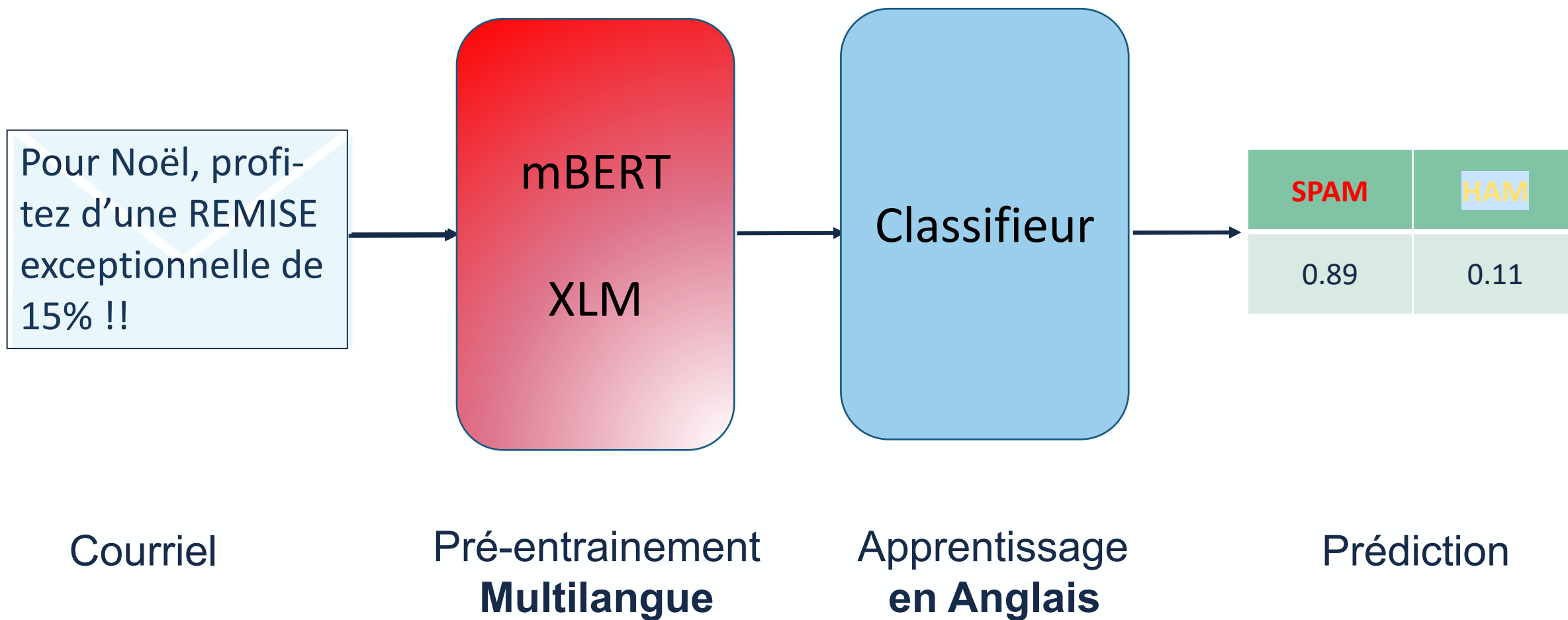
Prédiction
(probabilités)

SPAM	HAM
0.89	0.11

Apprentissage artificiel: reg. log., SVM, DT, NN,



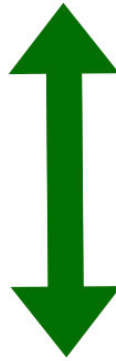
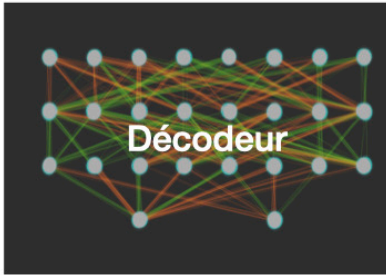
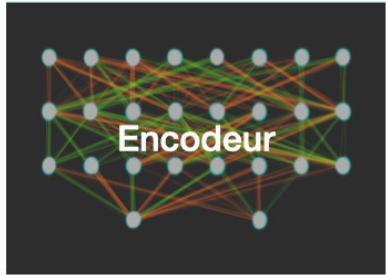




DES APPRENTIS PUISSANTS: TRANSFORMEURS

Plus le PDG de Titan, "le projet d'achat de Goodbeer Amiens démarre avec zéro employé". Après avoir parlé d'abord avec France en janvier pour la reprise partielle de la site Goodbeer Amiens (partie de la ferme, Maurice Taylor, le PDG du constructeur américain Titan, se déclare, vendredi) prêt à passer 333 emplois sur site à l'EST qui compte l'usine, Amont Montebourg, usine de démarrage produisant l'acier déjà annoncé, le 21 octobre. Après avoir annoncé au projet d'achat de l'usine en janvier, vous avez annoncé aujourd'hui. Vous avez annoncé de nouvelles usines et des machines, les parties de "non-déclarées", qui "travaillent plus heures" par jour de standards "Tian", en objet à l'EST. On ne comprend pas cette volonté. Et ce pour faire plaisir à M. Montebourg ? Le message de faire plaisir à personne. Seul à ma femme. M. Montebourg est un charmant jeune homme qui aime à sauver des emplois industriels par les moyens possibles, et il ne devient jamais par mes paroles, (en tous cas). Mais à France à un haut niveau d'emploi et de chômage aussi. La vérité est que l'usine ? Travailler sept heures par jour quand dans d'autres pays on travaille huit heures hebdomadaire en France. Ça n'est pas, en Chine, en Inde, en Chine et dans d'autres pays, on travaille de 4 à 6 heures par jour. Mais je n'ai pas de privilège sur la France. Ce que je vois, c'est une usine qui fabrique de bons produits agricoles, et de bons équipements, et dans d'autres pays on travaille huit heures hebdomadaire en France. Ça n'est pas, en Chine, en Inde, en Chine et dans d'autres pays, on travaille de 4 à 6 heures par jour. Mais, avec un prix correct et des travailleurs compétents, cela vaut le coup d'investir. Quel type d'accord attendez-vous entre la CGT et Goodbeer ? Si, après l'annonce de la fermeture de l'usine, Goodbeer avait offert aux salariés une bonne indemnité de départ, au moins que 100% des employés auraient accepté. Maintenant, imaginez que Tian achète à Goodbeer l'usine Amiens. Dès lors, Tian pourrait démanteler les machines vers la Pologne ou vers tout autre pays de l'Union européenne ou à l'étranger. Je pense que M. Montebourg sait cela. Or, à ma femme l'usine d'Amiens avec au moins 333 emplois bien payés. Tian accepte de ne recruter que les quelques 1000 salariés actuels de Goodbeer. Amont Montebourg n'est pas un engagement de Tian avant d'acheter à Goodbeer la CGT à l'usine à la base avec Goodbeer. Le premier étape est que la CGT et Goodbeer doivent un accord sur les indemnités de départ pour tous les employés. Dès lors, il n'y aura plus d'employés chez l'usine. M. Montebourg a dit que vous êtes prêt à garantir ces 333 emplois durant quatre ans. Le constructeur ? Le seul nombre que je mentionne au ministre, c'est 333. Je sais qu'il voudrait une garantie de quatre ans. Mais, comme je vous l'ai dit, la CGT et Goodbeer doivent un accord sur les indemnités de départ. Ce sont les conditions à accepter. Le projet d'achat de l'usine démarre avec zéro employé. Comment pourrions-nous donner des garanties de durée d'emploi quand il n'y a plus d'employés sur le site ? Si M. Montebourg parvient à ce que la CGT et Goodbeer se mettent d'accord que Tian achète l'usine, nous avons déjà l'intention de rester à Amiens. Nord plus de quatre ans.

Titan to buy Goodbeer Amiens with zero employees". Titan CEO says After dramatically throwing in the towel in January over the partial facility over the Goodbeer site in various parts which will be to close, Maurice Taylor, CEO of American manufacturer Titan, now says that he is ready to pass 333 of the factory's 1,137 employees, Amont Montebourg, Minister of Industrial Revival, had already announced on Monday October 21, after giving up the idea to buy the factory in January, you come back today. You had spoken of violent attacks and insults, speaking of "non-called workers", who "work three hours per day", "real" workers, the CGT. It's hard to understand this U-turn. It is hard to please M. Montebourg? I'm not trying to please anyone. Except my wife. M. Montebourg is a charming young man who is really a good person of the best paid industrial jobs. I'm sorry my words have offended anyone. But there are high levels of tax and unemployment in France too. Does the truth offend you? Working seven hours a day when people in other countries are working eight hours a day in France. In China, India and many other countries, people work to twelve hours a day. But I don't have prejudices about France. What I see is a factory which manufactures good agricultural items, has good equipment, a good location and plenty of room for growth. Why does Tian need this factory in Amiens? Tian doesn't need to buy this factory. But, if the price is right and the workers are qualified, it's worth trying. What sort of agreement do you expect between the CGT and Goodbeer? If Goodbeer had offered the employees a good severance package after announcing that the factory was closing, I think 100% of the employees would have accepted it. Now, let's imagine that Tian buys a closed factory from Goodbeer. At that point, Tian could choose to move the machines to Poland or any other country in the European Union that will have its own currency. I think M. Montebourg knows that. However, he wants to keep the factory in Amiens with at least 333 well-paid employees. Tian has agreed to recruit them from the L200 or so people currently working at Goodbeer. Also, M. Montebourg expects a commitment from Tian before going to get the CGT to do deal with Goodbeer. The first step is for the CGT and Goodbeer to sign an agreement on severance pay for all the employees. Then there won't be any employees left in the factory. M. Montebourg has said that you are prepared to guarantee these 333 jobs for four years. Can you confirm that? The only number I mentioned to the minister is 333. I know he would like a four-year guarantee. But, at least to you, the CGT and Goodbeer must first reach an agreement on the severance pay. If all the employees accept it, the project of buying Goodbeer's plant and people with zero employees. How can we give guarantees on the length of employment when there are no employees left on site? I think M. Montebourg gets the CGT and Goodbeer to reach an agreement and Tian buys the factory, we have enough employees in Amiens North for more than four years.



For the CEO of Titan, "the proposed purchase of Goodbeer Amiens starts with zero employees". Having thrown in the towel in January for the partial acquisition of the site Goodbeer Amiens (part of the farm, Maurice Taylor, the CEO of the US tire manufacturer Titan, expressed, today, ready to pass 333 jobs on the 1,137 in the factory, Jean-Michel Montebourg, Minister of Industrial Revival, had already announced on Monday 21 October. Having announced the proposed purchase of the plant in January, you come back today. You had spoken of violent attacks and insults, speaking of "non-called workers", who "work three hours per day", "real" workers, the CGT. It's hard to understand this U-turn. It is hard to please M. Montebourg? I'm not trying to please anyone. Except my wife. M. Montebourg is a charming young man who is really a good person of the best paid industrial jobs. I'm sorry my words have offended anyone. But there are high levels of tax and unemployment in France too. Does the truth offend you? Working seven hours a day when people in other countries are working eight hours a day in France. In China, India and many other countries, people work to twelve hours a day. But I don't have prejudices about France. What I see is a factory which manufactures good agricultural items, has good equipment, a good location and plenty of room for growth. Why does Tian need this factory in Amiens? Tian doesn't need to buy this factory. But, if the price is right and the workers are qualified, it's worth a try. What do you expect agreement between the CGT and Goodbeer? If Goodbeer had offered the employees a good severance package after announcing that the factory was closing, I think 100% of the employees would have accepted it. Now, let's imagine that Tian buys from Goodbeer plant closed. Therefore, Tian could move the machines to Poland or any other European Union country that will have its own currency. I think M. Montebourg knows that. However, he wants to keep the factory in Amiens with at least 333 well-paid jobs. Tian has agreed to recruit them from among the approximately 1,000 Goodbeer current employees. How can we give guarantees on the length of employment when there are no employees left on site? I think M. Montebourg gets the CGT and Goodbeer to reach an agreement and Tian buys the factory, we have enough employees in Amiens North for more than four years. How can we give guarantees on the length of employment when there are no employees left on site? I think M. Montebourg gets the CGT and Goodbeer to reach an agreement and Tian buys the factory, we have enough employees in Amiens North for more than four years.

Des algorithmes aux comportements **stupéfiants**

Les progrès de l'IA ont été si fulgurants qu'une machine peut maintenant pour la première fois être programmée pour exécuter un programme.

Des algorithmes aux comportements stupéfiant

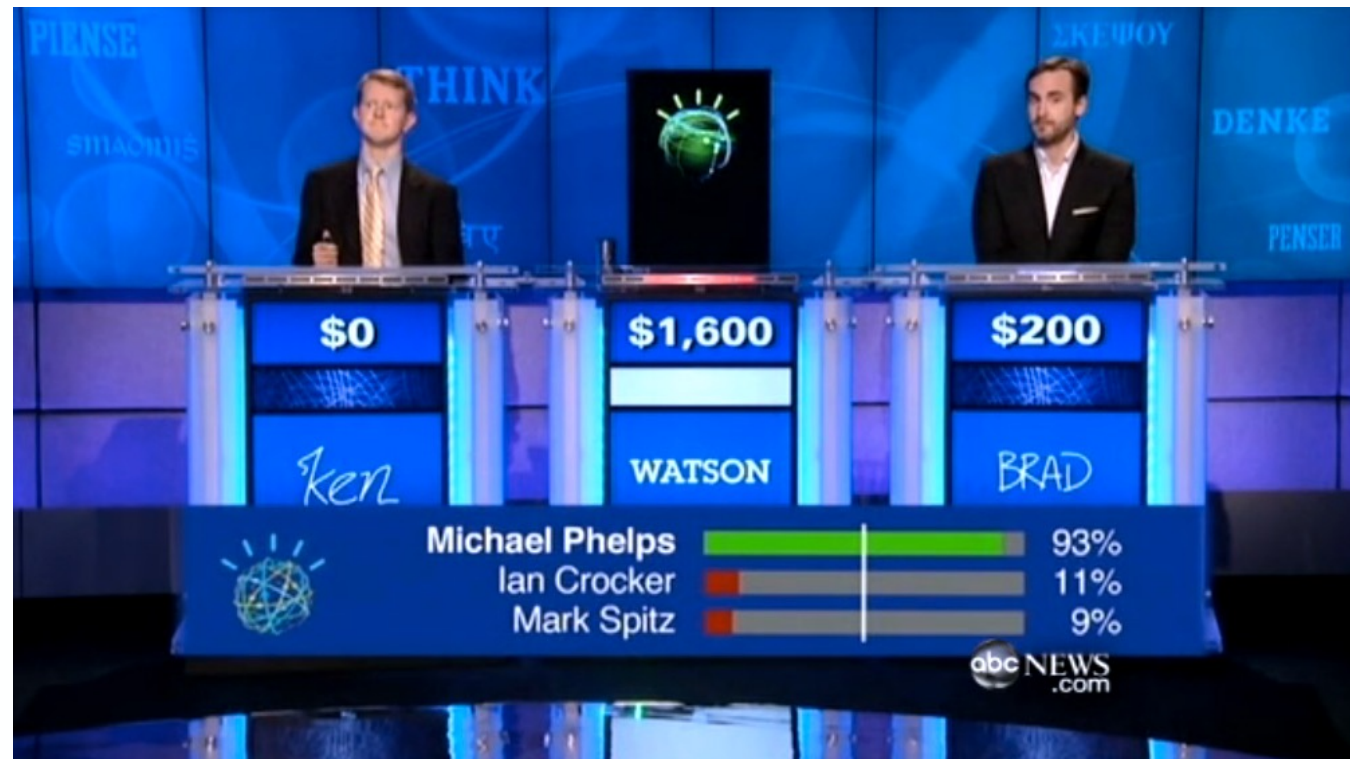


Nicolas Sarkozy arbitrera la semaine prochaine sur les amendements un bon cru pour le tourisme . Dès cet été des milices et les étrangers sont venus à Paris , bonsoir de fréquentation des hôtels et du tourisme d'affaires sur les six premiers mois de la vie .

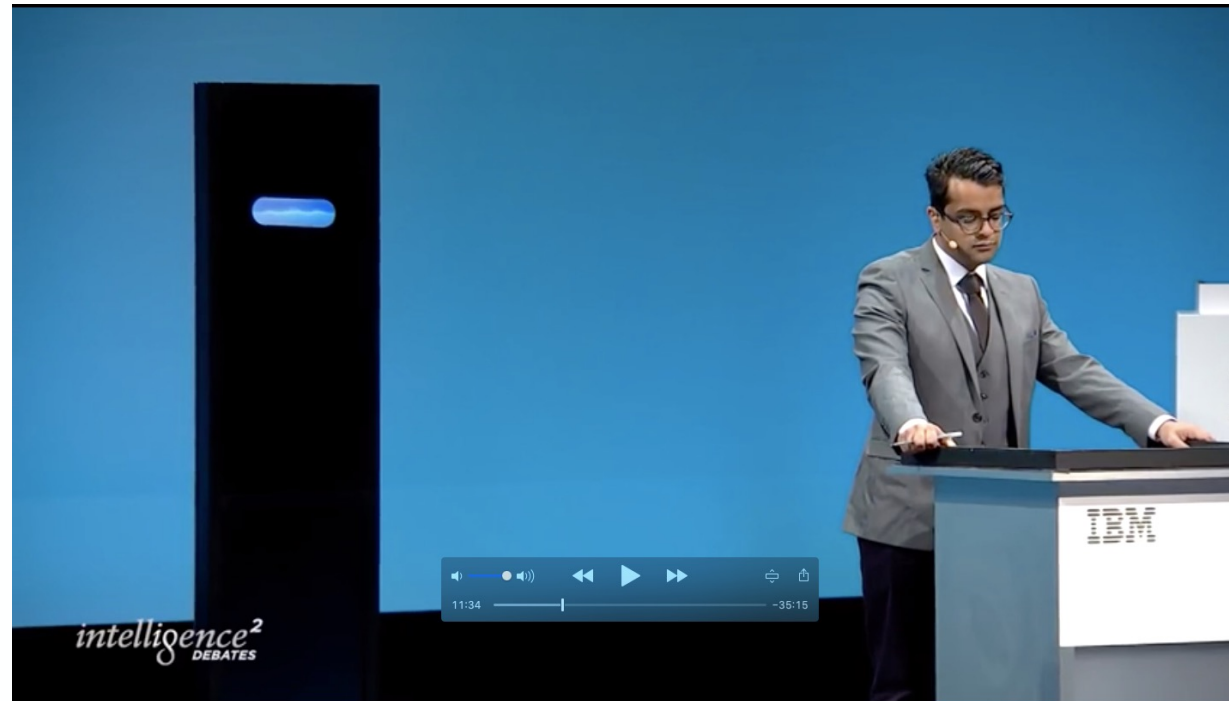
Et direction tout de suite . La Rochelle , où les socialistes font l'heure rentrée politique du PS , réuni au grand complet à l'exception de Dominique Strauss-Kahn , Amandine a -t-elle Hayat . Vous êtes sur place . Ségolène Royal est apparemment à très présente un depuis hier ..

Oui . Marianne et c'est clairement aujourd'hui la journée de Ségolène Royal , elle entend bien marquer des points . L'an dernier , elle avait survolé ses universités d'été avec des sauter , c'est bien cette fois , elle sera présente pendant les trois jours et cela commencé soir avec un dîner de presse . Cela s'est poursuivie ce matin avec de nouveau une foule de médias pour une conférence de presse sur la crise laitière et puis cet après- midi , ce sera bien sûr le traditionnel discours d'ouverture devant 4000 la militant discours qui s'annonce particulièrement offensif pour Ségolène Royal . avec en fil rouge les

Des algorithmes aux comportements stupéfiants



Des algorithmes aux comportements **stupéfiant**s



https://www.youtube.com/watch?v=3_yy0dnlc58

Des algorithmes aux comportements stupéfiants

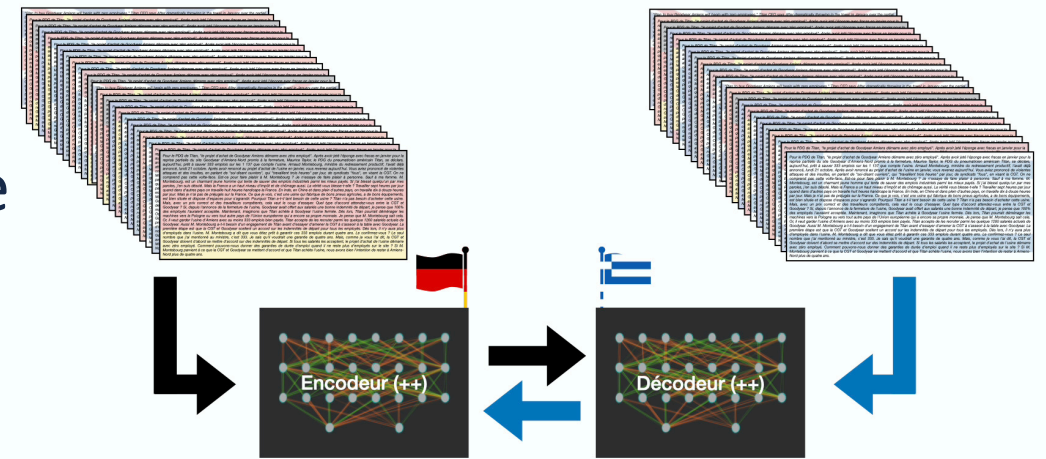


The screenshot shows a machine translation interface with two columns. The left column is labeled 'Anglais (langue détectée)' and contains the text: 'EU policy protects the environment and seeks to minimise risks to climate, human health and biodiversity. The European Green Deal aims to make Europe the world's first climate-neutral continent, in part by developing cleaner sources of energy and green technologies.' The right column is labeled 'français' and contains the translation: 'La politique de l'UE protège l'environnement et cherche à minimiser les risques pour le climat, la santé humaine et la biodiversité. Le "Green Deal" européen vise à faire de l'Europe le premier continent neutre sur le plan climatique, notamment en développant des sources d'énergie plus propres et des technologies vertes.' A double-headed arrow icon is positioned between the two columns. The interface also includes a 'Glossaire' button and a 'automatique' dropdown menu.

Mais subtilement trompeurs

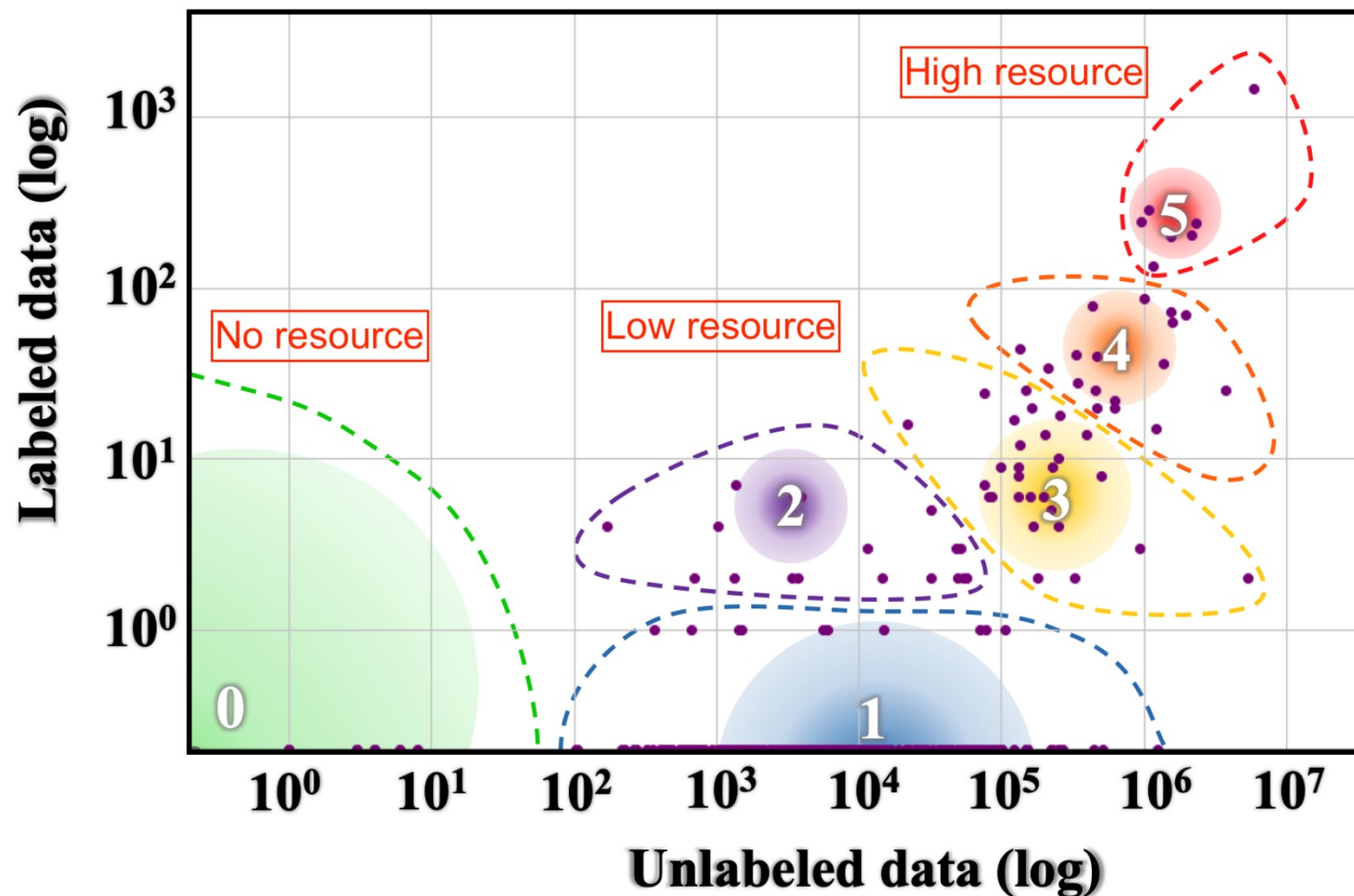
- Dépendances aux données et biais d'apprentissage
- Manque de robustesse face aux imprévus
- Opacité des décisions et des erreurs
- Ni compréhension ni raisonnement

- Puissance des modèles multilingues
- Pré-entraînement et supervision naturelle
- Complémentarité des modalités
- L'apprentissage « bout-en-bout »



- Les langues « peu dotées »
- Améliorer le transfert entre tâches, domaines et modalités

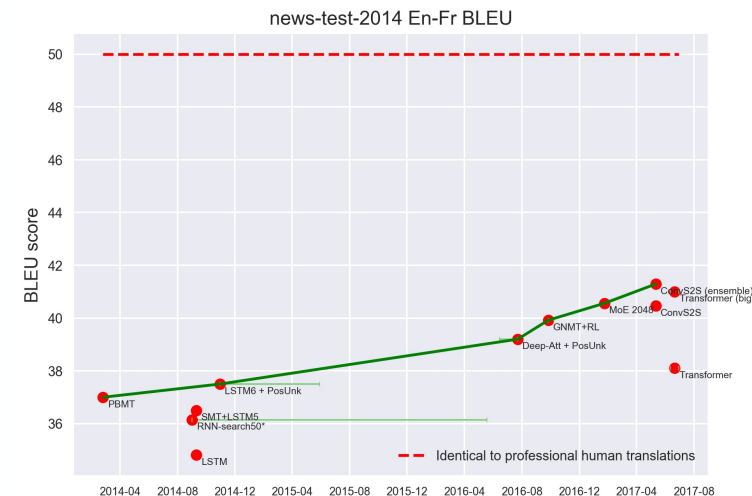
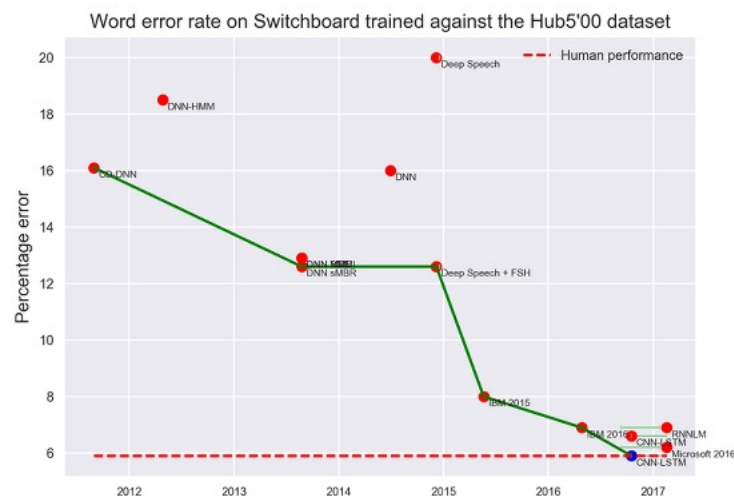
LES LANGUES PEU DOTÉES SONT MAJORITAIRES



- De multiples acteurs, couvrant tout **les spectre des technologies** dialogue, synthèse & transcription, traduction, génération, fouille de sentiments, QA
- Foison de **ressources généralistes** (dictionnaires, corpus, modèles)
- Un **tissu académique solide** : 3IA, CNRS, INRIA, Universités, R&D industrielle

- Trop peu de **domaines** couverts
- Pas assez de données **très annotées**
- Faible couverture pour **les langues régionales**
- Des **freins systémiques** au partage des données

- Métriques automatiques pour développer les systèmes
- Evaluations humaines pour l'utilisabilité & l'acceptabilité



- Des corrélations incertaines, des conclusions hasardeuses

Website: www.lr-coordination.eu
Twitter: @LR_Coordination
Email: info@lr-coordination.eu

