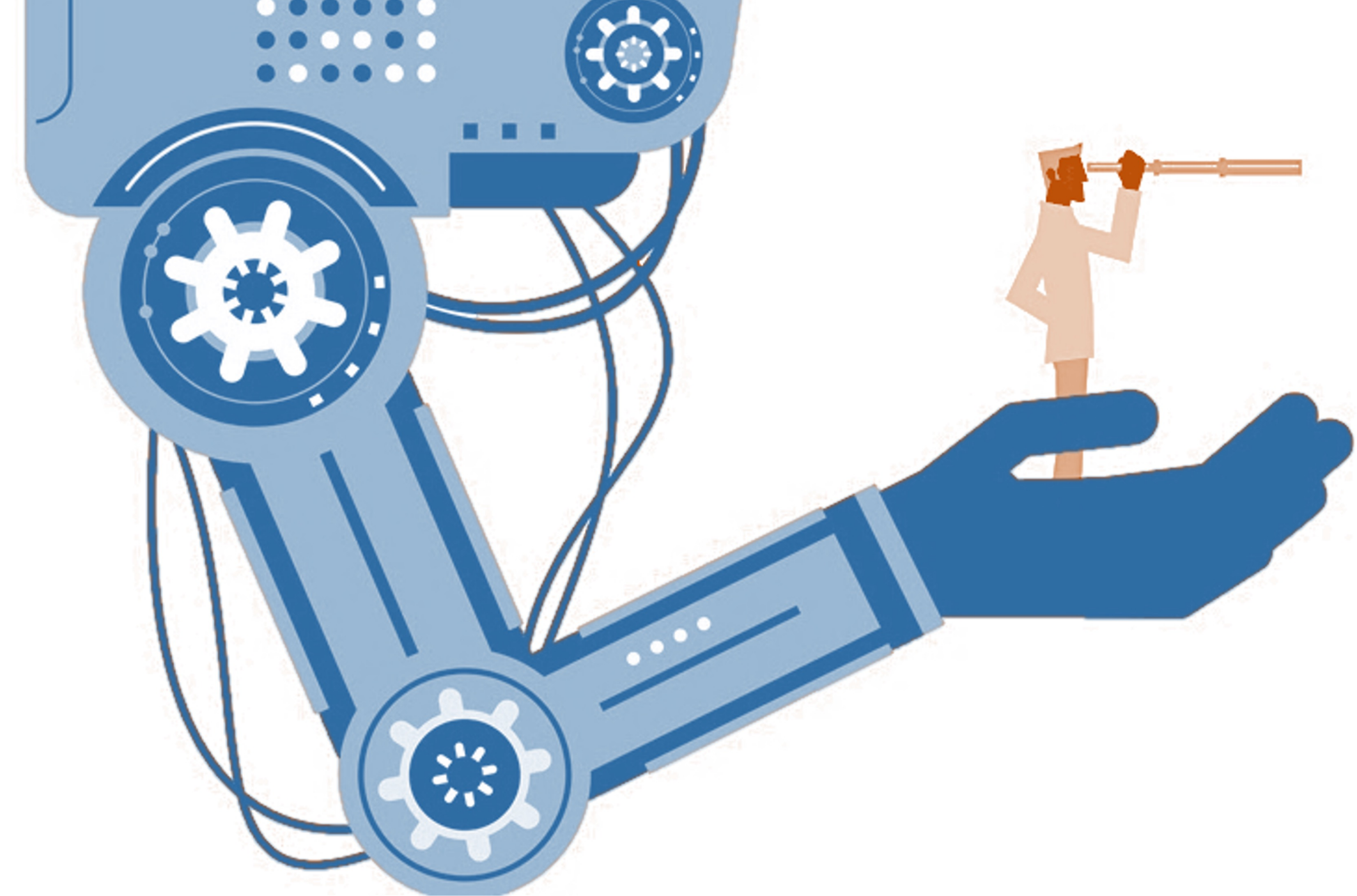




Mieux comprendre pour mieux classer

La linguistique et l'IA au service de l'explicabilité

Louis Escoufflaire – Jérémie Bogaert – Marie-Catherine de Marneffe
Antonin Descampe – François-Xavier Standaert – Cédrick Fairon

Dans le paysage journalistique actuel, la question de la place de l'**opinion** est de plus en plus importante. Afin de mieux comprendre et identifier ce qui rend un article en français plus **subjectif**, nous cherchons à créer un modèle de classification explicable.



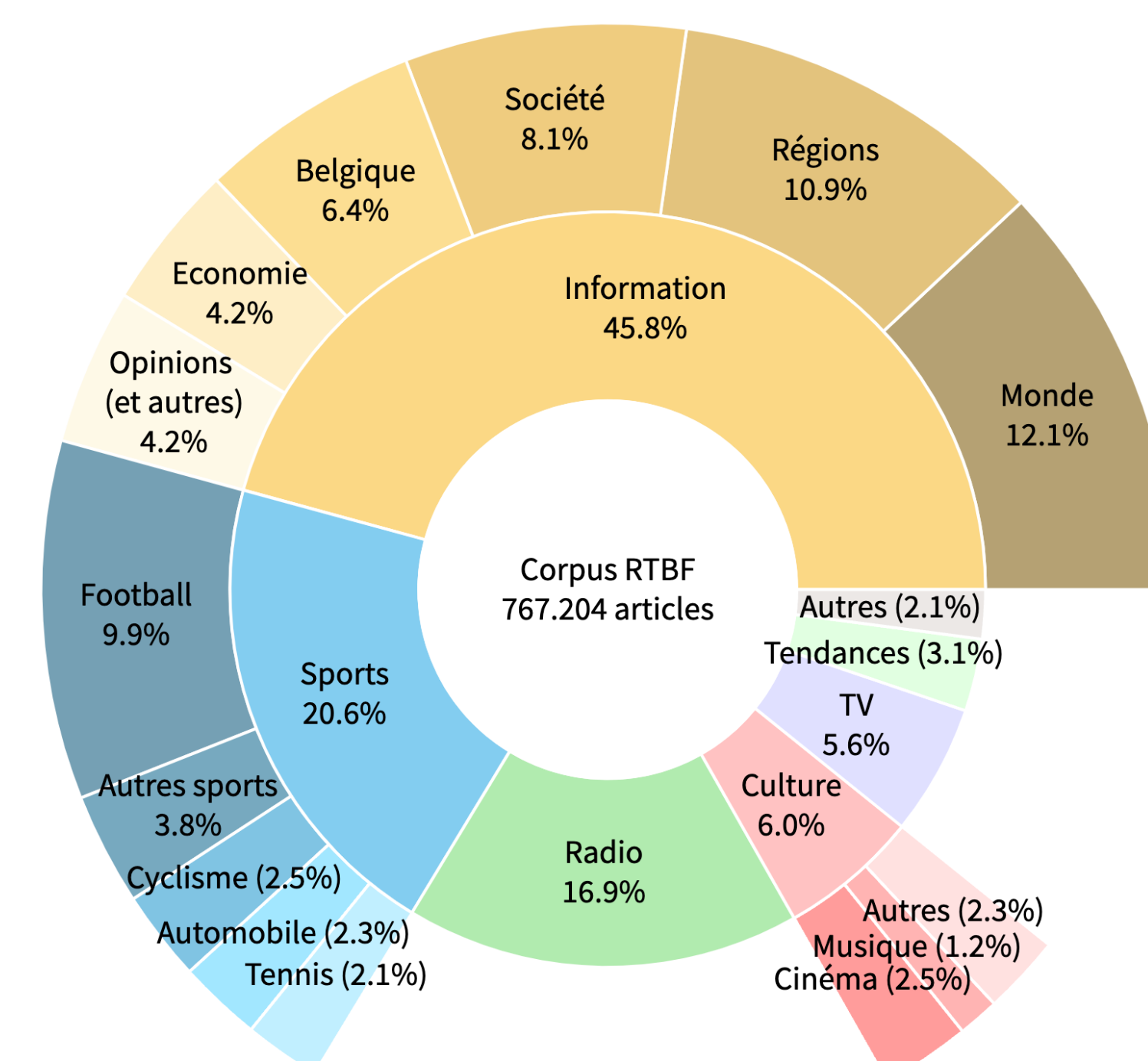
OPINION



VS. INFORMATION

Notre projet a pour but d'explorer la **complémentarité** pour la classification de textes subjectifs entre :

- et les **approches symboliques** (règles linguistiques)
- les **approches deep learning** (intelligence artificielle)



Le **Corpus RTBF** est le plus grand corpus francophone de textes journalistiques en libre accès. Les articles ont été publiés entre 2008 et 2021 sur le site web de la RTBF.

Le **Corpus InfOpinion** est un corpus de 10.000 articles :

- **5.000** articles **opinion** (chroniques, éditoriaux)
- **5.000** articles **information** (dépêches, actualités)

Modèle linguistique (par règles)



Explications LAM

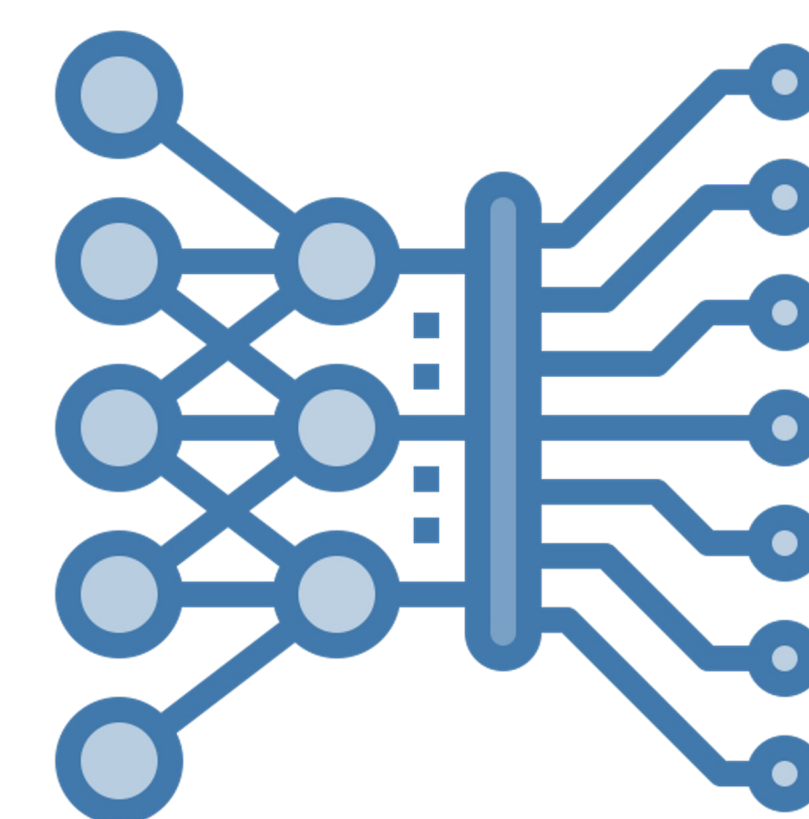
Linguistic attention mapping

Nous **comparons deux modèles** sur une tâche de **classification d'articles** de presse (**opinion vs. information**), puis évaluons un **modèle hybride** composé de connaissances issues du modèle linguistique et du modèle IA. Nous cherchons ensuite à **expliquer** sur quels éléments se base chaque modèle pour classer un texte.

- Le **modèle linguistique** se base sur des indicateurs lexicaux, grammaticaux et stylistiques. Il est rapide et interprétable, mais ne tient pas compte du contexte.
- Le **modèle IA** est basé sur une architecture *transformers* (CamemBERT). Il est long à entraîner et peu interprétable, mais prend en compte le contexte.

	Modèle LING	Modèle IA
Précision	86.9%	96.4%
Temps d'entraînement total	15 min.	13 jours
Temps de classification (pour 1000 articles)	1 sec.	9 sec.

Modèle IA (CamemBERT)



Explications LRP

Layerwise relevance propagation

Le gouvernement de Charles Michel est divisé avant un budget, rien d'étonnant en fait... Ce qui se passe est d'une banalité affligeante. On retrouve dans la séquence qui se déroule en ce moment une bonne partie des maux qui frappent la politique belge depuis au moins 15 ans, depuis le dernier gouvernement Dehaene. On retrouve des négociations marathons où telle taxe, telle coupe dans les soins de santé est décidée au bout de la nuit parce qu'il faut bien avoir quelque chose à livrer aux médias et au parlement.

Les **cartes d'attention** montrent quels éléments d'un texte ont influencé le plus la décision du modèle.

Similarité faible entre les deux modèles
 $Similarité\ cosinus = 0,23 \quad (-1 > s > 1)$

Un **article d'opinion** analysé avec le **modèle linguistique**.

Le **même article** analysé avec le **modèle IA**.

Le gouvernement de Charles Michel est divisé avant un budget, rien d'étonnant en fait... Ce qui se passe est d'une banalité affligeante. On retrouve dans la séquence qui se déroule en ce moment une bonne partie des maux qui frappent la politique belge depuis au moins 15 ans, depuis le dernier gouvernement Dehaene. On retrouve des négociations marathons où telle taxe, telle coupe dans les soins de santé est décidée au bout de la nuit parce qu'il faut bien avoir quelque chose à livrer aux médias et au parlement.

Motifs récurrents du modèle IA :

Articles d'**opinion** :

- adverbes et adjectifs
- marqueurs de discours
- Interjections
- concepts abstraits
- vocabulaire politique



Articles d'**information** :

- jours de la semaine
- verbes de parole et de citation
- références à des sources d'information
- vocabulaire de la guerre

En enrichissant le modèle linguistique avec 3 variables issues des motifs du modèle IA, nous parvenons à **augmenter sa précision de classification de 1,2%** (jusqu'à 88,2%), sans diminuer sa rapidité ni son explicabilité.

OPINION		INFORMATION	
LING	IA	LING	IA
...	langage	"	jeudi
;	imaginer	participé	souligne
?	époque	<NBR>	précise
sans	vrai	située	vendredi
ne	oui	arrêtés	explique
populiste	humour	indiqué	mercredi
lequel	etc	tué	ajouté
qui	eh	attendant	selon
féministe	idéologie	précisé	raconte
où	bref	blessé	AFP

Les 10 mots auxquels chaque modèle accorde le plus d'attention en moyenne pour chaque classe.

Perspectives de recherche futures

Insérer des couches de **règles linguistiques** dans le modèle de **classification IA** pour le rendre **plus explicable**.



Mettre en place une expérience d'**annotation humaine** de mots subjectifs et étudier la **perception** des cartes d'attention par l'humain.



Analyser les faiblesses de ces modèles face aux **exemples adversariaux** (qui exploitent les failles des modèles pour les tromper).

