

Textual and sentimental analysis

Application to the media coverage of the Right to be forgotten

Andrés MORALES, Data scientist, SESSTIM

Agnès DUMAS, Sociologue, SESSTIM

Contexte : L'accès à l'assurance emprunteur

- Crédit (à la consommation, professionnel, immobilier) doit être couvert par une assurance
→ Risque décès, risque invalidité : **Questionnaire médical**
- Difficultés d'accès pour les personnes avec un risque aggravé de santé
 - Surprimes (plus cher)
 - Exclusions de garanties (risques non couverts)
 - Refus
- 1991 : Naissance d'une approche « conventionnelle » : Etat + assureurs
→ Améliorer l'accès à l'assurance **décès** pour les prêts immobiliers aux personnes **séropositives**
- 2001 : Convention Belorgey : + banques + consommateurs + **malades**
→ Risque **invalidité**

Contexte législatif



- 2007 : Loi consacrant la convention AERAS
« s'Assurer et Emprunter avec un Risque Aggravé de Santé »
- 2016 : Adoption du Droit à l'oubli (**DAO**) : supprime l'obligation de déclaration de l'antécédent de **cancer 10 ans** après la fin des traitements
- 2022 : Loi Lemoine
 - **DAO : 5 ans** pour tous **cancers**
 - Fin du questionnaire **médical** si < 200 000€ / à échéance avant 60 ans
 - Mais aussi (et surtout) : résiliation infra-annuelle de l'assurance emprunteur immobilier, sans aucun frais et à tout moment

Contexte international

- Adoption du DAO en Europe
 - BE 2019, NL 2021, ...
 - 8 pays en 2024 (*Lawler et al., Lancet Oncol, 2024*)
 - Hétérogénéité : loi ou code de conduite, nombre d'années variable, ...
 - Volonté d'unifier au travers d'une réglementation Européenne

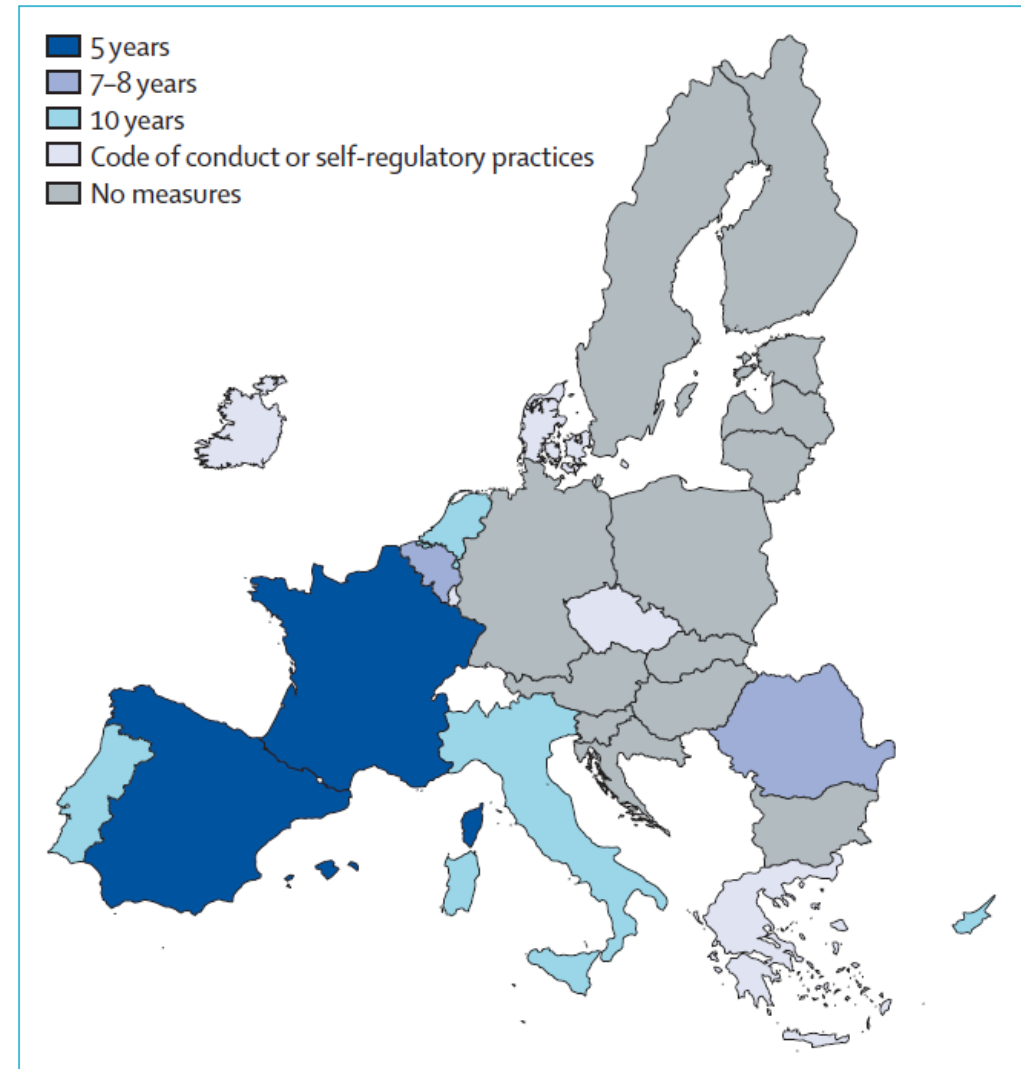


Figure: Right to be Forgotten—measures in place in the EU with timeline after which a patient with cancer can be declared cured and the Right to be Forgotten triggered to avoid declaring a previous cancer diagnosis

Objectif et méthodes du projet ELOCAN

(Effects of the right to be forgotten in accessing LOan-related insurance after CANcer)

Objectif : Comprendre la *mise sur agenda* du DAO en FR, BE et NL

- Saisir le poids des contingences, des logiques situationnelles, le travail des différents acteurs (*Kingdon, 1984*)

Méthodes :

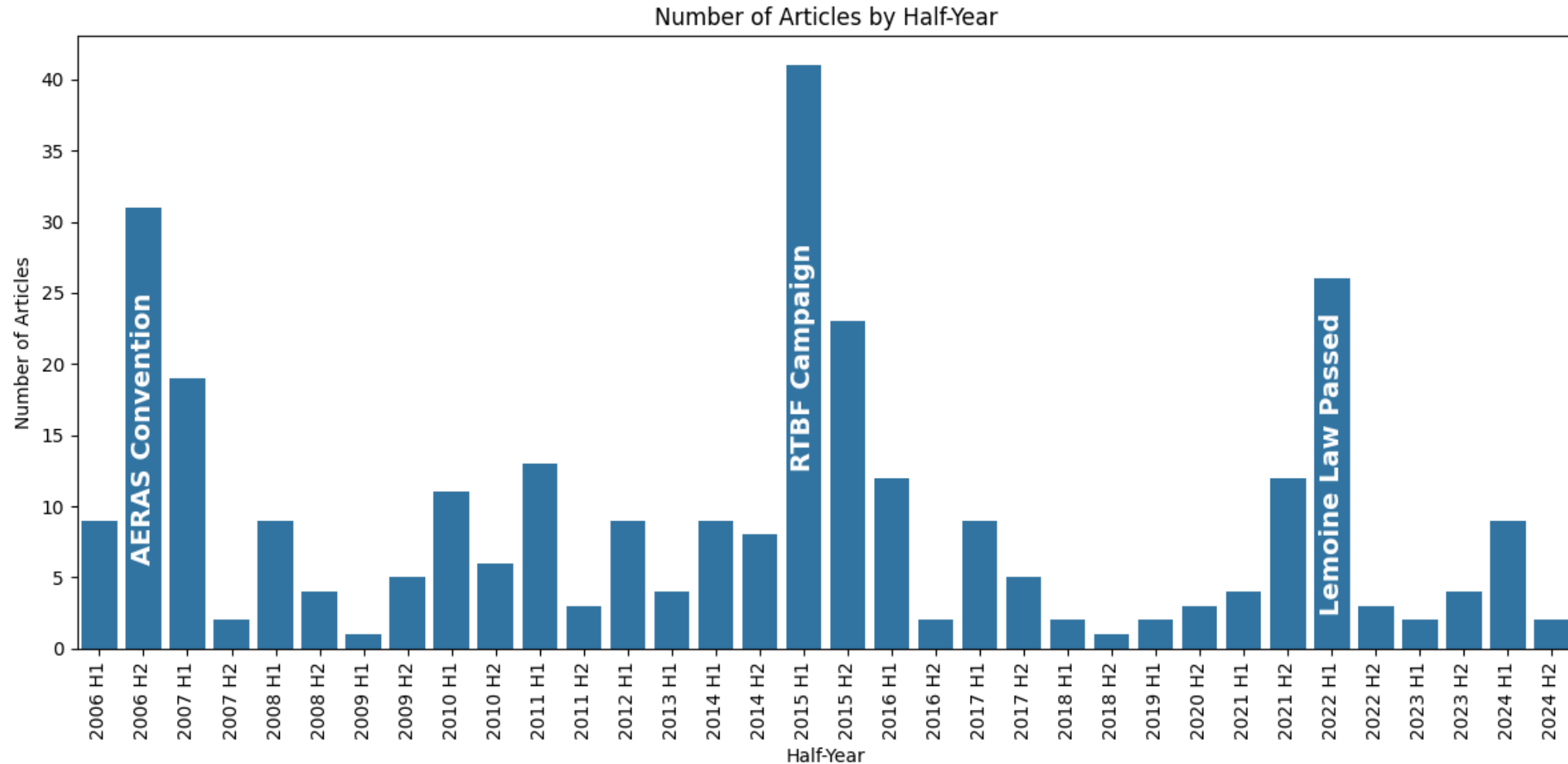
- **Entretiens** auprès d'associations, assureurs, institutionnels en FR, BE, NL (30 entretiens réalisés)
 - Stratégies diversifiées des *claims makers* selon les pays : stratégies **d'appel à l'opinion publique**
- Appuyer ces analyses qualitatives par une analyse de la **couverture médiatique (presse)**

Méthodologie de la revue de presse

- Algorithme de recherche
- Extraction des articles (Presse nationale FR) via la base de données Europresse (2004-2024)
- Conversion des fichiers html via EuroParser
- Data management → De 439 à 305 articles
- Description du corpus → nb articles, journaux, mots (*Voyant tools*) 
- **Identifier l'orientation des sentiments et les tendances de subjectivité dans les articles**

→ *Natural Language Processing (NLP) - Large Language Models (LLM)*

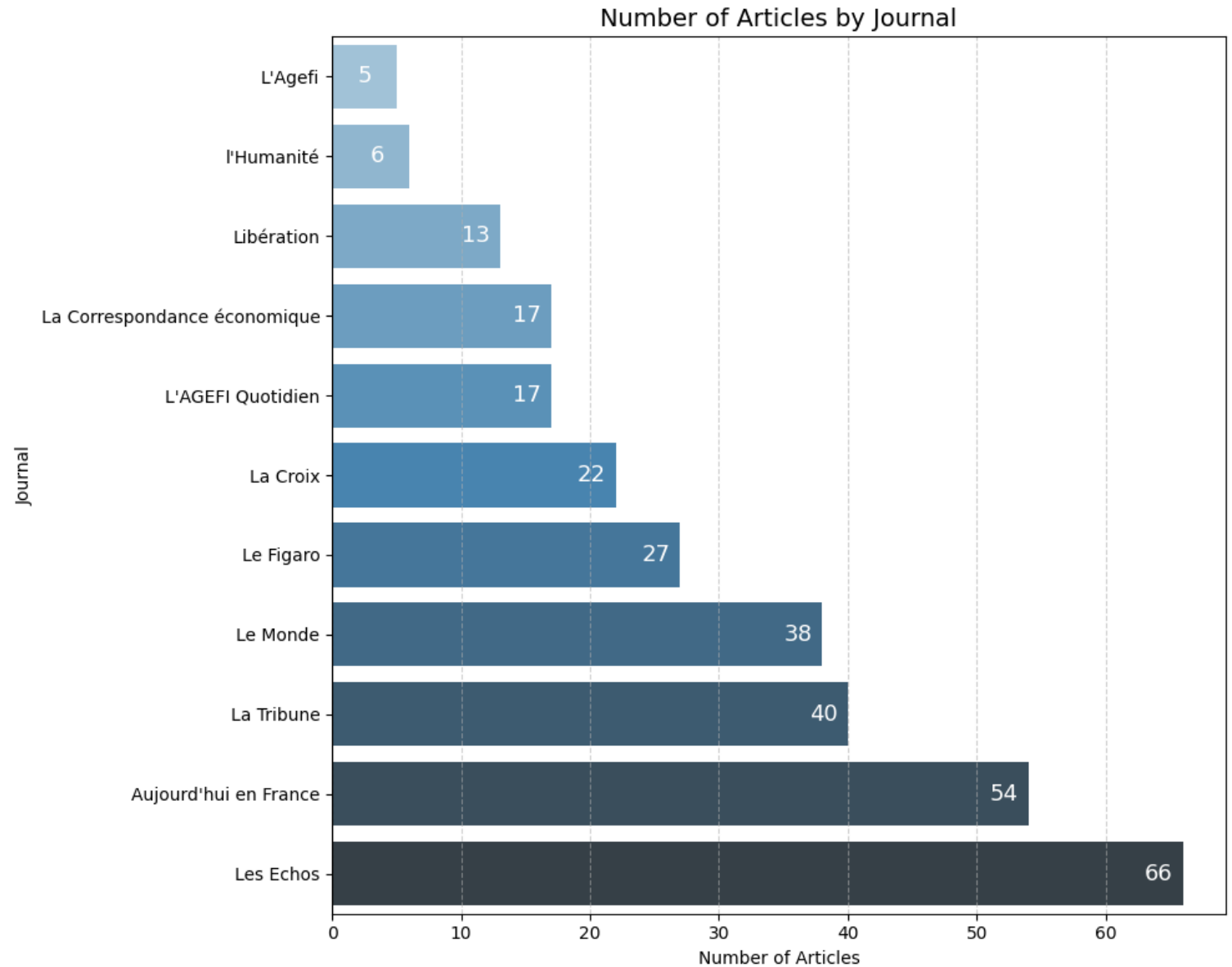
Analyse longitudinale : 3 temps législatifs



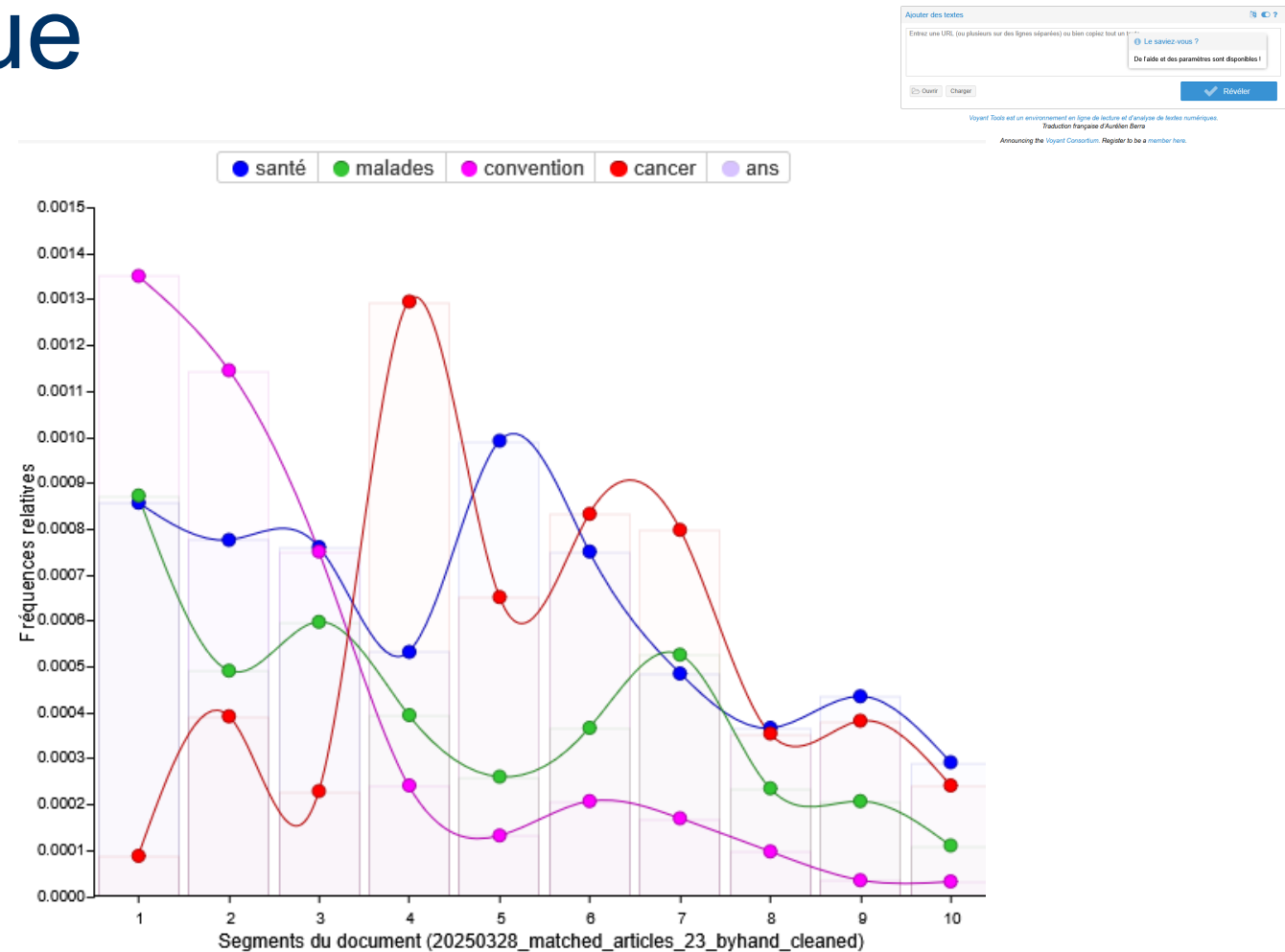
Type de journal

Renseigne sur le type d'arène
(sociale, politique, économique)
traitant le problème :

- Presse généraliste (N=160)
- Presse économique (N=145)



Analyse lexicométrique

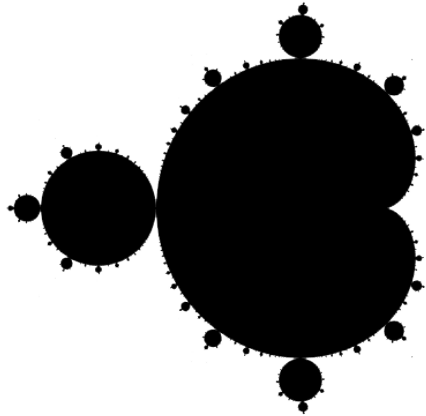


Sentiment Analysis of French Press Articles on the Right to Be Forgotten

Objective

Identify sentiment **orientation** and **subjectivity trends** on French news articles.

Methods



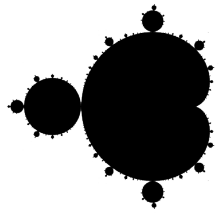
TextBlob

PatternTagger + PatternAnalyzer



Hugging Face

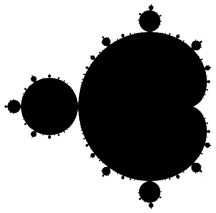
DistilCamemBERT- Sentiment



TextBlob

PatternTagger + PatternAnalyzer

"Ce film est absolument incroyable et très émouvant."

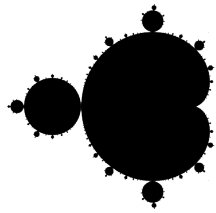


TextBlob

PatternTagger + PatternAnalyzer

"Ce film est absolument incroyable et très émouvant."

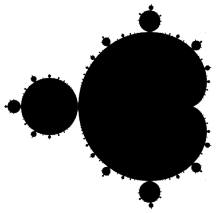
[('Ce', 'DT'), ('film', 'NN'), ('est', 'VB'), ('absolument', 'RB'), ('incroyable', 'JJ'), ('et', 'CC'), ('très', 'RB'), ('émouvant', 'JJ')]



TextBlob

PatternTagger + PatternAnalyzer

Word	Tag	Meaning
Ce	DT	Determiner (this)
film	NN	Noun
est	VB	Verb (is)
absolument	RB	Adverb
incroyable	JJ	Adjective (incredible)
et	CC	Conjunction (and)
très	RB	Adverb (very) - Intensifier
émouvant	JJ	Adjective (moving/emotional)



TextBlob

PatternTagger + **PatternAnalyzer**

Polarity

Quantifies sentiment orientation

Pessimistic

critical tone

-1

Factual

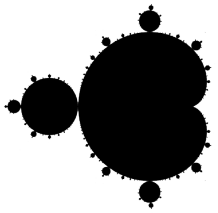
Objective reporting

0

Optimistic

favorable tone.

1



TextBlob

PatternTagger + **PatternAnalyzer**

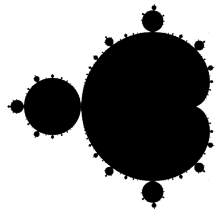
Subjectivity

measures how **opinionated** or **emotional** a sentence is

**Fully
objective**

**Fully
subjectivity**



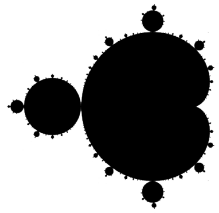


TextBlob

PatternTagger + **PatternAnalyzer**

1. Lexicon Lookups

Word	Polarity	Subjectivity	Type
incroyable	+0.7	0.8	Adjective
émouvant	+0.6	0.7	Adjective
absolument	Intensifier (+15%)	Rule-based	
très	Intensifier (+10%)	Rule-based	



TextBlob

PatternTagger + **PatternAnalyzer**

2. Polarity scoring:

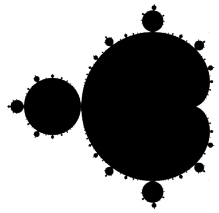
- Average of incroyable and émouvant $\rightarrow (0.7 + 0.6) / 2 = 0.65$

- Intensifiers:

- absolutement $\rightarrow +15\% \rightarrow 0.6 \times 1.15 = 0.69$

- très $\rightarrow +10\% \rightarrow 0.7 \times 1.10 \approx 0.77$

$$(0.69 + 0.77) / 2 = 0.73$$



TextBlob

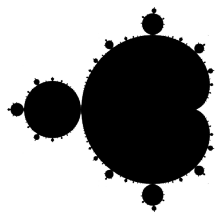
PatternTagger + **PatternAnalyzer**

2. Subjectivity scoring:

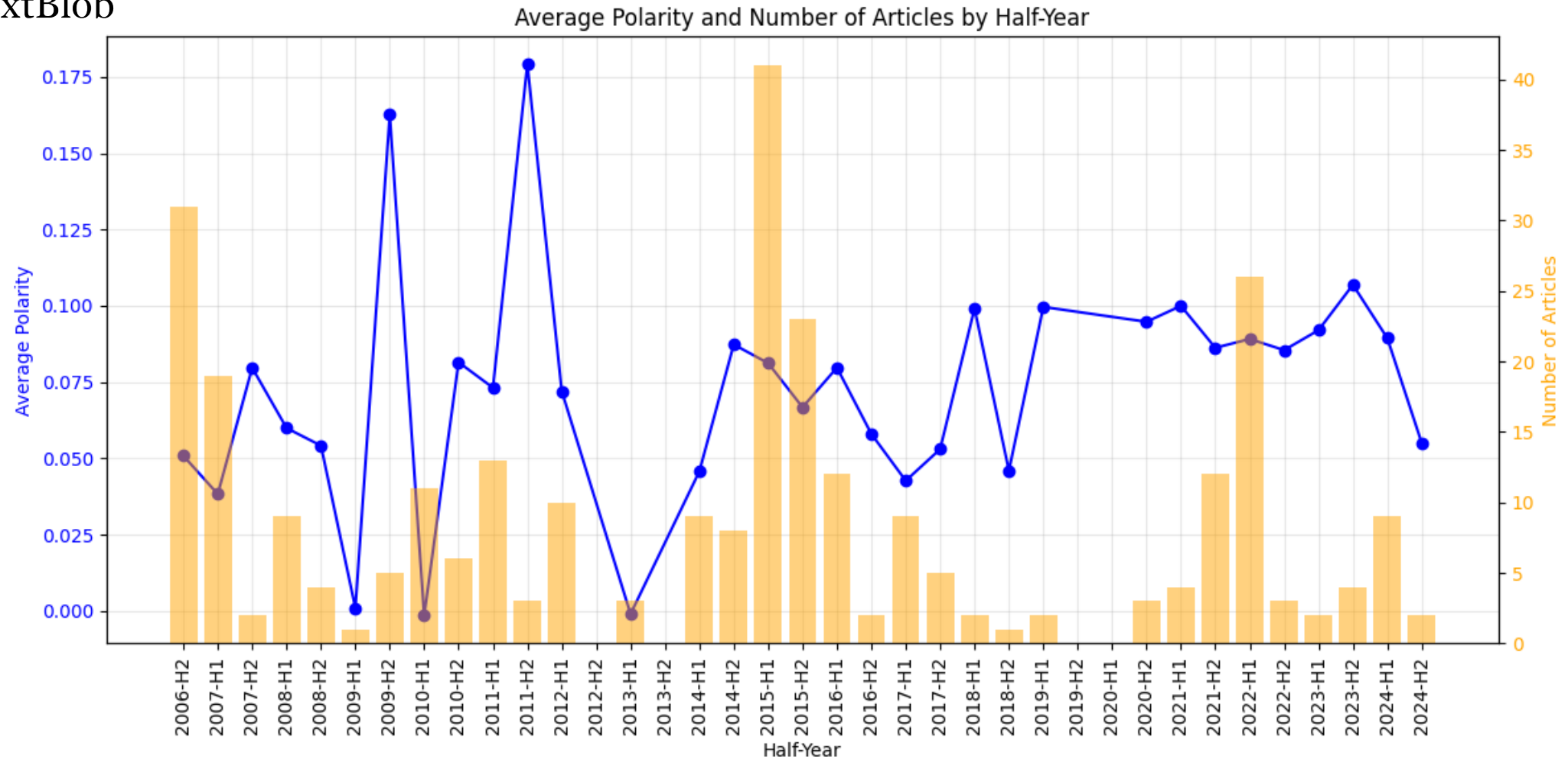
incroyable → subjectivity = **0.8**

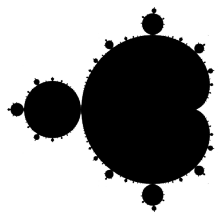
émouvant → subjectivity = **0.7**

$$(0.8 + 0.7) / 2 = \mathbf{0.75}$$

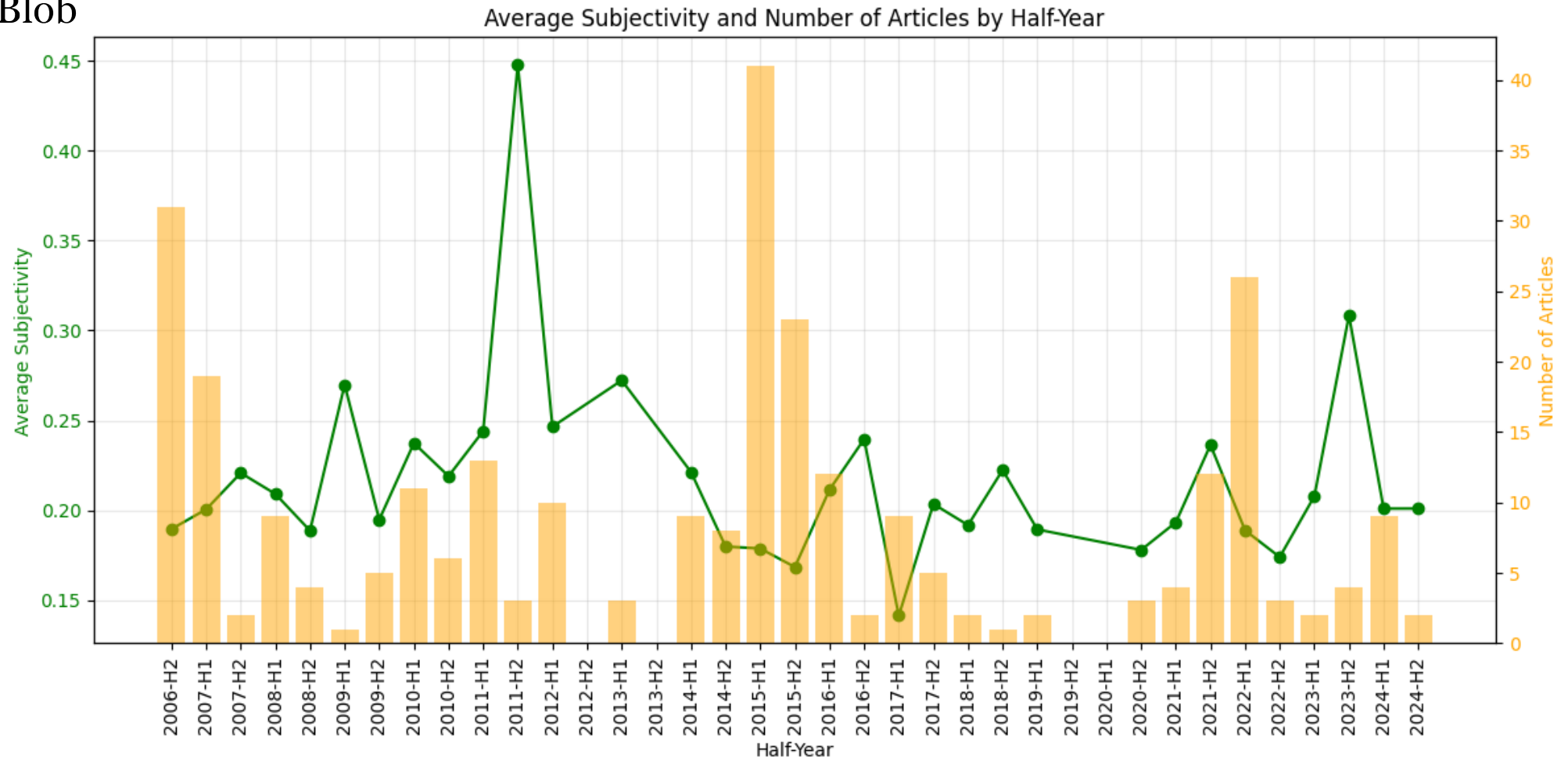


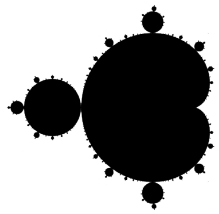
TextBlob



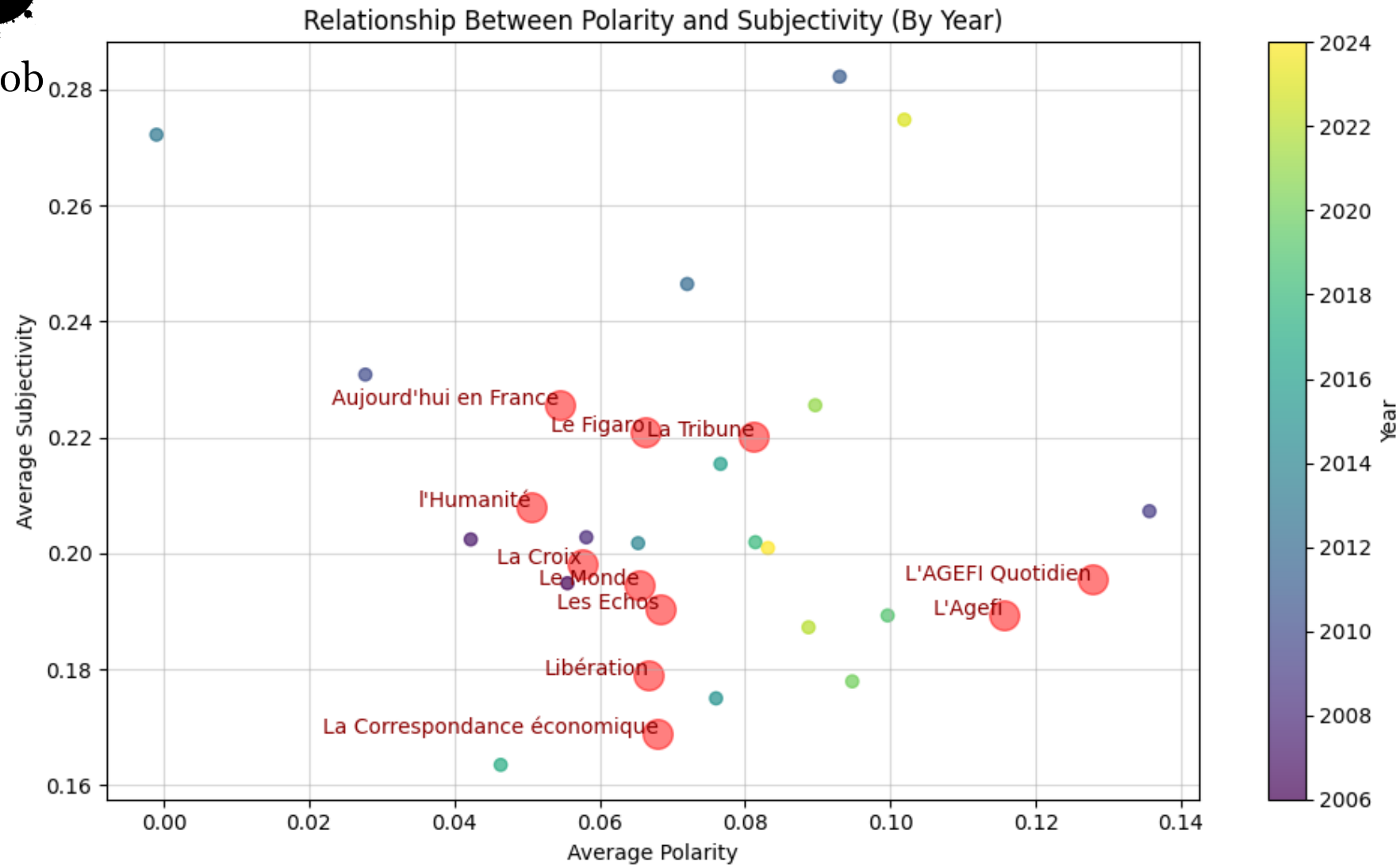


TextBlob





TextBlob





Hugging Face

DistilCamemBERT- Sentiment



CamemBERT



Irria



Hugging Face

DistilCamemBERT- Sentiment



210K reviews

+

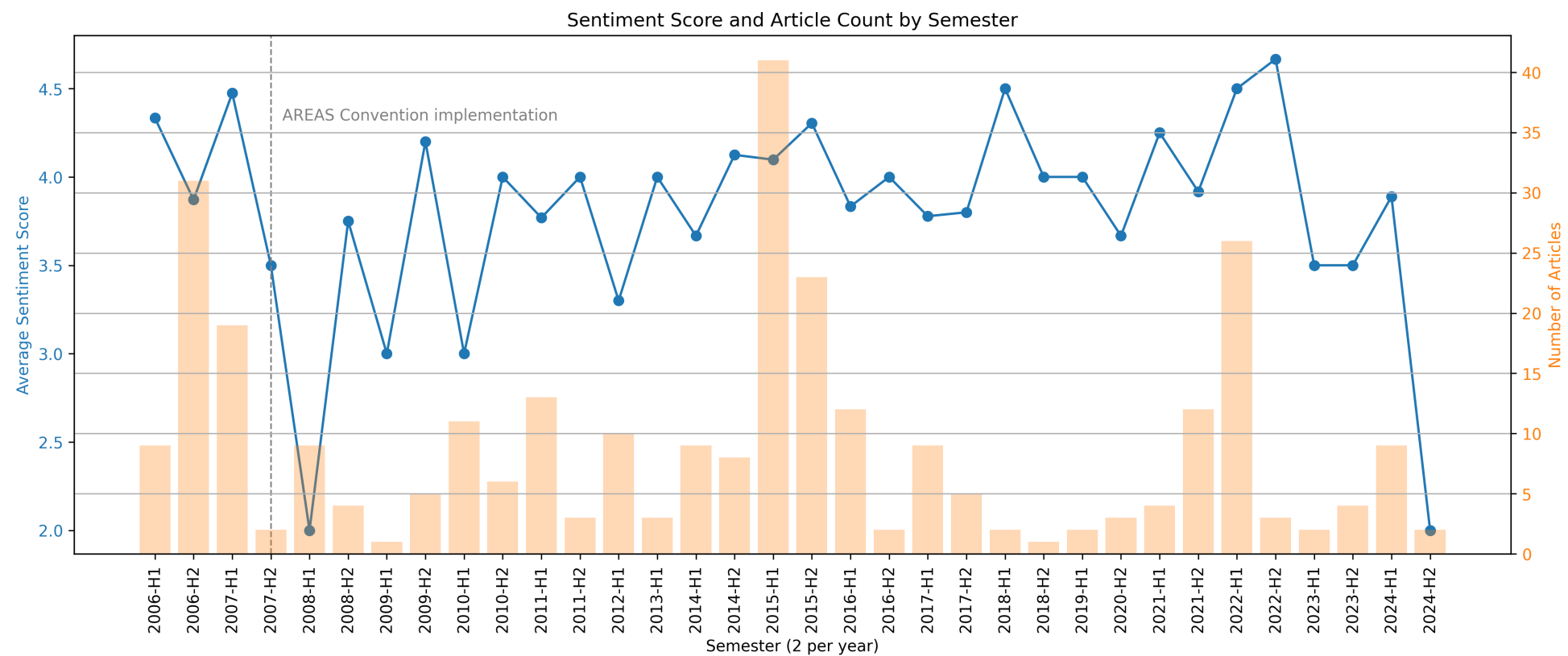


240K critiques



Hugging Face

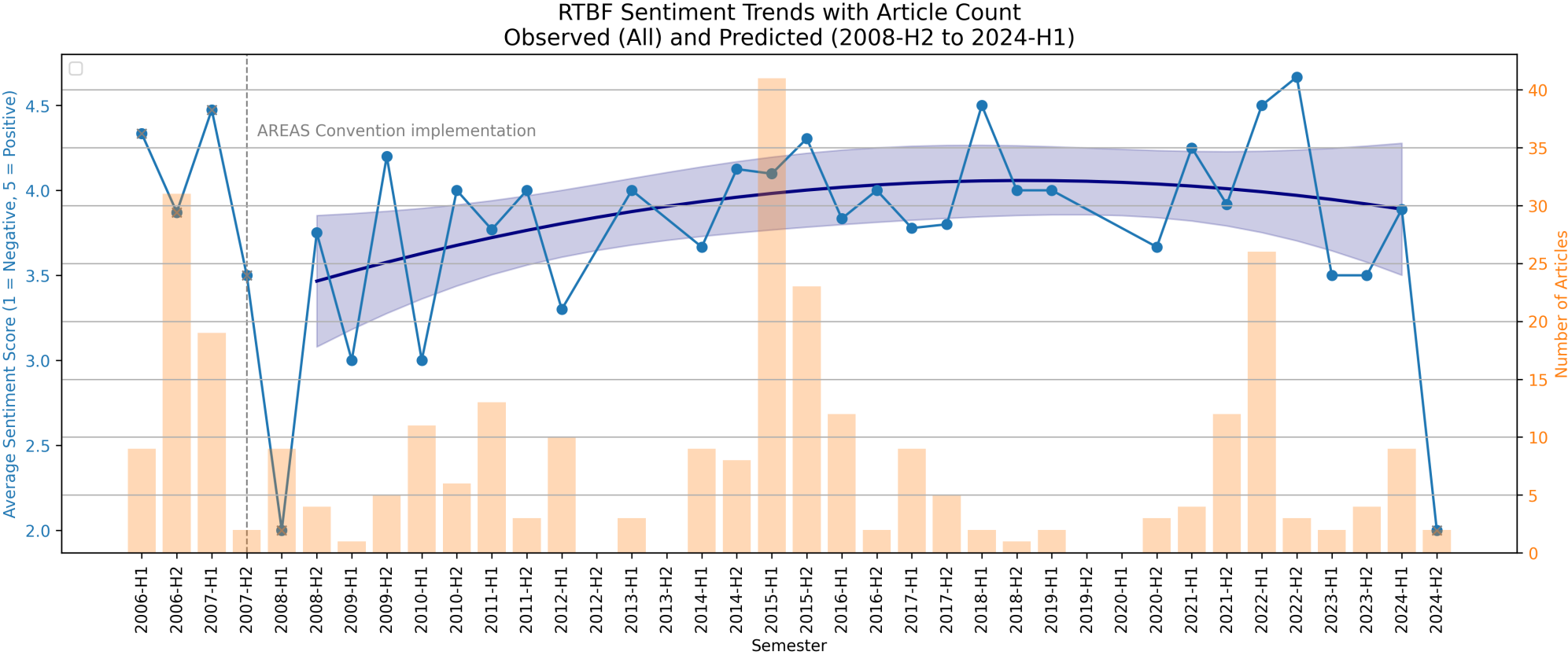
DistilCamemBERT- Sentiment

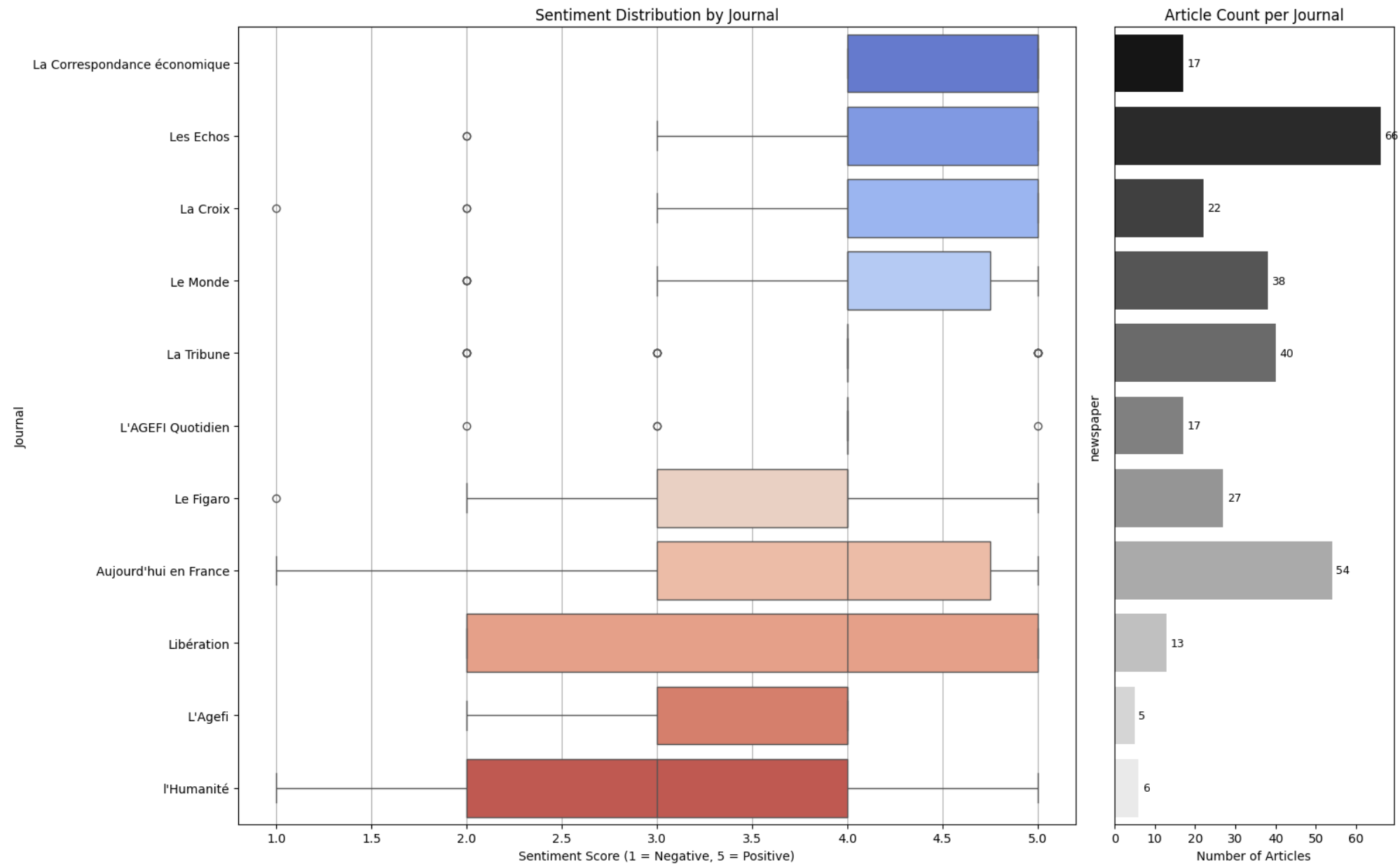




Hugging Face

DistilCamemBERT- Sentiment



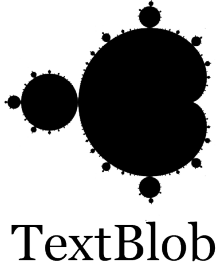




Hugging Face

DistilCamemBERT- Sentiment

Score	newspaper	Title	Explanation
Negative	1 Aujourd'hui en France	<u>on nous traite comme des citoyens de seconde zone</u>	The text appears critical or highlights negative experiences, justifying a low score.
Negative	1 l'Humanité	<u>depuis hier les anciens malades</u>	
Negative	1 Le Figaro	<u>ufcque choisir réclame une loi pour faciliter laccès des malades aux prêts immobiliers</u>	
Negative	1 La Croix	entretien celine lisraoux cofondatrice de rose magazine 1 <u>même guéris ceux qui ont eu un cancer restent des citoyens de seconde zone</u>	
Neutral	3 Le Figaro	<u>le droit à loubli fait ses premiers pas</u>	The content is descriptive without emotional charge, consistent with a neutral sentiment.
Neutral	3 Le Figaro	<u>les malades devraient pouvoir emprunter plus facilement</u>	
Positive	5 Aujourd'hui en France	<u>laccès des malades aux assurances et crédits sera amélioré</u>	The passage mentions improvements or hopeful outcomes, supporting a positive score.
Positive	5 La Croix	<u>les malades accéderont plus facilement aux prêts bancaires</u>	
Positive	5 Le Figaro	<u>accès au crédit facilité pour les malades</u>	



Insight and conclusions

- Mild positive sentiment with moderate subjectivity dominates.
- Outliers indicate opinion-driven neutral content and fact-driven neutral reporting.
- No strong linear relationship between polarity and subjectivity.

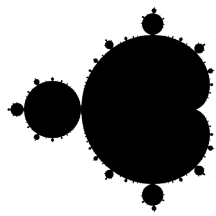


Hugging Face

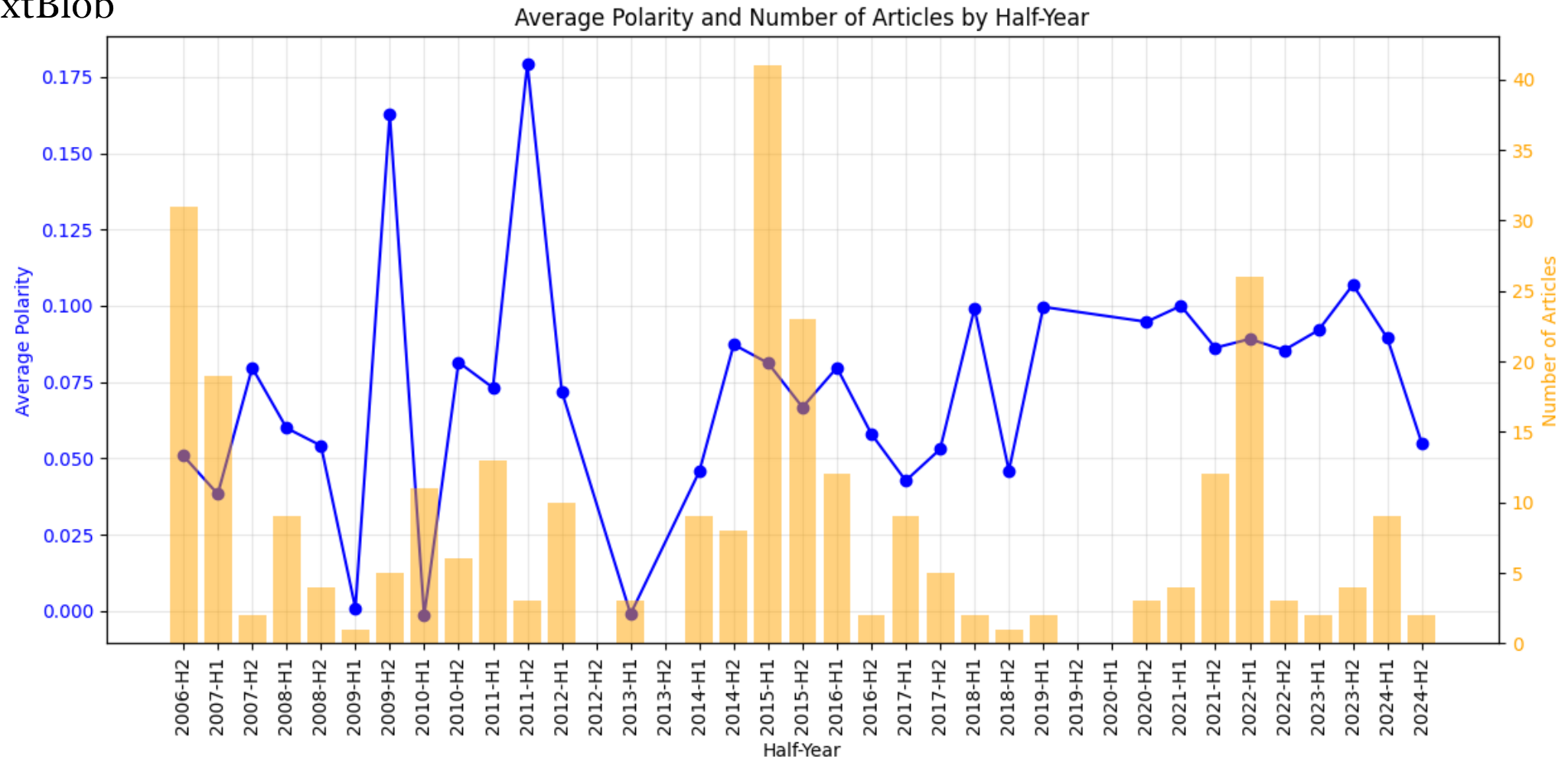
DistilCamemBERT- Sentiment

Insight and conclusions

- Initial high expectations for the AREAS convention were quickly disappointed due to maybe implementation delays / lack efficacy.
- The Lemoine law generated the most favorable sentiment scores.
- Overall, newspaper opinion shows a positive trend, reflecting an optimistic outlook.



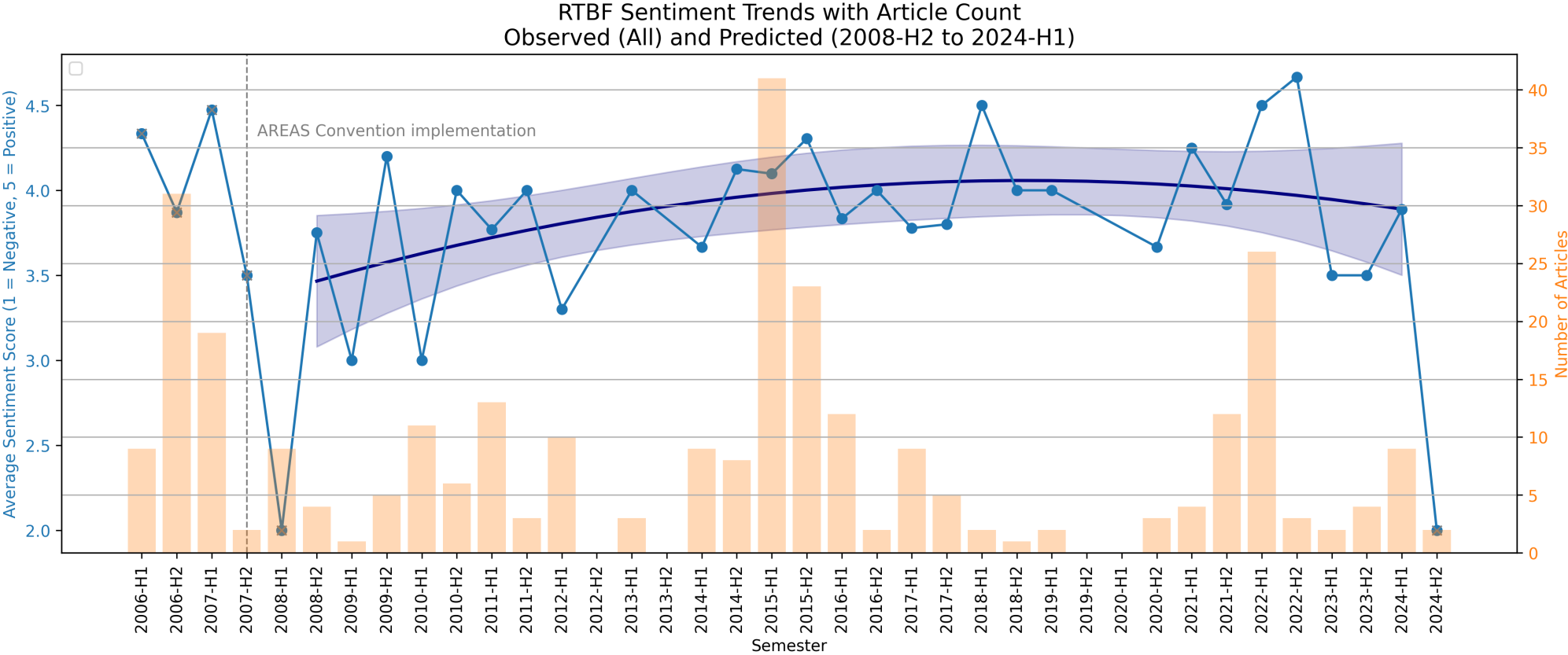
TextBlob

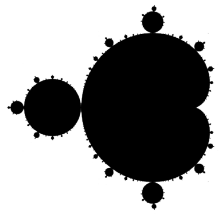




Hugging Face

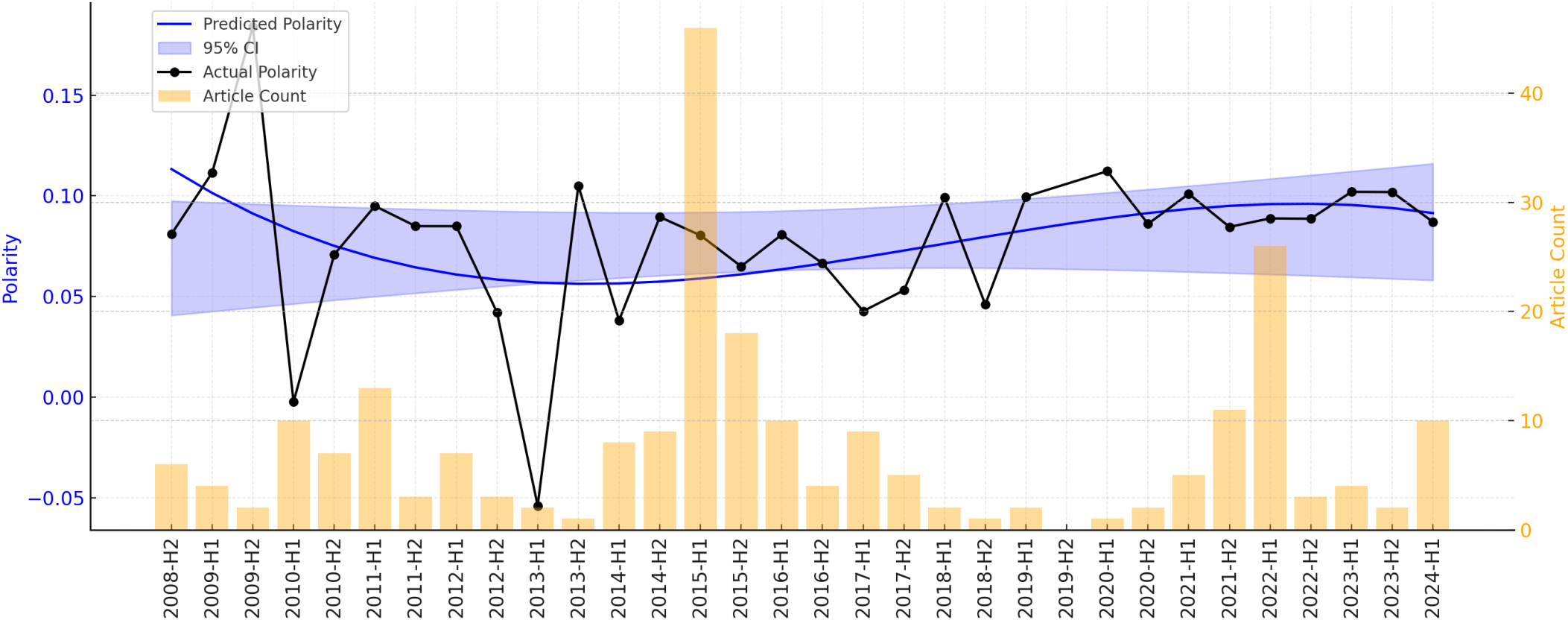
DistilCamemBERT- Sentiment





TextBlob

Polarity Trend with Article Count (2008-H2 to 2024-H1)

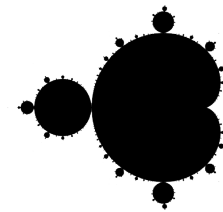




Hugging Face

DistilCamemBERT- Sentiment

Insight and conclusions



TextBlob

- **There is a modest correlation between the two methods in terms of polarity and sentiment scores.**
- **Both approaches consistently indicate a positive trend across most periods.**

Overall conclusion

- AERAS / RTBF / Lemoine law presented as social progress (global positive sentiment score)
- Low polarity (neutral discours) in the economical press
- Low individual embodiment of the feeling of discrimination in the French press (low subjectivity)
 - consolidation of qualitative data (interviews)

Merci à toute l'équipe ELOCAN !

Effets du droit à l'oubli pour l'accès à l'assurance emprunteur après un cancer

Effects of the right to be forgotten in accessing LOan-related insurance after CANcer

Marie-Sophie Desmarets, Gaël Lejeune (CERES)

Renaud Debailly, Hugo Jeanningros (GEMASS)

Corinne Alberti, Yann Benoist, Aurélie Bourmaud

Enora Le Roux, Morgane Michel, Asmaa Janah, Lisa Duconget (ECEVE)

Nadia Haddy, Florent De Vathaire (CESP)

Sandrine Dabakuyo, Emerline Assogba (CGFL)

Guillemette Jacob (Les Seintinelles)

Claude Forget (SiRIC CURAMUS)

Katia Roguin, Mehdi Aslam (LNCC)

Francoise Meunier, Grazia Scocca (ECPC)

<https://elocan.sorbonne-universite.fr/>



Merci pour votre attention !

Merci à toute l'équipe ELOCAN !

Effects of the right to be forgotten in accessing LOan-related insurance after CANcer
<https://elocan.sorbonne-universite.fr/>

Géraldine Cazorla, Mathieu Bourhis (SESSTIM)

Renaud Debailly, Hugo Jeanningros (GEMASS)

Marie-Sophie Desmarets, Gaël Lejeune (CERES)

Corinne Alberti, Yann Benoist, Aurélie Bourmaud

Enora Le Roux, Morgane Michel, Asmaa Janah, Lisa Duconget (ECEVE)

Nadia Haddy, Florent De Vathaire (CESP)

Sandrine Dabakuyo, Emerline Assogba (CGFL)

Guillemette Jacob (Les Seintinelles)

Claude Forget (SiRIC CURAMUS)

Katia Roguin, Mehdi Aslam (LNCC)

Francoise Meunier, Grazia Scocca (ECPC)



Algorithme de Recherche

<i>Nom des lois / convention</i>	<i>Assurance emprunteur</i>	<i>@ Santé</i>
Droit à l'oubli	Prêt(s), emprunt, crédit	Maladie, malade, santé, risque
AERAS	Assurance, assureur, assurer	Aggravé(s), accru(s)
Loi Lemoine, Loi du 28 février 2022	Banque, dépôt, mensualité, hypothèque, dette, débiteur, créancier	Hospitalisation, hôpital, patient, handicap

Champ : Presse nationale Française, 2004-2024

➔ 305 articles

Europresse >>> Europarser

EuroParser

Convertissez des fichiers html
Europresse vers d'autres formats

0

Récupérez vos fichiers html sur Europresse

Pour plus d'informations sur comment les télécharger, consultez le [tutoriel CERES](#)

1

Téléchargez les fichiers que vous souhaitez transformer



Déposez vos fichiers ici
ou cliquez pour les sélectionner

2

Choisissez les formats vers lesquels convertir

Iramuteq

TXM

JSON

CSV

Excel (xls)

Markdown

et/ou les statistiques à générer

Statistiques

Statistiques traitées

Graphiques interactifs

3

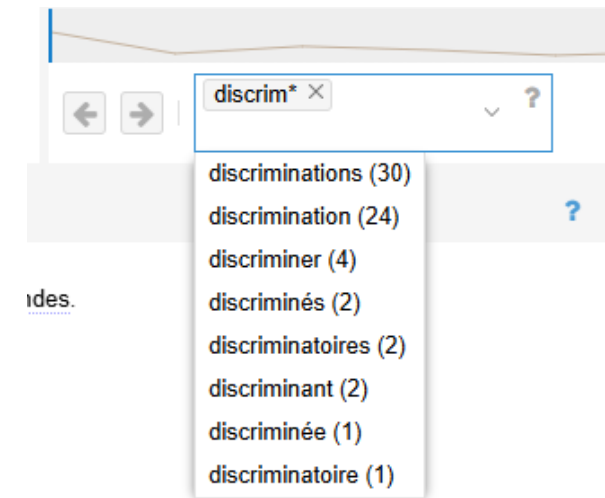
Téléchargez les fichiers transformés sur votre ordinateur



Convertir

Occurrence de mots à retravailler (Iramuteq)

- Progrès (85)
- Faciliter 82
- Avancées 67
- Bénéficiaire 65
- Lutte 58 (LNCC...)
- Problème 57
- Avancée 51
- Peine 45
- Survie 44
- Recherche 41
- Guéris 40
- L'hépatite 38
- Cause 38
- Négociations 37
- Guérisons 37 guéri 34
- CISS
- Sécurité 35
- Prévention 35
- Vih 34
- Exclusions 33
- Longue 32
- Jeune 32
- Améliorer 32
- Sociétés 31
- UF C31
- Discriminations 30
- Interassociatif 29
- Inégalités 29
- Réponse rechute parole 28
- Femmes enfants 28
- Vrai chroniques 26
- Victimes Peur mal facilité 25
- Positif 24
- Solution recours 22
- Mortalité 23
- Prostate 21
- Thyroïde 20
- Science 20
- Diabète 18
- Testicules 17
- Handicapés 17
- Surprise séropositifs 16
- Vaincu salué rémission 15
- Patient ménages 15
- Transparence 14
- Indispensable 14



VOYANT
see through your text



Le projet ELOCAN

Effets du droit à l'oubli pour l'accès à l'assurance emprunteur après un cancer

Effects of the right to be forgotten in accessing LOan-related insurance after CANcer

Un projet pluridisciplinaire et multi-équipes

<https://elocan.sorbonne-universite.fr/>



European Cancer
Patient Coalition



Financé par :