

**CENTRALE  
LYON**

# **MOS 4.4 - Evolution de l'utilisation des méthodes statistiques avancées dans le secteur financier**

---

Antonin ROUFFET

Tuteur : Emmanuel DELLANDREA

# Sommaire

---



- I. Contexte**
- II. Sentiment Analysis**
- III. Méthodologie de veille**

# I. Contexte

# Contexte

Le secteur financier : de multiples acteurs

## Assurances



Allianz 



## Sociétés de gestion de fonds

**Amundi**  
ASSET MANAGEMENT

J.P.Morgan

**BlackRock**

  
CRÉDIT AGRICOLE



BNP PARIBAS

  
Commission de Surveillance  
du Secteur Financier

AMF



## Banques

## Organismes régulateurs

# Contexte

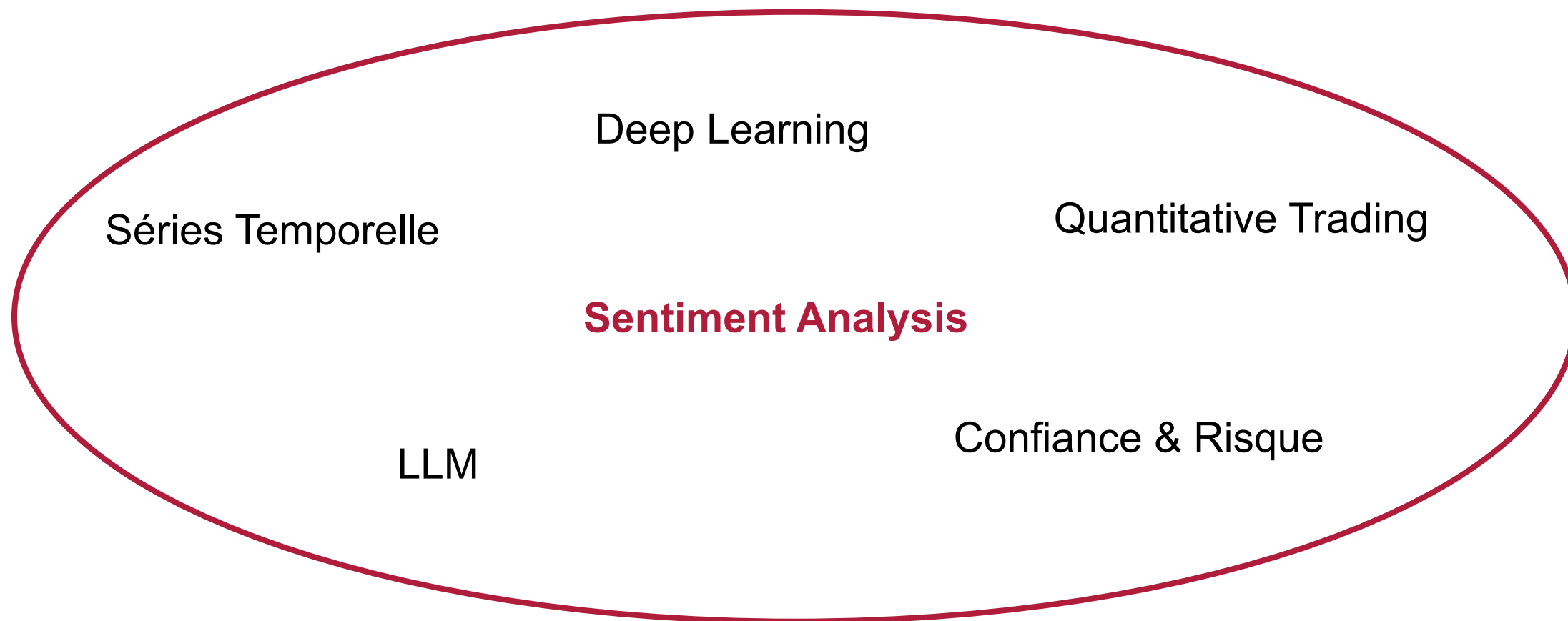
## De nombreux modèles statistiques



<b>Pricing des actifs</b>	Modèle de Black-Scholes, Modèle de Merton, Modèle de Heston
<b>Série temporelle</b>	Modèle ARIMA, Modèle GARCH, LSTM, Transformers
<b>Gestion des risques</b>	Simulations de Monte Carlo, Modèles Bayésien, Régression Logistique, Clustering
<b>Optimisation de trading</b>	Reinforcement Learning, Clustering, Modèle de Fama-French

# Contexte

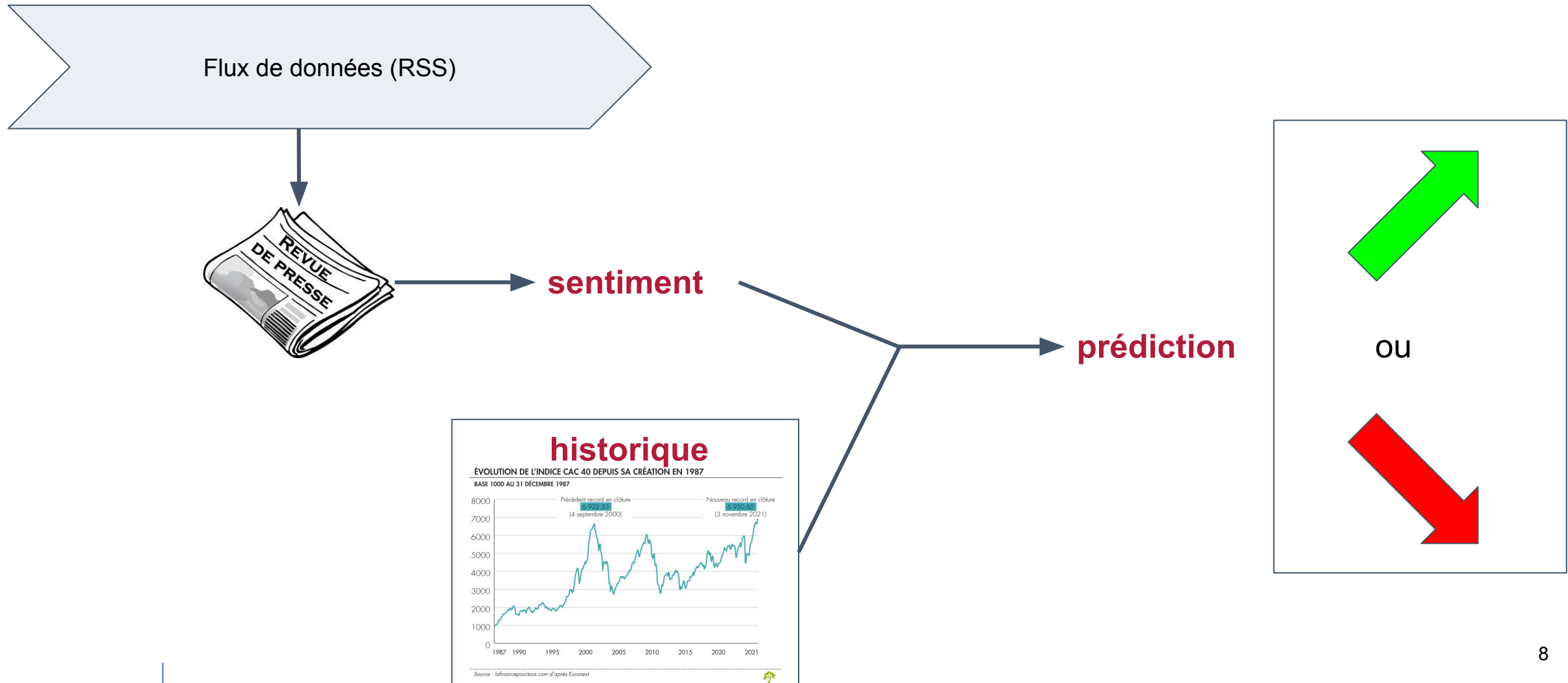
## Périmètre de l'étude de Veille



## **II. Sentiment Analysis**

# Sentiment Analysis

## Principe





# Sentiment Analysis

## VADER (Valence Aware Dictionary and sEntiment Reasoner)

- Dictionnaire de mot noté de -4 à 4
- Calcul un score moyen sur un texte en fonction des mots utilisés

**Sentiment journalier :**

$$Sent_d = \frac{N_d(pos) - N_d(neg)}{N_d(pos) + N_d(neut) + N_d(neg) + 3}$$

**Etude sur la Volatilité :**

$$r_t = \frac{P_t - P_{t-1}}{P_{t-1}} \quad \sigma = \sqrt{\frac{1}{N-1} \sum_{t=1}^N (r_t - \bar{r})^2}$$

# Sentiment Analysis

## VADER sur FTSE100 Index



- Sentiment négatif des titres corrélé avec les rendements du marché le même jour.
- Forte corrélation négative entre sentiment positif des tweets et volatilité du marché le jour suivant.
- Précision de 63 % dans la prédiction de la direction de la volatilité.

# Sentiment Analysis

## Heterogeneous multi-Agent Discussion (HAD)

- **Inspiration** : théorie de l'esprit et des émotions de Marvin Minsky
- 2 types d'agents :
  - **Linguistiques** : axés sur l'humeur, la rhétorique et les références temporelles
  - **Financiers** : simulent les perspectives des investisseurs institutionnels et individuels
- **Processus de discussion entre les agents**

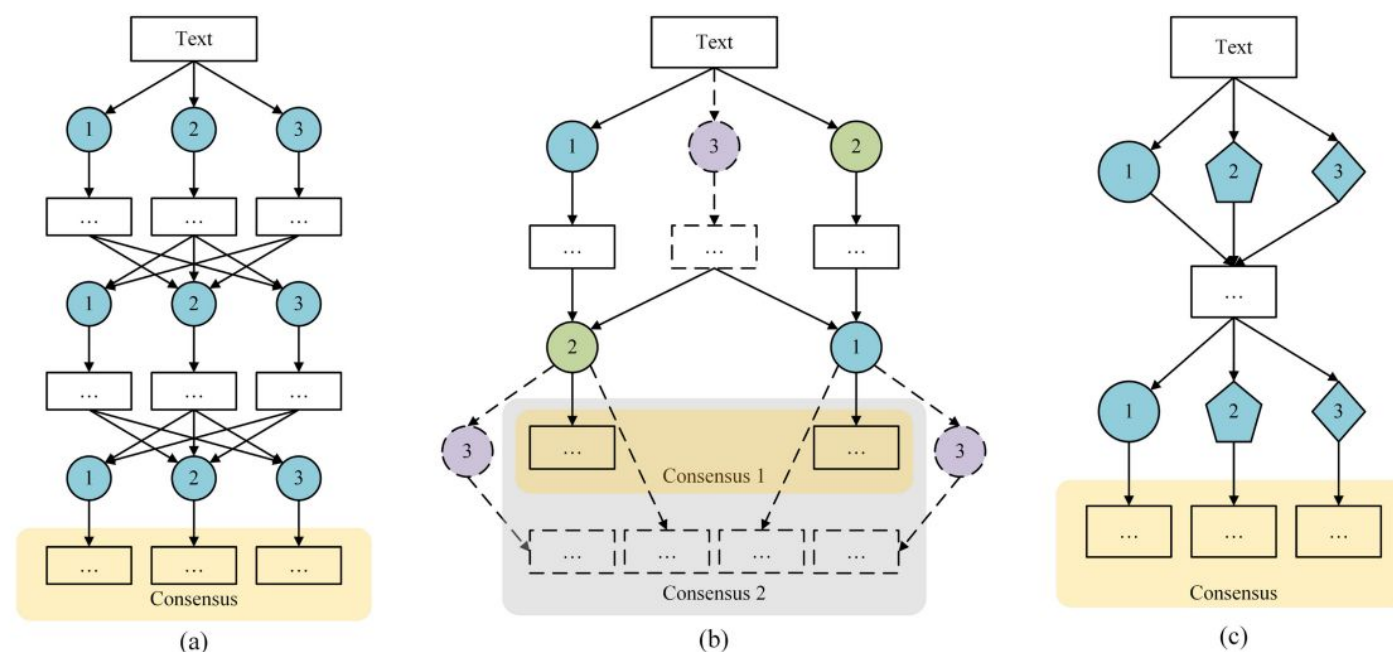


Fig. 1. Different multi-agent LLM frameworks for reaching a consensus: (a) homogeneous multi-agent debate [18], (b) multi-role multi-agent negotiation [42], (c) heterogeneous multi-agent discussion (HAD: the proposed framework). Colors denote different roles and shapes denote heterogeneous agents.

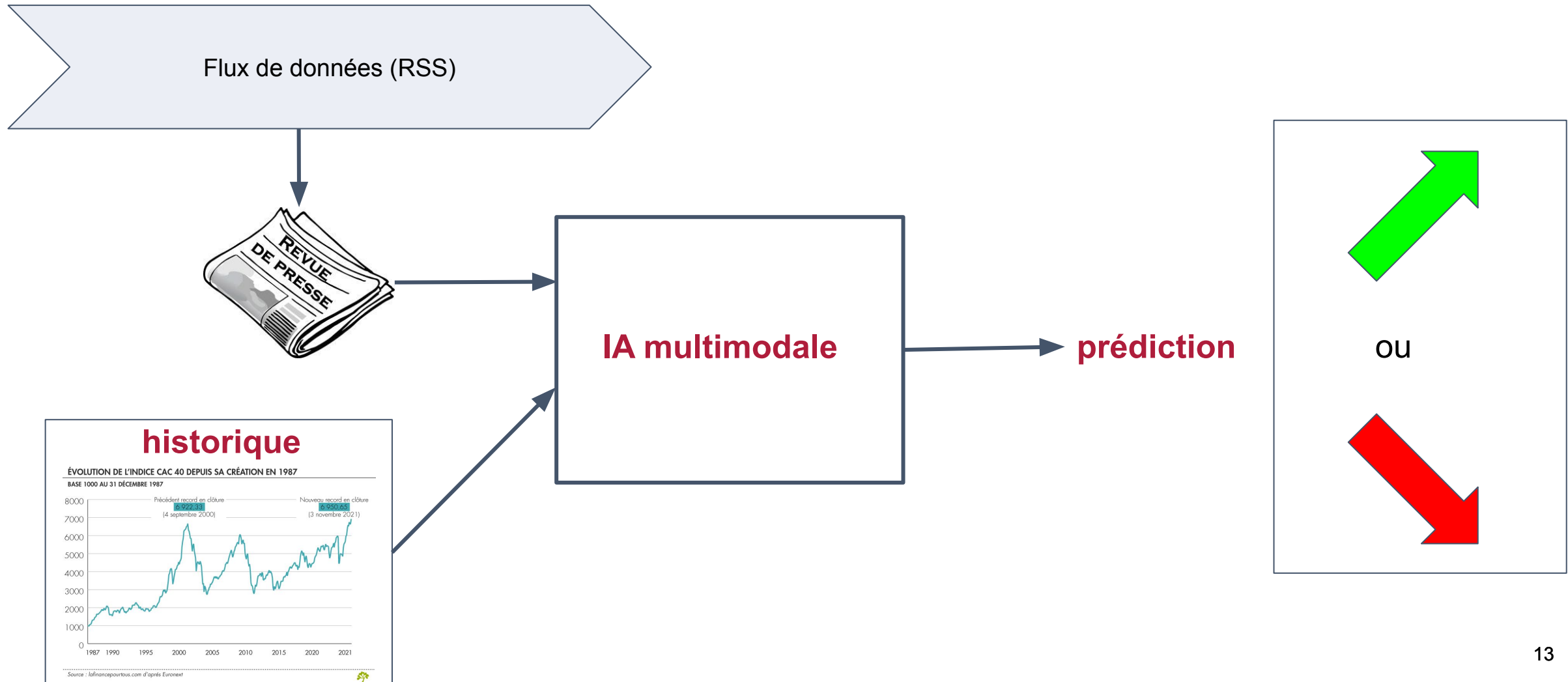
# Sentiment Analysis

## Heterogeneous multi-Agent Discussion (HAD)

Model\Dataset	FPB		StockSen		CMC		FiQA		SEntFiN		FinEntity	
	Acc.	F-1	Acc.	F-1	Acc.	F-1	Acc.	F-1	Acc.	F-1	Acc.	F-1
L & M Dictionary [31]	69.89	54.40	70.70	81.70	–	–	–	–	–	–	–	–
(Fin-)BERT	91.69	89.70	76.90	84.50	93.50	–	–	–	94.29	93.27	–	83.00
BloombergGPT	–	51.07	–	–	–	–	–	75.07	–	–	–	–
BLOOMZ-560m	34.63	32.90	63.65	72.47	87.16	92.62	78.33	83.64	<b>51.32</b>	<b>41.87</b>	32.43	26.90
BLOOMZ-560m (MSV)	29.57	30.13	65.46	74.41	–	–	77.87	82.83	–	–	32.57	27.98
BLOOMZ-560m (MD)	34.25	33.97	68.90	79.46	–	–	<b>78.80</b>	<b>84.21</b>	–	–	32.71	28.37
BLOOMZ-560m (HSV)	33.11	29.01	<b>69.09</b>	<b>83.47</b>	–	–	78.50	83.85	–	–	<b>46.36</b>	22.96
BLOOMZ-560m (HAD)	<b>34.89</b>	<b>38.40</b>	68.06	78.44	87.67	92.95	77.42	83.31	50.16	40.69	34.69	<b>32.93</b>
LlaMa3-70b	45.49	50.00	70.78	71.79	78.25	64.73	84.65	81.44	58.34	52.39	43.21	43.78
LlaMa3-70b (MSV)	42.23	46.87	75.26	<b>84.88</b>	<b>81.47</b>	<b>88.76</b>	85.73	89.38	57.70	50.36	42.75	42.78
LlaMa3-70b (MD)	46.27	51.03	76.35	73.61	77.05	63.00	84.96	84.55	55.36	47.77	45.83	44.69
LlaMa3-70b (HSV)	40.85	40.80	<b>77.67</b>	80.41	79.02	87.39	78.34	84.91	51.73	40.66	40.75	38.22
LlaMa3-70b (HAD)	<b>68.48</b>	<b>84.31</b>	74.25	76.12	71.91	82.06	<b>92.09</b>	<b>94.01</b>	<b>63.51</b>	<b>66.72</b>	<b>60.23</b>	<b>59.36</b>
GPT-3.5	78.58	81.06	67.64	73.93	85.31	91.05	90.53	92.41	67.99	63.21	55.84	56.00
GPT-3.5 (MSV)	72.22	74.68	68.62	76.08	86.31	92.25	87.81	90.29	74.06	64.62	<b>68.46</b>	<b>68.61</b>
GPT-3.5 (MD)	69.83	73.92	69.40	75.72	86.74	91.47	87.95	90.86	69.65	69.30	56.51	57.29
GPT-3.5 (HSV)	65.12	64.95	69.35	76.03	84.04	90.44	87.42	90.05	40.56	64.96	42.37	40.07
GPT-3.5 (HAD)	<b>81.25</b>	<b>87.10</b>	<b>70.01</b>	<b>77.97</b>	<b>90.91</b>	<b>92.69</b>	<b>95.07</b>	<b>96.20</b>	<b>78.16</b>	<b>77.72</b>	61.80	62.56

# Sentiment Analysis

## L'IA multimodale, une solution ?



# III. Méthodologie de veille

# Méthodologie de veille

## Objectifs et déroulé



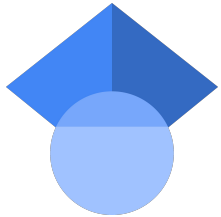
- Identifier les **méthodes et outils récents** en analyse de sentiment
- Détecter les **opportunités d'innovation** pour la prise de décision en investissement

1. **Cadrage de la veille**
2. **Identification des sources d'information**
3. **Collecte et filtrage des données**
4. **Analyse et structuration des informations**
5. **Synthèse et diffusion des résultats**

# Méthodologie de veille

## Mise en place de la veille

### Sources d'informations



R<sup>G</sup>



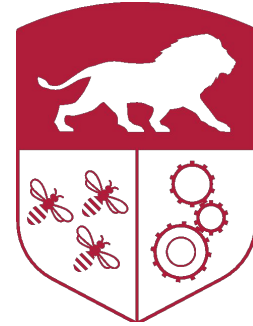
r/SentimentAnalysis  
r/LanguageTechnology  
r/quant  
r/MachineLearning

J.P.Morgan | CHASE 

### Mots-clés

- sentiment analysis
- finance
- Natural Language Processing
- financial text mining
- LSTM stock prediction sentiment





**CENTRALE  
LYON**

---

## **MOS 4.4**

36, avenue Guy de Collongue 69130 Écully  
[www.ec-lyon.fr](http://www.ec-lyon.fr) | [@centralelyon](https://twitter.com/centralelyon)