

## Introduction

**Les termes-clés** sont des locutions qui représentent le mieux un document.

**L'extraction de termes-clés** est utile dans de nombreux domaines :

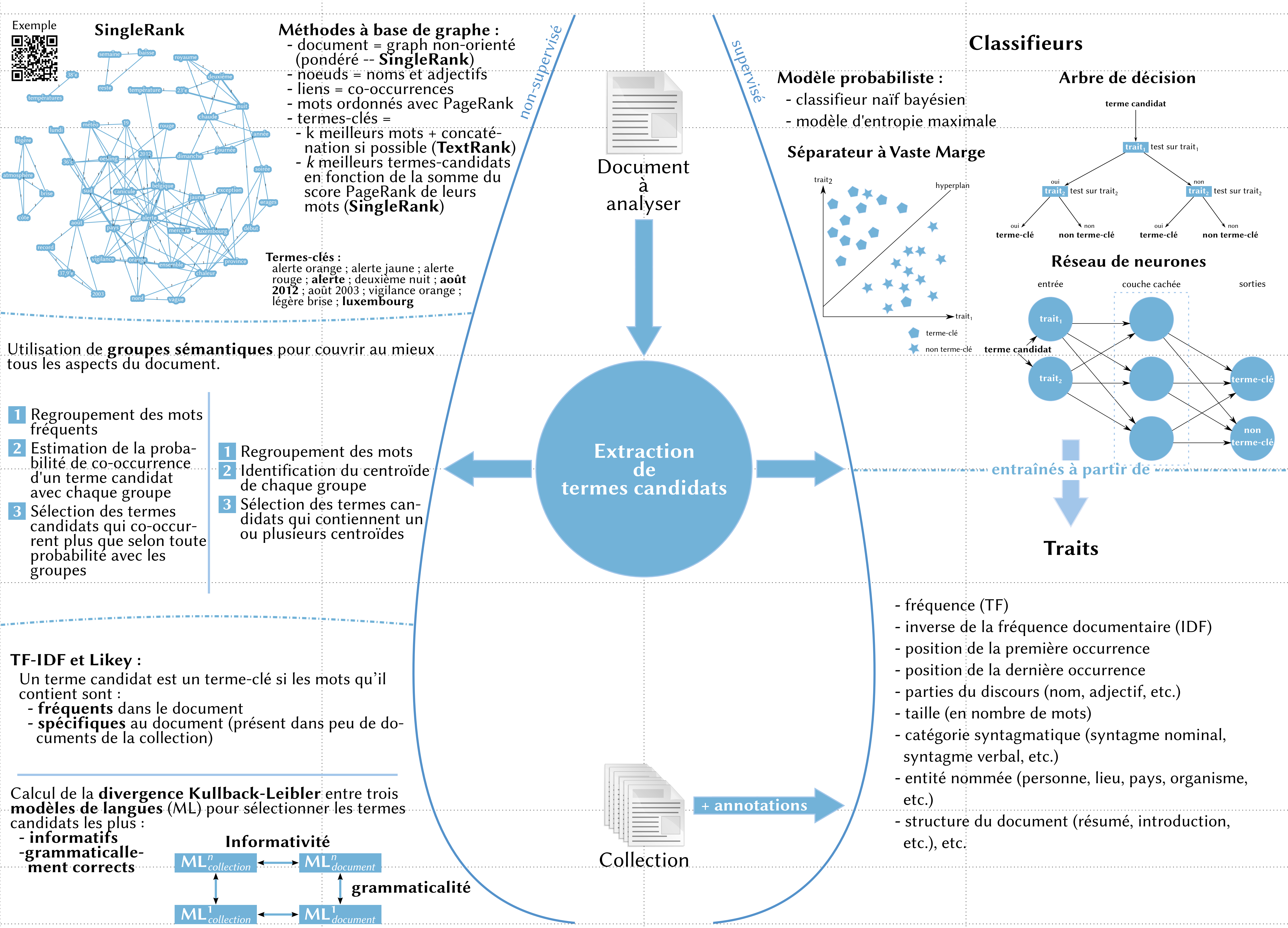
- indexation automatique
- résumé automatique
- classification de documents

## Données

**Des collections de documents** annotés en termes-clés sont utilisées pour :

- effectuer de l'apprentissage supervisé
- **ou** évaluer les méthodes d'extraction « à la Cranfield »

**Des collections de documents non-annotés** sont utilisées pour effectuer de l'apprentissage non-supervisé.



## Conclusion

De nombreuses méthodes.

→ Quelques points communs :

- pré-traitement
- extraction des termes candidats

→ Diverses approches :

- usage de groupes sémantiques
- usage d'un graphe
- entraînement d'un classifieurs, etc.

## Références

- [1] Jones K.S. : A Statistical Interpretation of Term Specificity and its Application in Retrieval.
- [2] Liu Z., Li P., Zheng Y. et Sun M. : Clustering to Find Exemplar Terms for Keyphrase Extraction.
- [3] Matsuo Y. et Ishizuka M. : Keyword Extraction from a Single Document Using Word Co-occurrence Statistical Information.
- [4] Mihalcea R. et Tarau P. : TextRank : Bringing Order Into Texts.
- [5] Nguyen T.D et Kan M.Y. : Keyphrase Extraction in Scientific Publications.
- [6] Paukkeri M.S et Honkela T. : Likey : Unsupervised Language-Independent Keyphrase Extraction.
- [7] Sarkar K., Nasipuri M. et Ghose S. : A New Approach to Keyphrase Extraction Using Neural Networks.
- [8] Sujian L., Houfeng W., Shiwen Y. et Chengsheng X. : News-Oriented Keyword Indexing with Maximum Entropy Principle.
- [9] Turney P.D. : Learning Algorithms for Keyphrase Extraction.
- [10] Wan X. et Xiao J. : Single Document Keyphrase Extraction Using Neighborhood Knowledge.
- [11] Witten I.H., Paynter G.W., Frank E., Gutwin C. et Nevill-Manning C.G. : KEA : Practical Automatic Keyphrase Extraction.