

PRÉSENTATION

Aujourd'hui, les outils informatiques font partie intégrante de l'analyse littéraire. Qu'il s'agisse de numérisation de corpus, d'analyses statistiques, ou encore de visualisation de données, ils permettent une nouvelle approche des textes, objective et exhaustive, ainsi que le perfectionnement d'études plus traditionnelles.

Les phrases de Marcel Proust sont-elles réellement plus longues que celles de ses contemporains ? Quels sont les prénoms que Balzac donne le plus souvent à ses personnages ? Comment peut-on caractériser le style de Pascal Quignard ? Quelles sont les couleurs dominantes dans Les Liaisons dangereuses ? Flaubert a-t-il un vocabulaire plus riche que celui de Malraux ? C'est le plus souvent en utilisant les textes sous leur forme numérique et des outils informatiques que l'on pourra répondre à ce type de questions, relevant d'hypothèses spontanées ou de la recherche scientifique. L'objet de cet ouvrage est de présenter les instruments, les ressources et les méthodes qui permettent de constituer un corpus de textes littéraires en vue de son étude informatisée.

Nous proposons ici d'exposer les différentes étapes qui permettent d'analyser un texte avec des outils numériques. Comment récupérer, sous forme numérique, les textes composant un corpus ? Comment préparer un texte ? Quelles corrections ? Quel(s) balisage(s) pour quel(s) logiciel(s) ? En quoi un logiciel de textométrie peut-il être utile ? Quels sont les logiciels de textométrie disponibles ? Ou encore, comment rendre lisibles et accessibles les résultats ainsi obtenus ? Quelles sont les meilleures représentations graphiques ? Quels outils pour mettre en ligne et diffuser les résultats ?

Autant de questions auxquelles cet ouvrage se propose de répondre.

Michel Bernard et Baptiste Bohet sont enseignants-chercheurs à l'Université Paris 3 – Sorbonne Nouvelle, où ils assurent des enseignements d'informatique, de culture numérique et de littérature. Ils animent notamment le centre de recherche Hubert de Phalèse, dont l'objectif est de faciliter l'accès des technologies numériques aux littéraires, à tous les niveaux du système éducatif et de la recherche.

TABLE DES MATIÈRES

SOMMAIRE

INTRODUCTION

CHAPITRE 1 - COMMENT CONSTITUER VOTRE CORPUS NUMÉRIQUE ?

Que signifie numériser un texte ?
Quelle(s) précaution(s) juridique(s) prendre ?
Comment obtenir des textes ?
Dactylographier le texte
Scanner le texte
Télécharger le texte

CHAPITRE 2 - QUELS LOGICIELS POUR ANALYSER MON CORPUS ?

Qu'est-ce que la textométrie, la lexicométrie, la logométrie ?
Que faire avec un traitement de texte ou un navigateur web ?
Comment utiliser une base de données textuelles en ligne ?
Quels logiciels spécialisés choisir ?

CHAPITRE 3 - COMMENT PRÉPARER VOTRE TEXTE POUR L'ANALYSER AVEC DES LOGICIELS DE TEXTOMÉTRIE ?

Comment corriger vos textes ?
Pourquoi baliser vos textes ?
Codages propres à certains logiciels
Un codage standard : la TEI
Le codage morpho-syntaxique
Automatiser le travail avec des macros

PARUTION LE 15 MAI



FORMAT 16X24
128 PAGES 10,50€
ISBN 978-2-87854-685-9

Michel Bernard - Baptiste Bohet

LITTÉROMÉTRIE

Outils numériques pour l'analyse des textes littéraires



Les fondamentaux
DE LA SORBONNE NOUVELLE

CHAPITRE 4 - QUELLES SONT LES FONCTIONNALITÉS DES LOGICIELS DE TEXTOMÉTRIE ?

Quelles unités pour la textométrie ?
Quelles fonctionnalités de recherche ?
Index
Recherche d'occurrences
Expressions régulières
Recherche de listes d'occurrences
Recherche de cooccurrences
Recherche de segments répétés
Quelles fonctionnalités statistiques ?
Principales caractéristiques lexicométriques
Spécificités
Cartographie des corpus et des occurrences
Analyse factorielle des correspondances (AFC)
Représentations arborées
Corpus alignés

CHAPITRE 5 - QUELLES SONT LES REPRÉSENTATIONS GRAPHIQUES LES PLUS PERTINENTES ?

Quels principes de base pour la représentation graphique ?
Comment aider à la lecture tabulaire ?
Quelle(s) représentation(s) pour faire apparaître quel(s) phénomène(s) ?

CHAPITRE 6 - COMMENT METTRE EN LIGNE, ET ORGANISER LA DIFFUSION DU CORPUS ET DES RÉSULTATS ?

Quel modèle pour les corpus en ligne ?
Comment améliorer la diffusion des résultats ?

CONCLUSION

Bibliographie
Les fondateurs
Ouvrages généraux sur la textométrie
Richesse lexicale
Représentations graphiques
Édition numérique et droits d'auteur
Études sur des corpus particuliers