



Bulletin de méthodologie sociologique

Bulletin of sociological methodology

85 | 2005
January

Des logiciels d'analyse textuelle au service de l'imagination sociologique

Didier Demazière



Édition électronique

URL : <http://journals.openedition.org/bms/978>

ISSN : 2070-2779

Éditeur

Association internationale de méthodologie sociologique

Édition imprimée

Date de publication : 1 janvier 2005

Pagination : 5-9

ISSN : 0759-1063

Référence électronique

Didier Demazière, « Des logiciels d'analyse textuelle au service de l'imagination sociologique », *Bulletin de méthodologie sociologique* [En ligne], 85 | 2005, mis en ligne le 08 juillet 2008, consulté le 01 mai 2019. URL : <http://journals.openedition.org/bms/978>

Ce document a été généré automatiquement le 1 mai 2019.

© BMS

Des logiciels d'analyse textuelle au service de l'imagination sociologique

Didier Demazière

- 1 Les sociologues qui souhaitent étudier d'importants corpus de textes ont aujourd'hui la possibilité d'utiliser plusieurs outils informatiques. Dans un contexte de développement et de diversification de l'offre logicielle, il n'est pas toujours aisé d'appréhender les caractéristiques des logiciels d'analyse textuelle et de faire un choix raisonné. C'est-à-dire un choix pertinent par rapport aux problématiques sociologiques mobilisées, aux points de vue théoriques privilégiés, aux interrogations de recherche formulées.
- 2 Une première méthode pour tenter d'éclairer la lanterne du sociologue pourrait consister à expliciter les caractéristiques de chaque logiciel : quelles opérations ils réalisent, quelles propriétés des corpus textuels leur sont les mieux adaptées, quelles sont les caractéristiques saillantes des résultats analytiques produits, quelles publications emblématiques sont issues de leur usage, etc. ? Cette piste pourrait déboucher sur la mise au point de fiches pratiques qui complèteraient utilement les documentations techniques en offrant des mémentos synthétiques, standardisés et comparatifs.
- 3 C'est une autre voie que le bureau du « Réseau thématique Méthodes » de l'Association Française de Sociologie a exploré en proposant ce dossier. Cette initiative s'appuie sur une idée directrice : la compréhension de ce que font les logiciels d'analyse textuelle est étroitement liée à leur mise en action sur des corpus. Autrement dit, la perspective consistant à comparer des logiciels comporte le risque de les réduire à des produits, analogues à des biens de consommation, affectés chacun de certaines utilités, propres à produire certains résultats. Autrement dit des produits dont les usages guideraient les utilisateurs, des outils qui en quelque sorte commanderaient le sociologue. Or si ces méthodes logicielles incorporent des procédures cadrées et spécifiques de traitement des matériaux, elles servent aussi des objectifs de connaissance variés et définis en amont par le sociologue. Si elles font des choses différentes inscrites dans leurs algorithmes, elles s'articulent à des approches sociologiques contrastées et élaborées avec d'autres outils, qu'on appelle habituellement théoriques, conceptuels, analytiques, etc.

- 4 Aussi nous avons opté pour un examen des logiciels en usage, c'est-à-dire à travers leur mise en œuvre dans l'étude d'un ensemble de textes. De nombreux exemples de telles études de cas sont disponibles, mais il est difficile de les rapprocher car chaque cas se singularise par des matériaux spécifiques. Aussi la formule adoptée ici consiste à étudier un même corpus en recourant à plusieurs méthodes logicielles d'analyse textuelle. Le corpus choisi est constitué de l'ensemble des résumés des communications proposées au premier congrès de l'Association Française de Sociologie, soit plus d'un millier de textes, répartis dans plus de quarante réseaux thématiques. Les logiciels utilisés sont les suivants : *Alceste*, *Calliope*, *Prospéro*, et *Trideux*. Toutes ces méthodes ont déjà fait l'objet de présentations détaillées dans le *Bulletin de Méthodologie Sociologique*. *Alceste* a été présenté par son auteur, Max Reinert, à deux reprises (Reinert 1987 ; 1990). *Calliope* a été présenté dans un article récent (Van Meter, Cibois et de Saint Léger, 2004). Il s'agit de la version Windows, mise au point par Mathilde de Saint Léger, de Leximappe-Lexinet, qui avait été précédemment présenté à deux reprises (Van Meter, Turner et Bizard, 1995 ; Van Meter et Turner, 1997). *Prospéro*, et sa version interactive, *Marlowe*, ont fait l'objet d'un dossier thématique publié dans le numéro 79 du *BMS* (Chateauraynaud, Reber et van Meter, 2003). *Trideux* a été présenté par Philippe Cibois, son concepteur (Cibois 1984 ; 1995). On peut aussi se reporter à un article de Jacques Jenny qui propose une classification des méthodes formalisées d'analyse de contenu et de discours utilisées dans la recherche sociologique française contemporaine et explicite les élaborations théoriques et méthodologiques au principe des logiciels disponibles (Jenny, 1997).
- 5 Le corpus choisi est particulier, est assez différent des corpus habituellement exploités avec tel ou tel de ces logiciels : il ne s'agit pas d'entretiens de recherche, d'un dossier de presse, d'un ensemble de documents d'archive, ou de toute autre composition de matériaux rassemblés ou produits dans le cadre d'une enquête, et donc constitué en fonction d'un questionnement sociologique. Ce corpus, improbable en quelque sorte, présente pourtant certains avantages. D'abord son caractère extravagant pousse à l'explicitation des modalités, conditions et contraintes associées à son analyse par tel ou tel logiciel ; il fait reculer les frontières des implicites qui accompagnent souvent les utilisations des logiciels, et qui sont justifiés parfois par l'existence de documents de présentation destinés à clarifier, en toute généralité, ces points. Sur un registre tout à fait différent, ce corpus constitue une trace de la production sociologique, essentiellement française, contemporaine. Certes cette trace ne peut être considérée comme représentative des productions de l'ensemble des sociologues, même si les thématiques couvertes sont très diversifiées et le nombre de contributeurs très élevé. Elle n'est pourtant pas sans signification par rapport aux intérêts de recherche, aux thèmes d'investigation, aux cadres analytiques qui parcourent et travaillent la discipline dans la période actuelle. Enfin, ce corpus est aussi une cristallisation d'un moment singulier de l'histoire de la sociologie en France, correspondant au premier congrès de la toute jeune Association Française de Sociologie.
- 6 Aussi, de manière plus latérale et conjoncturelle, le choix de ce corpus peut fournir un attrait supplémentaire à notre entreprise orientée vers la structuration et l'approfondissement de débats autour des méthodes logicielles d'analyse textuelle. Car, à partir du moment où les analyses proposées esquissent, à partir du même corpus, des descriptions plurielles de la production sociologique, elles peuvent rencontrer l'intérêt et la curiosité bien au-delà du cercle des chercheurs qui éprouvent quelque penchant pour ces méthodes, et peuvent, nous l'espérons, concerner tous les sociologues intéressés par

la sociologie de leur discipline. Nous rejoignons alors la problématique qui est au cœur de notre entreprise : il ne s'agit pas de comparer terme à terme des logiciels en les instrumentant comme des produits de consommation offerts aux sociologues. Car le risque serait alors important de fabriquer une fiction d'avancées méthodologiques, quand les usages mécaniques de ces logiciels conduiraient au contraire à un enfermement sur les techniques et à une régression de la discipline. Encore une fois, il s'agit de comparer des approches sociologiques différenciées d'un même corpus qui mobilisent des méthodes elles aussi différenciées ; il s'agit de montrer que les outils logiciels peuvent servir des lectures variées, et d'indiquer ainsi la voie à des usages raisonnés et fructueux.

- 7 Le dossier est constitué de trois articles.
 - Celui de Patrick Trabal (*Le logiciel Prospéro à l'épreuve d'un corpus de résumés sociologiques*) explicite les opérations de codage et d'interprétation nécessaires à la production d'une description des textes, en utilisant le logiciel Prospéro, et considère les résumés comme des réponses à des appels à communications, qui sont alors intégrés dans le corpus, ce qui conduit à examiner la proximité des appels et des réponses, mais aussi les singularités des réseaux thématiques.
 - Celui de Mathilde de Saint-Léger et Karl M. van Meter (*Cartographie du premier congrès de l'AFS avec la méthode des mots associés*) procède par une classification hiérarchique ascendante appliquée aux unités de texte codées à partir d'un dictionnaire de mots-clefs, réalisée avec le logiciel Calliope, ce qui amène à mesurer les ressemblances entre ces unités et à construire des classes, caractérisées par leur densité interne et par leurs liens avec d'autres classes.
 - Celui de Gérard Boudesseul (*De quoi parlent les sociologues réunis en congrès ? Eléments de complémentarité entre une analyse lexicale ouverte et le cumul de variables fermées*) est centré sur une analyse lexicale, conduite en utilisant les logiciels Alceste et Trideux, qui, ensuite, associe les mots employés à quelques attributs des auteurs, tels que le sexe, le lieu d'origine, et l'institution d'appartenance.
- 8 Ces trois articles correspondent à un premier état de l'analyse du corpus des résumés des communications présentées au premier congrès de l'AFS. Ils sont destinés à alimenter une journée d'étude organisée le 25 février 2005 par le Réseau thématique Méthodes de l'AFS autour de la question de « l'apport des logiciels d'analyse textuelle pour décrire les actes d'un congrès de sociologues ». Les présentations de ces travaux, les débats qu'ils susciteront, et les prolongements qui s'en suivront, donneront lieu, nous l'espérons, à d'autres publications destinées à favoriser chez les sociologues les usages des logiciels d'analyse textuelle, et, plus profondément sans doute, à diffuser un certain type de posture vis-à-vis de ces outils logiciels. C'est-à-dire une posture qui ne les réduit pas à des instruments d'objectivation et d'administration de la preuve dont les résultats s'imposeraient au chercheur, mais une posture qui les considère comme des ressources mobilisables, parmi d'autres, pour nourrir des interrogations sociologiques, tester des lectures interprétatives, éprouver des significations provisoires, enrichir des théorisations vivantes, et, finalement, stimuler l'imagination sociologique.

BIBLIOGRAPHIE

- Francis Chateauraynaud, Bernard Reber et Karl Van Meter (2003), Marlowe, Prospero et la technologie littéraire, *Bulletin de Méthodologie Sociologique*, 79, juillet, pp. 5-46.
- Philippe Cibois (1984), Méthodes post-factorielles pour le dépouillement d'enquêtes, *Bulletin de Méthodologie Sociologique*, 1, octobre, pp. 41-78.
- Philippe Cibois (1995), Trideux version 2.2, *Bulletin de Méthodologie Sociologique*, 46, mars, pp. 119-124.
- Jacques Jenny (1997), Méthodes et pratiques formalisées d'analyse de contenu et de discours dans la recherche sociologique française contemporaine. Etat des lieux et essai de classification, *Bulletin de Méthodologie Sociologique*, 54, mars, pp. 64-112.
- Max Reinert (1987), Classification descendante hiérarchique et analyse lexicale par contexte : Application au corpus des poésies d'A. Rimbaud, *Bulletin de Méthodologie Sociologique*, 13, janvier, pp. 53-90.
- Max Reinert (1990), Alceste - Une méthodologie d'analyse des données textuelles et une application: Aurélia de G. de Nerval, *Bulletin de Méthodologie Sociologique*, 26, mars, pp. 24-54.
- Karl M. van Meter, William A. Turner et Jean-Baptiste Bizard (1995), Cognitive Mapping of AIDS Research 1980-1990 - Strategic Diagrams, Evolution of the Discipline and Data Base Navigation, *Bulletin de Méthodologie Sociologique*, 46, mars, pp. 30-44.
- Karl M. van Meter et William A. Turner (1997), Representation and Confrontation of Three Types of Longitudinal Network Data from the Same Data Base of Sociological AIDS Research, *Bulletin de Méthodologie Sociologique*, 56, septembre, pp. 32-49.
- Karl M. van Meter, Philippe Cibois et Mathilde de Saint Léger (2004), Correspondence and Co-Word Analysis of Ten Years of BMS Articles 1993-2003, *Bulletin de Méthodologie Sociologique*, 81, janvier, pp. 48-57.

RÉSUMÉS

Text-Analysis Programs Serving Sociological Innovation: Beginning with the idea that what text analysis programs do is closely associated with their manner of application to a corpus, the Methods Thematic Network of the French Sociological Association (AFS) organized the analysis of the same corpus by different programs. This comparative perspective allows us to show that these programs are not simple instruments for generating proofs, but, on the contrary, constitute applicable resources – among others – for testing interpretations, for enriching theories, and thus for encouraging sociological innovation.

Partant de l'idée que la compréhension de ce que font les logiciels d'analyse textuelle est étroitement liée à leur mise en action sur des corpus, le bureau du Réseau Thématique Méthodes de l'Association Française de Sociologie a suscité des traitements d'un même corpus avec des logiciels différents. Cette mise en perspective permet de montrer que les logiciels ne sont pas de simples instruments d'administration de la preuve, mais constituent des ressources mobilisables

- parmi d'autres - pour tester des lectures interprétatives, enrichir des théorisations provisoires, et ainsi stimuler l'imagination sociologique.

INDEX

Mots-clés : Analyse textuelle, Méthodes logicielles, Prospéro, Calliope

Keywords : Text Analysis, Computer Programs, Trideux, Alceste

AUTEUR

DIDIER DEMAZIÈRE

Animateur du Réseau Méthodes de l'AFS ; CNRS, Laboratoire Printemps ;
didier.demaziere@printemps.uvsq.fr