



Bulletin de méthodologie sociologique

Bulletin of sociological methodology

100 | 2008
October

Un quart de siècle de méthodologie

L'émergence du longitudinal, des réseaux et de la statistique textuelle

Gérard Boudesseul



Édition électronique

URL : <http://journals.openedition.org/bms/3193>

ISSN : 2070-2779

Éditeur

Association internationale de méthodologie sociologique

Édition imprimée

Date de publication : 1 octobre 2008

Pagination : 34-41

ISSN : 0759-1063

Référence électronique

Gérard Boudesseul, « Un quart de siècle de méthodologie », *Bulletin de méthodologie sociologique* [En ligne], 100 | 2008, mis en ligne le 01 octobre 2011, consulté le 01 mai 2019. URL : <http://journals.openedition.org/bms/3193>

Ce document a été généré automatiquement le 1 mai 2019.

© BMS

Un quart de siècle de méthodologie

L'émergence du longitudinal, des réseaux et de la statistique textuelle

Gérard Boudesseul

- 1 La diversité des sujets abordés dans le BMS, l'origine internationale des contributions et la longue durée, un quart de siècle, offrent un bon plan d'observation de la réflexion méthodologique en sociologie.
- 2 À titre exploratoire sont traités ici les titres et les résumés en français de 318 articles publiés dans le *Bulletin de Méthodologie Sociologique (BMS)* de 1983 à 2008, ajoutés des mots-clés propres à chaque résumé. Suivant une procédure déjà expérimentée antérieurement et avec les précautions d'usage (Boudesseul, 2005, 2006a et 2006b), ce corpus se prête à une analyse textuelle par double classification hiérarchique, aboutissant à des « mondes lexicaux ». Ce résultat demande à être recoupé par d'autres traitements, dont une analyse factorielle de formes lexicales extraites du corpus.
- 3 Dans une première partie la méthode de classification descendante hiérarchique avec le logiciel Alceste produit quelques résultats qui peuvent être vus sous trois angles, selon que sont privilégiés les grandes oppositions entre classes lexicales, ou, à l'opposé ce qui semble unir chacun de ces mondes lexicaux tels qu'ils sont livrés par une classification ascendante hiérarchique à l'intérieur de chacune des classes précédentes. Enfin, ces mondes lexicaux peuvent être mis en rapport avec le sort réservé aux questions méthodologiques au cours des deux congrès précédents de l'Association Française de Sociologie (AFS), qui ont donné lieu à un traitement analogue.

Deux traditions : La méthodologie de l'enquête et le débat théorique

- 4 En premier lieu, il convient de se souvenir que la classification descendante hiérarchique opère les regroupements lexicaux par oppositions successives. Les points communs à l'intérieur des classes ne peuvent être induits qu'ensuite.
- 5 La première grande ligne de partage isole 36 % du corpus (plus exactement, 36 % des unités de contexte retenues par le logiciel, soit 88 % de l'ensemble du corpus). Les tout

premiers mots-clés sont *réponse, interview, qualité, enquête, vie, pauvre, panel, population, âge*. Ils dessinent un paysage de recherche de terrain à dominante empirique pour lesquelles une attention particulière est prêtée au recueil des données. Celui-ci est classique en sociologie et prend deux formes : l'enquête par questionnaire ou l'entretien. D'un autre côté, les 64 % du corpus traité restant paraissent plus distants du terrain, soit parce qu'ils sont conceptualisés au travers de catégories fréquentes comme celles de *lien, groupe, institution, discrimination*, soit parce que les contributions des acteurs portent sur des productions antérieures des sociologues au travers du patrimoine scientifique commun (ainsi apparaissent les notions de *sociologie, empirique, théorique, idéal type, champ, discipline*), ou enfin parce que les textes sont considérés en eux-mêmes comme l'objet d'un commentaire, d'une discussion voire d'un traitement informatisé : *corpus, analyse de correspondance, graphe, discours, logiciel*.

- 6 Au-delà de cette première ligne de clivage selon la distance au terrain, d'autres regroupements se dessinent. L'objet de la revue incite à choisir les options méthodologiques les plus visibles au sein de chacun des cinq classes. Le premier tiers à dominante empirique évoqué se scinde en fait en deux classes distinctes. La première porte spécifiquement sur la **Situation d'enquête**. Elle est alimentée par une réflexion sur le biais d'entretiens téléphoniques et d'interviews, en même temps que des questionnaires. Si la notion de terrain est mentionnée, on ne sait pas précisément lequel. En revanche, dans la classe connexe, nombre d'entre eux sont suggérés explicitement : pauvre, âge, sexuel, santé, vie, drogue, mais aussi ville, culture et chômage. Peut-on conclure à une séparation entre les auteurs préoccupés seulement des techniques d'entretien et ceux qui parlent de leurs objets privilégiés ? Non, car dans cette dernière classe certaines options méthodologiques ressortent spécifiquement : si la notion d'enquête est commune aux deux classes à un haut niveau de signification, la perspective longitudinale est nettement privilégiée. C'est pourquoi il est tentant d'attribuer le libellé d'**Analyse longitudinale** à cette classe, qui est structurée autour des analyses de panel, de périodes, de cohortes et de trajectoires. Il y aurait donc une spécificité méthodologique autour de ces objets de recherche. Nous verrons plus loin si la chronologie des publications intervient en ce sens.
- 7 Les deux tiers restants font l'objet d'un découpage plus complexe. La classe **Analyse des réseaux** la plus significative représente 17 % du corpus traité. Des indications méthodologiques apparaissent avec un khi-deux d'association à la classe très élevée (40 à 80) : les notions de liens, de réseaux, des groupes, d'élites, mais aussi d'inférence, d'équation, de distribution et de structure sont particulièrement saillants, dont la chronologie nous dira si elle est une tendance émergente de l'analyse des réseaux sociaux. Les objets de recherche sous-jacents sont ceux des institutions, des entreprises, des métiers, mais aussi du pouvoir et de la politique, et enfin des femmes et de la famille.
- 8 La classe **Débat théorique** présente une spécificité relative, mais représente une grande partie du corpus retenu, 32 %. Elle reflète sans doute une tradition du débat sociologique autour de la science, de la théorie et du paradigme, de la méthodologie et de l'épistémologie plus que de la méthode, mais aussi des oppositions empirique/théorique, qualitatif/quantitatif. Les outils mobilisés sont ceux de l'idéal type, du principe, du sens, de l'idée et du modèle. Il n'est pas fait référence significativement à des objets empiriques de recherche ou de terrains, mais à des champs, des domaines, des cadres. On peut ainsi penser que cette classe présente le moins de spécificité dans la revue au regard de la structure générale de la sociologie : elle en reflète une composante traditionnellement importante.

- 9 Enfin, une dernière classe, **Analyse textuelle**, présente à la fois une spécificité faible et une importance faible dans le corpus (15 %). Elle est centrée sur les analyses de correspondance et leurs représentations graphiques. Elle s'appuie sur des outils numériques et une réflexion sémantique à propos des corpus textuels et de l'analyse du langage, dont relève la présente contribution.
- 10 Dans l'ensemble, les cinq classes suggèrent bien cinq entrées méthodologiques préférentielles. Les classes 1 et 3 privilégient l'entretien et l'analyse longitudinale, alors que les classes 2, 5 et 4 sont centrées sur l'analyse de réseaux, la modélisation et l'analyse textuelle.
- 11 En troisième lieu, cette penta-répartition peut être rapprochée du sort réservé aux questions de méthode dans la sociologie française du moins telle qu'elle s'est traduite dans les congrès de l'Association Française de Sociologie.
- 12 Lors du congrès de Bordeaux de 2006, une classe sur cinq relevait spécifiquement de question de méthode. Même si celles-ci traversaient aussi les autres classes fondées sur une thématique de recherche, en particulier les réflexions sur la passation des questionnaires et des entretiens.
- 13 Trois sous-ensembles s'en dégageaient pourtant. Le premier portait sur la sociologie, la connaissance, la posture et les auteurs que l'on retrouve ici. Le second avait plus trait aux usages, aux situations, aux dimensions subjectives et cliniques, que l'on ne retrouve pas ici puisqu'on a vu que les récits de vie, par exemple, se retrouvaient rapprochés de l'analyse longitudinale. Le troisième sous-ensemble, enfin, était centré sur l'analyse et l'interprétation des énoncés, des discours et du langage, vus sous l'angle de la compréhension, des idées, des concepts, et de la portée heuristique, qui est, ici, déportée nettement vers les outils numériques de traitement des discours.
- 14 Dans l'ensemble, on note quelques distorsions qui peuvent aussi être considérée comme une certaine autonomisation du débat de méthodes : l'analyse des réseaux, l'analyse longitudinale et l'analyse textuelle se présentent comme des productions originales dans le présent corpus, alors qu'elles n'apparaissent pas significativement dans le congrès de 2004 et 2006 de l'AFS. En revanche, deux traditions de la recherche sociologique impriment leur marque sur le corpus du *BMS* : celles des précautions concernant le recueil des données empiriques, particulièrement par voie d'enquête et d'entretien, et celle du débat scientifique structuré autour des œuvres, des auteurs et des champs disciplinaires. Une absence notable est à relever : les manuels présentent souvent l'analyse de contenu comme une méthode importante du recueil de données et de leur traitement. Elle n'apparaît nulle part, ni dans les deux congrès de l'AFS, ni dans un quart de siècle de *BMS*.
- 15 En quatrième lieu, la datation des contributions au *BMS*, permet-elle de saisir des tendances déclinantes ou émergentes ? En dehors d'effets propres à certains numéros, la chronologie montre que certaines classes semblent atemporelles, alors que d'autres sont plus nettement datées. En effet, peu d'années ressortent de la première classe **Situation d'enquête** (1999), ce qui suggère qu'elle constituerait une préoccupation permanente, confirmant l'hypothèse précédente d'un reflet d'une tradition de précautions méthodologiques au moment du recueil de données. Cette hypothèse d'une méthodologie comme reflet d'une pratique permanente de recherche a aussi été formulée pour le débat conceptuel et le traitement des auteurs. La chronologie est en effet peu discriminante : tout au plus les années 1994 et 1997 sont elles un peu plus spécifiques de ces débats.

- 16 En revanche, les trois classes qui paraissent portées par une production méthodologique originale sont plus nettement datées : la classe **Analyse des réseaux** est très nettement datée à 1984 et 1992, ainsi que les années intermédiaires qui propulsent la période 1983-1992. La classe **Analyse longitudinale** est moins nettement datée mais mentionne toutefois l'année 1993 comme particulièrement significative. Enfin, la classe **Analyse textuelle** est à la fois nettement datée dans un premier temps, en 1983, et plus récemment pendant les années 2005-2007 et 2001-2008 apparaissant avec un degré d'association élevé à cette classe.
- 17 Ainsi la chronologie confirmerait l'hypothèse d'une production méthodologique propre dans ces trois dernières classes.

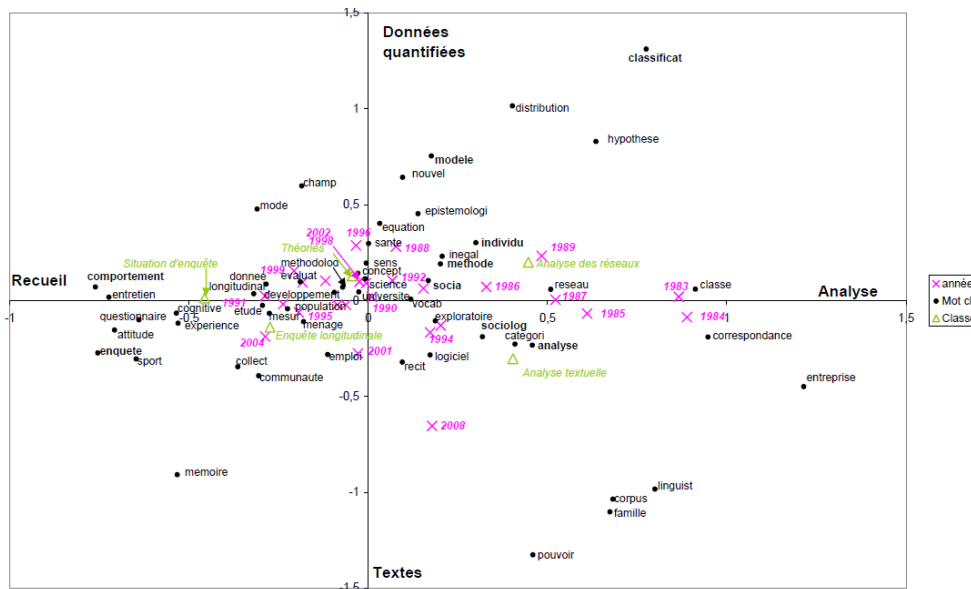
Lorsque l'analyse factorielle paraît réductrice

- 18 La sélection des mots-clés les plus fréquents par numéros de la revue est peu exploitable en vue d'une analyse factorielle. Un regroupement est nécessaire par proximité sémantique : ainsi *entretien*, *interview*, *réponse*, *répondre*, *téléphone* ont été regroupés dans **entretien**. Il a également été tenu compte de la proximité de ces formes lexicales dans la classification ascendante précédente. Ainsi, *enquête* qui apparaît dans deux classes n'a pas donné lieu à un regroupement, de même que *méthode* et *analyse*. Restent ainsi 56 **super mots-clés** projetés ci-dessous selon les 99 numéros de la revue. Les années de publication ainsi que les classes lexicales issues de la correspondance ascendante hiérarchique sont mentionnés en variables supplémentaires.
- 19 Le premier plan (16 % des contributions) présente en première approche un système d'opposition familier à la sociologie : l'axe horizontal semble structuré autour du recueil des données et de leur analyse, alors que le second opposerait des données plutôt quantifiées d'un côté ou des données textuelles d'un autre.
- 20 Horizontalement, vers la droite du plan, les contributions positives au premier axe sont apportées par *analyse*, *correspondance*, *catégorie* alors que les contributions négatives, représentées sur la partie gauche du plan se concentrent autour des mots *entretien*, *enquête*, *question*. La présence de nombreux mots ayant trait à la situation d'entretien à *attitude*, *expérience* révèle une attention particulière aux situations d'interaction dans cette phase de la recherche. A l'autre pôle, le type d'analyse est plutôt formalisé, en étant fondé sur la mise en évidence de correspondances et de catégories.
- 21 Verticalement, suivant le deuxième axe, les données textuelles sont illustrées du côté le plus interactif par la référence à la *mémoire* qui inclut *rétrospectif* et *événement*. De l'autre côté, plus formalisé, apparaissent les corpus textuel analysés à l'aide des outils numériques et les catégories de la linguistique qui sont mobilisées à cette occasion. Les données quantifiées apparaissent tout particulièrement dans les opérations de classification et de commentaires sur les distributions, qui incluent les références à *simulation*, *inférence* et *stabilité*.
- 22 Finalement, quatre classes sur cinq dégagées par Alceste se retrouvent ici assez aisément : la classe **Situation d'enquête** sur le premier axe *Recueil*, la classe **Analyse des réseaux** partagée entre le premier axe, côté *Analyse* et le second, côté *Données quantifiées*, suivie de la classe **Analyse textuelle** entre les pôles *Analyse* et *Données textuelles*. Enfin, la classe **Analyse longitudinale** se situe plutôt entre le pôle *Données textuelles* de l'axe 2 et le pôle *Recueil* de l'axe 1. Seule la classe **Débat théorique** est peu visible, à proximité de l'origine

des axes, ce qui peut être mis en rapport avec le fait qu'elle apparaissait comme relativement atemporelle selon la datation des publications.

- 23 La tendance chronologique des différents numéros du *BMS* présente une boucle ouverte qui part du pôle *Analyse* du premier axe pour ensuite se rapprocher de l'autre extrême du même axe, vers les conditions de *Recueil* des données, puis, à l'issue de quelques spirales, tendre vers les données textuelles, longitudinales, puis numériques.
- 24 Dans l'ensemble, l'analyse factorielle, dont les trois premières dimensions ont été testées, apporte peu d'informations supplémentaires au regard de la classification descendante et de la classification ascendante des formes lexicales. Un enseignement majeur de cette brève mise en perspective est que deux traditions de recherche semblent relativement atemporelles et donc faire partie intégrante de la sociologie et des débats de méthode qui la traverse : les précautions méthodologiques au moment crucial du recueil des données, mais aussi les débats autour des paradigmes, des disciplines proches et de l'épistémologie. Le second enseignement est que trois courants ont été particulièrement valorisés par le *BMS*, avec une datation assez précise : l'analyse longitudinale et les catégories qui l'accompagnent, en particulier autour du travail de la mémoire et des données rétrospectives, l'analyse des réseaux, armée d'un vocabulaire spécifique assez formalisé autour de graphes, de noyaux et de cibles, et enfin la statistique textuelle qui tente de concilier un usage raisonné des outils numériques et des emprunts à la sémantique et plus largement à la linguistique. Il est encore un peu tôt pour savoir si ces trois formes émergentes sont candidates à un renouvellement des traditions des méthodes de recherche en sociologie.

Graphique 1 : AFC des mots clés selon les numéros du *BMS* 1 à 99 ; plan des premier et troisième facteurs



BIBLIOGRAPHIE

Gérard Boudesseul (2005), « De quoi parlent les sociologues réunis en congrès ? Eléments de complémentarité entre une analyse lexicale ouverte et le cumul de variables fermées », *BMS*, n. 58, janvier.

Gérard Boudesseul (2006a), « Le sens des mots par la répétition ou en dépit d'elle ? Dimension sémiotique des statistiques textuelles » in Didier Demazière *et al.* (sous la direction de), *Analyses textuelles en sociologie. Logiciels, méthodes, usages*, Presses Universitaires de Rennes.

Gérard Boudesseul (2006b), « Brève sociographie du Second Congrès de l'AFS - Bordeaux septembre 2006 », *BMS*, n. 92, octobre.

RÉSUMÉS

En utilisant les titres, les résumés et les mots-clés des 318 articles des numéros 1 à 99 du *BMS*, l'auteur applique des programmes de classification hiérarchique ascendante et descendante, et une analyse factorielle des correspondances. Avec les résultats de ces programmes, il décrit en détail cinq classes thématiques majeures : Situation d'enquête, Analyse longitudinale, Analyse des réseaux, Débat théorique et Analyse textuelle. Leurs interrelations et leur évolution dans le temps sont aussi décrites.

A Quarter of a Century of Methodology - The Emergence of Longitudinal Analysis, Social Networks and Textual Statistics: Using the titles, abstracts and key words (in French) of the 318 articles in the *BMS*, issues 1 to 99, the author analyzes this corpus with hierarchically ascending and descending classification programs and with factorial correspondence analysis. With the results of these three methods, he describes in detail five major thematic classes: Survey Situation, Longitudinal Analysis, Social Network Analysis, Theoretical Debate and Textual Analysis. Their interrelationship and evolution over time is also described.

INDEX

Keywords : Bibliometrics, *BMS*, Classification Analysis, Factorial Correspondence Analysis, Thematic Evolution

Mots-clés : Analyse classificatoire, Analyse factorielle des correspondances, Bibliométrie, *BMS*, Evolution thématique

AUTEUR

GÉRARD BOUDESSEUL

CMH-CNRS-ENS-EHESS, gerard.boudesseul@caen.iufm.fr