Les méthodes « dictionnaires »

Parce que les méthodes de codage manuel demandent beaucoup de main-d'œuvre, une approche plus récente tend à remplacer les méthodes de codage manuelle des textes par des techniques informatisées (Kleinnijenhuis et Pennings 2001 ; Laver et Garry 2000). La plupart de ces techniques calquent le codage manuel, remplaçant le codeur humain par des programmes informatiques pour effectuer un codage basé sur des dictionnaires qui associent des mots spécifiques à des positions politiques prédéterminées.

Trois techniques distinctes qui emploient un dictionnaire seront discutées. La technique « *Positional Word Count* » construit un dictionnaire basé sur les mots les plus fréquemment employés qui sont représentatifs d'une position politique prise par un parti donné sur un enjeu. Une analyse de contenu des plateformes électorales de l'élection britannique de 1992 par Laver et Garry (2000) est un bon exemple de l'utilisation de cette technique. L'analyse fait correspondre les mots des dictionnaires prédéfinis conservateur, neutre et travailliste, avec les mots que l'on retrouve dans les plateformes électorales. À chaque mot est accordée une valeur de -1 s'il est également présent dans le dictionnaire conservateur, 0 dans le dictionnaire neutre et +1 dans le dictionnaire travailliste. Le même dictionnaire est ensuite utilisé pour analyser les plateformes des partis Conservateur et Travailliste lors de l'élection de 1997.

Kleinnijenhuis et Pennings (2001) ont développé une méthode qui rend le codage informatisé compatible avec le codage manuel CMP. Au lieu de donner à chaque mot dans une plateforme électorale une valeur de -1,0 ou +1, selon qu'ils soient présents ou non dans les dictionnaires conservateur, neutre ou travailliste, tel que Laver et Garry ont procédé dans leur étude des plateformes britanniques, Kleinnijenhuis et Pennings associent une liste de mots aux 54 catégories de la grille de codage CMP. Leur méthode consiste à calculer à quelle fréquence chaque mot du dictionnaire se retrouve dans les plateformes électorales analysées.

L'approche de la phrase naturelle développée par Ray (2001) utilise également le codage CMP-compatible. Par contre, contrairement à la méthode de Kleinnijenhuis et Pennings qui utilise le mot en tant qu'unité d'analyse, la technique

de Ray utilise la phrase (ou quasi-phrase) comme unité d'analyse. Le dictionnaire de Ray enregistre aussi la direction politique (« en faveur de », « opposé à », « appuie ») des énoncés de la plateforme électorale, en plus de l'enjeu. Les schémas de codage informatisés comme ceux décrits précédemment permettent de produire des estimés de positionnement politique beaucoup plus précis et valides à partir des mêmes textes lorsqu'ils sont comparés au codage manuel utilisé par le CMP. Cependant, il existe plusieurs limites associées au type de dictionnaire utilisé par ces méthodes :

1. Il faut construire un nouveau dictionnaire pour chaque échelle sur laquelle on souhaite situer les textes politiques. Il en résulte que, peu importe le niveau élevé d'automatisation, la méthode nécessite une implication humaine importante dans le développement et le testage de dictionnaires de codage qui correspondent à des époques et de cultures politiques particulières.
2. Tel que Laver, Benoit et Garry le mentionnent (2003 : 312), « l'implication humaine importante lors de la mise au point de dictionnaires de codage importe certains des désavantages méthodologiques des techniques traditionnelles basées sur des codes humains potentiellement biaisés ».
3. Une autre limite est le nombre peu élevé de données à partie desquelles un dictionnaire typique est construit. Ce faisant, additionner ou soustraire un mot du dictionnaire peu avoir un impact considérable sur le résultat final.
4. La méthode ne permet pas précisément de calculer une mesure d'erreur associée aux résultats générés par le dictionnaire. La validité de la méthode est donc impossible à établir.

Bibliographie

Kleinnijenhuis, J. et P. Pennings (2001). « Measurements of Party Positions on the Basis of Party Programmes, Media Coverage, and Voter Perceptions », dans M. Laver, dir., *Estimating the Party Position of Political Actors*, New York, Routledge, p. 162-182.

Laver, Michael, Kenneth Benoit et John Garry (2003), « Extracting Policy Positions from Political Texts Using

Words as Data », *The American Political Science Review*, vol. 97, n° 2, p. 311-331.

Laver, Michael et John Garry (2000), « Estimating Policy Positions from Political Texts », *American Journal of Political Science*, vol. 44, n° 3, p. 619-634.

Ray, Leonard (2001). « A Natural Sentence Approach to the Computer Coding of Party Manifestos », dans M. Laver, dir., *Estimating the Positions of Political Actors*, New York, Routledge, p. 149-161.