

Coupler Analyse de données textuelles et IA génératives pour « augmenter » le processus de recherche qualitative

Jean Moscarola

Université Savoie Mont Blanc – jmoscarola@gmail.com

Abstract

This paper is aimed at qualitative researchers. It develops the idea of coupling the use of ADT and generative AI to benefit from the complementarity between the statistical methods and the connexionist methods. In particular, it provides an answer to the difficulty of interpreting ADT results and to the limits and risks presented by AI. The paper will be developed in 4 parts: 1/ A presentation of ADT methods and how generative AI works, to show what differentiates them and what makes them complementary; 2/ Illustrations on 2 examples of non-directive interview analysis and bibliographic review analyzed with Sphinx and ChatGPT; 3/ An explanation of the notion of augmented qualitative research and its implementation through prompts built from lexicons resulting from ADT. 4/ A conclusion on the feasibility of this approach and the epistemological and ethical questions it raises.

Keywords : qualitative research, textual data analysis, generative AI, lexical statistics, semantics, ChatGpt, Sphinx

Résumé

Ce papier s'adresse aux chercheurs qualitatifs et développe l'idée de coupler l'usage de l'ADT et des IA génératives pour bénéficier de la complémentarité entre méthodes statistiques et connexionnistes. Il apporte en particulier une réponse à la difficulté d'interprétation des résultats de l'ADT et aux limites et risques présentés par l'IA. Le papier se développera en 4 parties 1/ La présentation des méthodes de l'ADT et du fonctionnement des IA génératives pour montrer ce qui les différencie et ce qui les rend complémentaires 2/ Des illustrations sur 2 exemples d'analyse d'entretiens non directifs et de revue bibliographique analysés avec Sphinx et ChatGPT 3/ Une explicitation de la notion de recherche qualitative augmentée et de sa mise en œuvre par des prompts construits à partir des lexiques résultant de l'ADT. 4/ une conclusion sur la faisabilité de cette approche et sur les questions épistémologiques et éthiques qu'elle pose.

Mots clés : recherche qualitative, analyse de données textuelles, IA générative, statistique lexicale, sémantique, ChatGpt, Sphinx