

Informations

Auteur.rice : PASTORELLI Benjamin, ESTEBAN Charlotte

Disciplines : sociologie et psychologie.

A laquelle des questions la soumission s'adresse : Expérimentation et observation : comment concevoir une expérience ?

Type de soumission : poster.

Résumé

« Je te comprends » ; proposition d'expérience pour un chatbot empathique

En 1966, l'informaticien allemand Joseph Weizenbaum publiait le célèbre article « ELIZA, a computer program for the study of natural language communication between man and machine », lequel restitue une expérimentation menée avec le programme d'automate conversationnel ELIZA (Weizenbaum, 1966). Cette expérience consiste à simuler un dialogue avec un programme, sur le modèle d'un échange avec un thérapeute. Si l'expérience est considérée comme pionnière dans le domaine, elle a l'intérêt de mettre en lumière deux dimensions des chatbots : d'une part, les chatbots comme dispositifs de compréhension linguistique, et d'autre part comme figure compréhensive. Aujourd'hui, les applications de dispositifs chatbots sont divers ; dans le domaine industriel, ceux-ci répondent actuellement à des contraintes de rationalisation et de réduction de coût des échanges. Pourtant, l'observation d'échanges avec un chatbot marchand montre tout de même que la figure de l'interlocuteur humain reste présente (Esteban & Beauvisage, 2020), ainsi que le souci d'échanger avec le « bon » interlocuteur— les usagers de chatbots pouvant réclamer « un humain » plutôt qu'un automate. Dans le secteur du service en particulier, l'empathie est considérée comme une compétence sociale, notamment dans les métiers relationnels où la gestion des émotions constitue une activité centrale (Grosjean & Weerdt, 2005), bien qu'invisibilisée (Hochschild, 2003) : amener des usagers d'un service à *se sentir compris* fait donc partie de ce travail émotionnel... Cependant, les automates conversationnels semblent dépourvus de cette qualité, pour deux raisons : d'une part, parce que ces dispositifs n'ont actuellement pas les ressources techniques pour comprendre les émotions des usagers, d'autre part, parce que les usagers ne projettent pas nécessairement cette qualité de compréhension émotionnelle sur les chatbots.

Le problème de la qualification ontologique de ces dispositifs, posé en 1950 par Alan Turing (Turing, 1950), reste ainsi une problématique d'actualité à l'heure où l'automatisation des échanges langagiers touche des domaines de plus en plus variés. Précisons toutefois qu'il n'est pas dans notre intention de trancher la question de savoir si une entité artificielle peut réellement ressentir de l'empathie ou non. Selon nous, cette question ontologique relève plutôt de la philosophie que de nos disciplines : ce qui nous intéresse est de savoir si un chatbot peut *faire preuve* d'empathie, c'est-à-dire *montrer* de l'empathie à un être humain.

Comment définir l'empathie ? Dans le domaine de la psychologie, plusieurs définitions de l'empathie existent (Davis, 2006). La plupart la présente comme la capacité à comprendre les états psychologiques d'autrui (Wispé, 1986), sans nécessairement partager ces états et ressentir à son tour les émotions associées. Ce deuxième cas concerne plutôt la « sympathie ».

S'engager sur cette voie suppose de caractériser l'empathie et de définir les réponses comportementales associées. À ce propos, le travail de Sharma et de son équipe (Sharma, Miner, Atkins & Althoff, 2020) identifie trois grands types de réponse : réaction, interprétation et exploration. Les réactions sont des démonstrations émotionnelles. Elles donnent un retour chargé de compassion, de chaleur ou d'inquiétude par exemple. Les interprétations traduisent la compréhension par la personne soutenant des états psychologiques de la personne soutenue. Les explorations mettent en avant l'intérêt de la personne soutenant envers la personne soutenue et sa volonté de mieux la comprendre. Ces trois réactions contribuent à raffermir le soutien et à créer un sentiment de connexion, propres à l'empathie.

En nous appuyant sur une approche pluridisciplinaire mêlant sociologie et psychologie, nous émettons l'hypothèse qu'un chatbot peut *se montrer* empathique en respectant ces règles. Mais nous envisageons une nuance. En effet, nous envisageons que l'identité artificielle du bot modifie la crédibilité accordée à ses comportements, notamment en fonction de s'il se montre empathique (compréhension de l'état psychique d'autrui) ou sympathique (partage de l'état psychique d'autrui). Ce deuxième cas pourrait être perturbateur et réduirait la confiance du partenaire humain. En effet, le bot, en tant que machine, n'est pas attendu pour éprouver des émotions. En émettant cette idée, il remet en question sa place relative à l'humain (voir Sundar, 2020).

Pour tester ces prédictions, nous nous reposerons sur l'expérimentation. Des personnes seront amenées à échanger avec un chatbot créé pour l'occasion et dont le comportement variera selon les facteurs étudiés (empathie ou sympathie, réactions impliquantes ou non, réactions fortes ou faibles). Cette interaction sera suivie d'un questionnaire. Plusieurs mesures, comportementales (durée de l'échange, nombre d'interactions, prolixité) et psychométriques (empathie perçue, confiance accordée, inclusion d'autrui dans le soi), permettront d'évaluer la situation.

Références

Borelle, C. (2018). Sortir du débat ontologique. *Rezeaux*, n° 212(6), 207-232.

Davis, M. H. (2006). Empathy. In *Handbook of the sociology of emotions* (pp. 443-466). Springer, Boston, MA.

Esteban, C. (2020). Construire la « compréhension » d'une machine. *Rezeaux*, N° 220-221(2), 195-222.

Esteban, C., & Beauvisage, T. (2020). Interacting with a Salesman Chatbot. In C. Stephanidis, M. Antona, & S. Ntoa (Éds.), *HCI International 2020 – Late Breaking Posters* (p. 308-314). Springer International Publishing.

Grosjean, V., & Weerdt, C. R.-V. D. (2005). Vers une psychologie ergonomique du bien-être et des émotions : Les effets du contrôle dans les centres d'appels. *Le travail humain*, Vol. 68(4), 355-378.

Hochschild, A. R. (2003). Travail émotionnel, règles de sentiments et structure sociale. *Travailler*, n° 9(1), 19-49.

Nass, C., Moon, Y., & Carney, P. (1999). Are People Polite to Computers? Responses to Computer-Based Interviewing Systems¹. *Journal of Applied Social Psychology*, 29(5), 1093-1109. <https://doi.org/10.1111/j.1559-1816.1999.tb00142.x>

Sharma, A., Miner, A. S., Atkins, D. C., & Althoff, T. (2020). A computational approach to understanding empathy expressed in text-based mental health support. *arXiv preprint arXiv:2009.08441*.

Sundar, S. S., Jia, H., Waddell, T. F., & Huang, Y. (2015). Toward a theory of interactive media effects (TIME) four models for explaining how interface features affect user psychology. In *The Handbook of the Psychology of Communication Technology* (pp. 47–86). Wiley. <https://doi.org/10.1002/9781118426456.ch3>

Turing, A. M. (1950). I.—COMPUTING MACHINERY AND INTELLIGENCE. *Mind*, LIX(236), 433-460. <https://doi.org/10.1093/mind/LIX.236.433>

Weizenbaum, J. (1966). ELIZA, a computer program for the study of natural language communication between man and machine. *Communications of the ACM*.

Wispé, L. (1986). The distinction between sympathy and empathy: To call forth a concept, a word is needed. *Journal of personality and social psychology*, 50(2), 314.