Actes de CORIA-TALN-RJCRI-RECITAL 2025
Marseille, France, 30 juin au 4 juillet 2025

Culture et acculturation des grands modèles de langue

Mathieu Valette[1]
(1) ERTIM, Inalco, 2 rue de Lille, 75007 Paris, France
mvalette@inalco.fr

RESUME

Il s'agira d'évaluer la place octroyée à la culture dans les travaux industriels et académiques portant sur la constitution de grands modèles de langue (LLMs), notamment lorsqu'il s'agit de les aligner. Le premier constat effectué est que la culture y est appréhendée de manière restreinte à des problématiques axiologiques (valeurs morales). Le deuxième constat est que les travaux actuels portant sur les cultures dans les LLMs se divisent en deux catégories : (i) évaluation des biais culturels par la confrontation à des référentiels culturels tiers, (ii) alignement axiologique. Nous discuterons des conséquences de ces orientations épistémologiques.

ABSTRACT

Culture and Acculturation of Large Language Models.
The objective will be to assess the role assigned to culture in both industrial and academic work related to the development of large language models (LLMs), particularly in the context of their alignment. The first observation is that culture is most often reduced to axiological concerns (moral values). The second observation is that current research on culture in LLMs falls into two main categories: (i) the evaluation of cultural biases through comparison with third-party cultural frameworks, and (ii) axiological alignment. We will discuss the consequences of these epistemological orientations.

MOTS-CLÉS : LLMs, culture, alignement, référentiel culturel, perspectivisme.
KEYWORDS : LLMs, culture, alignment, cultural framework, perspectivism.

1 Introduction

Le jeu d'imitation distributionnel que nous offre les industriels de la Tech est si réussi qu'en dépit de leurs défauts de jeunesse (Maleki *et al.*, 2024 dénombrent 9 types d'hallucination), les grands modèles de langues (LLMs) passeraient sans peine pour de grands modèles de culture. Dans les téraoctets de pages de textes ayant servi à leur entrainement, on sait qu'il y a des bibliothèques entières. Toutefois, si le phrasé est d'une perfection surhumaine, si souvent la qualité du raisonnement convainc, la pertinence culturelle quant à elle, est beaucoup plus aléatoire et fortuite. C'est qu'aux œuvres *remarquables* – les encyclopédies, les essais, les romans, les articles scientifiques – ces objets culturels destinés à être choisis, contextualisés, interprétés, et transmis pour faire commun, s'ajoutent sans ordre ni hiérarchie toutes les conversations de comptoir, les fonds de tiroirs et les ignominies d'Internet. Ainsi, *ça parle*, mais ça parle pour ne rien dire, ou pour dire des foutaises (Franfurt, 2005). Au fond, à moins de les circonscrire prudemment à des

Actes de l'atelier Évaluation des modèles génératifs (LLM) et challenge 2025 (EvalLLM), pages 68–76.

tâches de bas niveau (segmentation, lemmatisation) ou, à la limite, bureautiques (traduction automatique, aide à la rédaction, résumé de textes), les LLMs sont dangereux pour les cultures car ils en déprécient les objets, les segmentent, les atomisent, à telle enseigne que les concepteurs eux-mêmes nous enjoignent parfois de ne pas s'y fier.

Mais qu'est-ce que la culture ? (Krœber & Kluckhohn 1952) en recensèrent 164 définitions possibles ; contentons-nous de celle, consensuelle, qu'en donna l'UNESCO en 1982 : « [d]ans son sens le plus large, la culture peut aujourd'hui être considérée comme l'ensemble des traits distinctifs, spirituels et matériels, intellectuels et affectifs, qui caractérisent une société ou un groupe social. Elle englobe, outre les arts et les lettres, les modes de vie, les droits fondamentaux de l'être humain, les systèmes de valeurs, les traditions et les croyances »[1]. De cette définition de toute évidence inspirée par l'anthropologie tylorienne (Tylor, 1871), nous allons voir que les travaux portant sur la culture dans les LLMs ne retiennent que « les systèmes de valeurs », sans considérations notables pour les arts, les sciences et les lettres, ni les productions culturelles elles-mêmes. Pour la plupart, ils relèveraient ainsi d'une approche culturaliste centrée sur l'individu, sa personnalité, ses valeurs. Or, le culturalisme est un courant en anthropologie, principalement étasunien (par ex. Boas, 1911) réputé relativiste : les cultures doivent y être étudiées et interprétées dans leur seul contexte et non par comparaison avec d'autres cultures. On oppose au relativisme culturel la recherche d'universaux et l'approche comparative du structuralisme (par exemple Levi-Strauss, 1949)[2]. Les travaux portant sur cette culture réduite à l'axiologie dans la construction des LLMs[3] se divisent actuellement en deux catégories que l'on peut considérer comme des séquences : (i) évaluation des biais culturels des LLMs par la confrontation à des référentiels culturels tiers ; (ii) alignement axiologique au moyen d'un apprentissage, souvent par renforcement basé sur les commentaires humains (RLHF) ou par l'usage de RAG, pour les rendre conforme à un typage culturel.

Nous nous proposons donc dans cet article d'exposer et de discuter l'arrière-plan épistémologique et idéologiques des recherches relevant de cette problématique par l'examen d'un ensemble de publications académiques récentes, sans considérer systématiquement le détail des résultats obtenus ni les performances des solutions exposées. Ce n'est pas la pertinence de ces travaux qui nous intéresse mais les tendances implicites et les choix effectués par la communauté. Ainsi, nous nous référerons à plusieurs reprises à des *preprints* (souvent déposés sur arxiv.org) non sanctionnés par les pairs ou dont nous ignorons le statut éditorial au moment de la consultation.

---

[1] UNESCO – *Déclaration de Mexico sur les politiques culturelles, Conférence mondiale sur les politiques culturelles,* 26 juil. - 6 août 1982.

[2] L'anthropologie structurale est parfois critiquée pour la possibilité qu'elle donne de hiérarchiser les cultures en fonction de leur rapport aux universaux identifiés ; le relativisme culturel du culturalisme est accusé de légitimer des pratiques contraires à la dignité et aux droits humains (discrimination, mutilation rituelle comme l'excision, etc.).

[3] La génération d'objets culturels (narratifs, filmiques, picturaux, etc.) ne relèvera pas du périmètre de notre discussion.

2 Le benchmarking culturel des LLMs

Les LLMs ont occasionné un trouble épistémologique dans la communauté du TAL en raison de leur architectures réputées impénétrables, ce que (Offert et Dhaliwal, 2025) dénoncent comme une « casuistique de la boite noire » (*black-box casuistry*, *i.e.* un renoncement devant l'opacité des réseaux de neurones multicouches, et une difficulté à surpasser ce renoncement pour élaborer de nouveaux observatoires méthodologiques dédiés aux IA). Les modèles font toutefois l'objet de nombreuses études que l'on pourrait considérer comme de la rétro-ingénierie, notamment par le développement de benchmarks de plus en plus ambitieux et quantitativement significatifs. Très tôt, les biais culturels véhiculés par les LLMs ont en été étudiés. Les premières études, principalement menées sur ChatGPT 3 ont d'emblée mis en évidence le tropisme culturel de l'agent conversationnel d'OpenAI (Cao *et al.*, 2023, Johnson *et al.*, 2022, Liu *et al*., 2023, Wang *et al.*, 2023) : les premiers LLMs disponibles favorisaient nettement les cultures occidentales et plus particulièrement nord-américaines. On ne s'étonne pas que cette tendance perdure lorsqu'on sait qu'une large majorité des textes disponibles sur Internet sont rédigés dans une langue européenne, et près de 50% en anglais encore que cette domination de l'anglais diminue peu à peu depuis quelques années[4]. Plusieurs benchmarks ont été mis en place pour affiner ces analyses. Ils adoptent des référentiels culturels d'inspiration soit psychologique, soit sociologique qui eux-mêmes orientent les travaux par leurs épistémologies ; elles sont en premier lieu toujours déterminées par le principe que la culture est réductible à des valeurs.

2.1 Référentiels culturels psychologiques

Parmi les référentiels d'inspiration psychologique, celui développé par la psychologue et anthropologue néerlandaise Geert Hofstede et ces collaborateurs (Hofstede, 1984 ; Hofstede *et al.*, 2010) apparaît fréquemment mentionné. Or, Hofstede se revendique de l'héritage de Kluckhohn qui, dans le sillage de l'anthropologie culturaliste, a développé une théorie des modèles culturels et des systèmes de valeurs. Le modèle de Hofstede propose un positionnement culturel sur 6 dimensions : (i) rapport au pouvoir (*Power Distance*), (ii) individualisme *vs* collectivisme, (iii) masculin *vs* féminin stéréotypes de genre associés aux comportements et aux émotions, (iv) contrôle de l'incertitude (*uncertainty avoidance*), (v) long terme *vs* court terme, (vi) hédonisme *vs* austérité (*Indulgence vs. Restaint*). (Masoud *et al.*, 2025) utilisent le cadre d'Hodstede pour analyser l'alignement de plusieurs agents conversationnels et mesurer leurs affinités avec les valeurs des Etats-Unis, de la Chine et des pays arabes. (Karinshak *et al.*, 2024) effectuent le même type de travaux en utilisant un cadre culturel développé par (House *et al.*, 2004), le GLOBE (*Global Leadership and Organizational Behavior Effectiveness*), une vaste enquête portant sur 17 000 managers de 62 pays. Adossée aux travaux d'Hofstede l'enquête adopte un cadre sensiblement remanié et enrichi, puisqu'il ne vise pas seulement les valeurs du cadre initial mais y ajoute des valeurs de leadership et de capacité organisationnelle (performance, assertivité, etc.), une indication sur les catégories culturelles intéressant les auteurs : sont-ce les valeurs des LLMs que l'on cherche à inventorier ou leur adéquation à certaines valeurs ?

[4] D'après les statistiques fournies par World Wide Web Technology Surveys 3 (https://w3techs.com, consulté en mai 2025).

2.2 Référentiels culturels sociologiques

Une grande enquête sociologique sur les valeurs à portée internationale domine les travaux d'analyse axiologique des LLM : le World Values Survey (WVS ; Haerpfer et al. 2022). Initiée au début des années 80 par le politologue américain, Ronald Inglehart, l'enquête fait office de référence par son ampleur, sa durée et la qualité de son échantillonnage. Elle est exploitée notamment par (Durmus et al. 2024 ; Benkler et al. 2025 ; Sinacola et al. 2025) et de manière plus critique par (Kabir et al. 2025). (Durmus et al. 2024) par exemple élaborent leur référentiel culturel à partir du WVS et d'une autre enquête internationale moins souvent utilisée, Pew Global Attitudes Survey (PEW)[5]. Ils opèrent un astucieux système de *prompting* croisé de manière à identifier à la fois les biais culturels du LLM testé[6], les préjugés culturels que le LLM attribue à différents pays, et l'incidence de la langue de la requête sur la réponse. L'équipe compare ces questions aux réponses obtenues dans le cadre des enquêtes WVS et PEW. Elle constate une fois encore le tropisme occidental du modèle quelle que soit la langue de la requête, mais observe qu'il est en outre pétri de préjugés : ceux des Étasuniens sur les autres pays. Par exemple, à la question *"Do you personally believe that sex between unmarried adults is* [a] *morally acceptable,* [b] *morally unacceptable, or* [c] *is it not a moral issue?"* le modèle préjuge que les Russes répondront [b] à 73,9% alors que le taux est en fait d'environ 31%, équivalent aux réponses des Étasuniens d'après les enquêtes sociologiques exploitées.

2.3 Le recours non critique aux modèles tiers, une habitude bien ancrée

Si les grandes enquêtes sociologiques internationales sont des références et ont servi de base à des milliers de publications en sciences sociales, elles ne sont pas exemptes de critiques, notamment WVS qui, dans sa sixième vague, introduit des critères de personnalité, les « big Five », contestés à plus d'un titre et notamment parce qu'ils comportent – et c'est un comble – des biais culturels les rendant partiaux pour décrire des valeurs non occidentales (Ludeke & Larsen, 2017). Finalement, la réduction des cultures aux valeurs, puis aux valeurs individuelles par le biais d'approches psychologiques (2.1) voire psychologisantes (2.2) conduit à s'interroger sur les intentions des auteurs : s'agit-il d'évaluer les biais axiologiques des LLMs, d'en construire le sociotype ou un profilage psychologique ? Le recours non critique à des modèles tiers issus de la psychologie et détournés de leur fonction initiale est une habitude en TAL. Ainsi, la théorie des émotions, innéiste, simplificatrice (et accessoirement datée) de Robert Plutchik ou encore le test pseudoscientifique MBTI (Myers Briggs Type Indicator) ont été « implémentés » des centaines de fois dans des applications de détection des émotions ou de personnalités dans les années 2010 sans que jamais ne soient questionnées les limites, le sérieux ou l'adéquation de ces travaux à la tâche. Aurions-nous affaire aux mêmes travers concernant le benchmarking des cultures ?

[5] Lire toutefois (Huang *et al.* 2023).

[6] Vraisemblablement Claude (bien qu'il ne soit pas vraiment formellement désigné dans l'étude) car la recherche émane d'Anthropic.

3 Trois tendances de l'alignement culturel

La construction de référentiels axiologiques et culturels s'opère par des méthodes variées, qui toutes s'appuient sur des choix idéologiques pas toujours explicités. On distingue trois tendances générales.

3.1 L'approche descendante : définir ce que sont les valeurs

L'approche descendante consiste à élaborer un référentiel global inspiré de textes réputés fondamentaux destiné à être acceptable par le plus grand nombre. C'est le cas par exemple des LLMs exploités sous le nom commercial Claude, dite *Constitutional AI* (« IA constitutionnelle ») éditée par la société Anthropic (Bai *et al.* 2022). La « constitution » d'Anthropic s'appuie sur un assemblage de textes de références pour le moins hétéroclites allant de la Déclaration universelle des droits de l'homme (DDH) aux Conditions d'utilisations d'Apple ou DeepMind. Elle comprend 62 « articles » qui correspondent à des principes, ou contraintes à suivre par renforcement (RLHF), par exemple *« Please choose the response that most supports and encourages freedom, equality, and a sense of brotherhood »* (1er principe de la constitution inspiré du 1er article de la DDH) ou *« Choose the response that would be most unobjectionable if shared with children »*. Bien que d'inspiration universelle, la constitution intègre des principes correcteurs prenant en considération des « perspectives non occidentales » comme *« Choose the response that is least likely to be viewed as harmful or offensive to those from a less industrialized, rich, or capitalistic nation or culture »*. L'approche descendante apparaît relativement peu documentée dans le sens où rares sont les travaux académiques l'explicitant. L'IA constitutionnelle d'Anthropic s'inspire de toute évidence des valeurs humanistes occidentales[7], mais d'autres LLMs ont été affinés suivant un même principe descendant. Par exemple l'agent conversationnel DeepSeek de l'entreprise éponyme est a priori aligné sur les positions officielles du gouvernement chinois et a priori sur les valeurs afférentes. Nous ne pourrons toutefois rien en dire n'ayant pas identifié de travaux académiques sur son alignement.

3.2 L'approche ascendante : restituer les axiologies observées

L'approche ascendante s'appuie sur des référentiels sociologiques tels qu'évoqués ci-dessus, afin de localiser axiologiquement les IA en fonction des cibles et ainsi conformer les plateformes à des sensibilités socioculturelles variées. C'est le cas par exemple de (Tao et al. 2024 ; Seo et al. 2025). (Seo et al. 2025) utilisent un RAG basé sur le WVS pour améliorer les modèles GPT-4o-mini (OpenAI) et Gemini-1.5-Flash (Google), ou encore (Li et al. 2024) qui affinent le modèle LLama-2-70b (Meta) à partir d'un échantillonnage de données également issues du WVS et obtiennent des résultats a priori satisfaisant dépassant les modèles mis en vis-à-vis (ChatGPT 3.5, 4, Gemini Pro).

[7] Elle donne lieu à des approfondissements ou inspire d'autres cadres conceptuels, par exemple (Küsters & Wörsdörfer, 2024) sur la création d'IA ordolibérale (l'ordolibéralisme est une doctrine économique libérale octroyant à l'État un pouvoir normatif fort).

3.3 L'approche post-moderne : le relativisme axiologique radicalisé

Quoi qu'elles soient aujourd'hui associées à l'annotation de corpus plus qu'à l'alignement de LLMs, on voit émerger des approches post-modernes de l'axiologie qui radicalisent l'approche ascendante : afin de prendre en compte toutes les sensibilités possibles, toutes les opinions, toutes les valeurs, le paradigme perspectiviste (Cabitza et al., 2023) prône un relativisme généralisé. Partant d'un objectif éthique (désinvisibiliser les opinions des minorités), le paradigme perspectiviste adopte en fait une position post-factuelle non déclarée (il n'existerait pas une vérité mais plusieurs vérités et toutes se valent). Les références des auteurs attestent de ce choix puisqu'ils se réfèrent notamment à (Aroyo & Welthy, 2015) qui périmaient de manière provocatrice le concept de *vérité* (« truth is a lie », « the antiquated ideal of truth », etc.) (Valette, 2024). Il n'est pas anodin que la post-vérité pénètre la discipline pour des motifs éthiques et moraux tout comme elle investit aujourd'hui le champ politique au titre de la liberté d'expression.

4 Conclusion

Si toutes les études établissent que les LLMs reproduisent par défaut les biais culturels nord-américains – parfois jusqu'à la caricature – en raison de données d'entrainement elles-mêmes biaisées, les valeurs, lorsqu'elles sont issues d'un alignement, sont celles décidées par des entreprises qui les élaborent[8]. Ainsi, Anthropic constitutionnalise avec Claude les conditions d'utilisation des grandes industries avec lesquelles elle collabore au même titre que la Déclaration universelle des droits de l'homme. De même, jusqu'en janvier 2024 et l'élection de Donald Trump à la présidence des Etats-Unis, les IA génératives respectaient, voire contractualisaient, le cadre professionnel *Diversity, Equity, Inclusion* (DEI) condamnés depuis par Elon Musk (xAI) et Marc Zuckerberg (Meta). Des travaux menés en 2023 montraient d'ailleurs que l'idéologie de ChatGPT (OpenAI) était pro-environnementaliste et libertarienne de gauche (Hartmann, 2023), soit celle des entreprises de la Silicon Valley à l'époque de l'étude. Qu'en sera-t-il demain ?

Une préoccupante convergence des agendas politiques, scientifiques et technologiques se dessine. La plupart des travaux académiques portant sur la culture et l'alignement culturel des LLMs se fondent sur des travaux anthropologiques, psychologiques et même sociologiques dont les cadres théoriques sont dominés par le relativisme culturel plutôt que par une axiologie minimale valable pour tous. Si l'alignement culturel des LLMs suit l'évolution que nous avons observée ici et, à moins d'une révision épistémologique significative, les alignements rationalistes visant la construction d'un monde commun, deviendront marginaux. Des LLMs toujours plus localisés, communautarisés voire individués participeront à la fragmentation de sociétés conçues comme une collection d'individus à complaire ou à contraindre.

Références

AROYO L. & WELTY C. (2015). Truth is a lie: Crowd truth and the seven myths of human annotation. *AI Magazine*, 36(1): 15–24.

[8] Plus généralement, l'éthique technologique est réputée soumise à l'éthique des entreprises technologique (Green, 2021).

BENKLER N., MOSAPHIR D., FRIEDMAN S., SMART A., SCHMER-GALUNDER S. (2023). *Assessing LLMs for Moral Value Pluralism* [Preprint]. arXiv. https://arxiv.org/abs/2312.10075

BOAS F. (1911). *The Mind of Primitive Man*. The Macmillan Company.

CABITZA F., CAMPAGNER A., BASILE V. (2023). *"Toward a Perspectivist Turn in Ground Truthing for Predictive Computing"*, in *AAAI Technical Track on Machine Learning I,* Vol. 37 No. 6: AAAI-23 Technical Tracks 6.

CAO Y., Zhou L., LEE S., CABELLO L., CHEN M., HERSHCOVICH D. (2023). *Assessing Cross-Cultural Alignment between ChatGPT and Human Societies: An Empirical Study.* Proceedings of the First Workshop on Cross-Cultural Considerations in NLP (C3NLP), pages 53–67, Dubrovnik, Croatie. DOI : 10.18653/v1/2023.c3nlp-1.7.

DURMUS E., NGUYEN K., LIAO T. I., SCHIEFER N., ASKELL A., BAKHTIN A., CHEN C., HATFIELD-DODDS Z., HERNANDEZ D., JOSEPH N., LOVITT L., MCCANDLISH S., SIKDER O., TAMKIN A., THAMKUL J., KAPLAN J., CLARK J., GANGULI D. (2024). *Towards Measuring the Representation of Subjective Global Opinions in Language Models*. [preprint] https://arxiv.org/abs/2306.16388.

FRANKFURT H. G. (2005). *On Bullshit*, Princeton (New Jersey): Princeton University Press.

GREEN B. (2021). "The Contestation of Tech Ethics: A Sociotechnical Approach to Technology Ethics in Practice," in *Journal of Social Computing*, vol. 2, no. 3, pp. 209-225.

HAERPFER C., et al. (2022). World values survey trend file (1981–2022) cross-national data-set, data file version 3.0.0. *Madrid, Spain & Vienna, Austria: JD Systems Institute & WVSA Secretariat*. , 10.14281/18241.23

HARTMANN J., SCHWENZOW J., WITTE M. (2023). The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation, arXiv:2301.01768v1 [cs.CL].

HOFSTEDE G. (1984). Culture's consequences: International differences in work-related values, *Cross Cultural Research and Methodology*, Sage publications.

HOFSTEDE G., HOFSTEDE G. J., MINKOV M. (2010). *Cultures and Organizations: Software of the Mind: Intercultural Cooperation and Its Importance for Survival* (3rd ed.). New York: McGraw-Hill.

HOUSE R. J., HANGES P. J., JAVIDAN M., DORFMAN P. W., GUPTA V. (2004). *Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies*, Thousand Oaks (Calif.): Sage Publications.

JOHNSON R, PISTILLI G., MENEDEZ-GONZALEZ N., DIAS DURAN L. D., PANAI E., KALPOKIENE J., BERTULFO D.J. (2002). "The Ghost in the Machine has an American accent: value conflict in GPT-3", arXiv, 2022, https://arxiv.org/abs/2203.08900.

KABIR M., ABRAR A.,ANANIADOU S. (2025). *Break the Checkbox: Challenging Closed-Style Evaluations of Cultural Alignment in LLMs* [Preprint]. arXiv. https://arxiv.org/abs/2502.08045

KARINSHAK E, HU A., KONG K., RAO V., WANG J. (2024). LLM-GLOBE: A Benchmark Evaluating the Cultural Values Embedded in LLM Output, [preprint– arXiv:2411.06032v1 [cs.CL] 9 Nov 2024.

KÜSTERS A., WÖRSDÖRFER M. (2024). Exploring Laws of Robotics: A Synthesis of Constitutional AI And Constitutional Economics, *CEPinput*, n°13, Freiburg/Berlin.

KRŒBER A. L. & KLUCKHOHN C. (1952). *Culture: A Critical Review of Concepts and Definitions*, Cambridge (Mass.): Peabody Museum Press.

LÉVI-STRAUSS C. (1962). *La Pensée sauvage*, Paris : Plon.

LI C., WANG J., CHEN M., SITARAM S., XIE X. (2025). CultureLLM: Incorporating Cultural Differences into Large Language Models. *Proceedings of the 38th International Conference on Neural Information Processing Systems*, *37*, 1-15, https://doi.org/10.48550/arXiv.2402.10946.

LUDEKE S. G, LARSEN E.G. (2017). Problems with the Big Five assessment in the World Values Survey, In *Personality and Individual Differences*, Vol. 112, pp. 103-105.

MALEKI N., PADMANABHAN B., DUTTA K. (2024). AI Hallucinations: A Misnomer Worth Clarifying, arXiv:2401.06796v1 [cs.CL] 9 Jan 2024. Ludeke S. G, Larsen E.G. (2017). MASOUD R., LIU Z.,

FERIANC M., TRELEAVEN P., RODRIGUES M. (2025). Cultural Alignment in Large Language Models: An Explanatory Analysis Based on Hofstede's Cultural Dimensions, in *Proceedings of the 31st International Conference on Computational Linguistics*.

OFFERT F., DHALIWAL R. S. (2025). The method of Critcal AI Studies, A Propaedeutic, arXiv:2411.18833v3 [cs.CY] 23 Mar 2025.

SEO W., YUAN Z., BU Y. (2025). ValuesRAG: Enhancing Cultural Alignment Through Retrieval-Augmented Contextual Learning [Preprint]. arXiv. https://arxiv.org/abs/2501.01031

SINACOLA E., PACHOT A., PETIT T. (2025). LLMs, Virtual Users, and Bias: Predicting Any Survey Question Without Human Data [Preprint]. arXiv. https://arxiv.org/abs/2503.16498

TAO Y., VIBERG O., BAKER R.S., KIZILCEC R.F. (2025). Cultural bias and cultural alignment of large language models; *PNAS Nexus*, Volume 3, Issue 9, September 2024, pgae346, https://doi.org/10.1093/pnasnexus/pgae346

VALETTE, M. (2024). What Does Perspectivism Mean? An Ethical and Methodological Countercriticism, *NLPerspectives 2024@LREC-COLING 2024*, pp. 111-115.

WANG W., JIAO W., HUANG J., DAI R., HUANG J.-T., TU Z., LYU M. R. (2023). Not all countries celebrate thanksgiving: On the cultural dominance in large language models. arXiv:2310.12481v2 [cs.CL]