Regulation of AI? Comparing Czech and Portuguese Media Imaginaries with CADS

Raquel Amaro, Tibor Vocásek
CLUNL NOVA Lisbon, Portugal; ISS – FSV, CUNI Prague, Czechia
raquelamaro@fcsh.unl.pt; tibor.vocasek@fsv.cuni.cz

Paper Abstract

The recent buzz around Artificial Intelligence (AI) has raised significant debate in the EU on how to regulate it. Like any emerging technology, AI development depends on media-shaped public perception (Chuan et al., 2019). This research aims to contribute to the study of this perception through an interdisciplinary comparative perspective. It analyses AI regulation representations in the Czech and Portuguese online mainstream media, using Corpus Approaches to Discourse Studies (CADS) (Baker et al., 2008; Partington et al., 2013).

Czechia and Portugal are similar in area size, population, or GDP (Eurostat, 2024) but differ significantly regarding their tech sectors. While news headlines labelled Czechia the "sick man of Europe" due to its stuttering industrial economy (Willoughby, 2023), Portugal, with its vivid tech scene, was referred to as "Europe's Silicon Valley" (Böhnisch, 2021). Focusing on the EU debate, these countries' membership differs by almost two decades.

CADS combines corpus linguistics with traditional CDA while reflecting on its critiques (Orpin, 2005). It mitigates issues of data representativeness and interpretative transparency. CADS allows the investigation of the aggregate effects of language, highlighting typical discursive patterns. Conceptually, this study approaches AI regulation through "sociotechnical imaginaries", understood as "collectively held, institutionally stabilised, and publicly performed visions of desirable futures" (Jasanoff and Kim, 2015, p. 4).

Current studies using imaginaries are plagued by conceptual ambiguity (Richter et al., 2023). This research overcomes the issue by adopting a three-level imaginary concept (Sau, 2021) operationalised with CADS. It structures imaginary analysis by asking for representations of (1) social commentary, (2) vision of the future, and (3) means to achieve it. The analysis covers the period of discussions about the EU's "AI Act" regulation (3/2018-12/2023).

Comparable corpora, collected from digitally available national media in each country in this period, are compiled and explored using Sketch Engine (Kilgarriff et al., 2014) in two steps. Firstly, distinct imaginary layers are investigated via collocations of the topic-related keywords and analysed via concordances in each corpus. Secondly, the analysis focuses on keywords of the sub-corpus after the ChatGPT onset to reveal specifics of the "ChatGPT-moment". The results of both steps are then compared against each other.

Although not groundbreaking in terms of methodology, such research provides innovative, empirically rooted comparative insights into the current media debate on AI. It exhibits the usefulness of the imaginary concept for the CADS while providing a clearer perspective of sociotechnical imaginaries by grounding these to objective linguistic cues.

Keywords: Artificial Intelligence, Corpus Approaches to Discourse Studies, Media Discourse, Regulation of Digital Technologies, Sociotechnical Imaginaries

REFERENCES

Baker, P., Gabrielatos, C., KhosraviNik, M., Krzyżanowski, M., McEnery, T., & Wodak, R. (2008). A useful methodological synergy? combining critical discourse analysis and corpus linguistics to examine discourses of refugees and asylum seekers in the UK press. *Discourse & Society*, *19*(3), 273–306. https://doi.org/10.1177/0957926508088962.

Böhnisch, M. (2021). *Portugal: Europe's Silicon Valley?* Deutsche Welle. https://www.dw.com/en/portugal-europes-silicon-valley/video-57359402.

Chuan, C.-H., Tsai, W.-H. S., & Cho, S. Y. (2019). Framing Artificial Intelligence in American newspapers. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. https://doi.org/10.1145/3306618.3314285.

Eurostat. (2024). *Digital Economy and Society*. https://ec.europa.eu/eurostat/web/digital-economy-and-society.

Jasanoff, S., & Kim, S.-H. (2016). *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power*. The University of Chicago Press.

Kilgarriff, A., Baisa, V., Bušta, J., Jakubíček, M., Kovář, V., Michelfeit, J., Rychlý, P., & Suchomel, V. (n.d.). *The Sketch Engine: Ten Years on*. Lexicography. https://journal.equinoxpub.com/lexi/article/view/17961.

Orpin, D. (2005). Corpus linguistics and critical discourse analysis. *International Journal of Corpus Linguistics*, *10*(1), 37–61. https://doi.org/10.1075/ijcl.10.1.03orp.

Partington, A. (2010). Modern diachronic corpus-assisted discourse studies (MD-CADS) on UK newspapers: An overview of the project. *Corpora*, *5*(2), 83–108. https://doi.org/10.3366/cor.2010.0101.

Richter, V., Katzenbach, C., & Mike, S. (2023). Imaginaries of AI. In Simon Lindgren (Eds.), *Handbook of Critical Studies of Artificial Intelligence*. Preprint. Edward Elgar Publishing Ltd. https://doi.org/10.26092/elib/2190.

Sau, A. (2021). On cultural political economy: A Defence and constructive critique. *New Political Economy*, *26*(6), 1015–1029. https://doi.org/10.1080/13563467.2021.1879758.

Willoughby, I. (2023, December 11). *Economist Tomáš Dvořák on why Czechia is "Sick man of europe" – and how to move forward*. Radio Prague International. https://english.radio.cz/economist-tomas-dvorak-why-czechia-sick-man-europe-and-how-move-forward-8802580.

Acknowledgements

This study was supported by the Charles University, project GA UK No. 294723.

Part of this research is supported by the Portuguese national funding through the FCT – Portuguese Foundation for Science and Technology, I.P. as part of the project UIDB/LIN/03213/2020; 10.54499/UIDB/03213/2020 and UIDP/LIN/03213/2020; 10.54499/UIDP/03213/2020 – Linguistics Research Centre of NOVA University Lisbon (CLUNL).