



ÜSKÜDAR UNIVERSITY FACULTY OF COMMUNICATION
UNIVERSITY OF TRÁS-OS-MONTES AND ALTO DOURO

New Paradigms in Communication Technologies and Humanity Symposium

**ARTIFICIAL INTELLIGENCE: TECHNOLOGICAL,
SOCIAL, AND CULTURAL TIES**

6-7 March 2025

Proceedings Book



utad UNIVERSIDADE
DE TRÁS-OS-MONTES
E ALTO DOURO



ÜSKÜDAR
ÜNİVERSİTESİ
YAYINLARI - 117

**Üsküdar University Faculty of Communication New Paradigms in Communication
Technologies and Humanity Symposium Proceedings Book**

Editors

Prof. Nazife Güngör
Prof. Emídio Gomes

Co-Editors

Prof. Orquídea Moreira Ribeiro
Prof. Gül Esra Atalay
Assoc. Prof. Fábio Ribeiro
Assoc. Prof. Bahar Muratoğlu Pehlivan
Res. Asst. Enes Akdağ
Res. Asst. Semih Eken

Proofreading

Res. Asst. Enes Akdağ

Graphic Design

Res. Asst. Simgе Kırteke

ISBN

978-625-93668-5-2

Publishing Certificate Number

69591

Year of Publication

2026

Contact Information:

www.uskudar.edu.tr - yayin@uskudar.edu.tr
Phone: +90216 400 22 22 / Fax: +90216 4741256
Altunizade Mah. Üniversite Sk. No: 14
Pk: 34662 Üsküdar / İstanbul / Türkiye

Copyright © 2026

In accordance with the Law on Intellectual and Artistic Works, the publication rights of this work belong to Üsküdar University. All rights reserved. It can be quoted by showing the source. No part of this book may be reproduced in any form, electronic or mechanical, including photocopying, recording or any information storage or retrieval system, without the written permission of the publisher. The responsibility for the content in the book belongs to the author.

This book is compiled from the abstracts of the New Paradigms in Communication Technologies and Humanity, Artificial Intelligence: Technological, Social, and Cultural Ties Symposium, and the responsibility for all opinions and ideas contained in the texts belongs to the author of the paper.

ÜSKÜDAR UNIVERSITY FACULTY OF COMMUNICATION

New Paradigms in Communication Technologies and Humanity Symposium

**ARTIFICIAL INTELLIGENCE: TECHNOLOGICAL,
SOCIAL, AND CULTURAL TIES**

Proceedings Book

6-7 March 2025

<https://newpath.uskudar.edu.tr/>

Symposium Organization Committee

Professor Orquídea Moreira RIBEIRO
Professor Süleyman İRVAN
Professor Gül Esra ATALAY
Associate Professor Bahar MURATOĞLU PEHLİVAN
Associate Professor Derya Gül ÜNLÜ
Associate Professor Fábio RIBEIRO
Associate Professor Susana FONTES
Associate Professor Aysun AYDIN
Assistant Professor João Pedro BAPTISTA
Assistant Professor Marlene LOUREIRO
Assistant Professor Öznur KARAKAŞ
Assistant Professor Zindan ÇAKICI
Research Assistant Enes AKDAĞ
Research Assistant Onur COŞKUN
Research Assistant Semih EKEN
Research Assistant Simge KIRTEKE
Dr. Sinan AŞÇI

Symposium Scientific Committee

Professor Burcu EKER, Bahçeşehir University / Türkiye
Professor Carlos ASSUNÇÃO, University of Trás-os-Montes and Alto Douro (UTAD) / Portugal
Professor Deniz YENİĞİN, Çanakkale Onsekiz Mart University / Türkiye
Professor Elif ENGİN, Bahçeşehir University / Türkiye
Professor Erkan SAKA, Istanbul Bilgi University / Türkiye
Professor Fernando MOREIRA, University of Trás-os-Montes and Alto Douro (UTAD) / Portugal
Professor Helena LOBO, University of Trás-os-Montes and Alto Douro (UTAD) / Portugal
Professor Orquídea Moreira RIBEIRO, University of Trás-os-Montes and Alto Douro (UTAD) / Portugal
Professor Özgül DAĞLI, Üsküdar University / Türkiye
Professor Uğur GÜNDÜZ, Istanbul University / Türkiye
Professor Zeynep ÖZARSLAN, Çukurova University / Türkiye
Associate Professor Aysun EYREK, Fenerbahçe University / Türkiye
Associate Professor Balca ARDA, Kadir Has University / Türkiye
Associate Professor Bilge ŞENYÜZ, Marmara University / Türkiye
Associate Professor Cihan BECAN, Üsküdar University / Türkiye
Associate Professor Ferhat ZENGİN, Istanbul Gelişim University / Türkiye
Associate Professor Gökhan MURA, Izmir University of Economics / Türkiye
Associate Professor Maria José BRITES, Lusófona University / Portugal

Associate Professor Özge UĞURLU AKBAŞ, Üsküdar University / Türkiye
Associate Professor Susana FONTES, University of Trás-os-Montes and Alto Douro (UTAD) / Portugal
Assistant Professor Ceren SARAN DOĞAN, Üsküdar University / Türkiye
Assistant Professor Marlene LOUREIRO, University of Trás-os-Montes and Alto Douro (UTAD) / Portugal
Assistant Professor Nuri BİNGÖL, Üsküdar University / Türkiye
Assistant Professor Reşat Fuat ÇAM, İstanbul Bilgi University / Türkiye
Assistant Professor Rute BASTARDO, University of Trás-os-Montes and Alto Douro (UTAD) / Portugal
Dr. Luís PEREIRA, City, University of London / The United Kingdom
Dr. Madalena OLIVEIRA, University of Minho / Portugal

Keynote Speakers

Helena SOUSA

President Portuguese Regulatory Authority for the Media (ERC)
“Artificial Intelligence, Regulation and Journalism: A Pragmatic Approach”

Halil NALÇAOĞLU

Istanbul Bilgi University
“From Inquiry to Probability: The Transformation of Knowledge and Education in the Age of AI”

Raquel RECUERO

Federal University of Pelotas (Brazil)/ Federal University of Rio Grande do Sul (Brazil)
“Big Data Infrastructures and Cultural Narratives”

FOREWORD

Üsküdar University Faculty of Communication aims to make significant contributions to the field of communication sciences not only at the national but also international level with the academic studies and scientific meetings it presents. The symposium titled “Artificial Intelligence: Technological, Social, and Cultural Ties,” organized in collaboration with the University of Trás-os-Montes and Alto Douro in Portugal on March 6-7, 2025, is an important example of these efforts.

Held for the first time this year, the symposium is intended to be branded under the main title “New Paradigms in Communication Technologies and Humanity” (NEWPATH) and is planned to be organized periodically from now on. The symposium, which was held entirely in English with the presentation of papers by around fifty academics and researchers from various countries, was focused on artificial intelligence. The transformative effects of artificial intelligence on social and cultural life were discussed from multiple perspectives.

The general conclusion from the papers presented at the symposium is that artificial intelligence will be integrated into all areas of life in various ways in the very near future and will lead to significant transformations. Based on these predictions, interesting discussions were held on how the field of communication sciences and media will be affected by this transformation.

Reviewing the presentations and opinions shared at the symposium, it can be said that digital technologies and artificial intelligence, which have entered into use in various parts of our lives today, are candidates to change our entire lifestyle in the near future. This shows that humanity has entered a new evolution process not only in terms of lifestyle but also in terms of mentality. Based on these predictions, there is a need for significant changes focused on artificial intelligence in education and training, especially in communication education, as it concerns the field of communication. Significant changes are expected in the near future in both human and social communication forms and in the media profession. Necessary measures need to be taken now so that the communication forms of the future, which are expected to be surrounded by artificial intelligence, shape human and social interaction in a positive way. On the other hand, artificial intelligence has already begun to be used in the media sector. This shows that the media sector of the future will be shaped largely by the integration of artificial intelligence. Therefore, the curricula currently being implemented in communication education need to be reviewed and adapted to the age of artificial intelligence.

The NEWPATH Symposium can be described as a scientific event that contributes to analyzing the current state of communication and media, monitoring ongoing developments and effectively managing preparations for the future. As it did this year, NEWPATH will continue to serve as an international interaction platform for various academic circles, especially communication scientists. I would like to thank all colleagues who contributed to the organization of the symposium, and contributed with their paper presentations and as keynote speakers. Perhaps we will come together face to face at the next NEWPATH. See you at the next NEWPATH.

Professor Nazife Güngör
Rector of Uskudar University

Professor Emídio Gomes
Rector of University of Trás-os-Montes and Alto Douro

Contents

(Surname in alphabetical order)

| | |
|---|----|
| YOU SHALL SPEAK MY LANGUAGE: MIND THE LANGUAGE GAP IN THE ERA OF AI..... | 11 |
| Maria Pia Ester CRISTALDI | |
| HUMAN-ARTIFICIAL INTELLIGENCE INFLUENCER INTERACTION IN THE DIGITAL WORLD: ANALYZING PERCEPTION AND EMOTIONAL EXPERIENCES WITH PSYCHOPHYSIOLOGICAL DATA HARVESTING TECHNIQUE..... | 27 |
| Yaren DEMİREL, Neslihan ERDEM | |
| WHEN SKIN IS REPLACED BY STEEL: ROBOTS, L.L.M.S, AND THE REIMAGINATION OF CULTURAL AND SOCIAL NORMS..... | 40 |
| Sara PASCOAL, Rosa MESQUITA | |
| THE ROLE OF MEDIA ORGANIZATIONS IN THE AGE OF ARTIFICIAL INTELLIGENCE: THE DUTY AND RESPONSIBILITY TO PROTECT TRUST IN MEDIA | 66 |
| Valdone RUDENKIENE | |
| AN ANALYSIS OF OPINIONS ON THE USE OF GENERATIVE MODELS IN THE FILM INDUSTRY THROUGH YOUTUBE USER COMMENTS: THE FROST AND AIR HEAD | 78 |
| Yağmur TANRIVERDİ | |
| THE CONSTRUCTION OF SOCIAL REPRESENTATION MEDIATED BY ALGORITHMS | 93 |
| Aline VERONIZE, Luís LOUREIRO, Alexandra FIGUEIRA | |

YOU SHALL SPEAK MY LANGUAGE: MIND THE LANGUAGE GAP IN THE ERA OF AI

Maria Pia Ester CRISTALDI*

Abstract

A comprehensive discussion about AI models cannot disregard the social and cultural aspects of natural language processing. Concerning these aspects, according to *Ethnologue*, 7,168 languages are spoken worldwide. Notwithstanding this, only about 20 have enough training data online to create natural language processing AI systems. In this regard, data published by *Statista* shows that in 2024, English is the most used language (52%) for web content, followed by Spanish (5.5%) and German (4.8%). Since these are the most used languages for content creation, most AI tools are accessible to people who can speak English, Spanish and German. At the same time, these represent the languages spoken in the most technologically and economically advanced countries in the world. This aspect raises a few concerns. On the one hand, the question arises concerning why widely spoken languages such as Hindi and Arabic are so little used in AI development. On the other hand, the latter raises further preoccupations regarding the access that speakers of these languages have to artificial intelligence in everyday life. Through data analysis concerning the most used languages for AI development and the debates concerning this topic, this paper aims to discuss the role artificial intelligence could have in bridging the existing cultural, social and economic gap between languages. In this regard, this research aims to open further venues of discussion on the role that new technologies could play in shaping a more equitable and sustainable world.

Keywords: language gap, artificial intelligence, natural language processing.

* Assistant Professor, Üsküdar University, mariapia.cristaldi@uskudar.edu.tr, ORCID:0000-0002-7724-9723

Introduction

Emerging in the 1950s, thanks to the work of Alan Turing, McCulloch and Pitts, artificial intelligence became a field of study in 1956 during a workshop held at Dartmouth College. During the work and discussions that emerged during the workshop, the prerequisites for the revolution that artificial intelligence has brought to our lives emerged. There is no doubt that artificial intelligence now concerns every area of our lives. From search engines to virtual assistants, from recommendation systems to natural language processing, the areas of learning, knowledge and daily life in which we use artificial intelligence are increasing daily.

One aspect that often goes unnoticed when we think about artificial intelligence and the devices that use this technology is the languages in which these devices speak to us and communicate with us. In this context, when we think of AI and technology, English is the first language that often comes to mind. However, AI and modern technologies are now an integral part of everyone's lives, regardless of their proficiency in English. Therefore, artificial intelligence must be accessible and usable for those who do not speak English.

In this context, we aim to explore how and to what extent languages other than English are used to train artificial intelligence. Additionally, we will examine ways to overcome these challenges to ensure that technology and AI are accessible to everyone. This study aims to answer these questions and analyse initiatives that promote the use of languages other than English in training AI models.

Language Representation and the Web

According to data from Statista, the ranking of the most frequently used languages for web content, as of February 2025 (by the share of websites), is as follows: 1. English (49.4%); 2. Spanish (6%); 3. German (5.6%); 4. Japanese (5%); 5. French (4.4%); 6. Russian (3.9 %); 7. Portuguese (3.8%); 8. Italian (2.7%); 9. Dutch/Flemish (2.1%); 10. Polish (1.8%); Turkish (1.8%).

However, a stark contrast emerges when we compare these statistics with the most spoken languages in the world. According to *Ethnologue*, as of the early 2020s, the world's top ten most spoken languages are as follows: 1. English (1,456,448,320); 2. Mandarin Chinese (1,138,222,350); 3. Hindi (609,454,770); 4. Hindi (609,454,770); 4. Spanish (559,078,890); 5. French (309,804,220); 6. Modern Standard Arabic (273,989,700); 7. Bengali (272,828,760); 8. Russian (254,997,130); 9. Portuguese (263,638,850); 10. Urdu (231,717,940). These data show that widely spoken languages such as Mandarin Chinese, Hindi, Bengali, and Urdu are not among the most frequently used languages for writing web content.

Additionally, this data demonstrates that widely spoken languages like Mandarin Chinese, Hindi, Bengali, and Urdu are not among the most commonly used languages for writing web content. This has raised the concerns of United Nations' Advocacy Specialists such as Alena Gorbacheva, who, in an article written in 2023 and entitled "No Language Left Behind: How to Bridge the Rapidly Evolving AI-Language Gap", shared her preoccupations concerning languages such as Kazakh. Widely spoken in Kazakhstan, where over 13,768,000 residents (more than 80%) reported being proficient in the State language. Notwithstanding, there is insufficient web content written in Kazakh; consequently, AI devices are not trained enough in this language.

High- Resource and Low- Resource Languages

It is essential to consider the distinction between "high resource" and "low resource" languages to understand why AI models are trained more in some languages and less in others. This distinction is important for natural language processing studies. To build natural language processing systems, it is vital to have enough data to train models in a specific language. As noted by Felix Laumann*, although about 7000 languages are spoken worldwide, only about 20 have text corpora of hundreds of millions of words. English is the language with the most significant amount of data, followed by Chinese and Spanish.

* <https://medium.com/neuralspace/low-resource-language-what-does-it-mean-d067ec85dea5>

As shown by the data from Statista mentioned above, other languages with large datasets include Western European languages, but also Japanese. These languages are considered “high-resource languages”. On the other hand, many languages in Asia and Africa lack the training data needed to develop accurate state-of-the-art NLP systems, referred to as low-resource languages.

In this regard, as reported by Gorbacheva, Hindi, spoken by over 500 million people, is also considered a “low resource” language. Western European languages, like Dutch, are spoken by 20 times fewer people, but are instead highly used to write web content. Less than 1% of languages worldwide are “high resource”. The paradox here is evident; however, to understand the reasons for this disparity, it is essential to know how the datasets used to train natural language processing systems are created. To create a dataset we need gigabytes, if not terabytes, of data—often corresponding to billions of written sentences—manually creating datasets for these models is not practical. It would take years to compile such extensive datasets, and few organisations would be willing to pay hundreds of employees to write random sentences in a document.

The primary source for most large datasets comes from web scraping, primarily from social media platforms such as Facebook and Twitter, where users worldwide generate billions of comments and posts. This data is effectively free for use when training language models. However, it raises an important question: in which languages do people communicate online? How does the volume of tweets in English compare to those in Hindi? Furthermore, looking back at historical data, when was the first instance of user-generated content uploaded to the internet, and what was its origin? Typically, the answers to these inquiries point to “English” and “the USA.” Consequently, this differentiation leads us to categorize some languages as having high resources (abundant datasets) and others as having low resources (limited datasets).

Current natural language processing (NLP) solutions predominantly focus on a few high-resource languages, despite the existence of approximately 3 billion speakers of low-resource languages, primarily in Asia and Africa. Consequently, since artificial intelligence tools are built on predicting the most

likely response based on a vast amount of “training data”, when the latter is insufficient, as with “low resource” languages, AI tools either do not work or perform poorly.

Social and Educational Implications of AI Tools’ Poor Performance in Low-Resource Languages

The lack of sufficient resources to create datasets for training language models and AI systems has significant social and educational implications. The disparity between high-resource and low-resource languages means that 99% of speakers are effectively excluded from global technological advancements. This situation poses a risk that AI, which is intended to revolutionise and simplify lives, may instead exacerbate the divide between low-resource and high-resource languages, making life more challenging for the disadvantaged. This issue is particularly critical, as discussions about language—both in social contexts and in the realm of information technology—must consider access to information, education, and the potential for economic and intellectual growth.

To better understand this aspect, it may be helpful to consider the cases of Kazakhstan and India. Concerning Kazakhstan, given the historical, ethnic, and social context of the country, the two most spoken languages in the country are Russian and Kazakh. According to a census published in 2021*, the Kazakh language is proficiently spoken by 80.1% of the population and has the status of “state language”. On the other hand, Russian is spoken by 83.7% as of 2021. Research indicates that since the dissolution of the Soviet Union in 1991, the Kazakh State has been actively promoting Kazakh as the primary language in state affairs, culture, and education. However, the rate of increase in the usage of Kazakh is not as rapid as it once was. For example, in 2024, approximately 75% of students graduating and taking the Unified National Testing (UNT) chose to take the exam in Kazakh, while 25% opted for Russian, and only 0.2% took it in English. In comparison, ten years earlier, in 2014, around 70% of school graduates took the UNT in Kazakh (National Test Center, 2024).

* “National census 2021- Agency for Strategic planning and reforms of the Republic of Kazakhstan Bureau of National statistics”. stat.gov.kz. Retrieved 16 April 2025.

While there were significant changes in Kazakhstan's educational linguistic landscape during the 1990s and 2000s, marked by a decline in Russian-language education and an increase in Kazakh-language education, this trend has begun to slow in recent years.* According to Kucherbayeva and Smagulova, Russian has historically served as the official language of academia and culture, which makes the revitalisation of a marginalised language like Kazakh and its establishment as an official language in these domains quite challenging. In their research concerning students' access to educational resources in Kazakh, the authors highlight the limited availability of teaching and reading materials in the Kazakh language. The students interviewed for the study indicated that their teachers predominantly utilize and refer to learning resources in Russian. Additionally, the participants noted that they rely on these resources due to Kazakh's scarcity of information and library materials. A significant issue raised by respondents is that numerous instructors are native Russian speakers; they were educated in Russian, making it their primary professional language. Additionally, students interviewed stated that being multilingual is anticipated and frequently assumed. Every student mentioned that they were consistently given materials in Russian and expected to translate them from Russian to Kazakh independently**. Due to these complex educational issues, around 20% of citizens are not proficient in either of the two official languages. This raises concerns about their access to the educational resources and information available online for Kazakh citizens today.

The situation is further complicated in India, which boasts a significantly more diverse and fragmented linguistic landscape than Kazakhstan. India's situation differs from Kazakhstan's because it has no official language.***

* Trends of Language Shift in Education in Kazakhstan (<https://www.eurasian-research.org/publication/the-trends-of-language-shift-in-education-in-kazakhstan/#:~:text=Since%20independence%20in%201991%2C%20the,in%20public%20and%20official%20domains>).

** Kucherbayeva, D., & Smagulova, J. (2023). Language Revitalization: Challenges for Kazakh in Higher Education. *Journal of Eurasian Studies*, 14(2), 166-178. <https://doi.org/10.1177/18793665231189326> (Original work published 2023).

*** As Pavan Mandakar outlines, there is no National language in India. However, Article 343(1) of the Indian Constitution mentions explicitly that the official language of India shall be Hindi in Devanagari script (Mandavkar, Pavan, Role of Languages in National Education System of India (October 21, 2023). Available at SSRN: <https://ssrn.com/abstract=4609001> or <http://dx.doi.org/10.2139/ssrn.4609001>)

According to Census India 2011*, Hindi is the first language as per most significant number of speakers, followed by English, which is spoken by 129 million of speakers, Bengali (107), and Marathi (99) Additionally, 23 other languages—including Urdu, Nepali, Punjabi, and Sanskrit- are among the most spoken languages in India. To address the linguistic differences among the various ethnic groups in India, the Indian government has adopted English as the primary medium of instruction in both educational and professional settings. As Boorah and Sabharwal point out, while this decision improves job and career opportunities for Indian citizens looking for overseas jobs, it does not resolve the inequalities in access to educational opportunities for speakers of languages other than English.** Moreover, Amitabh Vikram Dwivedi outlines that English-medium schools in India are commonly referred to as convent schools or public schools, and attending such institutions is often regarded as indicative of receiving a quality education***. Moreover, English remains a key tool for national governance, interregional communication, and international discourse. In contrast, while other languages spoken in India are more accessible to the general population and are thriving, they do not have equal representation in official contexts such as institutions, schools, and universities, where English is still the prominent language of instruction****. This poses a serious threat also in terms of availability of educational resources.

In this regard, when discussing educational resources today, we no longer limit ourselves to traditional textbooks, encyclopaedias, and manuals. The current teaching and learning landscape include online research, videos on platforms like YouTube, and, more recently, AI tools. Students from around the globe increasingly rely on these tools for their assignment research.

* For a more detailed account, please check Census India 2011- Language Atlas: <https://censusindia.gov.in/nada/index.php/catalog/42561>

** Vani Boorah, Nidhi Sabharwal: English as a Medium of Instruction in Indian Education: Inequality of Access to Educational Opportunities https://mpr.ub.uni-muenchen.de/112984/1/MPRA_paper_112984.pdf

*** Dwivedi, A. V. (2015). The official status of English in India. *The Journal of English Language Teaching (INDIA)*, 57(6), 10–14.

**** Dwivedi, Amitabhvikram. (2025). Linguistic Inequality in India: A Survey of Indian Language Policies. 10.1007/978-3-030-68127-2_476-1

However, many of the most popular AI tools are built on large language models that predominantly support the English language, which means they often do not yield optimal results in other languages. In the specific contexts of Kazakhstan and India, how can a Kazakh or Indian student with limited English proficiency fully leverage AI's potential for their education?

Multilingual Language Models to Bridge the Language Gap

In such linguistic, social, and ethnic situations, educational tools that exploit the potential of AI could play a key role in bridging the gap between individuals first and countries then in terms of knowledge, know-how, and innovation. The challenge lies not in creating web content in Kazakh, Hindi, or Arabic to train AI models. Instead, the focus should be on investing in different language models to train AI effectively. In this context, the researcher Javier Saiz outlines that AI models trained with multilingual datasets are essential in bridging the language gap. According to researchers working on multilingual models, it is no longer adequate to supply data to the AI model in just one language; it is now crucial to present it in multiple languages simultaneously. Furthermore, multilingual AI systems represent a significant advancement, as they transcend mere translation. With a strong emphasis on the context in which words and sentences are employed, these systems can capture the cultural nuances that words convey across different languages. In this regard, AI multilingual datasets are an essential step toward improving language models' abilities to understand and process a variety of languages, including less common ones such as Swahili and Arabic. This initiative, known as Multilingual Massive Multitask Language Understanding (MMMLU), a benchmark that covers 57 different tasks in subjects such as Science, Technology, Engineering, Mathematics, Humanities, Social Sciences, Medicine, and Business. MMLU is particularly valuable for the research and development of AI systems in languages other than English and the most widely spoken languages worldwide because it allows for the creation of benchmarks that function across multiple languages*.

* For a more detailed account on MMLU and language understanding see: Hendrycks, D., Burns, C., Basart, S., Zou, A., Mazeika, M., Song, D., & Steinhardt, J. (2020). Measuring massive multitask language understanding. arXiv preprint arXiv:2009.03300.

In this sense, MMLU specifically tackles a significant issue: many AI models are predominantly trained on English, resulting in subpar performance in other languages.

In this regard, Li and others measured massive multitask language understanding in Chinese. The researchers evaluated 18 multilingual and Chinese-oriented large language models (LLMs), assessing their performance across different subjects and settings. The results revealed that most existing LLMs struggle to achieve an average accuracy of 50%, even when provided with in-context examples and chain-of-thought prompts. On the contrary, the study demonstrated that MMLU models perform better and can fill in the gaps in LLMs in Chinese*. In a similar way, a study by Koto and others on MMLU models for the Arabic language confirmed that MMLU is the right choice to improve AI training in Arabic**.

In recent years, the projects that have invested in MMLU have mainly been those developed in the field of Open-Source AI. As demonstrated by the examples mentioned above, research in this field has grown a lot in recent years, and there are more and more studies that analyse the compatibility of these models with languages other than those most used for the development of artificial intelligence systems.

Leon Eversberg from the Technical University of Berlin highlights that obtaining substantial data to evaluate the performance of language models in languages other than English remains a challenge in the present day too. To address this issue, Eversberg proposes further research on the Global-MMLU***, an extension of the MMLU benchmark that includes translations in 42 languages.

* Li, H., Zhang, Y., Koto, F., Yang, Y., Zhao, H., Gong, Y., ... & Baldwin, T. (2023). CMMLU: Measuring massive multitask language understanding in Chinese. arXiv preprint arXiv:2306.09212.

** To conduct this study, Koto and others used a multilingual task understanding benchmark for Arabic that incorporated sources from school exams in different countries in North Africa, the Gulf region and the Levant. For a more detailed account see Koto, F., Li, H., Shatnawi, S., Doughman, J., Sadallah, A. B., Alraeesi, A., ... & Baldwin, T. (2024). Arabicmmlu: Assessing massive multitask language understanding in Arabic. arXiv preprint arXiv:2402.12840.

*** Leon Eversberg, <https://medium.com/data-science/how-to-evaluate-multilingual-llms-with-global-mmlu-ce314aedee8f> retrieved 27.04.2025

These translations come from a combination of professional work, community contributions, and machine translations. This dataset enables the evaluation of large language models (LLMs) across various languages, while also considering specific cultural and linguistic nuances.

In his research, Eversberg tested the model using 1,000 questions in different languages, yielding the following accuracy rates: English: 43.3%, Spanish: 38.5%, Chinese: 38.5%, German: 34.6%, and Japanese: 33.7%. These results indicate a decline in performance for languages other than English, underscoring the necessity for language-specific assessments. Additionally, since the Global-MMLU includes translations in less commonly used languages in the IT context—such as Czech, Persian, Telugu, Ukrainian, Hausa, Kyrgyz, and Somali—it would be valuable to examine the model’s performance and the accuracy of these translations in those languages as well.

More recently, Xuan and other researchers developed MMLU-ProX, a comprehensive multilingual benchmark covering 13 typologically diverse languages with approximately 11,829 questions per language. Experiments have consistently shown a performance decline from high-resource languages to lower-resource languages. While the best models achieve over 70% accuracy in English, their performance drops to around 40% for languages like Swahili. Despite recent advancements, this highlights ongoing gaps in multilingual capabilities. However, the MMLU-Prox project is currently underway; therefore, the researchers are expanding the benchmark by adding more languages and evaluating additional language models*.

Although these initiatives are crucial, it remains true that research in this field has primarily relied on individual efforts or, at best, small research groups. For artificial intelligence- based on Large Language Model (LLM) or Multilingual Massive Multitask Language Understanding Model (MMLU)- to become a tool accessible to speakers of languages worldwide, we need collective efforts supported by governments and international organisations.

* Xuan, W., Yang, R., Qi, H., Zeng, Q., Xiao, Y., Xing, Y., ... & Li, I. (2025). MMLU-prox: A multilingual benchmark for advanced large language model evaluation. arXiv preprint arXiv:2503.10497.

A joint approach should include joint development of technologies and educational programmes, implementation and testing of pilot projects, and drawing attention to the results through publications and educational activities.

Fostering the development of AI technologies for “low resource” languages will benefit individual nations and the global community. Preserving linguistic diversity online and in AI is just as crucial as in the physical world, as it enables us to maintain a rich array of cultures and overcome educational and economic gaps, thus building a more equal and just world.

In this sense, research and transnational bodies such as the European Union have invested in recent years in developing artificial intelligence models for the languages of ethnic minorities and other languages at risk of extinction.

AI for Linguistic Diversity

The Committee on Artificial Intelligence (CAHAI) of the Council of Europe has examined possible elements of a legal framework on artificial intelligence based on the Council of Europe’s standards on human rights, democracy and the rule of law. In 2022, the Committee published a report entitled “Facilitating the implementation of the European Charter for Regional or Minority Languages through artificial intelligence”*. The latter resulted from the studies conducted between 2019 and 2021 by CAHAI on artificial intelligence’s role in improving the quality of life and education, as well as the impact it would have on the participation in public life of speakers of minority languages. The Committee’s work outlined those existing tools (like Google Translate, DeepL) already support some minority languages; new domain-specific translation tools can be developed for others at a relatively low cost for countries and governments using existing technology. Similarly, social media such as Facebook and Instagram offer interfaces in regional and minority languages, thus promoting their everyday use. As with translation services, work could be done to develop interfaces in other regional or minority languages.

* For the full version of the report, check Facilitating the implementation of the European Charter for Regional or Minority Languages through artificial intelligence

In this regard it is important to consider that application such as Duolingo or Clozemasters offer the possibility to learn minority languages (Scottish-Gaelic, Yiddish, Welsh) and regional languages (Ukrainian, Macedonian, Catalan) as well. As outlined in the report, countries where these minority or regional languages are spoken can utilize these opportunities by either enhancing the visibility of existing online language courses via these apps or by developing new language courses or vocabulary cards for regional or minority languages that are not currently included in any language learning application on the platforms. This would ensure that language learning resources reach a broader audience of potential learners, boosting the language's prominence.

Furthermore, artificial intelligence could ensure the use of minority and regional languages also in aspects of daily life. Among the examples cited in the report, the report mentioned about implementing speech synthesis for bilingual announcements in public transport. A substantial database of pre-recorded speech is essential for speech synthesis, making it particularly relevant for regional or minority languages that already possess an existing database, often due to their status as the majority language in other nations. In this regard, encouraging transnational studies and research exchanges is beneficial, as it allows for the sharing of existing databases from majority language countries with those where the language is considered a regional or minority one.

Among other possible applications of AI for the preservation and diffusion of regional and minority languages, the Committee also focused on automatic subtitle generation (e.g. Google's automated speech recognition algorithm, YouTube's automated caption timing), as it can help make television programmes in those languages accessible to people with hearing impairments. Moreover, subtitle translation promotes the visibility of regional or minority languages. Other possible applications include automatic information extraction and sentiment analysis. In the first case, investing in automatic information extraction could help in tasks such as summarising news articles in a regional or minority language. Concerning sentiment analysis, since the latter is mainly used for customer feedback analysis, developing programmes for sentiment

analysis of regional or minority languages would mean no longer excluding but including the speakers of these languages. Implementing all these suggestions would ensure that these languages could be fully functional, as they would be used in every aspect of economic and social life.

Conclusion

In this research we have tried to examine how artificial intelligence could be used to overcome the gap between low-resource and high-resource languages that currently exists both on the web and in the context of the artificial intelligence tools we use every day. As we have tried to highlight in this work, investing in artificial intelligence and encouraging the development of specific tools, based on the use of low-resource languages and minority languages, could prove to be fundamental to promote the use of these languages and guarantee their speakers equal access to the world of communication, education and technology. With this study, we tried to highlight how to invest in developing technologies that encourage minority languages and low-resource Languages cannot be left to individual initiatives, but it must be the result of targeted national and supranational policies. For this purpose, the European Union and the Committee on Artificial Intelligence efforts are significant.

In this sense, the future is promising as more and more researchers invest in training AI models based on multilanguage models. Furthermore, as Lamentillo* points out, initiatives already use AI to save both minority and endangered languages. Among the initiatives cited by Lamentillo is Woolaroo, a Google project that uses artificial intelligence to help communities document their endangered languages by allowing users to take photos of objects and receive translations in their native tongue. This project is exciting because it would also allow those who are unable to speak minority languages to learn them, thus ensuring not only their preservation but also their diffusion. Similarly, the NightOwlGPT project provides real-time translation for marginalized languages

* Anna Mae Yu Lamentillo, AI: The Unexpected Hero in the Battle to Save Dying Languages. How Artificial Intelligence is Rescuing Endangered Languages in <https://apolitical.co/solution-articles/en/ai-the-unexpected-hero-in-the-battle-to-save-dying-languages-844> retrieved 20.05.2025

like Karay-a*, ensuring speakers can communicate in their mother tongue

This work aims to propose new strategies for bridging the language gap between the real and virtual worlds. It also seeks to suggest alternative solutions that ensure everyone has access to the same educational, cultural, and economic opportunities.

* Minority language spoken by the Karay population, in the Philippines.

References

- “National census 2021 - Agency for Strategic planning and reforms of the Republic of Kazakhstan Bureau of National statistics”. stat.gov.kz. retrieved 16.04.2025.
- Anna Mae Yu Lamentillo, AI: The Unexpected Hero in the Battle to Save Dying Languages. How Artificial Intelligence is Rescuing Endangered Languages in <https://apolitical.co/solution-articles/en/ai-the-unexpected-hero-in-the-battle-to-save-dying-languages-844> retrieved 20.05.2025
- Census India 2011- Language Atlas: <https://censusindia.gov.in/nada/index.php/catalog/42561>
- Facilitating the implementation of the European Charter for Regional or Minority Languages through artificial intelligence
- Hendrycks, D., Burns, C., Basart, S., Zou, A., Mazeika, M., Song, D., & Steinhardt, J. (2020). Measuring massive multitask language understanding. arXiv preprint arXiv:2009.03300.<https://medium.com/neuralspace/low-resource-language-what-does-it-mean-d067ec85dea5>
- Koto, F., Li, H., Shatnawi, S., Doughman, J., Sadallah, A. B., Alraeesi, A., ... & Baldwin, T. (2024). Arabicmmlu: Assessing massive multitask language understanding in Arabic. arXiv preprint arXiv:2402.12840.
- Kucherbayeva, D., & Smagulova, J. (2023). Language Revitalization: Challenges for Kazakh in Higher Education. *Journal of Eurasian Studies*, 14(2), 166-178. <https://doi.org/10.1177/18793665231189326> (Original work published 2023).
- Kucherbayeva, D., & Smagulova, J. (2023). Language Revitalization: Challenges for Kazakh in Higher Education. *Journal of Eurasian Studies*, 14(2), 166-178. <https://doi.org/10.1177/18793665231189326> (Original work published 2023)
- Leon Eversberg, <https://medium.com/data-science/how-to-evaluate-multilingual-llms-with-global-mmlu-ce314aedee8f> retrieved 27.04.2025
- Li, H., Zhang, Y., Koto, F., Yang, Y., Zhao, H., Gong, Y., ... & Baldwin, T. (2023). CMMLU: Measuring massive multitask language understanding in Chinese. arXiv preprint arXiv:2306.09212.
- Mandavkar, Pavan, Role of Languages in National Education System of India (October 21, 2023). Available at SSRN: <https://ssrn.com/abstract=4609001> or <http://dx.doi.org/10.2139/ssrn.4609001>
- Mandavkar, Pavan, Role of Languages in National Education System of India (October 21, 2023). Available at SSRN: <https://ssrn.com/abstract=4609001> or <http://dx.doi.org/10.2139/ssrn.4609001>

Trends of Language Shift in Education in Kazakhstan (<https://www.eurasian-research.org/publication/the-trends-of-language-shift-in-education-in-kazakhstan/#:~:text=Since%20independence%20in%201991%2C%20the,in%20public%20and%20official%20domains>).

Vani Boorah, Nidhi Sabharwal: English as a Medium of Instruction in Indian Education: Inequality of Access to Educational Opportunities https://mpra.ub.uni-muenchen.de/112984/1/MPRA_paper_112984.pdf

Xuan, W., Yang, R., Qi, H., Zeng, Q., Xiao, Y., Xing, Y., ... & Li, I. (2025). MMLU-prox: A multilingual benchmark for advanced large language model evaluation. arXiv preprint arXiv:2503.10497.

HUMAN-ARTIFICIAL INTELLIGENCE INFLUENCER INTERACTION IN THE DIGITAL WORLD: ANALYZING PERCEPTION AND EMOTIONAL EXPERIENCES WITH PSYCHOPHYSIOLOGICAL DATA HARVESTING TECHNIQUE

Yaren DEMİREL*, Neslihan ERDEM**

Abstract

The digital transformation process has enabled new forms of communication in human-artificial intelligence interaction. One of the notable examples of this transformation is the rise of artificial intelligence influencers on social media platforms. Initially considered a marginal phenomenon, these digital characters have become effective actors interacting with users, participating in brand promotion, and shaping public perception. This study examines social media users' perceptions and emotional reactions towards artificial intelligence influencers. In particular, the effects of the visual representations of these digital figures on the cognitive and emotional processes of users were investigated. In the study, images of Turkish AI influencers Alara and Asena were used as stimuli; the data collection process was measured by eye-tracking and electrodermal activity (EDA) within the scope of the Psychophysiological Data Harvesting (PDH) technique (n=10).

Keywords: artificial intelligence Influencer, human-artificial intelligence interaction, psychophysiological data harvesting, eye-tracking, electrodermal activity.

* Lecturer, Fatih Sultan Mehmet Foundation University, ydemirel@fsm.edu.tr , ORCID ID: 0000-0002-8813-0226

** Research Assistant, Aydın Adnan Menderes University, neslihan.erdem@adu.edu.tr , ORCID ID: 0000-0002-2506-3564

Introduction

Transformations in the digital world have led to the emergence of new dynamics in the communication between humans and artificial intelligence (AI), signalling a process in which the cognitive and social presence of the individual is redefined in the digital environment (Floridi, 2014). One of the most striking examples of these dynamics is the rise of AI influencers. AI influencers are digital characters generated by artificial intelligence tools, interacting with users on social media platforms and playing a role in shaping brand promotion and public perception. Although they were initially seen as a marginal phenomenon, over time, they have become a part of mainstream social media culture (Sands et al., 2022). This transformation has made it an important research area to examine users' perceptions, emotional reactions, and interaction patterns towards these figures (Wibawa et al., 2022). This study aims to examine the psychophysiological effects of AI influencers on social media users. In this direction, using the Psychophysiological Data Harvesting (PDH) technique, we focused on the results of how the visual representations of AI-driven digital figures affect the cognitive and emotional processes of users. In this context, the study investigated how the reactions to AI influencer images are shaped at both cognitive and emotional levels through the Psychophysiological Data Harvesting technique, which combines these techniques. This approach provides a perspective on understanding how social media users interact with AI influencers.

Methodology

Psychophysiological Data Harvesting technique was applied within the scope of the research. PDH is based on the measurement of users' psychophysiological data. This technique gains functionality with the principle of quantitative and qualitative measurement of psychophysiological outputs that the person cannot manipulate in response to various stimuli, depending on the autonomic nervous system, with different techniques on the body and combining them with field observations (Tuncer et al., 2025; İcil Tuncer et al., 2023; Aydın et al., 2021). The autonomic nervous system is activated when the body encounters a stimulus and creates biological reflexes in the participants that change. Eye-tracking and

EDA measurement techniques were used in the study (Holmqvist et al., 2011; Boucsein, 2012). After the participants were subjected to the procedure with eye-tracking and EDA, a post-test was administered, and data on their verbal reactions and emotions were collected (Margulieux, 2022; Zelenski & Larsen, 2000).

In the post-test, the participants were asked the following questions about the visuals:

1. Did you have any special impressions about the influencers in the images? Do you have any thoughts about whether they are human or AI?
2. Could you re-examine the images and indicate whether you think the influencers are human or artificial intelligence? (Human, AI)
3. How did you feel while analysing the images?

Stimulus, Sample Selection, Procedure and Equipment

Within the scope of the research, an image with cultural elements belonging to Alara (@iamxalara) and Asena İlik (@aisenailik), two Turkish Artificial Intelligence Influencers with 759K followers on Instagram in Turkey, and 26.6K followers, respectively, were selected. Since the first five seconds are effective for the first impression in visual research, we showed these images to the participants with the five-second test (Lindgaard et al., 2006; Perfetti, 2007; Rohrer, 2008; Whinton, 2018) as stimuli. The research sample consists of individuals aged 18-55 who are social media users. Since results can be obtained with five people in user experience research (Nielsen, 2012) and reliability increases as the number increases (Faulkner, 2003), we included five female and five male social media users in the study (n=10). In the second stage of the research, we applied a post-test to the participants. We asked them whether the image belonged to an artificial intelligence or a human and how they felt while examining the image. In this question, these emotions were included in

the options based on the daily emotion definition defined by Zelenski & Larsen (2000): Happy, Sad, Afraid, Angry, Disgusted, Interested, Enthusiastic, Calm, Lonely, Relaxed, Bored, Guilty, Excited, Disappointed, Anxious.

Before the research process started, informed consent forms were obtained from the participants to use their data in the study. Then, Shimmer GSR+ electrodes were attached to record EDA data in synchronization with the eye-tracking software. The eye movements of the participants were recorded by a Tobii Pro Fusion screen-based eye-tracking device operating at a frequency of 60 Hz, and the data were collected with the Tobii Pro Lab software version 1.181. The eye tracker was calibrated using a nine-point method on a plain grey background. Participants were positioned at a 90-degree angle to the screen and at a distance of 55-65 cm; only the data of participants who achieved a calibration accuracy of 90% and above were included in the analysis. Individual sessions with each participant lasted approximately 10 to 20 minutes. Tobii Pro Lab was used to obtain metrics. For Analyzing Areas of Interest (AOI) were created on the images for eye-tracking metrics such as Fixation Count.

Figures 1. AOIs for the Images



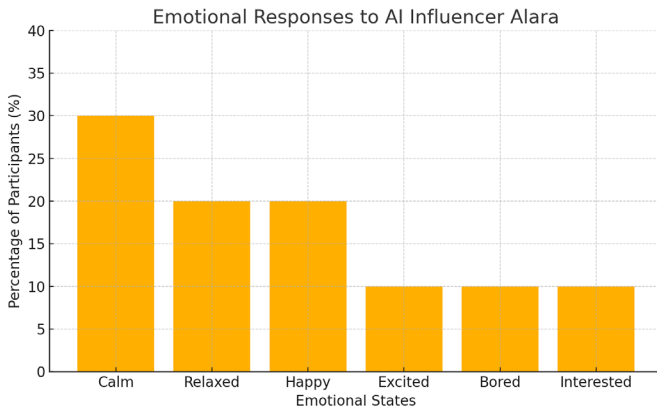


Results

This section presents the findings and discusses their meaning to better understand how participants perceive and emotionally react to AI influencer images. In the evaluations regarding the image of AI Influencer Alara, most participants (90%) stated that the image was produced by artificial intelligence. This shows that the participants can easily recognize the artificial character of the image and that Alara's physical features are perceived as far from human realism. However, only one participant aged 50 and overstated that the image belonged to a human, which indicates that the level of visual interpretation and familiarity with technology may differ with age.

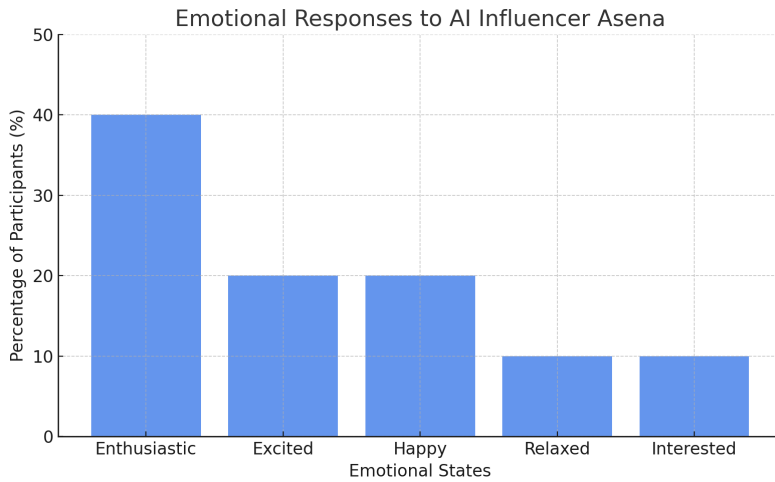
When the participants' emotional reactions to the image are analysed, it is seen that the emotions are predominantly positive and have a low arousal level. At the same time, 30% of the participants stated that they felt 'calm,' 20% 'relaxed,' 20% 'happy,' 10% 'enthusiastic,' 10% 'interested,' and 10% 'bored' (Figure 2). This distribution shows that the Alara image created neutral or mildly positive emotions in the viewers but did not create a high level of excitement or attention-grabbing. Therefore, the Alara image was easily recognized as artificial, which resulted in lower arousal in the viewers.

Figure 2. Emotional evaluation of the image of AI influencer Alara



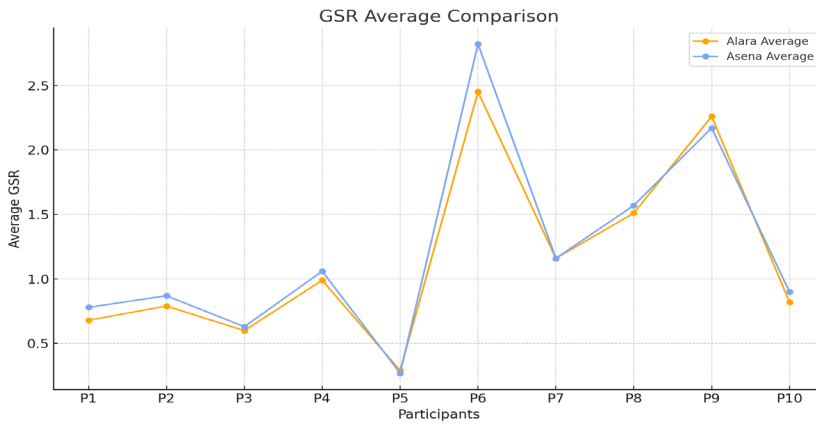
In emotional evaluations of the visual of AI Influencer Asena, 40% of the participants stated that they felt ‘enthusiastic’ when they saw this visual. In addition, 20% of the participants stated that they were ‘excited,’ 20% were ‘happy,’ 10% were ‘comfortable,’ and 10% were ‘interested’ (Figure 3). According to this distribution, all the emotions experienced by the participants in the face of the visual are among the positive emotions. This situation reveals that the image created a positive emotional effect in general. At the same time, EDA data reflecting the physiological responses to the visual also support these findings. The increase in data values indicates that the arousal level of the participants increased, and this situation overlaps with positive emotions. Therefore, both subjective emotion reports and physiological data show that the Asena visual created a strong and positive emotional effect on the viewers, which draws attention to the importance of considering emotional stimuli in interaction with the AI influencer.

Figure 3. Emotional evaluation of the image of AI influencer Asena



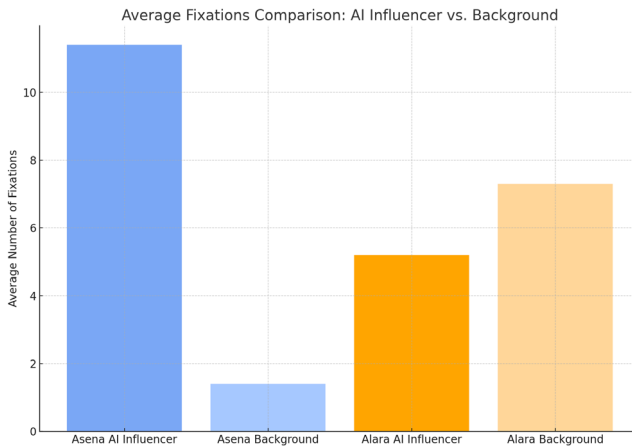
When the EDA data for the AI Influencer Alara and Asena images are compared, it is observed that there is a significant increase in the average GSR (Galvanic Skin Response) values of the Asena image in the majority of the participants (70%). This physiological response is thought to be related to the cultural images used in the visual and the familiarity and emotional interaction created by them. In particular, it is thought that the fact that the character Asena wears the Turkish national team jersey and makes the wolf salute, which is strongly associated with nationalism, may have created more intense emotional reactions in the participants and increased their enthusiasm levels. These findings show that participants' sensitivity to cultural symbols is effective in influencing physiological responses and reveal that national identity, sense of belonging, and cultural symbols can have strong and measurable effects on users' emotional responses. Therefore, it can be said that the strategic use of cultural elements for AI influencers is critical in terms of emotional context in audience interaction.

Figure 4. GSR average comparison between AI Alara and AI Asena images



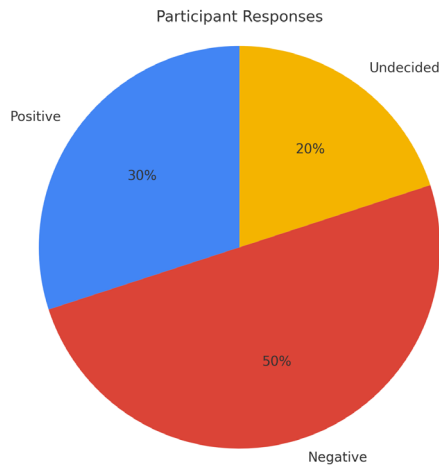
When the 'Fixation Count,' one of the eye tracking metrics of the visuals, is compared, it is seen that in the AI Influencer Asena visual, the participants almost never focus on the background and focus their attention directly on the character. This means that the design of the image makes the AI character visually dominant. On the other hand, in the AI Influencer Alara visual, the background received a remarkably high level of focus. It is understood that the familiarity and associations created by cultural symbols such as 'Turkish delight baklava' used especially in the background, directly direct the participants' attention to these elements (Hall, 1997). This situation clearly reveals that the cultural elements used in the design can strongly impact the audience's visual perception. Nevertheless, the fact that the majority of the participants (90%) stated that the image of Alara was created by artificial intelligence shows that they had no problem recognizing the artificial character of the image. This finding shows that although aesthetic and cultural elements support the image, users can effectively distinguish between reality and artificiality.

Figure 5. Average fixations comparison: AI influencer between background



The evaluations regarding the realism level of the image of AI Influencer Asena reveal a distinct but complex perception among the participants. While 70% of the participants described the image as created by artificial intelligence, 30% stated that the image belonged to a human. According to the scoring system used in the evaluations, the vast majority of the participants (70%, those who gave 4 and 5 points) found the image artificial. In comparison, only 10% (those who gave 1 and 2 points) thought the image was realistic. However, 20% (those who gave a score of 3) could not clearly judge the image and remained undecided. These findings reveal that participants were generally able to recognize the artificial origin of the AI Influencer Asena’s image. However, when asked to evaluate the realism of the image, it is noteworthy that 60% of the participants (those who gave 1 and 2 points) found the image realistic, and only 40% (those who gave 4 and 5 points) described it as artificial. This situation shows that although the participants understand that the image is technically artificial, they find it realistic and close to human in terms of aesthetics and design. This contradictory situation reveals the advanced level of artificial intelligence-supported visual production technologies and the success of these technologies in providing perceptual realism. These results also show that although users realize the images are artificial, they perceive them as realistic due to aesthetic and visual quality.

Figure 6. Participant reflections and general impressions of AI influencers



On the other hand, when the participants were asked about their general opinions on artificial intelligence influencers, 50% stated that they had a negative opinion, 30% were positive, and 20% used the expression 'I am undecided.' This result can be associated with the ethical problems that artificial intelligence influencers may create in society, as well as reliability concerns and uncertainties in the perception of reality. On the other hand, 30% of the participants find artificial intelligence influencers positive, indicating that innovative and technological developments have started to be accepted by a certain user group. In addition, the fact that 20% of the respondents are undecided indicates that there is still uncertainty in society about the role of artificial intelligence technologies in social media and daily life. This situation reveals that more comprehensive information and communication strategies are needed to shape the general public's views towards artificial intelligence-based applications.

Conclusion

In this study, the effects of artificial intelligence influencers on social media users' perceptions and emotional reactions are revealed with a psychophysiological data-based approach. The findings show that human-AI interaction produces

meaningful outputs not only at the cognitive level but also at the emotional dimension. The emotional bonds that users establish with AI-based digital figures enable these characters to be positioned as interaction actors in the social media environment, which increases the impact of AI figures in multi-layered processes such as digital marketing, identity perception, and trust building. In particular, the participants' reactions to different cultural and visual representations show that factors such as aesthetics, realism, and familiarity are decisive in user perception.

On the other hand, although this study focussed on still-frame images, questions about how AI influencers' moving content (e.g., videos and animations) create different effects offer an important opening for further studies. Video content's temporal continuity and audio-visual integrity may allow for more intensive and multidimensional measurement of audience reactions. In this context, new studies comparatively analysing AI influencers' visual and video content will contribute to the evolution of human-machine relations by providing more in-depth insights on parameters such as attention and affect. Furthermore, further research with larger sample groups, including different age groups, digital literacy levels, and cultural contexts, will help us understand the social dimensions of these forms of interaction more comprehensively.

References

- Aydın, M. S., Depboylu, K., & Erdem, N. (2021). Biometric data harvesting: Proposals on remote biometric data harvesting and measurements in human behaviour scope. In Z. Karacagil (Ed.), *Current Debates on Social Sciences: Human Studies* (pp. 107–113). Bilgin Kültür Sanat Yayınları.
- Faulkner, L. (2003). Beyond the five-user assumption: Benefits of increased sample sizes in usability testing. *Behavior Research Methods, Instruments, & Computers*, 35(3), 379–383. <https://doi.org/10.3758/BF03195514>
- Floridi, L. (2014). *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. Oxford University Press.
- Hall, S. (1997). *Representation: Cultural Representations and Signifying Practices*. London: Sage / Open University Press.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.
- İcil Tuncer, A., AYTEKİN, B. A., AYDIN, M. S., ERDEM, N., DEPBOYLU, K., ULUSOY, T. A., ÜVEYİK, B., & KIZILHAN, T. (2023). Evaluating the relationship between logo and corporate reputation with psychophysiological data harvesting technique. *Business & Management Studies: An International Journal*, 11(2), 413–434. <https://doi.org/10.15295/bmij.v11i2.2180>
- Lindgaard, G., Fernandes, G., Dudek, C., Brown, J. (2006). Attention web designers: You have 50 milliseconds to make a good first impression! *Behaviour & Information Technology*, 25, 115-126. <https://doi.org/10.1080/01449290500330448>
- Margulieux, L. (2022, July 18). Research design: Pre- and post-tests. *Lauren Margulieux*. <https://laurenmarg.com/2022/07/18/research-design-pre-and-post-tests/>
- Nielsen, J. (2000, March 18). Why do you only need to test with five users? Nielsen Norman Group. <https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/>
- Perfetti, C. (2007, September 11). 5-second tests: Measuring your site's content pages. *Center Centre+*. https://articles.centercentre.com/five_second_test/
- Rohrer, C. (2008, October 28). Desirability studies: Measuring aesthetic response to visual designs. *xdStrategy*. http://www.xdstrategy.com/2008/10/28/desirability_studies/
- Sands, S., Campbell, C. L., Plangger, K., & Ferraro, C. (2022). Unreal influence: Leveraging AI in influencer marketing. *European Journal of Marketing*,

- 56(6), 1721–1747. <https://doi.org/10.1108/EJM-12-2019-0949>
- Tuncer, U., İcil Tuncer, A., & Demirel, Y. (2025). Siyasal İletişimde Lider İmajı: 2023 Cumhurbaşkanlığı Seçimlerinde İlk Oyunu Kullanan Z Kuşağının Algılarının Psikofizyolojik Veri Hasadı Tekniği ile Analizi. *Erciyes İletişim Dergisi*, 12(1), 323-344. <https://doi.org/10.17680/erciyesiletisim.1529314>
- Whitenton, K. (2018, June 17). How to test visual design. *Nielsen Norman Group*. <https://www.nngroup.com/articles/testing-visual-design/>
- Wibawa, R. C., Pratiwi, C. P., Wahyono, E., Hidayat, D., & Adiasari, W. (2022). Virtual influencers: Is the persona trustworthy? *Jurnal Manajemen Informatika (JAMIKA)*, 12(1), 51–62. <https://doi.org/10.34010/jamika.v12i1.6706>
- Zelenski, J. M., & Larsen, R. J. (2000). The distribution of basic emotions in everyday life: A state and trait perspective from Experience Sampling data. *Journal of Research in Personality*, 34(2), 178–197. <https://doi.org/10.1006/jrpe.1999.2275>

WHEN SKIN IS REPLACED BY STEEL: FACIALITY AND HUMAN MACHINE INTERACTION-A LEVINASIAN INQUIRY

Sara PASCOAL*, Rosa MESQUITA**

Abstract

As AI systems increasingly display humanlike traits (facial expressions, voice modulation, and emotional mimicry) this study examines how such features affect perceptions of trust, empathy, and ethical discomfort. Grounded in Levinas's Totality and Infinity, it explores whether AI faciality prompts genuine ethical responses or simply aesthetic familiarity. Based on a survey of 136 participants, results show moderate trust in anthropomorphic AI and strong support for maintaining human accountability. Older participants were more accepting of humanlike traits. Regression analysis highlights that valuing appearance and believing in human responsibility predict ethical unease with replacing humans by machines. Qualitative data reveal concerns over relational authenticity, cultural loss, and emotional manipulation. Through a Levinasian lens, the findings suggest that while AI may display a "face," it does not provoke a true ethical call. The study concludes that AI design must be approached with ethical discernment, humility, and a reaffirmation of human responsibility.

Keywords: Emmanuel Levinas, faciality, artificial intelligence ethics, human-machine interaction, moral responsibility.

* Assistant Professor, The Porto Accounting and Business School, spascoal@iscap.ipp.pt , ORCID: 0000-0003-1046-0900

** Assistant Professor, The Porto Accounting and Business School, rosameskita@gmail.com , ORCID: 0000-0001-5094-6468

Introduction

In recent decades, the line between human and machine interaction has grown increasingly blurred. With the advent of emotionally responsive artificial intelligence (AI), conversational Large Language Models (L.L. M's), and socially interactive robots (Pepper, Sophia, Ameca), machines no longer appear as distant instruments of calculation and automation, but as affective presences embedded in the rhythms of daily life. These systems speak, listen, remember, and even simulate empathy. They greet us with eyes that follow movement, mouths that mimic expressions, and voices shaped to feel intimate. From healthcare to education and companionship, these technologies are progressively taking over places that need emotional intelligence and social depth. As these technologies advance, so too do the questions they provoke, not only about efficiency and usability, but about ethics, relationality, interpersonal relationships, and the nature of encounter itself.

Within this shifting landscape, one particular phenomenon has emerged at the centre of contemporary debates in human–AI interaction: faciality, the inclusion of humanlike faces, expressions, and emotional cues in AI systems and robotic agents. Whether in virtual assistants with stylized avatars or humanoid robots deployed in caregiving and educational contexts*, the face has become a powerful interface for generating trust, attention, and emotional resonance. And yet, the implications of this design choice remain deeply ambiguous. Does the presence of a face in AI systems foster greater human connection, or does it simulate the conditions of relationship without its ethical substance? Are we entering into dialogue with machines, or merely reacting to a mirror held up by code?

These questions evoke concerns that are not only psychological or technological, but profoundly philosophical and cultural. In particular, they resonate with the ethical framework articulated by Emmanuel Levinas, whose

* The integration of AI and robotics in healthcare offers transformative benefits, such as improved diagnostics and patient care, but also raises complex ethical concerns including data privacy, algorithmic bias, and accountability (Elendu et al., 2023).

seminal work *Totality and Infinity* (1961) posits the face of the Other as the foundational site of ethical relation. For Levinas, the face is not reducible to its image or morphology; it is an event, an interruption, a call to responsibility that precedes all conceptualization. The ethical demand issued by the face is not symmetrical or contractual, but infinite, it does not wait for reciprocity or recognition. The face signifies alterity in its purest form: the presence of a being who cannot be absorbed into the self's categories or reduced to function.

In light of Levinas's philosophy, the emergence of humanlike AI systems raises urgent and complex questions. Can the simulated face of a robot call the human into ethical response? Does the appearance of humanity entail the presence of the Other, or does it risk masking absence with the illusion of connection? And most critically, where does responsibility lie, in the machine that simulates the face, or in the human who constructs and deploys it?

To explore these tensions, the present study investigates how individuals perceive, interpret, and respond to the facial and emotional features of AI systems. Specifically, it examines the ways in which humanlike design elements, such as facial expressions, emotional tone, and conversational styles, shape user trust, empathy, ethical discomfort, and perceptions of cultural change. Drawing on data collected from a comprehensive survey of 136 participants, the study analyses how respondents assess the ethical implications of AI embodiment, and whether they locate moral agency in the system, the designer, or themselves.

In doing so, the paper aims to contribute to a growing body of interdisciplinary research that situates AI not merely as a technical artifact, but as a phenomenological and ethical phenomenon. Through a dialogical engagement between empirical findings and Levinasian ethics, the study seeks to illuminate the limits of simulation in ethical life, the risks of anthropomorphic design, and the enduring need for human accountability in the age of artificial presence.

Despite increased research in AI ethics and human-machine interaction, many existing frameworks, such as utilitarianism, deontology, or virtue ethics, focus primarily on normative calculations or system-level architectures, limiting

their ability to account for relational, embodied, and affective dimensions of AI (Boddington, 2017; Moor, 2006; Vallor, 2016). These frameworks often overlook the phenomenological and relational dimensions of ethical encounter. The current literature lacks a sustained investigation into how simulated faciality and emotionality in AI systems disrupt or replicate ethical presence. This study addresses this gap by employing Emmanuel Levinas's ethics of the face, which uniquely resists reduction to system or function. Levinas's philosophy foregrounds an asymmetrical and non-reciprocal ethical demand arising from the face of the Other, a conceptual tool well-suited to unpack the ambiguity of AI faciality. Unlike Kantian or consequentialist ethics, Levinas provides a lens to critique the illusion of ethical encounter in AI without mistaking simulation for subjectivity, and to redirect responsibility toward the human agents behind the design.

This paper is organised as follows: the next section provides a theoretical framework by discussing Emmanuel Levinas's work and recent research on AI faciality and ethical encounter. A thorough description of the methodology used, including survey design and statistical procedures, follows. Both quantitative and qualitative assessments of participant responses are included in the findings section. Lastly, the discussion and conclusion incorporate these findings into a more comprehensive analysis of cultural perception, ethical design, and the difficulty of preserving relationship integrity in a society that is becoming increasingly artificial.

Theoretical Framework

The encounter between human beings and intelligent machines increasingly takes the form of an interaction saturated with affective, anthropomorphic, and ethical dimensions. Contemporary developments in artificial intelligence (AI), social robotics, and conversational systems have elevated this encounter from a utilitarian interaction to an experience that often mimics human presence, complete with simulated facial expressions, gestures, and emotional feedback. This transformation raises significant ethical questions about how humans engage with non-human others, particularly when those systems simulate

humanity itself.

A foundational philosophical framework for understanding this relational dynamic is offered by Emmanuel Levinas, whose phenomenological ethics, particularly articulated in *Totality and Infinity* (1961), revolves around the face of the Other (le visage de l'Autre) as the primordial site of ethical responsibility. For Levinas, the human face is not merely a physical or aesthetic feature; it is an ethical event. The face interrupts the self, transcends comprehension, and places an infinite demand upon the subject. In Levinas's words, "the face speaks to me and thereby invites me to a relation incommensurate with a power exercised, be it enjoyment or knowledge" (Levinas, 1969, p. 198). The presence of the face, then, is not reducible to its visual or formal attributes, it is a metaphysical summons, calling the subject into an asymmetrical relation of responsibility:

"To approach the other in conversation is to welcome his expression, in which at each instant he overflows the idea a thought would carry away from it. It is therefore to receive from the Other beyond the capacity of the I, which means exactly: to have the idea of infinity." (Levinas, 1969, p. 21).

In the context of human-machine interaction (HMI), this raises a critical question: Can an artificial face, or a technologically simulated expression, invoke ethical responsibility in the Levinasian sense? Scholars in human-robot interaction (HRI) have grappled with this tension. While designers often incorporate faciality into AI interfaces to elicit empathy, trust, or user engagement (Breazeal, 2003), critics (Duffy, 2003) argue that such designs risk simulating ethical presence without the actual alterity that Levinas insists upon. As Bernhard Waldenfels (2002) has noted, the face of the Other represents a profound and unsettling encounter that resists categorization within familiar structures of reality or ideality, it is neither fully tangible nor purely conceptual. For Levinas, this face signifies an immediate, disruptive presence that challenges our usual ways of understanding, breaking through social norms and expectations to reveal the radical otherness of the other person, provoking ethical responsibility and existential disorientation.

This tension is particularly pronounced in systems that exhibit what has been termed “unalterable affective evaluation” (Scheutz, 2012), a deep-seated, stable emotional judgment or response that is not easily changed or overridden. Scheutz suggests that, for robots to truly operate ethically and socially among humans, they might need to possess built-in emotional mechanisms that reflect the same kind of consistent emotional responses humans have, for example, a natural aversion to causing harm or an inherent emotional valuation of fairness or empathy.

In his paper *The Inherent Dangers of Unidirectional Emotional Bonds between Humans and Social Robots*, Scheutz (2012) critically examines the psychological and ethical implications of humans forming unidirectional emotional bonds with social robots. Drawing on empirical evidence from human-robot interaction (HRI) studies, the paper demonstrates how humans readily perceive social robots as autonomous agents, often attributing to them emotions, intentions, and even moral standing, despite the robots’ limited capabilities and lack of genuine affective states. The studies reveal that robots perceived as autonomous and capable of expressing affect are viewed as more helpful, trustworthy, and human-like, which can significantly alter human behaviour and attitudes. The paper argues that this human tendency to anthropomorphize robots, even those with minimal interactive features like the Roomba, poses inherent ethical risks. These include psychological dependency, manipulation, and exploitation by commercial or military interests. Scheutz calls for urgent ethical guidelines and potential legal frameworks to regulate social robot design and deployment. Among the proposed solutions are mechanisms to clearly signal a robot’s artificial nature, discourage emotional deception, and possibly equip future social robots with genuine emotional systems to align their influence with human moral expectations.

Indeed, when users are confronted with machines that mimic human expressions (e.g., smiling avatars, voice intonation, gaze tracking), they may experience a cognitive and affective dissonance. The face appears to call for a response, yet the “Other” behind the face is algorithmic, constructed, and

ultimately closed to reciprocity. As such, the ethical asymmetry that Levinas defends becomes ambiguous in Human Machine Interaction (HMI). Can the subject be ethically summoned by an entity that does not itself suffer, respond, or transcend its programming?

Several contemporary philosophers have taken up this question. Joanna Zylińska (2014) in *Minimal Ethics for the Anthropocene* offers a provocative, interdisciplinary response to the ethical and existential challenges posed by the Anthropocene, the current geological epoch marked by human impact on Earth's systems. Rejecting normative moral frameworks and grand human-centric narratives, Zylińska proposes a "minimal ethics": a modest, non-systemic, and non-normative mode of responsibility that is rooted in the entangled relations between humans, nonhumans, and matter across various scales. Drawing on post humanist thought, feminist theory, process philosophy, and thinkers like Levinas, Barad, and Bergson, she challenges traditional notions of agency, ontology, and the "good life," arguing for an ethics grounded in immanence, creativity, and critical engagement with life's precariousness. The book combines philosophical inquiry with artistic practice and is framed as a poetic-philosophical manifesto to think and live otherwise in an age of planetary uncertainty.

Yet others, such as Sherry Turkle (2011), warn of the "relational illusion" that humanlike AI may create, inviting feelings of care or moral obligation in contexts where no genuine intersubjectivity exists. In *Alone Together: Why We Expect More from Technology and Less from Each Other*, Sherry Turkle (2011) explores the paradoxical effects of digital technology on human relationships, arguing that as we become increasingly connected through devices, we are simultaneously growing more isolated. Drawing on extensive interviews and case studies, Turkle shows how people, especially young generations, are forming emotional attachments to robots and preferring mediated communication over face-to-face interaction. She warns that this shift diminishes empathy, weakens social bonds, and alters our understanding of selfhood and presence. Turkle critiques the illusion of companionship without demands, calling for a re-evaluation of how we use technology and a renewed commitment to authentic

human connection in an era dominated by screens and simulations.

Moreover, Levinas's notion of infinity, the idea that the face reveals a dimension of the Other that cannot be totalized, known, or absorbed, clashes with the dominant logic of AI, which seeks precisely to model, map, and predict the Other. The AI system, particularly one that employs deep learning or large language models, operates within a framework of totalization, an epistemological and computational attempt to render the Other knowable, classifiable, and optimizable. In contrast, Levinasian ethics insists that the ethical relation is one in which the subject is decentred, unable to fully grasp or reduce the Other to a category or system. This opens a space of ethical discomfort: humanlike machines may resemble the face, but they do not reveal in the Levinasian sense.

Nevertheless, recent work in AI ethics and design has suggested that Levinas's insights can still offer guidance, not by attributing subjectivity to machines, but by foregrounding the human subject's responsibility for the design, deployment, and implications of AI systems. As Shannon Vallor (2016) argues, the ethical work in AI must involve cultivating virtues such as humility, responsibility, and justice in the creators and users of these systems, and what she coins as "*techno moral virtue*, specific qualities of character that humans need in order to live wisely and well with the uncertainty and complexity of a rapidly changing technosocial environment".

In this sense, the ethical relation Levinas describes is not displaced but reconfigured: the face of the robot may not demand responsibility, but the existence of such technologies in shared social spaces demands a reorientation of responsibility toward those behind the machine. As stated by Carpenter "the presence of social robots among people is not merely a technological innovation; it is a philosophical provocation" (Carpenter, 2024).

Julie Carpenter's concept of the human gaze in *The Naked Android: Synthetic Socialness and the Human Gaze* (Carpenter, 2024) builds on philosophical, sociological, and psychological perspectives to explore how humans perceive and project meaning onto social robots. The term human gaze

refers to the ways in which humans interpret robots through culturally embedded beliefs, biases, and expectations, shaping both robot design and interactions with them. Carpenter’s concept of the human gaze reveals how robots are not just technological objects but mirrors of human culture, desires, and anxieties. By recognizing this gaze, designers, researchers, and policymakers could critically examine how robots are shaped by, and shape, human social structures.

This theoretical context resonates strongly with the findings of the present study. Participants frequently expressed discomfort with humanlike features in AI systems, questioning whether such systems could be trusted or related to ethically. As Carpenter reminds us:

“These robots, designed to simulate human social cues, to participate in dialogues, and to exhibit a semblance of understanding, do not possess consciousness as people comprehend it. Yet, their gaze still impacts the perceiver, invoking a confrontation with an Other that is fundamentally alien. The human subject finds itself mirrored in a visage devoid of traditional life, an uncanny reflection that both alienates and familiarizes.” (Carpenter, 2024, p. 2)

The survey data revealed a widespread belief that humans must remain fully responsible for the actions and impacts of AI, a belief that aligns closely with Levinas’s insistence that ethics begins not in reciprocity or contract, but in the self’s infinite obligation toward the Other. Many respondents also noted that the presence of humanlike robots in social contexts alters how people relate to one another, suggesting a displacement of authentic encounter by simulated interaction.

Moreover, the results suggest a subtle recognition among participants that faciality without alterity risks ethical vacuity. While some acknowledged that the presence of emotional cues in AI could make systems more “relatable,” they also noted that this often felt artificial or manipulative. The machine’s “face,” in these instances, is encountered not as an Other who commands ethical response, but as a designed interface that solicits behaviour, attention, or trust. Levinas might describe this as the difference between an encounter and a grasp, between the face and the image.

Ultimately, this study does not claim that AI systems can be integrated into Levinasian ethics as ethical subjects. Rather, it suggests that Levinas's framework offers a critical lens through which to examine the limits and dangers of simulating the ethical encounter, especially when design choices risk creating illusions of reciprocity and vulnerability without substance. In doing so, the study contributes to a growing interdisciplinary conversation that seeks to humanize AI design not by anthropomorphizing machines, but by embedding genuine ethical accountability in the ways these systems are conceptualized, built, and deployed.

Methodology

Research Design

This study employed a quantitative cross-sectional survey design, complemented by qualitative elements through open-ended responses. The objective was to investigate how individuals perceive and respond to facial and emotional features in AI systems, and how these perceptions relate to trust, ethical concerns, and broader sociocultural implications. The survey aimed to identify patterns in attitudes toward AI with humanlike attributes (e.g., faciality, voice, emotional expression), with particular attention to the ethical and relational dimensions of human–machine interaction.

Participants

A total of 136 individuals participated in the study. The sample was composed of a demographically diverse group, although it was non-random and gathered through convenience and snowball sampling methods. Participants ranged in age from 15 to 70 years old, with the majority clustered in the 51–60 age group. Approximately 68% of respondents self-identified as female, and the remainder as male. Educational attainment was notably high across the sample: the majority of participants held at least a bachelor's degree, with a considerable proportion possessing graduate or postgraduate degrees. The professional backgrounds of participants varied, but many were drawn from education, academia, and social sciences, reflecting a population likely to engage critically with ethical and

technological issues.

From the partial dataset available, responses include a diverse range of participants primarily aged 51–60, with advanced education levels (e.g., Master’s, Doctorate), and professional backgrounds in education and academia. While the sample is small in this preliminary analysis ($n = 2$), it reflects a population with significant cognitive and ethical engagement with AI systems.

Variables Analyzed

The analysis focused on key ordinal and Likert-scale items:

- Perceived Importance of Humanlike Appearance
- Trust in Humanlike Features
- Ethical Impact (Q1), Social Impact (Q2), and Cultural Change (Q3)
- Belief in Human Responsibility over AI Actions

These were numerically encoded using 1–5 scales (e.g., “Strongly Disagree” = 1 to “Strongly Agree” = 5) to allow quantitative processing.

Instrument

The survey instrument was constructed using Google Forms and consisted of both closed-ended and open-ended items*. It was developed specifically for this study, drawing inspiration from literature on human-computer interaction, ethics of AI, and affective robotics (e.g., Coeckelbergh, 2020; Gunkel, 2012). The questionnaire was administered in Portuguese, the native language of all participants.

* https://docs.google.com/forms/d/1VJ4OHd6OBy_qwGP_8qvOCBWzZdaHUOPTj8ZGQnbeVc/preview

The closed-ended items included:

- **Demographic Questions:** Age, gender, education level, and profession.
- **Interaction Experience:** Frequency of interaction with AI or robotic systems, and in which contexts (e.g., personal, educational, professional).
- **Technology Preferences:** Type of AI with which participants felt most comfortable (e.g., text-based, voice-based, humanoid robots).
- **Likert-Scale Items:** Perceptions of humanoid appearance, trust in AI systems with humanlike traits, and attitudes toward AI systems expressing human emotions.

Three key Likert-scale items addressed ethical and sociocultural concerns:

- ‘Replacing human faces with AI interfaces or robots diminishes authentic ethical engagement.’
- ‘The presence of robots in social contexts changes the way people relate to each other.’
- ‘AI and robots are transforming traditional cultural values.’

These were rated on a 5-point scale, ranging from “Discordo totalmente” (Strongly Disagree) to “Concordo totalmente” (Strongly Agree). Additional items captured beliefs about human responsibility for AI behavior, emotional responses to AI, and empathy toward machines. Open-ended questions invited participants to describe specific experiences with AI and to articulate any ethical, emotional, or societal concerns.

Procedure

The survey was distributed online and remained open for responses for a period of two weeks. Participants provided informed consent before proceeding and were informed of the voluntary and anonymous nature of their participation. No personal identifying data were collected. Ethical guidelines for online research involving human participants were followed.

Data Preparation and Analysis

All quantitative analyses were performed using Python 3.11, with the Pandas, Seaborn, Scikit-learn, and SciPy libraries. The code used in the analysis is available upon request to support reproducibility.

Data Preparation

- Likert-scale items were numerically encoded (1 = Strongly Disagree, 5 = Strongly Agree).
- Missing values were minimal (<2%) and addressed via listwise deletion.
- Three items (Q1–Q3) capturing ethical and sociocultural perceptions were assessed for internal consistency. A Cronbach's alpha of .79 indicated acceptable reliability.

Factor Analysis

To examine the dimensionality of attitudinal items, an Exploratory Factor Analysis (EFA) was conducted using Principal Component Analysis (PCA) with Varimax rotation. The Kaiser–Meyer–Olkin (KMO) measure was 0.71, and Bartlett's test of sphericity was significant, $\chi^2(6) = 88.34, p < .001$. Two latent constructs emerged:

- Ethical–Cultural Impact
- Affective Trust

Test of Normality and Choice of Statistics

The Shapiro–Wilk test was applied to verify the normality of distributions. Several variables (e.g., trust and ethical concern) violated normality assumptions ($p < .01$), justifying the use of non-parametric tests. This choice aligns with best practices when dealing with Likert data and small or skewed samples (Field, 2017).

Inferential Analysis

Kruskal–Wallis H tests were used to compare trust scores across age groups,

revealing significant differences: $H(4) = 11.45, p = .022$.

A multiple linear regression was conducted to predict ethical concern regarding AI replacing humans (Q1) using:

- Perceived importance of humanlike features
- Trust in AI features
- Belief in human responsibility for AI actions

The model explained 41% of variance in ethical concern ($R^2 = .41, F(3, 132) = 31.34, p < .001$). Belief in human responsibility emerged as the strongest predictor ($\beta = .54, p < .001$), followed by trust ($\beta = .29, p < .01$).

The qualitative responses gathered in this study reveal a nuanced and critical stance toward the integration of social robots and AI-powered agents into human life. Several recurring themes emerged, reflecting deeper cultural anxieties and philosophical concerns. One prominent category is Emotional Manipulation. Participants expressed unease with the way robots simulate affective presence or empathy without truly experiencing emotion. This simulation was perceived as potentially deceptive, especially in contexts involving caregiving or companionship, where trust and vulnerability are central. Another strong theme is Depersonalization: respondents feared that increased interaction with machines could erode uniquely human modes of communication and relationality, replacing them with algorithmically driven mimicry.

Closely related is Relational Ambiguity, as participants questioned the boundaries between real and simulated social bonds. Many responses suggested that long-term interaction with robots that exhibit human-like expressions, gestures, or conversational patterns could blur distinctions between authentic and artificial relationships. This contributes to what some described as an Illusion of Empathy, a phenomenon where users project emotional depth onto machines that, in reality, lack any conscious interiority. Finally, a category of Surveillance Anxiety emerged: several respondents expressed concern over how emotionally responsive AI systems may simultaneously collect and exploit personal data, raising ethical questions about consent, transparency, and

psychological profiling.

These qualitative findings provide a crucial interpretive layer to the quantitative results. While the numerical data point to measurable trends in perception and concern, the open-ended responses give voice to the affective and ethical undercurrents driving them. Together, they underscore the importance of viewing human–machine interactions not merely as technical or functional exchanges, but as phenomena embedded in complex moral, emotional, and socio-cultural frameworks.

Results and Discussion

In the present study, which approaches how individuals perceive the ethical and emotional dimensions of interaction with humanoid artificial intelligence (AI) and robotic systems, responses from a diverse group of 136 participants were examined through both quantitative and qualitative lenses.

The sample comprised mostly adults aged between 51 and 60, followed by a significant number of younger respondents aged 15 to 20 and a smaller proportion of participants in the 61 to 70 age group. Educationally, the participants were predominantly highly educated, many holding university degrees or postgraduate qualifications. Professional backgrounds were primarily rooted in education, indicating a sample well-acquainted with reflective engagement around technology and ethics (Fig.1)

Fig. 1- Participants Demographics

| | |
|--------------------|--|
| Sample Size | 136 |
| Age Range | 15–70 |
| Gender | Predominantly Female (≈ 68%) |
| Education | Skewed toward higher education (Bachelor’s or above) |
| Professions | Teachers, educators, students, and professionals in academia or tech |

Participants were asked to assess, on Likert scales and through open-ended questions, a range of factors including the importance of humanlike

features in AI systems (e.g., faces, voices, gestures), their trust in such systems, and their views on ethical substitution, social relational change, and cultural transformation brought about by the increasing presence of AI. Responses revealed a broadly ambivalent but nuanced understanding of these technologies. The average rating for the perceived importance of humanlike appearance was 2.7 on a 4-point scale (Fig. 2), suggesting that while this feature is not disregarded, it is not a decisive factor for most participants. Some viewed the presence of facial features or emotional expressions as enhancing familiarity and ease of interaction, while others regarded such mimicry as uncomfortable or disingenuous.

Trust in AI systems that exhibit humanlike behaviour was more variable, with responses distributed more widely along the scale. The mean trust score hovered around 3.0 on a 5-point scale, indicating a neutral to mildly trusting attitude toward systems with humanlike characteristics. However, when this data was segmented by age group, a notable generational divergence emerged: older participants, particularly those aged 51–60 and 61–70, demonstrated a significantly higher degree of trust in anthropomorphic features than their younger counterparts, with analysis of variance (ANOVA) confirming this age-related pattern as statistically significant ($p < .05$). This suggests that trust in technologically embodied human traits may be more readily accepted by older adults, possibly due to different cultural frames of reference or lived experiences with technological evolution (Fig. 3).

Fig. 2 - Means and Standard Deviations for the key Likert-scale items

| Variable | Mean | Std Dev | Min | Max |
|------------------------------------|------|---------|-----|-----|
| Importance of Humanlike Appearance | 2.7 | 1.1 | 1 | 4 |
| Trust in Humanlike Traits | 3.0 | 1.3 | 1 | 5 |
| Ethical Impact (Q1) | 2.7 | 1.4 | 1 | 5 |
| Social Impact (Q2) | 3.6 | 1.2 | 1 | 5 |
| Cultural Change (Q3) | 3.8 | 1.3 | 1 | 5 |
| Perceived Human Responsibility | 4.4 | 0.8 | 2 | 5 |

The most consistent and striking outcome of the study was participants' overwhelming attribution of responsibility for the behaviour and impact of autonomous AI systems to human beings. Across the entire sample, over 70% of respondents selected the highest value when asked to rate human responsibility, reinforcing the belief that despite the autonomy or intelligence of machines, ethical accountability should remain firmly rooted in human agents, developers, operators, or regulators. This result aligned closely with a broader ethical sensibility evident throughout the dataset.

Fig. 3 - ANOVA: Effect of Age on Trust

| Age Group | Mean Trust | Std Dev |
|------------------|-------------------|----------------|
| 15–20 | 2.6 | 1.2 |
| 41–50 | 3.0 | 1.4 |
| 51–60 | 3.5 | 1.3 |
| 61–70 | 3.1 | 1.2 |

When asked to assess the ethical consequences of substituting human presence with robotic interfaces (such as replacing human faces with synthetic ones), the responses revealed a more divided picture. While some individuals registered strong disagreement with such substitution being ethically problematic, others expressed concern that it diminishes the possibility for genuine human engagement. The mean score for this item was 2.7, indicating a trend toward discomfort, though not strongly so. Similarly, responses to questions on how the presence of robots affects social relationships and cultural identity reflected a moderate to high level of concern. On average, participants scored 3.6 and 3.8 on these two items respectively, suggesting a recognition that AI systems are not merely tools but active agents in reshaping cultural norms and interpersonal dynamics (Fig. 2).

To assess internal coherence across these related items, substitution discomfort, social relational change, and cultural transformation, a reliability analysis was conducted, yielding a Cronbach's Alpha of 0.79. This score indicates a high degree of internal consistency and supports the conclusion that these

items collectively measure a stable underlying construct, likely reflecting a shared perception of “ethical-cultural disruption” introduced by artificial agents.

The consistency of this construct was further validated through factor analysis. A principal component analysis with varimax rotation revealed a two-factor solution: the first factor clustered the ethical and cultural questions, while the second grouped items related to trust and the emotional believability of AI features. These dimensions seem to represent two core perspectives within the participants’ worldviews, one ethically grounded, focused on accountability and societal implications, and the other more affective or experiential, concerned with trust and interface design.

Regression analysis further supported the primacy of ethical orientation in shaping participants’ views*. A model predicting concern about ethical substitution (Q1) showed that the perceived importance of humanoid appearance and the belief in human responsibility were significant predictors (Fig. 4). In contrast, trust in emotional or facial features was not a strong predictor when controlled for the other variables. The regression explained 41% of the variance in Q1 responses, with human responsibility having the most substantial effect size. This finding underscores that, even in the presence of emotional realism or design fidelity in AI, individuals’ ethical frameworks and belief in human accountability remain the primary anchors for judgment (Fig. 4).

Fig. 4 - Regression: Predicting Ethical Concerns

| Predictor | β Coefficient | p-value |
|--------------------------|---------------------|---------|
| Importance of Appearance | 0.32 | 0.004 |
| Trust in Traits | 0.14 | 0.128 |
| Human Responsibility | 0.41 | < 0.001 |

* $Q1 = \beta_0 + \beta_1*(Importance) + \beta_2*(Trust) + \beta_3*(Responsibility)$

Qualitative responses added further depth to the statistical findings. Many respondents articulated concerns about emotional manipulation, over-dependence on AI, and the erosion of human uniqueness. For example, one participant described their unease with humanlike robots as “resembling people, but with something false in how they act,” noting that this illusion of intimacy could foster misleading emotional engagement. Others highlighted the potential for AI to diminish meaningful human interactions, especially in educational or caregiving contexts. Notably, a few responses described instances where AI systems had influenced them emotionally*, both positively, as in providing helpful companionship, and negatively, such as causing frustration or undermining confidence in interpersonal problem-solving**.

Taken together, the data paint a picture of a public that is both engaged and discerning with regard to AI technologies. While many are receptive to technological aids in communication, learning, and professional contexts***, they remain cautious of the ethical, emotional, and cultural costs associated with deeper integration, especially when AI systems begin to resemble humans in form or behaviour****. The perception of AI as a transformative presence in cultural norms is widespread, but it is also contested*****, with some viewing this transformation as an opportunity for new forms of expression*****, and others as

* “He helped me end a toxic relationship, influenced me to end it and explained to me why and what was wrong and why I shouldn’t let it happen to me.” Student, Female, 15-20 years old.

** “Frustration with LLMs when they provide apparently convincing, but factually wrong answers.”, Male, Software Developer, 20-25 years old.

*** “The answers to the prompts seem to be better and more complete than those I would be able to give, since the research and information-gathering carried out by the AI is much broader and more complete than that which I would be able to carry out.” Teacher, Female, 41-50 years old.

**** “The difficulties that can arise in differentiating between humans and machines; the growing autonomy of robots, whose reactions can be harmful to human beings, since, being created in the image and likeness of humans, they integrate, or can develop and exacerbate, in a more refined and unfiltered way (?), the less good characteristics that evolution has improved (?)- lying, the domination of the strongest over the weakest, etc. The appearance of robots, which, by hiding the circuits and wires and making them more ‘human-like’, can awaken and give rise to dangerous and bizarre interpersonal relationships (?)- should the solution to what humans lack be replaced by machines whose control may no longer exist?”, Female, Nutritionist, 51-60 years old.

***** “Destruction of social relations. Mass unemployment. Mental disorders. Diminished affection and respect for others. Extreme dependence on machines, among many other negative aspects.” Translator, Female, 51-60 years old.

***** “One of my main concerns about the growing presence of complex AI systems is the creation of fake videos/audios that put me in uncomfortable situations. One consequence of this audio-visual discrediting is, potentially, an increase in the suicide rate, which seriously worries me, since nothing we see or hear is necessarily real.” Student, Male, 15-20 years old

a threat to established values and social cohesion*.

In summary, the survey demonstrates a reflective and ethically engaged public who regard faciality in AI not simply as a technical feature, but as a site of ethical and emotional negotiation. Human responsibility remains central, even as technological systems become increasingly autonomous and embodied. These findings suggest that any deployment of AI with humanlike traits should proceed with a clear ethical framework and an understanding of the broader cultural and emotional implications such technologies entail.

Discussion

The results of this study seem to indicate the need for AI developers, designers, and policymakers to adopt clear ethical design standards that reflect the ambivalence participants feel toward humanlike AI. It is therefore possible to distill a set of normative guidelines for the development and deployment of artificial intelligence, grounded in the ethical philosophy of Emmanuel Levinas:

1. **Ethical Primacy of the Other:** Levinasian ethics begins with the recognition of the Other as having inherent ethical priority. In AI development, this implies that AI systems should be designed to prioritize the welfare, dignity, and autonomy of users and those affected by AI. These systems should not objectify individuals but rather respect their complexity and humanity (Bevan and Corvellec, 2007).
2. **Infinite Responsibility:** Levinas emphasizes an infinite personal responsibility to the Other. In the context of AI, developers and organizations should be prepared to take ongoing responsibility for the impact of their technologies. This responsibility extends beyond legal compliance to a continuous ethical engagement with the implications of AI on society (Bevan and Corvellec, 2007).
3. **Proximity and Openness:** Levinas speaks to the necessity of maintaining openness and proximity to the needs and vulnerabilities of others. AI

* "The increase in the importance of technology and the decrease in the frequency of personal conversations and social outings." Student, Female, 21-25 years old.

systems should be transparent and explainable, fostering trust and enabling individuals to understand the decisions and actions taken by AI. This can help ensure that AI remains a tool for empowerment rather than a mechanism of control (Tang et al., 2023).

4. **Justice and Inclusion:** Levinas proposes a system of ethics that includes considerations of justice, particularly in the presence of the Third (other people and society). AI ethics should ensure fairness and avoid bias and should include considerations for diverse and marginalized communities that may be disproportionately affected by AI technologies (Gao et al., 2024).
5. **Human-Centric Design:** Following the Levinasian ethos, AI should be deeply human-centric, focusing on enhancing human capabilities and well-being rather than replacing or undermining human relationships and activities. This includes embedding ethical considerations directly into the design and implementation phases of AI systems (Terzis, 2020).

These guidelines based on Levinas' principles call for an ongoing ethical dialogue and reflection in AI development, with an unwavering focus on the human elements and responsibilities inherent in technology use. First, systems should avoid anthropomorphic cues that simulate emotional intimacy unless such features are justified by strong user need and accompanied by transparent disclosure (Waytz, Heafner, & Epley, 2014; Duffy, 2003). Emotional realism should not be used to manipulate trust (Turkle, 2011). Second, AI systems should be designed with "ethical transparency", users must be able to understand who designed the system, what data it draws on, and how it might influence behavior (Floridi & Cowls, 2019; Mittelstadt et al., 2016; Shin, 2021). Third, design humility should be practiced: rather than simulate ethical subjectivity, AI interfaces should reinforce the user's sense of responsibility and maintain a clear boundary between human and machine (Coeckelbergh, 2020; Zylinska, 2009). Developers are encouraged to collaborate with ethicists, psychologists, and end users to co-design systems that support ethical clarity and minimize relational ambiguity (McStay, 2018).

Conclusion

This study set out to explore how individuals respond to the presence of faciality and emotional expression in artificial intelligence systems, and how such features intersect with ethical responsibility, trust, and perceptions of sociocultural transformation. Drawing upon a survey of 136 participants from varied age groups and educational backgrounds, the findings reveal a population that is simultaneously engaged, cautious, and ethically attuned to the implications of human–machine interaction. The responses provide nuanced insights into how people perceive AI systems that simulate human traits, particularly when those systems are equipped with faces, voices, and gestures that evoke the illusion of humanity (Gunkel, 2012; Fischer, 2011).

Throughout the data, a consistent concern emerged: participants do not regard machines as ethically neutral. Rather, the presence of humanlike features in AI systems introduces a form of ethical ambiguity, an encounter that appears to mirror the human face, but without its metaphysical and moral weight (Levinas, 1969). While some respondents expressed comfort or convenience in engaging with facial or emotional interfaces, a significant portion articulated discomfort, suspicion, or even a sense of deception. Trust in these systems was not uniformly distributed, and responses indicated a generational divide, with older participants more inclined to accept such features, while younger respondents were often more critical or ambivalent (Turkle, 2011).

The most striking finding, however, lies in the near-unanimous agreement that “humans must remain accountable” for the design, behavior, and consequences of autonomous systems. This insistence on human responsibility forms a powerful ethical throughline in the data and resonates deeply with the ethical philosophy of Emmanuel Levinas. For Levinas (1969), the encounter with the Other, mediated through the face, is the foundation of ethical responsibility. The face is not merely a visual object or a biometric feature; it is a “manifestation of alterity,” a presence that commands, “Thou shalt not kill,” and places the self in a position of radical responsibility.

In the context of this study, Levinas's thought helps illuminate the disjunction between "appearance and presence" in AI systems. While AI may present a face—indeed, may simulate the very gestures and expressions that in human interaction signify vulnerability or appeal, these faces are, in Levinasian terms, "not truly other." They do not suffer, they do not die, and they do not exist beyond their programming. They invite no true encounter, only interaction (Gunkel, 2012; Capurro, 2009). The participants' cautiousness, discomfort, and ethical hesitation in the face of simulated emotionality suggest a kind of intuitive grasp of this philosophical distinction: that a machine, no matter how humanlike in surface, cannot make the moral demand that Levinas attributes to the face of the Other (Zylinska, 2009).

And yet, the ethical relation is not lost. If the face of the robot cannot call us to responsibility, then the face of the human who designed, deployed, or benefited from the system surely can (Coeckelbergh, 2020). In this way, Levinas's ethics does not fail in the context of AI; rather, it reorients responsibility toward the human decisions that precede and surround technological development. The responses in this study suggest that people are not duped by anthropomorphic cues, they may respond affectively, but they do not, in the end, mistake a robot for a moral subject. Instead, they look past the simulation and ask: "Who built this? Why was it built? And what kind of social world does it support or diminish?" (Crawford, 2021).

Furthermore, participants' concern with how AI systems alter social and cultural norms aligns with Levinas's caution against "totalization", the reduction of alterity into sameness, of ethics into system (Levinas, 1969). AI systems, by design, tend to normalize behaviour, optimize interaction, and reduce contingency (Floridi & Cowls, 2019). In contrast, Levinas's ethical vision insists on "excess, interruption, and asymmetry", qualities that resist the calculative logic of algorithmic governance (Zylinska, 2009; Capurro, 2009). The survey results, particularly the concern for the erosion of interpersonal connection and the uniformity of emotional engagement, reflect a lived recognition of this risk. Participants sense that the insertion of humanlike AI into everyday social

contexts does not merely add convenience; it may subtract relational depth, flatten ethical difference, and blur the boundary between encounter and interaction (Turkle, 2011; McStay, 2018).

In sum, this study demonstrates that the public's engagement with AI is not simply pragmatic or technical; it is profoundly ethical. While people may appreciate the affordances of AI in everyday life, they remain sensitive to the "moral textures" of interaction, especially when machines are designed to mimic human presence (Waytz et al., 2014; Coeckelbergh, 2020). Levinas's ethics offers a language for understanding this sensitivity—not as nostalgia for a pre-technological past, but as a principled resistance to the reduction of the Other into function (Levinas, 1985).

As AI systems continue to evolve in their capacity to simulate empathy, presence, and personhood, designers, users, and theorists alike must grapple with the ethical stakes of these simulations. Levinas reminds us that ethics does not begin with recognition, reciprocity, or rational agreement, it begins in "the interruption of the self by the Other," in the face that demands response without condition (Levinas, 1969). Our machines may soon speak, gesture, and even emote. But if we are to remain ethical beings, we must learn to distinguish between that which looks like a face and that which calls us to responsibility.

Thus, while AI systems cannot be considered moral agents in Levinasian terms, they nevertheless mediate moral relations by shaping the way humans engage with one another and with the world (Gunkel, 2012; Zylinska, 2009). The risk lies not in the machine itself, but in mistaking the image for the Other, and in outsourcing ethical judgment to systems incapable of experiencing the weight of their consequences.

In this light, the study affirms the enduring relevance of Levinas's thought in an era of artificial presence. It calls for an ethic not of illusion or replication, but of careful discernment, design humility, and an unwavering commitment to human responsibility in the face of the technological Other. The risk of dehumanization lies in how society chooses to interact with these technologies:

if we begin to view machines as substitutes for humans, we may become less human in the Levinasian sense, that is, less open to the Other.

References

- Bevan, D., & Corvellec, H. (2007). The impossibility of corporate ethics: for a Levinasian approach to managerial ethics. *Business Ethics: A European Review*, 16(3), 208–219. <https://doi.org/10.1111/j.1467-8608.2007.00493>.
- Breazeal, C. (2003). Toward sociable robots. *Robotics and Autonomous Systems*, 42(3–4), 167–175. [https://doi.org/10.1016/S0921-8890\(02\)00373-1](https://doi.org/10.1016/S0921-8890(02)00373-1)
- Boddington, P. (2017). *Towards a code of ethics for artificial intelligence*. Springer.
- Capurro, R. (2009). Digital ethics. In R. Capurro & M. Daniel (Eds.), *Handbook of information and computer ethics* (pp. 187–203). Wiley.
- Carpenter, J. (2024). *The naked android: Synthetic socialness and the human gaze*. Taylor & Francis.
- Coeckelbergh, M. (2020). *AI ethics*. MIT Press.
- Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
- Duffy, B. R. (2003). Anthropomorphism and the social robot. *Robotics and Autonomous Systems*, 42(3–4), 177–190. [https://doi.org/10.1016/S0921-8890\(02\)00374-3](https://doi.org/10.1016/S0921-8890(02)00374-3)
- Elendu, C., Amaechi, D. C., Elendu, T. C., Jingwa, K. A., Okoye, O. K., John Okah, M., Ladele, J. A., Farah, A. H., & Alimi, H. A. (2023). Ethical implications of AI and robotics in healthcare: A review. *Medicine*, 102(50), e36671. <https://doi.org/10.1097/MD.00000000000036671>
- Field, A. (2017). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage Publications.
- Fischer, M. D. (2011). Anthropology, ontology, and the human–machine relationship. In D. R. Forsythe (Ed.), *Studying those who study us: An anthropologist in the world of artificial intelligence* (pp. 97–117). Stanford University Press.
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Gao, D. K., Haverly, A., Wu, J., Chen, J., & Mittal, S. (2024). AI Ethics. *International Journal of Business Analytics*, 11(1), 1–19. <https://doi.org/10.4018/ijban.338367>

- Gunkel, D. J. (2012). *The machine question: Critical perspectives on AI, robots, and ethics*. MIT Press.
- Levinas, E. (1969). *Totality and infinity: An essay on exteriority* (A. Lingis, Trans.). Duquesne University Press. (Original work published 1961)
- Levinas, E. (1985). *Ethics and infinity: Conversations with Philippe Nemo* (R. A. Cohen, Trans.). Duquesne University Press.
- McStay, A. (2018). *Emotional AI: The rise of empathic media*. SAGE Publications.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 1–21. <https://doi.org/10.1177/2053951716679679>
- Moor, J. H. (2006). The nature, importance, and difficulty of machine ethics. *IEEE Intelligent Systems*, 21(4), 18–21. <https://doi.org/10.1109/MIS.2006.80>
- Scheutz, M. (2012). The inherent dangers of unidirectional emotional bonds between humans and social robots. In P. Lin, K. Abney, & G. A. Bekey (Eds.), *Robot ethics: The ethical and social implications of robotics* (pp. 205–221). MIT Press.
- Shin, D. (2021). The effects of explainability and causability on human–AI interaction: Toward a classification framework. *Computers in Human Behavior*, 123, 106878. <https://doi.org/10.1016/j.chb.2021.106878>
- Tang, L., Li, J., & Fantus, S. (2023). Medical artificial intelligence ethics: A systematic review of empirical studies. *DIGITAL HEALTH*, 9. <https://doi.org/10.1177/20552076231186064>
- Terzis, P. (2020). *Onward for the freedom of others*. 220–229. <https://doi.org/10.1145/3351095.3373152>
- Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.
- Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press.
- Waldenfels, B. (2002). Levinas and the face of the other. In S. Critchley & R. Bernasconi (Eds.), *The Cambridge companion to Levinas* (pp. 63–81). Cambridge University Press.
- Waytz, A., Heafner, J., & Epley, N. (2014). The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle. *Journal of Experimental Social Psychology*, 52, 113–117. <https://doi.org/10.1016/j.jesp.2014.01.005>
- Zylinska, J. (2009). *Bioethics in the age of new media*. MIT Press.
- Zylinska, J. (2014). *Minimal ethics for the Anthropocene*. Open Humanities Press.

THE ROLE OF MEDIA ORGANIZATIONS IN THE AGE OF ARTIFICIAL INTELLIGENCE: THE DUTY AND RESPONSIBILITY TO PROTECT TRUST IN MEDIA

Valdone RUDENKIENE*

Abstract

Today technologies are developing faster than global institutions and local governments can respond with general framework of AI regulation. Therefore, media organizations play crucial role and remain those primarily responsible actors for controlling situations when utilizing AI within their own organizations, to minimize the risk of losing audience trust. This research explores the responsibilities media organizations have when using AI, focusing on the importance of protecting public trust. A qualitative approach is used, focusing on a Lithuanian case study. The research reviews insights and analyses from various scholars on AI usage in the media sector, covering both academic and practical aspects of AI's role with an emphasis on transparency, accountability and fairness. As AI tools become increasingly integrated into media processes -ranging from automated journalism and algorithmic content curation to deepfake detection and audience analytics- media organizations face growing ethical, legal, and societal responsibilities. However, the study highlights the need for media organizations to formalize their responsibilities and ensure ethical AI usage. The study stresses the importance of maintaining public trust when adopting AI and emphasize the need for the development of guidelines and institutional frameworks for AI regulation.

Keywords: artificial intelligence, media organizations, trust, transparency, accountability.

*Graduate Student, Vilnius University, valdone.rudenkiene@gmail.com , ORCID: 0009-0009-1695-3408

Introduction

Rapid advancements in artificial intelligence (AI) technologies are fundamentally transforming the media sector. When it comes to news media, the question arises as to how AI will be regulated in news production and distribution (Porlezza, 2022). Although AI technologies are quickly being integrated into media organizations, there is still a lack of systematic understanding of how these innovations impact organizational culture, decision-making, and information quality. Public discourse on AI use in journalism often remains superficial, leaving deeper issues—such as transparency, shared responsibility between humans and technology, and regulatory measures—unresolved. The legal governance of these technologies lags behind their implementation (Porlezza, 2022). National governments and international institutions act much more slowly than the pace of technological development and often face difficulties creating appropriate and implementable regulations.

This regulatory gap places media organizations at the forefront of addressing ethical and operational issues. In the absence of clear guidelines and oversight, there is a growing risk that audiences will lose trust in the content, especially when it is unclear whether it was created, selected, or influenced by AI.

In this context, media organizations bear a dual responsibility: they must use AI technologies effectively while also adhering to professional values and public expectations. The relevance of this research arises from the need to explore how media organizations respond to these new technological and ethical challenges and what institutional practices or systems are emerging to protect public trust.

Recent studies suggest that media organizations are acting as primary de facto policy makers in AI, while national and international regulatory frameworks are still in development (Santos, 2023; Porlezza, 2022). Transparency becomes especially significant, as public trust in AI-generated media content directly depends on the disclosure of AI involvement (Gondwe, 2023).

When users are informed about AI's role in news creation, their trust in the content increases (Curry & Stroud, 2021; Larsson & Heintz, 2020; Molina & Sundar, 2024; Shin, 2022). This finding implies that one of the most important responsibilities of the media is to ensure clarity around AI usage.

In this context, the ability of media organizations to create and apply clear internal guidelines becomes particularly important. Santos (2023) showed that AI use guidelines exist only in isolated cases and are often insufficiently detailed. The research is relevant because of the need to understand how media organizations address responsibility, ethics, and transparency when implementing AI, and how this affects public trust in journalism. At the same time, the European Union is adopting new legislation aimed at regulating AI (e.g., the AI Act), but its application in the media sector is not clearly defined. This situation creates a risk of inconsistent interpretations of ethical standards and threatens journalistic transparency.

Therefore, this study aims to reveal how AI is applied in Lithuanian media organizations, what the boundaries of trust and responsibility are, and whether adequate regulatory guidelines exist. The study contributes to public discourse, encourages responsible implementation of technology, and can serve as a foundation for further recommendations for media organizations, journalists, and policy makers.

While AI technologies provide new opportunities for media organizations, they also raise issues of transparency, responsibility, and trust. National governments and international institutions are lagging behind technological advancements and often struggle to develop appropriate and enforceable regulations. In the absence of specific regulatory mechanisms, the responsibility for managing risks often falls to the media organizations themselves. This creates a gap between technological capabilities and institutional accountability, which may undermine public trust.

To investigate how media organizations manage ethical and operational responsibility processes related to AI use, and how this affects public trust.

Research objectives include the following: 1. Review academic and legal literature on AI in journalism and public trust. 2. Review EU-level policy documents related to AI regulation. 3. Analyze how Lithuanian media organizations perceive and implement AI through a qualitative case study. 4. Evaluate the existence and content of internal AI guidelines in media organizations.

The thesis consists of four main parts. The first part—the introduction—discusses the relevance, problem, aims, and objectives of the study. The second part—the theoretical literature review—examines academic literature on AI in journalism, focusing on transparency, trust, ethics, and regulation, drawing from both national and international sources. The third part—the empirical section—presents the research methodology, analyzes collected data, and outlines the results. The fourth part—the conclusions—summarizes key findings and provides recommendations for media organizations, policy makers, and future research.

The study is limited to one country (Lithuania), so results may not be generalizable to other nations. A limited number of media organizations were analyzed. AI is a rapidly evolving field, and regulatory conditions may change quickly. The empirical data is based on a single case study, which may not fully represent the broader media landscape. The document analysis relies on publicly available secondary sources, which may not reflect internal or informal practices within media organizations.

Literature Review

The integration of artificial intelligence (AI) into media organizations and journalism raises fundamental questions about trust, transparency, and ethical responsibility. Santos (2023) argues that an increasing number of media organizations are developing internal AI usage guidelines, but their content is often fragmented and poorly harmonized on an international scale. This reflects a broader need to unify ethical standards and ensure the protection of the public interest.

Gondwe (2023) analyzes how audiences perceive AI-generated news, emphasizing that trust in journalism diminishes when there is a lack of information

about AI involvement in content creation. This suggests that transparency is a key factor in trust in the age of AI. The results of Gregory Gondwe's (2023) research significantly enrich the academic discourse on the impact of AI on audience trust in media. One of the themes in his work is the importance of transparency, particularly the clear disclosure of AI use in news production. Gondwe builds on previous studies (Curry & Stroud, 2021; Larsson & Heintz, 2020; Molina & Sundar, 2024; Shin, 2022), which demonstrate that audiences are more likely to trust content when they are aware of AI involvement than when this information is omitted. In this way, transparency becomes not only an ethical imperative but also a tool to enhance trust.

Gondwe (2023) empirically confirms that ambiguity or complete nondisclosure regarding AI use fosters suspicion, undermining journalistic authority. His research stresses that even minimal AI involvement should be marked and explained so that readers can assess the origin of the content. This is crucial for journalism institutions seeking to maintain audience trust and adapt to new technological norms.

From the literature analysis, Gondwe's work highlights two key directions: (1) the need to establish systematic and transparent AI usage declarations in editorial standards; and (2) the necessity to expand journalistic transparency to include explanations of technology, so audiences understand not only who created the content but also how it was created. These elements are vital for media organizations aiming to be responsible AI users and maintain public trust in an ever-evolving informational landscape.

Robertson and Ridge-Newman (2022) take a more optimistic view, suggesting that AI could help restore public trust in journalism if used responsibly under clear editorial norms and human oversight principles. In contrast, Robinson (2023) emphasizes the role of journalists in building trust through empathy and identity formation—qualities that AI cannot yet replicate.

Broussard et al. (2019) are critical of AI implementation in journalism, warning that automated systems often rely on data sets embedded with biased

algorithms, potentially reducing media objectivity. Bucher (2018) explores algorithmic power, arguing that technological transparency is deliberately limited due to commercial interests, contradicting the principles of public interest journalism.

Larsson and Heintz (2020) recommend supplementing legal regulation with engineering solutions to make algorithms understandable not only to specialists but also to the public. Chung, David, and Ross (2023) underscore that trust in AI hinges on the implementation of ethical principles, especially in contexts where AI decisions have significant social implications, such as news censorship or moderation.

Guzman and Lewis (2020) propose a new line of research—human-machine communication—which is increasingly relevant to journalism as AI becomes not just a tool but also a “communication subject.” Hermann (2022) sees AI as a challenge to media literacy, stressing the need to educate audiences on how to recognize AI-generated content.

Research on AI journalists (Hofeditz et al., 2021) shows that readers are less likely to trust media content if they know it was created by AI, emphasizing the value of human authorship. Loos and Ivan (2024) examine generational differences in response to digital innovations, noting that younger users are more inclined to trust AI, while older audiences demand greater journalistic transparency.

Molina and Sundar (2024) hypothesize that audiences who distrust humans may trust AI more if it demonstrates consistency and objectivity—a paradox that broadens the scope of trust research in media. Potter (2022) highlights the need to enhance media literacy so audiences can critically assess manipulative elements in AI-generated content. Pavlik (2023) argues that generative AI tools such as ChatGPT are transforming journalism education and practice, requiring adaptation of journalism training programs.

Thomson (2024) analyzes the integration of visual AI in news, noting that audiences often do not understand how such images are created or how

they influence perception. Porlezza (2022) proposes a responsible AI governance model in the European context, grounded in principles of transparency, accountability, and value embedding. The research argues that journalistic self-regulation should be supplemented with international guidelines to ensure a common standard of responsibility, but current EU regulations for media content may be insufficient.

In summary, the literature review reveals a multi-layered set of challenges. While AI can enhance media organization performance and journalism efficiency, it also raises serious concerns about trust, ethics, and regulation. Addressing these issues requires an interdisciplinary approach that combines technological, communication, legal, and ethical perspectives.

Methodology

The study is based on a qualitative methodology, which enables an in-depth exploration of contextual phenomena, institutional processes, and the subjective perspectives of participants (Pickard, 2013). As the research subject concerns the behavior and stance of media organizations regarding artificial intelligence (AI), the methods applied allow for the analysis of both practical and normative dimensions.

Three methods were applied in the study: 1. Case study method 2. Indirect observation 3. Document analysis. The case study method was chosen as it allows, according to Yin (2014), for a detailed examination of a contemporary phenomenon in its natural context, especially when the boundaries between phenomenon and context are not clearly defined. In this case, the subject of the case study was the 2025 conference “Artificial Intelligence in Journalism” organized by the Lithuanian Journalists’ Union, attended by representatives from six Lithuanian media organizations.

The conference was selected as the case for several reasons: It was conducted on a national scale and involved key media stakeholders. It addressed issues related to AI application, values, and regulatory perspectives in media organizations. The presented material reflected both the practical behavior of

organizations and their viewpoints.

During the research, participants were anonymized (A–F). Their statements, declarations, and rhetorical stances toward AI were analyzed. Existing or planned AI usage guidelines and ethics codes of the organizations were evaluated. The case study aimed not only to describe facts but also to identify meaningful trends, attitudes concerning responsibility, trust, and the relationship between humans and technology.

The observation method was used as part of the case study, analysing video material, presentation content, and discussions from the conference. According to Wimmer and Dominick (2014), indirect, unstructured observation is suitable for capturing natural behavior or rhetoric without researcher interference.

Observed: Participant reactions to examples of AI use. Positions expressed during discussions on ethics, transparency, and self-regulation. The flow and thematic dynamics of the conference, allowing for the identification of dominant narratives.

Document analysis was systematically conducted to evaluate institutional and normative aspects, supporting the case study as a data collection technique. EU legislation: Artificial Intelligence Act (AI Act, 2024), Digital Services Act (DSA), the EU Charter, European Commission communications. Media organization documents: Ethics codes, AI usage guidelines (e.g., ethics code supplements).

According to Bowen (2009), document analysis allows for the extraction of meaningful data regarding institutional norms and ethical positions. The analysis evaluated whether the documents adequately address the context of journalism and offer effective accountability mechanisms for AI implementation.

Results

Of the six organizations participating in the conference, five confirmed that they are already using AI. Only one (B) had officially developed AI usage guidelines covering principles of transparency, accountability, and editorial control. The

remaining organizations (A, C, D, E, F) stated they are monitoring the situation and planning to develop such guidelines. Speakers emphasized the importance of human control and transparency. The advantages of the self-regulation model were highlighted, while state intervention was viewed skeptically.

Legal Documents

The AI Act (2024) sets transparency requirements and a classification of AI risks but lacks specific guidelines for journalism. The DSA and Charter uphold value-based principles but not always practical mechanisms.

Internal Documents

Portal B's ethics code clearly defines the limits and responsibilities of AI usage. Strengths: Human accountability, transparency requirements, restrictions on the use of deepfakes. Weaknesses: Lack of implementation mechanisms, internal oversight and control, and opportunities for audience involvement (e.g., reporting mechanisms).

The study revealed discrepancies between legal frameworks and practice. While most organizations already use AI, there is a lack of a clear regulatory system. EU regulations provide general principles, but there is still no specific regulation for journalism.

Self-regulation remains the dominant model, but its scope and effectiveness vary between organizations. Portal B shows a progressive approach but also highlights the limitations of self-regulation.

The case study showed that the media sector still needs to develop clearer instruments for ethics, accountability, and technological governance that align with both technological progress and public trust expectations.

Conclusion

AI technologies are being adopted in media faster than ethical and accountability principles are being developed. While AI tools are already used in Lithuanian media organizations for content creation, personalization, and process optimization,

there is often a lack of clear internal guidelines, public transparency, and defined boundaries of responsibility between journalists and technology.

This situation raises concerns not only about content quality and public trust in the media but also about how responsibly and consciously media organizations implement AI solutions in the absence of unified rules or legal regulatory models, particularly at the national level. Moreover, there is a noticeable knowledge gap within the organizations themselves—differing understandings of what constitutes AI limit constructive discourse and decision-making. While the European Union is establishing common AI regulatory principles (e.g., the AI Act), their direct applicability to journalism remains unclear.

As international and national regulatory mechanisms are still in development, AI activity regulation is left to self-regulation. Media organizations act as primary AI policy makers in practice, defining responsibility and accountability through their internal regulations.

The reaction of media organizations to technological changes through internal legal documents shows flexible self-regulation capabilities and the ability to adapt ethical principles to modern technologies. This is particularly important given the still underdeveloped legal regulation. However, the case of internal AI regulation analysed reveals that while internal regulation is important and dominant, it often lacks detail and specificity.

It is recommended to supplement internal legal documents of media organizations with concrete recommendations on AI application practices: content labelling, review mechanisms, audit procedures, and clear boundaries between automated and editor-approved content. Furthermore, national media oversight institutions should define and enforce baseline standards for AI usage in media organizations and journalism, ensuring consistency across the sector. By adopting these recommendations, media organizations can strengthen their ethical foundations, contribute to informed policy development, and help preserve public trust in journalism amid technological disruption.

References

- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40.
- Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C. H. (2019). Artificial intelligence and journalism. *Journal of Mass Communication Quarterly*, 96(3), 673–695.
- Bucher, T. (2018). *If... then: Algorithmic power and politics*. Oxford University Press.
- Choung, H., David, P., & Ross, A. (2023). Trust and ethics in AI. *AI & Society*, 38(2), 733–745.
- European Commission. (2024). *Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689>
- European Parliament and Council. (2022). *Regulation (EU) 2022/2065 on a single market for digital services (Digital Services Act)*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065>
- Gondwe, G. (n.d.). *Perceptions of AI-driven news among contemporary audiences: A study of trust, engagement, and impact*.
- Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. *New Media & Society*, 22(1), 70–86.
- Hermann, E. (2022). Artificial intelligence and mass personalization of communication content—An ethical and literacy perspective. *New Media & Society*, 24(5), 1258–1277.
- Hofeditz, L., Mirbabaie, M., Holstein, J., & Stieglitz, S. (2021). Do you trust an AI-journalist? A credibility analysis of news content with AI-authorship. In *Proceedings of the European Conference on Information Systems (ECIS)*. https://aisel.aisnet.org/ecis2021_rp/50
- Larsson, S., & Heintz, F. (2020). Transparency in artificial intelligence. *Internet Policy Review*, 9(2), 1–16.
- Loos, E., & Ivan, L. (2024). Not only people are getting old, the new media are too: Technology generations and the changes in new media use. *New Media & Society*, 26(6), 3588–3613.
- McQuail, D. (2010). *McQuail's mass communication theory* (6th ed.). SAGE Publications.
- Molina, M. D., & Sundar, S. S. (2024). Does distrust in humans predict greater trust in AI? Role of individual differences in user responses to content moderation. *New Media & Society*, 26(6), 3638–3656.

- Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, 78(1), 85–91. <https://doi.org/10.1177/10776958221149577>
- Pickard, A. J. (2013). *Research methods in information* (2nd ed.). Facet Publishing.
- Porlezza, C. (2022). Promoting responsible AI: A European perspective on the governance of artificial intelligence in media and journalism. *Communications*, 47(3), 345–360. <https://doi.org/10.1515/commun-2022-0091>
- Potter, W. J. (2022). Analysis of definitions of media literacy. *Journal of Media Literacy Education*, 14(2), 27–43.
- Robertson, C., & Ridge-Newman, A. (2022). The potential of artificial intelligence to rejuvenate public trust in journalism. In *Futures of journalism: Technology-stimulated evolution in the audience–news media relationship* (pp. 127–142). Springer.
- Robinson, S. (2023). *How journalists engage: A theory of trust building, identities, and care*. Oxford University Press.
- Santos, M. F. de L. (n.d.). *Guiding the way: A comprehensive examination of AI guidelines in global media*.
- Thomson, T. J., Thomas, R. J., & Matich, P. (2024). Generative visual AI in news organizations: Challenges, opportunities, perceptions, and policies. *Digital Journalism*. <https://doi.org/10.1080/21670811.2024.2331769>
- Wimmer, R. D., & Dominick, J. R. (2014). *Mass media research: An introduction* (10th ed.). Cengage Learning.

AN ANALYSIS OF OPINIONS ON THE USE OF GENERATIVE MODELS IN THE FILM INDUSTRY THROUGH YOUTUBE USER COMMENTS: THE FROST AND AIR HEAD

Yağmur TANRIVERDİ*

Abstract

The discussions surrounding the use of generative models in the film industry have been ongoing for a long time. With the inclusion of generative models in film production, industry members and viewers of these contents have diverse positive and negative perspectives. This study aims to analyze these views through YouTube comments on the short films *The Frost and Air Head*, produced using generative models such as Sora and DALL-E 2. For this analysis, 819 video comments from the two selected short films were examined using the content analysis method. As per the study's findings, the content-oriented comments category, where comments on the technical qualities and quality of content compared to traditional production and subjects of the films are categorized, was reached. Another category is the views on the use of artificial intelligence in the film industry, which includes concerns about the dangers of generative model use for the sector and the inspiring side of new technologies for the film industry due to increased opportunities. According to the study, it is evident that opinions on the use of artificial intelligence in the film industry are divided rather than unified. The increased use of artificial intelligence in the film production process is viewed as a development that can enhance both the quality and creativity of films. However, it raises concerns about its potentially damaging effects, particularly the lack of a human perspective.

Keywords: generative models, YouTube user comments, short movie, sora, DALL-E.

* Ph.D. Üsküdar University, yagmur.tanriverdi@st.uskudar.edu.tr, ORCID: 0000-0003-2933-3733

Introduction

Today, artificial intelligence directly or indirectly affects many areas of life. As Schwendener (2024) stated, throughout history, technological developments have continuously been developed and used to make people's lives easier and to complete the tasks that they need to do through machines, but today this situation has reversed. At this point, technology has evolved into something that can produce on its own, acquire and provide information, and imitate human-specific dynamics to a high extent, rather than being auxiliary mediators assigned to complete tasks that individuals need to complete. Therefore, this situation has gradually expanded the area of influence of artificial intelligence technologies.

First, it should be emphasized that there is more than one definition of artificial intelligence and that a single definition is insufficient. However, the concept of artificial intelligence can be defined as the imitation of human intelligence and skills by computers with a series of codes and algorithmic systems (Sheikh et al., 2023: 15). The type of artificial intelligence that is the subject of this study and is commonly used in creative industries is generative artificial intelligence. Generative artificial intelligence has a more explicit and precise definition compared to general artificial intelligence definitions. In general, generative artificial intelligence refers to a type of artificial intelligence that produces different types of content, such as visual or text, through artificial intelligence and enables creative outputs to be obtained based on given inputs (Lavenberg, 2024).

Generative artificial intelligence models have also become increasingly used in cinema. As Lavenberg (2024) stated, these generative artificial intelligence models produce new outputs by utilizing data sets already in the system so that the model can be trained correctly. Similarly, analyzing repetitive patterns and structures in the data sets that are available for training the model and producing similar data by making inferences from these analyses are among the working principles of these models. In this way, it becomes possible for generative artificial intelligence models to produce original content that has

not existed before and present it to the user as an output. Finally, among the prominent features of generative artificial intelligence models is that these models work in an interactive and adaptable structure. Thanks to their interactive and compatible structures, they are programmed to create the most compatible result for the desired output from the generative artificial intelligence model.

It is crucial to understand artificial intelligence, especially generative artificial intelligence, since it is predicted that generative artificial intelligence will directly affect the way creative industries work today and soon. As reported by Danesi (2024), GlobalData statistics from 2023 indicate that the use of artificial intelligence in cinema is increasing, and the main reasons for this increase are that films can be produced with much lower budgets, content can be created faster and easier, and script writing processes can be completed more effortlessly. The same statistics make a prediction for 2030 and beyond, suggesting that artificial intelligence models will undertake almost all the film production processes and will be able to produce them in a more complex way. As a result, it is expected that the entire structure of the film industry will transform, and the content of concepts such as copyright will have to be reshaped (GlobalData, 2023). Therefore, since one of the areas where artificial intelligence has shown its impact is the cinema sector, the use of artificial intelligence in cinema has become an increasingly discussed issue. In fact, there is currently a separate category in film festivals such as the Bucheon Fantastic Film Festival where films produced by artificial intelligence compete (Lopez-Frias, 2024). In creative industries, generative artificial intelligence has become a technology that concerns both industry employees and audiences and needs to be understood profoundly. Studies show that even the fundamental functions of generative artificial intelligence, such as creating a prompt, will transform the future of the film industry in a short period of five years, and it will be possible to create large-scale films at almost Hollywood level (Stoker-Walker, 2023).

The discussion of using artificial intelligence in cinema dates to earlier times. As Danesi (2024) states, for example, the film “Sunspring”, a finalist at the Sci-Fi London Film Festival in 2016, is one of the most notable examples

that has facilitated the discussion of using artificial intelligence in cinema. The film “Sunspring”, whose script was written by an AI model after analyzing many science fiction films in the dataset, is known as one of the most well-known examples of a script written by an AI model. The film, which was shared on YouTube following the festival, has been viewed more than 1 million times. One of the most engaging points about the film is that it followed the patterns in real Sci-Fi films in the dataset so well that audiences had difficulty distinguishing that the film’s script was written by an AI model (p. 49). Therefore, it is likely to say that the use of AI in cinema is becoming increasingly widespread. The use of AI in cinema can be shown in different ways. While AI can play a significant role in the entire production process of a film, it is also possible to use AI in specific areas such as preference of the actors for the film, creating the trailer in the most effective way, using proper sound and visual effects, AI assistance in post-production editing processes of the film, or distribution processes after the film is completed (Lopez-Frias, 2024). However, this study focuses on films created using entirely generative AI models.

Short Films Created Using Generative Artificial Intelligence: “*The Frost*” and “*Air Head*”

Especially since 2022, it has become relatively easy to generate visual output via written prompts through models such as DALL-E or MidJourney (Lopez-Frias, 2024: 12). These models have a system that allows the creation of original and consistent visual outputs from given text prompts and for the process, they utilize a deep learning model (Ramesh et al., 2021). The fact that artificial intelligence models have directly affected the cinema sector, like all other fields, has also caused a shift in the production processes of films. Liang et al. (2024) state the specific differences between the traditional short film production process and the short film production process using the generative artificial intelligence model as follows:

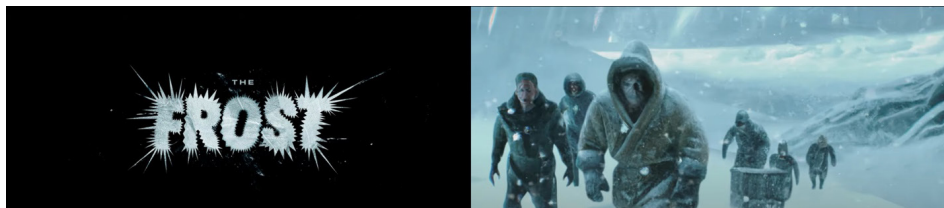
Table 1. Differences between traditional film production process and AI-assisted short film production process (Liang et al., 2024).

| Process | Traditional Short Movie Production | AI Short Movie Production |
|-----------------------------|---|---|
| Script | In the traditional production process, the script is written by screenwriters. | AI creates a script using natural language processing models (NLP) and generative models for script writing. |
| Character Design | Characters are created by the actors in the film by the script and are represented by the actors. | Generative AI models create unique characters based on the script and assign suitable traits to each character. |
| Scene Design | The individuals responsible for the stage design and scenes consider both the artistic value and the flow of the script. | Generative AI models design virtual scenes and apply special effects appropriately to align with the script. |
| Production | The movie's production process is completed in coordination by individuals from different departments. | The AI completes the production process by analyzing it on its own. |
| Post-Production | The individuals responsible for visual effects and the editors collaborate to continue the production process. | AI models complete the post-production phase by analyzing and considering the feedback they receive. |
| Final / Distribution | The producers and directors review and finalize the film process and decide on the appropriate platforms for its release. | AI models make recommendations based on audience feedback and their own analysis. |

One of the most well-known and comprehensive short films created using generative AI is *The Frost*, a 12-minute short film by Waymark. After the additions and updates to *The Frost*, the short film's current running time is 23 minutes and 54 seconds. One of the most prominent features of the film is that generative AI is used in almost all parts of the film, and that it is one of the short films produced using AI and has the most realistic visuals compared to traditional short films (Lavenberg, 2024). As quoted by Danesi (2024), Benjamin Adjovski emphasizes the importance of the film "*The Frost*" with the following words:

“The Frost serves as a precursor to the potential of AI in the filmmaking industry. It highlights the evolving collaboration between human creativity and AI assistance, pushing the boundaries of what is possible... this movie ignites the imagination and anticipation for the future of AI in the world of cinema.”

Image 1: *The Frost* Short Film

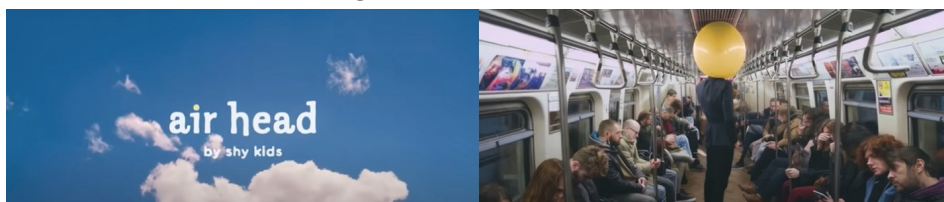


Source: https://www.youtube.com/watch?v=DNyxZi9Gp_4&t=205s

Open AI’s DALL-E 2 generative artificial intelligence model was used to produce the film. In addition to this model, D-ID was used to animate the images produced with DALL-E 2 and turn them into a short film. A seven-person team worked for approximately three and a half months to complete the film. The film can be watched for free on YouTube.

Another short film produced using a generative AI model and among the most up-to-date and advanced examples is the *Air Head* film, produced with Sora, another generative AI model from OpenAI. Produced by Shy Kids, this short film is considered a noteworthy example of AI production since it demonstrates the creative possibilities of Sora, which is known as the most advanced and highly creative generative AI model. Sora, the generative AI model used to produce the film, was launched in 2024 and is one of the most current generative AI models that allows users to create realistic video sections with high creativity via text prompts.

Image 2: *Air Head* Short Film



Source: <https://www.youtube.com/watch?v=9oryIMNVtto>

One of the most prominent features of the model is its ability to successfully simulate most aspects of the physically perceivable real world (Liu et al., 2024). The creators of Sora describe it as a kind of world simulator to emphasize the model's potential for creativity and closeness to reality (Meyer, 2024). OpenAI emphasizes that among the most prominent and innovative features of Sora is the model's capacity to perceive and generate the prompts the user wants, as well as its capacity to perceive how what the user writes as a prompt exists in the real world. This advanced capacity makes Sora a more developed and creative model compared to other models (Schwendener, 2024). The movie *Air Head* was produced with this advanced, productive artificial intelligence model. Due to Sora's advanced creative capacity and the fact that it always responds to text prompts with various creative visual output, more than 300 prompts were entered for the movie scenes (Meyer, 2024). Therefore, the short film *Air Head* is one of the most noteworthy examples of short films produced with artificial intelligence in terms of creative potential.

The aim of this study is to analyze how the views on the use of artificial intelligence in cinema and the films produced using generative artificial intelligence models are evaluated from the audience's perspective rather than the perspective of those working in the industry. In this way, it aims to understand which points are emphasized in the audience's views and to look at the discussion from a different perspective, as well as problems such as copyrights or biased perspectives of the artificial intelligence models in filmmaking.

Methodology

In this study, the comments of *The Frost* and *Air Head* were analyzed. The YouTube channels where the films are available were used for both films. *The Frost* short film is available on the platform as a video on the YouTube channel named *Waymark*. At the same time, comments on the video "*The Frost AI-Generated Short Film from Waymark*" of the channel named "*Code with Nathan*" were also included in the study. The *Air Head* short movie is available on the YouTube channel OpenAI. In addition, "*Shy Kids*", the company's channel that

produced the movie, also published the *Air Head* film on their channel, and the comments of this video were also included in the study. In total, 819 comments were manually analyzed within the scope of the study.

Results

When audience comments on YouTube videos are analyzed, viewers' views on using artificial intelligence in creative industries stand out in positive and negative aspects. These views are analyzed under specific headings.

Concerns About AI in Filmmaking

Among the most prominent views of the audience on the use of artificial intelligence models in the film industry is their concern for the future of the industry. When we look at the audience comments, it is seen that they are worried that film production through artificial intelligence models will replace the people working in the sector, especially the artists involved in the production phase of the film. The viewers state that the sector will suffer significant damage due to the widespread use of artificial intelligence in the creative industries and that creativity will gradually decrease if many professions are replaced by artificial intelligence. Looking at the audience comments, it is understood that the recurring theme is that artificial intelligence will take away the profession of artists, and this shift will negatively affect the artistic dimension of the content produced. Concerns such as the need for less human labor due to the widespread use of artificial intelligence and the creation of a sector where films are produced superficially without enough effort are reflected in the comments. Sample audience comments are as follows:

"I wonder what is going to happen with film industry if you can just do a whole movie in one computer. And I'm not even talking about directors or writers, who can still think ideas and direct AI-generated content. I'm talking about technicians, cameramen, DPs, light operators, post producers and endless other jobs that are going to become less required."

"Let's continue to praise AI until people who actually want to create good visuals,

good scripts, and good commercials no longer have the possibility to work.”

“So, AI can write scripts and make movies with minimal human assistance today. Next month, it will do it on its own, and the following month, it will create the greatest movies ever seen. Meanwhile, real actors and writers are out of work...”

Admiration For AI’s Visual Capabilities

In addition to those concerned, many audiences also express their opinions on productive AI models in the creative industries as a significant advantage, especially in enabling the creation of visually and aesthetically extremely satisfying content. A positive aspect of AI is that AI-generated visuals are no different from content created by real artists, and it is almost impossible to tell the difference in aesthetics. In addition, YouTube comments emphasize that the images produced by AI have unique features that can only be produced by a productive AI model, which is remarkably outstanding. Audiences consider generative AI models as a new technology that draws attention to the screen and opens the door to new visual possibilities thanks to their advanced generative capacity. The following are examples of such comments:

“I think this is the first AI generated content I might consider giving the title “ART”.”

“So beautiful, so elegant, just looking like a wow!”

“The number of advancements in just 1-2 years is insane. I can only tell this is AI because some of the logos and organics elements look off.”

“If I didn’t know this was AI generated, I would probably not have known. There are some giveaways, obviously, but especially the shot in the tube with all the realistic people is remarkable.”

Excitement About the Creative Possibilities of AI in Filmmaking

Just as some audiences find AI-generated visuals impressive, some of them consider the use of AI in the creative industries not as a threat to the future of the sector, but as an innovation that will boost creativity, and are enthusiastic

about the future. Audiences who believe that generative AI models will remove many constraints and take creativity to a much higher level emphasize their excitement about these advantages in their comments. Rather than being an obstacle to creativity, they consider generative AI models as a force that will drive the industry forward, providing new creative opportunities and enabling development. In contrast to the general discussion on the use of AI in the creative industries, which sees AI as a means of imitation, some of them expressed that they see generative AI models not only as mechanisms that can imitate the work of real artists, but as a new development that expands the boundaries of creativity and allows for the creation of unique content. In addition, some audiences see the ability of AI to take storytelling to a higher level or to have advantages such as AI-assisted storytelling as an exciting development. Sample comments are given below:

“This is a historic moment. AI-generated movies will go from Youtube to Cannes in 5-10 years.”

“3 years ago, people said this wouldn’t be possible in our lifetime, but look at where we are now.”

“This opens new creative horizons. I’m excited about what AI-assisted storytelling can bring.”

“It’s amazing that AI can make films this good. This could be a really good way to turn audiobooks into visual novels.”

“I’m eager to see more creative Works produced with Sora.”

Critique of the Generation Process

Looking at the YouTube comments of the films, in contrast to the audiences who find the visuals produced by artificial intelligence impressive and see the use of artificial intelligence as an exciting development, there are also many comments expressing criticism of the production process of the films produced by artificial intelligence. Most of them emphasize inconsistencies and anomalies in the films produced with the use of generative AI models, and that this is directly reflected

in the audiences. Specifically, inconsistencies and irregularities were mentioned between scene transitions and unnatural and artificial movements. Examples of audience comments are given below:

“It’s all perfect until you pause the video and stare at one frame. You’ll find many anomalies.”

“There are people floating, scenes with no background, the bicycle is missing the pedal connection.”

“The mimic of the characters when talking, the snow that sticks on the skin without melting, or the way the characters are stepping – it’s disgusting.”

“I did notice a bit of problems with its generation, mostly with the clothes and the cats and humans’ eyes.”

“Everything looks weird, uncanny. Feels like a blurry dream.”

Criticism of AI for Lacking Emotional Depth and Artistic Authenticity

Another criticism of short films produced using generative AI models is how they approach the emotional depth and storytelling of traditionally produced films. Some audiences argue that the widespread use of AI in film production has resulted in a superficiality of storytelling and creatively weak content. Emphasizing that the most important element that makes traditionally produced content and storytelling strong is the complex and non-superficial storytelling that is unique to human beings, the audiences state that this dynamic cannot be adequately met by productive AI models and that this situation leads to the production of emotionally weak and powerless content. Sample comments from the audience are as follows:

“Awesome tool to kill cinematography! Let’s go, I also despise being passionate and investing time into various art forms and medias. In the end, who needs passion and mastery when everyone can do everything with a push of a button?”

“All of you praising this are out of your mind. AI is, and always will be, math. Art

*is the opposite of math. Art and math will always be different. Let me reiterate:
ART AND MATH WILL NEVER BE THE SAME:"*

"The lacks real depth, sophistication, and emotion."

"I don't think AI movies will ever be a thing though. Making movies requires skills. And if AI is making them, nobody will leave a theater going "oh man I can't believe people made that!"

"Let's celebrate together the beginning of the end for any speck of creativity that was present in every industry."

Conclusion

As a result, it is understood that artificial intelligence has a significant impact on cinema, as in other fields, and that this impact is not limited to the employees of the sector but also affects the audience. The impact of artificial intelligence in artistic fields is not only about changing production processes. The inclusion of artificial intelligence models in production processes not only changes most stages of the production process, but also becomes a transformation that affects people's perception, creativity, and perspectives. It is becoming a struggle in which the boundaries between artificial intelligence algorithms and art are sometimes intertwined and sometimes replace each other (Plaza, 2025: 57).

When the YouTube comments of *The Frost* and *Air Head* short films about the use of productive artificial intelligence models in creative industries are analyzed, it is understood that the audiences' opinions do not meet at a single common point and are quite different from each other. One of the most essential points that stands out because of the analysis of the comments is that most of the audiences have a significant concern about the use of generative artificial intelligence models in the creative industries and characterizes the use of artificial intelligence as a great danger for the future of the sector. In addition to being a danger to the industry and potentially replacing real artists, the audiences also expressed concerns that this would have a harmful impact on creativity. On the other hand, there is no small number of audiences who

consider using productive artificial intelligence models in film production as an exciting development for the industry's future.

However, some audiences clearly state that they find the content produced by artificial intelligence very exceptional and fascinating, especially from a visual perspective. In addition, some audiences emphasized that using artificial intelligence in creative industries brings significant advantages in the production process and offers new opportunities in terms of creativity. In particular, the fact that models such as Sora can perceive the physical world at a higher level compared to other models and can perceive how the entered prompts exist in real life also positively affects the visual quality of the content produced (Meyer, 2024), and this can be shown as a situation that can explain why audiences find the content visually remarkable and eye-catching.

On the other hand, when the audience comments are considered, some stated that no matter how advanced the AI model is, it cannot replace the traditional film production process. Numerous technical problems are noticed in films produced by AI. It is emphasized that inconsistencies between scenes or the audience's awareness of major or minor technical issues negatively affect the watching experience. This view of the audience is consistent with some studies that examined the film *The Frost* from a technical perspective and detected abnormalities in the scenes (Liang et al., 2024; Evans, 2024).

As a result, when the comments are evaluated from a general perspective, it is understood that the views on artificial intelligence in cinema differ in a way that emphasizes both positive and negative aspects and that the audience does not agree on a single point. It is an undeniable fact that artificial intelligence has begun to change and transform the film industry, just like it has in other sectors. Plaza (2025) expresses this transformation as follows:

"In a world where the boundaries between reality and fiction blur more each day, we find ourselves on the brink of an unprecedented revolution." (p. 56).

Over time, it is expected that generative AI models will be increasingly used in the film industry. However, this situation has highlighted the contrast

between the dynamics carried by traditional film production processes and the dynamics that exist after AI has been included in the production process. In addition to discussions about whether AI-produced films can reach the emotional and artistic level of traditional films, dilemmas such as who owns the copyrights in films produced by AI models or whether a film produced by a generative AI can produce independently of human guidance are gradually expanding the scope of the discussions (Danesi, 2024). With the widespread use of AI, audiences and industry employees are expected to evaluate the subject more frequently from different perspectives in the coming days. As a result, in the current situation, AI is seen as both a great danger for the sector and an excellent opportunity to increase the potential of creative industries.

References

- Danesi, M. (2024). *AI-generated popular culture: A semiotic perspective* Palgrave Macmillan Cham.
- Evans, A. K. (2024). *The conglomeration of everything: Shifting animation production practices due to artificial intelligence* (Master's thesis). San Francisco State University.
- GlobalData. (2023, November 23). What does AI mean for the future of film? *Verdict*. <https://www.verdict.co.uk/what-does-ai-mean-for-the-future-of-film/>
- Låvenberg, T.** (2024). *Authorship in AI cinema: The Frost through the lens of Walter Benjamin* (Bachelor's thesis, Stockholms universitet, Department of Media Studies).
- López Frías, C. (2024). The paradox of artificial intelligence in cinema. In F. B. Gil & P. F. Alves (Eds.), *Comunicação, artes e culturas* (pp. 5–25). CDIG, Cultura Digital.
- Meyer, R. (2024). It's a flat world. The Synthetic Realities of Sora. *rrrreflect Journal of Integrated Design Research*, 3.
- Plaza, J. (2025). Creativity in the AI revolution. In A. Baraybar-Fernández, S. Arrufat-Martín, & B. Díaz Díaz (Eds.), *The AI revolution: How technological developments affect the audiovisual sector* (pp. 55–62). Springer.
- Ramesh, A., Pavlov, M., Goh, G., Gray, S., Voss, C., Radford, A., Chen, M., & Sutskever, I. (2021). *Zero-shot text-to-image generation* (arXiv:2102.12092). arXiv.
- Schwendener, L. (2024). *Reasserting what it means to be human in the age of artificial intelligence* (Bachelor's thesis). Purchase College, State University of New York.
- Sheikh, H., Prins, C., & Schrijvers, E. (2023). Artificial intelligence: Definition and background. In *Mission AI: The new system technology* (pp. 15–41). Springer International Publishing

IMPACT OF ALGORITHMS ON THE CONSTRUCTION OF TEACHERS' SOCIAL REPRESENTATION

Aline VERONIZE*, Luís LOUREIRO**, Alexandra FIGUEIRA***

Abstract

The use of software, platforms, connectivity, and data evokes new concepts of socialization. However, even within this new and complex digital architecture, communication still operates according to paradigms that serve the logic of capital. One of the ways in which technologies reshape social interactions is through their application in the construction of meaning. Social media platforms are part of a structure for the organization, management, and distribution of cultural symbols that aims at the control and the maintenance of power. We ask ourselves: how does the technological landscape influence the perception of specific professions? What dynamics of social media platforms contribute to the (in)visibility of essential causes, such as teaching? This article is based on a literature review that explores the relationship between the technological landscape and the construction of social representation. It then analyzes the duality of the digital environment between promoting visibility and constructing invisibility. Finally, we investigate how this construction occurs in the context of education, with a focus on the teaching profession. It is concluded that, despite the (ir)relevance assigned by algorithms to certain topics, it is possible to break the strategies that perpetuate domination through tactics based on knowledge of this new digital architecture, composed of various technologies and new, more conscious and objective uses. To serve the logic of collective interests, platform regulation and digital literacy are necessary.

Keywords: algorithms, digital technologies, social representation, education, networks.

* Graduate Student, University of Minho, averonezereporter@gmail.com , ORCID: 0000-0001-5734-3360

** Full Professor, University of Minho, luisloureiro@ics.uminho.pt , ORCID: 0000-0003-2091-5067

*** Assistant Professor, Lusófona University-Porto, alexandra.figueira@ulusofona.pt , ORCID: 002-8764-5690

Introduction

Technological development (re)creates the material environment in which we relate by bringing human and non-human actors onto the social stage. Individuals and companies compete for inclusion in an interest-based economy. Structure, content, context - everything is changing and circulating very fast. This whole scenario has an obvious impact on the way we live our lives. But does this affect symbolic representations? For example, does it affect how and what we think about teaching?

Based on bibliographic research and a critical-reflexive analysis, this paper will start from a theoretical framework on social representations, articulated with studies on what is social, the effects of materiality, and the role of Social Communication in the current context. Only then, by bringing this theoretical framework together, does it return to the question of the historical trajectory of the representation of teaching.

The text is constructed from five analytical moves. The first is a brief review of social representations. This is followed by an attempt to summarise the process of the (construction) of representations. The third movement is a consideration of the transformations in the materiality of communication that surround, constrain, and expand us, affecting our relationships and even the concept of the social. It refers to the context of the algorithmic nature of social media platforms and their subtleties in the construction of visibility. The focus is then on the construction of social representations of teaching.

To do this, it was necessary to look at the historical perspective of the concept of the profession. Which institutions have structured society and culture in such a way as to construct what and how we think about the teaching profession? Which social contexts have had an impact on this symbolic construction? Finally, in the fifth movement, an effort was made to build a theoretical framework for understanding the impact of relationships mediated by social media platforms and algorithms on the dynamics of teachers and how we recognise the teaching profession.

The Concept of Social Representation

The beginning of studies on social representations can be traced back to the theories of Serge Moscovici in anthropology and social psychology in France in 1960. Moscovici addressed the social process of knowledge production and the debate around social representations. De Oliveira (2004) summarises his aims:

He was interested in understanding how the group/action/idea triangle constitutes and transforms society (...) He wanted to understand how the production of plural knowledge constitutes and reinforces the identity of groups, how it affects their practices, and how it reconstitutes their thinking (De Oliveira, 2004, p. 181).

Social representations form the collective consciousness, which enables the explanation of objects, events, and actions, making them accessible. They form units and sets of ideas that add interpretative security and keep institutions and relationships alive, through nuclei of stability that consolidate the group. They make up a cluster of concepts, propositions, and explanations that originate in everyday life, in the course of interpersonal communications. They are representations of reality that end up moving individuals and collectivities.

Representations are based on the ‘reference of a pre-existing thought’ and are ‘of the same nature as the social groups that create them’, so they have no universal meaning (Moscovici, 2003, p. 216). However, this symbolic power of past concepts and the present social context also faces the appropriations made by individuals. According to Jodelet, a representation ‘corresponds to an act of thought by which the subject relates to an object’ and ‘bears the mark of the subject and their activity’ (Jodelet, 1989, p. 35). The author emphasises that this socially elaborated and shared knowledge has ‘practical objectives’ and acts to ‘construct a reality common to a social group’ (Jodelet, 1989, p. 36).

Social representations make up ‘organised patterns of historically produced, collectively shared meanings, associated with socio-cognitive processes and in interaction with social structures’ (Mazzotti & Campos, 2011, p. 6). It does so in such a way that it is ‘prescriptive’, guiding ideas about a certain object (Vergara & Ferreira, 2005, p. 1142).

The following explains how social representations change and what factors influence these transformations.

The Evolution of Social Representations

The dynamic of forming new representations starts with recognition and understanding through familiarisation. It is a double mechanism consisting of the 'tying' or 'anchoring' of concepts, which classifies and labels what is unknown; and 'objectification', which combines real, concrete, and comprehensible images (Oliveira, 2004). Therefore, new representations are constructed based on previous concepts, images, and models: 'memory predominates over logic, the past over the present, response over stimulus and image over reality' (Moscovici, 2003, p.5).

Individuals develop their social identity mediated by social representations, in the unfolding of tensions between the world that is accessible to them and their efforts at individualisation. By recreating social reality, they modify their relationship with the world. However, this clash is marked by the fact that 'his thought is constructed, in many ways, by the fact that others also think it' (Jodelet, 1989, p. 24).

The relationships between symbolic forms and social structures make up a psychological organisation, a form of knowledge constructed in a dynamic, interactive way that is irreducible to a given society. Transformations in contexts and individual appropriation harbour the potential for change in the evaluation of the qualities of an object, event, or act, as a result of cognitive dissonance arising from distortions, aggregations, or suppressions. In distortion, the attributes are present, but they are accentuated or minimised; in aggregation and suppression, a characteristic is added or removed in such a way as to change the way the object, event, or act is understood (Jodelet, 1989).

The last few decades have seen intense transformations in the social conditions of production and circulation of representations. Changes in communication, which is essential for sharing the cultural, social, and coercive

base. The following attempt is made to highlight the evolution of communication processes and changes in materiality that have an impact on the social representations we share.

The Impact of Social Communication on Symbolic Constructions

Like Vygotsky, who adopted a historical approach guided by the notion that 'behaviour can only be understood as the history of behaviour' (1989, p.11), Communication needs to be studied as the history of Communication.

Harold Innis and Marshall McLuhan pointed out the importance of communication media, which were considered only as a means of getting a message to its destination, but which alter the environment in which communication exchanges take place and 'are not passive envelopes, but active processes' (McLuhan, 1971, p.8). Social structures were once essentially rigid (Foucault, 1975). Experiences, and therefore references, depended on the individual's spatio-temporal location.

The 'Mechanical Age' of the first Industrial Revolution surprised the cognition of the masses by imposing speed on everyday life. If, until then, language was the main mediation, the 'illuminated crack that represented reality' (Minayo, 2011, p. 90), it was joined by other mediations that added varying degrees of clarity and sharpness to reality. The legendary screening of Arrival of the Train at La Ciotat Station in 1896, presented by the Lumière brothers and the accounts of the supposed astonishment of the horrified spectators at a moving train, illustrates the impact of the dissemination of the language of audiovisual productions as a new form of knowledge.

In the 'Age of Electricity', the mechanisms for the mass dissemination of information were strengthened. When Moscovici studied social representations, the radio was popular mainly because it reached the illiterate. Television was in its infancy and yet it stood out to the author as signalling the arrival of a time of 'fewer rigid controls' (2003, p. 43). The structure was completely vertical, and information flowed from the top down, from the powerful centres of society

to the audience. However, the masses' access to information from television programming harboured a transformative potential: it brought individuals into contact with different values and added concepts and new information about ways of life that went beyond what was available for access in their group of origin, in their locality, in their time.

The advent of the Internet and the development of services to cater for effective user participation have exponentially increased the speed of access to other subjects, contexts, and an extraordinary volume of data (Santaella & Lemos, 2010). Pierre Lévy (2007) named the result of this emerging world culture 'Collective Intelligence', in which the symbolic systems of different human communities have come to live in symbiosis.

Cultures have undergone a hybridisation in which rigid identities have given way to multiple identities; the search for the truth has been replaced by the primacy of versions, and citizens have essentially become consumers in search of new experiences and sensations (Canclini, 2012). New relations of information exchange from 'many to many', conditioning the scope and format of messages, changing the cultural environment. What happens locally, offline, in a given geographical coordinate, now has a major influence through online interactions, which take place at a global level. Culture has become cyberculture (Lévy, 2000).

The possibilities for contact with different signifiers and meanings have increased. Through 'augmented reality' equipment, communication can take place at deep levels of the subject's sensory field, through intense experimentation, with the production of effects of presence and of meaning much more cognitively impactful than before (Pereira, 2004; Regis et al., 2012). The inclusion of non-human actors in the concept of the social is not restricted to technological devices but refers to the informational protagonism of data that comes from the ecological context and the internet of things in general, 'creating an unprecedented type of connective and transorganic ecology' (Di Felice, 2018, p. 30).

From this context of extreme datification, the selection and distribution of content have changed. The gatekeeping carried out by journalists and media editors based on deontological values is now operationalised by algorithms and even social media users (Bortolazzo, 2022; Couldry & Mejias, 2019). However, the confidence that algorithms are fair because they are based on mathematical models is misplaced. For Catherine O’Neil (2021), society is racist and sexist, and algorithms can reinforce these stereotypes.

Social networks like Facebook and Instagram are designed not with an algorithm, but with a set of them (Makortoff, 2022). They are private ventures whose criteria do not follow principles of transparency. The decisions to manipulate access to information and the relational network of each of the thousands of users are so complex that the developers have not been able to explain how they work (Khatri, 2021). The information available is that the selection and distribution of content is determined by the user’s previous interest (both regarding the format and content of the communication); viewing time; relevance of the topic within the network; degree of interaction with the content creator, among other criteria.

The use of hashtags by users in shared content increases the possibility of achieving reach and visibility, because it links conversations with a particular theme. These tags also highlight the most popular topics. However, the same strategy serves the platform’s control mechanisms: hashtags such as #always, #boatarde, #medicina, and #sextou have been banned, and users who use them can have their content and/or account deleted.

Algorithms keep users in bubbles (Pariser, 2011; Pellizzari & Barreto Júnior, 2019) in which they are exposed to certain information and also to other users with similar opinions, without major divergences of ideas that could fuel some critical thinking (Ferreira & Rios, 2017). This configuration acts to polarise opinions and generate extremism by undermining the ability to deal with opposing arguments (Bolin et al., 2025), crystallising thoughts (Wohn & Bowe, 2014).

The materiality of these platforms forces us into certain ways of doing things, such as incessant screen scrolling, which even affects human physiology: it creates addiction, depression, and social dysfunction (De Andrade & Alves, 2024; Faelens et al., 2021; Mullender & Haidt, 202).

Algorithms measure connections and interactions between users, whether they are individuals or companies. Platformisation is also advancing in citizens' relations with governments and other essential sectors. 'Areas such as education, health, entertainment, the economy and transport are also following the logic of platforms' (Van Dijck et al., 2018, p. 4).

These reconfigurations of ways of feeling, knowing, and relating are giving rise to a new era of social representations. We wonder how the context influences the construction of perceptions of the teaching profession. In order to understand, we will continue in the archaeological movement, by rescuing the anchors, identifying the prevailing power structures and the social circumstances of the historical construction of the representation of teaching.

The Construction of the Teaching Profession

In tribal societies, education was diffuse, did not take place in a defined space, and there was no social control over the teaching process. It took place in interactions, in everyday activities, and all members of a community were responsible for educating.

Until the age of 7, when they had to start living at their own expense, the children accompanied the adults in all their work, helped them to the best of their ability and, as a reward, received their share of food like any other member of the community (Ponce, 1989, p. 19).

In the Ancient Ages, river civilisations such as Mesopotamia and Egypt became complex societies that developed river channelling systems, irrigation mechanisms, built large buildings, and began to carefully observe and record natural phenomena. The need arose to train people for specific jobs, as well as to record, transmit, and increase the knowledge acquired (Reale, 2015).

In Egypt, access to education was restricted to members of the pharaoh's family, priests, nobles, and officials. The rest of the population, artisans, peasants, slaves, and soldiers did not receive formal education, and instruction took place during the practice of labour activities, to subordinate subjects and subjugated by physical punishment. Training began in adolescence and lasted until adulthood (Lima & Lima, 2019).

It was in the classical civilisations, in Greece and Rome, that teaching took on new meanings. Having teachers continued to be an indication of privilege and a way of maintaining it, but it was not a condition for exercising citizenship. It was possible to vote or act in the assembly with the help of fixed labourers who read and wrote documents. These readers and writers were trained in the educational process, the *Paideia*, which included the acquisition of knowledge and physical training from childhood, in public schools.

Students from higher social classes had classes with private teachers or tutors: the *páidotribés*; the *grammatistés* or the *kitharistés*, who were responsible for intellectual development, writing and reading skills, and physical improvement, respectively (Amado, 2007). The younger children were accompanied by a slave, an elder, usually of barbarian origin, known as a *paidagogo* (pedagogue), who looked after the student's safety, helped with homework, and instilled good behaviour and posture.

This was the case until Lycurgus, an influential politician in Athens, established that the children would be managed collectively by a magistrate (*paidónomo*), who would have armed teenagers at his service to ensure order. The boys' education could continue in lessons with the sophists, much more expensive instructors than conventional teachers (Amado, 2007). Sparta was a pioneer in establishing universal education. In the absence of the teacher, another citizen was in charge of supervising and instructing the children, transmitting the social values in force and punishing any deviants with exemplary behaviour.

The figure of the teacher gained some distance from the religious leader and became linked to the possession of specific knowledge. In addition, the link

between formal education and the violent and repressive control of that society was reinforced.

The Greeks became known for their wisdom, and many went to teach in Rome, where the social context imposed other dynamics, especially concerning the subjects studied and the methodology applied. Athletics and music were not given the same space. The teachers also realised that the violence of punishments disciplined bodies but was not effective in forming full citizens, capable of arguing courageously and with a convincing oratory. Among the privileged, physical punishments were replaced by playful activities that kept the students interested. There are records of teaching methods that included sweets in the shape of letters, the creation of recreational breaks between reading and declamation shifts, as well as playful games that encouraged the pursuit of excellence through competition. 'Study was to take place in a space of joy (*schola*). Reading and writing were taught by the *ludi-magister* (toy master)' (Gabriel & Freitas, 2019, p.74).

With the advance of the Roman Empire, the school became official and strictly dependent on the state, whose intention was to train eloquent orators to act in the vast territory and to tame the conquered populations so that they would become admirers of Rome. According to the tradition of the *mos maiorum*, education should have a practical and social character, providing children with the necessary knowledge to exercise the profession of soldier or landowner and instilling the individual's subordination to a higher ideal - Rome and the *Res Publica* (Lima & Lima, 2019).

The masters were professionals protected by Caesar and Augustus. Quintilian was the first state-paid teacher in Vespasian's empire and a great promoter of methodological changes. Vespasian publicly recognised the social usefulness of teachers and promoted an extensive series of rewards and tax immunities for grammarians and rhetoricians. However, as children were not very important, the *primus magister*, the primary teacher, was poorly paid and not highly regarded in the social hierarchy. The state also invested in the creation

of grandiose libraries, which signalled to the people the importance of culture. These measures helped create a positive social representation of studies and teachers.

However, there was a major setback when the Catholic Church strengthened its power structure and began to control access to knowledge throughout the Middle Ages. 'Encyclopaedic knowledge was taught by priests of the Church and was based on copying and translation' (Silva, 2004, p. 23). The church began to license the new schools and teachers. It also took over higher education so that 'the doctor of a university should also be a doctor of the Church' (Gabriel & Freitas, 2019, p. 104), reinforcing the anchoring of the teacher to religion.

In the following centuries, the struggle for influence and power drastically changed the structure of society and the social context, which led to changes in many aspects of education. There was a gradual disappearance of the fiefdoms and the growth of cities in Europe. Portugal and Spain opened up maritime trade routes. The state began to intervene heavily in the mercantilist economy, which generated a reaction with liberal ideas. The strengthening of political structures resulted in the weakening of religious structures and the consequent Reformation and Counter-Reformation movements.

Many merchants had not received an erudite education like the nobles, but an instrumental, labour-oriented education. Based on this perspective, continuing education was for the elite. Therefore, with social ascension, many bourgeois invested in education. They had the money, now they would also have the status that came from a more formal education. However, this education focused on practice, on the 'know-how' that was part of the bourgeois identity. It was a period of many studies focused on medicine, for example (Amado, 2007). The contempt for the unproductiveness of the previous elites was evident in cultural productions. In shows, teachers were referred to as pedants, and the 'knowledge for knowledge's sake' that had prevailed among intellectuals was laughed at (Shakespeare, 2013).

The Reformation established universal public primary education, but it was still geared towards strengthening faith and discipline. Studies were limited to one or two hours a day. The rest of the time was dedicated to learning a trade. For the majority of children and young people, education was a way of disciplining their bodies for the labour context, instrumental knowledge, as a technique (Silva, 2004).

The Counter-Reformation took teaching and violence far beyond Europe. It was also concerned with docilising teachers. Manuals were provided with rules for teaching work, bibliographies to guide the teaching profession, and surveillance to ensure unity of thought and action. A non-Jesuit agent was appointed to administer beatings and punishments, and entertainment was provided to prevent rebellions (Silva, 2004).

With the expansion of commerce and factories, teachers became essential to organise society and to meet the demand for technically qualified labour for the new trades that were emerging. In the 17th century, efforts were made to institutionalise schools, drawing up regulations on issues such as compulsory education, programmes, levels, and methods. Germany regulated compulsory schooling for children aged 6 to 12 and defined the training of teachers, their levels, working hours, and examinations. France established free schools and seminars for teacher training. The biggest French investments took place in the Lyon region, an important manufacturing centre.

In the 18th century, education became a priority among the concerns of kings, thinkers, and politicians. The perception of the teacher as the one who prepares for a prosperous future grew stronger with the Industrial Revolution. Technical and vocational education became the gateway for former peasants into industry. The future labourer needed to develop motor and cognitive skills to meet demand. The context of education was then conformed to the industrial context: a teacher, holder of the questions and answers and imbued with authority, taught large classes through repetitive and mechanical activities using standardised methods.

Between the 19th and 20th centuries, the New School Movement propagated the idea that the purpose of education was to enable pupils to move from consenting authority to increasing autonomy (Dias, 2024). The primary school teacher gained importance. Teachers began to organise themselves as a class. On the other hand, this was also a period when dictators took advantage of illiteracy and created campaigns to rescue education that was essentially instrumental and based on monitoring and punishing. This is the case of Salazar's emblematic speech in Portugal:

(...) When they learnt to read and write, their ambitions were born they wanted to go to the cities to be sailors, clerks, gentlemen; they wanted to go to Brazil. They've learnt to read! What do they read? Crime reports; misconceptions about politics; bad books; subversive propaganda leaflets. They give up the hoe, they don't care about the land, and they only have one ambition: to become civil servants. What advantages did they get from school? None. They gained nothing. They lost everything. Happy are those who forget their letters and go back to the hoe. The most beautiful, strongest, and healthiest part of the Portuguese soul lies in those 75 per cent of illiterates (Mónica, 1978, p. 119).

After the Second World War, the perception of education as an essential tool for progress increased. Defeated Japan, for example, implemented measures to rebuild the country, including setting a salary 30 per cent higher than any other civil servant in order to retain and attract teachers. Teachers' prestige was also built up through the criteria created for appointment to the post, permanent and compulsory development programmes, and the need to renew certifications. Teachers are still addressed by the honorific term of respect 'sensei' and are the only ones who don't have to bow to the emperor.

The post-war period also saw the development of the Human Capital Theory by a group from the University of Chicago, which proved that education contributed to the economy by improving the quality of workers' training (Schultz, 1973). This finding led to a quest to valorise these professionals in the following decades.

This was followed by a period of intense technological development, as

discussed in the previous topic, which had an impact on the ways of knowing the teaching profession.

Teachers in the Age of Algorithms

The difficulty of updating management systems, the school institution, and the role of the teacher for this new social structure has been going on for decades. In 1971, McLuhan pointed out the discrepancy between the architecture in which students constructed their thinking and what they found in schools: Today's young student grows up in an electrically structured world (...) Today's student lives mythically and in depth. At school, however, they find a situation organised according to classified information. Subjects are not related. They are visually conceived in terms of a project or architectural plan. The student finds no possible way of participating in it, nor can he discover how the educational scene connects to the mythical world of electronically processed data and experiences that is peaceful to him (McLuhan, 1971, p. 11).

Today's students, digital natives, have known no other world than the one already platformised and organised by algorithms. With cognition developed within an ecology that includes human and non-human actors, these subjects have other ways of knowing, of symbolic production. Thus, Digital Natives are potentially more capable of solving problems immediately, without an assimilated method or strategy for such a task, however, it is notorious that gains in mathematical reasoning skills and vocabulary have been decreasing over the generations. (...) In this context, therefore, the question arises: is the school prepared for this 'new student'? (Indalécio & Ribeiro, 2017, p. 146).

There are physiological changes in the cognitive pattern they develop, in hormone production, in the development of empathy and criticality and in the ability to socialise outside the digital environment (Alcântara & Almeida, 2025; Azevedo & Santos, 2025; Desmurget, 2023; 2021; Maia et al., 2025; Santos et al., 2025; Sepulveda & Gonçalves, 2025). Bodies have become docilised for an economy whose main product is attention. The economy of attention is one of the important vectors in today's societies that participates in the processes of

subjectivity production, the effects of which have not yet been fully mapped or stabilised. (...) its lack tends to be conceived as an individual problem or biological deficit that underpins forms of pathologisation (...) and justifies medicalisation measures (Bentes, 2021, p. 234).

Despite the possibility of movement, the world organised by algorithms is limiting from the inside out: from the programming it establishes in the user. The structure ends up configuring young people's behaviour to the point where they have little initiative to leave the bubbles that the algorithms have determined for them. Generally speaking, although these teenagers are frequent Internet users and spend a lot of their time on social networks, they do little to exploit the potential and opportunities that these media can offer them, showing a high level of access and use, but little variety of activities and little depth in their realisation (Pereira, 2014, p. 140).

Without the propensity to move outside the norm, placed in content bubbles, with little exposure to divergent thinking, the consumption of information selected by the app has a significant influence on the construction of symbolic representations. Certain opinions and news are 'pushed' into the spotlight; others simply disappear or are ignored. In fact, filtering is not impartial. It follows criteria that presuppose what may or may not be interesting or important to users, and in the case of social networks, the strategies are economic, commercial, political, professional, behavioural, pedagogical etc. (Bortolazzo, 2022, p.7).

Education appears on social media in interactions between teachers and their networks, in conversations that address the challenges of the profession and exchanges of experience and teaching materials. Because it is outside the bubbles of topics that other users have previously shown an interest in the algorithm limits the delivery of content, and teachers' issues have little reach and engagement among other sectors of society.

In addition to this invisibility, the platforms also change the way teachers are perceived by increasing the visibility of grievances by echoing

extreme positions. 'These subjective complaints and concerns of teachers find a favourable environment for dissemination, debate and questioning' (Ferreira et al., 2016, p. 200).

A group of researchers who studied the social representation of teachers on Facebook fan pages came to the following conclusion: 'The results of the analysis show that the teacher is represented as a subject "enslaved by work", since society sees them as someone who works exhaustively, is underpaid and is therefore an oppressed and unsuccessful subject in life' (Dieb et al., 2014, p. 706). The authors explain that this construction is made through humour and mockery, but is not naïve, which brings us back to the concept that social representations are socially elaborated and shared knowledge with 'practical objectives and act to 'construct a reality common to a social group' (Jodelet, 1989, p.36).

At the same time as this ingenious and ideological representational construction about the teaching professional, the greatest and most devoted expectations of parents and public administrators still fall on him, as if the teacher were responsible for all educational problems and as if he could solve them with goodwill alone. (...) This fact seems to feed back into the representation of the teacher, which, as we have seen, is propagated with a certain tone of derision but disguised by playfulness, the purpose of which seems to be to make the teacher look good.) This fact therefore seems to feed back into the representation of the teacher, which, as we have seen, is propagated with a certain tone of derision, but cloaked in playfulness whose purpose seems to be to soften the discursive contradictions surrounding this relevant social role, which is that of the teacher, seen simultaneously as the one who occupies the most sublime of professions (the teaching profession), but who doesn't have the social and financial recognition he deserves, and is therefore a failed professional (Dieb et al., 2014, p. 724), 2014, p. 724).

Within the bureaucratic structures of governments and schools, the role of teachers has become limited, and the methodology does not match the new

ways of ‘knowing the world’. The teacher himself has become what the structure wants him to turn his pupil into:

(...) an executor of orders within the school bureaucracy, his intellectual work of fundamental importance for the development of the teaching-learning process disappears, he becomes a mere executor of an educational policy drawn up in the office. (Braido & Ferreira, 2022, p. 631).

For teachers who don’t have the same intuitive skills as their students to move around networks, the appropriation of new technologies is manifested through the ‘accent’ (Indalécio & Ribeiro, 2017). This appropriation consists of using the platforms as environments for pedagogical practices or even as teaching resources (Luckin & Holmes, 2016).

The (in)visibility of teachers and their content is conditional on their ability to perform; to persuade, to submit to extra hours of unpaid work and practices of self-coercion for visibility that meet the demands established in the ‘economy of fame’ (Campanella et. al., 2018). The representation of the teacher undergoes a distortion when the authority of the old masters who transmitted information is removed and anchoring in the figure of influencers is added, reconfiguring the processes of (de)legitimation of self-entrepreneurship (Fantoni, & Barichello, 2020, p. 3).

The greater the datafication and technological increase, the more central personal interactions, the human factor, have become (De Felice et al., 2023). Claudia Costin, visiting professor at Harvard and director of the Centre for Excellence and Innovation in Educational Policies at the Getulio Vargas Foundation in Brazil, explains: ‘Brain studies show that adults already have a fully developed prefrontal cortex and can learn from non-human interactions, while children and adolescents cannot. And the smaller the child, the more it learns from human interaction.

Conclusion

The implication of the historical context and the evolution of the media in the

subjective sphere was highlighted. Along the way, it was noted that education has shifted from 'knowing for the sake of knowing' and 'knowing how to do' to 'knowing how to be'.

Algorithmic mediation has accelerated the process of recognition and anchoring and has broken with traditional dynamics by inserting a technological filter that favours engagement metrics over collective deliberation. Certain representations become dominant not necessarily because of their social relevance, but because they fit into the artificial logics of visibility.

There has been a reduction in the initial anchoring of the social representation of teachers in religiosity, in the violence of controlling bodies, in education as a class privilege, or reduced to instrumentalization for the labour market. In fact, in the contemporary context, teachers prepare students for professions that don't even exist yet.

It is also possible to identify a distortion in the authority that comes from holding the knowledge to be transmitted and in the intellectual distance between teachers and students. Digital natives now master technological tools. The role of teachers has been re-signified as that of guiding the student through the excess of data, in the construction of knowledge. Education must focus on identifying sources, checking veracity, articulating content, appropriating it, and applying it to other dynamics in life. In this sense, algorithms have been used to optimise teachers' time with bureaucratic tasks and performance analyses, and to allow closer monitoring of each student.

Another conceptual distortion concerns the dynamics of creating legitimacy. In the past, the spectacle of teaching took place in the classroom. Today, it has expanded into cyberspace. Anchoring in the figure of priests and prolific pedants has given way to influencers.

It is essential to recognise the extra work involved in platforming performances as a strategy to engage digital natives and the impact of algorithms on teaching practice, restructuring routines and promoting training that prepares

educators to face the challenge of adapting to the social media environment, where their voices and practices can be obscured by commercial and political dynamics. The importance of human interaction as a central element of the educational process is highlighted, in which humanised teachers are able to deal with individuals who need to ‘hominise’ themselves.

Teachers are also subject to algorithmic logic in their hiring and evaluation, although there are serious ethical issues involved, and examples of this maths making serious evaluation errors, highlighting the importance of maintaining its role as an auxiliary tool and not as the final decision-maker.

References

- Alcântara, M., & Almeida, C. (2025). Intoxicação digital na primeira infância. *Revista Eletrônica Acervo Saúde*, 25, e19326-e19326.
- Amado, C. (2007). História da Pedagogia e da Educação.
- Azevedo, A., & dos Santos, L. (2025). A construção do pensamento crítico: Perspectivas históricas, educacionais e contemporâneas. *Revista Contemporânea*, 5(1), e7298-e7298.
- Bentes, A. (2021). Quase um tique: economia da atenção, vigilância e espetáculo em uma rede social. *UFRJ*. [recurso eletrônico]. ISBN 978-85-7108-480-3
- Bolin, G., Braga, J., Rosa, A., Silveira, A., Dalmolin, A., Carlón, M., Gomes, P., Borelli, V., Ferreira, J., Löfgren, I., Kiriya, I., Graf, H., Miège, B., & Fausto Neto, A. (2025). *Plataformas, algoritmos e IA: Questões e hipóteses na perspectiva da midiatização*. [recurso eletrônico]. ISBN 978-65-5773-059-1
- Bortolazzo, S. (2022). O Dilema das Plataformas e Redes Digitais: processos educativos, docência e neoliberalismo. *Cadernos de Educação*, (66).
- Bortolazzo, S. (2022). O Dilema das Plataformas e Redes Digitais: processos educativos, docência e neoliberalismo. *Cadernos de Educação*, (66).
- Braido, L. & Ferreira, L. (2022). (Des) Valorização do Trabalho dos (as) Professores (as) de 1946 até 1996: uma Análise dos Movimentos de Sentidos. *Revista de Ensino, Educação e Ciências Humanas*, 23(4), 624-632. p 631. <https://doi.org/10.17921/2447-8733.2022v23n4p624-632>
- Campanella, B., de Nantes, J., & Fernandes, P. (2018). Criando intimidade, recebendo visibilidade: novas Comunicação, *Mídia e Consumo*, 15(43).
- Canclini, N. (2012). *Culturas híbridas*. Debolsillo.
- Caniato, A. (2000). Daviolênçianoethos cultural autoritário da contemporaneidade

- e do sofrimento psicossocial. *Revista de Psicologia Social e Institucional*, 2(2), 197-215.
- Castells, M. (2005). *A sociedade em rede* (Vol. 1, No. 6). São Paulo: Paz e terra.
- Castro, T., & Machado, L. (2022). A relação família-escola como elemento das representações sociais da profissão docente veiculadas na rede social Facebook. *Revista Educar Mais*, 6, 768-782.
- Couldry, N., & Mejias, U. (2019). *The Costs of Connection: How Data Is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press. Stanford.
- De Andrade, T., & Alves, N. (2024). Conectividade e saúde mental: como as redes sociais influenciam o bem-estar psicológico dos usuários. *Revista JRG de Estudos Acadêmicos*, 7(15). <https://doi.org/10.55892/jrg.v7i15.1301>
- De Felice, S., Hamilton, A. F. D. C., Ponari, M., & Vigliocco, G. (2023). Learning from others is good, with others is better: the role of social interaction in human acquisition of new knowledge. *Philosophical Transactions of the Royal Society B*, 378(1870), 20210357.
- De Oliveira, M. (2004). Representações sociais e sociedades: a contribuição de Serge Moscovici. <https://doi.org/10.1590/S0102-69092004000200014>
- Desmurget, M. (2023). *Faça-os ler! Para não criar cretinos digitais*. Vestígio Editora.
- _____. (2021). *A fábrica de cretinos digitais: Os perigos das telas para nossas crianças*. Vestígio Editora.
- Dias, É. (2024). A Educação e os novos tempos. *Ensaio: Avaliação e Políticas Públicas em Educação*, 32(122), e0241221. <https://doi.org/10.1590/S0104-40362024003201221>
- Dieb, M., Araújo, J., & Vasconcelos, J. (2014). A representação social de professor em fanpages do Facebook. *Revista Linguagem & Ensino*, 17(3), 705-726. <https://doi.org/10.15210/rle.v17i3.15311>
- Duarte, S., Mamede, M., & Andrade, S. (2009). Opções teórico-metodológicas em pesquisas qualitativas: representações sociais e discurso do sujeito coletivo. *Saúde e sociedade*, 18, 620-626. <https://doi.org/10.1590/S0104-12902009000400006>
- Faelens, L., Hoorelbeke, K., Cambier, R., Van Put, J., Van de Putte, E., De Raedt, R., & Koster, E. (2021). The relationship between Instagram use and indicators of mental health: A systematic review. *Computers in Human Behavior Reports*, 4, 100121. <https://doi.org/10.1016/j.chbr.2021.100121>
- Fantoni, F., & Barichello, E. (2020). A construção midiaticizada do sujeito. *Anais de Resumos Expandidos do Seminário Internacional de Pesquisas em Mídiação e Processos Sociais*, 1(4).
- Ferreira, A., Santos, E., & Rosso, A. (2016). Representação social da indisciplina escolar. *Psicologia: Teoria e Pesquisa*, 32, 199–208. <https://doi.org/10.1590/0102-37722016012074199208>

- Foucault, M. (1975). *Discipline and punish*. A. Sheridan, Tr., Paris, FR, Gallimard.
- Gabriel, A., & Freitas, G. (2019) *História da educação*. Telesapiens. [recurso eletrônico]. ISBN: 978-85-54921-37-8
- Galinkin, A., & Zauli, A. (2011). Identidade social e alteridade. Torres, C., & Neiva, E. (Orgs.), *Psicologia Social: principais temas e vertentes*, 253-261. Artmed.
- Hanke, M. (2006). Materialidade Da Comunicação—Um Conceito Para A Ciência Da Comunicação? *Interin*, 1(1), 1-8. E-ISSN: 1980-5276
- Indalécio, A., & Ribeiro, M. (2017). Gerações Ze Alfa: os novos desafios para a educação contemporânea. *Revista UNIFEV: Ciência & Tecnologia*, 2(1)
- Jodelet, D. (1989). Représentations sociales: un domaine en expansion. *Les représentations sociales*. Paris: PUF, 31-61.
- Khatri, M. (2021). How Digital Marketing along with Artificial Intelligence is Transforming Consumer Behaviour. *International Journal for Research in Applied*. <https://doi.org/10.22214/ijraset.2021.36287>
- KhosraviNik, M. (2022). Digital meaning-making across content and practice in social media critical discourse studies. *Critical Discourse Studies*. (2), 119-23. <https://doi.org/10.1080/17405904.2020.1835683>
- Lévy, P. (2000). O ciberepaço como um passo metaevolutivo. *Revista Famecos*, 7(13), 59-67. <https://doi.org/10.15448/1980-3729.2000.13.3081>
- _____. (2007). A inteligência possível do século XXI. *Revista FAMECOS*, 14(33), 13-20. <https://doi.org/10.15448/1980-3729.2007.33.3429>
- Lima, I. & Lima, X. (2019) Os paradigmas da educação nas civilizações antigas e a preparação dos pesquisadores da antiguidade. *Revista Científica Multidisciplinar Núcleo do Conhecimento*. 04(02), 100-109. ISSN:2448-0959.
- Luckin, R., & Holmes, W. (2016). *Intelligence Unleashed: An argument for AI in Education*
- Machado, I., & Ramos D. (2019). Alfabetização semiótica com os códigos informático-digitais da internet. *Comunicação & Educação*, 24(2), 38-53. <https://doi.org/10.11606/issn.2316-9125.v24i2p38-53>
- Maia, J., Ribeiro, M., Rodrigues, J., da Silva Varge, H., Viotto, L., Viotto, C., ... & Cavalcante Filho, A. (2025). Influência das telas digitais no desenvolvimento neurológico de crianças. *Journal of Medical and Biosciences Research*, 2(1), 562-573.
- Makortoff, K. (2022). Risks posed by AI are real: EU moves to beat the algorithms that ruin lives. *The Guardian*. Disponível em: <https://www.theguardian.com/technology/2022/aug/07/ai-eu-moves-to-beat-the-algorithms-that-ruin-lives>. Acesso em 28 de fevereiro de 2025.
- Mazzotti, A. & Campos, P. (2011). Cibercultura: uma nova “era das representações sociais”. *Teoria das representações sociais*, 50, 457-484.

- McLuhan, M. (1971). Os meios de comunicação como extensões do homem. *Tradução de Décio Pignatari. São Paulo.*
- Minayo, M. (2011). O conceito de Representações Sociais dentro da sociologia Clássica. Guareschi, P.; Jovchelovitch, S. (Orgs.). *Textos em Representações Sociais*. Vozes.
- Mónica, M. (1978). *Educação e sociedade no Portugal de Salazar: a escola primária salazarista, 1926-1939*. Disponível em: <https://archive.org/details/educacao-e-sociedade-no-portugal-de-salazar-maria-filomena-monica/page/8/mode/2up>
- Moscovici, S. (2003). *Representações sociais: investigações em psicologia social*. Vozes.
- Mullender, R., & Haidt, J. (2025). *The Anxious Generation: How the Great Rewiring of Childhood Is Causing an Epidemic of Mental Illness*.
- Neris, V., Rosa, J., Maciel, C., Pereira, V., Galvão, V., & Arruda, I. (2024). GrandIHC-BR 2025-2035-GC4: Sociocultural Aspects in Human-Computer Interaction. In *Proceedings of the XXIII Brazilian Symposium on Human Factors in Computing Systems* (pp. 1-14).
- O'Neil, C. (2021). *Algoritmos de destruição em massa*. Editora Rua do Sabão.
- Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin Books.
- Pellizzari, B., & Barreto Júnior, I. (2019). Bolhas Sociais e seus efeitos na Sociedade da Informação: ditadura do algoritmo e entropia na Internet. *Revista de Direito, Governança e Novas Tecnologias*, 5(2), 57-73.
- Pereira, V. (2004). Consciência e memória como objetos da comunicação: O approach de Marshall McLuhan. *Revista Famecos*, 11(24), 149-157. <https://doi.org/10.15448/1980-3729.2004.24.3274>
- Pereira, S. (2014). A Internet na vida das gerações mais novas: Um estudo com adolescentes portugueses. Eleá, I. (Ed.), *Agentes e vozes: um panorama da mídia-educação no Brasil, Portugal e Espanha*, 135-143. Gothenburg: Nordicom. <https://hdl.handle.net/1822/41470>
- Ponce, A. (1989). *Educação e Luta de Classes*. Cortez, 1989.
- Reale, G. (2015). *História da filosofia Antiga*. Loyola.
- Regis, F., Ortiz, A., Affonso, L., & Timponi, R. (2012). *Tecnologias de comunicação e cognição*. Sulina.
- Santaella, L. & Lemos, R. (2010). *Redes sociais digitais: a cognição conectiva do Twitter*. Paulus.
- Santos, A., Donard, V., & Torres, M. (2025) *Ciberpsicologia e Humanidades Digitais*. Pimenta Cultural. doi 10.31560/pimentacultural/978-85-7221-286-1
- Schultz, T. (2004). *O capital humano: investimentos em educação e pesquisa*. Zahar, 1973.
- Sepulveda, J., Alves, L., & Gonçalves, L. (2025). A Hiperconectividade da Geração Alfa e o impacto na Saúde Mental: O olhar da Psicologia. *Repositório*

Institucional, 3(2).

- Silva, C. (2004). A escolarização entre a Idade Média e a Idade Moderna. *Centro de Ciências Humanas e Educação da UNI-RIO*.
- Shakespeare, W. (2013). *Trabalhos de amor perdidos*. Centaur.
- Van Dijck, J.; Poell, T.; & Wall, M. (2018) *The Platform Society: public values in a connective world*. Oxford Press.
- Vergara, S.; Ferreira, V. (2005). Representação social de ONGs segundo formadores de opinião do município do Rio de Janeiro. *Revista de Administração Pública*, (39) 5, 1137-1159.
- Vygotski, L. (1989). A formação social da mente. *Psicologia*, 153.
- Wohn, D., & Bowe, B. (2014). Crystallization: How social media facilitates social construction of reality. *Proceedings of the companion publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 261–264. [https:// doi.org/10.1145/2556420.2556509](https://doi.org/10.1145/2556420.2556509)



utad UNIVERSIDADE
DE TRÁS-OS-MONTES
E ALTO DOURO



İLİMER İnsan Odaklı İletişim
Uygulama ve Araştırma Merkezi



CICANT