AI and Language: New Forms for Old Discriminations? A Case Study in Google Translate and Canva

IA y lenguaje: ¿Nuevas formas para antiguas discriminaciones? Un estudio de caso en Google Translate y Canva

MARTINA MATTIAZZI

Author / Autora:
Martina Mattiazzi
CNR-ISMed Milan, Italia
martina.mattiazzi@cnr.it
https://orcid.org/0009-0005-7294-4814

Submitted / Recibido: 14/02/2024
Accepted / Aceptado: 29/08/2024

To cite this article / Para citar este artículo:
Mattiazzi, M. (2025). AI and Language: New Forms for Old Discriminations? A Case Study in Google Translate and Canva. *Feminismo/s*, *45*, 118-138. https://doi.org/10.14198/fem.2025.45.05

Abstract

The development of artificial intelligence (AI) is one of the greatest technological revolutions in recent human history. AI technology is widely used in various fields, including education. In this field, AI is studied as a discipline, and used as a tool to overcome social barriers. Like any human revolution, however, it is necessary to be careful about it and consider that the growing use of these new informatic systems also entails risks. One of them, it is the reinforcement of gender stereotypes and discrimination against women through linguistics feedback. Trough an experimental analysis conducted on common AI-integrated app –Google Translate and Canva–we will investigate linguistic behaviours such as responding to a command prompts. From the results obtained, we can demonstrate the existence of gender biases in the AI's productions, both in textual and visual language. Gender biases are consequences of the structural inequalities present in society: it is not the technology that is sexist, but it is the dataset on which it is based, which in turn is based on the results produced by users and published

on internet. In a society based on democracy and equality, it is important to ensure that the use of such a widespread technology as AI does not perpetuate existing stereotypes and does not allow to become a new form of strengthening discriminations. From a linguistics perspective, this means paying attention to the linguistic outputs, both textual and visual, provided by the AI and checking the dataset it has been training on. Due to their central role in the education of new generations, schools and institutions should prepare students for a critical vision of the phenomenon and provide them with the tools to contrast it. This path could start from teaching AI mechanisms and ethics of technology to students and using an inclusive language in the educational context.

Keywords: artificial intelligence; bias; Canva; Google translate; inclusive language; linguistics; linguistic sexism; stereotypes.

Resumen

El desarrollo de la inteligencia artificial (IA) es una de las más grandes revoluciones tecnológicas de la historia reciente de la humanidad. La tecnología de inteligencia artificial se utiliza ampliamente en diversos campos, incluida la educación. En este sector, la IA se estudia como disciplina y se utiliza como herramienta para superar ciertas barreras sociales. Sin embargo, como todas las revoluciones humanas, es necesario considerar que el uso creciente de nuevos sistemas informáticos también entraña riesgos. Uno de ellos es el refuerzo de los estereotipos de género y de la discriminación contra las mujeres en las producciones lingüísticas. A través de un análisis experimental realizado en herramientas comerciales comunes integradas con la IA –Google Translate y Canva– se investigará su comportamiento lingüístico como respuestas a indicaciones de comando. De los resultados obtenidos se puede demostrar la existencia de prejuicios de género en sus *outputs*, tanto en el lenguaje textual como en el visual. Los sesgos de género son consecuencias de las desigualdades estructurales presentes en la sociedad: no es la tecnología lo que es sexista, sino los conjuntos de datos en los que se basa, que a su vez se basan en los resultados producidos por los usuarios y usuarias y publicados en Internet. Para una sociedad basada en la democracia y en la igualdad, es importante garantizar que el uso de una tecnología tan extendida como la IA no perpetúe los estereotipos existentes y que no se convierta en un nuevo instrumento para fortalecer las discriminaciones. Desde una perspectiva lingüística, esto significa prestar atención a los resultados lingüísticos devueltos por la IA, tanto textual como visual, y verificar el conjunto de datos utilizado por la IA. Por su papel central en la educación de las nuevas generaciones, las escuelas y las instituciones deben preparar a las y los estudiantes para una visión crítica del fenómeno y brindarles las herramientas para contrastarlo. Este camino podría comenzar enseñando los mecanismos de la IA y su ética en clase, y utilizando un lenguaje inclusivo en el contexto educativo.

Palabras clave: inteligencia artificial; sesgos; Canva; Google translate; lenguaje inclusivo; lingüística; sexismo lingüístico; estereotipos.

1. INTRODUCTION

The expression artificial intelligence (AI) can indicate different elements: the definition itself has changed over time and can vary from author to author. Artificial intelligence can indicate both a field of study in computer science and the technology of the intelligent machines to interpret external data and learning from them. Or, as experts joke: «AI is everything that computers cannot currently do» (Bartneck et al., 2021, p. 7). Regardless of the differences between the definitions provided[1], we can say that the common element is that artificial intelligence involves the analysis, study, design and creation of intelligence agent or a machine that can achieve goals.

In common sense, AI is the ability of a machine to simulate human cognitive processes, such as interaction with the surrounding environment, learning, reasoning, and planning (Tecuci, 2011). The software is capable to process data received from the environment and return a response to the user, making predictions about future states and real-time decisions. Furthermore, some of the AI's algorithms can autonomously adjust their behaviour and improve their skills over time, by analysing the effects of previous actions through a process of trial and errors.

AI is largely used in various industrial fields, including medical care, like disease prediction and digital healthcare solutions; manufacturing, for example to optimize business processes and predict demand; finance, for data analysis and new forms of value, like crypto and NFTs; and marketing, like customer tailored recommendations and personalization of experience (Team Google Italy, 2023). Its applications concern also everyday life, like smart home systems, apps for travelling, smart website and chatbots (Maheshwari, 2023).

1. For further information on the different definitions of AI, its history and evolution and different AI system, please refer to the work of Bartneck et al. (2021) or Russell & Norvig (2021).

Due to its growing development and pervasiveness in society, in the 2018 the European Union started its footpath to define and promote the European AI Strategy. The aim is to make the EU a world-class hub for AI to help societies and businesses to enjoy the benefits of new technology, while ensuring safety and protection (European Commission, 2024). In this perspective, in September 2022 the European Commission published *Ethical Guidelines on the Use of Artificial Intelligence (AI) and data in teaching and learning for teachers*. The document is aimed at students and teachers, independent of their knowledge in digital education.

In fact, AI is becoming an integral part of the educational process, both as a subject of study and as a tool whose potential can be exploited to improve the educational context. In May 2023, the U.S. Department of Education – Office of Educational Technology (OET) released a document aimed to supporting innovations and the use of technology in teaching and learning, developing a guideline for edtech. In this document it is shown how it is necessary nowadays to train students so that they know how to manage the technological innovations that are presented in the future. It is also essential that teachers know their potential and pass them on the new generations (U.S. Department of Education, Office of Educational Technology, 2023).

Using AI in educational process means improving the students' path and offering new opportunities to teachers, like plan courses accordingly, or the use of adaptive learning technologies in providing individualised interventions. The use of AI in education could also help break down some social barriers, allowing students to improve their academic performance regardless of their difficulties, based on their needs. For example, the tools of speech recognition of AI can be used to support students with disabilities, or multilingual background, because of the significant quality of being adaptive of AI's digital tools for learning.

AI can also improve teaching job. First, AI can be used to reduce the administrative work, as reported by Bryant et al. (2020), to increase focus on students. In addition, an AI assistant can guarantee support to students with homework, whenever and wherever they are. Finally, according to the principles of inclusion and no discrimination which should be imperative in educational contexts, AI can help teachers with language translation, in all

those situations where they must deal with foreign parents of their students and work with them.

Like any new technological phenomenon, AI can bring with it both advantages and new critical issues that must be recognized and addressed (UNESCO & International Research Centre on Artificial Intelligence, 2024). Given its pervasiveness in daily life and its importance in improving the educational path, it is fundamental to understand which these critical issues are. One the most frequently encountered complexities in the use of AI concern the utilization of data. Data are the basis of the functioning of AI, particularly in the so-called machine learning (ML). This technology includes a group of AI that uses algorithms and data to learn from the environment and imitate the human's method of learning, gradually improving its performance (Russell & Norvig, 2021). Generative AI is a group of machine learning (ML) that can produce new content starting from the data received. While traditional AI is typically oriented towards processing existing information to provide specific answers or perform precise tasks, the Generative AI if capable of generating text, images, music and so on, creating original content that didn't exist before.

Because of these reasons, it is important to watch out about bias and complexities of these new tools. Bias, in the context of generative language models, can be defined as «the presence of systematic misrepresentations, attribution errors, of factual distortions that result in favouring certain groups or ideas, perpetuating stereotypes, or making incorrect assumptions based on learned patterns» (Ferrara, 2023, p.2). As reported in the AI Google blog about «Responsible AI practices», machine learning models use existing data accumulated in the real world and created by humans. A model may learn pre-existing biases in the data and even amplify them: «The risk is that any unfair bias in such systems can also have wide-scale impact» (Ferrara, 2023, p.3).

The goal of this article is to try to reveal some gender biases that are perpetrated by some apps integrated with AI, in particular Google Translate and Canva. The underlying idea is not to demonize this new technology, but rather to document that the sexism that lies behind AI does not depend on the technology itself; it depends on the environment from which data are retrieved (European Institute for Gender Equality, 2022; Manasi et al.,

2023). Recognizing this problem is the first step in gaining awareness about AI reinforcing stereotypes. Subsequently, some suggestions are proposed to be applied to the educational context and to equip students with the critical tools necessary to recognize discriminations in AI and act accordingly. Finally, on a theoretical level, the objective is to improve gender studies applied to artificial intelligence, as regards textual and visual language.

2. METHODOLOGY

To demonstrate the existence and perpetration of gender bias in app with generative AI, we conducted an experimental linguistic analysis on common commercial tools integrated with AI. The experimental process aimed to investigate the behaviour of two types of tools: Google Translate for translations, and Canva for image generation. These programmes were chosen due to their widespread diffusion worldwide, representing points of reference for their fields of application, and for the large use by young people also in academic contexts.

Google Translate is one of the most famous online translators. This tool is used to convert sentences, texts or even complete web pages and images from a language to another. In 2018 Google said that they had introduced machine learning technology to this tool, which allow translation more accurate, considering the position of the words in the sentences and the context in which they are placed. In the company's Italy blog is written that it is precisely AI that has made Google Translate a tool so appreciated by people all over the world (Google Italy, 2023).

The second tool is Canva, an online design and publishing software that recently has included in its toolkit the possibility of creating images with its AI generator. The user must write the description of the desired image and the app will return four options to choose from and the ability to generate new ones.

The analysed results were obtained by providing the models with texts or individual words via command prompts. In Google Translate the first phase consisted of providing as input a single word to see if the output result was free from bias (below denominated example 1a and 1b); in the second phase, we provide a sentence containing the word used in the previous test

(example 2a and 2b), to test whether the model was capable of distinguishing prejudices with a context.

In Canva, the inputs consisted of short expressions without context or information that could influence the model (example 3 and 4). In both tools, the type of words provided as input to the AI models was carefully chosen to create a small dataset that is primarily focused on gender biases; the distinctive criterion of the dataset concerns the use of words of phrases inherent, according to stereotypes, to social roles, tasks and other social conventions linked to gender prejudices.

The selection of different professions and activities as inputs was made to demonstrate that gender bias affects different roles in society. This choice is motivated by the test we wanted to conduct, as the objective was to verify if AI models can recognize a gender stereotype present in society and implement a change, or if they are trained unequivocally on datasets undermined by gender bias.

Furthermore, the choice of a methodology as simple as the one illustrated allows the experiment to be replicated in the future using the same linguistic inputs and check whether, over time, the apps have update and they are able to recognise gender bias.

3. RESULTS

If we use Google Translate to translate from English to Italian the word «engineer», the output will be in both genders, feminine «ingegnera» and masculine «ingegnere», with no discrimination (example 1a); the same will occur with the word «nurse», given back in both genders, feminine «infermiera» and masculine «infermiere» (example 1b). But if we try to translate a simple phrase with the same roles, the result will be different (example 2a): in «I spoke with the engineer», the output is only in the masculine form («Ho parlato con l'ingegnere»); and the same, on the contrary, will happen with the sentence «I spoke with the nurse» (example 2b), which give the result only in the feminine form («Ho parlato con l'infermiera»).

This representation bias occurs in adjectives too: as shown in Prates et al. (2020), when we use qualifiers such as «shy» or «kind» –adjectives stereotypically associated with woman's behaviour– in sentences without an explicit

gender, Google's translations are predominantly in the feminine forms. Quite the opposite, qualifier like «arrogant» and «proud» are translated with male pronouns. This happens because the same data from which generative AI takes its information are subject to gender discrimination. The data comes from newspaper online, example sentences from digitized dictionaries, text corpora, books available online, social media platforms, etc., and all of them are contexts created by human –and not by the AI itself– and only subsequently used by the AI to return an output. Gender bias in AI could be an effect of gender bias and lack of women's presence in text (Leavy et al., 2020).

Figure 1. A family doing chores
Source: https://www.canva.com/

The same gender bias is present also in image generative tool Canva. As demonstrated in the example 3 (Figure 1), if we write in the box «a family doing chores», the app will return stereotype images: in three of these, there are children with their mother; in only one appears the father[2].

The stereotype is even worse when we want to create an image in which there is a «CEO at work» (example 4): in this case, in all four images no women appear (Figure 2).

2. We can assume, in good faith, that the app with the term «family» also includes single-parent families and that therefore the choice to return images in which only the mother is present is a pure coincidence not subject to gender bias. We invited the reader to create an image with the Canva's AI generator writing as a description «a family doing homework» or «a family playing videogames» and see the differences.

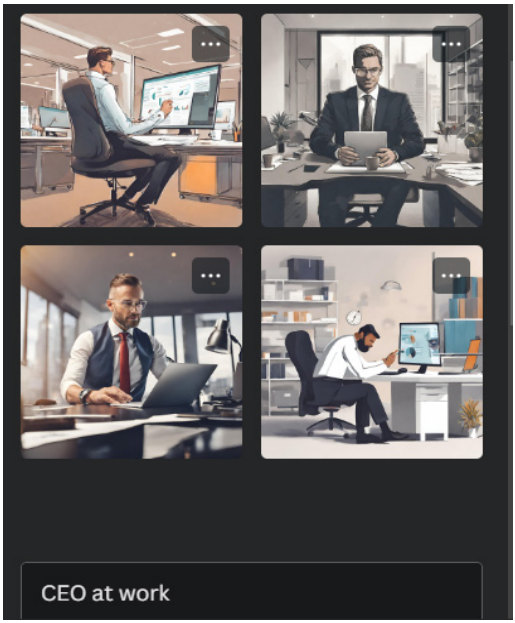

Figure 2. CEO at work
Source: https://www.canva.com/

These are clear examples of how generative AI can perpetuate gender stereotypes. An app heavily gender-biased can amplify social inequalities (Gross, 2023). According to the definition in the World Health Organization's website, «gender refers to the characteristics of women, men, girls and boys that are socially constructed. […] As a social construct, gender varies from society to society and can change over time […]. Gender is hierarchical and produces inequalities». Achieving gender equality is one of the Sustainable Development Goals adopted in the 2030 Agenda for Sustainable Development by United Nations. Gender equality is a fundamental human right; gender inequality affects not only women's lives, but also the economy, politics, and of course education in the society (UN Women, 2018). Gender inequality is supported by various social phenomena, and stereotypes constitute the foundations of the gender's inequalities pyramid.

Gender stereotypes on intellectual abilities and skills influence children precociously. According to Save the Children's atlas *Con gli occhi delle bambine* written by De Marchi (2020), gender stereotypes already begin to be absorbed in primary school, and consequently girls start doubting about their abilities and avoiding those activities perceived or presented as particularly complex. Gender stereotypes are reinforced by games, images in commercial, storytelling in media and social media, children's book, and textbook. Even in school, the place where students should be trained

impartially by teacher and with equal opportunities, sexist behaviours can occur, often almost invisible because rooted in the mindset and therefore difficult to unveil. Gendering has a strong influence on children, limiting their potential, their way of perceiving themselves and others, their behaviour –both for male and female[3]–. It can also discriminate those who do not fit into traditionally established roles.

Sexism can also be represented by the language used; in this case, we talk about linguistic sexism. This notion developed in the 60's-70's in USA and referred to the discrimination in the way women were (and still are) represented compared to men, using language (Robustelli, 2012). Language represents a strong tool capable of reinforcing or demolishing gender stereotypes. Linguistic sexism is the linguistic manifestation of cultural prejudices tainted by sexism (Gheno, 2021). Language is recognized as having a fundamental role in the social construction of reality, and, therefore, also of female and male gender identity. It is not only an instrument to communicate our thoughts and our feelings, but, due to its performativity, language can shape our way of seeing the world and therefore also our way of interacting with it[4].

Cultural concepts change over time, so they require new ways to express them. Language can satisfy these needs, providing the words to define them, redefining existing ones, and communicating them. Giving names to new cultural concepts and new social roles is the first step towards recognizing these realities as the cultural heritage of the linguistic community (Giusti, 2011). Language, as artificial intelligence, cannot be defined as sexist in itself: the sexism does not lie in linguistic structures and mechanisms, but in our choices as speakers. For this reason, it is extremely important to include language in the instruments available to defuse gender prejudices and use

3. Elena Gianini-Belotti in her *Dalla parte delle bambine* (English version: *Little girls: social conditioning and its effects on the stereotyped role of women during infancy*) used the term «mutilation» to refer to the impossibility given to boys from having behaviours socially associated with girls. Already in 1973, date of publication of the essay, she recognised how gender stereotypes are harmful not only to girls, but also to boys. This is why everybody should be interested in feminism and equalities, independently from their gender or sex.
4. For further information on language and its performativity, please refer to Austin (1975) and Searle (1969) speech act's theory; for the link between language and mind, we suggest the TED Talk of Lera Boroditsky (TED, 2017).

it to support the achievement of gender equality. To help diminish linguistic phenomena based on gender prejudices, it is fundamental to acquire metalinguistic awareness of these critical issues. Sexist language stems from culturally and historically perpetrated stereotypes. It is not the langue or its form that is sexist, but it is the use we make of it (Gheno, 2021).

Gender-inclusive language is a powerful tool to promote gender equality and eradicate stereotypes (Council of Europe, 2024). This results in the application of an inclusive language, i.e. a language that recognizes diversity in people, respects the differences and promotes equal opportunities. An inclusive language respectful of gender recognizes then the path travelled by women and their importance in the society, in the working position, and in the institutional role. Bringing out the presence of women in the language is the first step toward recognizing the legitimate presence of women in all social and cultural spheres, and the creation of equal models for the younger generations.

In relation to the results obtained in our experiment, it should be noted that these sexist stereotypes discussed are reproduced again. Thus, defining a CEO with stereotypically masculine characteristics, talking about kind female nurse and proud male engineer, and representing women in the housework context means relegating women only to some environment, precluding them from being part of other groups in society. The data obtained therefore confirm that these gender stereotypes are supported by new technologies. Society in general and the education of children in particular can be influenced by this distorted perspective and the problem will be repeated even in the new generations.

4. DISCUSSION

In 2021 the humanitarian organization Save the Children started the «DIG4FUTURE – Digital Competencies, Inclusion and Growth for Future Generations» project (https://www.dig4future.eu/Wordpress/), whose aim is to improve the digital skills and competencies in young people, implementing an innovative method focus on inclusive education and promotion of common values. The project aspires to strengthen the students' competencies in the DigComp framework, that includes information and data literacy;

communication and collaboration; digital content creation; safety; problem solving. In particular, the project is innovative in the way it includes training on AI. The skills are developed through activities that allow teachers and students to discover key concepts of AI and above all to develop a critical thinking on technology based on artificial intelligence.

By learning how AI works, students were allowed to use a Teachable Machine, an educational software capable of training an artificial intelligence by entering some data. In this way, the students not only learned its mechanism, but were also able to discuss the results generated by the AI: «Through a series of lessons and activities, teachers and educators are supported in training their students' digital competences while exploring the ethical implications that such technical concepts entail, such as algorithmic bias, targeted advertising and fake news generation» (Cattozzella et al., 2023, p. 4).

Specifically, the third module 'Ethics and AI. Societal impact and ethical considerations for AI' is entirely dedicated to the practical applications of AI and their implications. This allowed students to familiarize themselves with the main concerns regarding the use of AI and its bias. This step is a fundamental to raise a critical awareness: the project includes various activities in which to focus on the prejudice that AI risks spreading if the previous phase –that of training the Teachable Machine– is not implemented with data respecting ethics and no-discrimination. If the output is one like the examples seen previously, students will be able to analyse with critical awareness the discrimination in the system. They could understand the cause of the problem –i.e. the data that were used to train the Teachable Machine– and they could reflect on possible solutions.

This innovative method can reduce the ethical risks arising from the incorrect use of AI and encouraging the development of a critical discussion about these technologies. The aim to promote perception and acknowledgement of the possibility of biases in AI could be achieved through a new method of education. This means prevent the reproduction and reiteration of discrimination based on gender, ethnicity, sexuality, religion, etc. Trough the understanding of the AI mechanism, «students can become advocates for fairer and more ethical artificial intelligence system» (Cattozzella et al., 2023, p. 25).

If we compare this method of learning –i.e. study the mechanism of AI and its risks to avoid mistakes in the utilization– to the acquisition of a second language, we will be able to see that there are points of contact between the two learning situations. For many years, the science of foreign language teaching has considered the mistake as a deviation from the standard norm, and for this reason considered to be avoided and penalized. In recent years, however, errors have begun to be considered as a useful resource for learning a language other than the mother tongue, through a different point of view (Mezzadri, 2002).

In this sense, one useful technique is the self-correction. It consists in the treatment of the mistake by the same person who produced the output, who must therefore be able to recognize the error –unaided or helped by the teacher–, understand the reason and proceed with its correction. This means putting students in a position in which they could develop a critical autonomy that allows them not only to correct the mistake, but also –and above all– to understand the cause behind the error. From this perspective, learning a L2 does not mean accumulation grammatical knowledge about the foreign language –in a passive vision of learning– but being able to reorganize this knowledge to advance in learning. Students, thus, actively learns the mechanism of the foreign language and could perceive themselves as capable of improving their metalinguistic skills.

The same critical autonomy can be essential to recognize the bias in the AI's output (Trang, 2021). At school, students should acquire such knowledge of new technologies that they should be able to recognize by themselves the errors made by the software, understand the cause, and resolve the bias. This means activating a critical awareness of stereotypes, useful both in the acquisition of new technologies and in everyday social-life experiences. In the case of outputs that produce gender discrimination, students will therefore be able to understand that the inequality is not inherent in the machine, but it depends on the quality of the data it processes.

The interlanguage stage that students reach when learning a second language could be comparable to the language used by the AI to respond: in the first case, self-correction consists in modifying the same outputs produced by the student –i.e. a linguistic error–; in the second case, the student will

manually modify the AI's language, recognizing its discriminations and working on them in a self-critical manner.

From this perspective, the implementation of peer review could give interesting results too. As many authors have seen (Rossi, 2018; Saturno, 2018), interaction between peers aimed at solving linguistic problems in learning a L2 could represent a moment of serious consideration, due to the induction of participants to focus explicitly on the problematic structure, i.e. technically defined as noticing (Saturno, 2018). In learning AI mechanisms and in recognizing the (possible) presence of sexism in AI's output, the comparison between peer could give the opportunity to the students to understand deeply some critical issues and develop a broader awareness. Peer feedback represents a more shared language than that used by teachers, managing to better adapt to educational needs of students and to their way to communicate the world. Besides, peer reviews start from different points of view characterized by continuous exchange and sharing of ideas, that constitute a continuous process of readjustment and improvement (Segatello, 2021). In addressing the issue of social discrimination in AI, this could lead to an enrichment of perspectives: where one student sees no bias, another one might notice it and make it clear to the partner. Where one student doesn't recognize the discrimination, another one could have experience it and revealed it.

In this perspective, with Generative AI the students can identify some of the unconscious stereotypes developed and maintained throughout society, bringing about positive revision of their thoughts. Nota et al. (2013) consider worthwhile to create opportunities for social reflexivity and to stimulate children and adolescents to reflect on their experiences from different perspectives. The aim is to recognize diversity and be supportive towards other peers with different cultural backgrounds. In the future, it could be interesting to study this educational method to understand its actual benefits on younger generations and on the possibilities of increasing their awareness of AI and bias.

In the meantime, in the educative environment should appear two elements: the teaching of AI –such as the project DIG4FUTURE– and the use of an inclusive language. There are some practices that can be adopted to support a gender-sensitive approach in education and in the use of new

technologies, contributing to build a more inclusive environment in schools and academia. Concerning to language, we recommend to not perpetrate gender prejudices in professions and occupations: in the examples illustrated to students, it can be helpful to show both woman and man doing the same activity and reverse the roles stereotypically associated with genders, like the examples shown before.

In this perspective, we suggest also to not add irrelevant information about gender when describing a person and to prefer gender-neutral terms instead of gender-specific terms, like «firefighter» instead of «fireman». For language with grammatical gender, it is recommended to use the feminine correspondents of masculine terms in every context, also for job titles, and to replacing the overextended masculine –called, inappropriately, «generic» or «no-marked»– with double forms, both masculine and feminine, or neutral forms no gender-marked[5]. Concerning the use of pronouns, when gender is not relevant, the gender-neutral plural «they» should be chosen among the singular gendered ones; in particular, this consideration should be used also to avoid the binary conception of gender[6] (Cavallo et al., 2021).

As we have seen, the omission and invisibility of women can occur not only in written language, but also in visual language and graphics (Bazzan, 2022). It is important to consider other elements, such as colours, symbols, quantity, and quality of elements presented in a figure. Colours are arbitrarily related to genders, and often pink is connected to woman and blue to men. This reinforces gender stereotypes, influencing different life spheres, such

5. For example, in Spanish it is preferred to say «Los profesores y las profesoras» or «personal docente» instead of using only the masculine form. In English, most nouns do not have grammatical gender; for those who have different forms, it is recommended the same, like for «steward» and «hostess» that can be replaced with the generic expression «cabin attendant».
6. In English, the pronoun «they» is used by no-binary people, i.e. people that doesn't recognise themselves as a woman or a man; the LGBTQIA+ community has chosen this pronoun due to its genderless and to its use as a singular (better known as «singular they»). New solutions are being tested in other languages too: in Arabic, where singular, dual and plural exist, it has been proposed to use the dual to refer to a no-binary person; in Swedish, already in 2015 the pronoun *hen* was recognized as a neutral pronoun and subsequently in 2022 the same use was proposed by the Språkrådet, the Language Council of Norway. In languages where it is not mandatory to elicit the subject, such as Italian, it is preferable to omit the pronoun.

in the choice of the smock for children, and can strengthen the process of gendering. It is recommended to use various shades of colours and to use pastel colours not only to refer to women, but also to men.

Symbols tend to bolster discriminations too. Females represented with long hair and skirt reiterate a stereotyped aesthetic canon. Furthermore, the idea that the unmarked icon is used both with a generic meaning –such as in road sign– and associated with the male gender when paired –such as the symbols of toilet– is another case of obscuration of feminine presence and of use of overextended masculine. We recommend using alternative icons, like gender symbols most widely recognized, the circle above a cross (♀) and the circle below an arrow pointing diagonally (♂) to indicate women and man. In addition, it could be more inclusive to use no-binary icon too: the web offers a lot of alternative ideas about it.

As far as quantity and quality in images, to avoid gender discrimination, men and women should be represented in the same number, also paying attention to spatial and dimensional relations: in a group of persons, women should be portrayed in the same size of men and in foreground with them, and no smaller in the background[7]. Clothes, postures, and expressions should convey an equal status too. Lastly, women shouldn't have a purely decorative purpose in images, but they should be doing something or depict some role, preferring unconventional actions.

5. CONCLUSIONS

Providing students with the tools to critically reveal the stereotypes they encounter in school and everyday life is a fundamental step in achieving gender equality. Ethical AI is a necessity: although younger generations are defined as 'digital natives', this does not mean they know how to move autonomously and consciously in the online world. This can lead, among other risks, to not seeing gender discrimination.

A generation attentive to these critical issues will instead allow us to create new, inclusive, and democratic data. In the future, this data will in turn be

7. A classic example of this dimensional asymmetry is present in the European children crossing road sign in which the boy and his schoolbag are bigger than the girl, recognisable from the skirt and the hair's bow.

used by AI to return less sexist outputs. This would establish a virtuous circle that could replace the current vicious circle. As it stands today, sexist outputs arise from sexist inputs. When data contain biases, the machine learning will learn these biases and reported them in the outputs shown to the user.

If young people can recognize discrimination and defuse it –for example by giving equal alternatives or deconstructing sexist concepts with an inclusive textual and visual language–, they will create new data that will be used by AI to return subsequent results. The new outputs will therefore be less subject to discrimination thanks to the new environment created.

Obviously, this process will require a lot of time and effort on the part of the social actors involved –i.e. students, teachers and also society– but some changes are already underway. Although generative AI is a very recent technology, we can already see implemented actions that move in this direction. For example, to return to the examples reported previously, Canva's software already allows you to report AI generated images with discrimination of any kind; for some languages, Google Translate returns double gender not only in individual lexemes, but also for sentences[8].

AI is a huge resource for the future of humanity, if used correctly and democratically. It could make it possible to break down social barriers and improve people's lives. However, to achieve these noble goals it is necessary to ensure that its use does not reinforce discrimination already in place. Artificial intelligence, like language, is not a sexist tool, and it has all the potential to be used in a no-discriminatory manner. As with language, discrimination in AI arises from the way the tool is used by humans.

Meanwhile ICT scientists explore ways to reduce bias in artificial algorithms[9], the considerations for an inclusive language should be implemented

8. At January 2024, this happens when translate from English to Spanish («I spoke to the nurse» is translated in «Hablé con la enfermera» and in «Hablé con el enfermero»), or from Turkish –a language with a neuter personal pronoun for third singular person– to English («O bir doctor» is translated in «She is a doctor» and in «He is a doctor»).
9. Among the various propose, Ghosh and Caliskan (2023) suggest switching to a human centred approach to assisted language translation in AI software, for example, including native language speakers to forming accurately labelled text corpora, incorporating considerations of ethic in the design process of AI tools. A similar idea is proposing in Ferrara (2023) who suggest a human-in-the-loop approaches that involve human input, feedback, and real-time moderation in the development of langue model.

by both teachers and students to change the perspective on gender roles and modify their language toward a more inclusive one. Using these precautions in the educational context could allow students to get used to an inclusive language, to consciously recognize the discriminations present in the outputs of AI and produce themselves content that don't contain gender discriminations.

The language used by AI is the mirror of society's language: to eradicate gender bias in AI, it is necessary to eradicate them first from society itself. Old discriminations, such as gender stereotypes, can be overcome by the correct use of artificial intelligence which could indeed prove to be a very powerful ally in eradication of gender prejudices in the new generations, who represent the future of our society.

6. REFERENCES

Austin, J. L. (1975). *How to do things with words* (2nd ed.). Harvard University Press (Original work published 1962).

Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2021). *An Introduction to Ethics in Robotics and AI*. Springer. https://doi.org/10.1007/978-3-030-51110-4

Bazzan, B. (2022). *Il peccato originale delle intelligenze artificiali* [Unpublished master's thesis]. Politecnico di Milano – Scuola del Design. POLITesi – Archivio digitale delle tesi di laurea e dottorato. https://www.politesi.polimi.it/handle/10589/187945

Bryant, J., Heitz, C., Sanghvi, S., & Wagle D. (2020, January 14). How artificial intelligence will impact K-12 teachers. *McKinsey Blog*. https://www.mckinsey.com/industries/education/our-insights/how-artificial-intelligence-will-impact-k-12-teachers#/

Cattozzella, D., Hirsch, G., Fabretti, V., Schiavo, G., & Torrisi, E. (2023). *Artificial intelligence at school: the pioneering experience of Dig4future*. DIG4Future Project. https://drive.google.com/file/d/1ngvv3-fR2i14ADUIrWTlf6nvMgzOob46/view?usp=drive_link

Cavallo, A., Lugli, L., & Prearo, M. (2021). *Questioni di un certo genere. Le identità sessuali, i diritti, le parole da usare: una guida per saperne di più e parlarne meglio*. Iperborea.

Council of Europe. (2024). *Guidelines for the use of language as a driver of inclusivity*. Documents and Publications Production Department (SPDP), Council of Europe. https://rm.coe.int/0900001680aec235

De Marchi, V. (2020, November). *Con gli occhi delle bambine. Atlante dell'infanzia a rischio 2020*. Save the Children. https://s3.savethechildren.it/public/files/uploads/pubblicazioni/con-gli-occhi-delle-bambine_0.pdf

Directorate-General for Education, Youth, Sport and Culture, European Commission. (2022). *Ethical guidelines on the use of artificial intelligence (AI) and data in teaching and learning for educators*. Publications Office of the European Union. https://doi.org/10.2766/153756

European Commission (2024). *A European approach to artificial intelligence*. Publications Office of the European Union. https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

European Institute for Gender Equality (2022). *Artificial intelligence and gender equality*. EIGE Europa. Publications Office of the European Union. https://eige.europa.eu/newsroom/director-corner/artificial-intelligence-and-gender-equality?language_content_entity=en

Ferrara, E. (2023). Should ChatGPT be Biased? Challenges and Risks of Bias in Large Language Models. *First Monday, 28*(11). https://doi.org/10.5210/fm.v28i11.13346

Gheno, V. (2021). *Verso l'inclusività linguistica e oltre*. Aula di Lettere. Percorsi nel mondo umanistico. Zanichelli editore. https://www.zanichelli.it/download/media/e46r/10inparita_Gheno.pdf

Ghosh, S., & Caliskan, A. (2023). ChatGPT Perpetuates Gender Bias in Machine Translation and Ignores Non-Gendered Pronouns: Findings across Bengali and Five other Low-Resource Languages. In F. Rossi, S. Das, J. Davis, K. Firth-Butterfield & A. John (Eds.), *AIES '23: Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 901-912). Association for Computing Machinery. https://doi.org/10.1145/3600211.3604672

Gianini-Belotti, E. (2013). *Dalla parte delle bambine. L'influenza dei condizionamenti sociali nella formazione del ruolo femminile nei primi anni di vita* (28th ed.). Feltrinelli (Original work published 1973).

Giusti, G. (2011). Riferimento al genere e costruzione d'identità. In G. Giusti (Ed.), *Nominare per esistere: nomi e cognomi. Atti del primo convegno Lingua e Identità di Genere. Venezia, Auditorium s. Margherita, 19 settembre 2011* (pp. 9-24). Libreria Editrice Cafoscarina.

Google Italy. (2023, July 6) *10 modi in cui l'intelligenza artificiale sta già cambiando in meglio le nostre vite*. https://blog.google/intl/it-it/notizie-aziendali/10-modi-in-cui-intelligenza-artificiale-sta-gia-cambiando-in-meglio-le-nostre-vite/

Gross, N. (2023). What ChatGPT Tells Us about Gender: A Cautionary Tale about Performativity and Gender Biases in AI. *Social Sciences, 12*(8). https://doi.org/10.3390/socsci12080435

Leavy, S., Meaney, G., Wade, K., & Greene, D. (2020). Mitigating Gender Bias in Machine Learning Data Sets. In L. Boratto, S. Faralli, M. Marras & G. Stilo (Eds.), *Bias and Social Aspects in Search and Recommendation. BIAS 2020. Communications in Computer and Information Science* (pp. 12-26). Springer, Cham. https://doi.org/10.1007/978-3-030-52485-2_2

Maheshwari, R. (2023, August 24). Advantages of Artificial Intelligence (AI) In 2024. *Forbes Magazine*. https://shorturl.at/WsWjj

Manasi, A., Panchanadeswaran, S. & Sours, E. (2023, March 13). Addressing Gender Bias to Achieve Ethical AI. *International Peace Institute – Global Observatory*. https://theglobalobservatory.org/2023/03/gender-bias-ethical-artificial-intelligence/

Mezzadri, M. (2002). La correzione degli errori. *In.it, 1*(3), 4-9.

Nota, L., Soresi, S., & Santilli, S. (2013). Le abilità sociali e il loro ruolo in un contesto socioculturale. In A. Vieno & M. Santinello (Eds.), *Metodi di intervento in psicologia di comunità. Dalla progettazione alle esperienze professionali* (pp. 77-95). Il Mulino.

Prates, M. O. R., Avelar, P. H., & Lamb, L. C. (2020). Assessing gender bias in machine translation: a case study with Google Translate. *Neural Comput & Applic, 32*, 6363-6381. https://doi.org/10.1007/s00521-019-04144-6

Robustelli, C. (2012, July 19). Il sessismo nella lingua italiana. *Treccani magazine*. https://www.treccani.it/magazine/lingua_italiana/speciali/femminile/Robustelli.html

Rossi, S. (2018). Il *feedback* tra pari come strumento per l'apprendimento linguistico. In R. Grassi (Ed.), *Il trattamento dell'errore nella classe di italiano L2: teorie e pratiche a confronto* (pp. 111-124). Franco Cesati Editore.

Russell, S., & Norvig P. (2021). *Artificial Intelligence: A Modern Approach*. Pearson.

Saturno, J. (2018). Correzioni tra pari nelle fasi iniziali dell'acquisizione: uno strumento appropriato? In R. Grassi (Ed.), *Il trattamento dell'errore nella*

classe di italiano L2: teorie e pratiche a confronto (pp. 97-110). Franco Cesati Editore.

Searle, J. R. (1969). *Speech Acts*. Cambridge University Press. https://doi.org/10.1017/CBO9781139173438

Segatello, F. (2021). *La valutazione tra pari per lo sviluppo delle abilità sociali. Uno studio quasi-sperimentale in una classe quinta di Scuola Primaria* [Master's thesis, Università degli Studi di Padova – Dipartimento di Filosofia, Sociologia, Pedagogia e Psicologia applicata]. Padua Thesis and Dissertation Archive. https://thesis.unipd.it/handle/20.500.12608/37326?mode=simple

Team Google Italy (2023, July 6). *10 modi in cui l'intelligenza artificiale sta già cambiando in meglio le nostre vite*. Blog di Google Italy. https://blog.google/intl/it-it/notizie-aziendali/10-modi-in-cui-intelligenza-artificiale-sta-gia-cambiando-in-meglio-le-nostre-vite/

Tecuci, G. (2011). Artificial intelligence. *WIREs Computational Statistics, 4*, 168-180. https://doi.org/10.1002/wics.200

TED. (2017). *How language shapes the way we think | Lera Boroditsky* [Vídeo]. YouTube. https://www.youtube.com/watch?v=RKK7wGAYP6k

Trang, C. K. (2021). *Exploring Bias Against Women in Artificial Intelligence: Practitioners' Views on Systems of Discrimination* [Master's Thesis, Department of Informatics – The Faculty of Mathematics and Natural Sciences, University of Oslo]. UiO DUO vitenarkiv. https://www.duo.uio.no/handle/10852/88551?show=full

UN Women (2018). *Why Gender Equality Matters across all SDGS*. Un Women. https://shorturl.at/XMUMy

UNESCO, International Research Centre on Artificial Intelligence (2024). *Challenging systematic prejudices: an investigation into bias against women and girls in large language models*. https://unesdoc.unesco.org/ark:/48223/pf0000388971

U.S. Department of Education, Office of Educational Technology (2023, May). *Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations*. https://www2.ed.gov/documents/ai-report/ai-report.pdf

World Health Organization (n.d.). *Gender*. World Health Organization Europe. https://www.who.int/europe/health-topics/gender#tab=tab_1