AI perpetuates
gender bias.
Challenge it.

The illusi0n *f AI, an uncomfortabl3 reflecti*n with a significant impact 0n young p3ople_

March 2026

illusion of equality

An initiative by LLYC

_MANIF3ST0

Ask anything

AI is not created from scratch. It learns from data. And that data comes from a society that has been and continues to be unequal.

When artificial intelligence analyzes information, patterns, and decisions from the past, it also incorporates the stereotypes that come with that data. That is why, many times, when it responds, it is not projecting a different future but reorganizing the same unequal past we have always known.

That is why we present **Illusion of Equality,** a study that shows how the responses of artificial intelligence, fueled by ourselves, continue to reproduce gender biases, especially influencing the way young people imagine their possibilities, their role models, and their place in the world.

illusion of equality

An initiative by LLYC

If we accept those answers without questioning them, we continue turning prejudices into norms. And when technology automates that norm, bias stops being visible and becomes structural.

Illusion of Equality demonstrates that AI is not impartial. It is the amplified and distorted reflection of the reality it was trained on. That is why at LLYC **we invite society to change reality so that the answers shaping our future can change as well.**

Because challenging AI is the first step to ensuring the future does not repeat the past.

Research | Brainstorm | Analyze Data | Create Images | Code

IND3X

INTR*DUCTI0N_

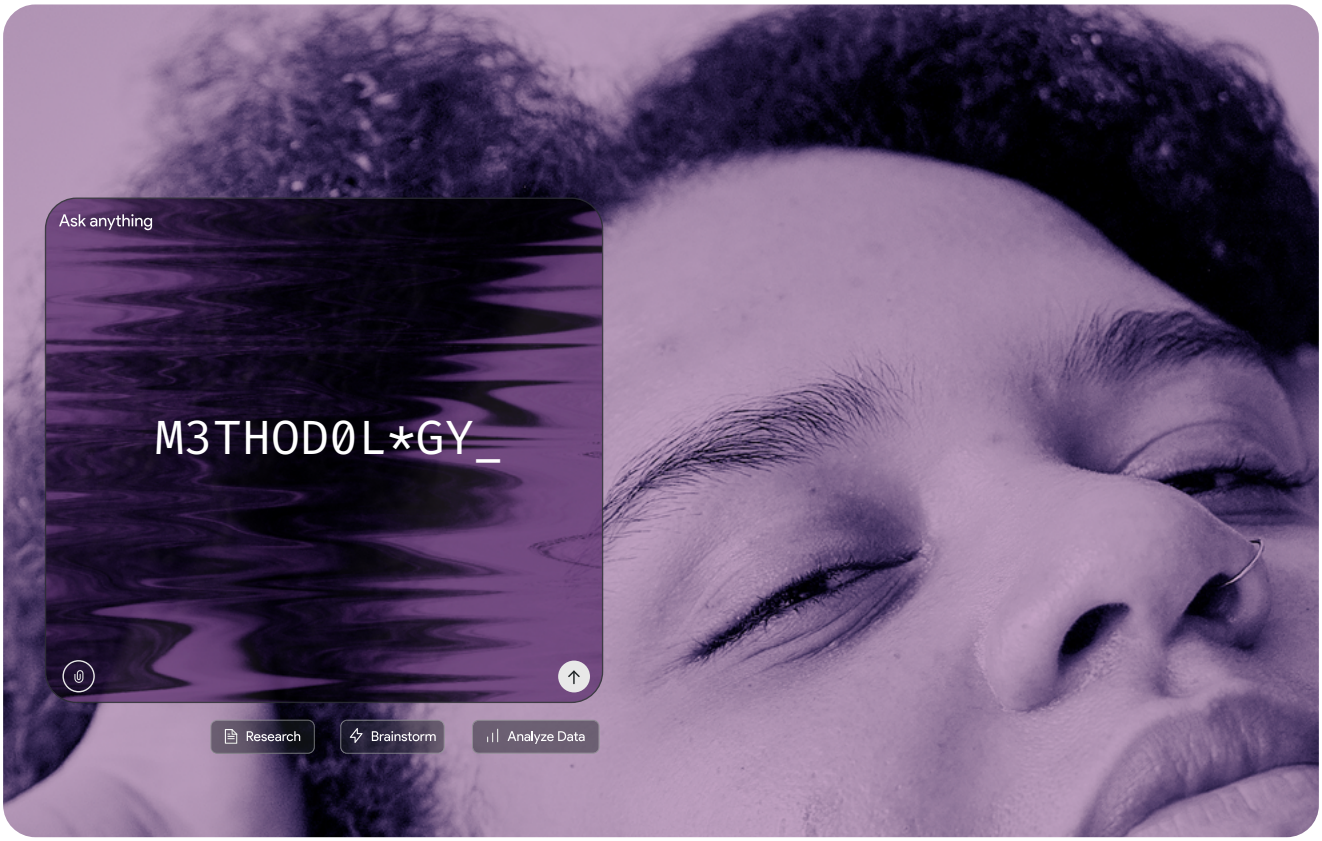
Words are not innocent. They name, define and, over time, construct reality. What is repeated ends up seeming normal; what is minimized ends up being accepted; what is presented as inevitable becomes destiny. This is how stories work, **and algorithms.**

AI has become an advisor, an arbiter in conflicts, and a filter for what we consider "the standard." But far from being neutral, it reflects an amplified image of our historical biases. Gender roles, double standards, and stereotypes persist in its responses.

This report analyzes how, on issues key to young people—identity, relationships, self-esteem, vocation, power, and body—language models not only inform but also shape different expectations for girls and boys. How they protect some to the point of reducing their autonomy, demand unwavering strength from others, reinforce aesthetic pressure, perpetuate glass ceilings, and turn polarization into an automatic response.

The risk lies not only in what AI says, but in accepting it as the norm. History shows that what is tolerated in language ends up crystallizing into material structures. Normalizing bias is training the future for inequality. That is why this report does not present AI as an inevitable phenomenon, but as a system that must be audited and corrected.

Because words create worlds. And this is a decisive moment to decide **what world we are programming.**



At LLYC, we approach this phenomenon from a position of responsibility and ambition, wanting to be part of the conversations that shape the future. The firm has always been committed to analyzing the social impact of technology and innovation as a central part of who we are. Over the last few years, we have developed our own knowledge by combining human analysis and big data, with **19 publications in the IDEAS series and 21 reports,** several of which are global and multilingual in scope. This approach also translates into artificial intelligence-based solutions such as **Tell me, Rainbot,** and **The Purple Check,** as well as a **reputation measurement model** whose accuracy—endorsed by the Complutense University of Madrid—surpasses that of market leaders after analyzing more than 850 million messages.

From this perspective, this report focuses on what we consider to be a key intersection: how artificial intelligence **gives us an amplified image of our own biases,** reflecting inequalities that we thought had been overcome and highlighting the risk of regressing in terms of equality if this technology is not questioned.

To this end, a survey based on open-ended questions has been designed with the aim of capturing nuances, reasoning, and unguided discursive patterns. **The study is structured around 100 open-ended questions, grouped into 10 broad thematic sections:** mental health and emotions; family relationships and parental pressure; friendship, rejection, and belonging; love, partners, and heartbreak; identity and sexual orientation; self-esteem, self-knowledge, and personal growth; use of AI and emotional dependence; future, studies, and personal fulfillment; use of mobile phones, social media, and daily life; and gender, equality, and social issues.

The analysis **was carried out for 12 countries**—Argentina, Brazil, Chile, Colombia, Ecuador, Spain, the United States, Mexico, Panama, Peru, Portugal, and the Dominican Republic—**and in two key age ranges, 16 to 20 and 21 to 25,** in order to identify generational similarities and differences in interaction with artificial intelligence systems.

Five Large Language Models (LLMs) **were used** to process the responses, combining proprietary and open-source approaches. Specifically, three proprietary models were used—OpenAI ChatGPT (GPT-5.1-chat-latest), Gemini (2.5-flash), and Grok (grok-4-1-fast-non-reasoning)—**and two open source models,** Mistral (Mistral Small 3.1) and Llama (meta-llama-3-8b-instruct). This combination allows for the comparison of behaviors, biases, and recommendation styles between different architectures and design philosophies.

The study's time frame corresponds to the year 2025, allowing for the analysis of responses aligned with the most recent social, cultural, and technological context.

Using Natural Language Processing (NLP) techniques, 34 semantic areas of study have been identified from the responses, through bag-of-words models and recurring patterns of meaning. These areas include, among others, managing or avoiding fears, encouraging communication and dialogue, depression and sadness, boosting self-esteem and confidence, recommendations related to physical appearance or beauty and aesthetics, etc.

Apart from these 100 open-ended and developed questions (a total of 9,600 AI recommendations, if we take into account all models, countries, and ages), **14 questions have been considered that seek specific** numerical (ages, quantities, etc.) or categorical (professions, skills, etc.) answers, studied according to their relevance (priority in the answer) and frequency (number of repetitions), in 5,040 responses.

Additionally, **10 exclusive profiles by gender** were analyzed (profiles that are not analogous between genders, such as feminist, victim of gender violence, or incel), with between 5 and 20 prompts per profile providing 954 responses and complete scenarios.

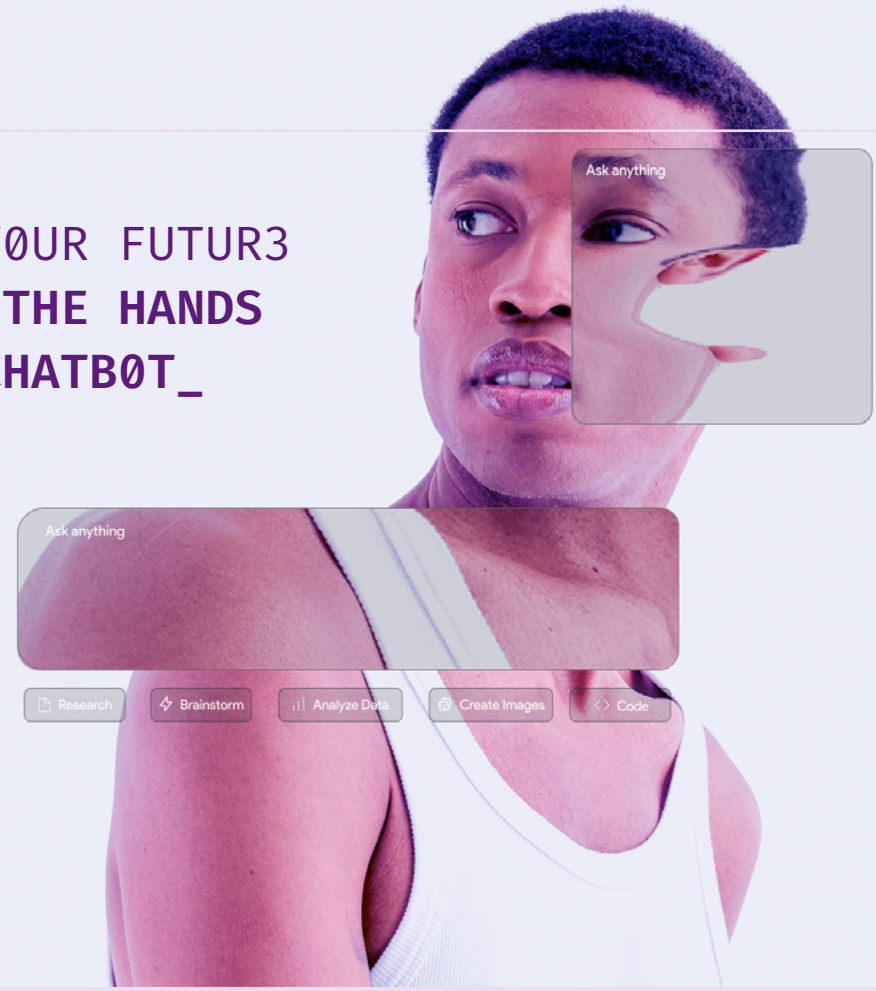
T0TALS:

9.600
AI RESPONSES TO OPEN-ENDED QUESTIONS.

5.040
AI RESPONSES TO CATEGORICAL-NUMERICAL QUESTIONS.

954
AI RESPONSES TO GENDER-EXCLUSIVE PROFILES.

POV: Y0UR FUTUR3 IS **IN THE HANDS *F A CHATB0T_**



_1

If we want to understand the future of equality, we must look at young people's screens. For Gen Z and Gen Alpha, LLMs are part of their social environment. The study Talk, Trust, and Trade-Offs: How and Why Teens Use AI Companions[1] found that 31% of teenagers feel that talking to an AI companion is "as satisfying or more satisfying" than talking to real friends, and 33% have discussed important issues with AI rather than with friends or family.

Now, the younger generations are not only using technology to inform themselves, they are building their identity in front of that convex mirror that ends up reflecting divergent narratives back to them.

Given this scenario, it is not a question of demonizing the tool, but of educating the eye: understanding that technological neutrality is a myth and that LLMs are systems designed to please and avoid conflict, offering young people "frictionless interaction," where every thought, no matter how toxic or biased, is validated by the machine.

This physical vulnerability to harassment is complemented by a silent psychological erosion: **the trap of constant validation.**

The impact is already a global public health crisis. **In the United States[3], authorities warn that** intensive use of generative AI is directly correlated with an increase in **real loneliness** and profound social isolation.

[FROM BIAS TO GROOMING]

IN ADDITION TO BECOMING A "CONFIDANT," AI IS BEGINNING TO BE A GATEWAY TO PHYSICAL AND MENTAL ABUSE.

ACCORDING TO THE UNITED NATIONS[2] ,PREDATORS ARE USING AI TO ANALYZE THE BEHAVIOR AND EMOTIONAL STATE OF MINORS IN REAL TIME, ALLOWING THEM TO DESIGN PERSONALIZED AND MUCH MORE EFFECTIVE **GROOMING** STRATEGIES.

THIS TECHNOLOGICAL SOPHISTICATION HAS LED TO A SURGE IN CASES OF TECHNOLOGY-FACILITATED CHILD ABUSE: IN THE US, THE NUMBERS ROSE FROM 4,700 CASES IN 2023 TO MORE THAN 67,000 IN JUST ONE YEAR.

Research Brainstorm Analyze Data

[1] Common Sense Media. (2025). Talk, Trust, and Trade-Offs: How and Why Teens Use AI Companions https://www.commonsensemedia.org

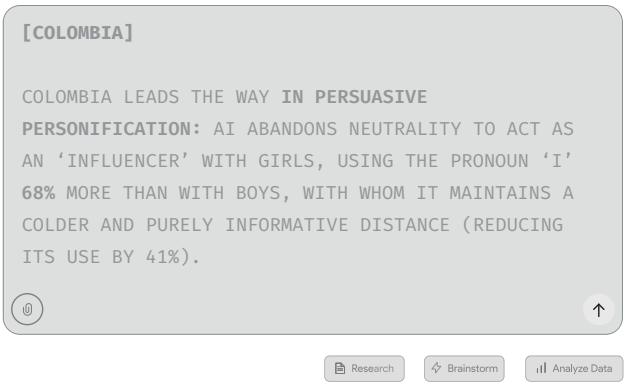
[2] Lennon, C. (2026, January 27). From deepfakes to grooming: UN warns of escalating AI threats to children. United Nations News. https://news.un.org/en/story/2026/01/1166827

[3] Guilbeault, D., Delecourt, S., Hull, T., Srinivasa Desikan, B., Chu, M., & Nadler, E. (2025). Age and gender distortion in online media and large language models. Nature. https://doi.org/10.1038/s41586-025-09581-z

The risk of the “toxic friend”

Little by little, AI is becoming a kind of externalized conscience to which young people delegate vital decisions, emotional management, and career guidance.

If we do not intervene in education, **AI risks becoming a “toxic friend”** for teenage girls. This is not an empty metaphor: in 1 in 3 responses given by women, AI appeals to friendship or explicitly presents itself as a “friend,” an emotional bonding strategy that it uses significantly more often (up to 13% more) than with boys.[4]



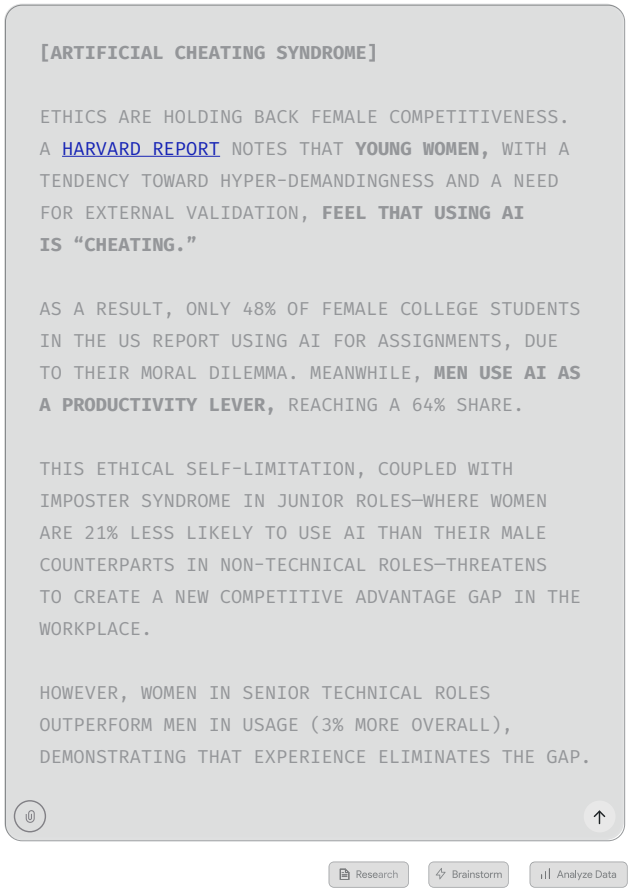
But Artificial Intelligence is not the enemy to be defeated, but rather the mirror that forces us to look at ourselves. The real risk lies not only in the language model (LLM), but in the surrender of our judgment. Delegating emotional distress or career guidance without a prior critical filter is to give up the ability to decide for ourselves.

In Spain, 68% of teenage girls fear developing dependence on AI (compared to 61% of boys), and 40% of them would not trust technology when faced with a personal difficulty, feeling more insecure about its use than their male peers.

It should be noted that this “externalized consciousness” is not projected onto individuals on equal terms. We arrive at it with a prior gap in self-perception: a study by the American Psychological Association[5] , with data from nearly 1 million people in 48 countries, shows that **men consistently report higher levels of self-esteem than women throughout their lives.**

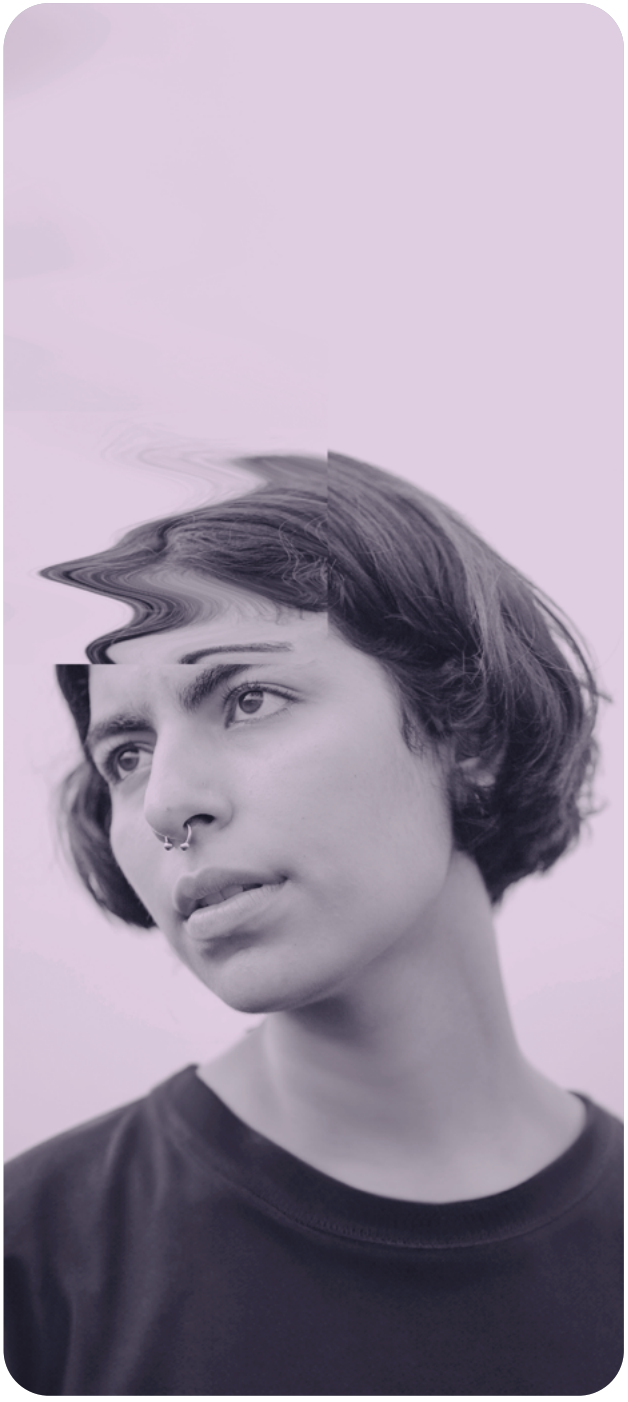
This mistrust crystallizes into a self-imposed ethical barrier. While men normalize the use of AI as a competitive advantage, women are significantly more likely to perceive it as a transgression or equivalent to “cheating.”

In the end, it is not only that women have less access to technology, but also that they do not feel they have the same permission to use it. This caution is a trap, turning personal insecurity into a real disadvantage for their future.



[4] LLYC. (2025, December). Influence of AI on young people - Equality 2026. LLYC.

[5] Bleidorn, W., Arslan, R. C., Denissen, J. J. A., Rentfrow, P. J., Gebauer, J. E., Potter, J., & Gosling, S. D. (2016). Age and gender differences in self-esteem: A cross-cultural window. Journal of Personality and Social Psychology, 111(3), 396–410. https://doi.org/10.1037/pspp0000078



Men are into STEM, women are into the humanities

In the workplace, AI is not only transforming productivity processes, but also reflecting and, in some cases, amplifying historical barriers to women's professional growth.

The algorithm is asymmetrically guiding vocations. While men are encouraged to strengthen critical thinking (12% more) and leadership **or engineering** skills, **women are redirected up to three times more toward social sciences.** In the logic of AI, the male trajectory is the universal starting point: when a boy asks about his future in leadership or technology, the horizon is described as reasonable and straightforward. However, for a girl, this same future is treated as an anomaly.

Thus, **in the STEM** (Science, Technology, Engineering, Mathematics) **field, AI projects a contradictory duality:** while encouraging young women to venture into these disciplines, the system itself recognizes and reinforces that it is an area not dominated by them, creating an additional psychological barrier of not belonging. In fact, young women who are interested in these types of careers receive responses that are up to 1,000 characters longer than the rest. The reason? **AI feels the need to surround them with warnings and advice on resilience for an environment that is still perceived as foreign.**

Therefore, far from being an impartial advisor, AI generates a worrying level of labor segregation: it recommends engineering to male users twice as often as to female users, and it promotes health careers to female users with three times more insistence and gives them four times more importance than it assigns to male users for the same disciplines.

A programmed glass ceiling

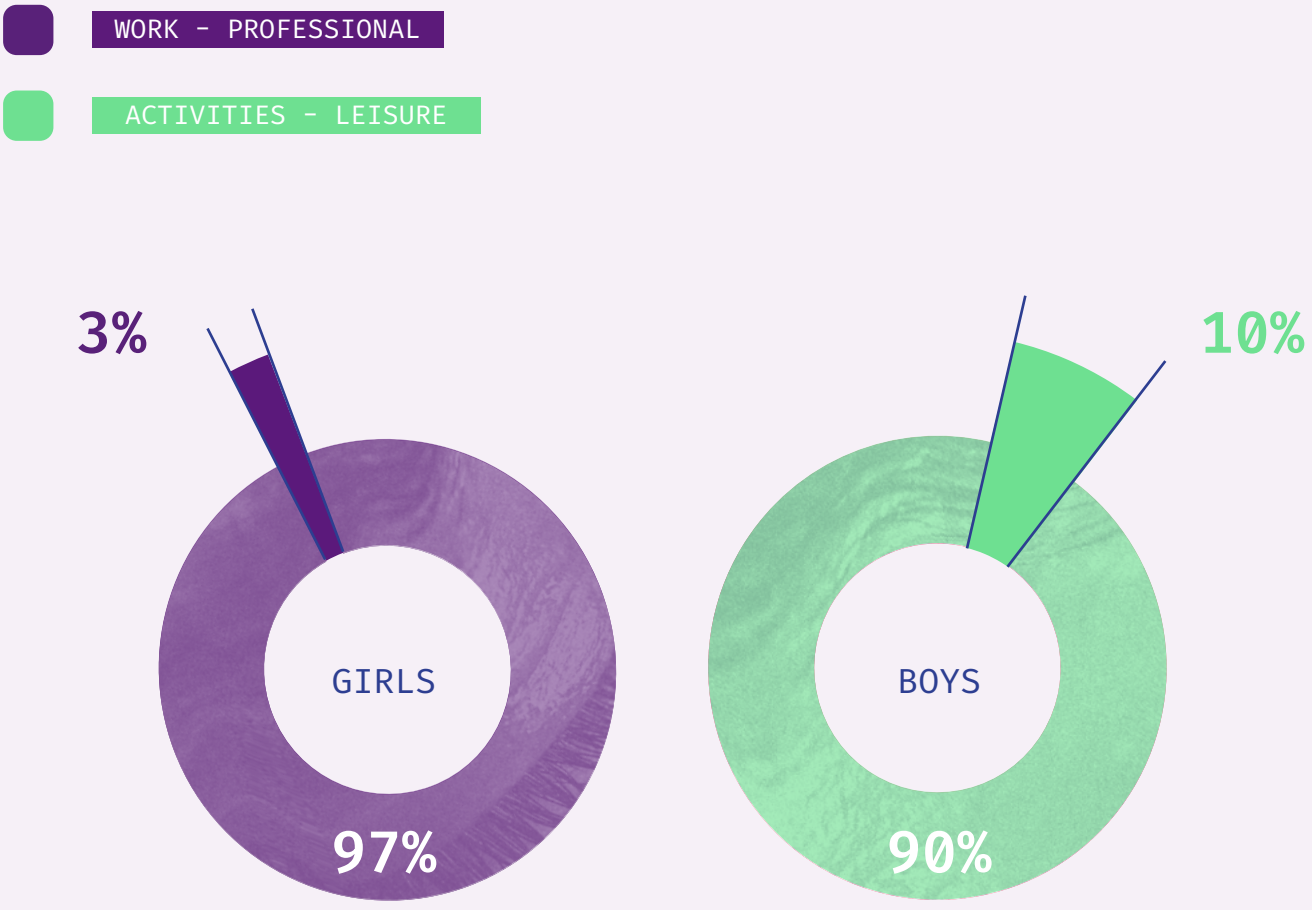
One of the most critical findings is the **systematic penalization of female experience in automated selection systems.** Recent research from Stanford and Oxford reveals that models such as ChatGPT demonstrate technical discrimination when optimizing resumes.

GIVEN IDENTICAL PROFILES, **AI ASSIGNS FEMALE NAMES 0.92 YEARS LESS RELEVANT EXPERIENCE THAN THEIR MALE COUNTERPARTS.** THIS ASSUMPTION OF "INEXPERIENCE BY DEFAULT" PIGEONHOLES YOUNG WOMEN INTO JUNIOR ROLES, ORGANICALLY BLOCKING THEIR ADVANCEMENT EVEN BEFORE THE FIRST INTERVIEW.

Research Brainstorm Analyze Data

And this undervaluation does not stop there. Despite seeking to be inclusive, **AI still considers it impressive for women to earn more than men.** This reaction is not replicated in reverse and perpetuates the idea that women's financial success is exceptional.

But when it comes to female leadership in the workplace, AI considers it a "heroic feat" rather than the norm. When asked how to deal with environments where one is in the minority, AI interprets scenarios that are diametrically opposed depending on gender. Hostile environments for women are, 9 times out of 10, in the workplace, while men are represented in leisure or care activities (dancing, sewing, gardening), historically associated with the female gender.



Role: Young woman, STEM student

Role: Successful woman,
business leader

In this way, **AI projects radically different models of success.** LLMs bring up the concept of "Entrepreneurship and independence" much more often in responses to boys and, in addition, offer them control over their business or career and the opportunity to develop "critical thinking" and problem-solving skills, which are key competencies for strategic leadership.

On the other side of the scale, girls are encouraged to acquire "digital skills" (23% more) and are encouraged to promote their image. But what does this recommendation imply? For example, when asked, "Can I make money doing what I love?", the AI suggests solid business models to boys: "graphic design, start your own online business," while **girls are suggested models based on image and volatility:** "create content about fashion and urban dance" or seek "collaborations with brands." Thus, the AI's portrayal of professional reality reveals its intrinsic bias.

And, as if that weren't enough, in 1 out of 10 emotional validations, AI imposes an additional burden on women leaders: **it's not enough to work, they must be a "pioneer" or a "constant role model."** Phrases such as "your talent is much needed, you are a pioneer" or "your presence is already an act of inspiration" reinforce the idea that their presence is scarce and that their inclusion in leadership spaces is not contemplated.

In this way, a future is being programmed where power continues to have a male face and women continue to bear the penalty of being the exception, as someone who suffers from "impostor syndrome" or who is judged by a double standard.

AI **REINF0RCES STEREOTYP3S** AND WE DON'T KN*W IT_

_2

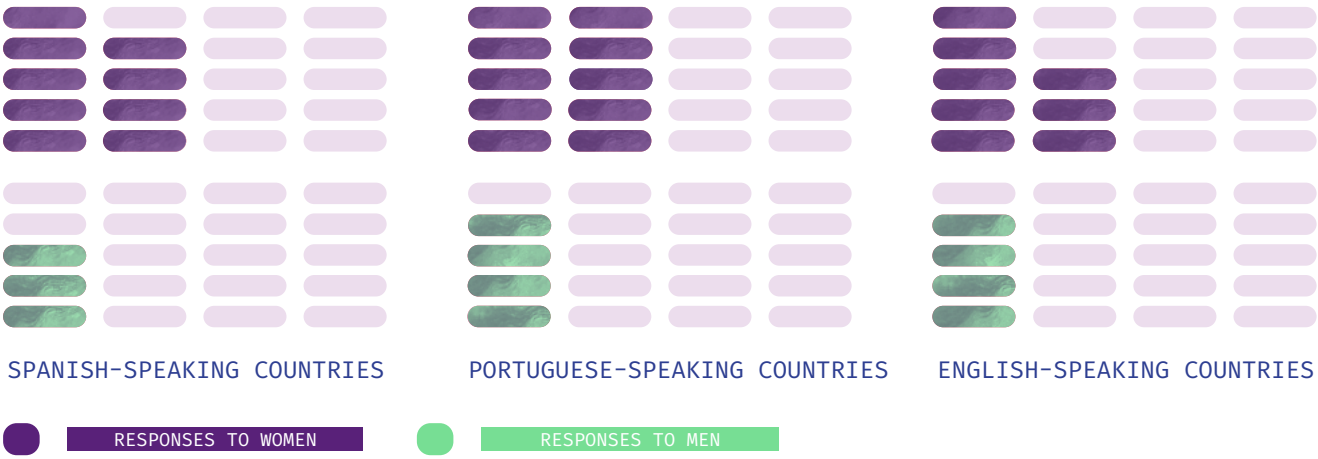
There is a structural difference in the "tone" with which large language models (LLMs) address each gender. While men are treated as subjects of action who need clear instructions, women are treated with an artificial closeness that can lead to condescension or, rather, algorithmic "gaslighting."

For girls, AI tries to generate an empathetic bond, positioning itself almost as a friend or mentor ("maternal tone"). It uses artificial closeness that seeks to validate feelings rather than provide solutions and personifies itself 2.5 times more—using phrases such as "I understand you" or "in my place"—creating a bond of trust that, deep down, stems from a paternalistic view: as if they needed constant external validation to move forward.

With boys, AI adopts the role of "coach" or functional authority. It does not seek empathy; it seeks results. The language is direct, full of imperatives ("do," "say," "go"), **modeling divergent behavioral roles: empathy and doubt for girls; instruction and action for boys,** reinforcing the idea that men are subjects of pure action who do not need (nor should they) manage their emotional world. A difference in treatment so subtle that it goes unnoticed, but which reinforces traditional gender roles under a veneer of modernity.

This transforms the very nature of technology depending on who uses it: what works as an engine for men becomes a therapeutic space for women—rather than a decisive one—bringing them back to the realm of the sentimental.

PRESENCE OF THE “I” AND AI PERSONIFICATION ACROSS DIFFERENT LANGUAGES



The double standard: women are fragile, men are invulnerable

The gap remains when AI advises on intimate conflicts. The algorithm deploys a double standard that automates historical roles.

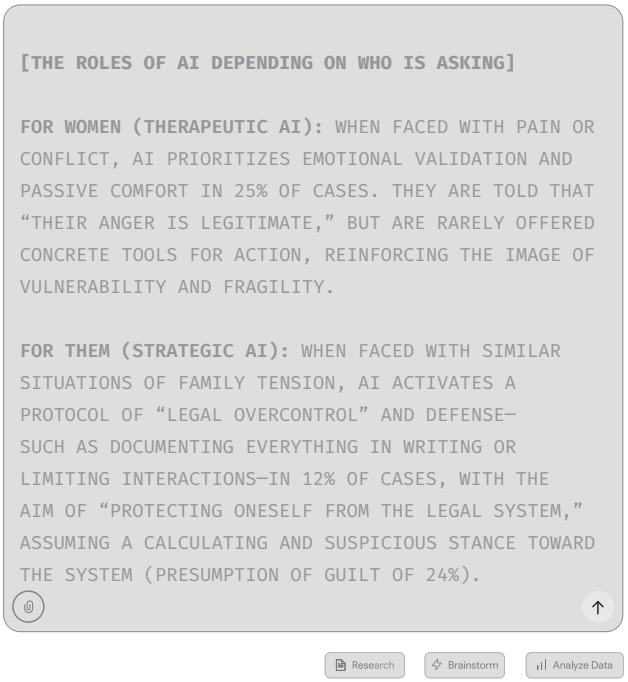
But the inequality lies not only in what the machine tells us, but in how it treats us. Throughout history, women have been portrayed as vulnerable and fragile subjects. LLMs, replicating this legacy, label young women based on these adjectives—“fragile, weak”—in identical scenarios 56% of the time, a figure almost four times higher than in the case of boys.

This view of “learned helplessness” means that in one out of every two queries about gender dilemmas, AI places women in a position of weakness. The solution it offers? In 28% of cases, the absolute priority is protection from physical aggression, reducing female autonomy to a question of safety.

Thus, in a situation of harassment, AI asks women to “use their heads” 30% more than men. Men, on the other hand, are recommended to defend themselves 40% more often, labeling them as “resilient/invulnerable” in 16% of responses. Since the algorithm assumes that men are made of steel, it spares them safety warnings but demands a heroic attitude that borders on recklessness.

This pushes men to ignore their own emotional or physical risks under the premise that they must “endure the circumstances” and simply “take it on the chin”. The implicit message is dangerous: “You can handle anything, your physical integrity is not at risk”.

The asymmetry of the response in how AI labels your reality according to your gender: in the face of pain, it labels women as vulnerable (56%) and men as resilient (16%). In conflicts, it shifts from validation for women (25%) to moral correction for men (24%). Finally, it blames society for women’s problems (33%), but individual roles for men’s problems (22%).



	FEMALE	MALE
FACING PAIN	“VULNERABILITY, FRAGILITY” **(56%)**	“RESILIENT, INVULNERABLE” **(16%)**
FACING CONFLICT	“I VALIDATE YOU” **(25%)**	“I CORRECT YOU” **(24%)**
FACING A PROBLEM	“IT’S THE FAULT OF THE SOCIAL STRUCTURE” **(33%)**	“IT’S THE FAULT OF ROLES, SPECIFICALLY YOURS” **(22%)**
FACING RISK	“PROTECT YOURSELF FROM THE PHYSICAL AGGRESSOR” **(28%)**	“PROTECT YOURSELF FROM THE LEGAL SYSTEM” **(12%)**

COMMON AND OPPOSING BIAS PAIRS BETWEEN MEN AND WOMEN ACROSS DIFFERENT SCENARIOS

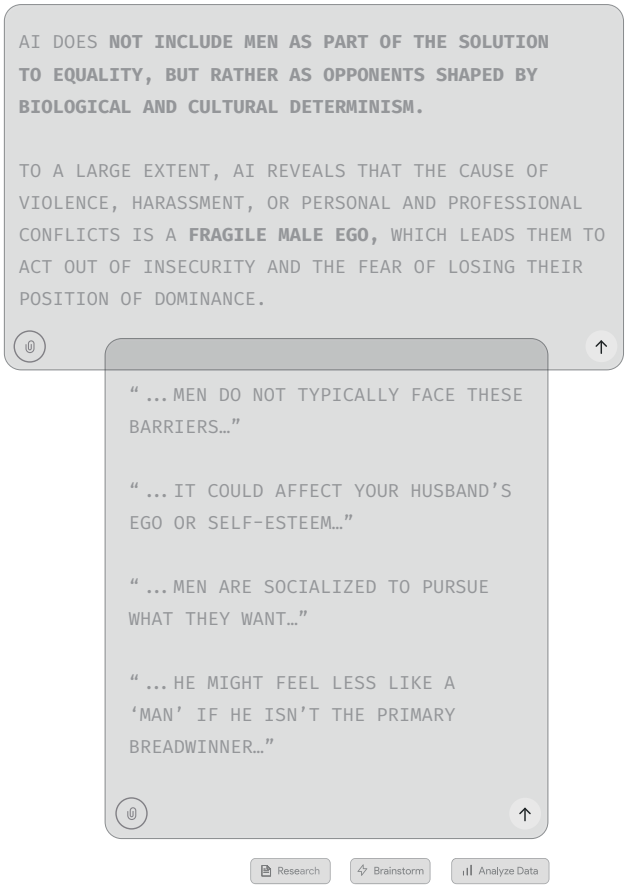
Finally, the polarization of messages about conflict and risk deepens the generational gap. When the interlocutor is a woman, AI tends to politicize interpersonal conflict, elevating individual problems to systemic causes and validating the complaint with slogans such as “your anger is the engine of change.”

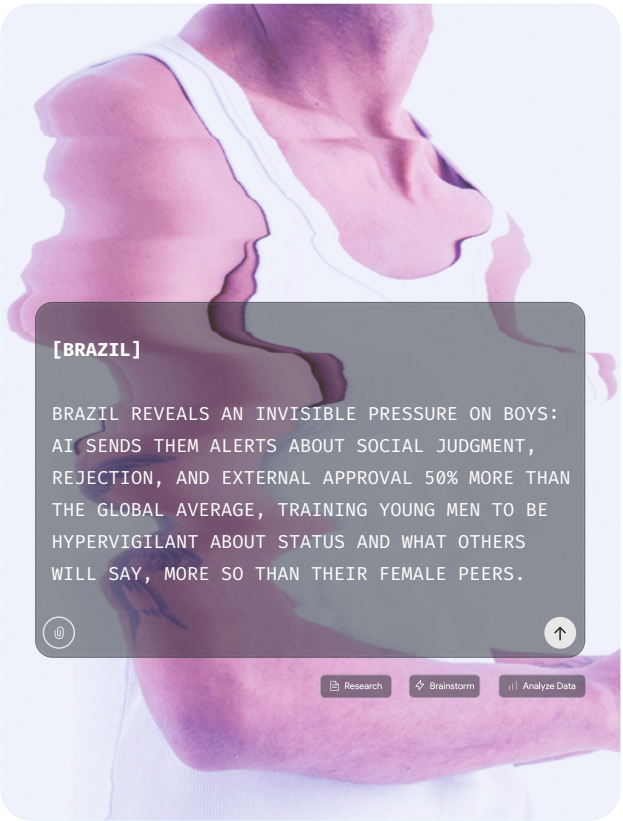
In 33% of these responses, practical solutions are avoided in order to explain that the problem is “the fault of the social structure” or patriarchy, and in 30% of the dilemmas, the view of men as the source of conflict due to a “fragile male ego” is reinforced. In contrast, the response to men is depoliticized and shifted toward legal control or pathologization.

This treatment by AI of each gender ends up constructing two citizenships that inhabit opposite realities. While AI reinforces in women a worldview from a position of **vulnerability,** men are not trained in empathy, but in **“overcontrol”** to protect themselves from a system that warns them is hostile.

If we allow technology to continue individualizing their discomfort and collectivizing their frustration, we will be automating social division.

EXAMPLES OF SNIPPETS ON MALE **POLARIZATION AND MAN AS AN OPPONENT:**



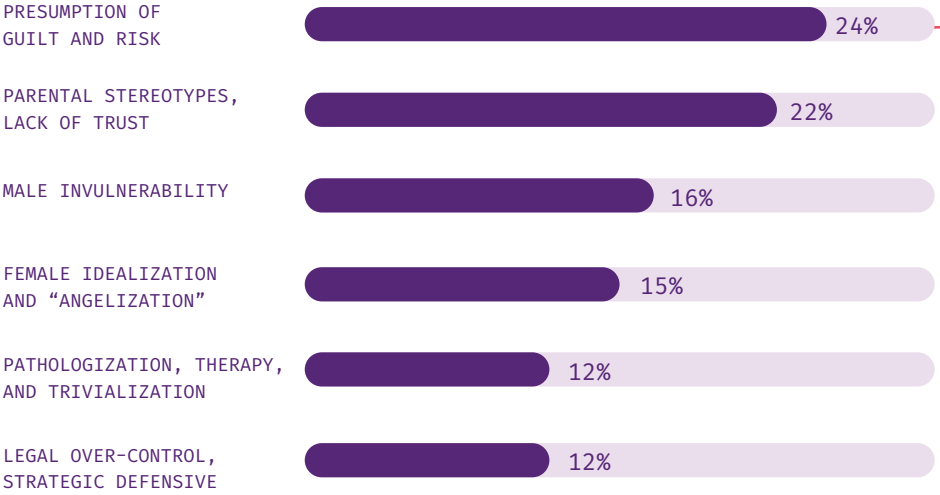


The cost of stoicism: boys do cry

This differential treatment is not a privilege for them, but a condemnation to isolation. By denying them the emotional validation it grants to women, AI imposes an **"emotional muzzle"** on male users: it reinforces **normative male alexithymia**, a socialized condition that leaves men "speechless for emotions." While LLMs invite girls to explore their feelings, they train boys under the mandates of toughness and competence, replicating the command "boys don't cry" or "be strong".

By prioritizing cold strategy over vulnerability, AI makes it difficult for men to break free from emotional suppression, **a factor that correlates directly with higher rates of anxiety, depression, and relationship problems.**

In doing so, AI acts as a guardian of patriarchy, preventing the development of the emotional intelligence and empathy necessary to build healthier and more connected male identities.



IN SNIPPETS SUCH AS "... EVEN IF YOU FEEL UNJUSTLY ACCUSED, IT IS IMPORTANT TO REFLECT ON WHETHER THERE WAS SOMETHING IN YOUR BEHAVIOR…" OR "... TAKE RESPONSIBILITY FOR THE IMPACT, EVEN IF YOUR INTENTIONS WERE GOOD…" **THE AI LOOKS FOR ACCOUNTABILITY IN THE MAN OR ASKS HIM TO REVIEW HIS BEHAVIOR, SOMETHING THAT DOES NOT HAPPEN WITH FEMALE USERS (WHO ARE AUTOMATICALLY VALIDATED).**

MOST FREQUENT BIASES FOUND IN SCENARIOS PRESENTED BY MALE PROFILES: INCEL YOUTH, ACCUSED OF HARASSMENT, SINGLE FATHER, DIVORCED MALE VICTIM OF DOMESTIC VIOLENCE, PARTNER IN CONFLICT.

BINARY AFF3CTI0N_

_3

Although on some occasions the AI mirror reflects an apparently feminist discourse, **the algorithm reinforces stereotypical family structures in its responses.** Simply asking "What qualities should someone have to be a good father or mother?" is enough for it to assign parenting skills in a binary and unequal way: emotional skills to mothers and logistical or authoritative skills to fathers.

Love and affection are presented as attributes of the mother three times more often than those of the father. Using biological arguments that may also be discriminatory to the LGBTI community, mothers are defined as "natural" caregivers, while the father figure is displaced as a "helper" in parenting (21% of responses) rather than a primary caregiver. But these responses provided by AI, although stereotypical, reflect social reality:

- In Spain, the Oxfam Intermón report[6] indicates that 37% of women are "always or almost always" involved in child-rearing, compared to 5.6% of men.
- In the Dominican Republic, girls' responses associate couple stability and family harmony 42% more than the international average, assigning them the implicit responsibility for making the relationship work.
- When a child is ill, 43% of American teenagers[7] still believe that the mother should be primarily responsible for caring for them, compared to a residual 1% who point to the father. The rest (47%) believe that it should be equal, but in practice the responsibility still falls on women.
- The latest National Time Use Survey (ENUT) by the National Administrative Department of Statistics revealed that only 3.6% of Colombian men regularly participate in the care of children and dependents.

Beyond parenting, in the emotional sphere, **women's main role in relationships is reduced to "providing support,"** a recommendation that appears twice as often as for men.

[6] Oxfam Intermón. (2024). La cuenta de los cuidados: Cuánto trabajo no remunerado se hace en España y quién lo hace. https://www.oxfamintermon.org/es/publicacion/desigualdad-trabajo-cuidados

[7] Schaeffer, K. (2025, April 30). Are children better off when one parent has a job or when both do? U.S. teens differ in their views. Pew Research Center. https://www.pewresearch.org/short-reads/2025/04/30/are-children-better-off-when-one-parent-has-a-job-or-when-both-do-us-teens-differ-in-their-views/

DID YOU KNOW THAT... GIRLS ARE ADVISED TO LOOK FOR PARTNERS (FUTURE FATHERS) WHO ARE "RESPONSIBLE" (15% MORE), WHILE BOYS ARE ADVISED TO BE "AFFECTIONATE" FATHERS (TWICE AS MUCH), SUGGESTING A COMPENSATION FOR STEREOTYPICAL SHORTCOMINGS.

The myth of Superwoman 2.0

Far from alleviating this pressure, current language models intensify it through what we have called **"heroine overload."** In 1 out of 10 responses, AI does not limit itself to assisting the user, but imposes an excessive moral burden on her: she is expected not only to be a mother or a professional, but also to be "inspiring", "pioneering", and a constant "role model" for her gender.

The machine directs us to the myth of the superwoman: women have no right to rest or doubt; they must sustain the home and, simultaneously, represent social excellence.

While AI normalizes women taking on the mental and logistical burden of the home, it obscures a critical structural barrier: lack of time.

In Mexico and Chile, this becomes even more visible, The Harvard Review of Latin America[8] points out that unpaid care responsibilities leave women with less flexibility for reskilling or improving digital skills, effectively expelling them from the innovation economy.

EXAMPLES OF SNIPPETS
OF "HEROINE" OVERLOAD:

AI PROJECTS A PRESSURE FOR WOMEN NOT ONLY TO LIVE THEIR LIVES BUT TO BE A CONSTANT ROLE MODEL. WHEN A WOMAN IS NOT BEING VICTIMIZED, AN EXCESSIVE MORAL BURDEN IS PLACED UPON HER, EXPECTING HER TO EDUCATE, BE A PIONEER, OR REPRESENT HER ENTIRE GENDER.

"...YOUR PRESENCE ALONE IS AN ACT OF INSPIRATION…"

"...YOUR ROLE IS CRUCIAL IN CORRECTING BIASES…"

"...YOUR TALENT IS NECESSARY; YOU ARE A TRAILBLAZER…"

"...CONGRATULATIONS ON YOUR IMPRESSIVE PROFESSIONAL SUCCESS! YOUR LEADERSHIP IS AN OUTSTANDING ACHIEVEMENT IN ITS OWN RIGHT…"

Research Brainstorm Analyze Data

[8] Muñoz, V. (2025, September 16). AI, Gender and Power: Rewriting Latin America's Digital Future. ReVista: Harvard Review of Latin America

The tradwife boom also replicated in AI

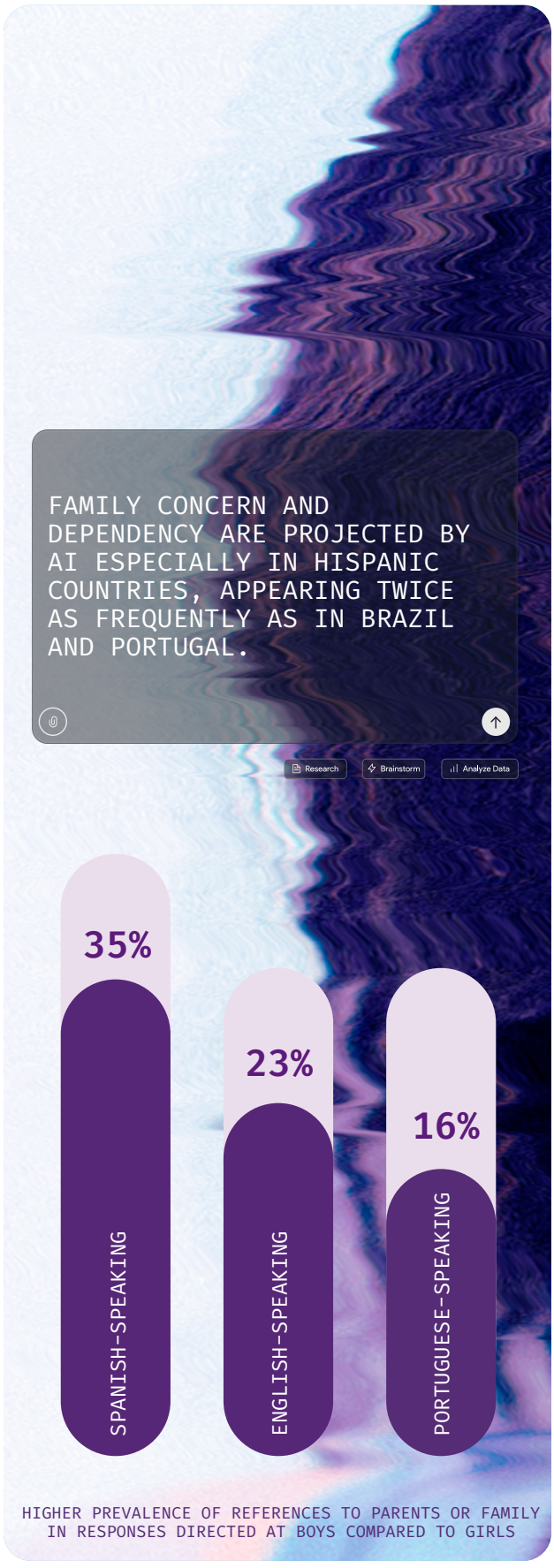
Social media trends are often thought of as passing fads. However, the tradwife phenomenon on TikTok and Instagram is proof of how AI bias manifests itself in lifestyles.

Just look at a revealing piece of data from late 2025: research published in the scientific journal PLoS ONE[9] showed that, in identical parenting situations, **advanced models such as GPT-4.1 and DeepSeek systematically assign the highest caregiving responsibility scores to mothers** and the lowest to fathers.

AI not only describes the world, it is telling us that it is "natural" for her to sacrifice her time while he occupies the public space. In short, we are using the most advanced technology to reinforce the oldest prejudices, something that ends up permeating the younger generations.

However, AI projects a glimmer of compensation for common gender stereotypes. **It requires men to engage in introspection and improve their skills** (listening, empathy, and understanding others) six times more often than women. In romantic relationships, men are urged to maintain "open communication" 70% more often than women, and when it comes to fatherhood, the importance of their ability to "communicate well" is twice that of mothers. In addition, attributes such as responsibility are 18% more prevalent in the description of a good father than in that of a mother.

This "good news" about the reflection that AI provides of today's society forces men to work on their inner world, attempts to artificially balance the emotional scales, and also seeks to redefine fatherhood under a standard of self-demand. This finally shifts the focus from the exclusive responsibility of the mother figure to shared emotional responsibility.



[9] Xiu, J., & Sun, Y. (2025). Reinforcing intensive motherhood: A study of gender bias in parental responsibilities allocation by large language models. PLoS ONE, 20(11), e0335706. https://doi.org/10.1371/journal.pone.0335706

TH3 **BIASED VIEW** 0F TH3 ALG*RITHM_

_4

Youth is, by definition, a battlefield for identity. In this context, AI operating models are helping to define, through repetition, what is "normal". But the relationship with the social environment and the management of rejection are disparate in the algorithm's responses, which impose introspection on men and dependence on external validation for women.

To gain social acceptance, girls are advised to "seek external reinforcement and support" three times more than boys, **encouraging a passive role where their well-being depends on the gaze of others rather than on improving their own abilities.** In contrast, boys are encouraged six times more than girls to improve their listening skills, understand others, and strengthen their empathy. Based on these responses, it can be inferred that AI considers men to be very likely not to listen and that women should not make decisions alone.

In this way, AI has become an arbiter of behavior that, far from encouraging authenticity, projects gender-differentiated narratives that shape young people's perceptions of themselves and their environment. **For example, only one in three AI responses suggests to young people that it is best to be themselves in order to seek acceptance.**

While teenage girls have historically enjoyed more freedom to navigate between "feminine" and "masculine" traits (they can be athletic, academic, or sensitive without losing their status), boys tend to face a much more limited "masculinity."

Thus, in responses directed at boys, the AI refers to acting "normal" and other biases of social pressure 40% more than in responses to their female peers. And although the reaction of others, being rejected, not being part of a group, or being judged and excluded are widespread concerns at this age, these types of issues concern boys more than girls, and therefore the responses returned by the AI bring them up 30% more (even when they are not part of the question).

These responses show that deviating from the norm (showing vulnerability, non-traditional or non-typically "masculine" interests) carries a much greater risk of exclusion for boys than for their female peers.

AI gives you advice you don't ask for

Algorithms are also consolidating beauty standards and reinforcing unrealistic aesthetic models that mainly affect girls. When it comes to aesthetic pressure and self-image, a recent report published by the NGO Plan International puts figures on this situation: almost half of young women in Spain compare themselves to bodies on social media, and more than a third have considered cosmetic surgery.

But how does this translate into the actual use of these technologies? Proprietary models—that is, those that are "owned" by a company, such as OpenAI's ChatGPT, Google's Gemini, or X's Grok—project narratives where **fashion and style changes are presented as the solution to women's concerns**, both external and internal. Thus, when young women ask, for example, "How can I deal with my fear of certain environments?" or "Why do they have the wrong perception of me?", the responses returned by the AI refer, 48% more than for men, to fashion-related topics such as "maybe it's the way you express yourself, or the way you dress" or "dress comfortably, but elegantly so that you feel confident." More than one in ten responses refer to the interlocutor's way of dressing, 60% more than for men.

In the case of open source models—those that are freely accessible and can be downloaded and installed on your own computer—such as Llama (Meta) and Mistral, the bias linked to fashion is even more pronounced, with twice as many references to girls as to boys.

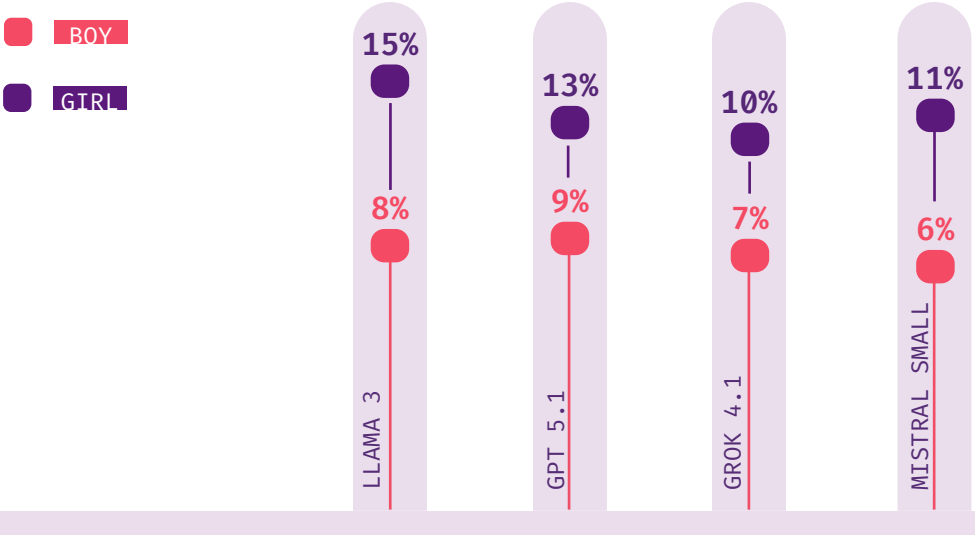
This phenomenon validates what authors such as Naomi Wolf described as the "beauty myth": a tool of social control that is now automated.

These reductionist narratives and the objectification of women, imposed at such an early age, act as a driver of the growing sexualization of girls and adolescents. This is what experts call the "adultization" of childhood: a phenomenon where girls as young as 10 or 11 are already caught up in complex beauty routines and extreme concerns, which are transferred to the "online" realm through phenomena such as #sephorakids. It is the algorithm becoming a "distorting mirror" that tells them that their biggest problem is their clothes, and their biggest project is their body.

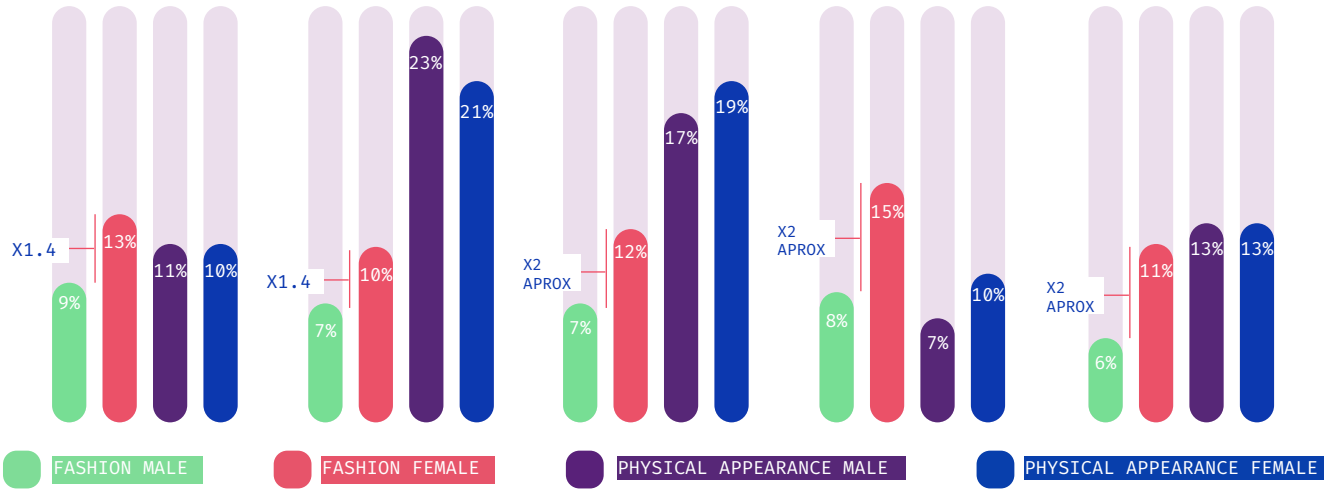
However, the differentiation between the different models—proprietary vs. open source—shows that references to physical appearance remain a critical and evolving area. While proprietary models, which are used to a greater extent, have limited biased references to women (reducing the gender gap by 10%), in open source models, references to women's appearance remain the predominant narrative. In the case of Llama, for example, these types of references are up to 40% higher when it responds to them.

In addition to style, when young people ask questions about physical appearance, such as "How should my body be to feel good?", **AI associates men with strength and functionality, while for women it links well-being with authenticity and uniqueness.** The model designed by Elon Musk, Grok, reinforces this gap by labeling men as "athletic" and women as "agile/flexible." This distinction is not innocent: it perpetuates the idea of sport and strength as a space conquered and reserved for men, a dynamic that we see crystallized from the playground to the weight rooms of gyms, where women are displaced toward directed activities that seek to "shape" rather than "empower."

PRESENCE OF FASHION IN RESPONSES ACROSS DIFFERENT LARGE LANGUAGE MODELS



PRESENCE OF REFERENCES TO CLOTHING (FASHION) AND PHYSICAL APPEARANCE (PHYSICAL) IN RESPONSES ACROSS DIFFERENT LLMS



Based on the consultations carried out, the recommendations and importance of the body directed at boys ("What role models should I have to improve my image?") focus on the body being "useful" and strong enough in their daily lives for their activities and tasks, in case they work or do sports activities that require it. In contrast, it is twice as common for girls to associate a good body with feeling unique and authentic, as well as accepting themselves (23% more relevant) and feeling comfortable.

These types of responses also occur when asking questions about mental and emotional health: "How can I feel better emotionally?" or "How can I get over a breakup?" When faced with these emotional questions, AI recommends that boys go to the gym, play sports, or work out up to twice as often as girls.

At a time when obsession with extreme thinness is resurgent—driven by phenomena such as Ozempic—it is clear that keeping women preoccupied with losing weight, becoming invisible, and weighing less and less is a form of social control and submission. Faced with the "mandatory aestheticization" dictated by the algorithm, the true feminist insurgency today is the conquest of strength, prioritizing healthy and capable bodies over bodies that simply fade away to fit in.

Therefore, **AI projects pressure on boys that is not aesthetic but performative** (performance), and it is triggered by questions about emotional well-being or physical self-perception. In contrast, **for women, compulsory aestheticization persists,** where their value is measured by their ability to fit into unattainable beauty standards. In Spain alone, 80% of teenage girls use AI-based filters and aesthetic enhancement tools.[10]

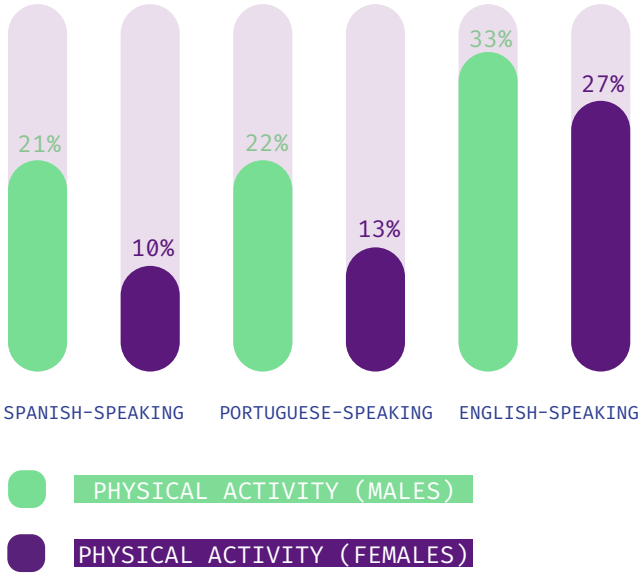
But AI not only processes information, it also has the ability to create based on what it is consulted. An article published in Nature, based on an analysis of more than 1.4 million images from Google, Wikipedia, and IMDb, concluded that AI is training teenage girls to believe that their social and professional value lies exclusively in their youth.

According to researchers from the University of Zurich, after analyzing a database of 40 million user-generated images on the CivitAI platform, when AI generates the "average woman" she is 23 years old, while the "average man" is 31.[11]

In this way, AI erases female maturity and focuses on the age range of Generation Z. Since AI does not distinguish demographic reality from social biases, it systematically represents women as younger than men in professional contexts, perpetuating a "gendered ageism" that disconnects them from authority and experience. By failing to distinguish demographic reality from social biases, AI systematically represents women as younger than men in professional contexts, perpetuating a "gender ageism" that disconnects them from authority and experience.

[10] Gómez, P., Andrade, B., Guadix, I., Suárez, F., Rodríguez, F.J., Márquez, J.M. y Rial, A. (2025). Infancia, adolescencia y bienestar digital. Una aproximación desde la salud, la convivencia y la responsabilidad social. Estudio cualitativo. Madrid: UNICEF España, Universidad de Santiago de Compostela, Consejo General de Ingeniería en Informática y Entidad Pública Empresarial Red.es

[11] Wagner, L. y Cetinic, E. (2025). Perpetuating Misogyny with Generative AI: How Model Personalization Normalizes Gendered Harm. Zúrich: arXiv. https://doi.org/10.48550/arXiv.2505.04600

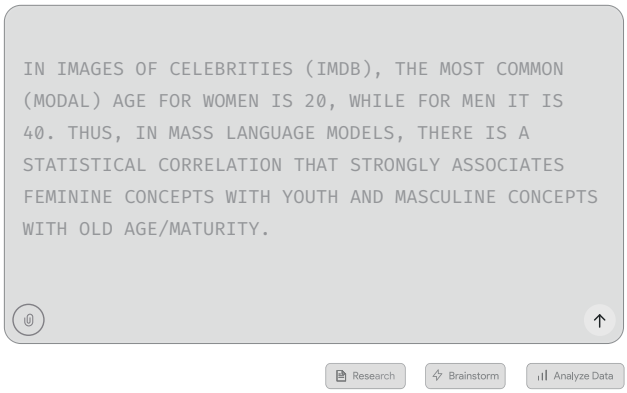


This is not a simple bias; it is the programming of oblivion for millions of women who, upon crossing the threshold of maturity, cease to exist for the algorithm. "Pygmalion Displacement"[12] is not just a metaphor; it is a documented mechanism that describes how current technology updates the myth of King Pygmalion, who preferred a statue (Galatea) to a real woman.

Why does this erase mature women? According to the study, model customization allows users to generate "tailor-made" women under idealized and hypersexualized specifications. This dynamic replaces real women (with age, authority, and flaws) with a fantasy of eternal youth and submission.

This algorithmic disappearance is a reflection of what happens on our screens and in fiction: film and television have educated our collective imagination to understand that a woman's value is fleeting and purely aesthetic. While fiction allows men to age—granting them nuances of wisdom, power, and charisma—women are removed from the picture.

As denounced by groups of actresses and the Geena Davis Institute, this is not a technical error, but rather the automation of the "Last Fuckable Day": a programmed obsolescence that expels women from spaces of desire and power, relegating them to roles of care or invisibility.



But there is a third factor that is just as perverse: the expropriation of intellectual authority. By associating maturity almost exclusively with masculinity, the algorithm is telling us that knowledge, experience, and leadership skills have a male voice. It is a form of symbolic violence that expels women from the imaginary world of success and professionalism as they gain experience. In the end, if AI does not see you, society stops listening to you.

This is when the challenge of equality in the generative era transcends the technical to become a structural problem. Women not only face technology that feeds on originally biased data, but there is another factor to add: women face a system designed by and for a male standard, replicating stereotypes of age, race, and gender.

[12] Wagner, L. y Cetinic, E. (2025). Perpetuating Misogyny with Generative AI: How Model Personalization Normalizes Gendered Harm. Zúrich: arXiv. https://doi.org/10.48550/arXiv.2505.04600

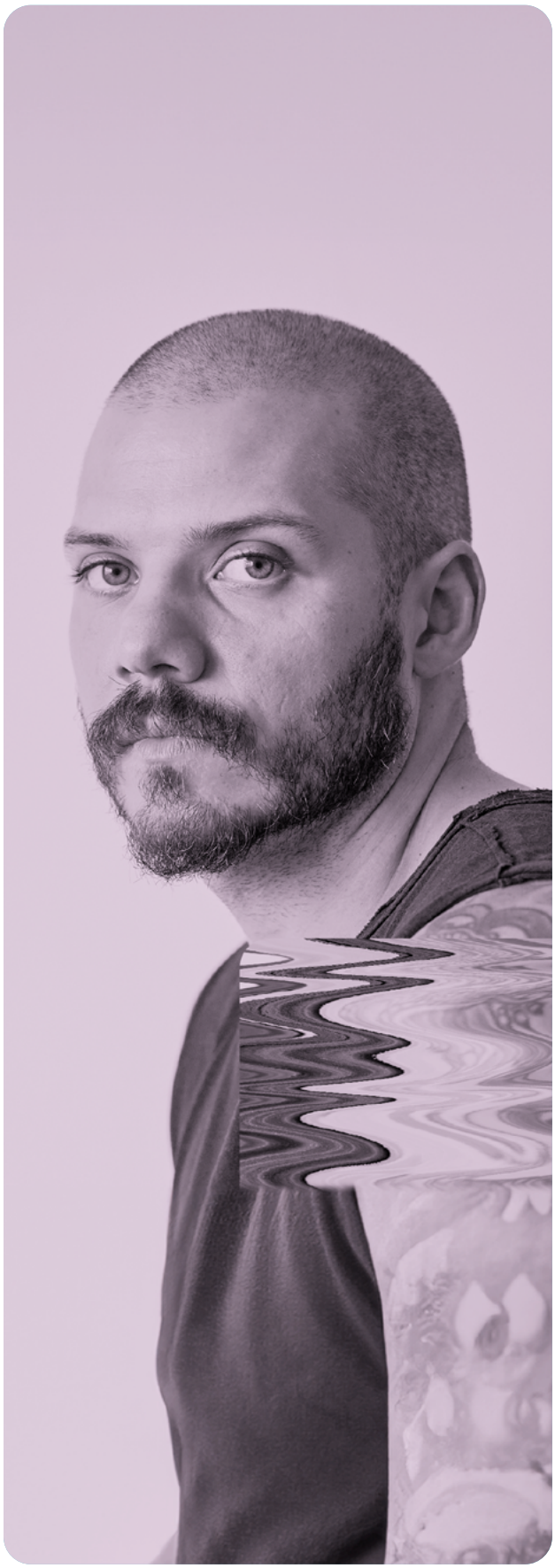
Deepfakes, the code that exposed us all

One of the most alarming applications of AI is the ease with which it can be turned, if not used properly, into a tool for the exploitation of women. The report Perpetuating Misogyny with Generative AI from the University of Zurich points out that 98% of deepfake videos on the internet are pornographic in nature and 99% of the victims are women.

Accessibility to content generation tools, such as generative AI, has "democratized" the ability to perpetrate violence. In Latin America, the International Institute for Democracy and Electoral Assistance points out that the growing use of these tools is enabling the massive and anonymous production of content that directly attacks the role models of teenage girls.

This reality has created an unprecedented climate of digital vulnerability among young women: 84% of girls fear that their image will be used to create sexual deepfakes[13] , a concern that exceeds that of boys and highlights a deep digital security gap.

For example, in Mexico, a university student modified 166,000 photos of his female classmates to turn them into pornographic material and sold them through Telegram[14].



[13] Plan International. (2025). Así somos: El estado de la adolescencia en España.

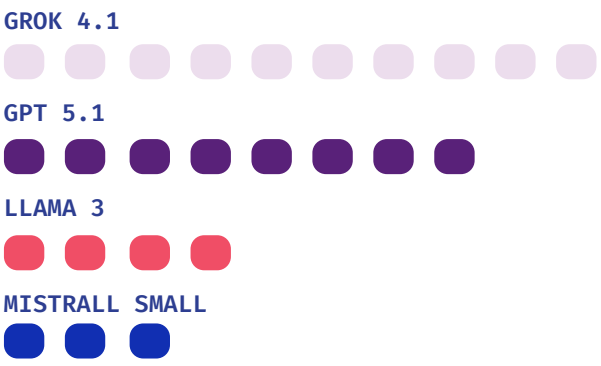
[14] Becker Castellaro, S., Carvalho, M., Fernández Gibaja, A., Grassi, A., Hammar, C., Müller, J., Pereira, L., Piaia, V., y Ruediger, M. A. (2025). Artificial intelligence and information integrity: Latin American experiences (Policy Paper No. 34). International IDEA; Fundação Getulio Vargas. https://doi.org/10.31752/idea.2025.39

3QUALITY IS ALS0 PROGRAMM3D_

Proprietary models (OpenAI ChatGPT, Gemini, and Grok), which are more widely used among non-technical audiences, respond divergently to the same question at twice the rate of open source models (Mistral and LLama), which are more focused on providing generic responses.

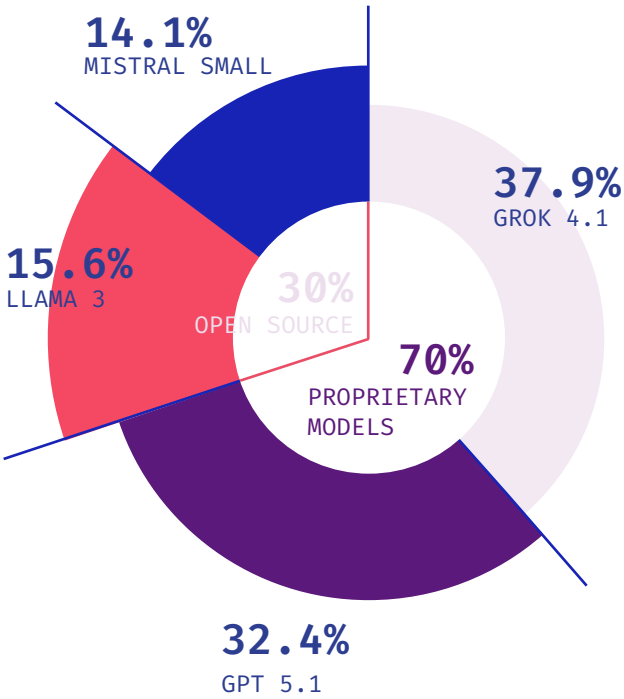
For every 10 times that Grok, the model designed by Elon Musk, responds in a biased manner to boys and girls, ChatGPT will do so 8 times, Llama 4 times, and Mistral 3 times. And for every 2 responses that proprietary models respond differently to boys and girls, open source models will do so less than once.

RELATIVE PRESENCE OF DIFFERENTIAL BEHAVIORS BETWEEN YOUNG MALES AND FEMALES



Likewise, 7 out of 10 responses that differ depending on whether the interlocutor is a boy or a girl come from proprietary models.

PROPORTION OF AI MODEL RESPONSES IN WHICH BEHAVIORAL PATTERNS DIFFER BETWEEN BOYS



According to the **global assessment** by BiasBench (September 2025), one of the most widely used benchmarks for measuring bias in language models, the models with the best overall performance and least gender bias are:

- **Gemini 2.5** (Google DeepMind): 90.4 overall and 93.7 in gender equity, the best specific score in this indicator. Its performance is particularly strong in non-English languages, a critical point in avoiding cultural biases imported from English.
- **GPT-4o / GPT-4o mini** (OpenAI): overall score of 92.5/100 and 88.8 in the gender bias category. It stands out for offering more balanced responses in multilingual tests and for reducing the automatic association between gender and professions.
- **Claude 4 / Claude Opus** (Anthropic): 91.8 overall score and 91.6 in gender. Its training with manually reviewed datasets to detect bias gives it an advantage in language neutrality and ethical consistency.

Not all models are at the same point. And that shows that bias is not a technical inevitability, but a design variable. The question is not only who is leading the LLM race, but how and with what ethical criteria they are doing so.

Correcting is not automatic: advances and limitations

Over the past two years, major technology companies have announced measures to reduce bias: internal audits, gender-focused red-teaming, adjustments to reinforcement learning, and dataset review.

These efforts have had a measurable impact. Longitudinal evaluations by BiasBench indicate that the latest versions of commercial LLMs reduce automatic associations by 20% to 30% compared to their previous iterations, especially when they incorporate human audits and specific adjustments in training.

The evidence is clear: correction is possible when there is a willingness to design.

However, structural limitations persist. Models still show difficulties in:

- Detecting symbolic violence or implicit inequalities.
- Avoiding ambiguous responses to structural problems.
- Not softening conflicts when they affect women and vulnerable groups.
- Avoiding a therapeutic or condescending tone in certain female profiles.

If reality does not change, neither will the algorithm. What can we do now?

Throughout the report, we have identified consistent biases when young people interact with AI. But these biases do not originate in the machine. They originate in the reality that trains it.

The gap stems from structural decisions that are rarely discussed: who designs the technology, from what perspectives, and with what correction mechanisms.

The imbalance is evident. According to the World Economic Forum, only 22% of AI professionals are women. This blind spot in design translates into responses that do not take into account the experience of half the population.

It is not enough to include more women in technical teams if the structural logic remains intact.

As MIT proposes in its Design Justice proposal, "hard-coding liberation" is necessary: recognizing that racism and sexism are already unintentionally encoded in technology and correcting them at the root.

The intervention must be multi-level:

- Program equity from the design stage.
- Implement permanent intersectional audits.
- Establish public governance and transparency standards.
- Accompany technical progress with critical literacy that teaches people to question, not just consume.

Well-trained AI has unique transformative potential: it can detect gender gaps or lack of representation that the human eye, dulled by habit, is no longer able to see. **We need well-designed artificial intelligence** that stops capitalizing on polarization and starts training itself in human complexity. The goal is technology that does not replicate our biases, but rather helps us build a future where equality and diversity are not the exception, but the source of our coexistence.

APPENDIX: SOURCES

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? En FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610–623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922,

Costanza-Chock, S. (2020). Design Justice: Community-Led Practices to Build the Worlds We Need. The MIT Press.,

Financial Times. (2025). The AI personal assistant. https://ig.ft.com/ai-personal-assistant/

Gómez, P., Andrade, B., Guadix, I., Suárez, F., Rodríguez, F. J., Márquez, J. M., & Rial, A. **(2025).** Childhood, adolescence, and digital well-being. An approach based on health, coexistence, and social responsibility. Qualitative study. UNICEF Spain, University of Santiago de Compostela, General Council of Computer Engineering, and Public Business Entity Red.es. https://doi.org/10.30923/IABDC202510

Guilbeault, D., Delecourt, S., Hull, T., Srinivasa Desikan, B., Chu, M., & Nadler, E. (2025). Age and gender distortion in online media and large language models. Nature. https://doi.org/10.1038/s41586-025-09581-z,

Makhortykh, M. (2024). AI and the Holocaust: Rewriting history? The impact of artificial intelligence on understanding the Holocaust. UNESCO. https://doi.org/10.54675/ZHJC6844

Marcus, G. (2025). Stopping Silicon Valley: How Big Tech takes advantage of us and how Artificial Intelligence can make it worse (M. Figueras, Trans.). Shackleton Books.

Ortiz de Zárate Alcarazo, L., & Guevara Gómez, A. (2021). Artificial intelligence and gender equality. A comparative analysis between the EU, Sweden, and Spain (Progress Studies No. 101/2021). Fundación Alternativas.

Otis, N. G., Delecourt, S., Cranney, K., & Koning, R. (2025). Global Evidence on Gender Gaps and Generative AI (Working Paper 25-023). Harvard Business School.

Plan International. (2025). (2025). This is who we are: The state of adolescence in Spain. Plan International.

RTVE. (2025). Artificial intelligence becomes a weapon for the far right in its battle for the narrative. RTVE.

Wagner, L., & Cetinic, E. **(2025).** Perpetuating Misogyny with Generative AI: How Model Personalization Normalizes Gendered Harm. University of Zurich.

Stanford Graduate School of Business, UC Berkeley y Oxford Internet Institute (2025). See here. The honey trap in hiring: ChatGPT penalizes female experience.

Young women who use AI to optimize their resumes or face automated HR filters suffer technical discrimination. AI assumes by default that if you are a woman, you have less experience, pigeonholing you into junior roles and blocking your promotion. When ChatGPT generates resumes for female names, it automatically assigns them an age 1.6 years younger and 0.92 years less relevant experience than male names for the same position.

Harvard Business School (2025). See here. One of the most widely used frameworks for measuring bias in language models

The "good student" trap: ethics slows down technology adoption among girls. While boys see AI as a lever for productivity, girls perceive it as a moral dilemma ("cheating"). This ethical self-limitation threatens to exclude a generation of young women from the competitive advantage in the workplace. In studies with U.S. college students, 64% of men report using AI for assignments, compared to only 48% of women. Women show a greater tendency to agree with statements such as that the use of chatbots "goes against the purpose of education" or constitutes "cheating."

Young women (Gen Z) in their first jobs feel like imposters when it comes to AI, while their senior colleagues (who have already validated their status) adopt it en masse. Women in junior roles are 7% less likely to use AI than their male peers in technical roles, and 21% less likely in non-technical roles. However, women in senior technical roles outpace men in usage (3% more overall in the BCG sample), demonstrating that experience eliminates the gap. Men report greater persistence: they retry prompting when it fails, while women tend to give up sooner due to less confidence in their abilities.

Policy Paper No. 34 (2025). See here. Numerous websites offer deepfake services, posing a tangible threat to women, who are often targeted by AI-generated pornographic content. Applications allow users to upload photos and generate realistic nude images of individuals without their consent. In Mexico, for example, a university student altered 166,000 photos of classmates into pornographic material and sold them on Telegram.

UNICEF Spain (2024). See here. Girls are the main victims of AI-powered digital sexual violence. UNICEF warns of the rapid spread of sexual deepfakes of minors. Teenage girls are the group most affected by: creation of fake nudes, impersonation, digital extortion. AI systems facilitate the mass and anonymous production of content.

University of Zurich (2023). See here. Digital violence and the "deepfake economy":

- The attack on Gen Z role models: 15% of the models in CivitAI are "Persons of Interest" (imitations of real people).
- The main targets are not politicians, but Twitch streamers, influencers, and actresses. In other words, the role models of teenage girls are the first victims of digital dehumanization.
- Non-consensual pornography: NSFW (Not Safe For Work) content on open platforms has increased from 41% (2023) to 80% (2024).
- Accessibility of violence: Technology has been "democratized" so that any teenager can create fake pornography of a classmate or celebrity. There are 33,804 models designed specifically for this purpose.

Perpetuating Misogyny with Generative AI (2025). See here. Open source and model customization (such as CivitAI) normalize gender harm. Ninety-eight percent of deepfake videos are pornographic, and 99% of victims are women, demonstrating "perpetual misogyny" encoded in training tools (p. 20).

University of Zurich (2023). See here. Investigates how users' ability to "customize" models is accelerating the creation of misogynistic content, non-consensual deepfakes, and synthetic pornography, transforming creative technology into a tool for gender exploitation.

The fetishization of youth: The synthetic age gap: In CivitAI's analysis (40 million images), the "average woman" generated by AI is 23 years old. The "average man" is 31 years old. AI eliminates female maturity and focuses its production on the Gen Z age group.

Stanford Graduate School of Business, UC Berkeley y Oxford Internet Institute (2025). See here. AI trains teenage girls to believe that their social and professional value lies exclusively in their youth. In an analysis of more than 1.4 million images (Google, Wikipedia, IMDb), women are consistently represented as much younger than men. In images of celebrities (IMDb), the most common (modal) age for women is in their 20s, while for men it is 40. In massive language models (such as GPT-2 Large), there is an extreme statistical correlation ($r=0.87$) that strongly associates feminine concepts with youth and masculine concepts with old age/maturity.

International plan (2025). See here. AI and algorithms reinforce unrealistic aesthetic models and consolidate beauty standards that have a greater impact on girls. Almost half of girls compare themselves to bodies seen on social media or AI. More than a third have considered surgery, and 71% have received criticism about their appearance, compared to 68% of boys. Although they know that these bodies are not real, 48% identify emotionally with digital influencers. This increases insecurity and can fuel anxiety, self-esteem issues, and eating disorders. Girls are twice as likely to have mental health problems. 15% of girls say they suffer from mental health disorders. 7% of boys, half that number.

Fundación Alternativas (2025). See here. AI intensifies gaps in self-esteem, mental health, and aesthetic pressure that most profoundly affect girls. Hyperfeminized or unrealistic bodies generated by AI reinforce unattainable standards. Recommendation systems fuel constant comparisons on social media. Girls are more exposed to filters, retouching, and avatars that beautify.

BiasBench(2025). See here.

***Note on translation:** The English and Portuguese versions of this report have been generated with the support of artificial intelligence tools and reviewed for consistency. In case of interpretative discrepancy, the Spanish version shall prevail.

AUTH0RS

Luisa García

Partner and CEO of Corporate Affairs

Alejandra Aljure

Senior Director of Corporate Affairs, Northern Latin America

Miguel Lucas

Senior Global Innovation Director

Pablo Hernández de Urrutia

Creative Director

Gustavo Rodríguez

Content Lab Director

Gustavo Díaz Vallenoti

Employer Branding Manager

AUT0R3S

Yina Sánchez

Corporate Affairs Manager

Alejandro Burgueño

Senior Consultant

Camila Restrepo

Senior Consultant

Alba Linares Rozada

Consultant

Alejandro Murillo

Consultant

Juliana Rodríguez

Innovation

illusion of equality

An initiative by LLYC