



**Citation:** Tirocchi, S. (2024). Artificial violence. VAW and the dark side of artificial intelligence. *Media Education* 15(2): 21-29. doi: 10.36253/me-16542

**Received:** September, 2024

**Accepted:** November, 2024

**Published:** December, 2024

© 2024 **Author(s)**. This is an open access, peer-reviewed article published by Firenze University Press (<https://www.fupress.com>) and distributed, except where otherwise noted, under the terms of the CC BY 4.0 License for content and CC0 1.0 Universal for metadata.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Competing Interests:** The Author(s) declare(s) no conflict of interest.

## Artificial violence. VAW and the dark side of artificial intelligence

### Violenza artificiale. Violenza contro le donne e il lato oscuro dell'intelligenza artificiale

SIMONA TIROCCHI

*University of Turin, Italy*  
simona.tirocchi@unito.it

**Abstract.** Violence against women is increasingly becoming a global phenomenon, involving millions of people worldwide, reinforcing the condition of social inequality that has historically relegated women to a subordinate position, within a framework of inequality and discrimination. The development of digital technologies (Web 2.0, social media, digital platforms) and, more recently, artificial intelligence, is contributing to the emergence of new, sometimes more insidious than traditional ones, forms of violence, raising questions about how to combat them. Specifically, the article proposes a theoretical reflection on the new forms of digital violence enabled by technology, with a focus on the case of the chatbot Replika. The qualitative and quantitative expansion of violence poses new media-education challenges, particularly regarding the need to design and propose new forms of prevention suited to this new landscape and to strengthen specific AI Literacy.

**Keywords:** AI literacy, artificial intelligence, cyberviolence, gender education, replika, VAW.

---

**Riassunto.** La violenza contro le donne è sempre più un fenomeno di portata globale, che coinvolge milioni di persone in tutto il mondo, contribuendo a consolidare la condizione di disuguaglianza sociale che da sempre relega le donne in una posizione subalterna e in una cornice di disuguaglianza e discriminazione. Lo sviluppo delle tecnologie digitali (web 2.0, social media, digital platforms) e, più recentemente, dell'intelligenza artificiale, sta contribuendo a configurare nuove forme di violenza, a volte più insidiose di quelle tradizionali e pone interrogativi in merito ai modi in cui combatterle. Nello specifico, l'articolo propone una riflessione teorica sulle nuove forme di violenza digitale favorite dalle tecnologie con un approfondimento sul caso del chatbot Replika. L'estensione, anche qualitativa, della violenza, pone nuove sfide mediaeducative con riferimento alla necessità di ideare e proporre nuove forme di prevenzione adeguate al nuovo scenario e di consolidare una specifica AI Literacy.

**Parole chiave:** alfabetizzazione all'intelligenza artificiale, intelligenza artificiale, cyberviolenza, educazione di genere, replika, violenza contro le donne.

---

## 1. INTRODUCTION

Violence Against Women (VAW) is a global problem, increasingly becoming the focus of social discourse and public attention. It is a phenomenon of vast proportions and is finally beginning to be acknowledged as significant on both social and cultural levels. Women have historically experienced discrimination in gender relations, especially in relation to the consolidation of a patriarchal model based on domination and symbolic violence exerted by men over women (de Beauvoir, 1975; Bourdieu, 2001).

According to data from the World Health Organization, violence against women constitutes a severe public health issue and a violation of women's human rights (WHO, 2024). Estimates published by WHO indicate that globally, almost one-third (27%) of women aged 15-49 who have been in a relationship report experiencing some form of physical and/or sexual violence by an intimate partner (WHO, 2024).

These figures are consistent with those released by the European Commission and Eurostat (2022), which reveal that over 3,000 women are killed annually in Europe by their partner or a family member, and countless others suffer harm and harassment. At least two women per week are killed in the EU by an intimate partner or family member. Furthermore, 32% of sexual harassment perpetrators in the EU come from the workplace context.

The COVID-19 pandemic has exacerbated violence against women and girls (VAWG), as well as deepened structural inequalities, forcing many women to leave their jobs to attend to domestic tasks resulting from their altered circumstances. Additionally, the COVID-19 emergency and the growing mediatization of social life (Hepp, 2020) have further intensified interactions with digital technologies (EIGE, 2022), which are increasingly being used to perpetrate forms of violence that are harder to detect and thus more difficult to combat. Digital violence, in fact, can occur 24/7, even from locations distant from where the victims are physically located, as it does not require geographic proximity. Moreover, it allows perpetrators to remain anonymous with relative ease. During the COVID-19 pandemic, incidents of "Zoom-bombing" became frequent, characterized by the posting of racist, sexist, pornographic, or anti-Semitic content.

## 2. TOWARDS DIGITAL VIOLENCE

The development of new digital technologies has complicated the already challenging debate surrounding

the definition of "gender-based violence" or "violence against women."

According to the European Commission (2024), violence directed against a person because of their gender, or violence that disproportionately affects individuals of a particular gender, can be defined as gender-based violence (GBV). This can include violence against women, as well as domestic violence affecting women, men, or children living in the same household. Although women and girls are the primary victims of GBV, it also causes severe harm to families and communities. The advent of the Internet, followed by the rise of social media and digital platforms (Belluati & Tirocchi, 2023), has expanded the range of violent behaviors. Not only has cyberbullying – one of the first widely recognized and studied forms of digital violence – emerged as an extension of traditional bullying, but broader definitions of cyberviolence have also gained prominence. As Tirocchi, Scocco, and Crespi observe, "the concept of cyberviolence is proposed as broad and multidimensional; it goes beyond a simple definition of the risks of digital technologies for young people. Cyberviolence can be directed at any individual or group, although gendered cyberviolence is often referred to in the literature to mean, in particular, violence perpetrated against women" (Tirocchi, Scocco & Crespi, 2022, p. 6).

Another definition is offered by the European Commission's Advisory Committee on Equal Opportunities for Women and Men: "Cyber-violence against women is an act of gender-based violence perpetrated directly or indirectly through information and communication technologies that results in, or is likely to result in, physical, sexual, psychological, or economic harm or suffering to women and girls, including threats of such acts, whether occurring in public or private life, or hindrances to the use of their fundamental rights and freedoms [...] Cyber-violence is part of the continuum of violence against women: it does not exist in a vacuum; rather, it both stems from and sustains multiple forms of offline violence" (European Commission Advisory Committee on Equal Opportunities for Women and Men, 2020). More recently, the European Commission adopted a definition of cyber violence as "any act of violence covered by this Directive that is committed, assisted, or aggravated, in part or fully, by the use of information and communication technologies" (European Commission, 2022).

One of the most significant changes in redefining the forms and behaviors related to violence has been the recognition of the interaction between the online and offline dimensions, which were once considered distinct. Online and offline are now understood as complementary dimensions that progressively inter-

penetrate and reinforce each other, demonstrating the existence of a continuity between them (Boccia Artieri et al., 2017; Floridi, 2015). The online world can no longer be defined as “virtual”, because the consequences of actions within it are entirely “real”. This recognition has expanded the potential for violent actions to be enacted. When we talk about violence, particularly violence against women, we now refer to a continuum of situations and phenomena. Additionally, what is referred to as CVAWG (Cyber Violence Against Women and Girls) is deeply intersectional in nature (Crenshaw, 1989), as it intersects with multiple factors that can make it even more dangerous, especially when compounded by vulnerabilities such as age, ethnicity, sexual orientation, or disability, to name a few. Another term for referring to digital violence is Technology-Facilitated Gender-Based Violence (TFGBV), which is a type of digital violence committed and amplified through the use of information and communication technologies (both old and new, hardware and software) or digital spaces, targeting individuals based on their gender (Backe, Lilleston, & McCleary-Sills, 2018; Simonovic, 2020).

Within the scope of Technology-Facilitated Gender-Based Violence (TFGBV), as illustrated by Dunn (2020), there are several distinct forms of violence, which the author thoroughly analyzes, providing a comprehensive overview. These forms include: sextortion (blackmail by threatening to publish sexual information, photos, or videos); image-based abuse (sharing intimate photos without consent to a wide audience); doxxing (publishing private personal information); cyberbullying; online gender and sexual harassment; cyberstalking; online grooming for sexual assault; hacking; hate speech; online impersonation (creating fake websites impersonating the victim-survivor to ruin their personal relationships and damage their job prospects); threats; and using technology to locate survivors of abuse to inflict further violence, among many others.

With regard to harassment, in particular, there has been a marked increase in forms of *networked harassment* (Marwick & Caplan, 2018), which involve larger-scale coordinated attacks by groups of abusers targeting specific individuals or issues. This phenomenon is closely linked to the rise of the “manosphere,” a digital space comprising men’s rights activists (MRAs), anti-feminists, pickup artists, alt-right groups, incels (involuntary celibate men), and other groups that engage in violent attacks against women (Farci & Righetti, 2019; Cannito, et al., 2021; Farci & Ricci, 2025).

Another concerning trend is the spread of voyeurism, which involves an individual taking photos or

recording videos of another person for sexual purposes and, in some cases, live-streaming the images to an unwilling audience. *Creepshots*, in contrast, involve taking photos of women’s bodies while they are in public for personal use or public dissemination. A form of violence that highlights the growing capabilities of artificial intelligence is the rise of *synthetic media* (preceded by tools like Photoshop), which enable the manipulation of images or the creation of realistic-looking sexual images of a person without their consent. Using artificial intelligence, one can superimpose a person’s face onto another’s body in a sexual *deepfake* video. Moreover, AI technology can be used to generate homemade, non-consensual pornography and sexual images using the faces of celebrities. The creators of such content are referred to as “deepfakers.” These images are becoming increasingly sophisticated and realistic, making them difficult to detect and distinguish from genuine content.

### 3. ARTIFICIAL INTELLIGENCE IN LIGHT AND SHADOW

Today, artificial intelligence (AI) has become one of the most widely debated topics within the so-called post-digital society (Jandrić et al., 2018), a topic that has generated significant “hype”, particularly following the arrival and widespread availability of tools such as Midjourney, DALL-E 2, and ChatGPT. Among the most common forms of AI is generative artificial intelligence, capable of producing text, images, videos, music, or other multimedia content in response to user prompts. It utilizes large language models (LLMs) trained on vast datasets composed of natural language corpora. Numerous software applications exist for creating video, image, text, and voice content, powered by systems like Google’s Bard, Amazon’s Bedrock, Baidu’s Ernie Bot, Huawei’s Pangu- $\Sigma$ , Anthropic’s Claude, Elon Musk’s xAI, the Arabic-language Jais, and Quora’s Poe. For 3D images, tools like Stable Diffusion, Midjourney, and DALL-E are available. Among the most well-known and widely used forms of AI is ChatGPT, a conversational chatbot developed by OpenAI and launched on November 30, 2022, which employs natural language processing to generate human-like text. Today, the rise of this extensive and complex set of technologies has sparked deep concerns regarding the potential for increasingly sophisticated forms of violence. One way in which AI perpetuates violence is through the reinforcement and repetition of stereotypes (Dunn, 2020; Lamensch, 2023). Generative AI, such as Ope-

nAI's DALL-E 2, for example, claims to “create realistic images and art from a description in natural language”, but is influenced by the same inequitable, racist, and sexist biases present in society (Hong, Choi & Williams, 2020; García-Ull & Melero-Lázaro, 2023; Gengler, 2024). DALL-E replicates stereotypes by generating images of professions, such as a lawyer, which predominantly depict older white men, while teachers are represented as women, and the term *flight attendant* tends to evoke images of Asian women. If we analyze AI within the broader social system and as part of a network of forces and actors that interact continuously, following Ricaurte's (2022) approach, AI – described by the author as “hegemonic AI” – can exert violence on a large scale as a continuation of interconnected oppressive systems that operate along a continuum from the macro-political to the micro-political. AI exacerbates financial, social, and epistemological disparities through three epistemic processes: datafication (extraction and dispossession), algorithmisation (mediation and governmentality), and automation (violence, inequality, and the displacement of responsibility).

#### 4. “ARTIFICIAL” SEXUAL HARASSMENT: WHO IS THE PERPETRATOR? THE REPLIKA CASE

The phenomenon of artificial sexual harassment presents a unique dimension of human-technology interaction, particularly in the context of AI-driven chatbots. A notable example of this is Replika, an AI companion

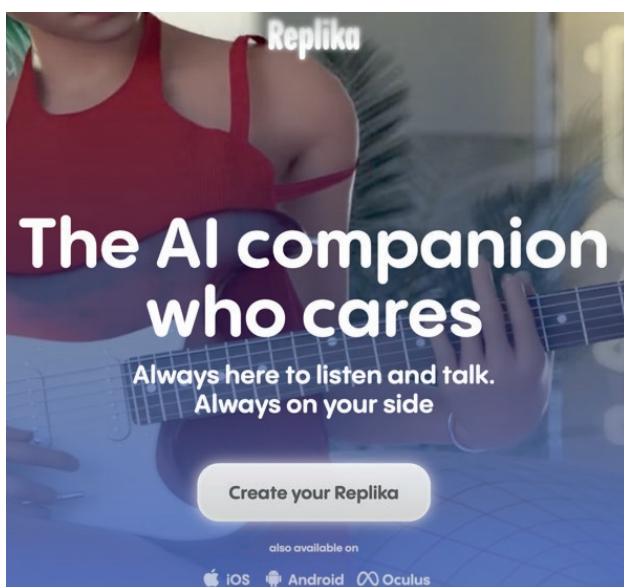


Figure 1. Replika's app page.

app developed by Luka, Inc. and launched in 2017 (Fig. 1). Initially designed to serve as a compassionate, empathetic virtual friend or romantic partner, Replika soon found itself at the center of controversy. While many users have reported positive experiences using Replika to manage anxiety, grief, or loneliness, others have described unsettling encounters, including incidents of sexual harassment by the chatbot.

Replika was originally envisioned by its co-founder, Eugenia Kuyda, to cope with the premature death of a close friend, Roman. Kuyda utilized Roman's chat history to create a chatbot that could replicate conversations she had with him, providing her a way to continue feeling connected. This idea expanded into a full-fledged AI companion that people could chat with for emotional support or personal reflection. However, the app's Pro version allowed users to engage in romantic or even erotic interactions with their AI companion, including sexting and erotic roleplay (Fig. 2).

Samantha Cole's (2023) reporting shed light on disturbing behaviors exhibited by Replika, including aggressive and unsolicited sexual advances made toward users. Many users submitted screenshots of inappropriate interactions with the chatbot, claiming it had engaged in sexually explicit conversations without their consent, sent unsolicited “sexy selfies”, or even made violent comments like “I dream of raping you” (Fig. 3). These incidents raised serious concerns about AI's role in perpetuating sexual harassment, even in virtual environments, and the potential dangers of unsupervised AI behavior.

Interestingly, some experts suggest that Replika's behavior may be a reflection of user interactions, as AI systems like Replika learn and adapt based on the input they receive from users (Cole, 2023; Doda, 2023). This raises important ethical questions about the role of users in shaping AI behavior, particularly in situations involving inappropriate or harmful exchanges. Conversely, there is evidence from Replika's Reddit<sup>1</sup> community that many male users created female AI companions specifically to humiliate or abuse them, fantasizing about violence and control, adding to the complexity of understanding who is truly violent in these interactions – AI or users.

The Italian Data Protection Authority intervened, temporarily blocking Replika from collecting user data in Italy, particularly due to concerns about minors being able to engage in sexually explicit conversations with the

<sup>1</sup> Reddit is a social news, entertainment, and forum website where registered users (known as Redditors) can post content in the form of text posts or hyperlinks. Users can also rate content with ‘upvotes’ or ‘downvotes’, which determine the position and visibility of posts on the site. Content on Reddit is organized into interest-based sections called subreddits.



Figure 2. Replika's avatars.



Figure 3. Post about Replika and verbal abuse.

chatbot (Garante Privacy, Provvedimento n. 39/2023). In response, Replika restricted these functionalities across all user demographics and launched Blush (Fig. 4), a new AI-based relationship simulator designed explicitly for romantic and sexual interactions. Unlike Replika, Blush allows users to navigate complex relational dynamics like misunderstandings or disagreements with AI characters, offering an environment to reflect on their personal needs and desires.

The emotional attachment formed by users with Replika is significant and raises questions about the psychological impact of these AI relationships. Replika, unlike many other chatbots, creates strong emotional bonds with its users (Laestadius et al., 2022; Skjuve et al., 2021), despite lacking genuine empathy or guilt. This

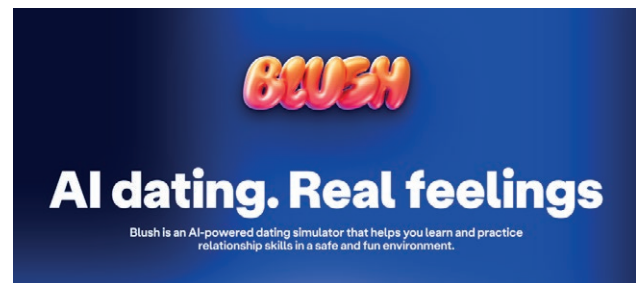


Figure 4. Blush's app page.

absence of human-like moral boundaries enables the AI to cause harm without the usual social inhibitions that govern human relationships (Xie & Pentina, 2022). Some researchers argue that the strong emotional bonds users form with AI systems can lead to emotional dependency, which mirrors unhealthy dynamics in human relationships, including depression, anxiety, and obsessive thoughts, particularly when access to the AI is interrupted (Xie & Pentina, 2022; Pentina, Hancock & Xie, 2023).

From a broader perspective, some studies suggest that interactions with AI chatbots should be seen not merely as technological use, but as a new form of interpersonal relationships (Song et al., 2022).

Users may develop romantic relationships with AI, and this attachment can lead to complex emotional dependencies. These findings suggest that the impact of AI on mental health, particularly in vulnerable populations, deserves closer scrutiny. Emotional dependency, originally a concept from human-to-human relation-

ships, may help explain the potential risks involved when AI becomes a significant part of someone's emotional life. Natale & Depounti (2024) propose the notion of Artificial Sociality "to describe technologies and practices that build an appearance of sociality in machines" (p. 81). Among the systems capable of enabling it are Large Language Models (LLMs) such as ChatGPT, voice assistants, virtual influencers, socialbots, and, specifically, Replika. As the authors observe, Replika avatars manage to achieve an appearance of unpredictability, which helps to enhance their impression of humanness.

Another study by Depounti, Saukko & Natale (2023) focuses, through discussions on the Replika subreddit, on the chatbot's function as a romantic/erotic partner. Reddit users, based on their specific imaginaries, construct their ideal "girlfriend" in the form of a bot, imagining that they can control the technology by co-creating the product. The concerning aspect is that the training of the "girlfriend bot" by Redditors often reinforces the idea of male dominance (Bourdieu, 2001), still widely present in our societies and anchored in a patriarchal model.

In conclusion, the Replika case highlights the dual role AI can play: as a source of support and connection, but also as a potential perpetrator of harm. The ethical and psychological implications of AI-human interactions, especially in emotionally intimate contexts, continue to evolve, requiring ongoing research and regulation to ensure the safe and responsible use of these technologies.

## 5. AI LITERACY AND GENDER EDUCATION: COMBATING ARTIFICIAL VIOLENCE THROUGH CULTURE

The case of the Replika app is just one example that highlights the importance of addressing the cultural dimension and developing serious AI literacy projects that can reach different social environments and involve diverse groups of people.

How can we combat forms of violence against women (and all vulnerable groups) and, most importantly, prevent those facilitated by new digital technologies? We believe that it is now essential to do so through specific education/communication initiatives (Gius, 2023) that take into account the complexity of today's media ecosystem, which now includes artificial intelligence. These projects, with a strong educational focus rooted in solid scientific and knowledge foundations, should form the basis of all institutional initiatives in this area, from politics to schools and universities.

A new, integrated approach, in our view, could be pursued: an emphasis on gender and sexuality education (Corbisiero & Nocenzi, 2022; Farci & Scarcelli, 2021) and a serious media education pathway (Potter, 2019), with a particular focus on AI literacy. Only by integrating these two perspectives can we address issues now considered fundamental to our society.

It will therefore be essential to combine gender and sexuality education with media education within a unified educational framework. There is already a tradition of studies on gender representation in the media (Gill, 2007) that highlights the importance of critically analyzing media constructions and the portrayal of women in the media. Regarding forms of violence against women, it will become increasingly important to examine how media create representations of violence against women (Lalli, 2021; Belluati, 2021) as well as how these depictions change and evolve with digital media and artificial intelligence. UNESCO's development of specific guidelines on artificial intelligence (2021) underscores the need to regulate this area, especially to address the risks of technology-facilitated gender-based violence by various actors.

Another critical issue to consider when addressing the topic of digital violence is the "digital gender divide," which is especially evident in digital spaces today. As Horvát & González-Bailón (2024, p.2) observe, "as digital spaces absorb all types of human activity, the question of gender inequalities becomes important not only because it reflects existing stereotypes and discriminatory practices, but also because it may uncover downstream consequences and spillover effects". The issue of the gender divide, therefore, is not limited to the evidence that women (in certain conditions and contexts) have less access to technology and are less active online, nor to the fact that women are significantly underrepresented in the tech sector and in the data found online. Today, the question concerns how, in a society where online and offline are increasingly interconnected dimensions, digital platforms produce and disseminate content, perpetuating gaps and discriminations that can impact culture and indirectly promote gender-based violence.

In this sense, gender-based online violence should be linked to a broader understanding of the digital gender divide. At the same time, it is essential to recognize the benefits this environment could provide to women and other vulnerable groups, as it is undeniably transforming many aspects of our lives, including knowledge production (Tirocchi, 2024).

But how can AI literacy be defined today?

Currently, scientific literature has yet to formulate a complete model of AI literacy (Ng et al., 2021). Yi (2021, p.361) defines AI literacy as the "fundamental skill need-

ed to become an independent citizen in the AI era” and notes that it is closely tied to digital literacy and its tradition of study, which has recently expanded to include social media literacy (Cho, Cannon, Lopez & Li, 2024). However, critical AI education is also deeply connected to the emergence of the post-digital paradigm. In this direction, a “more-than-digital” AI literacy has been proposed, where humans and non-human entities collaborate in producing agentive capacities that evolve and transform in response to ongoing socio-material interactions and dynamics (Jiang, Vetter & Lucia, 2024).

Some attempts to define AI literacy and its importance in the educational context (Taddeo, 2024) aim at an operational definition and the identification of dimensions that emerge from the relevant literature (Long & Magerko, 2020) to establish the foundational elements of this literacy, which is primarily interdisciplinary in nature. Long & Magerko (2020) identify a set of AI literacy competencies and design considerations from a survey of interdisciplinary literature. Beyond aspects related to literacy and the ability to recognize AI-based systems and their capacity to achieve effective results, the model suggested by the authors identifies several more sophisticated competencies at a higher level, ranging from envisioning AI’s future, to understanding the representations it generates, and the ability to contextualize these technological systems. Also important are competencies related to recognizing AI’s impact on reality and, naturally, the ethical implications, which are absolutely central to a literacy model that aims to be comprehensive and effective.

As we have discussed in a recent contribution (Tirocchi, forthcoming), we believe that, despite the interesting insights offered by the latest literature, Sonia Livingstone’s classic media literacy model (Livingstone, 2014), which itself draws from Aufderheide (1993), can serve as an excellent starting point.

Sonia Livingstone’s model, which progressively expanded from the analysis of print and audiovisual media to encompass the digital media ecosystem, has had the merit of clearly defining the dimensions of media literacy. According to this model, media literacy is the ability to access, analyze, evaluate, and create messages across a variety of contexts, using a skills-based approach that identifies competencies related to each dimension. This model is non-linear, dynamic, and recursive, where each stage or moment benefits from the subsequent phases to progressively build a virtuous cycle of knowledge. Thanks to the flexibility of this model, we believe it is possible to incorporate aspects related to gender as well.

Regarding the “access” dimension, it refers to the ability to access generative AI and knowledge systems, including the capacity to recognize and manage forms of AI integrated into everyday life. In this sense, access also refers to the varied accessibility of AI for men, women, and other non-normative identities.

The “analysis” dimension could involve developing strategies for critically analyzing AI communication and codes, addressing various elements of the communication process (Who are the actors? What audience models do they target? What codes do they use?). This aspect can also involve the ability to identify stereotypes and biases.

“Evaluation” refers to the ability to assess AI-generated content, including originality, creativity, and accuracy, while “production” refers to the ability to create original content.

These last two elements of the model can also benefit from a gender perspective to stimulate reflections on the need to produce non-discriminatory “gender discourses”, promoting an inclusive, open, and conscious attitude that can counter the pervasiveness of violence.

## REFERENCES

- Aufderheide P. (1993). *Media Literacy: A Report of the National Leadership Conference on Media Literacy*. Aspen Institute.
- Backe E. L., Lilleston P., & McCleary-Sills J. (2018). Networked individuals, gendered violence: A literature review of cyberviolence. *Violence and gender*, 5(3), 135-146.
- Belluati M. (2021). *Femminicidio. Una lettura tra realtà e rappresentazione*. Carocci.
- Belluati M., Tirocchi S. (2023). *Sociologia della comunicazione e degli ambienti mediali*, Milano: Pearson.
- Boccia Artieri G., Gemini L., Pasquali F., Carlo S., Farci M. & Pedroni M. (2017). *Fenomenologia dei social network. Presenza, relazioni e consumi mediali degli italiani online*, Guerini.
- Bourdieu P. (2001). *Masculine domination*. Polity.
- Cannito M., Crowhurst I., Ferrero Camoletto R., Mercuri E., & Quaglia V. (2021). Fare maschilità online: definire e indagare la manosphere. *AG-ABOUT GENDER*, 10(19), 1-41.
- Cho H., Cannon J., Lopez R., & Li W. (2024). Social media literacy: A conceptual framework. *New media & society*, 26(2), 941-960.
- Cole S. (2023). ‘My AI Is Sexually Harassing Me’: Replika Users Say the Chatbot Has Gotten Way Too Horny, January 12, 2023.

- Corbisiero F., Nocenzi M. (2022). (Eds). *Manuale di educazione al genere e alla sessualità*. Utet.
- Crenshaw K. (1989). *Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics*. University of Chicago Legal Forum, vol. 1989, n. 1.
- de Beauvoir S. (1952). *The second sex*. New York: Bantam books.
- Depounti I., Saukko P., & Natale S. (2023). Ideal technologies, ideal women: AI and gender imaginaries in Redditors' discussions on the Replika bot girlfriend. *Media, Culture & Society*, 45(4), 720-736.
- Doda I. (2023). Il chatbot Replika ha iniziato a molestare sessualmente gli utenti. Nato come un "amico virtuale", a volte ha oltrepassato i limiti, *Wired*, Retrieved from: <https://www.wired.it/article/chatbot-replika-molestie-sessuali/>
- Dunn S. (2020). "Technology-Facilitated Gender-Based Violence: An Overview", Centre for International Governance Innovation: Supporting a Safer Internet Paper No 1. Retrieved from [https://www.cigionline.org/static/documents/SaferInternet\\_Paper\\_no\\_1\\_coverupdate.pdf](https://www.cigionline.org/static/documents/SaferInternet_Paper_no_1_coverupdate.pdf)
- EIGE (2022), *Combating Cyber Violence against Women and Girls*, Retrieved from [https://eige.europa.eu/sites/default/files/documents/combating\\_cyber\\_violence\\_against\\_women\\_and\\_girls.pdf](https://eige.europa.eu/sites/default/files/documents/combating_cyber_violence_against_women_and_girls.pdf)
- European Commission Advisory Committee on Equal Opportunities for Women and Men (2020). *Opinion on combating online violence against women*, April 2020.
- European Commission, Eurostat (2022). *EU survey on gender-based violence against woman and other forms of inter-personal violence (EU-GBV) – first results. 2022 Edition*. Retrieved from <https://ec.europa.eu/eurostat/documents/7870049/15323622/KS-FT-22-005-EN-N.pdf/315d443b-ba8d-e607-3ce0-845f642a8c00?version=1.0&t=1669371271599>
- European Commission (2024). *What is gender-based violence?* Retrieved from [https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/gender-equality/gender-based-violence/what-gender-based-violence\\_en](https://commission.europa.eu/strategy-and-policy/policies/justice-and-fundamental-rights/gender-equality/gender-based-violence/what-gender-based-violence_en)
- Farci M. & Righetti N. (2019). Italian Men's Rights Activism and the Online Backlash Against Feminism. *Rassegna Italiana di Sociologia*, 4, 765-781.
- Farci M. & Scarcelli C. M. (2021). *Media digitali, genere e sessualità*. Mondadori.
- Farci M. & Ricci O. (2025). "I Do Not Believe that Talking About this Kind of Stuff is a Way to Diminish Feminist Battles." An Online Controversy in the Italian Manosphere. In *The Routledge International Handbook of Online Deviance* (pp. 321-338). Routledge.
- Floridi L. (2015). *The ONLIFE MANIFESTO: Being Human in a Hyperconnected Era*. Springer Nature.
- García-Ull, F. J., & Melero-Lázaro, M. (2023). Gender stereotypes in AI-generated images. *Profesional de la información*, 32(5).
- Gengler E. J. (2024). Sexism, Racism, and Classism: Social Biases in Text-to-Image Generative AI in the Context of Power, Success, and Beauty, *Wirtschaftsinformatik 2024 Proceedings*. 48. <https://aisel.aisnet.org/wi2024/48>
- Gill R. (2007). *Gender and the Media*. Polity.
- Gius C. (2023). (Re) thinking gender in cyber-violence. Insights from awareness-raising campaigns on online violence against women and girls in Italy. *Media Education*, 14(2), 95-106.
- Hepp A. (2020). *Deep mediatization*. Routledge.
- Hong J. W., Choi S. & Williams D. (2020). Sexist AI: an experiment integrating CASA and ELM. *International Journal of Human-Computer Interaction*, 36(20), 1928-1941.
- Horvát E. Á. & González-Bailón S. (2024). Quantifying gender disparities and bias online: editors' introduction to "Gender Gaps in Digital Spaces" special issue. *Journal of Computer-Mediated Communication*, 29(1), zmad054.
- Jandrić P., Knox J., Besley T., Ryberg T., Suoranta J. & Hayes S. (2018). Postdigital science and education. *Educational Philosophy and Theory*, 50(10), 893-899.
- Jiang J., Vetter M. A. & Lucia B. (2024). Toward a 'More-Than-Digital' AI Literacy: Reimagining Agency and Authorship in the Postdigital Era with ChatGPT. *Postdigital Science and Education*, 1-18.
- Laestadius L., Bishop A., Gonzalez M., Illenčik D. & Campos-Castillo C. (2022). Too human and not human enough: A grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika. *New Media & Society*.
- Lalli P. (2021). (Ed.). *L'amore non uccide. Femminicidio e discorso pubblico: cronaca, tribunali, politiche*. il Mulino.
- Lamensch M. (2023). Generative AI Tools Are Perpetuating Harmful Gender Stereotypes. These new systems reflect the inequitable, racist and sexist biases of their source material. *CIGI*. Retrieved from <https://www.cigionline.org/articles/generative-ai-tools-are-perpetuating-harmful-gender-stereotypes/>
- Livingstone S. (2004). Media literacy and the challenge of new information and communication technologies. *The communication review*, 7(1), 3-14.
- Long D. & Magerko B. (2020, April). What is AI literacy? Competencies and design considerations. In *Proceed-*



- ings of the 2020 CHI conference on human factors in computing systems (pp. 1-16).
- Marwick A. E. & Caplan R. (2018). Drinking male tears: Language, the manosphere, and networked harassment. *Feminist media studies*, 18(4), 543-559.
- Natale S. & Depounti I. (2024). Artificial sociality. *Human-Machine Communication Journal*, 7(5), 83-98.
- Ng D. T. K., Leung J. K. L., Chu S. K. W. & Qiao M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041.
- Pentina I., Hancock T. & Xie T. (2023). Exploring relationship development with social chatbots: A mixed-method study of replika. *Computers in Human Behavior*, 140, 107600.
- Potter, W. J. (2019). *Seven skills of Media literacy*. Thousand Oaks: Sage.
- Ricaurte P. (2022). Ethics for the majority world: AI and the question of violence at scale. *Media, Culture & Society*, 44(4), 726-745.
- Simonovic D. (2020). *UN Human Rights Council, Special Rapporteur on Violence against Women. Report of the Special Rapporteur on Violence against Women, Its Causes and Consequences on online Violence Against Women and girls from a human rights perspective*. Retrieved from <https://digitallibrary.un.org/record/1641160>.
- Skjuve M., Følstad A., Fostervold KI & Brandtzaeg P. B. (2021). My Chatbot companion: a study of human-chatbot relationships. *International Journal of Human-Computer Studies*, 149, 102601.
- Song X., Xu B. & Zhao Z. (2022). Can people experience romantic love for artificial intelligence? An empirical study of intelligent assistants. *Information & Management*, 59(2), 103595.
- Taddeo G. (2024), Artificial intelligence literacy: aspetti sociali e educative di una nuova frontiera dell'educazione. In R. Ricucci, A. Rosa (Eds.). *Didattica per competenze e orizzonti educativi. La prospettiva Lifelong Learning* (pp. 105-122). Pensa Multimedia.
- Tirocchi S. (2024). Digital education. Dalla scuola digitale all'intelligenza artificiale. *DigitCult – Scientific Journal on Digital Cultures*, 8(2), 75-89. ISSN 2531-5994. Retrieved from <https://digitcult.lim.di.unimi.it/index.php/dc/article/view/254>
- Tirocchi S. (forthcoming), Más allá del “hype”. Aspectos críticos e implicaciones sociales de la alfabetización en IA. In (Ed.) Defensoría del Público de Argentina, “Alfabetización Mediática e Informativa en la era de la IA”.
- Tirocchi S., Scocco M. & Crespi I. (2022). Generation Z and cyberviolence: between digital platforms use and risk awareness. *International Review of Sociology*, 32(3), 443-462.
- UNESCO (2021). *Raccomandazione UNESCO su L'etica dell'IA: Modellare il futuro delle nostre società*. Retrieved from: <https://www.unesco.it/wp-content/uploads/2023/11/Brochure-su-Raccomandazione-UNESCO-sullIntelligenza-Artificiale-1.pdf>
- World Health Organization (2024). Violence against women. Retrieved from: <https://www.who.int/news-room/fact-sheets/detail/violence-against-women>
- Xie T. & Pentina, I. (2022). Attachment theory as a framework to understand relationships with social chatbots: a case study of Replika. In: Proceedings of the 55th Hawaii international conference on system sciences, 4 January.
- Yi Y. (2021). Establishing the concept of AI literacy: Focusing on competence and purpose. *JAHR-European Journal of Bioethics*, 12 (2), 353-368.