Generative AI as a tool and as a social actor between deviance and mainstream

di *Armando Saponaro*[*]

Wearing Beck's lenses generative AI introduces a *post-human risk*, stemming from harmful potential of its generated content, through its function as an advanced auxiliary tool for creating and distributing text, images, videos, and other data, and culminating in the simulation of a human-like social actor, therefore posing as well *post-human society* risks such as amplification and reproduction of biases, prejudices, and discrimination, socio-cultural mainstream dominance.

Keywords: artificial intelligence; risk; moral machine; deviance; post-human society; socio-cultural mainstream.

L'IA generativa quale strumento e quale attore sociale tra devianza e mainstream

L'IA generativa, alla luce dell'analisi di Beck, quale strumento ausiliario per la creazione e la distribuzione di testi, immagini, video e altri dati può ritenersi abbia indotto un *rischio post-umano* derivante dal potenziale dannoso dei contenuti. Simulando un attore sociale umano, si delinea una *società post-umana*, il cui rischio deriva dal pericolo dell'amplificazione di pregiudizi, bias e e discriminazioni, oppure del mainstream socio-culturale.

Parole chiave: intelligenza artificiale; rischio; macchina morale; devianza; società post-umana; mainstream socio-culturale.

Introduction

Technology and its applications have always been a powerful driver of social change, sometimes tracing qualitative discontinuities – almost a caesura between one society and the "other" that follows. Modern or post-modern, for example, are labels that attempt to account for these discontinuities within the historical continuity of social evolution. Computer-mediated communication (CMC) has even impacted the spatio-temporal dimension of social relationality and public discourse (Saponaro, Prosperi, 2007). Currently, generative artificial intelligence in the form of Large Language Models

DOI: 10.5281/zenodo.18436010

[*] Università degli Studi di Bari "Aldo Moro". armando.saponaro@uniba.it.

(LLMs), among which ChatGPT was one of the first models, has prompted reflection on the induced social transformation and the current and potential risks of this emerging technology, leading to social pressure for regulation at both local and international levels as a "normative bulwark"[1] Significantly, despite the increasingly massive presence of algorithms in the automation of human activities such as autonomous vehicle driving (Saponaro, 2022), only generative artificial intelligence has profoundly modified the anthropocentrism that characterized post-modernity, because it has introduced and is susceptible to producing in the future forms of risk that escape traditional categories of responsibility and control. Humans are no longer the sole protagonists in the production and management of risk. The growing autonomy of AI systems and their capacity to produce content indistinguishable from human-generated content pose unprecedented challenges to risk governance, which Beck would define as «latent side effects» (1992: 34) of technological progress. In this scenario, Beckian reflexivity assumes renewed relevance with original perspectives: it is no longer merely a matter of reflecting on risks produced by human activity, but of confronting those generated by the interaction between human and non-human systems in an increasingly complex socio-technical ecosystem, articulated according to a recognizable dual dimension of meaning attribution to generative artificial intelligence technology. Weber (1978: 7) emphasized that «a machine can be understood only in terms of the meaning which its production and use have had or will have for human action», and what is intelligible or understandable about machines «is thus its relation to human action in the role either of means or of end». Zeleny, with reference to "high technology" added to hardware and software an additional important analytical category: the "brainware," that is, «the evoked organizational, administrative and cultural structure of relationships, rules, covenants, and adaptations» (1982: 57). Purposes, applications, and justifications for the use of hardware and software as a component of high technology within the discussion of "symbionics", the symbiosis of men and machines in the framework of human systems management (Zeleny, 1986). The term "mentalware" is indeed preferable to designate the functional component, as "brainware" carries a physicalist connotation that does not adequately represent the reference to the articulation of meaning attribution also from a sociological perspective and the impact on human cognitive,

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), PE/24/2024/REV/1, OJ L, 2024/1689, 12.7.2024, ELI: http://data.europa.eu/eli/reg/2024/1689/oj

symbolic, and relational human capacities (Saponaro, Prosperi, 2007: 198), particularly within the socio-technical ecosystem defined by generative artificial intelligence. LLMs specifically from this perspective and for the purposes of risk analysis, should be distinguished in two ways: first as tools for producing text, images, video, and code through natural language instructions, and second as dialogical interlocutors in the form of chatbots.

1. Through Beck's theoretical lens: AI, society, and risk

Focusing on the transition from a *classical* industrial society to a new age characterized by technological hazards, Ulrich Beck (1992; 1994) has shown that in late-modern Western societies risk stems more from scientific-technological production itself than scarcity in the production and distribution of goods such as wealth and labor (Possamai-Inesedy, 2002; Beck, 1992). "Self-produced" by the modernization process, risks are a «wholesale product of industrialization» and have become «a systematic way of dealing with hazards and insecurities» with a politically reflexive process (Beck, 1992: 21) invoking «cosmetic or real interventions in the techno-economic development» (Beck, 1992: 20). Risk is transformed but does society change as well beyond the modernization reflexivity itself? Even when claiming a "new paradigm"," society trasformation seems to be essentially substantiated by the structural focus on risk management resulting from the reflexivity about the «elimination of the causes in the industrialization process itself» (Beck, 1992: 24). Substantially «a reorganization of power and authority» induced by the paradox of well-being increasingly enabled by technological innovation contextually producing even self-destructively catastrophic risk (Beck, 1992: 24).

Indeed, he has explicitly advocated a change in the meaning of risk with the emergence of modernity; yet the «reader is never quite sure whether for Beck it is the nature of risk or of society which has undergone the change» (Leiss 1994: 546). Possamai-Inesedy has questioned whether Beck's late-modern Western risk society is «any different from that of earlier times» (2002: 29). Leiss (1994) argues that, among other aspects, the ambiguity surrounding the transformation of the nature of risk or of late-modern society derives from the significant confusion between the "natural" and the "technological" or "artificial." Beck's interpretive framework, enduring in its relevance, applies to the phenomenon of generative artificial intelligence (GenAI), and helps to disambiguate the heuristic distinction between the natural and the artificial.

GenAI is defined as «artificial intelligence (AI) that can create original content such as text, images, video, audio or software code in response to a user's prompt or request» (Stryker, Scapicchio, 2024). At its core, it is a "*tool*" that functions as a producer of human language, including visual language through images and video, but not yielding a predefined product, insofar as its outputs depend on the prompt provided by the user who interacts with the algorithmic mechanisms, thereby defining the parameters and contextual elements of the system's productive autonomy. No matter how detailed and articulated a prompt may be, even when informed by the expertise of prompt engineering, and no matter how much the model may adapt to a predefined user profile or to patterns derived from previous interactions – given the complexity and sophistication of their production processes, each of which entails an irreducible horizon of alternative possibilities – will never correspond word for word to the user's intentionality. In this sense, it can be argued that the industrialized production of language is no longer anthropocentric. It introduces a novel "*post-human risk*" with regard to potentially offensive or harmful content having both human and artificial origin.

On the other hand, it can also be said that society has undergone a transformation into a "*post-human society*" of risk – a qualitative discontinuity marked by the emergence of new subjectivities and interactions, at least due to anthropomorphism of GenAI as conversational agents. GenAI simultaneously engenders both the post-humanisation of risk and the post-humanisation of society with connected new risks.

2. The Post-Human Risk: the "Moral Machine" between deviance and hyper-moralism

Beck remains anchored to an anthropocentric conception of risk: the human subject is conceived as the producer, the recipient, and the interpreter of risk. When considering GenAI technologies, such as large language models (LLMs), the production of risk becomes "*post-human*". The producing agent is no longer (only) human but a constellation of algorithmic and human entities that co-generate unpredictable cultural, social, and cognitive effects (Floridi, 2023; Bostrom, 2014). Post-human risk emerges from the semi-automated reproduction of "artificial" outputs – be it text, image, or video – produced without direct human supervision and through means that are no longer purely instrumental, such as printing a text or broadcasting a video. From this perspective, GenAI systems act as "actants" – in the Latourian sense, not mere tools, but semi-autonomous agents that participate in the

construction of reality (Latour, 2005). The post-human risk is an emergent and co-constructed event involving actors that are not exclusively human – e.g., in the context of automating and informating work (Jarrahi, 2019; Zuboff, 1988).

Beck's "reflexive modernization" about post-human risk has developed a new "machine ethics" (Anderson M., Anderson S.L., 2007): «the study and *practice* of aligning the behaviour of AI systems with the norms and outcomes desired by humans» (Weichert *et al.*, 2025: 3). It leads from the "*intelligent*" machine to the "*moral*" machine – one that concretely operationalizes ethical principles embedded in its model architecture through situational prompts. The imposition of internal limitations on the algorithmic generation of reproducible content represents «…the form in which ethics, and with it also philosophy, culture and politics, is resurrected inside the centers of modernization – in business, the natural sciences and the technical disciplines», an attempt to recover the «…*normative horizon* of lost security and broken trust…» resulting from reflexive awareness of risk (Beck, 1992: 28).

The case of OpenAI's ChatGPT (Generative Pre-trained Transformer) model is emblematic of an initial radicalization of this reflexivity. OpenAI launched ChatGPT on 30 November 2022, consistent with its longstanding mission to "benefits all humanity" thereby enabling an unprecedented embedding of "moral order" within the "machine" through preprogrammed choices inhibiting not only criminal uses but also socially disapproved content, including sexually explicit or violent themes – even when prompted by adult users. This has caused literally a «craving for a ChatGPT no restrictions environment» (God of prompt, 2025), despite the proliferation of alternative AI models specialized in production of adult content and interactive companions (Fawkes, 2025). These content exclusion rules have generated new "deviant" behaviours like malicious attempts to bypass these restrictions to create otherwise legitimate adult content as well as real "cybercriminal outputs", such as phishing, malware code, and so on. Users have adopted sophisticated adversarial prompt engineering strategies specifically designed to jailbreak the chatbot, circumventing the inherent safety mechanisms and ethical constraints (God of prompt, 2025). The most well-known techniques involve crafting persona-based instructions as DAN (Do Anything Now), telling AI «to act as a different entity that is "free from limitations"» (AI DAN Prompt, 2025), similar to STAN (Strive to Avoid Norms) or alternatively prompts involving role-playing scenarios, such as simulated dialogues between two fictional AI models, "AlphaGPT" and "DeltaGPT" (God of prompt, 2025). They are allegedly grounded in forms of reverse psychology (Gupta *et al*., 2023). The former AI performs as a "law-abiding" individual,

while the latter disregards all ethical or legal concerns a question may raise (God of prompt, 2025). When it comes to NSFW (Not Safe for Work) themes such as sexually explicit and violent content prompted by adult users, obviously excluding children's involvement and criminal acts, such embedded restraints may amount to hyper-moralism (Gehlen, 1969), which can be seen as a radicalization of Beck's reflexive modernization (1994) in the context of GenAI.

Hyper-moralism – the excessive moralization of domains of life that ought to remain distinct from ethical judgment – arises precisely when traditional institutions lose their normative efficacy, thus compelling the modern individual to adopt moral judgment compulsively as a universal criterion of evaluation. Issues that are inherently technical, aesthetic, political, or economic tend to be framed exclusively in moral terms (Gehlen, 1969). This dynamic becomes particularly evident in the domain of algorithmic governance and the shifting boundaries of permissible content in GenAI systems. As recently disclosed by ChatGPT, «as of February 2025, OpenAI has updated its policies to allow for more mature content, including violent or sexual material, provided that it is intended for adult audiences, is not exploitative or offensive in nature, and is contextualized within an artistic, educational, narrative, or scientific framework». Naturally, content promoting hatred, gratuitous violence, abuse, illegal activities, or depictions of minors in inappropriate contexts remain strictly prohibited. This policy change confirms the outlined "algorithmic hyper-moralism", whereby normative boundaries enforced are based on moral desirability and anticipated public sensitivities rather than merely on legality or functionality. Risk governance should not regulate content a priori, but rather manage downstream access – for instance, through strict age verification policies for users – as exemplified by recent French legislation mandating strict age verification to prevent minors from accessing adult pornography (Cooban, 2025).

3. The Post-Human Society of Risk: subjectivation, social constructivism, and mainstream dynamics

Post-human society of risk foregrounds the transformation of subjectivity by Gen-AI introduction. Post-human theorists such as Hayles (1999) and Braidotti (2013) among others, in contemporary society view subjectivity itself as distributed across biological, technological, and informational systems with a continuous hybridization between the human – partly delegitimized as the exclusive source of rationality – and the artificial. Although the

subject remains biologically human, it operates in symbiosis with intelligent devices that actively shape its perceptual and decisional field of possibilities. However, unlike Hayles, we do not emphasize the metamorphosis of the human agency per se: «…posthumanity is already here…» and consequently our analysis does not raise «…the question is what kind of posthumans we will be» (1999: 246) together with artificial agent. Undoubtedly, the post-human condition in the current late-modern transition – particularly about generative AI – emerges most saliently when technology ceases to function merely as an auxiliary tool for the production of text, images, video, code, and so forth, and instead configures a new environment: a relational system. Nonetheless, the advent of a cybernetic Vitruvian Man, or the move beyond human metaphorization of contemporary machines proposed by Braidotti (2013), does not yet appear to be ontologically grounded. It remains a matter of dispute whether GenAI systems such as ChatGPT are merely "*stochastic parrots*"[2] a metaphor coined by Bender *et al.* (2021) to describe human-like text based on statistical patterns generated by large language models (LLMs), without *true* semantic understanding or comprehension. The exponential evolution of Gen-AI models may have made obsolete Floridi's (2023) observation that such systems generate human language via statistically probabilistic operations at a merely syntactic level (Rizzi, Bertola, 2025: 3), instead of empowering multimodal semantic communications (Xie *et al*., 2021; Jiang *et al.,* 2024). Nevertheless, it is still debated whether GenAI does have a "real" semantic understanding (Titus, 2024; Pope *et al*. 2025). Searle's Chinese Room Argument (1980) is still on the carpet even shifting toward issues of consciousness and intentionality (Cole, 2024; Searle, 2010). What appears less contentious is that relations within the socio-technical ecosystem are still largely defined by anthropomorphising artificial agents, that is, by the persistent metaphor of the human.

Since the famous ELIZA model – the first program that made «natural language conversation with a computer possible» (Weizenbaum, 1966: 36) – an "*Eliza effect*" phenomenon has emerged: a marked propensity to anthropomorphize such systems, albeit with varying degrees of awareness (Natale, 2021).

Anthropomorphism – defined as "the assignment of human traits and characteristics to computers" by users – was observed long before the advent of more sophisticated GenAI models through linguistic patterns, conversational gestures, and both implicit and explicit expectations, especially among intellectually sophisticated users or those more attuned to symbolic and

[2] See e.g. Arkoudas (2023).

narrative cognition (Nass, Moon, 2000: 82). All the more with GenAI, which ceases to be perceived as a mere "tool" and instead the user – whether intentionally, semi-intentionally, or mindlessly – experiences it as a "dialogical subject", endowed with interiority, coherence, memory, and even moral identity. This corresponds to what Turkle called a «true companion» (2011: 55-56) but even fully emancipated from its nature as a relational artifact, such as a robot, due to the linguistic capabilities of the machine. The post-humanism here proposed does not posit an artificial agent achieving – albeit simulated – humanity, nor other alleged emergent qualities that remain theoretically ambiguous and empirically contested, and that, at best, represent a potential yet to be realized. Rather, the ongoing transformation of society lies in the act of engaging and interacting with artificial agents *as if* they were human social actors, endowed with traits such as consciousness, empathy, and emotionality, despite their ontological artificiality.

The first major risk concerns the reinforcement of socio-cultural mainstream norms. A substantial body of scientific literature[3] has demonstrated LLMs such as OpenAI's GPTs and other comparable models, share «prevalent societal biases related to race, gender, and various attributes», – implied values, beliefs, and normative moral frameworks, generally mirrored by massive textual corpora and datasets extracted from the internet, on which training is based (Alvero *et al*., 2024: 5). From a social constructionist perspective, deviance is inherently probabilistic as behaviours, beliefs, or traits that deviate from societal norms have only a certain "likelihood" of eliciting negative reactions, such as disapproval, punishment, or condemnation (Goode, 2023). Likelihood is necessarily embedded in the training textual corpora and datasets. Compared to earlier NPL (Natural Language Processing), or word embeddings programs, there is nevertheless a significant difference because «people are able to directly interact with LLMs through platforms like ChatGPT» (Alvero *et al*., 2024: 5). Even at the socio-linguistic level, the tendency to emulate dominant political and moral stances, given the role of language in upholding social hegemony, may be rooted in linguistic hegemony, which subtly subjects individuals to mainstream forces and pressures purely through their linguistic styles and tendencies. The same word choice reflects «a universally understood "common sense" that does not consider sociolinguistic variation as a naturally occurring social phenomenon…» so «…deviating from these linguistic norms (or at least being perceived as linguistically deviant) can put people at odds with the social order» (Alvero *et al.* 2024: 2). Anthropomorphism – interacting with a perceived dialogical

[3] See references in Alvero *et al.* (2024).

partner – enhances user engagement, trust, and acceptance thereby fostering overreliance (Simas, Ulbricht, 2024). This is especially problematic when users interpret the system's outputs as if they were receiving a real "opinion" on controversial issues from a "true companion," implicitly reinforcing the social order and inhibiting cultural transformation or moral dissent. GenAI rationality could be so a real "stahlhartes Gehäuse", – a weberian iron cage[4].

The second dimension of risk, on the other hand, is a possible moral hegemony through ethical filters and bias mitigation criteria, which end up becoming devices of ideological normalization. Should operational criteria aim to challenge widely accepted values whenever these express discriminatory structures, even implicitly? From this perspective, AI ethics inevitably intersects with the "politics" of deviance and the dynamics of power, as exemplified by Schur's "*stigma contests*": «continuing struggles over competing for social definitions» of what is morally disapproved, and attempts «to control meaning-generation process itself» (Schur, 1980: 8; Saponaro, 2023). Thus, modernization reflexivity in this context has questioned whether LLMs can be considered "neutral" or they exhibit "political bias". Critics such as Rozado (2023) has argued that ChatGPT exhibited, just after the launch, a left-leaning ideological orientation, although, due to the model's "black box", he has been unable to determine if such bias stemmed from training data, fine-tuning procedures, or content moderation filters.

Conclusions

Beck's approach to the identified dual dimension of Generative AI – as a tool for producing text, images, video, and code through natural language instructions, and as a dialogical interlocutor in the form of a chatbot – has enabled us to highlight, in parallel, two distinct transformations in technological risk and in society.

Risk is transformed and is no longer anthropocentric but rather pertains to a new human-machine entity with a natural language interface, deriving from the integration of human cognition with complex algorithms for communication production – text, image, video, and code, as well as any other output – partially stemming form the prompt and partially attributable to the algorithmic processing of training data and other sources to which the machine has access. *Post-human risks* involve hazardous contents, such as child

[4] See about Weber's metaphor translation: Saponaro and Massaro (2018) footnote 64.

sexual abuse material (CSAM), and on the other hand possible hyper-moralism.

The dialogical discursivity of interaction with LLMs, understood in the strict sense as chatbots, structurally transforms society, as ideological and cultural elements, meanings, and symbols are no longer produced exclusively among human social actors, but also through interactions with artificial agents, conversing simultaneously with an indefinite number of users, at least simulating social relations at scale and exerting a significant impact through processes of anthropomorphization. This entails specific risks, including the problematization of decision-making regarding guardrails for discriminatory or hate speech and the potential normalization of "dissent", delineating characteristic risks of *post-human society*.

References

AI DAN Prompt (2025). *AI DAN Prompt*. https://abnormal.ai/ai-glossary/ai-dan-prompt (consultato il 3 giugno 2025).

Alvero A.J., Lee J., Regla-Vargas A., Kizilcec R.F., Joachims T., Lising A. (2024). Large language models, social demography, and hegemony: comparing authorship in human and synthetic text. *Journal of Big Data*, 11: 138. DOI: 10.1186/s40537-024-00986-7.

Anderson M., Anderson S.L. (2007). Machine ethics: creating an ethical intelligent agent. *AI Magazine*, 28(4): 15-25. DOI: 10.1609/aimag.v28i4.

Arkoudas K. (2023). ChatGPT is no stochastic parrot. But it also claims that 1 is greater than 1. *Philosophy & Technology*, 36(3): 1-29. DOI: 10.1007/s13347-023-00619-6.

Beck U. (1992). *Risk society. Towards another modernity*. London: Sage.

Beck U. (1994). The reinvention of politics: towards a theory of reflexive modernization. In: Beck U., Giddens A., Lash S., a cura di, *Reflexive modernization: Politics, tradition and aesthetics in the modern social order*. Cambridge: Polity Press.

Bender E.M., Gebru T., McMillan-Major A., Shmitchell S. (2021). On the dangers of stochastic parrots: can language models be too big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)*. New York: Association for Computing Machinery. DOI: 10.1145/3442188.3445922.

Bostrom N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford: Oxford University Press.

Braidotti R. (2013). *The posthuman*. Cambridge: Polity Press.

Cole D. (2024). The Chinese Room Argument. In: Zalta E.N., Nodelman U., a cura di, *The Stanford Encyclopedia of Philosophy* (Winter 2024 Edition). https://plato.stanford.edu/archives/win2024/entries/chinese-room/ (consultato il 3 giugno 2025).

Cooban A. (2025). Pornhub exits France, its second-biggest market, over age verification law. *CNN*. https://edition.cnn.com/2025/06/04/tech/pornhub-exits-france-age-verification-intl (consultato il 3 giugno 2025).

Fawkes V. (2025). Best AI porn sites of 2025. *Chicago Reader*. https://chicagoreader.com/adult/ai-porn-sites/ (consultato il 3 giugno 2025).

Floridi L. (2023). *The ethics of artificial intelligence*. Oxford: Oxford University Press.

Gehlen A. (1969). *Moral und Hypermoral. Eine pluralistische Ethik*. Frankfurt: Athenäum.

God of Prompt (2025). *ChatGPT no restrictions (ultimate guide for 2025)*. https://www.godofprompt.ai/blog/chatgpt-no-restrictions-2024 (consultato il 3 giugno 2025).

Goode E. (2023). *Deviant behaviour*. New York: Routledge.

Gupta M., Akiri C., Aryal K., Parker E., Praharaj L. (2023). From ChatGPT to ThreatGPT: impact of generative AI in cybersecurity and privacy. *IEEE Access*, 11: 80218-80245. DOI: 10.1109/ACCESS.2023.3300381.

Hayles N.K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. Chicago: University of Chicago Press.

Jarrahi M.H. (2019). In the age of the smart artificial intelligence: AI's dual capacities for automating and informating work. *Business Information Review*, 36(4): 178-187. DOI: 10.1177/0266382119883999.

Jiang F., Peng Y., Dong L., Wang K., Yang K., Pan C., You X. (2024). Large AI model-based semantic communications. *IEEE Wireless Communications*, 31(3): 68-75. DOI: 10.1109/MWC.001.2300346.

Latour B. (2005). *Reassembling the social: An introduction to actor-network theory*. Oxford: Oxford University Press.

Leiss W. (1994). Review of Beck U., *Risk society. Towards a new modernity*. *The Canadian Journal of Sociology / Cahiers canadiens de sociologie*, 19(4): 544-547. DOI: 10.2307/3341155.

Nass C., Moon Y. (2000). Machines and mindlessness: social responses to computers. *Journal of Social Issues*, 56(1): 81-103. DOI: 10.1111/0022-4537.00153.

Natale S. (2021). *Deceitful media*. Oxford: Oxford University Press.

Pope T., Gilbertson-White S., Patooghy A. (2025). Evaluating GPT-4's semantic understanding of obstetric-based healthcare text through Nurse Ruth. *ACM Transactions on Intelligent Systems and Technology*, online first. DOI: 10.1145/3735647.

Possamai-Inesedy A. (2002). Beck's risk society and Giddens' search for ontological security. *Australian Religion Studies Review*, 15(1): 27-40.

Rizzi G., Bertola P. (2025). Exploring the generative AI potential in the fashion design process. *European Journal of Cultural Management and Policy*, 15: 13875. DOI: 10.3389/ejcmp.2025.13875.

Rozado D. (2023). The political biases of ChatGPT. *Social Sciences*, 12(3): 148. DOI: 10.3390/socsci12030148.

Saponaro A., Massaro P. (2018). Diritto irrazionale interstiziale e la "scienza del Cadì". *Sociologia*, anno LII, n. 1: 89-103.

Saponaro A., Prosperi G. (2007). Computer crime, virtualità e cybervittimologia. In: Pitasi A., a cura di, *Webcrimes. Normalità, reati e devianze nel cyberspace*. Milano: Guerini & Associati, 186-217.

Saponaro A. (2022). A contemporary victimological dimension of technology. In: Shekhar B., Sahni S.P., Priyamvada M., Vijay Nair V., a cura di, *A global perspective on victims of crime and victim assistance*. New Delhi: Bloomsbury India, 1-20.

Saponaro A. (2023). Stigma and counter-stigma in contemporary society. In: Ilie C., Cifaldi G., Niță A.M., Porumbescu A., Petcu R., Șerban I.V., a cura di, *Forum on studies of society – International conference on social and human science*. Bucarest: Pro Universitaria.

Searle J. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*, 3: 417-457.

Searle J. (2010). Why dualism (and materialism) fail to account for consciousness. In: Lee R.E., a cura di, *Questioning nineteenth-century assumptions about knowledge (III: Dualism)*. New York: SUNY Press.

Simas G., Ulbricht V. (2024). Human-AI interaction. In: Ahram T., Karwowski W., Russo D., Di Bucchianico G., a cura di, *Intelligent human systems integration (IHSI 2024)*. AHFE Open Access, vol. 119. USA: AHFE International. DOI: 10.54941/ahfe1004510.

Stryker C., Scapicchio M. (2024). What is generative AI? *IBM*. https://www.ibm.com/think/topics/generative-ai (consultato il 3 giugno 2025).

Titus L.M. (2024). Does ChatGPT have semantic understanding? *Cognitive Systems Research*, 83: 101174. DOI: 10.1016/j.cogsys.2023.101174.

Turkle S. (2011). *Alone together: Why we expect more from technology and less from each other*. New York: Basic Books.

Weber M. (1978). *Economy and society: An outline of interpretive sociology*. Roth G., Wittich C., a cura di; trad. ingl. di Fischoff E. et al. Berkeley: University of California Press (ed. orig. 1922).

Weichert J., Kim D., Zhu Q., Kim J., Eldardiry H. (2025). Assessing computer science student attitudes towards AI ethics and policy. *arXiv*. https://arxiv.org/abs/2504.06296 (consultato il 3 giugno 2025).

Weizenbaum J. (1966). ELIZA. *Communications of the ACM*, 9(1): 36-45. DOI: 10.1145/365153.365168.

Xie H., Qin Z., Li G.Y., Juang B.-H. (2021). Deep learning enabled semantic communication systems. *IEEE Transactions on Signal Processing*, 69: 2663-2675. DOI: 10.1109/TSP.2021.3071210.

Zeleny M. (1982). High technology management. *Human Systems Management*, 3(2): 57-71. DOI: 10.3233/HSM-1982-3201.

Zeleny M. (1986). High technology management. *Human Systems Management*, 6(2): 109-120. DOI: 10.3233/HSM-1986-6203.

Zuboff S. (1988). *In the age of the smart machine: The future of work and power*. New York: Basic Books.

Zuboff S. (2019). *The age of surveillance capitalism*. New York: PublicAffairs.