# SYNTHETIC MEDIA FOR SOCIAL GOOD: UNLOCKING POSITIVE POTENTIAL

**DOI**
https://doi.org/10.18690/um.feri.2.2026.6

**ISBN**
978-961-299-109-8

ANGELO TUMMINELLI,[1] LUCY CONOVER,[2]
CALOGERO CALTAGIRONE,[1] GIUDITTA BASSANO,[1]
GJON RAKIPI,[3] TOMMASO TONELLO[2]

[1] LUMSA University, Department of Human Sciences, Rome, Italy
a.tumminelli@lumsa.it, c.caltagirone@lumsa.it, g.bassano@lumsa.it
[2] Utrecht University, Freudenthal Institute, Utrecht, the Netherlands
l.a.conover@uu.nl, t.tonello@uu.nl
[3] Albania Institute of International Studies (AIIS), Tirana, Albania
gjonrakipi@aiis-albania.org

This chapter examines the ethical, communicative, and societal dimensions of artificial intelligence for social good (AISG) through a series of participatory workshops conducted in collaboration with the European Citizen Science Association (ECSA). The workshops engaged 44 participants from 18 national backgrounds, selected according to age, gender balance, and domain expertise, and addressed emotionally and epistemically sensitive domains, including climate change communication, the visibility of women in science, and AI-mediated psychological support. The analysis identifies four determinants shaping perceived impact: narrative–intentional coherence, technical–mimetic realism, ethical transparency, and contextual adequacy. Together, these dimensions inform a preliminary set of ethical and design guidelines for socially engaged and educational media. The chapter further proposes a methodological framework that combines semiotic modelling with iterative user testing to evaluate AI-generated content beyond criteria of realism or imitation. By foregrounding communicative function, ethical clarity, and cultural resonance, the findings suggest that synthetic media can meaningfully contribute to socially oriented and educational contexts when designed with participatory and ethically grounded approaches.

**Keywords:**
social good,
literacy,
citizen science,
education,
ethics

University of Maribor Press

## 1 Why AI for Good?

The purpose of ethics is to promote the full flourishing of people in their deepest relational openness and in their aspiration to meaning. Ethics of AI is, therefore, called to not only define the normative criteria within which to place the interaction between AI and human beings, but above all, to identify the strategies with which the use of the former is placed at the service of personal fulfilment and the common good. Thus, the ethics of AI goes well beyond a merely deontological approach, constituting itself, rather, as a fundamental tool for promoting human beings in the face of the challenges imposed by the digital revolution and the advent of AI.

We follow Aristotle, who in the *Nicomachean Ethics* argued that within society, the common good must be pursued as a supreme ethical task to which individual action is called to contribute significantly. According to Aristotelian teleology, every being is oriented toward an end (*telos*) and evaluates actions based on how well they realize the human good. (Aristotle, 2012, I, 1094a, pp. 1–3)

In contemporary AI ethics, this idea reappears when defining the desirable goals of intelligent systems and the criteria for judging their alignment with human values.
In both perspectives, what matters is determining which end should guide action to direct AI development and use toward the common good.

In the present-day debate, Luciano Floridi also explains the potential of the political use of AI for the common, or social good (AI4SG), highlighting how its ethical use necessarily implies "the design, development and implementation of AI systems in order to (I) prevent, mitigate or solve problems that negatively impact human life and/or the well-being of the natural world and/or (II) allow socially preferable and/or environmentally sustainable developments" (Floridi, 2022, p. 223).

Luciano Floridi (with Josh Cowls) proposes five fundamental ethical principles for AI, often referred to as the "Unified Framework of AI Ethics":

*Beneficence* – AI should promote well-being and generate social value.

*Non-maleficence* – it should avoid harm, undue risks, and abusive uses.

*Autonomy* – it should respect individuals' decision-making capacity without manipulating them.

*Justice* – it should be fair, non-discriminatory, and distribute benefits and burdens appropriately.

*Explicability* – it should ensure transparency, intelligibility, and traceability of decisions.

Floridi bases his reflection on the ethics of AI on these five principles, borrowed from an accredited approach in bioethics, to combine the use of AI and the promotion of the individual and the common good of humanity. In compliance with the principle of Beneficence, according to Floridi, it is necessary to create an AI technology that is beneficial for humanity and that puts the promotion of the well-being of people and the planet at its centre, thus safeguarding the human dignity of the present and the future as a common good.

The principle of non-maleficence, on the other hand, is based on the need to prevent violations of personal privacy to avoid improper use of AI technologies that could harm humanity as a whole. The principle of Autonomy, then, is the one that is called to safeguard the freedom of individuals as a shared heritage (Floridi, 2022): if it is true that when AI and its intelligent action are adopted, the individual voluntarily gives up part of his decision-making power to machines, affirming the principle of Autonomy in the context of AI means reaching a balance between the decision-making power that the individual retains within himself and that which he delegates to artificial agents. Starting from this, not only should human freedom be promoted, but also the autonomy of machines should be restricted and made intrinsically reversible.

Floridi's perspective is particularly interesting because it places the social good and the possibility that it can be achieved through personal freedoms at the centre of an ethical use of AI (Floridi, 2022; Floridi et al., 2020). Only when this happens in a society can the common good be achieved: this is not a utopia but an ethical task that awaits all human beings in the face of the challenges of their time.

If ethics aims to guide human action toward personal flourishing and meaningful relationships, then AI ethics must not only set the norms governing human-AI interaction, but also determine how AI can genuinely support human fulfilment and the common good. Thus, AI ethics goes beyond a purely deontological framework: it becomes a key instrument for fostering human development in the face of the digital revolution and the rise of AI.

## 2 Positive Applications in Citizen Science, Community Engagement, and Education

Since 2015, the United Nations Sustainable Development Goals (UNSDGs) have been endorsed by all UN Member States to tackle the most pressing social, environmental, and economic issues by 2030. Citizen science, as "a form of research collaboration involving members of the public in scientific research projects to address real-world problems" (Wiggins & Crowston, 2012) has proven its contribution to the SDGs. Citizen science is an "umbrella term" to include various participatory approaches where non-professional scientists contribute to research (ECSA, 2015; 2020), such as participatory monitoring, crowd-sourced science, or participatory action research. Indeed, participatory approaches leveraging public involvement have demonstrated to significantly enhance data collection, foster community empowerment, and drive progress toward achieving the SDGs (Ballerini & Bergh, 2021; Fraisl et al., 2023; Gaventa & Barrett, 2012; Huttunen et al., 2022; Loeffler & Martin, 2015; Müller et al., 2023). In this section, we show how AI is used in citizen science initiatives, community engagement and education to support the Sustainable Development Goals. This section will present a short background of different types of AI-supported citizen science initiatives and learnings from the SOLARIS project, which constitute the bedrock of the activities carried out during Use Case 3 (UC3).

In citizen science, AI-driven tools can enhance data analysis, pattern recognition, and predictive modelling, not only improving the efficiency and accuracy of citizen science projects, but also expanding their scope and scalability (Fraisl et al., 2025; Hayes et al., 2025; Sinha et al., 2024). Among citizen science projects, the most common way of integrating AI is by having participants train algorithms (Chandler et al., 2025; DeSpain et al., 2024; Duerinckx et al., 2024, p. 3; Jia et al., 2025; See et al., 2025). This is sometimes called "hybrid intelligence (HI) systems" (Chen et al., 2024) or "Crowd AI" (Palmer et al., 2021), as citizen scientists provide data and support machine classification tasks, for example in monitoring efforts such as high-tide flooding (Golparvar & Wang, 2020), vector-borne diseases (Saran & Singh, 2024), or harmful mosquitos or snails (Chan et al., 2024). AI use in citizen science also enhances challenges such as the mitigation of algorithmic biases (Vinuesa et al., 2020) and inclusive, accessible technological designs that ensure broad participation (Fortson et al., 2024). Questions remain in terms of data privacy, hence emphasizing the importance of adopting ethical frameworks that prioritize transparency, accountability, and fairness in citizen science projects (Ceccaroni et al., 2019; Fortson et al., 2024; Vinuesa et al., 2020). In citizen science biodiversity research, for instance, AI can be used for species identification (Hogeweg et al., 2024), such as mammal species in the FOOTPRINTS-CITSC project,[1] or diseases on potato crops in the PataFest project.[2] Additionally, AI chatbots on biodiversity monitoring platforms have also been shown to enhance engagement, as contributors use the bot as a "dialogic partner" to discuss the pictures of bumblebees they upload (Sharma et al., 2024). And yet, power asymmetries in current data governance still fail to properly acknowledge citizen scientists as relevant stakeholders for drafting and implementing data principles, which in turn inform data storage and data use. Nonetheless, the same public engagement values that support citizen science would appear to benefit ethical data governance: there already exist positive initiatives, especially in relation to citizen science as undertaken within indigenous communities, to inquire into local knowledge. By fostering data justice processes – e.g., through the promotion of data commons and cooperatives – and the enhancement of multi-stakeholder data governance processes through its participatory principles, citizen science represents a relevant tool to also enhance accountability mechanisms and to democratise data governance (Borda & Greshake Tzovaras, 2025; Sterner & Elliott, 2024). In the educational sector, the Smartschool project,[3] Supporting teachers and pupils through a smart signal, is currently working on an AI tool for teachers to identify their teenage students' learning needs on a learning platform. The project is a collaboration between students, parents, education professionals, and Hasselt University.[4] Furthermore, the Monumai project[5] citizens participate in data collection and training algorithms to recognize architectural styles from photographs of monuments, whereby they also learning to recognize the characteristics. In the care sector, the project "Machine learning as a citizen science tool to improve the quality of life of older people and their caregivers'[6] wants to make psychology and computer science research accessible to

[1] See link: https://footprints.citizenscience.no/
[2] See link: https://www.patafest.eu/
[3] See link: https://citizenscience.eu/project/488
[4] See link: https://www.uhasselt.be/en/faculties-and-schools/school-of-social-sciences
[5] See link: https://monumai.ugr.es/
[6] See link: https://citizenscience.eu/project/72

the wider society and support the early detection of loneliness, social isolation, and stress in older adults. Data is provided by volunteers, who will analyse it before feeding machine learning algorithms for training.

The aforementioned projects show how, across disciplines, citizen science initiatives are increasingly using AI tools to address various SDGs. “AI for good”, in the context of UC3, means AI to achieve the SDGs. By promoting citizens’ participation in the co-creation of AI-generated content for educational purposes, UC3 aimed to promote AI to achieve the SDGs, or “AI for good”. It supported SDG 4 - Quality Education, in two ways: first, participants co-created content for awareness raising – on topics such as climate change; second, the workshops fostered participants’ digital literacy and enabled individuals to better understand and navigate the complexities of AI technology. UC3 also played a significant role in advancing SDG 16 - Peace, Justice, and Strong Institutions, by pushing for pro-democratic values and promoting transparency and accountability in AI governance. The participatory governance model inherent in UC3 encouraged citizens to take an active role in decision-making processes, thereby ensuring that AI systems align with societal values. In practice, we selected three SDGs to promote “AI for good”:

- SDG 3: Good Health and well-being, focusing on mental health,
- SDG 5: Gender equality, especially with regards to the inclusion of women in science, and
- SDG 13 Climate Action, focusing on the effects of climate change.

SOLARIS project member created eight videos on these themes. During the workshops part of SOLARIS UC3 activities, we therefore contributed to an acceptable or desirable approach for awareness raising of artificially generated content. We framed possible answers to the question: “what could “good” AI-generated content look like?” By enabling citizens to co-create AI-generated content with experts, the workshops contributed to the transparency, inclusivity, and accountability that are fundamental to democratic governance. The workshops were also based on the value-sensitive design approach (Umbrello & Van De Poel, 2021, p. 284), which takes “values of ethical importance into account”, considering “a tripartite methodology of empirical, conceptual and technical investigations”.

## 3 Semiotic at the service of AI for Good

Use Case 3 explored the civic and communicative potential of "positive deepfakes," that is, synthetic texts generated by AI for educational, memorial, scientific, and civic engagement purposes, rather than for manipulative or deceptive purposes. UC3 adopted a semiotic and processual approach. Its goal was not to evaluate persuasion or misinformation, but to understand how artificial texts[7] are constructed, which dimensions guarantee their credibility, or conversely, reveal their artificiality, and how workshop participants interpret such products by attributing meaning to them.

Within this framework, "semiotics", understood as the science of meaning-making forms and of the conditions of their production and interpretation (Eco, 1976; Greimas, 1983; Greimas & Courtés, 1982; Hjelmslev, 1961) was considered a useful framework to complement the ethical perspective of AI4SG. UC3, therefore, sought to approach deepfakes as semiotic objects whose analysis requires decomposition into levels of textual articulation and reconstruction of the pragmatic conditions of reception. Hence, there is a need for a multilevel analysis integrating discursive, narrative, enunciative, axiological, and plastic components to map how synthetic contents acquire meaning and produce social effects. From a semiotic perspective, each artificially generated video can be analysed as a text articulated on multiple levels:

- Discursive level: any audiovisual text, even a static one, "speaks" of something, projects figures, situates them in space and time, and constructs a coherent discursive universe.
- Narrative level: concerns the characters' actions, the transformations that occur, and the evolution of the storyline. It is the level at which conflicts, changes of state, and narrative programs can be observed.
- Enunciative level: includes the traces indicating the relationship between sender and receiver, the contracts of truth, and the framing regimes (fiction, testimony, document, hybrid, etc.).
- Axiological level: relates to the explicit or implicit values conveyed by the text, such as truth, authority, empathy, transparency, or responsibility.

[7] In semiotics, "text" is a generic term that can refer to audiovisual contents too.

To these levels, we add the specificity of visual and audiovisual texts. According to Polidoro (2008), visual semiotics distinguishes two areas of analysis:

- Figurative semiotics, which analyses meaning derived from the recognition of objects and scenes.
- Plastic semiotics, which investigates the significance of visual configurations such as shapes, colours, textures, and lighting.

This dual articulation suggests that the plausibility of visual content does not depend solely on perceptual accuracy but is mediated by cultural codes and cognitive competencies. Visual literacy is built over time through familiarity with communicative genres, aesthetic codes, and narrative conventions. This was the ground on which UC3 developed its investigation.

The eight videos produced in UC3 were designed to systematically and creatively test a set of variables.[8] The language used in all videos was English, and the videos covered the following themes:

- **SDG3: Women Scientists– Marie Curie**: three videos presented the scientist as an authoritative witness, capable of reflecting on the role of women in science.
- **SDG5: Climate Crisis– Amina:** Two videos narrated the experience of a woman forced to leave her homeland near Lake Chad due to desertification.
- **SDG13: Mental Health– Casey**: Two videos explored the use of synthetic avatars in psychological therapy.

The design logic was to combine predefined variables to observe thresholds of acceptability and mechanisms of suspended disbelief. Eight variables were initially identified, derived from narratological frameworks already adapted in previous research on synthetic media and video analysis (Bassano & Cerutti, 2024; Genette, 1982; Greimas, 1988). Their articulation allowed us to operationalise classical narrative dimensions, namely actoriality, focalization, setting, and modality within an experimental design suited to AI-generated content:

[8] The videos are safely stored in the SOLARIS archive and can be accessed upon request, but they are not publicly accessible.

1) famous vs. an anonymous person.
2) realistic vs. decontextualized/abstract setting.
3) monologue vs. dialogue.
4) focus on detail vs. overall view.
5) blurred face vs. AI-generated (deepfake) face.
6) first-person vs. third-person narration.
7) artificial landscape vs. artificial person.
8) serious vs. entertainment context.

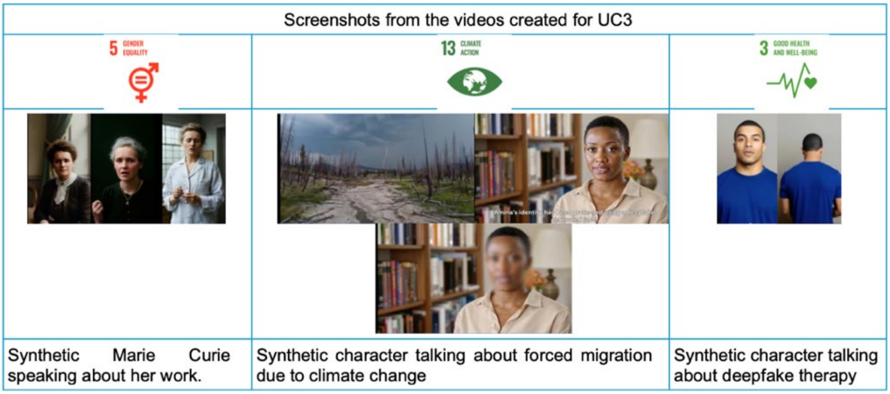


**Figure 6.1: Screenshots from the videos created for UC3**
Source: SOLARIS

For practical reasons, the deepfakes focused on five of these variables (1, 2, 5, 6, 7), which were articulated across the three themes described above. The scripts were initially proposed by ECSA, then further developed and conceptually authored by Giuditta Bassano (LUMSA), and finally produced by the partner CINI, in particular by Michele Brienza.

### 3.1 The Textual Taxonomy of UC3

Based on this theoretical framework, and on the analysis of data collected during the workshops, we propose a textual classification of the positive deepfakes used during UC3 along three principal axes: (i) their discursive form, (ii) their identity function, and (iii) their destination. These three axes, intertwined with one another, enable the distinction of how synthetic actors acquire meaning and produce communicative effects. This taxonomy, specifically developed for the purposes of this project and constituting an original contribution of this chapter, indicates that the evaluation of *positive* AI-generated contents cannot be based solely on technical quality. Instead, they must be read as complex textual configurations capable of combining different degrees of discursive involvement, identity strategies, and forms of destination. In this section, the term "textual" refers to the intrinsic configuration of the deepfake as a discursive object: its narrative structure, identity work, and intended destination. This level concerns the organization of meaning within the text itself, independently of how it is received. By contrast, the interpretive taxonomy presented in the following section focuses on the modes of reception activated by audiences, showing how viewers make sense of the same textual features through different perceptual, cognitive, and ethical frameworks.

The first axis (discursive form) concerns the degree of personal involvement that the narrator assumes in the account. We can imagine a continuous spectrum with two opposite poles. On one side, we would place the evocative or illustrative pole. This occurs when the narrative voice remains external, minimally engaged in the first person, limiting itself to evoking facts or presenting issues. This is the case of Marie Curie: even when referring to her own biography, the scientist appears rather detached, informing us of "public" events, already known and of common interest, thus functioning more as an exemplary figure than as a subject testifying in the first person to a personal experience. On the opposite side, we find the testimonial pole, a position that entails the highest degree of intimacy and subjective implication. Casey's narrative could have been placed here, especially if the synthetic actor had gone so far as to describe concrete details of his anxiety disorder.

The second axis (identity function) concerns the way in which deepfakes handle the identity of the subject being represented. We distinguish between passive and active functions. The passive function consists in covering and protecting a real identity by concealing its individual traits. This is the case of Amina and Casey, whose faces were blurred or withheld from view, to safeguard anonymity or reduce exposure. The active function, instead, corresponds to the maximum degree of identity affirmation, when the deepfake serves a memorial function, bringing historical figures back to life to prolong their presence. This is the case of Marie Curie, who appears or is evoked in the three videos as a historical and symbolic figure, whose identity is not concealed but reaffirmed and consolidated.

The third axis (destination) concerns intended use of deepfakes. Here, too, we can imagine a continuum. On one end lies the public pole, meaning texts designed for a broad, general audience, such as the Dalí deepfake (evoked during the UC3 workshops) in a museum setting. The videos of Marie Curie also share this orientation: they are meant to convey collective values and educational messages. On the other end lies the specific pole, which refers to contents designed for situated, personalized, or dialogic use. This is the case of the videos about Casey, which evoke an individual therapeutic context, as well as the workshop discussions about chatbots as personal assistants capable of establishing a unique relationship with a single user. By combining the three axes, it is possible to position the UC3 cases within a textual matrix:

- Marie Curie: *evocative, active, public*;
- Amina: *evocative/testimonial, passive, public*;
- Casey: *testimonial, passive, specific*;

Considered together, the three cases display different types of balance across the proposed axes. Marie Curie, as a historical and already public figure, clearly occupies an evocative position on the first axis, rather than a testimonial one, since the narrative mobilizes shared and well-known events without direct personal involvement. On the second axis, her deepfake performs an active identity function, reinforcing and extending her symbolic presence. Finally, its destination is unmistakably public, oriented toward broad educational dissemination. Amina occupies a more nuanced position: her discourse is predominantly evocative, yet certain passages introduce elements of testimonial engagement. Her identity, however, remains passively configured, as the message protects and obscures individual traits; her destination is likewise public, given that the content is framed as a general appeal. Casey stands at the opposite corner of the matrix: his deepfake is grounded in a strongly testimonial mode, openly engaging personal experience; his identity is passive, since his face is concealed for privacy reasons; and the destination is specific, as the video aligns with therapeutic or relational contexts rather than with broad public dissemination.

## 3.2 The Interpretive Taxonomy

While the textual taxonomy has made it possible to classify civic deepfakes according to their formal and discursive configuration, an interpretive taxonomy allows us to understand their modes of reception. The UC3 workshops showed that the credibility of deepfakes does not depend solely on technical realism but unfolds through different interpretive registers activated by the audience when encountering the texts. We can distinguish five primary modes of reception:

1) *Plastic interpretation*: this is the most immediate threshold of access, linked to visual and auditory perception. Details such as lip-sync, frame rate, coherence of lighting and textures, movement rhythm, or the quality of the synthetic voice constitute decisive clues for acceptance or rejection. In the workshops, younger participants proved particularly sensitive to this level: for them, plastic realism represented a non-negotiable condition of credibility. This emerged clearly in reactions to Marie Curie's slightly imperfect lip-sync, which younger participants immediately flagged as a credibility break.

2) *Discursive interpretation*: beyond the plastic level, viewers assessed the content based on narrative and thematic coherence. Here, the effects of meaning emerge, tied to the construction of plausible stories, the consistency of the conveyed values, and the text's ability to articulate a meaningful account. Older participants tended to prioritize this dimension, paying greater attention to the quality of discourse than to technical perfection. For instance, when the video on climate-change consequences was shown, participants focused on the coherence between the verbal text and the visual depiction of environmental impacts.

3) *Ethical-cognitive interpretation*: the reception of civic deepfakes also implies a judgment about the appropriateness of their use in specific contexts. The workshops revealed that a deepfake may be deemed acceptable in a museum or classroom, yet disturbing in a promotional or commercial setting. This level thus concerns the audience's ability to relate synthetic content to social and ethical frameworks, evaluating its legitimacy and transparency. For example, in Casey's case, participants noted that it would be inappropriate to use an avatar of someone with mental health disorders in a pharmaceutical advertisement or in promotional material for medical services. They also stressed, however, that

this is very different from the experience of a patient with mental health conditions who wants to educate and inform others through a deepfake.

4) *Passional interpretation*: a fourth register concerns the emotional dimension. Reception depends on the alignment between sensible form and narrated content: a smiling face recounting a trauma generates discomfort, whereas an empathetic tone strengthens the text's acceptability. This aspect became evident when participants discussed the quality of Amina's video, noting that her expression appeared too cheerful compared to the dramatic nature of what she was describing.

5) *Metareflective interpretation*: finally, a more sophisticated mode arises when participants thematize the deepfake itself as an object of reflection. Co-creation fostered this level: citizens discussed the contents and the cultural, ethical, and political implications of the technology, highlighting their active role as critical interpreters. This mode emerged directly from the workshop discussions, as a recurrent interpretive pattern observed among participants. In UC3, this mode surfaced when participants discussed the broader implications of using deepfakes of figures like Marie Curie, Amina, and Casey in civic contexts.

The intersection between the textual and interpretive taxonomies shows how the three strands of UC3 were received in different ways. For Marie Curie, the public dimension seemed to strengthen acceptability, even though workshop participants still emphasized discursive and ethical-cognitive interpretation (given the educational context). For the synthetic character of Amina, identity protection and blurring weakened the testimonial effect; participants oscillated between plastic rejection (the synchronization of body and facial movements was judged unconvincing) and passional discomfort, while nevertheless paying attention to significant metareflective aspects, such as the synthetic actress's voice. For the synthetic character of Casey, the testimonial effect appears to have failed altogether, as participants mainly interpreted the video in plastic and passional terms, discussing evident artificiality and a sense of detachment. The analysis of the workshops provided a rich picture of how citizens interpret and evaluate synthetic content, offering empirical validation for the two taxonomies developed. The results extend beyond observing individual reactions, as they demonstrate how participants employed complex interpretive strategies, combining plastic, discursive, ethical-cognitive, passional, and metareflective evaluations.

Despite the richness of its findings, UC3 presents certain structural limitations tied to the online workshops' format. The videos were shown in standardized, decontextualised conditions, far removed from the communicative ecosystems in which synthetic content circulates typically. As already noted, a deepfake never exists in isolation: its meaning depends on the discourses that accompany it, the users' comments, the platforms that host it, the viewing devices, and the intertextual frameworks into which it is inserted – this is the network approach developed by SOLARIS project (see McIntyre et al., 2025, Bisconti et al., 2024).

## 4 Concluding Remarks

Our findings bring to the fore the theme of "Digital education". Digital education plays a crucial role in developing skills for digital citizenship and democracy, as it trains individuals capable of interacting consciously, responsibly, and actively in a digital context. These skills are essential to navigate the online world and to participate in democratic life with critical thinking and respect, promoting open and inclusive dialogue. Digital education promotes skills such as critical thinking, responsibility, respect for privacy and digital rights, the fight against disinformation, and active participation. In this regard, starting from the interplay between empirical findings and theoretical models, Panciroli and Rivoltella (2023) speak of "algorithmic pedagogy", meaning the set of strategies that make use of technological and digital devices used in educational contexts to promote learning and the integral formation of the person. The two scholars refer to three possible configurations of algorithmic pedagogy, and distinguish: 1. "AI in education", which involves the teacher being supported by a humanoid robot available to answer students' questions based on profiling and individualized programming processes (here the reference is to robots used in co-teaching for feedback management and personalized tutoring); 2. "AI by education", or the provision of pre-established and predetermined ethical criteria for devices in the design phase (in this regard, the responsibility of the computer designer comes into play, who, already in the creation of the algorithm and in the writing of the code, establishes limits and ethical criteria); 3. "AI for education", which consists of the task of digital education, aimed at arousing critical thinking in students. This awareness implies distancing from the technological artefact, which is recognized in its functional utility and not as a substitute for interpersonal educational relationships. An ethical digital education, in the context of the infosphere, thus becomes an essential basis for the promotion of humanity and the construction of the common good.

Overall, we see that AI has the potential to promote social good, if it is developed and used responsibly. By maintaining thoughtful reflection about the complexities of AI in the context of education and social good, the technology could be used to provide a positive lens in these fields. However, future endeavours need to avoid the deficit model, which considers the general public as only lacking skills to interact with AI: while education has a crucial role to play, focusing only on digital education tends to reinforce systemic barriers to participation and inclusion (Patel, 2025). Instead, we need to ensure that diverse voices are included and can participate in the development of tools and technologies influencing society. Future research should focus on participatory co-design of educational AI tools.

**End notes**

Angelo Tumminelli and Calogero Caltagirone conceptualized the chapter, coordinated the writing of the text, and conducted the literature review. Angelo Tumminelli and Calogero Caltagirone wrote the abstract, conclusions, and section "Why AI for good?". Gjon Rakipi and Tommaso Tonello wrote Section "Positive applications in citizen science, community engagement and education"; Giuditta Bassano wrote Sections "semiotic at the service of AI for Good", "The Textual Taxonomy of UC3", and "The Interpretive Taxonomy". Lucy Conover harmonized the whole text and coordinated revisions, including reference formatting and editing for grammar and style of the entire chapter. All authors reviewed and approved the final version. We thank the colleagues from the European Citizen Science Association (ECSA) – Romane Blassel, Nefertari Nachtigall, and Carolina Doran for helpful comments and suggestions on earlier versions of this chapter, and for their contributions to the activities described and discussed in this chapter.

**References**

Aristotle. (2012). *Nicomachean ethics* (R. C. Bartlett & S. D. Collins, Trans.). University of Chicago Press.

Ballerini, L., & Bergh, S. I. (2021). Using citizen science data to monitor the Sustainable Development Goals: A bottom-up analysis. *Sustainability Science, 16*(6), 1945–1962. https://doi.org/10.1007/s11625-021-01001-1

Bassano, G., & Cerutti, M. (2024). Posthumous digital face: A semiotic and legal semiotic perspective. *International Journal for the Semiotics of Law / Revue Internationale de Sémiotique Juridique, 37*(3), 769–791. https://doi.org/10.1007/s11196-023-10067-2

Bisconti, P., McIntyre, A., & Russo, F. (2024). Synthetic socio-technical systems: Poiêsis as meaning making. *Philosophy & Technology, 37*(3), Article 94. https://doi.org/10.1007/s13347-024-00778-0

Borda, A., & Greshake Tzovaras, B. (2025). *Perspectives on crowdsourced citizen science and the data governance of AI applications* (SSRN Scholarly Paper No. 5291262). Social Science Research Network. https://doi.org/10.2139/ssrn.5291262

Ceccaroni, L., Bibby, J., Roger, E., Flemons, P., Michael, K., Fagan, L., & Oliver, J. L. (2019). Opportunities and risks for citizen science in the age of artificial intelligence. *Citizen Science: Theory and Practice, 4*(1), Article 29. https://doi.org/10.5334/cstp.241

Chan, K. H., Tumusiime, J., Jacobs, L., & Huyse, T. (2024). The potential of deep learning object detection in citizen-driven snail host monitoring to map putative disease transmission sites. *Citizen Science: Theory and Practice, 9*(1), Article 25. https://doi.org/10.5334/cstp.724

Chandler, C. O., Sedaghat, N., Oldroyd, W. J., Frissell, M. K., Trujillo, C. A., Burris, W. A., Hsieh, H. H., Kueny, J. K., Farrell, K. A., Borisov, G. V., DeSpain, J. A., Bernardinelli, P. H., Kurlander, J., Magbanua, M. J. M., Sheppard, S. S., Mazzucato, M. T., Bosch, M. K. D., Shaw-Diaz, T., Gonano, V., … Dukes, C. J. A. (2025). AI-enhanced citizen science discovers cometary activity on near-Earth object (523822) 2012 DG61. *Research Notes of the AAS, 9*(1), Article 3. https://doi.org/10.3847/2515-5172/ada368

Chen, V. Y., Lu, D.-J., & Han, Y.-S. (2024). Hybrid intelligence for marine biodiversity: Integrating citizen science with AI for enhanced intertidal conservation efforts at Cape Santiago, Taiwan. *Sustainability, 16*(1), 454. https://doi.org/10.3390/su16010454

DeSpain, J. A., Chandler, C. O., Sedaghat, N., Oldroyd, W. J., Trujillo, C. A., Burris, W. A., Hsieh, H. H., Kueny, J. K., Farrell, K. A., Magbanua, M. J. M., Sheppard, S. S., Mazzucato, M. T., Bosch, M. K. D., Shaw-Diaz, T., Gonano, V., Lamperti, A., Da Silva Campos, J. A., Goodwin, B. L., Terentev, I. A., & Dukes, C. J. A. (2024). Discovery of Jupiter family comet 2011 UG104 through AI-enhanced citizen science. *Research Notes of the AAS, 8*(5), Article 140. https://doi.org/10.3847/2515-5172/ad4d9c

Duerinckx, A., Veeckman, C., Verstraelen, K., Singh, N., Van Laer, J., Vaes, M., Vandooren, C., & Duysburgh, P. (2024). Co-creating artificial intelligence: Designing and enhancing democratic AI solutions through citizen science. *Citizen Science: Theory and Practice, 9*(1), Article 43. https://doi.org/10.5334/cstp.732

Eco, U. (1976). *A theory of semiotics*. Indiana University Press.

Floridi, L. (2022). *Etica dell"intelligenza artificiale: Sviluppi, opportunità, sfide* (M. Durante, Ed.; 1st ed.). Raffaello Cortina Editore.

Floridi, L., Cowls, J., King, T. C., & Taddeo, M. (2020). How to design AI for social good: Seven essential factors. *Science and Engineering Ethics, 26*(3), 1771–1796. https://doi.org/10.1007/s11948-020-00213-5

Fortson, L., Crowston, K., Kloetzer, L., & Ponti, M. (2024). Artificial intelligence and the future of citizen science. *Citizen Science: Theory and Practice, 9*(1), Article 32. https://doi.org/10.5334/cstp.812

Fraisl, D., See, L., Fritz, S., Haklay, M., & McCallum, I. (2025). Leveraging the collaborative power of AI and citizen science for sustainable development. *Nature Sustainability, 8*(2), 125–132. https://doi.org/10.1038/s41893-024-01489-2

Gaventa, J., & Barrett, G. (2012). Mapping the outcomes of citizen engagement. *World Development, 40*(12), 2399–2410. https://doi.org/10.1016/j.worlddev.2012.05.014

Genette, G. (1982). *Figures of literary discourse* (A. Sheridan, Trans.). Columbia University Press.

Golparvar, B., & Wang, R.-Q. (2020). AI-supported citizen science to monitor high-tide flooding in Newport Beach, California. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Advances in Resilient and Intelligent Cities* (pp. 66–69). Association for Computing Machinery. https://doi.org/10.1145/3423455.3430315

Greimas, A. J. (1983). *Structural semantics: An attempt at a method* (R. Schleifer, D. McDowell, & A. R. Velie, Trans.). University of Nebraska Press.

Greimas, A. J. (Ed.). (1988). *Maupassant: The semiotics of text: Practical exercises* (P. Perron, Trans.). John Benjamins Publishing Company. https://doi.org/10.1075/sc.1

Greimas, A. J., & Courtés, J. (Eds.). (1982). *Semiotics and language: An analytical dictionary*. Indiana University Press.

Hayes, S., Jandrić, P., la Velle, L., Earle, S., Šrajer, F., Dragić, Z., Kubat, S., Peraica, A., Švraka, D., Popović, S., Mumelaš, D., Pospiš, D., Vujanović, B., Lugović, S., Jopling, M., Tolbert, S., & Watermeyer, R. (2025). Postdigital citizen science and humanities: Dialogue from the ground. *Postdigital Science and Education, 7*(1), 188–223. https://doi.org/10.1007/s42438-024-00514-z

Hjelmslev, L. (1961). *Prolegomena to a theory of language* (F. J. Whitfield, Trans.; Rev. English ed.). University of Wisconsin Press.

Hogeweg, L., Yan, N., Brunink, D., Ezzaki-Chokri, K., Gerritsen, W., Pucci, R., Ghani, B., Stowell, D., & Kalkman, V. (2024). AI species identification using image and sound recognition for citizen science, collection management and biomonitoring: From training pipeline to large-scale models. *Biodiversity Information Science and Standards, 8*, e136839. https://doi.org/10.3897/biss.8.136839

Huttunen, S., Ojanen, M., Ott, A., & Saarikoski, H. (2022). What about citizens? A literature review of citizen engagement in sustainability transitions research. *Energy Research & Social Science, 91*, 102714. https://doi.org/10.1016/j.erss.2022.102714

Jia, P., Lv, J., Li, Y., Song, Y., Fu, M., Li, Z., Liu, Y., Li, N., Shen, S., Wang, T., Li, R., Zhou, Z., Ren, J., He, Z., Li, S., Tao, Y., & Cui, C. (2025). Galaxy Circus: A new paradigm for anomalous galaxy discovery with artificial intelligence and citizen science. *Preprint in review*. https://doi.org/10.21203/rs.3.rs-6397547/v1

Loeffler, E., & Martin, S. (2015). Citizen engagement. In *Public management and governance* (3rd ed.). Routledge.

McIntyre, A., Conover, L., & Russo, F. (2025). A network approach to public trust in generative AI. *Philosophy & Technology, 38*(4), Article 137. https://doi.org/10.1007/s13347-025-00974-6

Müller, M., Lorenz, J., Voigt-Heucke, S., Heinrich, G., & Oesterheld, M. (2023). Citizen science for the Sustainable Development Goals? The perspective of German citizen science practitioners on the relationship between citizen science and the Sustainable Development Goals. *Citizen Science: Theory and Practice, 8*(1), Article 34. https://doi.org/10.5334/cstp.583

Palmer, M. S., Huebner, S. E., Willi, M., Fortson, L., & Packer, C. (2021). Citizen science, computing, and conservation: How can "crowd AI" change the way we tackle large-scale ecological challenges? *Human Computation, 8*(2), 54–75. https://doi.org/10.15346/hc.v8i2.123

Panciroli, C., & Rivoltella, P. C. (2023). *Pedagogia algoritmica: Per una riflessione educativa sull"intelligenza artificiale*.Scholé.

Polidoro, P. (2008). *Che cos'è la semiotica visiva* (1st ed.). Carocci.

Saran, S., & Singh, P. (2024). Systematic review on citizen science and artificial intelligence for vector-borne diseases. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-4–2024*, 397–402. https://doi.org/10.5194/isprs-archives-XLVIII-4-2024-397-2024

See, L., Chen, Q., Crooks, A, Laso Bayas, J. C., Fraisl, D., Fritz, S., Georgieva, I., Hager, G., Hofer, M., Lesiv, M., Malek, Ž., Milenković, M., Moorthy, I., Orduña-Cabrera, F., Pérez-Guzmán, K., Schepaschenko, D., Shchepashchenko, M., Steinhauser, J., & McCallum, I. (2025). New directions in mapping the Earth's surface with citizen science and generative AI. *iScience, 28*(3), 111919. https://doi.org/10.1016/j.isci.2025.111919

Sharma, N., Colucci-Gray, L., Lakeman-Fraser, P., Robinson, A., Newman, J., Van der Wal, R., Rueger, S., & Siddharthan, A. (2024). Image recognition as a "dialogic AI partner" within biodiversity citizen science – An empirical investigation. *Citizen Science: Theory and Practice, 9*(1), Article 35. https://doi.org/10.5334/cstp.735

Sinha, R. K., Kumar, R., Phartyal, S. S., & Sharma, P. (2024). Interventions of citizen science for mitigation and management of plastic pollution: Understanding sustainable development goals, policies, and regulations. *Science of The Total Environment, 955*, 176621. https://doi.org/10.1016/j.scitotenv.2024.176621

Sterner, B., & Elliott, S. (2024). How data governance principles influence participation in biodiversity science. *Science as Culture, 33*(3), 366–391. https://doi.org/10.1080/09505431.2023.2214155

Umbrello, S., & Van de Poel, I. (2021). Mapping value sensitive design onto AI for social good principles. *AI and Ethics, 1*(3), 283–296. https://doi.org/10.1007/s43681-021-00038-3

Wiggins, A., & Crowston, K. (2012). Goals and tasks: Two typologies of citizen science projects. In *2012 45th Hawaii International Conference on System Sciences* (pp. 3426–3435). IEEE. https://doi.org/10.1109/HICSS.2012.295