SOCIAL NETWORKS AND ROBOT COMPANIONS

TECHNOLOGY, ETHICS, AND SCIENCE FICTION

CARME TORRAS

Information technologies have become part of our everyday lives and are increasingly acting as intermediaries in our workplaces and personal relationships or even substituting them. This growing interaction with machines poses several questions about which we have no previous experience, nor can we reliably predict how they will influence the evolution of society. This has led to the convergence of technoscience and humanities in an ethical debate that is starting to bear fruit, not only with the setting of regulations and standards, but also with educational initiatives in university teaching, professional improvement, and the conformation of public opinion. Interestingly, science fiction often plays a prominent speculative role in highlighting the pros and cons of potential scenarios.

Keywords: social networks, assistive robotics, ethics, science fiction, science and humanities.

◼ NETWORKS AND ROBOTS BECOME «SOCIAL»

The invention of the Internet and the emergence of mobile phones have led to the creation of social networks, a phenomenon that was not easy to foresee only a few decades ago. It was also difficult to imagine that robots would stop being limited to the work environment and that so-called *social robots* would appear. We will increasingly see these robots in our everyday surroundings: assisting elderly or disabled people, working as receptionists or as shopping centre clerks, as guides at festivals and museums, acting as playmates for young people and adults, and even working as babysitters and teaching assistants.

Computer and robotic technologies represent a further step in a social transformation that started with the agricultural revolution and continued with the industrial revolution, but they introduce a qualitative difference. It is not merely a matter of mechanising heavy and repetitive tasks in farms and industries, or the fact that electrical appliances free up people's time for use in more creative and enjoyable ways. The difference lies in these new technologies affecting our circle of social relationships and entering the sphere of emotions and feelings.

This poses a series of ethical questions that were not relevant for other types of machines. For instance, cognitively impaired people may think that the robots looking after them actually care for their well-being and consequently, may delegate all decision-making to them. Or children might think that their robotic companions are real friends. A business laptop could know when and where its user was most productive and share this information with other devices in order to optimise performance or for other purposes, with or without the user's consent. Not to mention the manipulation skills of so-called influencers and the potential misuse of the large amount of information we share over social networks.

«NEW COMPUTING AND ROBOTICS TECHNOLOGIES AFFECT OUR CIRCLE OF SOCIAL RELATIONSHIPS»

◼ OPPORTUNITIES AND DANGERS IN SOCIAL NETWORKS

An influencer is an individual who can influence the decision-making processes of others. They can

influence people to obtain power, and vice versa, they can use power to influence people. The Internet has created enormous changes in persuasion patterns, multiplying the possibilities for influencing people and placing them within anyone's reach. Michael Wu, an expert in social network analysis, has identified six determinant factors of such social-media influences (Wu, 2012). On the one hand, there are general aspects such as the credibility of the influencers and the audience they have; on the other, are aspects connected to those on the receiving end of this influence: how much they rely on the influence, the relevance of the information, as well as its opportuneness in time and space (factors known as «the right person, information, time, and place»). Regarding the first two, there are no universal influencers or general media for them. An influencer is only an influencer in their field of expertise and through the right channels. The timeliness of the message is crucial because there is a golden window of opportunity during which the receiver is interested in making a decision and is open to considering different options. This window is detected by the user's actions, such as them joining a thematic group or asking for information in a forum. In fact, the intention is not to become a good influencer per se, but to establish a network of influencers to send each message to the appropriate recipients at the appropriate times.

Influencers have become powerful opinion creators, sometimes with negative consequences such as the dissemination of fake news and the introduction of bias during political campaigns. But, is it possible to regulate their ability to influence without threatening their freedom of speech? We often feel manipulated and bombarded by the presence of customised advertisements in every application we use, yet we naively give them our profile data. This is especially so among young people, who too often are not aware of the potential misuse of the information they share over social networks, which can be used to bully, blackmail, or intimidate them, or might simply play against them when they are looking for a job. Most computer attacks and social network threats are made anonymously and they obviously have a much farther reach than attacks that might occur in the physical world.

But these dangers cannot cover up the immense possibilities offered by social networks. For instance, in collaborative work. Take time banks, in which different professionals exchange their services, as an example; or Google translate, which, after a translation, asks the

Scientific Culture Unit of the University of Oviedo - FICYT

Little by little, so-called social robots are entering our everyday environment, assisting elderly or disabled people, as guides at festivals and in museums, or even as babysitters and teaching assistants, among other tasks. Above, an Eldertoy test session with elderly people, focused on improving the psychosocial, affective, and cognitive abilities of its users. The picture on the right shows a workshop where children can learn to program the NAO robot (Softbank Robotics).

Pixabay

Many social network users, especially young people, are not aware of the potential misuse of the information they share, which can be used to bully, blackmail, or intimidate them, or can simply play against them when they are looking for a job.

Then, what guidelines should be followed to design a game leading to social change? Swain (2007) suggests the following: precisely defining the expected results, including experts in the discipline in the design process, association with like-minded organisations, building a sustainable community, addressing problems without clear rules, maintaining journalistic integrity, measuring the transfer of knowledge, and making the game enjoyable.

However, most video games only try to look for entertainment without appealing to the players' genuine interest in solving problems. According to data from the Entertainment Software Association, a 21-year-old American will have spent an average 10,000 hours playing computer games. This is equivalent to five years working 40 hours a week. Unsurprisingly, someone saw an opportunity to benefit from so many hours spent playing. Luis von Ahn and Laura Dabbish were the first to design compelling games that, as a by-product of the gameplay, solved problems or generated data to train machine-learning algorithms (Von Ahn & Dabbish, 2008). An example of this is *ESP*, which led to Google's picture-labelling technology. In this game, when presented with a picture, players must type the names of the objects they see as fast as they can; coincidences between players determine the best tags and rankings, which were used to assign them rewards. An average of 233 tags per hour were submitted during an average 91 minutes per session. Several design strategies were used to motivate players: time limits, levels, scoreboards for different categories, and randomisation of the pictures presented.

Mechanisms to ensure the quality of the resulting tags, and to prevent cheating, have also been implemented; for example, to prevent several users scheming to obtain the maximum number of tag coincidences without being faithful to the picture content.

These types of games, called *games with a purpose*, have led us to today's ever-present gamification of society. This raises a question: is it ethically acceptable to design games or technology to create addictions and then make profits from them?

Stephe Chin

user for improvement feedback which is added to its statistical data so that similar future translations can be perfected. Social-impact games such as the popular *Evoke*, which urges players from around the world to solve challenges together, are also noteworthy, especially because of their influence on African communities. Participants are encouraged to learn about developing countries, act on that knowledge, and use creativity to imagine a better future for the planet. Using a comic to advance the storyline, players face challenges such as fighting against world hunger, using renewable energies, empowering women, or designing a plan for equal access to drinking water. To achieve the goal, players can use «superpowers» such as collaboration, courage, ingenuity, and entrepreneurship. Other similar games are *A force more powerful*, in which players must invent nonviolent resistance strategies to overcome oppression in a community, or *Participatory Chinatown*, which tries to help the residents of a Boston district to improve the future development of their neighbourhood. Many of these games allow users to learn about social problems, but until now very few have achieved measurable results in the real world.

«SOME OF THE TOPICS ADDRESSED IN CLASSICAL WORKS BY ASIMOV, DICK, OR BRADBURY HAVE GAINED TRACTION TODAY THANKS TO THE DEVELOPMENT OF SOCIAL ROBOTS»

While in the past, technological evolution tended to be ahead of its social implications, now that innovations are constant and become integrated into our daily lives in the blink of an eye, we could say that we are participating in a worldwide experiment without any prior impact study. It is difficult to predict – in a substantiated way – the influence that hyperconnectivity and our growing interaction with machines will have on the evolution of society, economy, and people's lives. Therefore, when trying to establish an ethical debate, we often resort to science fiction. If we do not have accurate models to make reliable predictions, a reasonable option is to imagine potential future scenarios, discuss their pros and cons, and then make our minds up in a well-argued way.

There are several works of fiction that successfully deal with emotional, psychological, and social issues connected to computer technology, and they all contribute to the discussion. I would highlight the TV series *Black Mirror*, the movie *Her,* and the novel *The lifecycle of software objects*, to name but a few. I also wanted to contribute to the debate with the novel *Enxarxats* ("Networked", Torras, 2017), touching upon some of the topics mentioned above, such as the persuasion strategies used by influencers and the design of computer games with social impacts, as well as others related to the Internet of things, emotional problems derived from excessive exposure to mechanical companions, and the creation of an avatar compiling our contributions to the network in order to guarantee a kind of digital immortality (Bearne, 2016). Aside from the novel's fiction content, it also includes an appendix with links to websites and abstracts of interest to readers who would like to know more about these topics.

Together, these are tools that we have at hand and that can, in a short time, flip someone's reputation, transform a district, modify the job market, and change our relationships – not just at work, but also within our families and our personal relationships – or change what a person leaves behind after they die, which now includes a digital footprint. We must consider that every contribution we make to the world wide web has an impact, and that programs that learn from people, just as Google translate does, transfer the responsibility for them functioning well or poorly from

Annapurna Pictures / Warner Bross
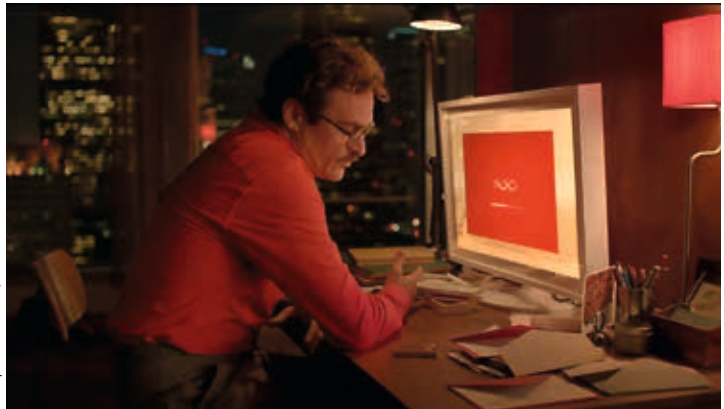
Annapurna Pictures / Warner Bross

Zeppotron / House of Tomorrow

Science fiction is often used to try to establish an ethical debate about the influence that hyperconnectivity and our growing interaction with machines will have on the evolution of society. Several works of fiction successfully deal with emotional, psychological, and social issues connected to computer technology, and contribute to the discussion. The film *Her* (above) is a good example. In the film, the main character falls in love with a computer operating system. The TV series *Black Mirror* has also explored this topic in some of its episodes, like «*Nosedive*» (on the right), in which people evaluate each other through a mobile phone app, and their rating directly affects their economic and social status.

«IS IT ETHICALLY ACCEPTABLE TO DESIGN GAMES OR TECHNOLOGY TO CREATE ADDICTIONS AND THEN MAKE PROFITS FROM THEM?»

These so-called «social robots» are the product of a series of advances in mechanics, visual and acoustic processing, adaptive control, computation, and artificial intelligence (Torras, 2016) that enable them to interact with people in a friendly and safe way, an aspect which is critical and very technical, especially when the interaction involves physical contact. One of the greatest current challenges is to provide these social robots with the ability to learn so they can adapt to different users and changing environments, and to unforeseen situations. Advances in this direction will surely yield more useful and versatile robots but will also intensify the debate about whether robots should be given more autonomy and decision-making abilities, not only in critical contexts such as the military or medical fields, but also in assistive and educational tasks. As we mentioned in the introduction, an elderly person with a mild cognitive impairment might think that the robot looking after them actually cares for their well-being and consequently, delegate all decision-making tasks to it. Or a child who is too attached to a robotic companion might not correctly develop empathy.

To address this sort of questions, the robotics community has approached humanities fields and many initiatives have started in two major areas: legal regulation and ethical education. Regarding the first, institutions such as the European Parliament, the South Korean Robot Ethics Charter, the IEEE Standards Association, and the British Standards Institution are developing regulations for robot designers, programmers, and users.

Ethical education is a broad field ranging from texts for high school students and online courses for the general public to materials for professional improvement and, especially, university books and papers. Prestigious associations such as the Institute of Electrical and Electronics Engineers (IEEE) and the Association for Computing Machinery (ACM) include a subject devoted to the ethics applied to technology in their engineering curricula, in which they increasingly deal with ethical issues in robotics, in a discipline known as *roboethics* (Veruggio, Operto, & Bekey, 2011; Lin, Abney, & Bekey, 2016). The most hotly debated area are aspects that affect

the programmers to the users. In this sense, Microsoft's experience with the Tay *chatbot* – based on machine learning techniques – in Twitter is revealing. In less than 24 hours they had to remove it because of the racist and sexist comments it had learned by chatting with humans (Hunt, 2016).

■ SOCIAL ROBOTS AND THE TECHNOLOGISTS' APPROACH TO HUMANITIES

Half a century ago, robots were first used in the production lines of the automotive industry to perform repetitive or dangerous tasks, often caged for security reasons and after being programmed by experts. Currently, the need for labour has moved from factories to the sectors of health and services, and robots are being designed to perform a variety of tasks in human environments, instructed by non-experts.
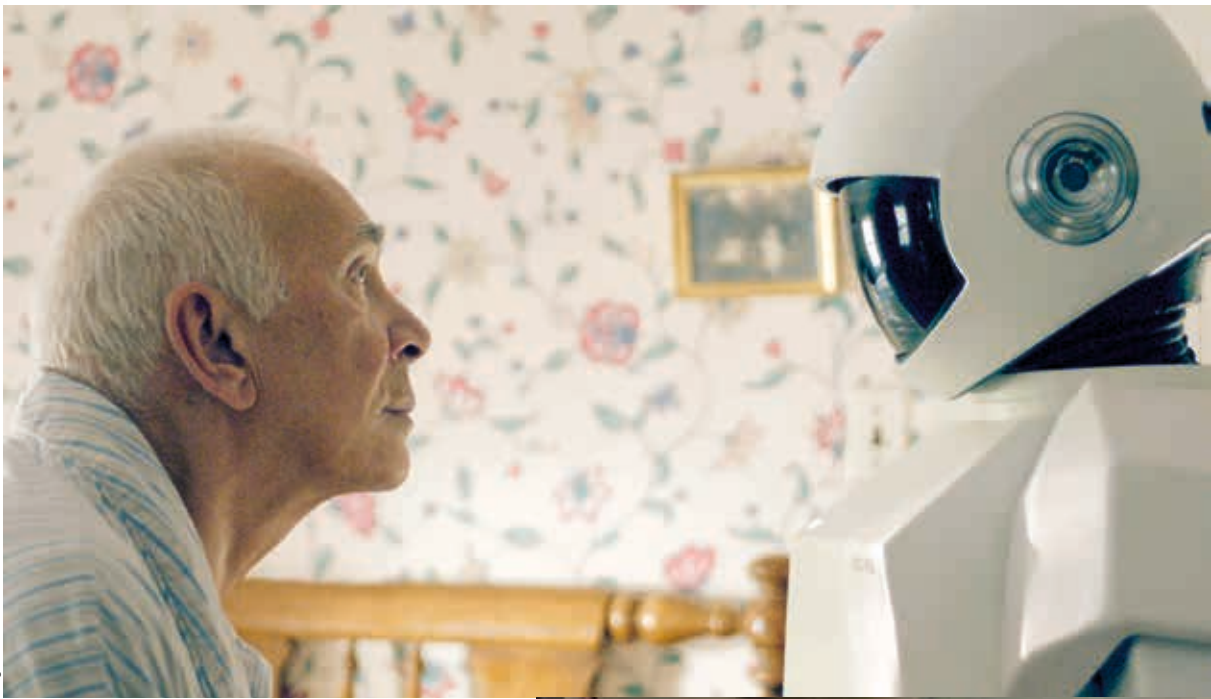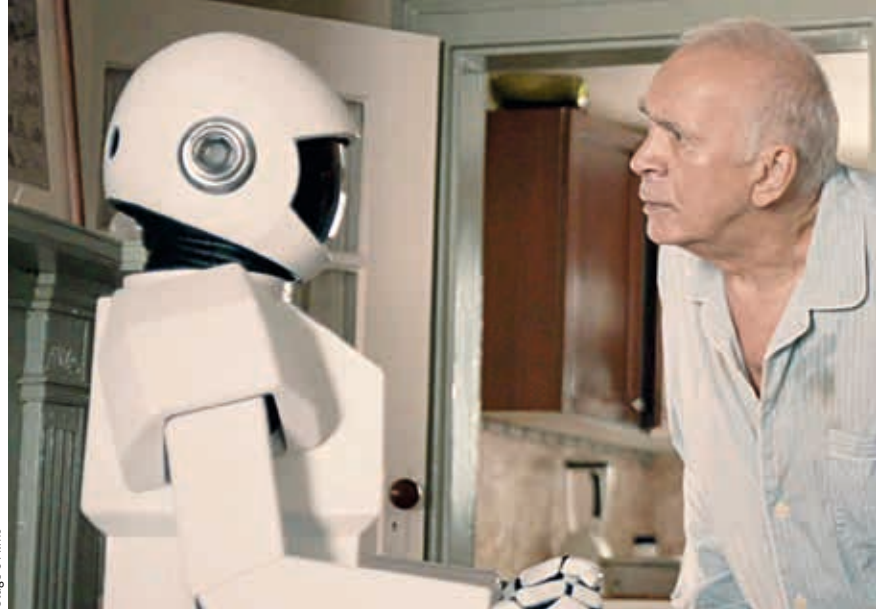
Stage 6 Films

Fiction has also dealt with ethical questions related to the introduction of social robots in everyday life. For instance, in the film *Robot and Frank*, the son of a retired and senile jewellery thief decides to buy a robot caretaker, so he does not have to watch his father all the time. Frank becomes more and more attached and trusting of the robot and the film presents different reactions to this in the old man's environment.

the military and medical fields, as well as privacy, legal responsibility, and the digital divide. Emotional, psychological, and social questions like the ones mentioned above are only now starting to arise with the emergence of assistive robotics.

As occurred with the Internet and social networks, science fiction is also often used in this area when trying to establish an ethical debate or teach a course. Some of the topics addressed in classical works by Asimov, Dick, or Bradbury such as the three laws of robotics, mechanical nannies, or humanoid replicas, have gained traction today thanks to the development of social robots. Recent films and TV series have also debated roboethics topics and are used in courses, both online and in person. I would like to highlight the TV series *Real humans* (where almost human-like robots coexist with people and often compete with them), the film *Surrogates* (in which every citizen has an avatar controlled from home that moves around the city and interacts with people), and the novel *The windup girl* (in which a robot becomes aware that it

Stage 6 Films

«AN ELDERLY PERSON MIGHT THINK THAT THE ROBOT LOOKING AFTER THEM ACTUALLY CARES FOR THEIR WELL-BEING AND DELEGATE ALL DECISION-MAKING TASKS TO IT»

was built to serve people and wonders about its rights and duties). The film *Robot and Frank* – showing the relationship between an old man, Frank, and its robotic caregiver – deserves a special mention for its realism and educational value and was the basis for an online course on the *Teach with movies* website, among other places.

In the context of university education, my novel *La mutació sentimental* (Torras, 2008) was translated into English with the title *The vestigial heart* (Torras, 2018) and has been published together with some ethics materials to teach a course about ethics in social robotics and artificial intelligence. The objective was to provide useful guidelines for students and professionals (robot designers, manufacturers, and programmers), as well as for users and the general public. It addresses six major issues: how to design the «perfect» assistant; the importance of robot appearance and the simulation of emotions for the acceptance of robots; automation in work and educational environments; the dilemma between automatic decision-making and human freedom and dignity; and civil responsibility related to programmed «morals» in robots. Each topic was developed based on scenes from the novel which depicts a future society in which each person has a robotic assistant and tells the story of a teenager, cryopreserved in our time because of an incurable disease and brought back to life in the future. This leads to conflicts with future humans who have been raised by artificial nannies, have learned from robotic teachers, and share work and leisure time with humanoids.

■ CONCLUSION

The growing interconnectivity between people in social networks and with all sorts of devices and robots in everyday life poses several challenges – both technical and scientific as well as in the humanities and social sciences – with a lot of potential to substantially shape the future and which are fostering an interesting social and ethical debate.

For example, some of the issues that we must consider are: how to prevent *chatbots* and robots from being misled by living beings or the most vulnerable groups from blindly relying and delegating all decision-

making to them? Will the excessive exposure of children to mechanical companions make it difficult for them to develop empathy and other affective skills? In the case of medical or security data, when must the common good prevail over personal data privacy? How can automatic decision-making be made compatible with human freedom and dignity? Is it permissible to allow the design of devices or programs that create addiction and dependence? How can we prevent such devices from being used to control people? Should we limit the power and opportunities for manipulation that influencers have? How can we close the digital divide and other social divides resulting from it?

Philosophy, psychology, and law are providing perspectives and prior knowledge to the debate, while science fiction allows us to freely speculate upon potential scenarios and the role that humans and machines will play in the *pas de deux* that irredeemably connects us. ⊙

«PROGRAMS THAT LEARN FROM PEOPLE TRANSFER THE RESPONSIBILITY FOR THEM FUNCTIONING WELL OR POORLY FROM THE PROGRAMMERS TO THE USERS»

REFERENCES

Bearne, S. (2016, 14 September). Plan your digital afterlife and rest in cyber peace. *The Guardian*. Retrieved from https://www.theguardian.com/media-network/2016/sep/14/plan-your-digital-afterlife-rest-in-cyber-peace

Hunt, E. (2016, 24 March). Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter

Lin, P., Abney, K., & Bekey, G. (2011). *Robot ethics: The ethical and social implications of robotics*. Cambridge, MA: MIT Press.

Swain, C. (2007). Designing games to effect social change. In A. Baba (Ed.), *Proceedings of the 2007 Digital Games Research Association Conference* (pp. 805–809). Tokyo, Japan: Digital Games Research Association.

Torras, C. (2008). *La mutació sentimental*. Lleida, Spain: Pagès Editors.

Torras, C. (2016). Service robots for citizens of the future. *European Review, 24*(1), 17–30. doi: 10.1017/S1062798715000393

Torras, C. (2017). *Enxarxats*. Barcelona: Males Herbes.

Torras, C. (2018). *The vestigial heart. A novel of the robot age*. Cambridge, MA: MIT Press.

Veruggio, G., Operto, F., & Bekey, G. A. (2016). Roboethics: Social and ethical implications of robotics. In B. Siciliano, & O. Khatib (Eds.), *Handbook of robotics, 2nd edition* (pp. 2135–2160). Berlin-Heidelberg: Springer.

Von Ahn, L., & Dabbish, L. (2008). Designing games with a purpose. *Communications of the ACM, 51*(8), 58–67. doi: 10.1145/1378704.1378719

Wu, M. (2012, 15 September). My chapter on influencers. *Lithium Community*. Retrieved from https://community.lithium.com/t5/Science-of-Social-Blog/My-Chapter-on-Influencers/ba-p/8213

Carme Torras. PhD in Computer Science and Research Professor at the Robotics Institute (CSIC-UPC, Barcelona, Spain), where she leads a research group in assistive and collaborative robotics. She has directed sixteen European projects, among which CLOTHILDE, a European Research Council Advanced Grant dealing with wearable robotics, stands out. She believes science fiction can help promote ethics in robotics and new technologies. She wrote the novels *Enxarxats* (Males Herbes, 2017) and *La mutació sentimental* (Pagès Editors, 2008) which won the Pedrolo and Ictineu awards and was translated into English with the title *The vestigial heart* (MIT Press, 2018). ✉ torras@iri.upc.edu