Elements of Roboethics

Fiorella Operto

Abstract Roboethics analyzes the ethical, legal and social aspects of robotics, especially with regard to advanced robotics applications. These issues are related to liability, the protection of privacy, the defense of human dignity, distributive justice and the dignity of work. Today, roboethics is becoming an important component in international standards for advanced robotics, and in various aspects of artificial intelligence. An autonomous robot endowed with deep learning capabilities shows specificities in terms of its growing autonomy and decision-making functions and, thus, gives rise to new ethical and legal issues. The learning models for a care robot assisting an elderly person or a child must be free of bias related to the selected attributes and should not be subject to any stereotypes unintentionally included in their design. As roboethics goes hand in hand with developments in robotics applications, it should be the concern of all actors in the field, from designers and manufacturers to users. There is one very important element in this—albeit one that is related indirectly—that should not be overlooked: namely, how robotics and robotic applications are represented to the general public. Of the many representations, the legacy of mythology, science fiction and the legend still play an important role. The world of robotics is often marked by icons and images from literature. Exaggerated expectations of their functions, magical descriptions of their behavior, overanthropomorphization, insistence on their perfection and their rationality compared to that of humans are only some of the false qualities attributed to robotics.

Keywords Service Robotics · Field Robotics · Roboethics · ELS (Ethical · Legal · Societal) Issues

F. Operto (✉)
Scuola Di Robotica, Genoa, Italy
e-mail: operto@scuoladirobotica.it

D. Scaradozzi et al. (eds.), *Makers at School, Educational Robotics and Innovative Learning Environments*, Lecture Notes in Networks and Systems 240,
https://doi.org/10.1007/978-3-030-77040-2_10

1 The Birth of Roboethics

This article outlines some of the main lines of development and application in roboethics—that is, applied ethics in advanced robotics—which examines the ethical, legal and societal issues (ELS) inherent to the field. Roboethics was born—as a term and as applied ethics—in 2004, during the First International Symposium on Roboethics [2]. Roboethics analyzes the ethical, legal and social aspects of robotics, especially in relation to service and field robotics applications. These issues are related to the protection of privacy, the defense of human dignity, distributive justice and the dignity of work. Today, roboethics is the subject of hundreds of studies, applications, research, and is becoming an important component in international standards for advanced robotics, and also in various aspects of artificial intelligence. Robots are certainly formidable tools. There is no aspect of our private and social lives that cannot be improved by the introduction of robots. However, technology applied to human life always raises ethical questions. In the case of robotics, especially service robotics, these ELS issues are novel, emerging, complex and involve several disciplines.

2 A New Science?

Robotics is a field of research and application, or a new science, still in its infancy, born from the fusion of many disciplines within the humanities and natural sciences. Here, the sum is greater than the parts [2]. It is a very powerful tool for studying and increasing our knowledge, not only of the universe around us—space, oceans, our body—but also our brain/mind. This is why robotics can lead to a convergence of the so-called *two cultures*: human sciences and natural sciences. Robotics covers the following disciplines: mechanics, automation, electronics, informatics, cybernetics, physics/mathematics, artificial intelligence, and draws contributions from (and, in fact, is *invading*) logic/linguistics, neuroscience/psychology, medicine/neurology, biology/physiology, psychology, anthropology/ethology, art/industrial design.

This complexity—involving novel aspects such as ESL issues, which drive research and robot production—is giving the subject the caliber of a science, with its laws, coherent and comprehensive understanding of nature and predictive capabilities. For an overview of the state of robotics today we recommend the *Springer Handbook of Robotics* [3].

Moreover, the object of the research and development of advanced robotics, an autonomous robot endowed with learning capabilities, shows specificities in terms of its growing autonomy and decision-making functions.

3 What Ethics Should Be Applied in Roboethics?

We have different *versions* of roboethics depending on which ethical theories are adopted (utilitarianism, deontology, virtue ethics, rights ethics, Rawls' theory of justice). Yet, in all these versions, a logical and critical framing of ethics is needed, one that reveals the implicit, uncovered assumptions, and analyzes the reasons, the pros and cons, and their origin. This frame also allows us to define (the extent and limits of) human liability and machine autonomy, in cases of damage caused by a learning robot.

In addition, in the light of more complete ethical theories, the ethical framing should assess whether, according to distributive justice principles, the actors involved should be socially duty-bound to compensate for the dramatic changes caused by a heavy, rapid and unnegotiated introduction of robots to our society: job displacement and loss; privacy issues and encroachment on personal life; technological dependency; robotics divide (in terms of generations, social status, and areas of the world).

Finally, roboethics should cover a series of positive recommendations and rules that would be implemented in all contexts where robots are introduced. These should be along the same lines as the recent prescriptions being adopted for the production and use of plastics, energy, and other industrial sectors. In roboethics, analysis of ELS issues often leads to recommendations which, in many cases, have been submitted to the UN, European governments, the European Commission and the European Parliament [4].

In light of the lessons learned from the COVID-19 pandemic, we cannot afford to introduce large-scale technological applications into society without offsetting the ensuing imbalances in the environment and the disruption to social groups.

4 Emerging and Novel Roboethical Issues

Since 2004, several authors have intervened in the debate on roboethics to highlight certain ELS issues that have arisen over a very long period of time. Issues such as the rights of and the payment of and for robots and their status as moral agents can be interesting to discuss, but too far-reaching. They also do not consider the urgency of and the need for addressing ethics-related technical issues.

Given the rich and complex debate on roboethics and the sometimes unknown developments that could come over the next two decades, the author, the partners and experts of the Ethicbots European Project [5] adopted a triaging system to analyze the following issues:

Novelty: Issues that have never been looked at; the *absentia legis* and the lack of regulations is, in many cases (bionics and military robotics), evidence of a severe responsibility gap.

Emerging: Issues that are not planned for, since robotic prototypes are the result of different forces: research and business.

Complexity: Issues lying at the intersection of several disciplines (robotics, AI, moral philosophy, psychology, anthropology, law).

Social pervasiveness: Issues related to current and yet-to- be-released robotic products [6].

The sectors most directly and urgently interested in robotic applications are the military and certain areas of biomedicine (invasive prosthetics) [7, 8].

5 The Risk of Unintended Machine-Learning Bias

Issues of bias in artificial intelligence are well-known to scientists. Machine-learning models are developed to be predictive, when large datasets teach the robot learning models to predict the future, based on the past. Trained models can read and use an incredible amount of data (texts, pictures, software, other models), consuming it to identify the data patterns considered most suitable for carrying out the mission. In this way, predictions can be more accurate than with simple built-in models. The bias issue is related to the fact that machine-learning models can predict precisely what they have been trained to predict, and their predictions are as accurate as the data used to train the machine. Any errors are explained in the maxim "garbage in, garbage out." In fact, many cases of bias detection, which range from the light to the heavy involving AI products, stem from human bias intervening during the creation of data models. Either the data collected were unrepresentative of reality—as in the *portability issue*, when a model is employed out of context—or they reflect existing human prejudices—for instance, when certain attributes in the model are selected or ignored [9].

The learning models for a care robot assisting an elderly person, a child, or in a hospital must be free of bias related to the selected attributes (e.g., culture, gender, social or economic status, linguistic attributes) and should not be subject to any stereotypes unintentionally included in their design. It is easy to imagine how complex this process could be in a learning robot, especially since bias detection cannot be performed at the expense of the assisted person.

6 Ethical Guidelines for All Robots

In a review article written by Matthew Studley and Alan Winfield on the ESL aspects of industrial robots [10], the authors came to an interesting conclusion after reviewing around 84 papers on the topic: today, even robots used in industrial production are subject to similar ELS problems to those found with service robots, which means that the gap between industrial robots and other types is narrowing.

Industry is changing; converging technologies have ushered in a fourth Industrial Revolution, where new collaborative robots, or *cobots*, work alongside humans on common tasks. Unlike more common industrial robots, which largely work alone and unsupervised, collaborative robots are programmed and designed to work with humans, responding to human behaviors and actions. The authors of that review article point to the increasing importance of human–robot interaction (HRI) and the reduced differentiation between industrial robotics and other robot domains affected by the definition and range of ELS issues. In this, advanced industrial robotics may be affected by the same sorts of concerns that are faced in assistive robotics: predicting and interpreting human intentions and future actions in order to perform efficiently. Here, the interactions between humans and robots involve teaching rather than programming. The ELS issues that affect learning cobots include psychological and sociological impacts, liability, data and privacy. Cobots can be programmed for the speed, tasks and precision to which humans have to adapt. In addition, cobots can be reprogrammed rapidly for another task, forcing humans to make rapid changes with no time to adapt. Cobots can gather data about the pace of work, abilities and needs of their human co-workers. These data may be processed in cloud services and could be used by other organizations. Use of the resulting data profiles could breach European privacy legislation.

7 Representation of Robots with the General Public and Agnotology Issues

A full analysis of roboethics should not overlook another element that is indirectly related, but has a great impact on it: how robotics and robotic applications are represented to the general public. Among the many representations, the legacy of mythology, science fiction and the legend still play an important role [12]. The world of robotics is often marked by icons and images from literature. Exaggerated expectations of their functions, magical descriptions of their behavior, over-anthropomorphization, insistence on their perfection and their rationality compared to that of humans are only some of the false qualities attributed to robotics.

The word ‘agnotology’ was coined to refer to the study of culturally induced ignorance, and specifically its implications for individual and collective decision-making [13]. The continuous and massive dissemination of inaccurate information about robots in the media can hinder the public’s understanding of the fundamental ELS aspects of robotics that already affect quality of life. The obfuscation and omission of basic, relevant, and accurate information about robotics and roboethics issues can only result in wrong assumptions which affect the public’s ability to take part in the collective decision-making process.

8 Conclusions

We have briefly described only some of the elements of roboethics that should be studied by those who program, produce, market and use robots. Roboethics will have an impact on the design, programming, shape and use of robots. It should be included in engineering and architecture programs, as well as in the various disciplines of the humanities. Also, it is important to build trust between the general public and robotics laboratories through honest, concerted information campaigns.

References

1. Veruggio, G., Abney, K.: Roboethics: the applied ethics for a new science. In: Lin, P., Abney, K., Bekey, G. (eds.) Robot Ethics: The Ethical and Social Implications of Robotics. MIT Press, Cambridge, MA (2018)
2. Veruggio, G., Operto, F.: Roboethics: social and ethical implications of robotics. In: Handbook of Robotics. Springer, Heidelberg (2008). www.roboethics.org
3. Siciliano, B., Khatib, O.: Springer Handbook of Robotics. Springer, Berlin, Heidelberg (2016)
4. Resolution of the European Parliament: A comprehensive European industrial policy on artificial intelligence and robotics (2019): https://www.europarl.europa.eu/doceo/document/TA-8-2019-0081_EN.html.
5. ETHICBOTS project: Emerging Technoethics of Human Interaction with Communication, Bionic, and robOTic systems. (2005, Nov. 1–2008, Apr. 30). Coordination action, FP6—Science and society. Available: http://ethicbots.na.infn.it/ (2008).
6. Operto, F.: Ethics in advanced robotics. IEEE Robot Autom Mag Special Issue on Robot Ethics **18**(1) (2011)
7. Sharkey, N.: Automating warfare: lessons learned from the drones. J. Law Inf. Sci. **21**(2), 140 (2012)
8. Sharkey, A.: Autonomous weapons systems, killer robots and human dignity. Ethics Inf. Technol. **21**, 75–87 (2019). https://doi.org/10.1007/s10676-018-9494-0.(2019)
9. Battistuzzi, L., et al.: Embedding ethics in the design of culturally competent socially assistive robots. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, 1–5 Oct 2018, Madrid, Spain (2018)
10. Studley, M., Winfield, A.: ELSA in industrial robotics. Curr Robot Rep (2020). https://doi.org/10.1007/s43154-020-00027-0.(2020)
11. Tamburrini, G.: On the ethical framing of research programs in robotics. J. AI Soc. (2020). https://doi.org/10.1007/s00146-015-0627-2
12. Operto, S.: Evaluating public opinion towards robots: a mixed-method approach. Paladyn **10**(1):286–297 (2019). https://doi.org/10.1515/pjbr-2019-0023
13. Proctor, R.N.: Londa Schiebinger. Agnotology: The Making and Unmaking of Ignorance. Stanford University Press (2008)