SOCIAL ROBOTS: MAIN ETHICAL CHALLENGES AND ISSUES

Áurea Subero-Navarro, María Alesanco-Llorente, Jorge Pelegrín-Borondo y Mario Arias-Oliva

Universidad de La Rioja (Spain), Universidad de La Rioja (Spain), Universidad de La Rioja (Spain), Universidad Complutense de Madrid (Spain)

aurea.subero@unirioja.es; maria.alesanco@alum.unirioja.es; jorge.pelegrin@unirioja.es; mario.arias@ucm.es

INTRODUCTION

In recent years, the great advances that have been made in different disciplines such as computer science and mechanics have enabled the development of robots created not only for the realization of industrial tasks, but also to interact with people in healthcare environments, providing different services such as care of the elderly people, advisory tasks in commercial environments, medical tasks, etc. (Torras, 2014). These types of robots are called social robots and their rise has boosted the interest of researchers.

Today the implementation of robots is taking place practically in every area of society. However, the discussion over the use of robotics and artificial intelligence (AI) has increased as the possible consequences of the use of robots for the economy, employment and society are beginning to be seen (Huang and Rust, 2018). These future changes in society generate both, expectations and fears (Mick and Fournier, 1998). That is why in a society where AI is becoming more relevant, it is necessary to define a number of limits in relation to the mode of use of AI and the way of interaction with human beings (Santos González, 2017).

Technoscientific and technological development implies the introduction of improvements in human life but also implies the emergence of numerous risks in which it is necessary to influence. One of the relevant issues to address is the impact of ethics on technoscience, that is to say, how to apply ethics to emerging new technologies. This is because any cultural change requires, in turn, an ethical rethink to avoid unwanted situations in the future (Valls Prieto, 2019).

It is therefore necessary not only to anticipate the possible scenarios that may occur but also to identify the future moral problems that will arise from them. There are numerous initiatives and projects related to roboetics that aim to answer all these questions (Torras, 2014).

The future will be conditioned by the decision-making that takes place at this time and not only by the legislator, but also by each of the agents who have the capacity to influence them. It is therefore necessary to regulate both individual and collective behaviour in relation to minimum requirements and to carry out a periodic evaluation of it. How the sharing of benefits and costs is appreciated will obviously influence the blessing of consumers. The introduction of these technological developments does not necessarily imply inequality, as long as these technological tools are subject to values and standards (Grau Ruiz, 2019).

Another main challenge of ethics or theft is based on the use of language. Currently, the terms used can be applied exclusively to humans and not to robots. Decision-making is completely different and contingencies arise in terms such as "consciousness" or "empathy" that constitute human realities and are mis-applied to machines. It will be necessary to introduce terms that adapt to robotics. The risk of this is very great because if robots are equated to humans, there is a risk that machines will resemble people (Noeo Tech, 2017).

Therefore, it will be essential to implement an ethical system that promotes the ethical behavior of the different agents participating in it. However, the main problem in relation to this issue is that human beings usually know what is ethical, but not how to achieve it, so this supposes the main challenge in the field of Philosophy and Law (Grau Ruiz, 2019).

In this regard, the European Parliament in 2017 approved the robotics report establishing an ethical code of conduct. Its purpose is to serve as: "an ethical guidance framework for the design, production and use of robots". All this has led to the emergence of a new term coined by Gianmarco Veruggio: robotics (Noeo Tech, 2017).

In addition, it is also necessary to influence the opinion of the population in relation to the use of these robots. Ethics reflects the plurality of legal, moral and religious norms that govern in a community (Berger et al., 2008). Reidenchach and Robin (1990) consider that individuals use more than one basis to make ethical judgments so that the employment of multiple dimensions is necessary to capture the meaning of ethical judgment. Therefore, after the revision of literature dedicated to moral philosophy, they developed the Multidimensional Scale of Ethics (MES) which includes five dimensions based on contemporary normative moral philosophies: moral equity, utilitarianism, relativism, selfishness and contractualism.

This theoretical framework is taken as a reference. The objective of the present research is to find out if the ethical action of consumers in relation to the use of robots takes some place in the minds of the consumers involved in it and analyze the main ethical issues of the introduction of social robots in our lives. Based on this theoretical background, the authors propose to advance in the knowledge of the impact of ethical judgment on the intention of using social robots in commerce. To do this, the respondents were asked their opinion about the advantages and disadvantages of using social robots in commerce.

MAIN ETHICAL ISSUES

The robotics industry is growing at a dizzying pace. Social robots are increasingly present in all areas: education, health, care of elderly people, customer service (Beer et al.2011; Conti et al.2019; Alaiad & Thou, 2014) and their adoption rates have accelerated exponentially reaching sales of service robots at annual rates above 30% (International Federation of Robotics, 2018).Mckinsey'report (2018) says that of the nineteen major industries, the retail industry has the greatest potential to create value through the use of autonomous technologies and AI which translates to more than 600 billion dollars annually.

In the light of these facts, it seems important to analyze the ethical issues raised by the use of social robots. These ethical issues regarding the use of robots and their impacts on society are the main subject of roboethics (Demir, 2017). Advanced robotics can create problems if people don't understand the consequences that can result from introducing an increasingly intelligent technology (Alsegier, 2016). Therefore, addressing the key principles of roboethics as they arise is essential to guarantee a correct symbiosis in human-robot interaction (Tsafestas, 2018). The need for ethical considerations in the development of these intelligent systems is becoming one of the main areas of research giving rise to different initiatives such as the IEE on the ethics of autonomous systems, the foundation for responsible robotics and the association on AI, among others (Dignum, 2018).

In this section, the main ethical problems of human use of social robotics will be described and the problems arising from the use of this emerging technology will be examined. The design of these systems is not only relevant in terms of their responsible use (Houkes & Vermaas, 2010), but also requires a responsible design. Among the main ethical dilemmas are:

- **Privacy:** Currently there is a discussion about the privacy and surveillance of information technology (Macnish, 2017), that is, access to private and sensitive data. So, security and data protection has become a relevant issue. The introduction of technology has accelerated exponentially while regulation has taken a long time to respond (Proposal for a Regulation laying down harmonised rules on artificial intelligence-Artificial Intelligence Act, 2021) (European Commission, 2021).

  However, currently robotic systems have not yet played an important role in the area of privacy. Nevertheless, this will change when they are introduced in different scenarios. Data protection is at risk with the rapid development of AI since it's use involves the processing of large amounts of data.

  Therefore, the Spanish Data Protection Agency (AEPD) has published a guide to adapt to the general data protection regulation (RGPD) for products and services that include AI components (AEPD, 2021).

- **Behavior manipulation:** Ethical issues in relation to AI are not only reduced to the use of data, but also include the use of that information to modify and manipulate human behavior (Dezfouli et al.2020). In this sense, a study conducted by Data61 researchers from the Commonwealth Scientific and Industrial Research Organization or CSIRO has concluded that AI can find vulnerabilities within human decision-making (Dezfouli et al.2020).

  According to Steinert, 2014 robots are mere extensions of human capabilities and can be used as tools to modify a situation according to human wishes. Robots are now considered amoral systems since technology is neutral in relation to use (Westerlund, 2020). That is, a robot can be used to perform surgery and save a life or, on the contrary, it can be used to kill someone as a result of human will (Steinert, 2014).

  In fact, robots could be used as killer weapons (Demir, 2017). However, even if a robot is built as an autonomous and intelligent system, the ethical concerns that arise from its usage are linked to the human design (Westerlund, 2020). Some researchers point out that robots are analogous to domestic animals in terms of liability (Kelley et al., 2020). In other words, if a robot is involved in an accident or a problem, the responsibility lies with the owner (Westerlund, 2020). For this reason, the protection of humanity is necessary in relation to protection against possible manipulation and in terms of responsibility.

- **Opacity of IA systems:** opacity and biases constitute the main ethical challenges of AI (Floridi & Taddeo, 2016; Mittelstadt & Floridi, 2016). One of the characteristics of algorithms is the opacity ("black box") (Monasterio Astobiza, 2017). Although they are hidden, they are invisible, because they are inscrutable between the layers and sub-layers of computer programming. They are opaque in the sense that they are hermetic to interpretation (Monasterio Astobiza, 2017).

  So, there is concern regarding the use of AI systems for automated decision support and predictive analytics in relation to the lack of process and community involvement and auditing (Whittaker et al. 2018).

  Because of this, people fail to understand the basis for making a decision. To alleviate this darkness, a new discipline has been created: explainable AI that enables human beings to understand the decisions made (Barredo Arrieta et al.,2019).

For its part, the EU Regulation 2016/679 has taken this opacity into account and provides that when consumers are faced with a decision based on algorithmic processing, they have the right to a legal explanation of the decision.

- **Bias in decision systems:** Automated systems that carry out decisions and predictive analysis operate on the data and, based on it, generate a decision as an output. Algorithms influence important decisions in people's daily lives, so mistakes should be avoided and they should be infallible and ethical. Choi et al., 2010 analyze biases and group them into three broad categories:

  A) Biases derived from problems with the drafting of the question.

  B) Biases derived from problems with the design and layout of the questionnaire.

  C) Biases derived from problems with the use of the questionnaire.

  However, when this problem is limited to the context of AI, it's necessary to specify it.

  According to Hao (2019) bias can occur at different times in the process, although he highlights the three stages where it usually arises: in the definition of the model, in the data collection and in the preparation of data for its use. Therefore, if the algorithm that's going to make the decision of a development is biased, the result will also be biased (Salazar & Escribano Otero, 2021). There are numerous examples of theses biases that have been seen in the press such as the following: "The Amazon algorithm that does not like women". This article refers to the Amazon algorithm that was used in the staff selection that rewarded men over women (Salazar & Escribano Otero, 2021).

  There are constant efforts to remove biases from AI systems. However, they are in the early stages of learning (Brownsword et al., 2017). Technological solutions have their limits in the sense that a mathematical notion of justice is required, which is very difficult to achieve (Selbst et al.2019).

To avoid these ethical dilemmas, the European Parliament has approved a report of guidelines for the use of AI in both, civil and military fields. The European Parliament has urged the European Commission to create a framework of principles and legal obligations on artificial and robotic intelligence. The proposal is based on the fact that AI must be governed by a series of requirements that guarantee security, transparency and accountability with safeguards against bias and discrimination. As well as the right to redress, social and environmental responsibility, privacy and data protection (Madiega, 2019).

METHOD

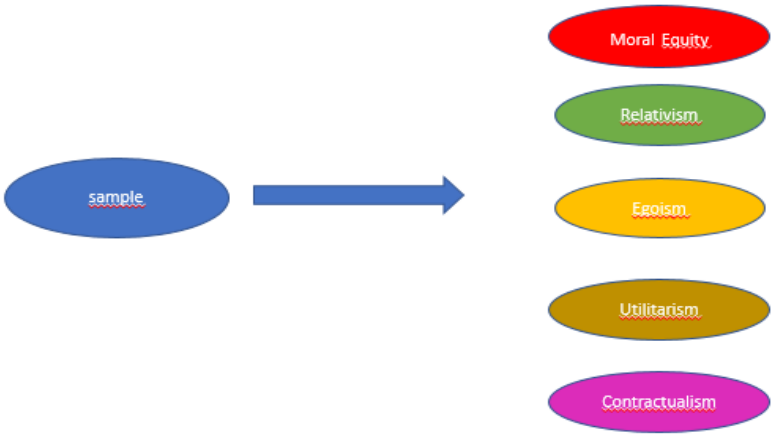
A personal survey has been applied to a sample of 100 individuals over 18 years of age residing in Spain. Regarding the characteristics of the sample, they are: Men: 48% and Women: 52% from 18 years to more than 65 years. The survey was based on two open questions about the advantages and disadvantages about the use of social robots in the commerce.

In terms of mechanics, the survey takers located the participants according to gender and age quotas. The interviews were personal and recorded to ensure the quality of the field work and the information.

Once the readings corresponding to the review of the survey applied to the sample had been carried out, the responses were categorized. In this sense, they were classified into categories according to

the thematic criteria related to the objectives of the research. The categories were the following: moral equity, relativism, selfishness, utilitarism and contractualism (Shawner & Senneti, 2009).

Model 1.



Source: Self-Elaboration.

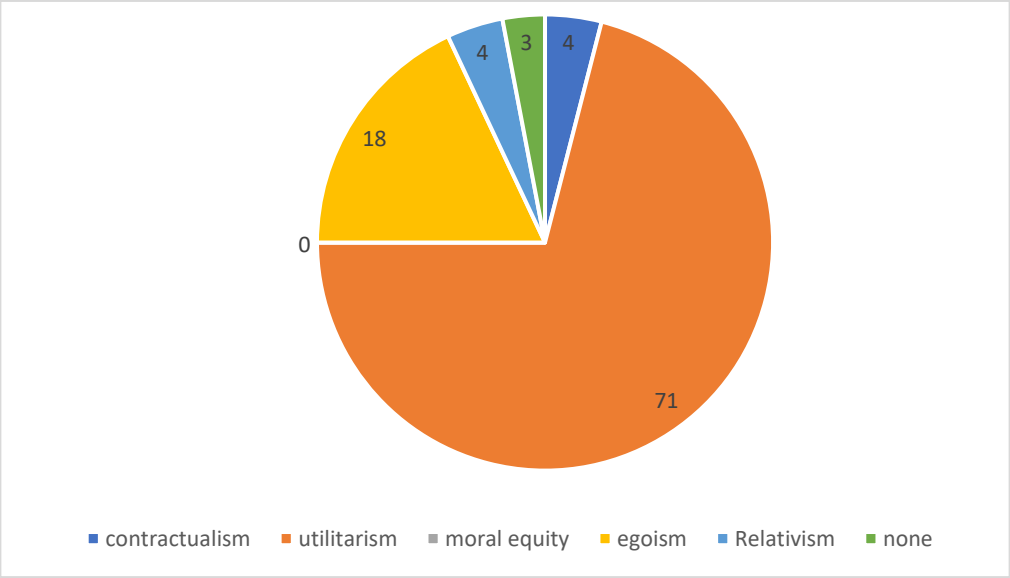
RESULTS

To achieve this goal, one hundred people have been interviewed. These people have been asked open questions about what they consider to be the main benefits and disadvantages of the use of social robots in retail (Shawver and Senneti, 2009). The results show the following:

- Utilitarianism: a dimension based on consequential theories that consider morality to be measured by the consequences of the actions performed (Reidenbach and Robin, 1990). The results show that when claiming advantages of the use of social robots in retail, the respondents made a balance of their costs and benefits and 71% of them were guided by utilitarian reasons such as speed of service, handling large amounts of information and the ability to easily locate products. On the other hand, only 31% alleged utilitarian reasons when alleged inconveniences.
- Moral Equity: dimension based on the theory of justice (Rawls, 1971). The results show that at the time of stating the advantages of the use of social robots in retail, none alleges reasons related to moral equity, while, for their part, in stating the disadvantages 59% alleged these reasons, such as the dehumanization of work, destruction of employment and increase in the number of people in unemployment.
- Selfishness: dimension also based on consequential theories, but focusing exclusively on the individual consequences (Reidenbach and Robin, 1990). The results show that 18% used selfish arguments such as "not having to be kind to dependents". For its part, when it comes to stating inconveniences, none alleged selfish reasons.
- Relativism: it is defined as the "perception of what is correct versus incorrect based on guidelines/parameters of the social/cultural system (Nguyen and Biderman, 2008). The results of the research show that when declaring the benefits of the use of robots in retail 4% of respondents alleged relativistic motivations by comparing the right and wrong of the use. They argued, for example, that it would be good because of the increased of efficiency while it wouldn't be correct because of the destruction of employment. For their part, when they exposed the drawbacks of the use of robots in retail, 9% made relativistic arguments.

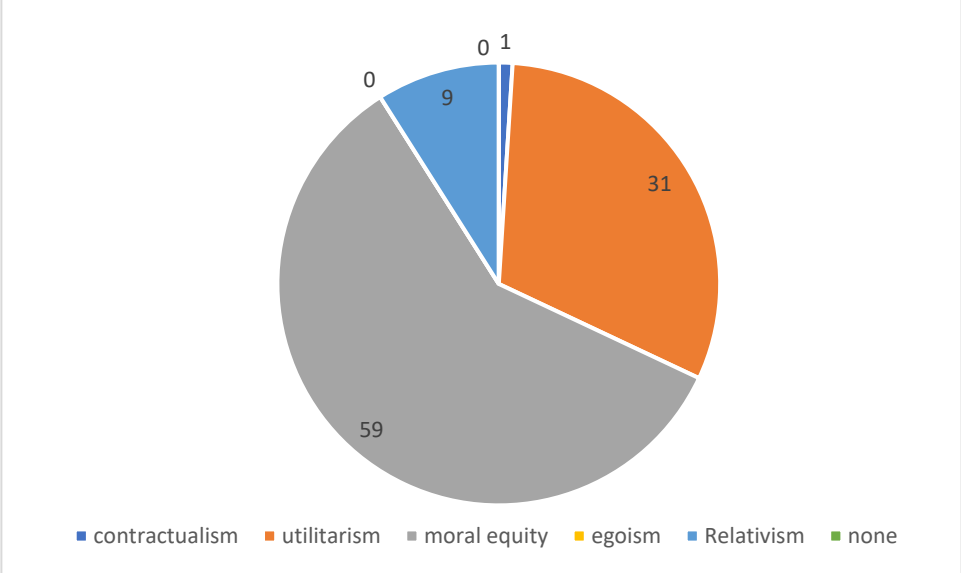
- Contractualism: a deontological dimension that encompasses different notions such as implicit obligation, contracts, duties and rules (Reindenbach and Robin, 1990). The results show that when declaring the benefits of the use of robots in retail 4% claimed contractual reasons, while when it came to stating the inconveniences 1% alleged such reasons.
- For their part, 3% of respondents said they found no benefit from the use of robots in retail.

Graphic 1. Results Advantages.



Source: Self-elaboration.

Graphic 2. Results Disadvantages.



Source: Self elaboration.

The influence of ethical judgment on the intended conduct is a conclusion from the investigation. When respondents were asked about the benefits of using robots, most of them alleged utilitarian reasons. It is clear that one of the strengths of robots are the great possibilities they offer us: they

enable economic growth and the development of traditional businesses, they accelerate and optimize business processes, they are a source of information and recommendation and constitute a source of entertainment for users. However, in exposing the inconveniences of their employment, 59% of people alleged moral equity reasons. This is mainly due to the fear generated by the introduction of robots into our lives due to process automation and job destruction. However, robots can be used with the aim of complementing humans and freeing them from the most routine tasks, rather than replacing humans.

ACKNOWLEDEGEMENTS

This research was funded by the bridge grants for research projects awarded by the University of La Rioja (PROYECTOS PUENTE PP-2020-02), subsidized by Banco Santander, the COBEMADE research group at the University of La Rioja and ayudas para estudios cientificos de tematica riojana del IER año 2020-2021.

KEYWORDS: social robots, ethics, artificial intelligence, ethical judgment.

REFERENCES

AEPD (2021). Requisitos en auditorías de tratamientos que incluyen Inteligencia Artificial. *AEPD*. Retrieved from: https://www.aepd.es/es/documento/requisitos-auditorias-tratamientos-incluyan-ia.pdf.

Alaiad, A., and Zhou, L. (2014). The determinants of home healthcare robots adoption: An empirical investigation. *International Journal of Medical Informatics*, 83(11), 825-840.

Barredo Arrieta, A., Díaz Rodríguez, N., Del Ser, J., Bennelot, A., Tabik, S., Barbado González, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, V.R., Chatila, R., & Herrera, F. (2019). Explanaible Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward responsable AI, Information Fusion, 58, 2-73.

Beer, J.M., Prakash, A., Mitzner, T.L., & Rogers, W.A. (2011). Understanding robot acceptance, technical Report HFA-TR-1103, Atlanta GA: Georgia Institute of Technology School of Psicology. Rerieved from: https://smartech.gatech.edu/bitstream/handle/1853/39672/HFA-TR-1103-RobotAcceptance.pdf;accessed

Berger, C., González-Franco, M., Ofek, E. & Hinckley, K. (2018). The uncanny valley of haptics. *Science Robotics,* 3 (17), 1-2.

Brownsword, R., Scotford, E., & Yeung, K. (2017). The Oxford Handbook of Law, Regulation and Technology, Oxford: Oxford University Press. http://doi.org/10.1093/oxfordhb/9780199680832.001.0001

Choi, B., Granero, R., & Pak, A. (2010). COMITÉ EDITORIAL Comunicación Especial. Revista Costarr Salud Pública, 19(2), 106–118.

Chui, M., Manyika, J., Miremadi, M., Henke, N., Chung, R., Nel, P and Malhotra, S. (2018). Notes from the AI Frontier: Applications and Value of Deep Learning. *Mckinsey & Company*. Retrieved from: https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-applications-and-value-of-deep-learning.

Conti, D., Cattani, A., Di Nuovo, S., and Di Nuovo, A. (2019). Are future psychologists willing to accept and use a humanoid robot in their practice? Italian and English Students Perspective. *Frontiers in psychology*, 10 (2138), 1-13.

Demir, K. A. (2017). Research questions in roboethics. *Mugla Journal of Science and Technology*, 3(2), 160- 165. http://doi.org/10.22531/muglajsci.359648

Dezfouli, A., Nock, R., & Dayan, P. (2020). Adversarial vulnerabilities of human decision-making. Data61, Commonwealth Scientific and Industrial Research Organization (CSIRO), 1-8. Retrieved from: https://www.pnas.org/content/pnas/117/46/29221.full.pdf

Dignum, V. (2018). Ethics in artificial intelligence: introduction to the special issue. *Ethics and Information Technology*, 20 (1-3), 1-3.

European Comission (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence-Artificial Intelligence Act. Retrieved from: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELLAR%3Ae0649735-a372-11eb-9585-01aa75ed71a1

Floridi, L. & Taddeo, M. (2016). What Is Data Ethics? *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2083). http://doi.org/10.1098/rsta.2016.0360

Grau Ruiz, M. A. (2019). La importancia de la ética en el mundo de la inteligencia artificial y la robótica. *Universidad Complutense de Madrid*, 76-77.

Hao, K. (2019). This is how AI bias really happens—and why it's so hard to fix. *MIT Technology Review.* Retrieved from: https://www.technologyreview.com/2019/02/04/137602/this-is-how-ai-biasreally-happensand-why-its-so-hard-to-fix/.

Houkes, W. & Vermaas, P.E. (2010). Technical Functions: On the Use and Design of Artefacts, (Philosophy of Engineering and Technology 1), Dordrecht: Springer Netherlands.

Huang, M-H. & Rust, R.T. (2018). Artificial intelligence in service. *Journal of Service Research*, 21(2), 155-172.

International Federation of Robotics. (2018). "World Robotics report 2018", Press conference summary, Tokyo.

Kelley, R., Schaerer, E., Gomez, M., & Nicolescu, M. (2010). Liability in Robotics: An International Perspective on Robots as Animals. *Advanced Robotics*, 24(13), 1861-1871. http:// doi.org/10.1163/016918610X527194.

Macnish, K. (2017). *The Ethics of Surveillance: An Introduction,* London: Routledge.

Madiega, T. (2019). EU guidelines on ethics in artificial intelligence: Context and implementation. European Parliamentary Research Service, 1-13. Retrieved from: https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf.

Mick, D.G. & Fournier, S. (1998). Paradoxes of Technology: Consumer cognizance, emotions, ad coping strategies. *Journal of Consumer Research,* 25(2), 123-143.

Mittelstadt, B. D. & Floridi, L. (2016). The Ethics of Big Data: Current and Foreseeable Issues in Biomedical Contexts. *Science and Engineering Ethics*, 22(2): 303–341. http://doi.org/10.1007/s11948-015-9652-2

Monasterio Astobiza, A. (2017). Ética algorítmica: Implicaciones éticas de una sociedad cada vez más gobernada por algoritmos. ILEMATA,24,185-217.

Nguyen, N. T., & Biderman, M. D. (2008). Studying ethical judgments and behavioral intentions using structural equations: Evidence from the multidimensional ethics scale. *Journal of Business Ethics*, 83(4), 627–640.

Noeo Tech (S.f.). Roboética, una necesidad urgente. *Edita Noeo Tech*. Retrieved from: https://noeotech.com/

Rawls, J. (1971/1999). *A theory of justice (revised edition).* Cambridge: Harvard University Press.

Reidenbach, R. E., & Robin, D. P. (1990). Toward the development of a multidimensional scale for improving evaluations of business ethics. *Journal of Business Ethics*, 9(8), 639–653. http://doi.org/10.1007/BF00383391

Santos González, M. J. (2017). Regulación Legal de la Robótica y la Inteligencia Artificial: Retos de futuro.

Salazar, I., & Escribano Otero, J. J. (2021). ¿Necesitan tener ética los robots? UEM STEAM essentials, 1-7.

Selbst, A.D., Boyd, D., Friedler S.A., Venkatasubramanian S., & Vertesi, J. (2019). Equidad y abstracción en sistemas sociotécnicos, en Actas de la Conferencia sobre equidad, responsabilidad y transparencia — FAT*'19, Atlanta, GA: ACM Press, 59–68. http://doi.org/10.1145/3287560.3287598

Shawver, T. J., & Sennetti, J. T. (2009). Measuring ethical sensitivity and evaluation. *Journal of Business Ethics*, 88(4), 663–678. Retrieved from: https://doi.org/10.1007/s10551-008-9973-z

Steinert, S. (2014). The Five Robots - A Taxonomy for Roboethics. *International Journal of Social Robotics*, 6, 249-260. http://doi.org/10.1007/s12369-013-0221-z

Torras, C. (2014). Robots sociales: Un punto de encuentro entre ciencia y ficción. *Revista Mètode*, 1-5. Retrieved from: https://dialnet.unirioja.es/servlet/articulo?codigo=4765325

Tsafestas, S. G. 2018. Roboethics: Fundamental Concepts and Future Prospects. *Information*, 9(6), 148. http://doi.org/10.3390/info9060148

Valls Prieto, J. (2019). El reto de una robótica e inteligencia artificial honesta con las personas. *Infobae*.

Westerlund, M. (2020). An ethical framework for smart robots. *Technology Innovation Management Review*, 10 (1), 35-44. http://doi.org/10.22215/timreview/1312

Whittaker, M., Crawford, K., Dobbe, R., Fried, G., Kaziunas, E., Manthur, V., Myers West, S., Richardson, R., Schultz, J., & Schwartz, O. (2018). AI Now Report 2018, AI Now Institute, 1-62. Retrieved from: https://ainowinstitute.org/AI_Now_2018_Report.pdf.