Human Factors in Robots, Drones and Unmanned Systems, Vol. 138, 2024, 128–135
https://doi.org/10.54941/ahfe1005016

Robotics and Autonomous Systems in Public Realm: An Exploration of Human, Ethical and Societal Issues in Emergency First Response Operations

Gabriella Duca, Raffaella Russo, and Vittorio Sangermano

Institute for Sustainable Society and Innovation, Naples, Italy

ABSTRACT

In the face of increasing threats from climate change and natural hazards, the need for faster, safer, and more effective first response operations has become paramount. This has led to a growing focus on the potential of robotics aids and autonomous systems to support first responders in their duties. While these technologies hold promise for more efficient onsite operations and reduced risk exposure for first responders, there are emerging concerns about their adaptability to real environment constraints, usability, and societal impacts. Scientific literature only mention high-level concerns about human-centric approach and generic ethical issues, but these are worthy to be identified and elicited in parallel with the evolution of technical requirements and specification, to build capacity of estimating the extent of new operating methods and procedures impact on victims and responders, but also on other stakeholders. Guidelines to steer choices of emergency personnel already exist, for instance in the case of medical personnel, but first response automation might imply unknown or indefinite dilemmas on aspects such as fairness and discrimination, false or excessive expectations, privacy, physical and psychological safety, liability. The paper proposes a review of the current status of human and societal issues in robotics and automation, eliciting human factors and ergonomics specific issues to foster the human-centric approach claimed by European Union.

Keywords: Human factors, Usability, End-users requirements, Inclusion

INTRODUCTION

In a world facing increasing threats from climate change and natural hazards affecting increasing densely populated areas, making first response faster, safer and more effective and efficient is a crucial aspect for saving lives and reduce losses from disasters. In order to make first response operations more adequate to increasingly challenging needs, great expectations are put towards autonomous systems (including robots and autonomous vehicles, either ground and aerial) supporting first responders in their duties, with the twofold aim to carry out more efficient onsite operations and reduce their exposure to risk.

All technologies can have human impacts as they are introduced into society and when it comes at emergencies, where decisions are taken in a scarcity

of resources and information, ethical and societal impact can be greater or disproportionate.

Scientific literature only mention high-level concerns about acceptance and generic ethical issues, but these are worthy to be identified and elicited in parallel with the evolution of technical requirements and specification of new operating methods and procedures to estimate consequences and weight on victims and responders, but also on other stakeholders. Guidelines to steer choices of emergency personnel already exist, for instance in the case of medical personnel, but first response automation might imply unknown or indefinite dilemmas on aspects such as fairness and discrimination, false or excessive expectations, privacy, responsibility, physical and psychological safety.

THE PROBLEM OF PROPERLY FRAMING HUMAN-CENTREDNESS OF ROBOTICS AND AUTONOMOUS SYSTEMS

Critical thinking towards innovation is somehow embodied in human kind. Throughout history, technological innovations have often been met with fear and scepticism; these fears have been driven by concerns about job loss, social unrest, and the potential consequences of technological advancement for human health, safety and capability. With rise of expectations generated by scientific advances in autonomous systems and AI (Artificial Intelligence) the public discourse on these topics started to encompass a wide range of concerns and language, with a focus on ethical, legal, and societal implications, as well as the need for meaningful human control over these systems (Johnson and Verdicchio, 2017). Especially with reference to the word "ethics", it seems that it is becoming a buzzword, having acquired so many different meanings in the documents and discourses on AI that it has almost become an empty signifier (van Dijket al., 2021). Therefore, it is clear that practical approaches or actionable methodologies bringing human centeredness of autonomous systems to practical evidences seem still lacking, especially with reference to the ethics sphere, with the possible consequence of a cloudy claim for human-centeredness that can be misleading or foster manipulation (Casiraghi, 2023).

The European Union (EU) has a strong focus on human-centered AI and aims to build trustworthy AI that puts people first. The EU has proposed the first-ever legal framework on AI, which addresses its risks and positions to play a leading role globally. The Coordinated Plan on AI aims to accelerate investment in AI, implement strategies and programs, and align policy to prevent fragmentation within Europe. The EU's approach to AI involves fostering excellence, enabling its development and uptake, and ensuring that AI works for people and is a force for good in society. The EU is also engaged in international outreach for human-centric AI, promoting its vision on sustainable and trustworthy AI globally through policy dialogue, joint initiatives, and public outreach (European Commission, 2024, January 29; European Commission, 2024, February 6).

The EU's approach is underpinned by the work of the High-Level Expert Group on AI, which has developed Ethics Guidelines for Trustworthy AI and

policy recommendations to guide reliable AI towards sustainability, growth, competitiveness, and inclusion (AI HLEG, 2019).

The EU's proposed AI regulation aims to provide clear requirements and obligations for AI developers, deployers, and users, with a strong focus on addressing the risks associated with AI and ensuring the safety and fundamental rights of people and businesses.

The following paragraphs try to elicit the human related issues in a more actionable perspective, with the aim to foster the development of human-centred autonomous systems in the context of public safety during the emergency management.

Fairness and Discrimination in Autonomous Decision-Making

One of the key aspects of ergonomics and human factors perspective on robotics and autonomous systems in first response operations deals with the opportunity and capability of a comprehensive consideration of human variability and social inclusion, to leave none behind. This topic opens the issue of fairness perceptions of algorithmic decision-making and the analysis of discrimination, bias, fairness, and trustworthy AI in the context of social systems (Starke et al., 2022; Varona and Suarez, 2022). Research has been conducted on the fairness and generalizability of AI models, highlighting concerns about model discrimination and bias, particularly against certain ethnic and socioeconomic groups (Röösli et al., 2022), to not to mention gender. Inequalities can be generated by the underrepresentation of specific social groups in civil society, resulting for instance in "biased" database. When using autonomous system for real time data collection prior the rescue operations, it has also to be considered that biased data set can be intrinsically unjust (Challen et al., 2019), thus leading to inequalities in risk prediction models, and unfair allocation of technological and human resources to be deployed, as in the case of estimating the extent of material of human losses or damages in a given area to prioritize intervention. As practical example, irregular migrants might not be in official databases and then their presence in a building or area might be overlooked by the autonomous decision making tools. Also in terms of assessing risks to individuals, including the balance between human agency and machine autonomy, autonomous systems in emergency response raise unprecedented questions (Pflanzer et al., 2023).

The development of privacy-respecting autonomous systems is an active area of research, aiming to address the privacy concerns associated with the use of autonomous systems, including the collection of personal data (Such, 2017) that might affect real time decisions by both humans and machines. Autonomous systems, including autonomous vehicles and drones, together with the use of wearables technologies by first responders can raise ethical concerns about the collection, processing, and use of personal data, posing legal and ethical issues related to privacy, confidentiality, and data ownership when processing real-time data for emergency management and response (Damaševičius et al., 2023).

These discussions and studies reflect the complex practical ethical considerations involved in human variability and inclusion, especially in the context

of resource allocation during a public safety crisis. On the other side, we can expect that elicitation and early consideration of such aspects in real-like scenarios to train autonomous systems represents a unique opportunity to steer first interventions towards a more responsible civic response, fostering the allocation of resources based on central policies rather than individual beliefs and judgment by people on the scene. Therefore, for the integration of ethical frameworks into disaster decision making by autonomous systems, human factors researchers and professionals should foster the protection of individual rights and well-being during crisis situations focusing on human variability and inclusion data.

Calibrating Expectations by Citizens and First Responders

In general, the aim of saving lives and rescuing affected people in an emergency scenario generate positive attitudes to autonomous technologies in both citizens and first responders. As practical example, the use of autonomous vehicles in emergency medical services is expected to bring efficiency, cost reduction, and increase life savings and it is therefore considered a typical application where high social acceptance could drive more pervasive use of autonomous technologies (Duca et al., 2023). Thanks to autonomous systems able to collect and merge more reliable data on crisis scenario, first responders would benefit of a better situational awareness allowing, for instance a faster localization of someone entrapped under debris or in a flood, or the identification of the safest approaching route to a target point. Moreover, robotics can bring items or execute actions keeping the first responders far from dangerous environments (i.e. smoky environment or collapsing buildings). It is clear that such benefits create high expectations in both first responders, able to improve their safety and efficiency, and general public, expecting to receive more efficient and effective aids. Despite that, when available technologies are experimented in real-like disaster scenarios, robotics and autonomous systems show reduced benefits due to lack consideration of real environment constraints, poor adaptability to case by case needs, to not to mention overall usability issues (Avinaash et al., 2021). A further issue is represented by false positive or incomplete data set feeding the data processing, with the consequence of inefficient allocation of resources.

Inadequate effectiveness, negative prior experience on reliability, and the degree of risk involved in the task with robots and autonomous systems can affect trust (Oksanen et al., 2020). Ability of autonomous and robotics systems can be considered as a facet of trustworthiness, as recent studies demonstrate the relationship between usability and willingness to use a robot in future workplaces (Kim et al., 2020). Trust can be conceptualized as a multidimensional psychological attitude involving beliefs and compliance with the robot's recommendations and autonomous systems decisions (Lewis et al., 2018).

Real life trials underscore the need of a robust methodological approach in end-users requirement definition for autonomous system, to avoid false or excessive expectations and build trust towards such technologies.

Responsibility, Accountability, Liability

The relationship between autonomy and responsibility in autonomous systems is a key consideration. The criticism of autonomous operations in emergency situations revolves around several key concerns. One major issue is the difficulty in determining accountability and liability when autonomous systems make critical decisions that might be harmful for citizens or first responders, as it may be challenging to hold individuals or entities responsible for the outcomes of autonomous decisions (United States Army, 2017).

Actually, responsibility, accountability, and liability are distinct concepts relevant to the development and deployment of autonomous systems, entangling human actors in different ways (Goetze, 2022). Responsibility refers to the ethical and legal obligation of an agent, which can be moral or professional in nature. It encompasses the idea of being accountable for one's actions and decisions. In the context of autonomous systems, responsibility rests upon various agents involved, such as developers, users, and regulatory bodies (Futurium, n.d.). Accountability is the expectation that an individual or entity will be evaluated against specific performance standards or outcomes. It involves answerability and enforcement of consequences for responsibilities not fulfilled. As AI systems become more autonomous, it becomes increasingly difficult to determine who should be held accountable for their actions, raising complex questions about the nature of responsibility and the extent to which autonomous systems can be held accountable. Liability, on the other hand, is a legal concept that refers to the legal responsibility for one's acts or omissions. It involves the obligation to compensate for damage caused by an action and to follow rules and regulations laid down in law. In the context of autonomous systems, liability legislation will require thorough deliberation, especially as these systems are based on automated decision-making in social and economic domains (Goetze, 2022).

It is clear that under the perspective of complex socio-technical systems the three concepts are intertwined and represent a crucial research topic in the field of tasks and decisions-making allocation between human and machine. Standardized Human Factors techniques can help to put forward the understanding of who can be responsible for what, specifying steps and rules in the autonomous processes. This also deals with the topic of explainable and transparent AI (Ghaffarian et al., 2023).

Finally, also the topic of possible interferences of autonomous systems normally operating the public realm, such as autonomous cabs, is worthy to be mentioned. Incidents have been reported where driverless vehicles have obstructed emergency response zones, interfered with fire equipment, and caused delays in first responders' activities, calling for better training of autonomous vehicles to properly behave in disaster scenarios and effectively interact with first responders (City & County of San Francisco, 2023, August 7).

Physical and Psychological Safety of Citizens and First Responders

Autonomous systems are expected to operate with a high level of safety and efficiency, but the limitations of classical safety approaches to this field are well acknowledged (McDermid et al., 2019).

There are concerns related to safety risk sources, such as lack of situational awareness, control failures and access to forbidden area, unexpected autonomous operating mode initiation, and incomplete or improper system updates (Malm et al., 2022), which become even less manageable under the changeable and uncertaint context that characterize a disaster scenario. This is linked to another concern: the unpredictability of complex autonomous systems, which can lead to unanticipated failures and potentially catastrophic consequences.

Even if emergency situations usually generate positive attitude towards autonomous systems by citizens and end-users, we should consider that human actors in a disaster scene can be under emotional or cognitive pressure. First responders co-operating with autonomous systems or robots could experience haltered level of danger finally worsening their level of psychological comfort (Akalin et al., 2023). Under the perspective of assisted or rescued people, we can observe a twofold implication on psychological safety. On one side, being helped by an inanimate technology could deprive the victim of the psychological first aid provided by human first responders. On the other hand, autonomous systems can convey the information of human presence to missing people that would not be reached by any other human contacts, bringing a sort of psychological relief.

It is clear that robots and autonomous systems are necessarily deployed within complex human, social, and organizational systems and sociotechnical sources of risk and failure should be taken into account in risk assessment (Macrae, 2021) to develop robust safety functions and to define overall system requirements able to control potential safety drawbacks in a disaster scenario.

CONCLUSION

As organizations increase their use of artificial intelligence, people are questioning the extent to which human biases have made their way into AI systems. Biases tend to persist as they necessitate a profound grasp of data science methods and a comprehensive understanding of social dynamics, such as data collection, to be recognized and mitigated. Overall, debiasing stands out as one of the most challenging hurdles, particularly due to its complex social implications (McKinsey & Company, 2024).

The paper has provided a review of the current status of human and societal issues in robotics and autonomous systems. It highlights how ergonomics and human factors can play a crucial role in achieving the "AI for good" goal. It underscores the need to consider human variability and social inclusion to ensure fairness in algorithmic decision-making and address potential biases in the use of autonomous systems during first response operations. It also emphasizes the ethical implications related to privacy, confidentiality, and data ownership in the context of real-time data collection for emergency

management and response. While the objective of preserving life and aiding those hurt in crisis situations fosters a favourable perception of autonomous technologies among both the general public and emergency personnel, we can conclude the a robust methodological approach in defining end-user requirements for autonomous systems is still needed. This is also crucial to avoid creating false or excessive expectations and to build trust in these technologies. Responsibility, accountability, and liability are separate but interconnected critical concepts for autonomous systems, which could benefit from standardized Human Factors techniques to determine who can be held responsible for what. Finally, the paper acknowledges the limitations of classical safety approaches to autonomous systems and draws the attention to unpredictability of autonomous systems in uncertain scenarios, the psychological safety implications for both people in need and first responders, as well as the need of a better understanding of the socio-technical sources of risks.

REFERENCES

Akalin, N., Kiselev, A., Kristoffersson, A., & Loutfi, A. (2023). A taxonomy of factors influencing perceived safety in human–robot interaction. International Journal of Social Robotics, 15(12), 1993–2004.

Avinaash, M., Andrew, D., Visnu, S., Abrar, M., & Seenu, N. (2021, November). Design and simulation of disaster mitigation robot using machine vision. In Journal of Physics: Conference Series (Vol. 2115, No. 1, p. 012009). IOP Publishing.

Casiraghi, S. (2023). Anything new under the sun? Insights from a history of institutionalized AI ethics. Ethics and Information Technology, 25(2), 28.

Challen, R., Denny, J., Pitt, M., Gompels, L., Edwards, T., & Tsaneva-Atanasova, K. (2019). Artificial intelligence, bias and clinical safety. BMJ quality & safety.

City & County of San Francisco. (2023, August 7). CPUC Status Conference: Safety Issues Regarding Driverless AV Interactions with First Responders. In City & County of San Francisco (Ed.), Driverless AV Safety and First Responders.

Damaševičius, R., Bacanin, N., & Misra, S. (2023). From sensors to safety: Internet of Emergency Services (IoES) for emergency response and disaster management. Journal of Sensor and Actuator Networks, 12(3), 41.

Duca, G., Trincone, B., Dziugiel, B., Liberacki, A., Russo, R., Sangermano, V., & Witkowska-Konieczny, A. (2023). Urban Air Mobility in the next future: Main findings from a technical survey. Transportation Research Procedia, 72, 965–972.

European Commission. (2024, February 6). A European approach to artificial intelligence. Digital-strategy. https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

European Commission. (2024, January 29). International outreach for human-centric artificial intelligence initiative. Digital-strategy. https://digital-strategy.ec.europa.eu/en/policies/international-outreach-ai

Futurium. (n.d.). European AI Alliance - Accountable Artificial Intelligence Builds upon Responsibility and Liability. Retrieved from https://futurium.ec.europa.eu/en/european-ai-alliance/best-practices/accountable-artificial-intelligence-builds-upon-responsibility-and-liability.

Ghaffarian, S., Taghikhah, F. R., & Maier, H. R. (2023). Explainable artificial intelligence in disaster risk management: Achievements and prospective futures. International Journal of Disaster Risk Reduction, 98, 104123.

Goetze, T. S. (2022, June). Mind the Gap: Autonomous Systems, the Responsibility Gap, and Moral Entanglement. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency (pp. 390–400).

High-Level Expert Group on Artificial Intelligence (AI HLEG). (2019) Ethics guidelines for trustworthy AI, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419.

Johnson, Deborah G. & Verdicchio, Mario (2017). Reframing AI Discourse. Minds and Machines 27 (4):575–590.

Kim, W., Kim, N., Lyons, J. B., & Nam, C. S. (2020). Factors affecting trust in high-vulnerability human-robot interaction contexts: A structural equation modelling approach. Applied ergonomics, 85, 103056.

Lewis, M., Sycara, K., Walker, P. (2018). The Role of Trust in Human-Robot Interaction. In: Abbass, H., Scholz, J., Reid, D. (eds) Foundations of Trusted Autonomy. Studies in Systems, Decision and Control, vol. 117. Springer, Cham.

Macrae, C. (2022). Learning from the failure of autonomous and intelligent systems: Accidents, safety, and sociotechnical sources of risk. Risk analysis, 42(9), 1999–2025.

Malm, T., Tiusanen, R., Heikkilä, E., & Sarsama, J. (2022). Safety risk sources of autonomous mobile machines. Open Engineering, 12(1), 977–990.

McDermid, J. A., Jia, Y., & Habli, I. (2019, August). Towards a framework for safety assurance of autonomous systems. In Artificial Intelligence Safety 2019 (pp. 1–7). CEUR Workshop Proceedings.

McKinsey & Company. (2024, March 29). What AI can and can't do (yet) for your business. Retrieved from https://www.mckinsey.com/capabilities/quantumblack/our-insights/what-ai-can-and-cant-do-yet-for-your-business.

Oksanen, A., Savela, N., Latikka, R., & Koivula, A. (2020). Trust toward robots and artificial intelligence: An experimental approach to human–technology interactions online. Frontiers in Psychology, 11, 568256.

Pflanzer, M., Traylor, Z., Lyons, J. B., Dubljević, V., & Nam, C. S. (2023). Ethics in human–AI teaming: principles and perspectives. AI and Ethics, 3(3), 917–935.

Röösli, E., Bozkurt, S., & Hernandez-Boussard, T. (2022). Peeking into a black box, the fairness and generalizability of a MIMIC-III benchmarking model. Scientific Data, 9(1), 24.

Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. Big Data & Society, 9(2), 20539517221115189.

Such, J. M. (2017, August). Privacy and Autonomous Systems. In Ijcai, 4761–4767.

United States Army. (2017, May/June). Pros and Cons of Autonomous Weapons Systems. Military Review. https://www.armyupress.army.mil/Journals/Military-Review/English-Edition-Archives/May-June-2017/Pros-and-Cons-of-Autonomous-Weapons-Systems

van Dijk, N., Casiraghi, S., & Gutwirth, S. (2021). The 'Ethification' of ICT Governance. Artificial Intelligence and Data Protection in the European Union. Computer Law & Security Review, 43, 105597.

Varona, D., & Suárez, J. L. (2022). Discrimination, bias, fairness, and trustworthy AI. Applied Sciences, 12(12), 5826.