The emergence of ethics engineering in Industrial Cyber-Physical Systems

Paulo Leitão*, and Stamatis Karnouskos†
* Research Centre in Digitalization and Intelligent Robotics (CeDRI), Instituto Politecnico de Bragança, Campus de Santa Apolónia, 5300-253 Bragança, Portugal, Email: pleitao@ipb.pt
† SAP, Walldorf, Germany, E-mail: karnouskos@ieee.org

***Abstract*—Artificial Intelligence (AI) is being successfully adopted in innovative industrial applications in different sectors, e.g., manufacturing, logistics, healthcare, smart grids, and smart transportation. However, the potential direct or indirect threats of using AI are not sufficiently investigated in these domains, and as a result, unwanted vulnerabilities are introduced in industrial applications, e.g., related to business info leakage, privacy, and even undesired/false training and learning of personnel and automated processes. The ethical dimension is new ground for most industrial applications and poses a significant challenge in AI systems overall, and especially for complex Industrial Cyber-Physical Systems (ICPS), which are based on collective intelligence and exhibit autonomy, emergence and self-organization. This work discusses the importance and implications of ethics in industrial AI-based systems such as the ICPS. For this purpose, the concept of ethics in such intelligent systems is surveyed and characterized using a bibliometric literature review, and then, areas linked to ethics, as well as the main challenges for engineering ethically compliant ICPS are discussed.**

I. Introduction

The advent of the fourth industrial revolution is transforming the way people, systems and processes are interacting and operating, with digitalization and digital transformation affecting not only a few sectors but rather having a major impact on industry societal fabric, effectively transforming how the world operates. In such an environment, Industrial Cyber-Physical Systems (ICPS) [1] act as the backbone infrastructure, combined with emergent ICT technologies like Artificial Intelligence (AI), Internet of Things (IoT) and cloud computing, to develop innovative systems and processes based on a network of integrated cyber and physical counterparts. In these systems, the intelligence and data analysis capabilities are distributed and emerge through cooperation and collaboration among stakeholders. These collective intelligence capabilities are embedded as AI algorithms that allow to extract insights and derive knowledge in a distributed manner, which is then utilized to enhance business and operational processes e.g. with intelligent monitoring, diagnosis, prediction, optimization and planning.

Most efforts asymmetrically focus on the benefits that AI can provide for industry and, to a lesser degree, on the increasing footprint of associated security, trust, privacy and governance aspects of such systems. This has lately raised criticisms and discussions that go beyond the technological perspective, namely related to social, ethical and law aspects of the industrial AI-based scenarios. This is due to the capability of these systems to operate autonomously and in combination with the capability to learn in order to better support future decision-making processes. These issues are even more important and complex to be handled in the case of collective artificial intelligent systems, such as ICPS, where decentralization, emergence, and self-organization phenomena are also present.

In particular, ethics is a crucial aspect to take into consideration in the engineering of intelligent based ICPS. The *ethics* term is derived from the Greek word *ethos* which can mean custom, habit, character or disposition and is concerned with good practices for individuals and society. In fact, ethics, also called moral philosophy, is a set of moral principles [2], [3], *"concerned with what is morally good and bad and morally right and wrong"* [2]. More recently, applied ethics refers to the will to "embed" ethics into designed systems [4], and can be defined as *"an academic discipline that inquiries about the correctness of certain practical human activities, primarily using philosophical methods"* [5]. However, the realization of ethics in ICPS is a complex and challenging undertaking, highly dependent on interdisciplinary factors, as contemporary efforts, e.g. in self-driving cars have shown [6].

While in science fiction, Isaac Asimov has considered the ethical aspects of intelligent systems in his 1950s book "I, Robot", where the Three Laws of Robotics were proposed to regulate intelligent robotic systems, such laws are limited and do not sufficiently cover modern systems [7], [8]. Ethical aspects are not only concerned with the operation process of an intelligent system due to its capability to learn and autonomously take their decisions, but also during the design and engineering process of such intelligent systems, due to the application (or not) of code of conduct by the system developers. Therefore, the whole lifecycle of intelligent ICPS needs to be considered, and this does not only refer to engineering such ethical industrial systems [9] but also consider their governance and potential future societal impact [8] they might have.

Having this in mind, the objective of this paper is to study and analyze the importance and impact of the ethics aspects in AI-based systems, particularly in ICPS. For this purpose, the concept of ethics associated with morality and conscience in such intelligent systems is surveyed and characterized using a bibliometric literature review, and then, the factors that emerge

978-1-7281-9023-5/21/$31.00 © 2022 IEEE

as important to ethics and the main challenges for engineering ethical compliant ICPS are discussed.

II. A Bibliometric view of Ethics in Autonomous and Intelligent Systems

A bibliometric literature review of ethics in autonomous and intelligent systems has been carried out by retrieving a corpus from the Scopus scientific repository that addresses the following criteria (searched in the title, abstract and keywords): *[("Ethics" OR "Ethical") AND ("Artificial Intelligence" OR "Cyberphysical system" OR "Cyber-physical System" OR "Intelligent system")]*. The search was filtered for papers written in English, resulting in a dataset of 3815 documents. The dataset analysis was conducted using automated methods based on Natural Language Processing (NLP) combined with human review analysis to obtain preliminary insights regarding the impact of ethics in AI and CPS systems, while we also focus on the direct or indirect applicability in industry and commonly refer to them as ICPS throughout this work.

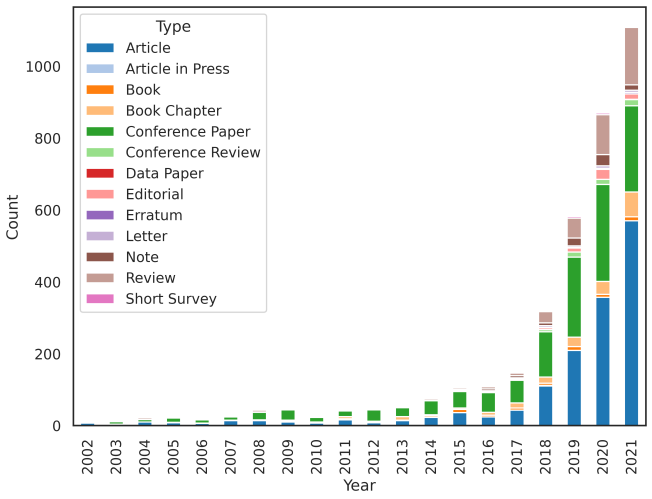

Figure 1. Evolution of papers along the last 20 years focusing ethical aspects in AI or ICPS.

An initial bibliometric analysis is shown in Figure 1 which illustrates the evolution of the number of publications along the years focusing on the ethical aspects in AI or ICPS, grouped by the different forms where these appeared. While in the dataset such efforts are evident over 40 years ago, these were limited in number, while the increasing focus is evident in the last 20 years as shown in Figure 1. Continuous growth over the years has been noticed, which is exponential in the last five years, denoting the sharp rise of interest in ethical aspects of AI and ICPS. However, although the ICPS community is increasingly considering ethical aspects, the published works still constitute a fraction of the overall area as approximately 1.8% of the retrieved papers focus on ethics and ICPS topics (since 2013). The trends may imply that this growing interest in ethical considerations from the AI community will continue in the next years and that the interest by the ICPS community may start to also increase, especially considering that visions such as that of Industry 5.0 promote sustainability and human integration.

In terms of the geographical areas, the distribution of the authors is shown in Figure 2 for the top 25 countries. It becomes evident that this research field is more centered at the moment in the US and UK communities, followed by Germany, Canada, Australia, etc., which shows that the countries where AI and ICPS based systems are stronger disseminated are also those that are more concerned to the ethical aspects (note that the geographic distribution considers the affiliation of the authors of each article using a composite score coming out from all authors listed in the article, with linearly decreasing weights).

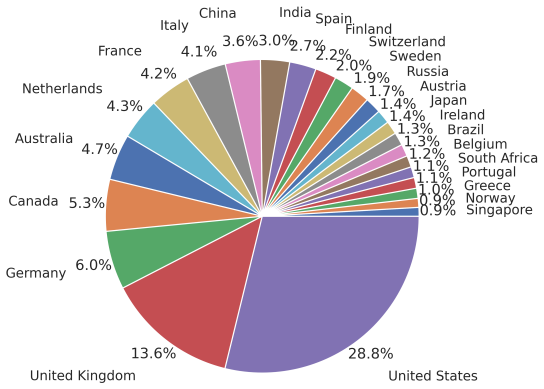

Figure 2. Geographical distribution of the authors in the top 25 countries

Analyzing the uni-grams and bi-grams extracted from the retrieved dataset, it is also possible to observe the most frequent aspects related to the use of ethics in AI and/or ICPS. Figure 3 shows that several of them are related to inherent characteristics of such intelligent systems (e.g., autonomy, learning, emergence), related to data (e.g., privacy, sharing and protection), as well as related to the behavior (e.g., law, responsibility, conscience, moral, trustworthy). It is important to notice the relevance of the education aspect, which signifies the need to consider these ethical aspects in the educational and training programs.

An interesting analysis is to observe which aspects emerge when narrowing the analysis to the papers directly related to ethics and ICPS (only 68 papers), which are illustrated on the right side of Figure 3. Generally, the major aspects are those usually found in AI-based systems, with security assuming a prominent relevance.

Two main application sectors appear as mainly associated to the ethics in AI: medical and autonomous driving (including robotics). This is understandable as the use of AI in these two domains is becoming widely adopted and has an impact to their behaviours. These domains are also highly linked to human values and safety as typically in both the technology results directly impact human lives, e.g. in assistive robotics [10] or self-driving cars [6]. Hence it comes as no surprise that societal, ethical and governance/law concerns related to

Preprint version of doi:10.1109/icps51978.2022.9816931

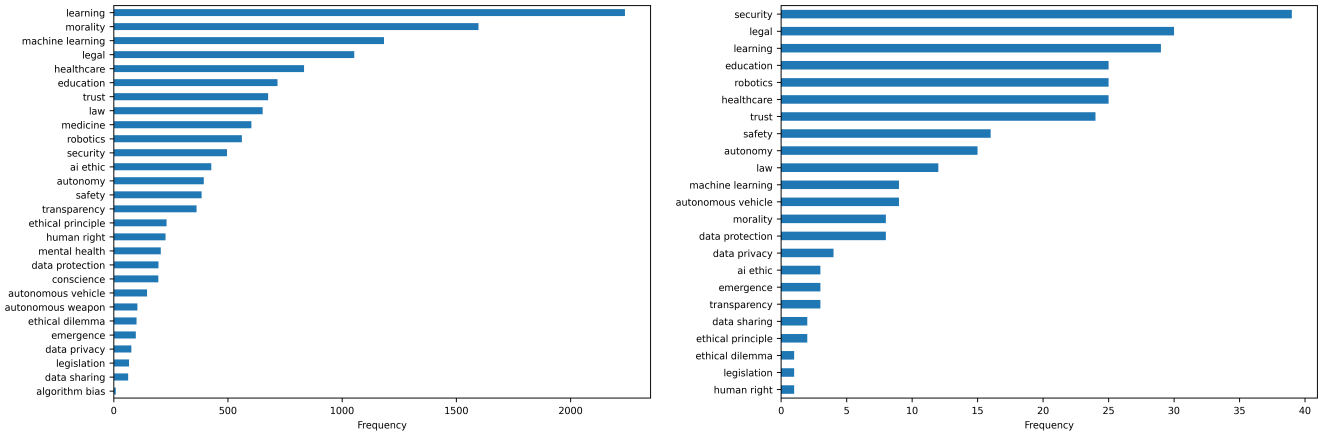

Figure 3. Main aspects associated to the ethical behavior in AI and ICPS (left) and ICPS only (right).

the operation of such systems arise. Over the last years, also weaponization and military usage of autonomous AI systems have been on the rise [11], which, however, may not be sufficiently reported in the literature due to confidentiality concerns.

III. Key Research Areas for Ethics in ICPS

A grouping of the areas that have emerged from literature has been carried out, and which has led to consider as key, holding the potential to impact ethics directions in ICPS.

A. Situational and Cultural Context Awareness

AI systems should exhibit characteristics such as transparency, justice and fairness, non-maleficence, responsibility, privacy, beneficence, freedom and autonomy, trust, sustainability, dignity and solidarity, to be accomplished with ethical aspects [12]. A challenge is the need for AI systems to be transparent so that humans can be aware that they are interacting with a product or service that is empowered with AI capabilities, as well as be aware of its limitations and data usage. The same applies to the trustworthiness dimension, which would be challenging to establish since, for instance, nowadays a machine is not trusted in the same manner as the humans are. Expectation management on the behaviour of AI-enabled ICPS is also complex, as AI systems are expected to behave perfectly and are more harshly judged when they fail than their human peers [8].

Ethics are also influenced by cultural, social and situational context awareness, which are also linked to the notion of morality. As an example, different cultures have distinct perspectives regarding morality, namely eastern and western countries have different cultural traditions and ethical principles, and the law and responsibility are treated in different manners according to the local context. In the same manner, data management and the acceptance of technology are handled differently. ICPS will need to operate in a manner that is in compliance with the law in each country while also being sensitive to these cultural and social contexts, something that is challenging.

B. Explainable AI Behavior and Evolution

The use of AI has both challenges and opportunities. On the one side, it can bring substantial benefits, e.g., to assist in industrial scenarios with prediction, optimization and diagnosis tasks, as well as to assist developers of such applications and services in developing safe and reliable systems. However, on the other side, its usage can become dangerous if its behaviour can not be properly predicted or explained, as it may be producing false commands and undesired decisions, e.g., using industrial systems in a mode not originally considered by its developers or making strategic military decisions autonomously [13] without prioritizing human ethics, which may result in devastating consequences.

In industrial settings with well-defined scenarios, the AI can outperform humans, e.g., in medical image analysis, and make faster and better decisions, which in turn increases the effectiveness of such services. However, this assumes that there is sufficient training data and effective algorithms, as well as that the data used is free of fallacies, e.g., underlying bias. Implementing ethical decision making in the lifecycle (from training to inference in the real world) is a challenging issue, especially for borderline cases. Because the learning/inference processes rely on the used data, over the last years, attacks have been demonstrated that manipulate the data in order to lead to misbehaviour of the AI system. Such attacks, e.g. image manipulation, might be undetectable by humans, which, however, could have devastating effects on AI ICPS as the system may misbehave and can no longer be trusted to operate as designed.

The definition of autonomy boundaries during the learning process is important to ensure adherence to moral values, ethical considerations and overall governance. As an example, a well-trained autonomous vehicle initially doesn't drive faster than the speed limit allows. However, as it also continues to learn by observation, it may observe other misbehaving vehicles, and may decide that it can also violate the set speed rules. How continuous learning for such systems can be realized while also adhering to safety as well as validation and certification processes set in each domain is challenging.

Preprint version of doi:10.1109/icps51978.2022.9816931

C. Autonomy and Emergence

In ICPS, based on the interaction of autonomous and intelligent individuals, the collective system intelligence emerges from the cooperation and collaboration among these individuals placed at the local/micro level, resulting in a global behaviour that is unknown to them [14] and is more flexible, robust and responsiveness. The emergence concept defines the way the macro level system arises out from a multiplicity of interactions among entities exhibiting simple behaviour (micro level), without the guidance of a central entity and only when the resulted behavior of the whole is greater and much more complex than the simple sum of the behaviors of its parts [15].

The presence of emergence in such systems makes them much more powerful as they can easily adjust to unpredictable situations, but at the same time also much more challenging, e.g., due to the cognitive capabilities of individual entities, the non-linearity of the multiplicity of interactions and the sensibility of initial conditions (also known as the butterfly effect). In ICPS, the ethical aspects arise at local and global levels that emerge from the interaction among individual entities. As an example, one particular entity could have a good ethical behaviour, but when included in collective conduct, as ICPS is, it can exhibit a different ethical and moral behaviour due to emergence and context-awareness. To what extent the ethical behaviour of individual systems can also provide a guarantee for the ethical behaviour at system of systems level is in need of further research.

D. Data Governance

AI-based systems take advantage of the huge amount of data that they train the learning models. Some of these data can be sensitive, e.g., clinical or personal data, and their use should be properly managed, e.g. anonymized. Particularly in distributed systems, where data is shared between different individuals and systems, or in cloud-based systems, where the data is sent and stored in cloud platforms, data sensitivity and privacy raise important issues. In such cases, improper data protection measures or security mechanisms make the system vulnerable and may lead to privacy loss issues, not to mention the incompatibility with specific legislative frameworks.

In industrial environments, stakeholders are traditionally constrained to share their data, even with their partners (customers and suppliers) adopting the usage of private clouds instead of public cloud platforms. Because data is seen as a business advantage, the right level of secure, privacy-preserving sharing needs to be established, which is challenging. Not sharing or significantly restricting the data sharing is not a realistic solution, as high quality and large sets of data is needed to properly train industrial systems, something that might be constrained by sharing data among stakeholders and, therefore, scenarios that could benefit all may not be easily realizable.

Due to the importance of data to the digital economy and to ICPS, effective data governance should be legally regulated to a certain degree by independent certified organizations. The establishment of the European General Data Protection Regulation (GPDR) guidelines is an illustrative example that can be adopted during the design and operation of such systems. However, the establishment of strict regulations may be counterproductive and may affect the way the system will learn. Liability issues are also important, particularly in cases of ICPS misbehavior, which is coupled with considerations of agency and the role of intelligent machines in society [8].

E. Conscience and User Acceptance

Conscience can be defined as the *"personal sense of the moral content of one's own conduct, intentions, or character with regard to a feeling of obligation to do right or be good"* [2]. Humans are intelligent and behave according to their conscience and moral guiding principles. Since ICPS will also feature advanced intelligence and operate in a connected manner, they also need to exhibit moral behavior bound to an ethical framework upon which they can make decisions. It is challenging to define the level and kind of conscience needed and how it may impact their decision-making processes. Of paramount importance for such intelligent systems is user acceptance. In industrial environments, stakeholders are usually pessimistic and conservative regarding the usage of autonomous and evolvable intelligent systems, as they put more emphasis on strictly controlling their behavior rather than value autonomy. However, as systems evolve, their acceptance may increase, especially if they substantiate the added value they bring. Even then though, understanding the factors that may lead to AI system acceptance is not a trivial task and might be domain-dependent [6].

IV. Challenges for Ethics Engineering

Driven by the bibliometric results in section II and the discussion of areas in section III, some main challenges concerning the engineering of ethics in ICPS have been identified. An overview is shown in Table I, which includes also actions for each challenge and an estimated difficulty and priority for its implementation.

A. Managing the Learning and Emergence Processes

In ICPS systems, the global behavior emerges from the interaction among individuals, raising the question related to in which way the interactions between individuals and the cultural context affects the ethical dimension of decision-making, especially in modern multi-cultural contexts. The non-linear interaction and the emergence phenomenon may provoke the emergence of offensive or discriminating behaviors. Hence it needs to be investigated how such multi-goal objectives are handled by AI algorithms to successfully operate in complex environments from a technical and societal point of view.

In autonomic systems, self-* features are inherently present, e.g., self-organization, self-adaptation and self-healing, enhancing the system performance but at the same time introducing potential problems, e.g., regarding myopia, chaos and nervousness [16]. This requires the establishment of boundaries that limit the way the industrial system is operating,

Preprint version of doi:10.1109/icps51978.2022.9816931

Table I
CHALLENGES FOR SELECTED ETHICAL ASPECTS IN ICPS

Challenges	Example actions	Est. Difficulty	Est. Priority
Managing learning and emergence processes	Navigating the emergence in complex contexts	high	medium
	Define boundaries for autonomic and self-* features	medium	medium
	Validation/Certification for the training data of AI systems	medium	high
Regulation and Legislation	Define legal requirements and penalties context of usage	low	high
	Define & enforce the law and operational code of ethics (conduct)	high	high
	Address liability/responsibility for AI system actions	high	high
Data Governance (incl. privacy and security)	Security mechanisms	high	medium
	Adoption of GDPR by AI systems	low	high
	Effective data sharing and data lifecycle control mechanisms	high	high
Methodologies to engineer ethics in ICPS	Standardization & certification	medium	high
	Development of trustworthy AI systems	high	medium
	Evolution & Transfer learning and evolvement of AI ICPS	high	high
User interaction and acceptance	Education/training to enable effective interaction with ICPS	medium	high
	Integration of humans in the AI based decision-making process	medium	high
	Build-up user trust in interaction with AI systems	high	high

aiming to constrain it in a well-defined context and maintain it under control.

An additional problem is related to the bias that may be implicitly evident in AI systems, e.g. introduced by the systems designers or training data. Therefore rigid validation and certification processes need to be in place for the data used for training as well as the algorithms of such systems in order to increase trustworthiness and explainability.

B. Regulation and Legislation

Aiming to enhance the ethics and law aspects in the engineering of ICPS, it is necessary an effort to develop proper legislation that regulates the engineering and operation of such industrial systems. A proper legislative framework needs to be in place that clearly states the legal requirements they have to adhere as well as the penalties for violation. In addition, the regulatory framework should guarantee via proper processes the monitoring and enforcement of the law. For this purpose, it is important to understand the operational context of such systems, including their limits and expectations, as well as not engineer them but also provide the tools to monitor their behavior and compliance with the law and the defined operational code of ethics. Finally, it is crucial to understand the liabilities linked to accountability under the judicial system. This is not easy to address, as responsibilities in case of accident or harmful decisions taken must be clear. Nowadays, several discussions are ongoing about the agency of such systems that extend potentially to their owners or designers. Potentially the adoption of some traditional mechanisms such as reviewing, voluntary labelling, certification by standard organizations, responsible AI license and third-party audits could help in the mid term.

C. Data Governance

The exchange and sharing of industry data among different applications and organizations can be beneficial to businesses and fuel the data economy. However, it also requires the implementation of security mechanisms to guarantee that such data exchange can be done in a safe, privacy-preserving and secure way to not increase the attack surface and averse malicious attack scenarios to the industrial systems. Due to the new attacks possible on AI systems (e.g. manipulated data), new AI-ready security mechanisms need to be in place. In addition, proper governance frameworks for the whole lifecycle of the data (from cradle to grave) also need to be developed. As an example, the ICPS in Europe need to adhere to the GDPR regulation and EU as well as national laws for data protection. However, what is needed is more than the mere implementation of data protection mechanisms towards considering the insights and knowledge generated from that data and how it may be used along with the ethical behavior of the ICPS.

D. Methodologies to Engineer Ethics in ICPS

The engineering of ICPS that consider ethical aspects requires the development of solutions that go beyond the simple use of technology and approaches the issues with truly interdisciplinary teams in all of its phases. The European AI strategy [17] reinforces this need by claiming that *"Interdisciplinarity should also be supported (by encouraging joint degrees, for example in law or psychology and AI).".* To this end, new standards and certification approaches need to be developed for all lifecycle phases of ICPS, e.g. from engineering to operation and recycling.

Automatic and intelligent machines that exhibit learning and cognitive characteristics should also adhere to a code of ethics that they should respect during their operation, which is challenging, especially considering the highly dynamic environments they have to operate, e.g. a smart city, which is not as controlled as some industrial settings, e.g. a factory shop floor. This is also linked to the trustworthiness of AI systems, which may be a critical factor for their wide adoption. Trustworthy AI requires the combination of technical (e.g., software tools) and non-technical methods (e.g. governance mechanisms), being necessary the mathematical/logical formalization of target problems to avoid ambiguity, as well as the development of a common language between AI researchers and ethics/laws experts [18]. Finally, the adoption of methodologies and frameworks that simplify the development

Preprint version of doi:10.1109/icps51978.2022.9816931

of such symbiotic ICPS solutions assumes crucial relevance, e.g., the one proposed by [19] that considers transparency, equity, safety, accountability, privacy and trust, as primary principles in the Operator 4.0 typology. Along with these issues of evolution of AI systems, including transfer learning without compromise of their intended purposes and code of conduct, is another challenging issue.

E. User Interaction and Acceptance

Due to the interdisciplinarity involved in developing the future AI-based ICPS, new more holistic education and training is needed for both students and employees in order to learn how to effectively develop, interact and collaborate with such systems. The proper usage and the mindset of using AI require the implementation of training actions and practical learning for users of digital technologies, especially regarding the basic concepts and the benefits related to the usage of these technologies. The European AI strategy puts forward that the importance of ethics in the development and usage of new technologies should be incorporated into the educational programs and courses [17], particularly introducing ethics training in engineering and computer science courses.

User acceptance for industrial stakeholders is paramount to the success of ICPS in these settings. Users' awareness of the benefits and threats associated with these systems also contributes to the trustworthiness of AI-based systems. Similarly, the degree of explainability of the operations and decisions made by these AI systems contributes to the user's understanding and acceptance. Hence actions that build up user trust in the interactions, co-existence and collaboration with ICPS should be carried out. In addition, a symbiotic [8] integration of the human in the decision-making process since it is the most flexible piece in an automation system [20]. This integration should not be only implemented at the operational level but essentially at the strategical level.

V. Conclusion

AI technologies provide several benefits inevitably to develop smart products and systems, offering a means for data analytics aiming for monitoring, diagnosis, prediction, optimization and planning. In this work, a bibliometric view on the ethical aspects related to the development of AI ICPS has been taken. The trends show an increasing interest, especially in the last years, as AI systems become a reality rather than lab-only developments. Some of the major areas identified with respect to the engineering of ethics are related to situational and cultural context awareness, the explainable AI behavior and evolution, the autonomy and emergence, the data governance, and the conscience and user acceptance. It is argued that the ethical aspects should be considered along the lifecycle of the ICPS, focusing on its interaction with a variety of human stakeholders, each one with a different scale of implication. Finally, based on the identified areas and trends shown in the bibliometric analysis, a discussion on some selected challenges in the engineering of ethical ICPS is performed, including the actions to address them. Future work can be devoted to understand the societal factors better as well as how these can be realized successfully with the state of the art technologies available in order to address some of the identified challenges.

Acknowledgments

This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the Project Scope UIDB/05757/2020.

References

[1] P. Leitão, A. W. Colombo, and S. Karnouskos, "Industrial automation based on cyber-physical systems technologies: Prototype implementations and challenges," *Computers in Industry*, vol. 81, pp. 11–25, Sep 2016.

[2] Britannica Encyclopedia, [Accessed: 08 Dec 2021]. [Online]. Available: https://www.britannica.com

[3] Oxford Learner's Dictionaries, "Ethic," [Accessed: 08 Dec 2021]. [Online]. Available: https://www.oxfordlearnersdictionaries.com/definition/english/ethic

[4] V. Bonnemains, C. Saurel, and C. Tessier, "Embedded ethics: some technical and ethical challenges," *Ethics and Information Technology*, vol. 20, no. 1, pp. 41–58, Jan 2018.

[5] J. Tham, "Applied ethics," in *Encyclopedia of Global Bioethics*. Springer International Publishing, 2014, pp. 1–12.

[6] S. Karnouskos, "Self-driving car acceptance and the role of ethics," *IEEE Transactions on Engineering Management*, vol. 67, no. 2, pp. 252–265, May 2020.

[7] R. Murphy and D. D. Woods, "Beyond Asimov: The Three Laws of Responsible Robotics," *IEEE Intelligent Systems*, vol. 24, no. 4, pp. 14–20, Jul 2009.

[8] S. Karnouskos, "Symbiosis with Artificial Intelligence via the prism of law, robots, and society," *Artificial Intelligence and Law*, vol. 30, no. 1, pp. 93–115, May 2021.

[9] D. Trentesaux and S. Karnouskos, "Engineering ethical behaviors in autonomous Industrial Cyber-Physical Human Systems," *Cognition, Technology & Work*, vol. 24, no. 1, pp. 113–126, Mar 2021.

[10] T. Vandemeulebroucke, B. D. de Casterlé, and C. Gastmans, "Ethics of socially assistive robots in aged-care settings: a socio-historical contextualisation," *Journal of Medical Ethics*, vol. 46, no. 2, pp. 128–136, Dec 2019.

[11] T. de Swarte, O. Boufous, and P. Escalle, "Artificial intelligence, ethics and human values: the cases of military drones and companion robots," *Artificial Life and Robotics*, vol. 24, no. 3, pp. 291–296, Jan 2019.

[12] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nature Machine Intelligence*, vol. 1, no. 9, pp. 389–399, Sep 2019.

[13] S. Umbrello, P. Torres, and A. F. D. Bellis, "The future of war: could lethal autonomous weapons make conflict more ethical?" *AI & Society*, vol. 35, no. 1, pp. 273–282, Feb 2019.

[14] E. Bonabeau, M. Dorigo, and G. Theraulaz, *Swarm Intelligence*. Oxford University Press, Oct 1999.

[15] J. H. Holland, *Emergence: from Chaos to Order*. Oxford University Press, 1998.

[16] J. Barbosa, P. Leitão, E. Adam, and D. Trentesaux, "Nervousness in dynamic self-organized holonic multi-agent systems," in *Advances in Intelligent and Soft Computing*. Springer, 2012, pp. 9–17.

[17] European Commission, "Artificial Intelligence for Europe," Apr. 2018, COM(2018) 237 final. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN

[18] European Commission and Directorate-General for Communications Networks, Content and Technology, *Ethics guidelines for trustworthy AI*. Publications Office, 2019.

[19] D. Burnett, N. El-Haber, D. Alahakoon, S. Karnouskos, and D. De Silva, "Advancing an artificial intelligence ethics framework for operator 4.0 in sustainable factory automation," in *Service-oriented, Holonic and Multi-Agent Manufacturing Systems for the Industry of the Future (SOHOMA'21)*. Springer, 2022.

[20] P. Fantini, P. Leitão, J. Barbosa, and M. Taisch, "Symbiotic integration of human activities in cyber-physical systems," *IFAC-PapersOnLine*, vol. 52, no. 19, pp. 133–138, 2019.

Preprint version of doi:10.1109/icps51978.2022.9816931