Beyond 100 Ethical Concerns in the Development of Robot-to-Robot Cooperation

Author(s): Rousi, Rebekah; Vakkuri, Ville; Daubaris, Paulius; Linkola, Simo; Samani, Hooman; Mäkitalo, Niko; Halme, Erika; Agbese, Mamia; Mohanani, Rahul; Mikkonen, Tommi; Abrahamsson, Pekka

Title: Beyond 100 Ethical Concerns in the Development of Robot-to-Robot Cooperation

Year: 2022

Version: Accepted manuscript

Please cite the original version:

Rousi, R., Vakkuri, V., Daubaris, P., Linkola, S., Samani, H., Mäkitalo, N., Halme, E., Agbese, M., Mohanani, R., Mikkonen, T. & Abrahamsson, P. (2022). Beyond 100 Ethical Concerns in the Development of Robot-to-Robot Cooperation. *2022 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics)*, 420-426. IEEE. https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData-Cybermatics55523.2022.00092

Beyond 100 Ethical Concerns in the Development of Robot-to-Robot Cooperation

Rebekah Rousi*, Ville Vakkuri†, Paulius Daubaris‡, Simo Linkola‡, Hooman Samani§, Niko Mäkitalo‡,
Erika Halme †, Mamia Agbese†, Rahul Mohanani† Tommi Mikkonen†, and Pekka Abrahamsson†
*School of Marketing and Communication, University of Vaasa, Finland
†Faculty of Information Technology, University of Jyvškylä, Finland
‡Department of Computer Science, University of Helsinki, Finland
§School of Engineering, Computing and Mathematics, University of Plymouth, United Kingdom
Emails: rebekah.rousi@uwasa.fi, ville.vakkuri@jyu.fi, paulius.daubaris@helsinki.fi, simo.linkola@helsinki.fi,
hooman.samani@plymouth.ac.uk, niko.makitalo@helsinki.fi, erika.a.halme@jyu.fi, mamia.o.agbese@jyu.fi,
rahul.p.mohanani@jyu.fi, tommi.j.mikkonen@jyu.fi, and pekka.abrahamsson@jyu.fi

***Abstract*—This paper investigates ethical issues implicated in multi-robot cooperation development. The authors explore the possibilities of applying ECCOLA—a tool and method for ethical AI software development in the context of robot-to-robot cooperation. The method entailed a workshop of researchers (N=10) who ideated multi-robot ethical concerns (N=149). These were analysed by an external researcher and emergent themes were compared to the ECCOLA framework. The results show the need for expansion of the ECCOLA framework. The paper makes three important contributions: 1) application of ECCOLA in multi-robot cooperation; 2) identification of ethical concerns for robot-to-robot cooperation; and 3) urgency for attention towards mutual cooperation and confrontation in multi-robot cooperation.**

***Index Terms*—multi-robot systems, robots, artificial intelligence, ethics, agile software development**

I. Introduction

The release of artificial intelligent (AI) systems in human societies demands that consideration for their impact on the dynamics and well-being of human entities is well understood, not simply from the point of implementation, but also that of the design and development itself [8]. AI ethics is a field that while related to robotics, has been treated somewhat separately to robot ethics [6]. Reasons for this include differences in the embodied nature of the systems, as well as in relation to the operation logic. Traditionally, robots have not necessarily been reliant on learning algorithms. Yet, the embodied nature of many forms of robotics imposes numerous concerns regarding human-beings that operate on the levels of the social, physical and psychological. Moreover, the ability of material machinery to infiltrate physical environments, also adds to considerations in terms of both gains (e.g., accessing places that are dangerous to humans) and risks (e.g., permeating sensitive sites that may cause risk to human safety and security).

While these differences exist, there are still baseline ethical (or moral) factors that can be brought into development processes. These arise for instance, when considering ethics from normative and applied perspectives in relation to humans and human systems (social and cultural) [21]. Significant efforts have been placed in constructing frameworks, guidelines and principles that reflect the key concerns of humanity and human integrity within the development of AI systems, e.g., IEEE's Ethically Aligned Design, Ethics Guidelines for Trustworthy AI (European Commission) (see also, [21][8][6]). These concerns, or factors, are certainly one step towards developing technology that supports human and societal well-being. However, the matter of understanding how to translate these theoretical factors into practical action points remains a challenge. The challenge comes not only in terms of knowing how to convert the lateral level of ethical construct to the literal level of action and behavior, but also in comprehending how to structure methods in which this conversion and application becomes a natural part of development and programming practices and processes [3] [33].

In this short paper we explore the application of ECCOLA, a method and tool for ethical AI software development, to processes of cooperative robot development - CACDAR [26]. We aim to demonstrate how the combination of implementing ECCOLA in conjunction with multi-robot system development can help identify and address obvious as well as dormant ethical issues within robot-to-robot (R-R) cooperation. The paper begins by plotting the background of key concepts, starting with a brief explanation of AI ethics. Robot ethics follows in which emphasis is placed on the differences between the two forms of ethics. Multi-robot cooperation is then described in order to establish understanding of the application context in question. ECCOLA is outlined in the next section, citing its theoretical underpinnings and practical application. This is followed by a description of the CACDAR project for R-R cooperation development. The method of scoping more than 100 related ethical concerns is explained with a presentation of the emergent ethical themes arising from the analysis that are subsequently triangulated with the ECCOLA framework. These concerns are placed back into context through the presentation of an ethical use case relating to CACDAR robots. The use case is then broken down into identified ethical concerns and implications for future development actions. The study as a whole gives rise to several future research directions.

II. BACKGROUND

The field of AI ethics has relatively long-established origins in areas such as information ethics [18], machine ethics [1] and is related to the domain of Fairness, Accountability and Transparency (FAT) [9]. Societies are witnessing a sharp turn from mere research interest in AI ethics towards a growing urgency for applied ethical considerations in relation to issues already manifesting in the systems. From algorithmic biases, such as Amazon's recruitment application that favoured male applicants [13], to accountability matters seen in the self-driving Uber that killed a pedestrian, whereby the 'safety driver' was rendered liable for negligence [14], ethical issues have arisen in all shapes and sizes. Other prominent matters include the possible threats to human employment, black box dilemma of inexplicable algorithmic development [2], cyber security challenges [23], and technological singularity [10].

Contemporary scholarly efforts in AI ethics represent a practical departure from the philosophical roots of ethics. Applied ethics (pragmatic consideration of morals in action) [11], and utilitarian ethics (determining right from wrong through analysing outcomes) [15], and even virtue ethics (traditionally also connected to business ethics - relating to the characteristics of a person/entity in terms of morality and honesty) [19] are some of the ethical schools that are currently being brought into predominant AI ethics discussions. In AI ethics there is the understanding that these ethical traditions allow for clear exemplification of ethics in action. This enables effective translation of concepts, when designers and developers can see what the issues are in practice. Each of these fields has a deeper philosophical basis that may be further explored as AI systems become more complex.

A. Robot Ethics

AI and robot ethics are undoubtedly related. Yet, while there are commonalities and intersections, there are additionally distinctions between the two [28]. There are several reasons for this. One of the main differences between the two is that robots and robotics in general can be characterised through their embodiedness [7]. Even though there are numerous forms of software-based robots in existence (chatbots, virtual assistants, expert systems etc.), they are understood as entities in themselves. Where AI from a unit perspective may be undefinable, robots have clear boundaries [25].

Rapid development in robotics involves a host of social, legal and political issues. This development also gives rise to concern and interest in the ethical challenges that new robotic technologies pose. The robotics revolution promises a range of enticing and innovative benefits, but, like other emerging technology, it also brings with it challenges and new questions for society to address. The traditional dyadic model of human-robot interaction (HRI) serves as a useful tool for conceptualising the interaction between humans and robots. Yet, it fails to account for the complexity of the systemic framework within which it operates and exists. Even though robots are definable as units or entities, they do not exist as single-standing (self-contained) or autonomous objects per se. Rather, they function and embody highly complex systems of connectivity [35]. Thus, HRI should be seen as rather human–robot–system interaction [34].

Robot units draw on numerous nodes of collective knowledge (information sources) in order to be able to operate and learn [30]. This matter becomes even more complicated when considering the development of robots that are intended to cooperate (work in tandem with others) and collaborate (co-produce and jointly affect) with other robots. Indeed, contrary to what may be thought of in terms of for instance, swarm robotics, multi-robot systems can be understood as systems or situations in which individual robots – robots with separate operating systems and identities – interact with other individual robots [22]. This brings another dimension to robot ethics, in that no longer do scholars purely discuss cybernetic ethics (robots in relation to biological organisms) but additionally, robots in relation to robots [16].

B. Multi-robot Cooperation

Multi-robot cooperation is an established area of research and development. Significant advancements were made during the 1990s (see e.g., [29]) with software such as ALLIANCE that allowed robots to learn and behave in an adaptive and seemingly social way in relation to other robots. The technology has been developed as a means of addressing the needs for entities to undertake tasks that would otherwise be dangerous or hazardous to human health, safety and well-being [31]. One could say that an antecedent behind the development of multi-robot teams has been concern for the humane treatment of human beings [29]. As with human beings, complex operations are difficult if not impossible to undertake via single entities even if often the work of one stands out [4]. This means that if robots would not be working in tandem with other robots the inevitable co-team members would be humans. This is not to say that humans are not a part of multi-robot cooperation, rather, from the embodied understanding of robotics, it means that humans may be physically removed from hazardous environments.

III. ECCOLA

ECCOLA is both a method and a tool for transforming ethical principles into tangible action points within AI development processes [33]. As a method, it aggregates extensive research from the past decades. In particular, three main sources have been influential in ECCOLA's development: 1) IEEE's Ethically Aligned Design [5]; 2) Ethics Guidelines for Trustworthy AI [32]; and Jobin et al's [21] comprehensive review of AI ethics guidelines. It comprises a deck of 21 cards featuring eight themes: 1) analyze, 2) transparency, 3) data, 4) agency and oversight, 5) safety and security, 6) fairness, 7) well-being, and 8) accountability. Each card is broken down into designated sections: a) motivation – explaining the rationale for the theme; b) a list of several questions assisting developers in the consideration of certain ethically motivated questions; and c) a practical example highlighting an illustrative scenario and that highlights ethically motivated concerns.

The themes in turn are divided into explanatory sub-themes: 1) analyze – stakeholder analysis; 2) transparency – types of transparency, explainability, communication, documenting trade-offs, traceability, system reliability; 3) data – privacy and data, data quality, access to data; themes 4 and 5 feature sub-themes that correspond with the thematic titles (human oversight and system safety and security); 6) fairness – accessibility, stakeholder participation; 7) well-being – environmental impact, societal effects; and 8) accountability – auditability, ability to redress, minimizing negative impact.

ECCOLA can be applied in a number of ways. The most popular being in a workshop setting where professionals choose a sub-set of cards (e.g., 3-5 cards) that are meaningful for the projects in question. The workshop process is undertaken in sprints in which case elements are probed through the contents and instructions of the cards. Within workshops, other techniques may be adopted such as card sorting [27] and speed dating [17]. Another technique that is widely used in Agile software development is that of user stories [12]. User stories are informal descriptions of software systems and features from the perspective of human users. Halme et al. [20] demonstrated that through combining the ECCOLA method with user stories, development teams were able to generate actionable ethically concerned human-centered information that could readily be applied to development processes.

The main emphasis of this paper is on exploring the potential for ECCOLA as a tool to enhance the development of R-R cooperation in the CACDAR project from an ethical perspective. Here, we focus on the preliminary stage of mapping the ethical concerns that emerge when considering multi-robot cooperation, comparing these to the ECCOLA themes and then hypothetically examining them via a CACDAR R-R use case.

IV. CACDAR: Project for Robot-to-Robot Cooperation Development

In this section, we describe the CACDAR [26] project on R-R cooperation, through which we apply ECCOLA. CACDAR is a project that focuses on collaboration and cooperation between diverse robots. Cooperation refers to the process of working together in general, while collaboration delineates the process of working together to create something. Here, we mostly refer to cooperation for its generic quality, as the robots in the CACDAR project possess their own goals. Some of the goals (both between robots or within a single robot) may conflict with other goals. To this end, the CACDAR project has two main elements: (1) a *proof of concept software* (simply the software from now on) for individual robots and (2) a *three world cooperation development process*. The software enables cooperation between the robots and the three world development process. This enables the incremental additions of new software features, and thus, robots. Next, we briefly describe the main properties of both.

The three world development process is used to tackle challenges in building the software and testing the R-R cooperation processes. The main idea is that the software should run on three different world types:

- BW: simple 2D Block World with discrete time,
- VW: 3D Virtual World with simulated physics, and
- RW: Real World with real, physical robots.

During the development process, all the worlds can be utilized at the same time and focus on specific aspects of the collaboration: BW concentrates on testing the base collaboration logic, communication procedures, etc., VW verifies collaboration procedures developed in BW and acts as a bridge between BW and RW, and RW provides case-by-case tests, e.g., in actual robot deployment settings. Multiple simulations on both BW and VW can be run at the same time with different robot composition and other parameters while an individual robot operates only in one of the worlds.

In the project, the core software needs to operate in all three worlds. To accommodate for this, the software has two main parts: the main cooperation "brain", which works in all three worlds and all robot platforms (ROS, ROS2, etc.), and the platform (and world) specific code which allows the robot to operate in its current deployment world (e.g. VW) using specific platform (e.g. ROS2). The overall idea is that when the main brain works in all the worlds, the designers can develop, e.g., joint plan formation in BW and then immediately bring the developed algorithm to VW/RW to test how it works. In a similar manner, the lessons learned from RW can be transmitted to BW/VW, either by changing the robot code or by adding some new aspects to the simulated environments (BW/VW) that had not been previously modeled. Currently, the project uses ROS2 as its main platform.

The robots cooperate by constructing *plans* for specific *goals*. The plans are (partially ordered) sequences of *tasks* where each task is appointed to a specific robot that can complete it by executing one or more *actions*. Plans and tasks are platform independent. The brain operates on them and they can be communicated and agreed between robots. Actions are platform (and world) specific, they tell the robot how to use its, possibly simulated, sensors and motors to achieve the tasks. That is, two robots with different platforms might not be able to understand each others' actions that are implemented to complete the same task.

The plan formation and knowledge of "how the world works" is enabled by ontology-based knowledge and reasoning. The project utilizes DUL (DOLCE+DnS Ultralite) ontology[1] as top-level ontology and minimally extends it to accommodate for the collaboration requirements. The robots can exchange information about the world, and update their own world views based on the observed/acquired information and ontological structures. The robots do not share the whole world state with each other, but rather aim to only exchange information that is relevant for the current "context", i.e. plan or goal. For example, is the robot able to do a certain task, e.g. open a door, and, in case it is able, how long it would estimate the task would take (given that the robot might have some

[1] http://ontologydesignpatterns.org/wiki/Ontology:DOLCE+DnS_Ultralite

previously agreed responsibilities that it needs to complete before it can move on to the task). The robots also maintain models of themselves and their peers, called *peer models*. The peer models hold, e.g., capabilities of the robots and are used when reasoning which of the peers are capable of doing a certain task or from which peer a particular piece of information is requested.

To circumvent the symbol grounding problem, the environment is expected to provide cues about the ontological information physical objects within it hold. To this end, in VW and RW, the objects of interest (doors, trashcans, service stations, other robots, etc.) need to have QR codes which the robots can read to identify the object and access information relevant for it, such as its unique ID and type, as well as update its knowledge about the physical object's current location in case it can move. Some of this object detection might be changed to more robust machine learning-based methods, but as the project currently (1) operates on very limited hardware and (2) requires the computation happen in the physical robots and not in the cloud, the number of machine learning models the individual robots can utilise is limited.

V. Method - gathering and analysing robot-to-robot ethical concerns

ECCOLA has been applied to plentiful software development cases. Yet, the domain of robotics and particularly multi-robot and R-R cooperation are still unexplored through the frames of ECCOLA. Before contaminating the territory of CACDAR development with notions of the themes already presented in ECCOLA, we decided to conduct a workshop to collect at least 100 perceived ethical concerns regarding multi-robot cooperation. The motivation behind collecting 'raw' ethical considerations was to examine the match between the pre-existing ECCOLA themes and the multi-robot cooperation context. The intention was to reveal the strengths and relevance of this AI ethics tool, and what would need to be added or adjusted for this new development context of multi-robot cooperation.

Ten researchers took part in the workshop. An understanding of CACDAR was used as a base point for conceptualising multi-robot cooperation. Yet, ideas were not limited to this strict context. Rather, the activity was intended to benchmark the ethical landscape of multi-robot cooperation. In total, N=149 ideas[2] were collected during the workshop. Afterwards, the ideas were compiled on an excel spreadsheet. An external researcher with no previous experience of either ECCOLA or CACDAR was recruited to perform a blind analysis of the data in relation to emergent themes. Short descriptions were provided with the emergent themes. The final stage of this light-weight applicability measure was to connect the emergent themes to the ECCOLA framework. An extra workshop was used as a synergy platform for connecting ECCOLA to the actual CACDAR development and exploring possibilities for further co-development of the framework and the software.

[2]bit.ly/ethical_concerns

A. Emergent ethical themes in multi-robot cooperation

As a result of the blind analysis 12 emergent themes were identified. These themes were: defect, empowerment, ensuring data privacy, confrontation or incompatibility, human inefficiency, mutual cooperation, data driven inefficiency, ensuring ethical concerns, ensuring security, ensuring safety, designing and accountability. Figure 1 illustrates the study design and distribution of themes in relation to concern frequency.

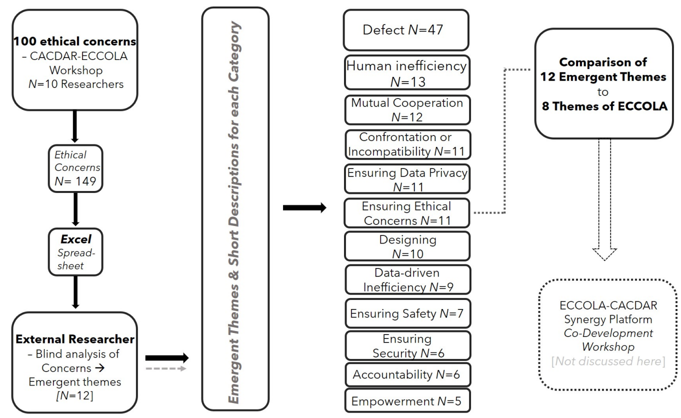

Fig. 1. Study design and emergent themes

The concerns mostly raised in the workshop pertained to defects (47) of the system. Thus, there was foresight and concern regarding the quality of the system. The second greatest ethical concern for R-R cooperation was human efficiency (13). This was roughly equal to human inefficiency (12) – i.e., "lack of human control", and "lack of system understanding" (also giving rise to 'blackbox' algorithmic issues). Ensuring data privacy, confrontation or incompatibility, and ensuring ethical concerns were the next most frequent mentions (each 11). Accountability (6) was associated with ownership and actor networks, problematising the effects of already complex techno-accountability relationships.

Interestingly, from the raw data some other themes emerged that were not included in the external researcher's coding – yet would cause reason to iterate the thematic coding process. These themes included: economic matters and issues pertaining to intellectual property; regulatory challenges and aligning system design to conform with regulations and standards, while also accounting for extra ethical layers (potentially relating to contextual sensitivity); and the clash between false promises and expectations. In fact, this final concern causes great challenges in human relations to emerging technology as a whole [21].

The comparison of multi-robot ethical concern themes with ECCOLA's framework revealed some overlaps as well as points for adjustment (see Table I). It must be emphasized that ECCOLA has been developed over years of iteration. The present categorization by one researcher in this current study does not match the cross validation of many in ECCOLA. The theme of 'Designing' for instance, in the multi-robot concerns could refer to considerations for the

TABLE I
COMPARISON OF MULTI-ROBOT CONCERNS WITH ECCOLA

Themes	ECCOLA
Accountability	Accountability
Confrontation OR Incompatibility	-
Data Driven inefficiency	Data
Defect	Safety & Security
Designing	-
Empowerment	Fairness
Ensuring data privacy	Data
Ensuring ethical concerns	-
Ensuring safety	Safety & Security
Ensuring Security	Safety & Security
Human inefficiency	Agency & Oversight
Mutual cooperation	-
-	Transparency
-	Well being
-	Analyze

designs themselves or even design related matters such as process, designer relationship to artifact, or consideration for stakeholders and ethics. 'Ensuring ethical concerns' is rather vague. Yet, two thematic constructs are particularly interesting: 1) mutual cooperation – alluding to exchange and sharing of workload; and 2) confrontation or incompatibility – conflict, clash of values or logic. These are both specific to multi-robot cooperation and could prove a valuable edition to ECCOLA when considering all types of cooperative systems.

VI. ETHICAL USE CASE OF CACDAR

In order to observe how the adjusted ECCOLA framework may operate in the context of CACDAR we illustrate a use case scenario. The use case description is combined with the presentation of ethical concerns on a general level. We then outline these preemptive CACDAR development concerns in relation to ECCOLA. These are outlined in terms of what they mean for the subsequent development process. ECCOLA and the additional concerns identified are further discussed regarding their qualities for early intervention and systematization of matters such as accountability.

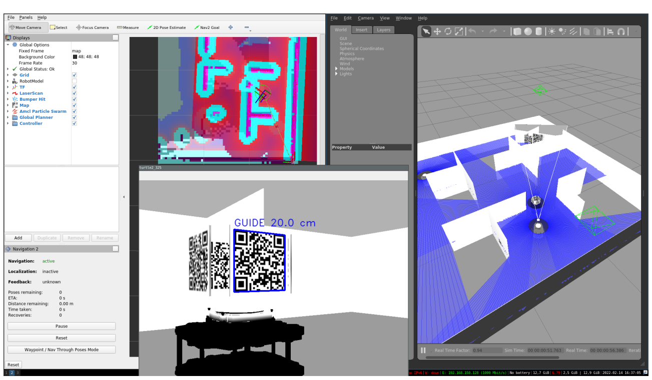

Fig. 2. Cooperation scenario running in the virtual world (Gazebo).

The use case scenario is of delivery robots (see Figure 2) for which we also provide a video link[3]. The example is of a situation in which multiple robots co-exist in an arbitrary environment (e.g., postal office). Each robot has its own goals that prescribe the tasks the robot will perform. In this case, the goals include, e.g., package transfer or floor cleaning. Besides robots native to the environment, a completely new robot might be introduced. This robot may not necessarily be familiar with the environment and therefore, will require guidance to achieve its own goals. The guidance in this case is achieved by a robot that is native to the area that acts as a guide for the newly introduced machine enabling it to complete its task. Note however, that being a guide might not be the sole duty of the native robot; it only acts as a guide in situations where it is able to. The decision of whether a robot is suitable as a guide is based on agreement between the robots regarding how the cooperation should be achieved. If cooperation can be achieved, the assisting robot must temporarily reconsider its goals and proceed with helping the foreign machine.

[3] https://tinyurl.com/cacdar-demo

A. *Identified concerns of the CACDAR case*

Taking a general approach to identifying the ethical concerns present in this use case places our analysis at the level of the identified multi-robot ethical themes. 'Ensuring data privacy' is the first theme that emerges in this hypothetical use case. Already from the outset, a significant ethical concern in CACDAR is the collection of irrelevant data. Information, such as floor maps in which robots operate, is often collected via sensors that prevent robots from colliding with arbitrary objects. However, the collected location data can be sensitive and thus, require appropriate mechanisms ensuring that such data does not leak to the general public and does not expose private information. Such a requirement is hard to satisfy, given the fact that the CACDAR project leverages ROS 2, which embraces the usage of third-party packages. Thus, the second theme present in this use case is 'Ensuring security'. As a result of third-party packages, the code responsible for handling sensitive data might include malicious behavior causing undesired consequences.

Another ethical concern is related to unfulfilled agreement between the robots. In CACDAR, a robot with an empty task queue may begin to work on a goal of its own or a requested goal invoked by another robot. However, a robot might not abandon a task it has already started; it is programmatically ensured that the robot always completes at least one goal at a time and does not start anything else in the meantime. Nonetheless, it is possible to have a continuous goal that might take a long time to complete. Therefore, the robot that asks for assistance may need to wait before the request is completed.

The last problem we consider concerns the security aspects of collaborative R-R interaction, which once more refers to the 'Ensuring Security' theme. The underlying communication protocol of ROS 2, Data Distribution Service (DDS), implements some security mechanisms that make the communication more secure. However, several shortcomings have been identified in one DDS implementation regarding its security mechanisms [24]. This implies that the robot can be subject to cyber-attacks, potentially risking the integrity of the system and safety of its surroundings.

B. ECCOLA analysis for multi-robot collaboration

In order to verify the ECCOLA framework in relation to CACDAR cases, and synergize understandings of the logic of respective projects an additional workshop was held. During this workshop the CACDAR team understood the ECCOLA process in a five card sprint. Thus, within the presented use case, concerns mentioned resonate with the original ECCOLA framework. For instance, data related concerns such as ensuring data privacy and data driven inefficiency were mentioned. Safety and security were included in relation to the threats caused by third-party packages and cyber-attacks. Furthermore, in a detailed use case scenario a thorough application of ECCOLA as a tool for analysis and backlog creation would illustrate aspects of CACDAR that relate to most, if not all ECCOLA themes. Aspects such as human agency and oversight, and reinforcing transparency within the multi-robot processes would be elaborated upon in respect to broader stakeholder and agent concerns. This would also give rise to opportunities for exploring the systematization of accountability. Yet, as noted from the comparison of themes, application of ECCOLA in the context of CACDAR and R-R cooperation gives rise to new opportunities and deeper considerations for ethics in multi-agent, connected, and collaborative systems. These can be applied to a range of other contexts exhibiting similar attributes (i.e., IoT, swarm robotics etc.).

VII. Conclusion

In this paper we described ECCOLA as a tool and method for developing ethically aligned AI. We illustrated a new application case - the CACDAR project. A workshop intended to identify 100 ethical concerns of multi-robot collaboration resulted in 149 identified concerns. An analysis of the concerns by an external researcher revealed 12 emergent themes. Here, we focused on presenting a pragmatic approach to instil ethical considerations in the development of multi-robot cooperation. The tool (and method) used was ECCOLA. ECCOLA has been developed over several years. Originating as a synthesis of several critical sources of AI ethics research and principles, then being iterated by scholars in applied design and development processes, ECCOLA is gradually advancing towards a stage of high reliability. Through shifting systemic contexts from AI in general to the more specific domain of multi-robot cooperation, we have seen that some themes should be added to the ECCOLA framework. With this said, three significant contributions have arisen from this study:

- Application of ECCOLA in multi-robot cooperation.
- 149 ethical concerns for multi-robot cooperation.
- Emergent awareness of matters such as mutual cooperation and confrontation in multi-robot cooperation.

The study was limited by the use of only one external analyst and should be repeated with more groups of researchers and developers for validity. Future research steps involve further research and probing into the ethical concerns that may arise within the field of multi-robot cooperation. This will be followed by rigorous analysis of themes which are iterated, validated (construct validation) and incorporated into ECCOLA.

Acknowledgment

The work was supported by the Academy of Finland (project 328729), JYU StartUp Lab, University of Jyväskylä, Finland, and the School of Marketing and Communication and Digital Economy Platform, University of Vaasa, Finland.

References

[1] Colin Allen, Wendell Wallach, and Iva Smit. "Why machine ethics?" In: *IEEE Intelligent Systems* 21.4 (2006), pp. 12–17.

[2] Mike Ananny. "Toward an ethics of algorithms: Convening, observation, probability, and timeliness". In: *Science, Technology, & Human Values* 41.1 (2016), pp. 93–117.

[3] Jani Antikainen et al. "A Deployment Model to Extend Ethically Aligned AI Implementation Method ECCOLA". In: *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*. IEEE. 2021, pp. 230–235.

[4] Margaret Arnd-Caddigan. *Sherry Turkle: Alone together: Why we expect more from technology and less from each other*. 2015.

[5] IEEE Standards Association et al. *Ethically Aligned Design, Version 2 (EADv2)*. 2018.

[6] Christoph Bartneck et al. *An introduction to ethics in robotics and AI*. Springer Nature, 2021.

[7] Nicolas Bredeche, Evert Haasdijk, and Abraham Prieto. "Embodied evolution in collective robotics: a review". In: *Frontiers in Robotics and AI* 5 (2018), p. 12.

[8] Joanna Bryson and Alan Winfield. "Standardizing ethical design for artificial intelligence and autonomous systems". In: *Computer* 50.5 (2017), pp. 116–119.

[9] Vincent Chiao. "Fairness, accountability and transparency: notes on algorithmic decision-making in criminal justice". In: *International Journal of Law in Context* 15.2 (2019), pp. 126–139.

[10] Mark Coeckelbergh. *AI ethics*. MIT Press, 2020.

[11] Andrew I Cohen and Christopher Heath Wellman. *Contemporary debates in applied ethics*. John Wiley & Sons, 2014.

[12] Mike Cohn. *User stories applied: For agile software development*. Addison-Wesley Professional, 2004.

[13] Jeffrey Dastin. "Amazon scraps secret AI recruiting tool that showed bias against women". In: *Ethics of Data and Analytics*. Auerbach Publications, 2018, pp. 296–299.

[14] Madeleine Clare Elish. "Moral crumple zones: Cautionary tales in human-robot interaction (pre-print)". In: *Engaging Science, Technology, and Society (pre-print)* (2019).

[15] Amitai Etzioni and Oren Etzioni. "Incorporating ethics into artificial intelligence". In: *The Journal of Ethics* 21.4 (2017), pp. 403–418.

[16] Zhi Feng et al. "An overview of collaborative robotic manipulation in multi-robot systems". In: *Annual Reviews in Control* 49 (2020), pp. 113–127.

[17] Eli J Finkel and Paul W Eastwick. "Speed-dating". In: *Current Directions in Psychological Science* 17.3 (2008), pp. 193–197.

[18] Luciano Floridi. "Information ethics: On the philosophical foundation of computer ethics". In: *Ethics and information technology* 1.1 (1999), pp. 33–52.

[19] Thilo Hagendorff. "The ethics of AI ethics: An evaluation of guidelines". In: *Minds and Machines* 30.1 (2020), pp. 99–120.

[20] Erika Halme et al. "How to write Ethical user stories? Impacts of the ECCOLA Method". In: *International Conference on Agile Software Development*. Springer, Cham. 2021, pp. 36–52.

[21] Anna Jobin, Marcello Ienca, and Effy Vayena. "The global landscape of AI ethics guidelines". In: *Nature Machine Intelligence* 1.9 (2019), pp. 389–399.

[22] Ameer Tamoor Khan, Xinwei Cao, and Shuai Li. "A survey on blockchain technology and its potential applications in distributed control and cooperative robots". In: *arXiv preprint arXiv:1812.05452* (2018).

[23] Zarina I Khisamova, Ildar R Begishev, and Elina L Sidorenko. "Artificial intelligence and problems of ensuring cyber security". In: *International Journal of Cyber Criminology* 13.2 (2019), pp. 564–577.

[24] Jongkil Kim et al. "Security and performance considerations in ros 2: A balancing act". In: *arXiv preprint arXiv:1809.09566* (2018).

[25] Bert-Jaap Koops, Mireille Hildebrandt, and David-Olivier Jaquet-Chiffelle. "Bridging the accountability gap: Rights for new entities in the information society". In: *Minn. JL Sci. & Tech.* 11 (2010), p. 497.

[26] Niko Mäkitalo et al. "Towards Novel and Intentional Cooperation of Diverse Autonomous Robots: An Architectural Approach". In: *Companion Proceedings of the 15th European Conference on Software Architecture*. 2021.

[27] Juan Antonio Morente-Molinera et al. "A dynamic group decision making process for high number of alternatives using hesitant Fuzzy Ontologies and sentiment analysis". In: *Knowledge-Based Systems* 195 (2020), p. 105657.

[28] Vincent C Müller. "Ethics of artificial intelligence and robotics". In: (2020).

[29] Lynne E Parker. *Heterogeneous multi-robot cooperation*. Tech. rep. Massachusetts Inst of Tech Cambridge Artificial Intelligence Lab, 1994.

[30] Akhlaqur Rahman et al. "Motion and connectivity aware offloading in cloud robotics via genetic algorithm". In: *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE. 2017, pp. 1–6.

[31] Mac Schwager et al. "A multi-robot control policy for information gathering in the presence of unknown hazards". In: *Robotics research*. Springer, 2017, pp. 455–472.

[32] Nathalie Smuha. "Ethics guidelines for trustworthy AI". In: *AI & Ethics, Date: 2019/05/28-2019/05/28, Location: Brussels (Digityser), Belgium*. 2019.

[33] Ville Vakkuri et al. "ECCOLA—A method for implementing ethically aligned AI systems". In: *Journal of Systems and Software* 182 (2021), p. 111067.

[34] Aimee Van Wynsberghe and Shuhong Li. "A paradigm shift for robot ethics: From HRI to human–robot–system interaction (HRSI)". In: *Medicolegal and Bioethics* 9 (2019), p. 11.

[35] Shaohua Wan, Zonghua Gu, and Qiang Ni. "Cognitive computing and wireless communications on the edge for healthcare service robots". In: *Computer Communications* 149 (2020), pp. 99–106.