Business and Ethical Concerns in Domestic Conversational Generative AI-Empowered Multi-robot Systems

Rebekah Rousi[1(✉)], Hooman Samani[2], Niko Mäkitalo[3], Ville Vakkuri[1], Simo Linkola[4], Kai-Kristian Kemell[4], Paulius Daubaris[4], Ilenia Fronza[5], Tommi Mikkonen[3], and Pekka Abrahamsson[6]

[1] University of Vaasa, Wolffintie 32, 65200 Vaasa, Finland
rebekah.rousi@uwasa.fi
[2] University of Arts London, 272 High Holborn, London WC1V 7EY, UK
[3] University of Jyväskylä, Mattilanniemi 2, 40100 Jyväskylä, Finland
[4] University of Helsinki, Yliopistonkatu 4, 00100 Helsinki, Finland
[5] Free University of Bozen-Bolzano, Sparkassenstraße 21 - via Cassa di Risparmio, 21 39100 Bozen-Bolzano, Italy
[6] University of Tampere, Pohjoisranta 11A, 28100 Pori, Finland

Abstract. Business and technology are intricately connected through logic and design. They are equally sensitive to societal changes and may be devastated by scandal. Cooperative multi-robot systems (MRSs) are on the rise, allowing robots of different types and brands to work together in diverse contexts. Generative artificial intelligence has been a dominant topic in recent artificial intelligence (AI) discussions due to its capacity to mimic humans through the use of natural language and the production of media, including deep fakes. In this article, we focus specifically on the conversational aspects of generative AI, and hence use the term Conversational Generative artificial intelligence (CGI). Like MRSs, CGIs have enormous potential for revolutionizing processes across sectors and transforming the way humans conduct business. From a business perspective, cooperative MRSs alone, with potential conflicts of interest, privacy practices, and safety concerns, require ethical examination. MRSs empowered by CGIs demand multi-dimensional and sophisticated methods to uncover imminent ethical pitfalls. This study focuses on ethics in CGI-empowered MRSs while reporting the stages of developing the MORUL model.

Keywords: Multi-robot cooperation · Business · Ethics · Conversational Generative AI · Large Language Models

1 Introduction

Generative Artificial Intelligence is currently in the spotlight, drawing both praise and criticism. Conversational AI, on the other hand, has been studied for several years and refers to chatbot technologies which are somehow considered to make the interactions

S. Hyrynsalmi et al. (Eds.): ICSOB 2023, LNBIP 500, pp. 173–189, 2024.
https://doi.org/10.1007/978-3-031-53227-6_13

with the chatbot intelligent. In this article, we use the term Conversational Generative Artificial Intelligence (CGI) to refer specifically to the combination of generative and conversational artificial intelligence (AI). It has permeated every corner of society, revolutionizing communication between humans and machines using natural language. Two fields significantly impacted by this technology are business and robotics. Integrating CGI into organizational operations can yield substantial business value [1]. Similarly, employing CGI in robotics enhances usability, accessibility, and the market potential of robotic systems [2]. However, embracing these cutting-edge technological developments is not without risks. Recent headlines in major media outlets have underscored the potential consequences of mishaps in sophisticated data-driven systems for humans, technology, and businesses alike.

One of the primary contexts for deploying these complex emerging products and services is the home. For instance, the global smart home market is projected to grow from $93.98 billion in 2023 to $338.28 billion by 2030 [3]. This rapid growth in the market introduces a complex landscape, integrating multi-layered Systems of Systems (SoSs) into the traditionally private and sacred space of the home [4, 5]. Everyday products such as refrigerators, vacuum cleaners, and toasters are transforming into intelligent devices with the potential to function as discreet communicators [6]. Consequently, ethical considerations are intertwined with all levels of technological implementation in the home due to the changing dynamics in human-object relationships [7].

The presence of CGI-embedded Multi-Robot Systems (MRS) in domestic settings raises a multitude of ethical concerns for businesses [8, 9]. The development of CGI-embedded MRSs has predominantly focused on industrial and business applications [10]. These systems aim to automate tasks and enhance efficiency in various industries, including manufacturing, healthcare, and customer service. As a result, the ethical dimensions of CGI-embedded MRSs have often been overlooked. Businesses engaged in the development or deployment of CGI-embedded MRSs must carefully consider these ethical concerns and take steps to address them. This paper adopts an applied ethics approach to explore potential ethical issues arising from the development and deployment of data-driven multi-robot cooperative systems. Applied ethics, in this context, refers to a case-specific approach that examines how social ethical dilemmas manifest practically when specific technical and social-technical elements (involving a blend of human and technological factors) are put into operation in specific contexts [11].

Instead of seeking to already *solve* problems, this study primarily focuses on *identifying* potential ethical challenges during the development, deployment, and implementation of multi-robot cooperative systems for implementation in the home. As this is a novel context in the area of AI ethics, we consider such problem identification important at this stage. In this respect, we consider the concept of moral awareness essential in order to go beyond the concerns voiced in existing literature on AI ethics. *Moral awareness* is defined as the ability to identify ethical aspects in a given context [12]. In this paper, a scenario-based approach is employed to investigate the potential ethical concerns and moral implications of introducing heterogeneous multi-robots into domestic spaces.

More specifically, the authors aim to develop a model for promoting moral awareness in multi-robot systems (MRSs) – the MORUL model. Furthermore, the authors recognize that not all ethical issues and related interventions can be addressed during the pre-development phases. In the emerging MORUL model, ethical concerns are mapped and predicted in relation to stages at which analyses should be conducted. These analyses are carried out with regard to the dimension affected by the ethical concern, such as safety, security, or societal impact. This paper contributes to and builds upon previous efforts that sought to establish ethical practices and frameworks for the development of artificial intelligence (AI) [13].

2 Background

2.1 Large Language Models (LLMs) in Multi-robot Cooperation

Large Language Models (LLMs) and Generative Artificial Intelligence represent some of the latest developments in machine learning that have gained widespread public attention. OpenAI's Generative Pre-training Transformer architecture (ChatGPT) has been at the center of headlines and public debates since around 2018 [1]. LLMs are part of the recent trend in the growing popularity of chatbot development [14], which make Conversational Artificial Intelligence stand out as an advancement towards higher AI development goals such as Artificial General Intelligence (AGI). Hence, we use the term 'Conversational Generative Artificial Intelligence' (CGI) in this article to be specific about the technology we are referring to. In the case of chatbots, Natural Language Processing (NLP) is employed to interact with users by providing optimal responses from the information system. ChatGPT can be viewed as an advanced form of chatbot, enhancing earlier versions by combining deep learning and LLMs [15]. LLMs focus on predicting word sequences commonly used in human communication. However, this process introduces biases and discrimination due to the reliance on neural network transformer architectures and deep learning, which depend on representative data [16]. For instance, ChatGPT combines supervised fine-tuning with unsupervised pre-training to generate responses that appear to be human-like, thus expanding the social dimension of human-data interaction and improving data accessibility for non-experts.

Currently, engaging in prompt-based conversations with AI-based chatbots can be relatively expensive, considering the number of prompts typically required for a single task and the widespread usage of these models. Tech companies like OpenAI, Microsoft, Alphabet, and Meta are striving to capitalize on this emerging technology by building businesses around AI-based applications for personal and professional use. Given the costs associated with training and running these models, companies are competing with diverse business strategies. OpenAI, for example, offers its GPT model as a service via an API, allowing new AI-based applications to be developed on top of their models. Meanwhile, new open-source LLMs with various capabilities and licenses are being released on the internet. Meta, for instance, provides its advanced LLAMA 2 model as open source, with limited commercial use.

Multi-robot cooperation involves two or more robots, regardless of brand, model, or type, working together to achieve shared goals [17]. While each robot may have unique objectives, there should be a common overarching goal among them, such as ensuring a safe and clean home or delivering timely and effective services in a hospital. The ultimate goal in such scenarios is typically the well-being of the human owner.

Multi-robot cooperation primarily addresses complex tasks that are nearly impossible to accomplish successfully without a team effort [17, 18]. At all stages, human involvement is a constant factor, whether it's in programming, giving commands, or collaborating with the robots. Consequently, multi-robot cooperation should always be considered in relation to humans and their varying levels of involvement in different processes [19]. Considering human factors in working with multi-robot systems introduces different levels of complexity, as identified by Simões and colleagues [20]: 1) the human operator and the technology itself; 2) recommendations and guidelines affecting the performance of human-robot teams; and 3) complex holistic approaches guided by recommendations and guidelines that influence human-robot interaction.

In any case, it is essential to recognize that the human dimension in multi-robot cooperation is always the result of complex negotiations between integrated systems, diverse operational goals, varied corporate strategies, governed by standards, laws, and recommendations. Therefore, the starting point for examining such systems always begins at Level 3 [20]. Preempting ethical issues during the pre-development phase elevates the investigation to Level 4, involving systemic ethical forecasting in cybernetic systems. This forecasting requires an understanding of how Multi-Robot Systems (MRSs) operate within human contexts, with communication playing a crucial role [21]. Communication not only involves the functional aspects of human interaction with multi-robot systems but also encompasses the social-emotional components of Human-Robot Interaction (HRI) [21, 22]. As a result, CGI in forms such as ROSGPT or ChatGPT has significantly impacted the ways people interact with machine learning systems [23].

ROSGPT [24] introduces an innovative approach that leverages the full potential of LLMs to enhance human-robot interaction significantly. This framework integrates ChatGPT into ROS2-based robotic systems, creating a synergy between language understanding and robotic control. ROSGPT's advantage lies in its effective prompt engineering, utilizing ChatGPT's versatile capabilities, from information elicitation to coherent train of thought, to convert unstructured natural language commands into precise, contextually relevant robotic instructions. ROSGPT capitalizes on the inherent learning capabilities of LLMs to effortlessly extract structured commands from unrefined language inputs. The proof-of-concept demonstration, highlighting the translation of human language into actionable robotic instructions, underscores ROSGPT's potential across a range of applications. Beyond its immediate utility, ROSGPT's open-source implementation on ROS 2's platform not only fosters collaboration between the robotics and natural language processing fields but also represents a significant step toward the realm of AGI.

2.2 Business Effects of AI Ethics, CGI and Multi-robot Cooperation

Ethics in the domains of AI have been hot topics for decades now, and this is becoming increasingly more so as AI is deployed widely in society. Earlier discussions applied the terms 'information ethics', 'machine ethics' and 'computer ethics' [13, 25] to describe the field of examining ethical and moral implications of IT. With the broadening adoption of AI technologies in a multitude of domains, various practical incidents have highlighted diverse risks associated with AI.

The existing discussion on AI ethics, which far predates recent incidents, has served to identify and understand many of the risks already in the past - before they unfolded in actuality. Now, these predicted risks are becoming real, meaning that they present practical issues enabled by recent progress in ML. These risks are typically approached in research and development through *principles* in AI ethics [13]. For instance, racism, which is often associated with the principle of fairness, not only manifests through abuse and degradation, but also false accusation (see e.g., [26]). There is a sense of urgency spurred from the already emergent incidents involving machine learning (ML) technology utilization [25]. Whether the incidents involve matters of accountability and responsibility as witnessed in accidents in which human life has been harmed or damaged. The AI Incident Database [26] reported 90 incidents in 2022 alone, of AI-caused accidents, 45 already at the beginning of 2023. The rate of AI incidents seems to be increasing at a comparative pace to Moore's Law - doubling every year, similarly to the compounding capacity of computing speed [27]. These not only incur substantial costs in damages and potential insurance premiums, but pose serious problems from basic issues of human respect, safety, and dignity, to the severe tarnishing of reputation for businesses who do not embrace humane factors as a part of their data-driven business strategy [28].

The 2018 self-driving Uber accident in which a pedestrian was fatally wounded (see e.g., [29]) incurred irreparable immaterial damage. This no doubt contributed to loss of income, hindered self-driving vehicle development (and brands), tarnished Uber (now owned by Aurora Innovations) as a transportation service, and the operator who was responsible for monitoring the vehicle. While the human operator has been found guilty of negligence, the repercussions of the accident in terms of legal expenses and loss of consumer trust are remarkable. Not only were the direct implicated actors affected, but the US Federal Government was also accused of not properly regulating the industry. Moreover, had the accident led to a total abandonment of self-driving vehicles by companies such as Uber, profit trajectories would be thrown off course, because drivers account for 80% of all costs - self-driving units being evaluated at 7 billion United States dollars already in 2020 [3029].

Business intercedes on many dimensions of AI and robot ethics. From privacy-related issues and dark practices of the surveillance economy, to platform economy logic, and 'login – lock-in' cultures, business needs to be considered from both back and front-end perspectives. When it comes to ethics, business itself can be its own worst enemy. The logic that may pave the way to patents and trade secrets, may be guilty of fostering ethical potholes such as black box systems diminishing customer and user trust, and even simply, bad user experience with greater social repercussions. The dance between ethics and business is like a temptation-filled devil's tango. The appeal of fast profits blinds many of careful foresight in business strategy. Effective management of ethics in AI and robotic development would not just mean better business strategy, but also longevity [31].

3 Method

In the present study the researchers employed a qualitative exploratory method via two workshops. A scenario-based approach was used to contextualize the inquiry that entailed imagining that several robots of different use purpose, brand and type, utilizing CGI technology were implemented in the home (see Fig. 1). In the scenario, two cleaning robots of the same brand and make have been used in the home for quite some time. The new addition of a robot arm from a different brand and manufacturer elicits ethical concerns when considering the need for all robots to cooperate in order to perform tasks to reach certain goals. The goal of the workshops was to spark moral awareness in the participants in order to recognize ethical concerns and compare the identified concerns to those existing via previous research, and found in AI ethics guidelines and principles. The workshops were held at separate times: Workshop 1 (W1) was held during February, 2023, for two days face-to-face at a lab hosted by one of the participating research institutions; and Workshop 2 (W2) was held in June, 2023, for one hour via Zoom. The idea behind the separate timing was to allow for the analysis of W1 results, in order to synthesize and construct a preliminary framework for W2. The preliminary framework was seen as the basis for modeling a matrix that eventually will serve as a scaffolding for ethical multi-robot development. The matrix would include facets starting from ethical business strategy (understanding the influence of economic superstructures in molding the logic of technological products), to hardware and software, human-technology interaction, larger societal repercussions, and back again to business impact.

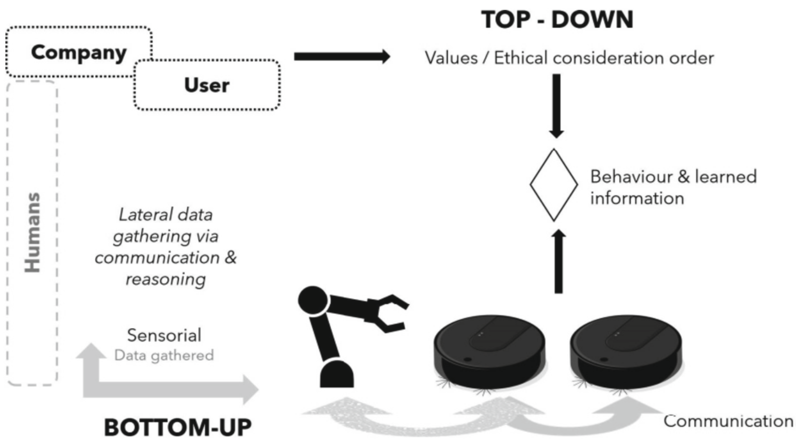

Fig. 1. Domestic scenario of two cleaning robots and one robot arm - understanding relations between layers and domains of multi-robot cooperation from a techno-corporate perspective

Qualitative data was collected in the form of brainstorming drawings and notes. The material from W1 was originally in paper versions, which were subsequently photographed and digitally archived. The material from W2 was produced on Google Jamboard. During processing of the data - transferal from the drawing boards to excel and image files - preliminary thematic categories were established. Extra rounds of thematic analysis [32] were performed by the research team in an excel document. The study was conducted via a constructivist grounded theory [33] approach in order to build on previous AI ethics principles, guidelines and methods (see e.g., [18], while allowing for deeper examination of specific details and dimensions that are phenomenologically unique to the domain of multi-robot cooperation.

3.1 Ethical and Responsible Research

As this is a novel space of research that deals with ethics across a range of levels, from basic practical levels to higher levels of abstraction, the research team deemed the safest and most responsible approach to be that of internal inquiry. To avoid physical or psychological harm, the team of experts maintained the empirical component outside the realm of physical human-robot or robot-robot interactions. Rather, the researchers deliberated through discussion, illustration and writing. All researchers involved in the workshops were willing participants, agreeing the use of their data, exercising scholarly agency as experts within their respective fields. In compliance with the General Data Protection Regulation (GDPR), all data is stored in secure password-protected digital locations to which only two main researchers have access. No personal data is stored with the research data.

3.2 Participants

Each workshop comprised eight participants, rendering $N = 16$ contributions in total. Five participants participated in both workshops ($N = 10$ contributions) while six participants only participated in one of the workshops. This meant that the overall total of individual participants was $N = 11$. All participants possessed a higher tertiary degree, starting at PhD level researchers and higher. The gender distribution was two females and nine males. The fields of expertise that the participants represent are: software engineering and computer science; robotics and software for robotics; edge intelligence; computing education; information systems; cognitive science; human computer interaction; communication; and social ethics.

3.3 Procedure

The workshops were planned and agreed upon in a series of online meetings. In these meetings the strategy was deliberated, goals were set, as well as timing, procedure and locations were established. The context for the scenario was decided upon via several brainstorming sessions in which the team examined areas, environments and situations in which ethics and moral conduct would be considered as most sensitive [5]. After identifying several domains including education, healthcare, elderly care, and the home, the team selected the home, both for its intimate framing of privacy, as well as its diversity [4]. While there are central features defining a home - living space, kitchen, bedroom etc. - the ways in which people appropriate, populate, and utilize their spaces is quite eclectic [4]. This is as opposed to public institutions such as hospitals that are laden with rules, standards and top-down regulations.

Workshop 1

Workshop 1 took place in person, on location at the lab of one of the participating research institutions. The lab is designed as an innovation space with a central meeting area equipped with audio-visual and teleconferencing equipment, as well as traditional tools such as flipcharts, post-it notes, colored pens. One participant contributed via Zoom for logistical reasons. The workshop was held over a two-day interval. The procedure entailed a round of introductions and articulating our interests in relation to the topic for the participants who had not been involved in the previous online planning sessions. The workshop proceeded as seen in Table 1.

Table 1. Workshop 1 procedure.

Step No.	Step label	Description
1	*Re-cap of use context and scenario*	Narrative unfolds in the home. Two similar robots (vacuum cleaners) and a newly introduced robot arm
2	*Independent mind-mapping of ethical concerns [unstructured]*	Independent work (30 min.), focus on ethical concerns
3	*Group discussion and comparison of findings*	Discussion of mind-maps, sharing ideas and introducing new concerns that arose in the group discussion
4	*Identification of the layers*	Identifying layers implicated in LLM-enabled multi-robots
5	*Model formulation*	Deliberation of actionable models of ethics in multi-robot collaboration that could be utilized within the programming process

Workshop 2

Workshop 2 was carried out via Zoom to allow for international collaboration while some members of the study were traveling. The duration of the workshop was two hours and held on Google Jamboard. Building on the findings of Workshop 1, Workshop 2 was structured according to a matrix of multi-robot cooperation domains and layers: Human-Interaction; Sensorial Layer (robot hardware); Deliberation (robot brain); Behavioral (robot hardware); Communication and Networking (robot-to-robot interaction); and System of Systems (network or systems). From the human perspective, considerations of ethical aspects were encouraged to be thought of through the frames of: 1) safety, 2) security, and 3) societal dimensions. The procedure of Workshop 2 is observed in Table 2.

Table 2. Workshop 2 procedure.

Step No	Step label	Description
1	*Instructions & breakdown of procedure + use-context recap*	Use context is the home and workshop members are encouraged to think of all potential ethical issues and scenarios arising from the introduction of LLM-powered multi-robot cooperation in domestic spaces
2	*Independent mind-mapping of ethical concerns [unstructured]*	Independent work (30 min.), focus on ethical concerns
3	*Group discussion & comparison of findings*	Groups progressed through the domains and layers of multi-robot cooperation as well as the human dimensions of the concerns
4	*Layer and domain refinement*	Group reflected on the earlier version of the layers and domains based on new findings arising in W2
5	*Model refinement*	MORUL model for ethical CGI-enabled multi-robot development further refined

3.4 Analysis

Thematic analysis [32] was employed to analyze the data of both workshops. In the case of Workshop 1, the researchers transcribed mind-maps, notes and illustrations that had been expressed on large flip chart sheets into excel sheets. From Workshop 2, the Google Jamboard notes were transferred into excel. The analysis took place in three steps: 1) sorting data into themes; 2) refining the themes; and 3) performing frequency analysis to determine which themes arose in relation to which layer of the multi-robot systems. The themes were compared between both data sets, and cross-validated among the research team to ensure consensus of the themes and labels. The themes were again reviewed according to the technological layers, as well as the domains (i.e., safety, security, and society) that they are implicated with. The business dimension of the multi-robot ethical concerns has been positioned as a superstructure (economic and logic base) during and after analysis to make sense of the influence that corporate competition through technological design has on the ethical implications from conceptualization to implementation of the multi-robot systems.

4 Results

In total, 21 themes arose from the data. The themes and their quantities varied from Workshop 1 (W1) to Workshop 2 (W2). In W1, the emergent themes from 61 constructs (expressions) were: data security and privacy (3–4.9%); corporate dominance (3–4.9%); communication (17–27.9%); cooperation (10–16.4%); reliability and recover (1–1.6%);

logic and standards (2–3.3%); human oversight (5–8.2%); prioritization/hierarchy (2–3.3%); trustworthiness/virtue (5–8.2%); executive function (2–3.3%); maleficence (3–4.9%), user experience (UX, 6–9.8%); and legislation (2–3.3%). The distribution of frequencies can be seen in Fig. 2.

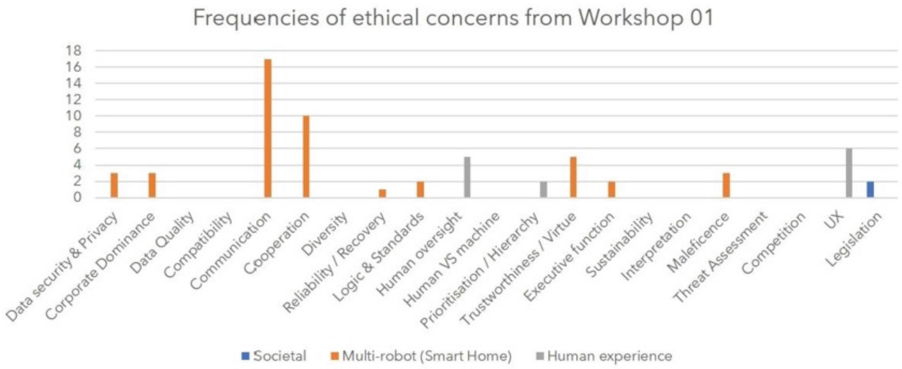

Fig. 2. Frequencies of ethical concerns expressed in Workshop 1

All themes in addition to the legislation theme are displayed in Fig. 1. Based on the percentage of frequencies, *communication* (27.9%) was by far the most mentioned theme. Attributes associated with communication included communication failure between brands and makes of robot - corporate strategy and/or mere incompatibility. Communication was additionally connected to maleficence in cases whereby robots of competing companies may deliberately offer each other misleading communication. Another concern raised in relation to communication was the potentiality for a black box scenario in which human users, via CGI, communicate on one level with the robots, yet the robots themselves communicate and operate on a different level to humans. This may lead to various aspects of data collection and sharing of data that human users are unaware of. Following communication is *cooperation* (16.4%). Both through communication as well as strategic behavior, robots may either withhold crucial information and task sharing from one another, placing obstacles in robots of competing brands' pathways (including themselves). While these tactics may seem childish, one may only look towards current and recent world leaders to understand that people (and companies) will do anything to ensure an advantage over competition. Thus, other thematic aspects can be seen as related to (*corporate dominance*, *trustworthiness/virtue*, and *maleficence*), intertwined with (*prioritization/hierarchy*, *executive function*, *legislation*, *logic & standards*), and resulting from (*UX*, *human oversight* and *data security & privacy*) ethical concerns in *communication* and *cooperation.*

W2's results follow a factor logic that connects the themes strongly to related domains or layers (see Fig. 3). Thus, issues of *diversity* (8–10%) including matters of accessibility and linguistic input preference (capabilities) were mentioned mostly in relation to the layer of human interaction. Diversity was also mentioned in reference to the sensorial hardware, other systems and behavioral hardware, and these can be understood as intertwined with the *communication* theme. While *communication* was mentioned six (7.5%)

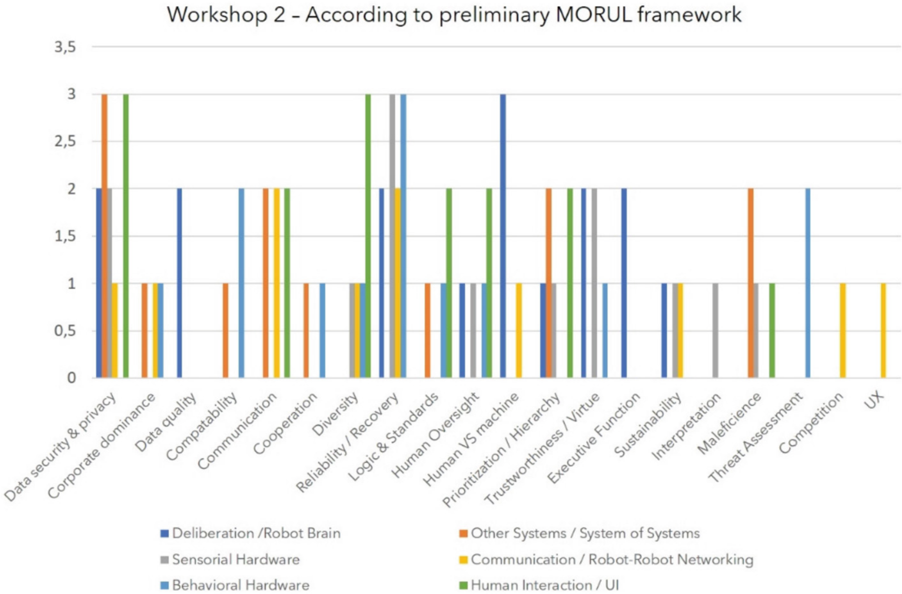

Fig. 3. Frequencies of ethical concerns expressed in Workshop 2

times in reference to other systems, robot-to-robot networking, and human interaction, other themes rose to the fore. *Interpretation* (1–1.3%) resonates with communication, and was mentioned in conjunction with the sensorial hardware. *Human versus machine* (4–5%) manifested in comments regarding the logic of deliberation/robot brain and communication/robot-robot networking. Perhaps related to the theme of *human oversight* (4–5%) and the ability of humans to keep pace of what is happening within the systems, and as such, maintain a certain level of control *human versus machine* radiates an element of techno-paranoia and the prospect of developing systems that eventually humans may not be able to control. *Logic & standards* (4–5%) were mentioned in relation to the system of systems, behavioral hardware layer, as well as the human interaction layer. These may be seen as both enablers of CGIs in multi-robot cooperation (standardizing and coordinating cooperation between and across robots, with humans), and gray areas when considering built-in logic that differs across language boundaries, and standards.

The *executive function* (2–2.5%), was noted and linked to the robot brain, which should not be surprising. Yet, in relation to this layer, there were thoughts that could be connected to the *human versus machine* theme, as well as *trustworthiness & virtue* (5–6.3%). This is considered from the perspective that the goals, and hierarchy of goals guided by the executive function could very easily be dictated by corporate objectives rather than the concerns of human users. *Maleficence* was mentioned more (4–5%) in relation to other systems, yet was also connected to the sensorial hardware and human interaction domains. This theme connected with the intention of the company or developer (for instance, the Amazon ownership of Roomba was raised often in discussion) and reasons for particular types of ownership in light of potential data collection, data sharing (sales), and 'lock-ins' (need to be locked/logged into certain systems at all times). *Sustainability* (3–3.8%) was a theme connected to the deliberation/robot brain layer, sensorial hardware, and robot-to-robot networking. Issues of programmed obsolescence and consideration for corporate responsibility in relation to the production of components, as well as recycling and disposal of non-working devices were raised.

The results led to the deliberation of a diagram that organized themes in relation to how they were represented within the workshops (see Fig. 4). The authors of the current paper acknowledge the role of culture in shaping not only society, but all the socio-technical and corporate aspects of any technological development. This said, the *cultural* domain is nestled next to the *systems and artefacts* domain due to their interwoven relationship that spans from tribal rituals and hand tools to complex AI and multi-robot systems. The *societal domain* is seen here as a holistic framework that is characterized by standards, regulations and general governance. As mentioned earlier, the researcher workshop participants were highly critical regarding the effectiveness of current regulatory frameworks (including the recently released draft of the EU AI Act, see [34] as it seems that the development is by far outpacing the speed of governance [35] over the technology in society.

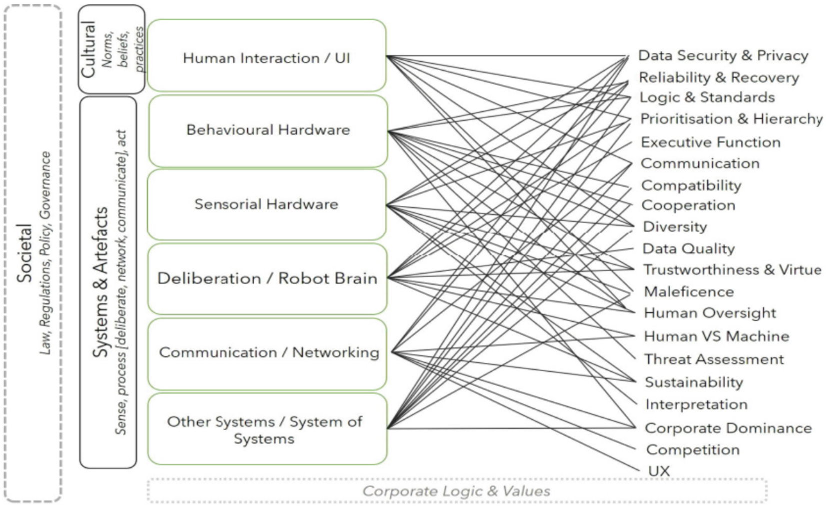

Fig. 4. Organization of domains, layers and themes

The layers are subsequently arranged from the 'top' layer of human interaction or user interface (UI) layer to the behavioral hardware - the observable action layer that both undertakes tasks and interacts with humans. Both processes and layers are interwoven and interdependent - they are SoSs. CGI was interpreted as the buffer between non-expert humans and functionality. It is not simply a UI component in itself, yet provides a substantial logic that feeds into the SoSs via provision of training data collected from users, cross-robot communication (additionally with robots or bots not directly present within the domestic setting), and above other things, has the capacity to establish affinity between human beings and robots through its seeming intelligence.

The behavioral hardware is more directly attached to the understanding of the robot unit's actions. However, as understood in the case of adding CGI, more than one unit is already present within the seemingly single-standing robot. Sensorial hardware, while embedded within the physicality of the robots, also connects with what we can understand as the 'robot brain' - the central processing unit utilized for deliberation. Once again, this lends to gray area territory due to the interconnected nature of the robots with similar, and also *other* robots. The SoS entails the complex systems supporting the robots, yet additionally connects with the broader system of domains (societal, artifactual, and corporate). Figure 4 sheds light on the thematic findings of the workshops in respect of the layers they predominantly attach with.

5 Discussion

The integration of CGI-embedded Multi-Robot Systems (MRSs) into domestic environments raises several ethical concerns that businesses need to address. Historically, the development of CGI-embedded MRSs has been primarily oriented toward industrial and business applications, with limited consideration given to the ethical implications and design choices throughout the production process [10, 22]. These systems have been created to automate various tasks and enhance efficiency across industries like manufacturing, healthcare, and customer service. Consequently, ethical considerations related to CGI-embedded MRSs have often been sidelined. Businesses involved in the development or deployment of CGI-embedded MRSs must diligently evaluate a spectrum of ethical concerns, spanning safety, security, liability, accountability, societal impact, and the implications for their own operations.

While the field of human-computer interaction emphasizes the importance of considering all aspects and stakeholders from the outset, this research underscores that not all ethical issues can be fully accounted for during the conceptualization phase. For instance, the ethical dilemmas associated with social media platforms became apparent only after widespread adoption. CGI-embedded MRSs follow a similar trajectory, where ethical concerns may not become fully evident until they are widely deployed. It is conceivable that these systems could be exploited for spreading misinformation, propaganda, or discriminatory practices against specific groups. In navigating the realm of the unknown, prudent business strategy entails anticipating the chronological stages and various components, domains, and potential impacts where ethical issues may surface, or should, at the very least, be evaluated.

For example, if concerns revolve around bias resulting from Large Language Model (LLM) training data, a multi-pronged approach involving the adoption of multiple LLMs within the systems can be considered. In cases where machine learning (ML) processes in the backend of the robots are expected to occur rapidly, incorporating checkpoints, communication protocols, and designated "pit-stops" (pauses in system operation) becomes essential. These mechanisms enable both general users and experts to observe and comprehend the actions taking place within the learned data, thereby ensuring transparency and human oversight. There are numerous other actionable strategies and operations that both businesses and developers can proactively anticipate for intervention and management, such as data offloading.

5.1 Limitations

The current study presents a number of limitations. Firstly, the empirical study presents a conceptual scenario-based investigation of CGI-empowered MRSs in the home. There was a limited number of participants, and the expert sample could have been strengthened with more research from the disciplines of law, software engineering and robotics, as well as psychology. Future steps would entail including experts from these disciplines, in addition to delving more specifically into the traits and problematics that CGI pose for MRSs – deep fakes and anthropomorphism are two areas that challenge the ethical use of CGI by its very nature. May people see Britney Spears or their *favorite* neighbor sweeping their floors any time soon? Where are the boundaries and/gray areas of privacy and intellectual property concerns when personalizing personal consumer CGI-empowered MRSs? Other limitations include the fact that this study to date has almost strictly focused on front-end issues, ignoring the back-end realm in which matters such as accuracy can severely impinge on the operations of the systems. In turn, the corporate influence and affects multiple LLMs defining the logic of the systems need to be critically examined.

6 Conclusion

As for long-term strategy, social responsibility and corporate reputation, businesses should develop clear policies and procedures that preempt and avoid foreseeable issues already at the strategy phase of innovation. This includes instilling transparency and clarity regarding privacy policies and practices, as CGI-empowered MRSs are constantly collecting, utilizing and disclosing data. By addressing these ethical concerns, businesses further ensure that CGI-embedded MRSs are used in responsible and ethical ways, potentially preventing incidents that cost business and society millions if not billions in damages. Indeed, ethical coverage of CGI-empowered MRSs may be worth billions in added-value.

It is important to start considering the ethical implications of CGI-embedded MRSs now, before they are widely deployed. This will help ensure that these systems are used in a responsible and ethical manner. Steps must be taken to mitigate ethical issues. Yet, the timing and level upon which mitigation takes place varies according to the nature of the concern itself, its cause, and how it manifests within the systems. Ethics permeates the entire hardware and software development process from design to operations. It is far cheaper to make changes during design and far more expensive, and maybe even nigh impossible, to fix ethical issues in production. While issues like bias can be may be tackled with model re-training that can be done even after deployment, if the goal or purpose of the system itself is the problem (e.g., social credit scoring with facial recognition on the streets), it may be very hard to tackle – due to its short-term business value (i.e., attractiveness for places and business such as airports).

In terms of practical implications, the issues already identified within this paper may form the platform upon which organizations may be guided. In particular, the MORUL framework for ethical multi-robot cooperation has its basis in the dual process presented in the workshop scenario method reported here. The authors would also like to emphasize two fundamental challenges that AI ethics per se, repeated face: 1) a lack of consensus regarding what AI and AI-robot ethics *is* – requiring a framework to generate broad shared understanding among communities; and 2) *how* to engage in AI, and AI-robot ethics – how can attributes such as fairness, transparency, and privacy etc. be instilled in data-driven systems? Once more, a framework is needed. Future papers will document the progress of MORUL, and will present its application with working demos and prototypes. At this time, we may consider MORUL as a *call to action* to gear business up for considering ethical issues from the outset, as a part of best practice, and as an *essential* salespoint.

Acknowledgements. We acknowledge the support of: the Research Council of Finland's funding, Emotional Experience of Privacy and Ethics in Everyday Pervasive Systems [BUGGED] project (decision number 348391); School of Communication and Marketing, and Digital Economy, University of Vaasa; Faculty of Information Technology, University of Jyväkylä; Creative Computing Institute, University of Arts, London; Department of Computer Science, University of Helsinki; GPT Lab, Faculty of Information Technology and Communication, University of Tampere; and the AI Forum project funded by the Finnish Ministry of Education and Culture.

References

1. Dwivedi, Y.K., et al.: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. Int. J. Inf. Manage. **71**, 102642 (2023). https://doi.org/10.1016/j.ijinfomgt.2023.102642
2. Stige, Å., Zamani, E.D., Mikalef, P., Zhu, Y.: Artificial intelligence (AI) for user experience (UX) design: a systematic literature review and future research agenda. Inf. Technol. People. **ahead-of-print** (2023). https://doi.org/10.1108/ITP-07-2022-0519
3. Fortune Business Insights: The global smart home market size was valued at $80.21 billion in 2022 & is projected to grow from $93.98 billion in 2023 to $338.28 billion by 2030. https://www.fortunebusinessinsights.com/industry-reports/smart-home-market-101900. Accessed 10 Sept 2023
4. McClain, L.C.: Inviolability and privacy: the castle, the sanctuary, and the body. Yale JL Hum. **7**, 195 (1995)
5. Lutz, C., Newlands, G.: Privacy and smart speakers: a multi-dimensional approach. Inf. Soc. **37**(3), 147–162 (2021). https://doi.org/10.1080/01972243.2021.1897914
6. Yao, Y., Basdeo, J.R., Mcdonough, O.R., Wang, Y.: Privacy perceptions and designs of bystanders in smart homes. In: Proceedings of the ACM on Human-Computer Interaction, vol. 3(CSCW), pp. 1–24 (2019). https://doi.org/10.1145/3359161
7. Atlam, H.F., Wills, G.B.: IoT security, privacy, safety and ethics. In: Farsi, M., Daneshkhah, A., Hosseinian-Far, A., Jahankhani, H. (eds.) Digital Twin Technologies and Smart Cities. IT, pp. 123–149. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-18732-3_8
8. Nascimento, N., Alencar, P., Cowan, D.: Self-adaptive large language model (LLM)-based multiagent systems. arXiv preprint arXiv:2307.06187 (2023)
9. Rousi, R., Piispanen, J.R., Boutellier, J.: I trust you dr. researcher, but not the company that handles my data –trust in the data economy. In: Proceedings of the Hawaii Conference on System Science (HICSS) (2024, Forthcoming)
10. Mandi, Z., Jain, S., Song, S.: RoCo: dialectic multi-robot collaboration with large language models. arXiv preprint arXiv:2307.04738 (2023). https://doi.org/10.48550/arXiv.2307.04738

11. Beauchamp, T.L.: The nature of applied ethics. In: Frey, R.G., Wellman, C.H. (eds.) A Companion to Applied Ethics, pp. 1–16 (2003)
12. Reynolds, S.J., Miller, J.A.: The recognition of moral issues: moral awareness, moral sensitivity and moral attentiveness. Curr. Opin. Psychol. **6**, 114–117 (2015). https://doi.org/10.1016/j.copsyc.2015.07.007
13. Vakkuri, V., Kemell, K.K., Jantunen, M., Halme, E., Abrahamsson, P.: ECCOLA—A method for implementing ethically aligned AI systems. J. Syst. Softw. **182**, 111067 (2021)
14. Lokman, A.S., Ameedeen, M.A.: Modern chatbot systems: a technical review. In: Proceedings of the Future Technologies Conference, pp. 1012–1023 (2018)
15. Radford, A., Narasimhan, K., Salimans, T., Sutskever, I.: Improving language understanding by generative pre-training. Preprint, pp. 1–12 (2018)
16. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, vol. **30** (2017)
17. Mäkitalo, N., Linkola, S., Laurinen, T., Männistö, T.: Towards novel and intentional cooperation of diverse autonomous robots: an architectural approach. In: CEUR Workshop Proceedings (2021). https://ceur-ws.org/Vol-2978/casa-paper1.pdf
18. Rousi, R., et al.: Beyond 100 ethical concerns in the development of robot-to-robot cooperation. In: 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), pp. 420–426. IEEE (2022)
19. Zhang, S., Chen, Y., Zhang, J., Jia, Y.: Real-time adaptive assembly scheduling in human-multi-robot collaboration according to human capability. In: 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 3860–3866. IEEE (2020)
20. Simões, A.C., Pinto, A., Santos, J., Pinheiro, S., Romero, D.: Designing human-robot collaboration (HRC) workspaces in industrial settings: a systematic literature review. J. Manuf. Syst. **62**, 28–43 (2022)
21. Gervasi, R., Mastrogiacomo, L., Franceschini, F.: A conceptual framework to evaluate human-robot collaboration. Int. J. Adv. Manuf. Technol. **108**, 841–865 (2020)
22. Zhou, S., Tian, L.: Would you help a sad robot? Influence of robots' emotional expressions on human-multi-robot collaboration. In: 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 1243–1250. IEEE (2020)
23. Che, Y., Okamura, A.M., Sadigh, D.: Efficient and trustworthy social navigation via explicit and implicit robot–human communication. IEEE Trans. Rob. **36**(3), 692–707 (2020)
24. Koubaa, A.: ROSGPT: next-generation human-robot interaction with ChatGPT and ROS (2023). https://doi.org/10.20944/preprints202304.0827.v3
25. Jobin, A., Ienca, M., Vayena, E.: The global landscape of AI ethics guidelines. Nat. Mach. Intell. **1**(9), 389–399 (2019)
26. Hill, K.: Wrongfully accused by an algorithm. In: Ethics of Data and Analytics, pp. 138–142. Auerbach Publications (2022)
27. AI Incident Database. https://incidentdatabase.ai/about/. Accessed 10 Sept 2023
28. Minto, R.: AI accidents are set to skyrocket this year. News Week (2023). https://www.newsweek.com/ai-accidents-set-skyrocket-this-year-1795928. Accessed 10 Sept 2023
29. Smiley, L.: The legal saga of Uber's fatal self-driving car crash is over (2023). https://www.wired.com/story/ubers-fatal-self-driving-car-crash-saga-over-operator-avoids-prison/. Accessed 10 Sept 2023
30. Hawkins, A.: Uber's fraught and deadly pursuit of self-driving cars is over (2020). https://www.theverge.com/2020/12/7/22158745/uber-selling-autonomous-vehicle-business-aurora-innovation. Accessed 10 Sept 2023
31. Payne, D.M., Raiborn, C.A.: Sustainable development: the ethics support the economics. J. Bus. Ethics **32**, 157–168 (2001)

32. Terry, G., Hayfield, N., Clarke, V., Braun, V.: Thematic analysis. In: The SAGE Handbook of Qualitative Research in Psychology, vol. 2, pp. 17–37 (2017)
33. Glaser, B.G.: Constructivist grounded theory? Historical Social Research/Historische Sozialforschung. Supplement, pp. 93–105 (2007)
34. EU Parliament. https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence. Accessed 10 Sept 2023
35. Kavanagh, C.: New tech, new threats, and new governance challenges: An opportunity to craft smarter responses? Carnegie Endowment for International Peace. https://carnegieendowment.org/2019/08/28/new-tech-new-threats-and-new-governance-challenges-opportunity-to-craft-smarter-responses-pub-79736. Accessed 09 Nov 2023