2018 IEEE Workshop on Advanced Robotics and its Social Impacts (ARSO)
Genoa, Italy, September 27-28, 2018

Machine agency in socio-technical systems: A typology of autonomous artificial agents

Setareh Zafari[1] and Sabine T. Koeszegi[1]

Abstract— ***Robots are expected to become a significant part of our lives in the future. In this conceptual paper, we discuss how (social) agency is ascribed to robots and how this form of machine agency can be conceptualized for an analysis of socio-technical systems. Further, we develop a typology of autonomous artificial agents which can be used to address their social impacts within a socio-technical system. We conclude by outlining implications for designers and advancing an agenda for future research.***

I. Introduction

Autonomous artificial agents are being assigned more social roles in today's society – with professional service and leisure robots penetrating the private spheres, and consequently, our everyday lives. This presents a need to better understand the impact of said agents on society and its social fabric. To explore the threats and challenges that are associated with their integration into society, an analysis of the domains in which they are already implemented, and the domains in which they potentially can or should be used is required.

Over the last decade, advances in the field of artificial intelligence (AI) have made certain passive entities active and agents. Automation, which was principally focused on the physical functions has now found his way to impact cognitive functions as well [1]. Social robots, self-driving cars, smart ambient home systems are all examples of cognitive computing systems that are designed to work through tasks without human intervention. These agents have embedded capabilities such as learning, adapting, reasoning and interacting naturally with humans. They facilitate the decision-making process by analyzing a variety of data, offering suggestions or performing automated actions. Nevertheless, it is generally accepted that these systems have a certain degree of autonomy and can make decisions independently [2]. For example AIBO, a social robot developed by Sony, can move and communicate autonomously with humans [3]. Taking into account the autonomy of these robots, the previously applicable forms of accountability need to be reconsidered. Indeed, developments in the field of artificial intelligence suggest an increase in the agency of machines, as we are assigning them roles and thereby control that would have been in human hands. However, the unpredictability of the actions undertaken by autonomous artificial agents while they are not fully controllable by a human leads to situations where such agency becomes an issue [4].

Scholars have long debated the impact of machine agency on the diffusion of responsibility. Boos et al. [5] argued that users can only be held accountable if they can understand, predict and influence the work process. This approach emphasizes the fit between an actor's control capabilities and the capacities of machines While attributing responsibility is the social function of being in control (i.e. sense of agency) [6], the current robots cannot be held responsible. In addition to legal accountability of the robots, machine agency is relevant for improving the users' acceptance of human-machine interactions [2], [7], building trust in this relationship [8] and analyzing the ethical implications of smart technologies [9]. Overall, for successful cooperation in such a socio-technical system, where humans are supported by technologies, a mindful consideration of the social components of human-artificial agent interaction is necessary. Currently, there is only anecdotal evidence of the social impact of robots, and the concept of machine agency has not been clearly defined [10]. All these arguments exhibit a need for a fresh look at concepts such as machine agency.

In this conceptual paper, we have two objectives. The first objective is to improve the general understanding of the notion of machine agency and critically analyze how robots affect possibilities of the agency. To this end, we build on the work of Rose and Turex [11] to develop a typology that identifies four types of autonomous artificial agents. By classifying the agency of machines according to certain dimensions, one may be able to understand the ontological difference between human and autonomous artificial agents. For the second objective, we use the proposed typology to explore social issues encountered in practice. The academic literature on computers-are-social-actor (CASA) has established that social responses to computers fall under natural reactions to social situations and thus the principles drawn from sociology and social psychology are relevant for user interface design [12]. Therefore, we seek to assess whether machine agency can be considered social affordance, by investigating the role of the autonomous artificial agents in the context of social interaction.

Throughout this paper, the terms autonomous artificial agent and robots are used interchangeably to refer to physical robots. They are developed based on the "sense-think-act paradigm" [12, p.164]. As such they are able to sense the environment through their sensors, decide how to respond through their processors and act on it through their actuator to pursue its own agenda for an extended period of time without human intervention [13]. Therefore, software robots that can autonomously perform a virtual task are outside the scope of this typology. Furthermore, our analysis focuses on socio-technical systems, a term which has been used to describe systems that contain complex interaction between human,

[1] Institute of Management Science, Vienna University of Technology, Vienna, Austria (+43-1-58801-33070, {setareh.zafari, sabine.koeszegi}@tuwien.ac.at).

machine and the present context [14]. Examples of such a system could be found in work settings where human agents interact with robots to get a given task done but also at home where people interact with smart technologies.

II. AGENCY

A. Essence of agency

The integration of machine and human networks have led to a renewed interest in the notion of agency due to the advances in cognitive computing. Disagreements persist about the domain of agency and adding the prefix “machine” further exacerbates this definitional debate. The Merriam-Webster Dictionary defines agency as ‘the capacity, condition, or state of acting or of exerting power; a person or thing through which power is exerted or an end is achieved’ [15]. While previous research has addressed several properties of an agency such as self-organizing, self-re-activeness, self-reflectiveness [16], forethought and self-consciousness [9], there is no single attribute that can make a distinction between agents and non-agents. A possible explanation for this might be that the concept of agency is rooted in various disciplines along with social sciences. In economics, agency is referred to as a common mode of social interaction. The theory of agency [17] describes an agent as the party that is authorized to perform an action under the control of another party, namely the principal. In sociology, agency is used as opposed to “structure”, and an agent is defined as an entity that acts independently[18] [19]. In philosophy, goal-orientation is the main domain in studying agency. For instance, Aristotle’s theory of change defines agency as the precondition of change. Change is a directed process when an agent acts on a patient. According to the standard theory of action [20], an action is intentional when it is caused by certain mental states. The event-causal relationship between intentional action and mental state of an agent is the most widely accepted view in the contemporary philosophy of mind and action [20]. However, explaining agency in terms of desire, intention and beliefs, or in general as mental states, does not imply that the possession of a mental state is identical to having agency. What is crucial for human agency is the capacity to process the contents of these mental states in a way that not only the intentional action is caused by the right mental states, but also the agent is able to justify it [20]. Moreover, the causal effect of mental states needs to be interpreted with caution, since an action is not necessarily caused by conscious intention but, according to Wegner’s model of apparent mental causation, the intention leads the action[20]. Intentionality drives and directs our actions and accordingly is an important property of an agency.

It is now well established from a variety of studies that agency of machines differs from human agency. Robots neither possess certain traits such as trust, altruism and irrationality [8] nor are they yet able to set their own goal and reason about consequences of their actions [21]. As a result, the concurrence of action and self-conscious intention starts to disappear, and autonomous artificial agents still rely strongly on human interaction and their societal intentions.

Many theoretical concepts have been developed to understand aspects of what we call machine agency. While some researchers agree that robots can also have some sort of agency, the interpretation of the machine agency varies. Current literature on socio-technical systems particularly focuses on the interrelated development of social and technical aspects of systems such as our organizations [22] [23] . Using such technological determinist and social constructivist approach, researchers have been able to provide important insights into the complex interactions between humans and technology. For example, the sociomateriality theory[24] focuses on the inseparability of the social and the material features and views agency as the capacity for action that is distributed through socio-material practices. While Gidden’s structuration theory considers agency as a property that belongs only to humans, Latour’s actor-network theory, however, considers it as a collective capacity for actions by humans or machines and makes no difference between entities in a socio-technical system [25]. This similarity between human and machine actors is not attributed to the character of the agency but their contribution to the agency of a network. Whatmore also describes agency as ‘a relational achievement, involving the creative presence of organic beings, technological devices and discursive codes’[26]. The double dance of agency model of Rose and Jones also focuses on the relation and interaction between human and machine subsequently indicating that human agency outlines and reacts to machine agency [25]. The intended action, whether performed by autonomous artificial agents or human agents, requires the interaction between a technical function and intentional behavior. In brief, machine agency is not an external pre-determined capacity that can be embodied in a technological system; it is rather an emerging phenomenon in the socio-technical systems where the human and machine are interacting in accordance to their own roles [27].

Furthermore, empirical evidence suggests that people attribute agency to non-human entities[7], [28]. This highlights the role of our perception about the agency of an entity. A seminal study in this area is the work of Rose and Turex [11] which relates perceived agency of machines to the human tendency towards anthropomorphism and describes machine agency as the extent to which machines are perceived by humans as having autonomy. In view of these arguments, it seems appropriate to analyze the problem in terms of when and how (social) agency is ascribed to machines.

People interacting with artefacts such as Wilson (Chuck’s imaginary friend in the movie Castaway) or Buddy (companion robot developed by Blue Frog robotics) as they were actual human is nothing new. Perhaps this relates to our tendency to apply what we already know to make sense of the situation. For instance, it is argued that the ideal interface for a human is a human [29]. Anthropomorphizing robots aims to enrich the quality of human-robot interaction. Indeed, we have the tendencies towards attributing human-like characteristics (such as agency) to anything that is imprecisely lifelike. Previous studies have found different features that are related to our perception of machine agency, such as adaptability [13], purposeful-looking movement [30], complementary personalities [3], and humanlike appearance [2]. Overall, autonomy in actions of machines seems to be the important anthropomorphic cue for attributing agency to them.

B. Autonomy

Autonomy is a combination of two Greek terms, auto (self) and nomos (governance) and is expressed in two dimensions of self-directedness and self-sufficiency [31]. While former describes the agent's capability to take care of itself and creates its own agenda, the latter describes the extent an agent is interdependent on external control. Taking these concepts into account, we can conceptualize what an autonomous agent is. However, the concept of autonomy is often confused by free will and free act that are strongly associated but not necessarily the same. According to Frankfurt [32], a person acts freely when the desire on which she acts is the desire she wants to be action-causing. In other words, she is not acting freely when her action is caused by none of her own desires. This applies to the situations where people claim that they were forced to act in a certain manner. For instance, even if an action is not directly based on the desire to perform that action but corresponds to a desire to have a certain desire (second-order desire), they have acted freely. Frankfurt also suggests that in addition to freedom of action, one can also have freedom of will. A person has free will when she is able to make any of her desires effective. When, due to external forces; for instance, a neurophysiologic implant, the effective desire has been manipulated, one has no free will. For example, even if a person acts according to the advice of a robot, as long as this suggestion corresponds to the desire that one wants to have, one has free will [33]. Therefore, we can conclude that the notion of autonomy rests on both free will and free action.

In the field of AI, the dominant understanding of autonomy is behavioural, i.e. refers to the extent to which robots can perform tasks independently [34] [35]. Bekey defines autonomy in robots as 'the capacity to operate in the real-world environment without any form of external control, once the machine is activated and at least in some areas of operation, for extended periods of time'[36, p. 18]. Therefore, their capability to operate or make decisions autonomously is relative to a set of pre-defined tasks and to the situation at hand. To assess the perceived autonomy of a system, the notion of "level of autonomy" has been used [34]. For instance, self-driving capacity of a car is measured on a 0-5 scale of autonomy in which 0 = All major systems are controlled by humans, 1= Certain systems, such as cruise control or automatic braking, may be controlled by the car, one at a time, 2= The car offers at least two simultaneous automated functions, like acceleration and steering, but requires humans for safe operation, 3= The car can manage all safety-critical functions under certain conditions, but the driver is expected to take over when alerted, 4= The car is fully-autonomous in some driving scenarios, though not all, and 5= The car is completely capable of self-driving in every situation[37]. The different degrees of autonomy is, therefore, a matter of technological progress. Conceptualization and design of a machine and are ascribed to it. In other words, autonomy is a subjective notion. In case of an autonomous artificial agent, autonomy is *granted* to the agent, guided through pre-defined goals, set by a principal and restricted through certain boundary conditions. Within these boundaries, autonomous artificial agents have some level of autonomy in their operation to achieve the goal that cannot be transferred onto other purposes or contexts.

In summary, these arguments support the idea that, within a socio-technical system, humans attribute agency to robots as a delegated authority. The human may distribute certain tasks to a robot and may grant it a certain degree of autonomy to make certain decisions, thus, within certain boundaries (i.e. the specific context and purpose the autonomous artificial agent was designed for which is guided through the goal function). However, this delegation of decisions to autonomous artificial agents does not exempt the human from responsibility for the consequences of the delegated decisions, as it would be the case when the decisions were *distributed* among the two agents. Clearly, there exists a hierarchical relationship between the human and the machine, in which the final control and responsibility always remain with the human; be it designers, operators, manufacturers, or owners. By integrating these reviews, we suggest the following definition:

Machine agency is the capacity attributed to machines to evoke changes within a socio-technical system by autonomously carrying out a task to reach a certain goal.

III. Typology

Based on the proposed definition of machine agency, we can develop a typology in which different types of autonomous artificial agents can function as a theory [38]. In this paper, we conceptualize autonomous artificial agents in four different types: Non-AI marginally autonomous agents, AI marginally autonomous agent, AI semi-autonomous agent, and AI pseudo-autonomous agent. These types attempt to capture the potential antecedents, processes and consequences among the different types of autonomous artificial agents using several dimensions. Though these agents are technology-based systems, major differences exist among them in terms of how they control the input-output cycle, pursue the goal and the extent to which they can perform tasks independently. Table I summarizes the main differences among four autonomous artificial agents.

A. Non-AI marginally autonomous agents

Non-AI marginally autonomous agents are programmed to carry out repetitive series of movements to achieve the user's goal until they are deactivated. Hence, the human agents have the full control over their execution. The autonomous capability rests on their ability to perform the required task without any input after being programmed. The tasks are not complex, and therefore these agents are not equipped with capabilities such as learning, reasoning or interacting with human agents. For instance, the simple Cobots [36] is programmed to carry an object and place it somewhere. Their response selection is, therefore, context-specific, a fixed repetitive series of movements in a predefined manner. For example, a thermostat can neither turn the lights down nor forecast the temperature as it doesn't fall under its typical use scenario.

B. AI marginally autonomous agents

AI marginally autonomous agents react to its environment and run without continuous direct supervision to perform more complex tasks. It relies on AI algorithms or rules to solve a problem such as identifying a pattern in a large data

TABLE I. TYPOLOGY OF AUTONOMOUS ARTIFICIAL AGENTS

Type	*Non-AI marginally autonomous agents*	*AI marginally autonomous agents*	*AI semi-autonomous agents*	*AI pseudo-autonomous agents*
Goal setting	No AI is required to achieve the goal	AI is required to achieve the goal	AI can adapt the goal within certain limits	AI can fundamentally change the goal, if necessary
Control structure	Rule-based, non-AI programming	AI algorithms	Heuristic algorithms, Fuzzy	Virtual neural networks, autonomous control
Autonomous capability	Auto-moving	Auto-analysing	Auto-learning	Auto-organizing
Range of decision making by agent	-	No uncertainty with human override	Low uncertainty with human override	Uncertain with human override
The extent of Human control	Human full control	Human pre-defined limits	Human pre-defined limits	Human guidance
Response selection	Context specific	Context specific	Context specific / Variable	Variable
Accountability	Human	Human	Human	Human(?)
Example	Classic thermostat “Cobot”	Weather station “Wizard-chess”	Current Self-driving cars “Gita”	“Survivor”

set or playing games such as chess and maze[39]. They have autonomous considerations built into their design to analyze the context and modify their action within pre-defined limits set by humans. Their performances do not expand to another context without human intervention, as they are not able to re-encode into another representation. However, they can control their action so as to maintain their effectiveness according to the rule or regulation that are set [40]. The type of decision made by the agents contains no uncertainty, as they are used in a known environment.

C. AI semi-autonomous agents

AI semi-autonomous agents are designed based on heuristic algorithms or fuzzy control technique which provides a mechanism for incorporating human-like reasoning capabilities and computationally in control systems [41]. Therefore, they can adapt the goal setting within certain limits. They can direct the user to act in a certain way by preventing them from doing certain acts or performing the act on their behalf. The type of decisions made by semi-autonomous agents can contain a low level of uncertainty, as they are equipped with context awareness and learning capability to improve from their experience and user feedback. For example, Gita robots [42] learn how to navigate complex spaces by trailing the user. Their decision making occurs in pre-delegated modes, in which the human agent authorizes the automated action to be taken based upon certain prescribed conditions and criteria [43]. They can identify, process relevant information about a situation, and make sensitive determinations about what should be done.

D. AI pseudo-autonomous agents

AI pseudo-autonomous agents are able to organize themselves. This requires the ability to self-construct the means and laws of operation in order to imply satisfactory performance under unexperienced situations and make up for system breakdown without any need for external control [44]. Therefore, they should be able to utilize their know-how to a different context and adapt their goal setting. Reflective capabilities on goal setting require integration of autonomous artificial agents into the system of human agents, combined with developing AI capabilities that aim for a cyber-subconscious level of decision making on behalf of the human agent [42]. They are designed based on autonomous control; a decentralized decision-making process which assumes interacting elements in non-deterministic systems and that are able to interact with other objects, process information, evaluate alternatives and thereby make decisions on their own. [35]. Examples of this type of autonomous artificial agent are in their research and development phase (see [45] for more details and examples). Survivor [46], for instance, is an autonomous UAV that has to some extent self-assembly feature in which it can heal itself.

IV. DISCUSSION AND OUTLOOK

Attributing agency to machines caused a vast amount of discussion about the social consequences of the application of autonomous artificial agents in socio-technical systems. Evidence suggests that robots are expanding their roles in the society and due to this social penetration, new ways of interactions between human and robots are taking place. From a social psychological perspective, the subjective willingness of humans to accept this interaction is more relevant than the objective capabilities of the robots [47]. This typology may have implications for theories addressing the outcomes of the human interactions with autonomous artificial agents on human-human interaction.

Primarily, it is necessary to consider what makes interaction with an autonomous artificial agent ‘social’. In practice, interaction with robots is considered social when the human-robot interaction is the crucial constituent of that autonomous artificial agent’s behavior [9]. For example, Buddy and Jibo are made to serve as social companions as they are able to communicate with voices or gestures, have better sensors and actuator to be aware of their surroundings. Therefore, interaction with them is considered social. From a theoretical point of view, an interaction is social when the intention, emotion and action of two parties are mutually oriented with each other [48]. We interact with others according to our interpretation of the stimulus we receive from them [49]. The interpretation is a flexible social construct which depends on the context and the party involved [50]. It helps us to clarify what to expect from the other party and goes on to be the basis for our future interactions. Autonomous artificial agents are equipped with different formats, dynamics and functions, which permit and encourage different kinds of interactions. In terms of intention, autonomous artificial agents (except *AI pseudo-autonomous agents*) are not self-governed, but the goal or purpose is programmed and defined for them. Further, they do not have internal emotional states but are able to use synthetic emotions to communicate with people [36]. Even when a robot says that it is happy for you, it does not necessarily mean that it feels happiness, but only that it is using an emotional language to interact with humans. According to Searle [51], the current type of artificial intelligence that we see in robots is "weak AI" as it only simulates the ability to understand. They are interactive and able to identify additional information to adjust the meaning and provide a proper reaction [52]. To conclude, the perception that is interacting with an autonomous artificial agent is a social interaction increase from *non-AI marginally autonomous agents* to *AI pseudo-autonomous agents* as they take human agents (i.e. their intention, emotion and action) more into consideration and fit their reactions accordingly.

This raises the concerns of how to design effective human-robot teams where autonomous artificial agents are not tools but teammates with different capabilities and dependencies to accomplish a joint work. Following the emergence of robots in work settings, it should be evident that this process is not only changing the work we do but also who does the work. The concept of management and leadership in organizations is based on the idea that in the cases of novel situations or situations characterized by uncertainty and, in case of conflicting goals or interests among agents, human decision makers give directions and take over responsibility for the consequences of their decisions. Therefore, in the design of collaborative socio-technical systems accountability has to remain in the hand of the human [26]. However, if *AI pseudo-autonomous agents* are able to fundamentally change the goal, holding a human agent responsible is far from fair. If teamwork requires human agents to be able to observe, understand and predict the state and action of the autonomous artificial agents [31], to explore the setting within which *AI semi-autonomous or AI pseudo-autonomous agents* adapt or change the goal functions and how human agents can accurately extract these modifications can enrich theory building on organizational study. To sum up, technology, work and organization are constitutively entangled and cannot be treated as separable analytical constructs [23], and greater attention has to be given to social dimension of the collaboration between agents.

Finally, it is necessary to examine potential changes in the social fabric of society. Changes in roles and forms of interaction could pose fundamental problems. For instance, interacting with *non-AI or AI marginally autonomous agents* who are always obedient might develop an unsociable personality. Or when human agents who are constantly exposed to *semi-autonomous agents* which only fake social capabilities might lose their own social capabilities. This requires consideration whether we shall apply the same rules and mental models that we learn from childhood on or we require a different "script" for human-robot interaction.

This paper outlines several criteria by which different types of autonomous artificial agent could be conceptualized. Given that, the concept of autonomy is key to our definition of machine agency; future researchers should clarify other dimensions. The use of this typology and the social impact of autonomous artificial agents provide valuable insights not only for those who design the robots but also the ones who rely upon them. Though we have discussed several key social issues that are associated with their integration into society, further research on engendering intelligence and autonomy into machines are essential. Even if developments in the field of AI are mostly driven by forces of global economic and political markets, we need to deal with these forces consciously to influence the route of technological evolution. For instance, a number of experiments (e.g.[53][54]) are using gender cues as a factor to affect the social perception of robots. But why can't we accept that robots are neutral and by that try to wipe out the stereotypes? Developments in AI require proceeding based on critical engagement and interpretation of their impact on the society.

REFERENCES

[1] E. Hollnagel, "Cognitive functions and automation: principles of human-centred automation," *Adv. Hum. Factors/Ergonomics*, vol. 20, no. C, pp. 971–976, 1995.

[2] J.-G. Lee, K. J. Kim, S. Lee, and D.-H. Shin, "Can Autonomous Vehicles Be Safe and Trustworthy? Effects of Appearance and Autonomy of Unmanned Driving Systems," *Int. J. Hum. Comput. Interact.*, vol. 31, no. 10, pp. 682–691, 2015.

[3] K. M. Lee, W. Peng, S. A. Jin, and C. Yan, "Can robots manifest personality?: An empirical test of personality recognition, social responses, and social presence in human-robot interaction," *J. Commun.*, vol. 56, no. 4, pp. 754–772, 2006.

[4] R. Weber, Â. G. Pereira, F. Dechesne, J. Timmermans, R. K. Van, and H. Vom Lehn, "Fact sheet - Ethics Subgroup IoT - Version 4 . 0," *Delft Univ. Technol. Chair Ethics Subgr. IoT Expert Gr.*, pp. 1–21, 2013.

[5] D. Boos, H. Guenter, G. Grote, and K. Kinder, "Controllable accountabilities: The Internet of Things and its challenges for organisations," *Behav. Inf. Technol.*, vol. 32, no. 5, pp. 449–467, 2013.

[6] C. D. Frith, "Action, agency and responsibility," *Neuropsychologia*, vol. 55, no. 1, pp. 137–142, 2014.

[7] K. J. Kim, "Interacting Socially with the Internet of Things (IoT): Effects of Source Attribution and Specialization in Human-IoT Interaction," *J. Comput. Commun.*, vol. 21, pp. 420–435, 2016.

[8] V. Engen, J. Brian Pickering, and P. Walland, "Machine agency in human-machine networks; impacts and trust implications," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9733, no. July, pp. 96–106, 2016.

[9] P. Lin, K. Abney, and G. A. Bekey, *Robot ethics: the ethical and social implications of robotics*. 2012.

[10] H. Jia, M. Wu, E. Jung, A. Shapiro, and S. S. Sundar, "Balancing human agency and object agency: an end-user interview study of the internet of things," *Proc. 2012 ACM Conf. Ubiquitous Comput. - UbiComp '12*, pp. 1185–1188, 2012.

[11] J. Rose and D. Turex, "Machine agency as perceived autonomy: an action perspective," in *Org. and Social Perspectives on Info. Tech., R. Baskerville et al. (eds.), New york: Springer Science+Business Media*, pp. 371–388, 2000.

[12] C. Nass, J. Steuer, and E. R. Tauber, "Computers are social actors," *Conf. companion Hum. factors Comput. Syst. - CHI '94*, pp. 204, 1994.

[13] S. Franklin and A. Graesser, "Is It an Agent , or Just a Program?: A Taxonomy for Autonomous Agents" *Int. Workshop on Agent*

Theories, Architectures, and Languages, pp. 21-35,1996.

[14] G. Baxter and I. Sommerville, “Socio-technical system: From design methods to systems engineering,” *Interact. with Comput.*, vol. 23, no. 1, pp. 4–17, 2011.

[15] Merriam-Webster, “Agency” Internet: www.merriam-webster.com/dictionary/agency, Feb.27, 2018 [March. 07. 2018].

[16] Bandura, “Social cognitive theory: Agentic perspective,” *Annual review of psychology, 52(1),* pp. 1-26. 2001.

[17] S. A. Ross, “The Economic Theory of Agency : The Principal ’s Problem,” *Am. Econ. Assoc.*, vol. 63, no. 2, pp. 134–139, 2016.

[18] M. S. Archer, “Structure, agency and the internal conversation,” *Struct. Agency Intern. Conversat.*, no. June, pp. 1–372, 2014.

[19] R. Holton, “The Act of choice,” *Philosophers’ imprint,* vol. 6, no. 3, pp.1-15, 2006.

[20] M. Schlosser, “Agency”, *Stanford Encyclopedia of Philosophy,* 2015.

[21] W. Rammert, “Distributed Agency and Advanced Technology", *Agency Without Actors? New Approaches to Collective Action*, pp. 98-112, 2011.

[22] D. Shin, “A socio-technical framework for Internet-of-Things design: A human-centered design for the Internet of Things,” *Telematics. Informatics*, vol. 31, no. 4, pp. 519–531, 2014.

[23] W. J. Orlikowski, “The sociomateriality of organizational life: considering technology in management research,” *Cambridge J. Econ.*, vol. 34, no. November 2009, pp. 125–141, 2009.

[24] P. M. Leonardi, “Materiality, sociomateriality, and socio-technical systems: What do these terms mean? How are they different? Do we need them. in Materiality and organizing: Social interaction in a technological world, P. M. Leonardi, B. A. Nardi, and J. Kallinikos, Ed.,” *Oxford university press*, 2012.

[25] J. Rose and M. Jones, “The Double Dance of Agency : A Socio - Theoretic Account of How Machines and Humans Interact,” *An Int. J. Commun. , Inf. Technol. Work*, vol. 1, no. 1, pp. 19–37, 2005.

[26] C. Knappett and L. Malafouris, *Material agency: towards a non-anthropocentric approach*. 2008.

[27] W. J. Orlikowski, “Using Technology and Constituting Structures: A Practice Lens for Studying Technology in Organizations,” *Organ. Sci.*, vol. 11, no. 4, pp. 404–428, 2000.

[28] D. Ullman, I. Leite, J. Phillips, J. Kim-Cohen and B. Scassellati, “Smart Human , Smarter Robot : How Cheating Affects Perceptions of Social Agency,” *Proc. of the Annual Meeting of the Cognitive Science Society,* vol. 36, no. 36, pp. 2996–3001, 2013.

[29] H. Ishiguro, “Interactive humanoids and androids as ideal interfaces for humans,” *Proc. 11th Int. Conf. Intell. user interfaces - IUI ’06*, p. 2, 2006.

[30] B. J. Scholl and P. D. Tremoulet, “Perceptual causality and animacy,” *Trends Cogn. Sci.*, vol. 4, no. 8, pp. 299–309, 2000.

[31] J. M. Bradshaw, R. R. Hoffman, D. D. Woods, and M. Johnson, “The seven deadly myths of ‘autonomous systems,’” *IEEE Intell. Syst.*, vol. 28, no. 3, pp. 54–61, 2013.

[32] H. G. Frankfurt, “Freedom of the Will and the Concept of a Person,” *J. Philos.*, vol. 68, no. 1, pp. 5–20, 1971.

[33] M. Rapoport, “Persuasive robotic technologies and the freedom of choice and action.” *Social Robots: Boundaries, Potential, Challenges,* pp. 219–238, 2017.

[34] M. Harbers, M. M. M. Peeters, and M. A. Neerincx, “Perceived autonomy of robots: Effects of appearance and context,” *Intell. Syst. Control Autom. Sci. Eng.*, vol. 84, pp. 19–33, 2017.

[35] D. Uckelmann, M. Isenberg, M. Teucke, and H. Halfar, “Autonomous Control and the Internet of Things : Increasing Robustness , Scalability and Agility in Logistic Networks”, *Unique radio innovation for the 21st century*, pp. 163-181, 2010.

[36] P. Lin, K. Abney, and G. A. Bekey, *Robot ethics: the ethical and social implications of robotics*. Cambridge, MA. MIT Press, 2012.

[37] Union of Concerned Scientists, “Self-Driving Cars Explained.” Internet: www.ucsusa.org/clean-vehicles/how-self-driving-cars-work#.Wyt7wlUzaHu, Feb. 21, 2018 [March. 01, 2018].

[38] P. C. Fiss, “Building better causal theories: A fuzzy set approach to typologies in organization research,” *Acad. Manag. J.*, vol. 54, no. 2, pp. 393–420, 2011.

[39] S. Sarker, “Wizard chess: An autonomous chess playing robot,” *2015 IEEE Int. WIE Conf. Electr. Comput. Eng. WIECON-ECE 2015*, pp. 475–478, 2016.

[40] T. Smithers, “Autonomy in robots and other agents,” *Brain Cogn.*, vol. 34, no. 1, pp. 88–106, 1997.

[41] S. Zein-Sabatto, A. Sekmen, and P. Koseeyaporn, “Fuzzy Behaviors for Control of Mobile Robots,” *Syst. Cybern. Informatics*, vol. 1, no. 1, pp. 68–74, 2003.

[42] Piaggio Fast Forward, “Gita.” Internet: www.piaggiofastforward.com/gita [March. 02, 2018]

[43] A. Bolton, L. Goosen, and E. Kritzinger, “Enterprise Digitization Enablement Through Unified Communication & Collaboration,” *Proc. Annu. Conf. South African Inst. Comput. Sci. Inf. Technol.*, pp. 6, 2016.

[44] P. J. Antsaklis, K. M. Passino, and S. J. Wang, “An introduction to autonomous control systems,” *IEEE Control Systems*, vol. 11, no. 4. pp. 5–13, 1991.

[45] D. Vasilescu and M. Filzmoser, “Machine Invention Systems : A Fundamental Paradigm Shift in the Field of Innovation and its Societal Implications,” *Proc. Inno. Sust. Int. Scientific Conf.,* vol. 3, pp. 6–13, 2017.

[46] Guardian News and Media Limited, “3D printer drones will take to skies by 2040, claim BAE scientists". Internet: www.theguardian.com/world/2014/jul/06/3d-printer-drones-2040-bae-futurist-aircraft, Jul. 6, 2004 [March. 1, 2018].

[47] G. Echterhoff, G. Bohner, and F. Siebler, “‘Social Robotics und Mensch-Maschine-Interaktion,” *Zeitschrift für Sozialpsychologie*, vol. 37, no. 4, pp. 219–231, 2006.

[48] R. Rummel (1976). *Understanding conflict and war*: vol. 2: the conflect helix, ch.9; Social behavior and interaction [On-line]. Available: www.hawaii.edu/powerkills/TCH.CHAP9.HTM [March 1, 2018].

[49] H. Blumer. *Symbolic Interactionism: Perspective and Method*. englewood cliffs, NJ: Prentice Hall.1969

[50] T. J. Pinch and W. E. Bijker, “The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology might Benefit Each Other,” *Soc. Stud. Sci.*, vol. 14, no. 3, pp. 399–441, 1984.

[51] J. R. Searle, “MINDS , BRAINS , AND PROGRAMS,” *Behavioral and Brain Sciences,* vol. 3, no. 3, pp. 417–457, 1980.

[52] C. Bartneck, M. Verbunt, O. Mubin, and A. Al Mahmud, “To kill a mockingbird robot,” *Proc. ACM/IEEE Int. Conf. Human-robot Interact. - HRI ’07*, p. 81, 2007.

[53] J. Otterbacher and M. Talias, “S/he’s too Warm/Agentic!,” *Proc. 2017 ACM/IEEE Int. Conf. Human-Robot Interact. - HRI ’17*, no. December, pp. 214–223, 2017.

[54] F. Eyssel and F. Hegel, “(S)he’s Got the Look: Gender Stereotyping of Robots,” *J. Appl. Soc. Psychol.*, vol. 42, no. 9, pp. 2213–2230, 2012.