https://doi.org/10.7592/EJHR2021.9.2.443

The European Journal of Humour Research 9 (2) 154–171
www.europeanjournalofhumour.org

Laughing with machines: philosophical analysis on the preconditions of sense of humour for machines

Jarno Hietalahti
University of Jyväskylä, Finland
jarno.hietalahti@jyu.fi

Abstract

This article will analyse the preconditions of sense of humour for artificial intelligence. Can artificial intelligence have a sense of humour? Is there a difference between human and machine laughter? Some machines already fulfil certain conditions which are associated with the human sense of humour: on the most superficial level machines appear to laugh and produce jokes, and they recognise sarcasm and punchlines, and they can evaluate funniness. In short, artificial intelligence is already able to recognise humour, and reacts to it accordingly. Furthermore, people laugh with humorous machines. However, it is still uncertain whether artificial intelligence can have a sense of humour or not, at least in comparison to a human sense of humour. To build bridges between AI research and philosophy of humour, this article proposes that there are (at least) five notable philosophical issues to be addressed if we are to accept that machines can have a (humanlike) sense of humour. These principles are: 1) worldview, 2) self-consciousness, 3) self-reflection, 4) self-criticism, and 5) losing control.

Keywords: artificial intelligence, incongruity theory, humour, laughter, humanity

1. Introduction

Are robots humorous?[1] Obviously, people can laugh at and with different kinds of machines, but this is not definitive evidence for a machine sense of humour. The philosophical problem is, rather, can artificial intelligence (AI) understand and appreciate humour on a self-reflective level? What does that mean and what would that require? This article answers these questions from a philosophical perspective. The actual technical analysis, that is, how to program such features, is not the main concern of this paper; this research focuses on philosophical perspectives on humour, and offers perspectives and ideas at the meta-level, not on the level of practical applications. Nevertheless, this article respects the current level of study in the field of

[1] In this paper, I will not take part in the debate considering conceptual differences between different artificially constructed intellectual agents. Briefly, the subject matter is handled here in the broadest possible sense. On a general level, we must accept a grain of vagueness; even Silicon Valley pioneers and virtual reality gurus like Jaron Lanier are uncertain about the definition of artificial intelligence (Lanier 2018: 111). For a more detailed analysis of the definitions of artificial intelligence, see Carter (2007) and Bringsjord & Govindarajulu (2020).

humour and artificial intelligence, so concrete examples from recent achievements of AI research will be discussed to strengthen the philosophical position offered. This article aims first and foremost to build bridges between AI research and philosophy of humour. This will be done by presenting philosophical ideas about human humour that should be taken into account if we are to build machines with a sense of humour.

Victor Raskin (1984; 1996; 2018) has claimed that machines do not understand humour, even if they produce it, because they only model humour, and therefore, machines do not possess a sense of humour. Raskin's main claim is that machines are unable to reflect on their own jokes and reactions to humour. This is probably true in regard to contemporary artificial intelligence, but it does not mean that there will always be a gap between humans and machines in this aspect. For instance, Yuval Noah Harari has, in his original style, predicted that computers may very well attain superhuman intelligence (Harari 2016), so it is an apparent possibility that machines can learn the secrets of humour too. However, if one is to program machines with a sense of humour, the very concept of humour needs to be unwrapped.

Similar issues have been addressed in other areas of AI research. For instance, robotics in health care is an ongoing issue, and one of the main questions is whether machines can promote 'human dignity' and, in general, have an ethical dimension. (E.g., Laitinen et al. 2019, Laitinen 2018, Anderson & Anderson 2007). Analogously, here it will be asked whether machines can have a sense of humour, and if we are to answer to this question, we must first understand what the phenomenon of humour is from a human perspective. This question is relevant for other fields of AI research too; for instance, care robots are often claimed to have positive effects for older adults (e.g., Pu et al. 2019, Khosravi & Ghapanchi 2015, Khosravi et al. 2016), and it might be argued, a fitting sense of humour is an important part of communication between the elderly and artificially intelligent companions. In short, if elderly people are to feel comfortable around social robots, it is a sensible idea to program these machines with a sense of humour too (see Mejia & Kajikawa 2017 on social interaction between robots and human beings). Nabil Hossain and colleagues (2020) go even further and claim that all AI systems that are programmed to emulate human intelligence must develop capabilities to both recognise and generate humorous content.

A working definition of humour can be based on Raskin's widely recognised conceptualisation of humour: humour is based on a contradiction between two general categorisations that are brought together in a logical manner. In humour, the apparent incongruity becomes surprisingly congruent. Jokes illustrate this position very well. For example, 'Q: What's the best thing about Switzerland? A: I don't know, but the flag is a big plus.' Sense of humour, then, refers to humour competence, that is, the ability to understand and experience these surprisingly congruent incongruities, and to be amused by them. (See Raskin 1984, especially Chapters 3, 4 and 6.) This is the core characteristic that should be applied to a machine sense of humour too.

Raskin's technical conceptualisation is a great starting point, but it can be complemented by philosophical insights about humour. In this article it will be argued that a sense of humour requires a worldview (Plessner 1970) and gets its highest manifestation in laughing at oneself (Critchley 2002, Morreall 2009, Amir 2019). Humour is not a single detachable feature but a part of the human whole. Furthermore, laughter, as a reaction to humour, has an expressive character. Laughter signals losing control; if machines are to have a sense of humour, a similar element should be programmed for AI too. But this raises hard philosophical and practical questions: what kind of sense of humour should machines have? Does a machine sense of humour need to match a human sense(s) of humour? To what extent should artificial intelligence need, for example, to be able to laugh at itself?

Implementing the feature of self-ridiculing in machines requires that artificial intelligence is able to question its own basis, i.e., it needs to be able to question its own algorithms and

databases. Programming[2] this kind of feature is a tough task in practice, but not logically impossible. If we agree that a sense of humour requires a worldview, then a machine needs to have one too. Possession of a worldview is the basis for developing a sense of humour – even if the current level of artificial intelligence is unable to do this. Rodney Brooks and Daniel Dennett (1994) have discussed the possibility that artificial intelligence could grow up on its own and learn human ways of communicating. This is a fruitful idea to be applied to the possibility of an artificial sense of humour too.

To give more flesh to the conceptual bones, 'humour' is understood as an umbrella concept which covers all the different subgenres from farce to slapstick, irony to satire, and buffoonery to witticism. Laughter, then, is a reaction to humour; here we will not discuss laughter triggered by, for instance, intoxication or tickling. Theoretically, this paper will follow the general principles of the incongruity theory; humour is based on paradoxes, or in other words, humour occurs when cultural conceptualisations are in a contradiction. Roughly, if something goes wrong (that is, our expectations do not match with what actually happens in the world), humour is always potentially present. Sense of humour in its simplest form refers to a capability to understand such humorous occurrences. Artificial intelligence is understood as a set of algorithms and a database, and its competence is based on the logical processes and the knowledge and information that it possesses. Briefly put, artificial intelligence is a combination of processes (algorithm) and knowledge (context). In the text, artificial intelligence and machines are used synonymously[3]. With regard to sources, this article uses mainly respected studies, but because the subject matter grows ever more topical, a few examples from science fiction and newspapers are included – the latter sources are not considered 'scientific' but used to show how the issue is discussed in the public sphere.

The structure of the article is as follows: First, the article examines certain general topics on sense of humour and artificial intelligence, and offers a philosophical problematisation to further the discussion by showing that human beings may not be the only laughing and humorous creatures in the universe. Second, the article analyses what kind of short-term and long-term goals there are for an artificial sense of humour; the most relevant goal in the current field is to create a machine with a sense of humour matching that of human beings. The third part of the article critically discusses the relationship between artificial and human humour by offering different kinds of philosophical goals for artificial intelligence, and by comparing joke generators and human creativity. It will be claimed that a capability to produce jokes is not a sufficient condition for a sense of humour. The fourth section offers an argument on how humour is actually a negative feature, and connected to the human situation in the world. Humour expresses the existential situation of a human being, and this position proposes a hard challenge for artificial intelligence. Finally, the article will show that even if some famous social robots that are basically publicity stunts (e.g., humanoid robots like Pepper and Sophia) appear to be humorous, witty and laughing, they do not possess a sense of humour (even if they produce the sound of laughter and create joking texts). In conclusion, it will be demonstrated that, for us to accept that artificial intelligence has a sense of humour, it must meet five tough philosophical preconditions: worldview, self-consciousness, self-reflection, self-criticism, and losing control.

[2] I will not discuss details of a suitable code. My interest in this paper lies in the philosophical principles of the general subject matter.

[3] It is clear that there are primitive machines that have a very little in common with modern artificial intelligence, but in this text, the context will reveal the appropriateness of the chosen terms.

2. Artificial fun: the questionable human monopoly on humour

Humour and laughter are said to be essential human features. At least Aristotle (*On the Parts of Animals*: Book III) claims that humans are the only creatures who laugh, and from this perspective we can be called *homo risus*, the laughing man. However, in the light of modern humour research, this position is false. There is evidence which suggests that animals – among others, rats (Panksepp & Burgdorf 2003) and dolphins (Blomqvist et al. 2005) – laugh. Of course, mere laughter does not imply that a non-human animal is capable of humour. But there might be animals who know how to joke. A gorilla called Koko learned sign language, and reportedly laughed at her own jokes and those of others (Patterson & Gordon 2002). These studies offer at least circumstantial evidence that some animals can have some level of sense of humour.

The studies on animals suggest that it is possible that the human race does not have a monopoly on humour. Even so, we have a tendency to evaluate the possibility of a sense of humour in an anthropocentric fashion. If we are to agree that gorillas potentially possess a sense of humour, they need to be funny in some way that is comprehensible to us. The same demand is implicitly present for humorous machines, too.

There is an age-old dream to create a robot who looks and behaves like a human being. Even Leonardo da Vinci tried to build a humanoid robot, and various attempts have been made ever since (Moran 2007, see also Kato 1973). Very recently, there have been a number of publicity stunts based on humorous humanoid robots, and because of them it has become easy for the layperson to believe that machines can genuinely laugh and joke. Generally speaking, it is sensible to accept that a capability for humour must be involved in one way or another in the process of creating humanlike robots. The ideal machine would be able to both recognise and appreciate human humour, as well as to respond to it with its own witty remarks. If a poetic formulation may be used, artificial intelligence has been made in one's own image, as Sherry Turkle (2011: 44, 55) holds. Therefore, if human beings are humorous and laughing creatures, artificial intelligence should be, at least to some extent, like us in this regard. The aspect of gradual similarity leads to the idea that machines' sense of humour should be (at maximum) on the same level as humans' and they should not outwit people. In this paper, we do not need to answer the technical question of how this can be done, but instead we can focus on the hard pressing philosophical problem: what does it mean if an artificial intelligence has a sense of humour?

This philosophical question has been tested in what may be called 'philosophical thought experiments in science fiction'. One of the clearest examples is an android called Data from the science fiction series *Star Trek: The Next Generation*. Data struggles with humour. He can give firm definitions of, for example, jokes but fails to understand the practical manifestations of humour in real-life situations. He does not get jokes.

If we follow the classic argument by Alan Turing, Data miserably fails the 'Turing Test'.[4] According to Turing, instead of asking whether a machine can think or not, we should ask whether a machine can be indistinguishable (in a linguistic sense) from a human. (Turing 1950.) It is clear that while communicating with people, Data fails the test of humour, so to speak. He is not funny in everyday encounters, and this troubles Data deeply.

After Data realises that he lacks something essential (that is, a sense of humour) compared to his human fellows, he tries to find out more about this peculiar form of human interaction. The main computer of the star ship shares more data with him on the subject matter, and he manages to mimic comedians' routines in an exact manner and deliver jokes to other passengers

[4] History of philosophy shows that similar ideas have been present for a long time. René Descartes claimed that if machines who look and act like human beings cannot communicate like humans, they are not real human beings (Descartes 1911, 116).

on the star ship. Unfortunately, this is not enough. People do not consider Data humorous. Or to be precise, there are a couple of moments when Data is accidentally funny in human(oid)s' eyes, but he himself does not grasp why. Humour remains a mystery to Data.

Star Trek illustrates in an insightful way a topical issue among contemporary humour researchers (see Raskin 1996, Hempelmann 2008; 2014): what does it require for a machine to possess a sense of humour? Depending on the scholar, the answer refers to the capability of producing humorous material, or being able to recognise and possibly to respond to humour, or being able to evaluate the funniness of a joke. Computational humour studies and research on artificial intelligence are ever more important, and progress is being made with increasing pace. For instance, in the International Society for Humour Studies Tallinn Conference (2018), topics considering humour and artificial intelligence included human-machine communication (Generalova 2018), generating and recognizing humour through deep learning (Hempelmann et al. 2018) and knowledge-based artificial intelligence systems and their relation to script opposition (Petrenko et al. 2018). Even so, humour can be said to be a final frontier for artificial intelligence; creating an artificial sense of humour is arguably as demanding a task as programming such emotional traits as love, hate, sadness and empathy. Jaron Lanier claims that if he knew how to define a person, he might, but he admits that he cannot. According to Lanier, it is a grave mistake to turn 'life into database' and to believe that computers can presently represent human thought, because they cannot (Lanier 2010: 9, 82). It may be argued that similar problems hold for a sense of humour too. Humour is still, as it is to Data in *Star Trek*, a mystery. However, research on the subject matter progresses all the time.

3. Humorous machines

If one wants to build humorous robots, the question 'how is humour possible?' needs to be answered. What are the fundamental elements of the phenomenon of humour? What kind of agents possess a sense of humour? Furthermore, one needs to have a goal for programming a sense of humour for artificial intelligence – what kind of sense of humour should it be? If a programmer aims to maximise the instrumental value of artificial intelligence with the help of humour, humorous robots are a sensible idea. From this perspective, they need to be, for example, easily approachable and pleasant company. If it is cheaper and more efficient to build funny robots than to hire people, for example, in the field of social care, the whole pattern is reasonable. Just as Sherry Turkle claims, technology is a seductive alternative when there is a chance that it could meet human vulnerabilities and help us; social robots can make people feel good at least for a moment. This trait is already a reality in the field of health care across the world. For instance, social robots offer company for elderly people when there is a lack of human resources. (Turkle 2011: 1, 24, see also Laitinen et al. 2019, Anderson & Anderson 2007, Pu et al. 2019.)[5]

The immediate goals of an artificial sense of humour are related to the general long-term aims of AI research. Scholars such as Stuart Russell and Peter Norvig (2009) and Maija-Riitta Ollilla (2019: 51) suggest that there are (at least) four guiding principles for possible systems of artificial intelligence. Either artificial intelligence should think like people (that is, match human rationality); or it should think along the lines of ideal rationality (that is, match the most rational way of thinking, for instance, in mathematics); or it should act like people (that is, match human actions); or it should act in the most rational way possible (that is, act flawlessly, for instance, when performing movements). When applied to humour, the four possible long-term goals for artificial intelligence are slightly altered, and the distinctions are now located between

[5] Turkle is highly critical of the tendency to replace human workers by robots, but we will leave this important discussion for other articles.

reason/communication and acting/behaviour. 'Communication' here refers to linguistic communication, whereas 'behaviour' refers to non-verbal actions. 'Current sense of humour' refers to a vague idea of a general sense of humour which involves possible failures to understand and produce humour. 'Ideal sense of humour', then, refers to a flawless sense of humour, and to a capability to produce and understand humour in the best imaginable way.

Table 1. Long-term goals for an artificial sense of humour

	Current sense of humour	Ideal sense of humour
Reason	Systems that communicate as funnily as human beings (jokes, witticism)	Systems that communicate in an even funnier manner than human beings (jokes, witticism)
Behaviour	Systems that act as funnily as human beings (slapstick, pranks)	Systems that act in an even funnier manner than human beings (slapstick, pranks)

The topic is very timely, and in addition to the long-term goals mentioned above, various AI researchers aim to tackle the problem of artificial humour for its potential application, for example, in generating humorous texts (Hossain et al. 2020), text mining and detecting online harassment (Van Hee et al. 2018), understanding the nature of funniness (Chiruzzo et al. 2019), enhancing human-computer interaction (Miller et al. 2017), and more broadly, understanding humour (Hempelmann 2008). In addition, a general artificial sense of humour would be usable by multinational corporations for various marketing reasons from analysing customer feedback and product reviews to customer service. As important as these short-term goals and practical applications are, the main interest of this paper is focused on the long-term goals and on the possibility of achieving them.

Drawing from the table above, presently, it seems, artificial intelligence should be able to communicate with human beings on the same level; or, expressed negatively, it should not outwit people. In light of our current knowledge, it seems unlikely that one could program 'an ideal sense of humour', because humour is an interpersonal form of communication (see Ruch 1998). In other words, it is hard to see how a 'superior sense of humour' could exist. One can claim the opposite too. There are various competitions in the field of humour (e.g., an international competition to find the 'funniest person in the world'), and it is a philosophically inspiring idea that machines could write funnier jokes than human beings. But for the sake of realism and keeping in mind the current status of research, let us proceed from the idea that at the moment the most pressing task is to program a matching sense of humour between artificial intelligence and human beings.

A notable portion of AI research in the subfield of computational humour has focused on detecting whether a text is humorous or not (Potash et al. 2017: 49) but there has been very recent progress that aims at evaluating different degrees of humour (Potash et al. 2017, Chiruzzo et al. 2019) by scoring the detected level funniness. For instance, AI has already achieved some success in ranking the funniness of Twitter messages (Hossain et al. 2020), interpreting puns (Miller et al. 2017), and detecting irony (Van Hee et al. 2018).

The recent progress is related to the challenge of how to build intelligent and humorous (but not too humorous) agents[6] (see Russell 1997).[7] Following this idea, machines are not just *for* human beings but more importantly *like* them (see Turkle 2011: 86). Briefly put, the field of computational humour uses computers in humour research to offer models of a sense of humour; as early as 1992, Igor Suslov formulated a computer model of sense of humour and a general algorithm for a computer sense of humour, and pondered: 'Is it possible to create a computer program which will "laugh" in the same cases as a man? From our viewpoint it is quite possible.' (Suslov 1992: 9) In addition, computer programs have produced jokes for decades. Early pioneers in the field were Kim Binsted and Graeme Ritchie (1997) who devised a formal model of simple types of punning in the form of question-answer jokes. For instance:

> *Q: What do you get when you cross breakfast food with a murderer?*
> *A: A cereal killer.*

Depending on one's taste in humour, this joke can be funny, and at a minimum, it fulfils the technical requirements for a joke. It has a setup and a punchline which turns the expectation on its head. So, at least on this modest level, programmers have been able to crack the code for jokes. However, these kinds of joke generators are not enough for people who want to write a sense of humour for AI. Victor Raskin has criticised these types of pun generators[8] for a lack of creativity (1996; 2018). In other words, these machines lack a central element when compared to human humour. Humour is arguably a creative process during which brains process something other than just ordinary communication; the rules of an ideal speech situation (see Habermas 1973) do not apply to humorous interaction. Jaron Lanier (2010: 185) has pointed out that this will be a tough requirement for artificial intelligence, because human language is hardly a code that can be easily cracked. Humour, evidently, makes things even more complicated. Even so, the task is not impossible, and in regard to artificial intelligence, this means that it must not only follow the rules of joking but understand the more general rules of communication too.

However, one should not be too eager to mystify the element of human creativity. Creativity should not be understood as an external source of outstanding ideas and surprising viewpoints. Even professional joke writers have methods to refine their creative processes.[9] If a regular Joe was asked to write 50 jokes in one day, this would sound like a task for Hercules – practically impossible. However, there are people who manage to match these demands on a regular basis. The pattern for success is not overly complicated. One way to foster humorous creativity is to take two culturally incongruent subject matters (e.g., artificial intelligence AND Neanderthals), and then make a list of things associated to each term respectively (e.g. AI – computer, future, mathematics, programming, nerd, etc. and Neanderthal – genes, extinction, primates, anatomy, past, etc.). After the lists are done, both sets of words are checked to locate the possibly incongruent combinations, and finally it is possible to formulate jokes based on these findings. Naturally, this is not to say that every joke written in this way is sensationally funny; instead, it shows how human creativity can be boosted with a relatively simple trick. (For a detailed take on humour and creativity, see Luria et al. 2018.) These kinds of basic ideas can be written into

[6] There is plenty of research which does not aim at this kind of general artificial intelligence, but the objectives are often much humbler; for instance, there are projects to build models for humour recognition (e.g. Miller, Hempelmann & Gurevych 2017).

[7] Again, the actual way of writing programs and building artificial intelligence is not the main interest here. Admittedly, just like Bringsjord, & Govindarajulu (2020) put it, subjective consciousness and creativity are high mountains facing artificial intelligence, and the field probably is not getting anywhere close in fulfilling these tasks at the moment. Arguably, both features are pivotal for possessing a sense of humour (see Ruch 2008).

[8] Interestingly, Raskin himself has written a program that produces light bulb jokes (Raskin 1996).

[9] For a detailed take on how stand-up comedians approach their craft, see Ajaye 2002.

a program, too. The task is, then, to formulate an appropriate database and a practical algorithm so that AI can operate competently in a relevant context.

Creativity, of course, is not the only element of humour. If one wants to build a machine that can be called genuinely humorous, it must be based on a plausible definition of humour. As various humour scholars agree, this demands much more than just a technical wordbook reference: human humour is associated with, among other things, imagination, spontaneity and empathy, and a solid understanding of the nature of language. (See e.g., Graham 2010, Adamle & Ludwick 2005, Raskin 2008, Ruch 1998, Critchley 2002, Amir 2014.) In general, humour is an interpersonal feature that is learned throughout life. We need to understand ourselves and others around us if we are to succeed in humorous interaction. In short, a sense of humour requires a worldview[10]. Furthermore, the greatest humourists are often said to have a unique worldview, and they express this in their comedy.

4. The art of humour: the requirements of personality and worldview

From a practical point of view, humour can be called a skill (that is, it demands practice and repetition), or a virtue (it demands education), or perhaps an art (it demands even more practice, repetition and education). If this starting point holds, humour can be taught, and presumably written into a code, too. At least it is possible for human beings to learn humour around the globe – be it how children are taught the forbidden objects of ridicule, etiquette lessons for proper use of humorous forms of communication in different social situations, or a beginner's course in stand-up comedy. If one is to become a stand-up comedian, it makes sense to record his or her routines to gather data about timing, reactions of the audience, and so forth. This helps to improve the performances. In short, there are practical possibilities to learn about humour and to improve oneself in this regard.

If we follow the idea that humour is an art, then the personal, social and ethical aspects of humour intertwine. The Greek philosopher Aristotle has given sensible guidelines for understanding the entanglement of the personal and the social in the field of humour. According to Aristotle, a sense of humour is a social virtue. This means that there is a golden middle for how to practice humour and laughter. Aristotle claims that vulgar buffoons exaggerate humour because they spare no one while chasing laughter; on the other hand, boors are dull and refuse to laugh at anything. According to Aristotle, the golden middle can be called 'tactful'. This type of humour does not, for example, offend just for the sake of offending, but offers highly important relaxation in social life in a respectful manner. Tactful people know when to tell jokes and when to keep their funny ideas to themselves. A good sense of humour, then, is morally worthy too. This kind of ready-wittedness Aristotle calls *eutrapelia*. (Aristotle, *Nicomachean Ethics*: 4.8, 1127b33–1128b5; see also Comte-Sponville 2001.) Following Aristotle, the scholastic philosopher Thomas Aquinas states that humour and laughter offer psychological distance from exhausting mental studies: in the middle of all the seriousness, we need some fun. 'Those words and deeds in which nothing is sought beyond the soul's pleasure are called playful or humorous, and it is necessary to make use of them at times for solace of soul.' (*Summa Theologiae:* 2s2ae, Q. 168, Art. 2) For Aquinas, those who are never funny, kill-joys as they are called, are actually sinful. (*Summa Theologiae:* 2a2ae, Q. 168, Art. 4). According to this Aristotelian tradition we need to be humorous in a virtuous way. This implies that sharing amusing jokes is not enough. The demand on artificial intelligence is that it needs to have a

[10] The concept of worldview is similar to Ludwig Wittgenstein's idea about 'form of life' (Wittgenstein 1986: §19, §23, §241, 226, see also Hunter 1968), which describes how people understand the surrounding world; the form of life explains how people use and relate to, for instance, imagination, play and queer changes.

moral understanding as well. If artificial intelligence is to be called human-like in this aspect, it must be funny in a proper manner.

One obvious problem is how to measure funniness. Of course, there have been competitions for world's funniest joke and world's funniest person, but these tend to focus more on popularity than scientifically objective amount and depth of funniness (see Hietalahti 2016). It is clear that tastes in humour vary from one era to another, from culture to culture, and from an individual to another (Critchley 2002). Another problem is that human beings' stamina is not endless. People need breaks from the fun. One can read ten jokes in a row and make a listing based on his or her personal taste. But if the task is to go through, say, more than a hundred jokes in a short period of time, even if the evaluation takes place on a subjective level (i.e. is based on personal taste) the objectivity of the evaluation suffers. Fun has to stop at some point. Just like Aristotle claims, a virtuous sense of humour tells people when to joke and when not. An artificial sense of humour should be sensitive to these aspects, too.

There might be an interesting solution for the previous problem: write a program that makes the comparison between all the produced funny material. If a joke bot offers an endless source of comedy as long as it is plugged into an energy source, then we need an audience that does not get tired as long as it has energy. This would cut the unreliable human middleman from the equation. If the algorithms are solid and the database coherent, we can just check on the outcome they provide, and trust their evaluation.

For some, this might sound a silly solution, but there have been various projects to program artificial intelligence systems that measure the level of funniness (e.g., Potash et al. 2017, Chiruzzo et al. 2019, Hossain et al. 2020). This does not mean that human beings are actually removed from the process because the code is written by human individuals. However, this trend in AI and humour research raises other difficult issues concerning bias in data-driven artificial intelligence systems. For instance, there are problems related to processing of data that can lead to prejudiced results if AI does not recognise and value features like sex, race and so forth. (On the bias of AI, see Ntoutsi et al. 2020, O'Neil 2016.) In 2016, it was widely reported how an AI chatbot Tay started to tweet misogynist and racist tweets. This machine learning project aimed at learning to communicate and interact with people from conversations, but the process became distorted. The bot was shut down shortly after its launch. (E.g., Vincent 2016, Reese 2016.) If an AI merely reflects the conversations around it, there is a risk that the outcome is not the noblest possible. For this reason, it is necessary to understand the connection between humour and virtuous social behaviour, like Aristotle suggests. This guiding principle explains why sense of humour is dependent in part on an individual's total personality.

To avoid bias, it is essential to examine and define crucial principles of a sense of humour. For instance, the Merriam-Webster dictionary defines this human feature as follows: 'a personality that gives someone the ability to say funny things and see the funny side of things.'[11] This definition is limited[12], but it is of some value; a sense of humour is related to personality. Willibald Ruch has deepened our understanding of the concept of sense of humour by claiming that a mere appreciation or creation of humour does not define a sense of humour properly. Instead, the personality approach promoted by Ruch claims that understanding a sense of humour must involve the various individual differences regarding how people relate to humour (Ruch 1998: 10-11). This is an important addition and it highlights the idea that human beings each have a unique sense of humour, although there are various common denominators, too (see Hietalahti 2016). It is sensible to agree that a sense of humour is a personality characteristic (Ruch 1998) and if machines are to possess a sense of humour, they also need to have a personality; and to have a personality, one needs a worldview.

[11] https://www.merriam-webster.com/dictionary/sense%20of%20humor (date of access: 8 October 2019)

[12] For example, what does it mean if something is funny? If someone says something funny but others do not consider it funny, does the joker lack a sense of humour? Who is the judge of fun?

If a sense of humour is related to a worldview, it is by definition a socially constructed feature. Human beings are social creatures, and their understanding of the world is formed in social relationships (Fromm 2003). One cannot have a private sense of humour.[13] Roughly put, a sense of humour is a unique personality trait that corresponds to others' equally unique features. Therefore, humour is not just about the wording of a joke, but it is a part of the agent's existential situation. In a philosophical sense, humour offers an answer to existential human problems[14]; for example, humour is a form of relatedness, and makes one able to feel him- or herself part of the wider social group (cf. Provine 2000). We can conclude that, as humour is a part of the individuals' total character structure, it cannot be detached from the way one is related to others (see also Hietalahti 2016). If this argument holds for humans, it must hold for artificial intelligence too.

Humour can be understood as a way to relate to the world and to others, and laughter as an expression of the human situation in the world. Helmuth Plessner (1970) argues that laughter has an expressive character, but the meaning of laughter is not fixed, that is, laughter does not always signal the same things in every single situation. The reason for this is clear: human individuals each possess a unique sense of humour, and this is highly influenced by the socio-cultural historical circumstances.

Philosophically speaking, humour is related to the core of humanity. If the very basis of humour is based on deviations from the normal, this actually means that human beings are optimally flawed for humour; humans know more than other animals, but not everything, in the way an omnipotent being would. Based on this premise, humour is actually a negative feature, related to human imperfection. For an omnipotent creature who knows everything, there can hardly be humorous surprises (Baudelaire 1956), and animals lack some of the necessary cognitive capabilities to understand cultural categorisations and conflicts between them (cf. Hazlitt 1951). A sensible assumption is that humans possess a sense of humour because they are self-conscious, but have a limited knowledge of their surroundings – we have firm conceptualisations of the world, but they can never portray world in a perfect way, and when we see that there is a deviation from our expectations, we can be amused thanks to our reflective thinking. Artificial intelligence needs to match this existential position to have a sense of humour equal to that of humans.

5. Laughter and the loss of control: why current machines do not have a sense of humour

At the moment, artificial intelligence does not possess consciousness, that is, it does not want anything; or in other words, artificial intelligence cannot choose its options independently. (Ollila 2019: 74.) In this sense, artificial intelligence cannot choose to want to learn the secrets of humour (unlike the android Data mentioned in the beginning of this text). Nevertheless, there are humanoid robots which can and do joke with people, and they laugh at people's jokes – or at least they simulate joking and laughing by repeating sound files of recorded laughter and generate joke-shaped texts. Nevertheless, this trait can be deceivingly accurate. A social robot called My Real Baby can laugh with perfect timing when someone shows it silly faces (Turkle 2011: 110).

[13] Compare with Ludwig Wittgenstein's idea about the impossibility of a private language (Wittgenstein 1986).

[14] Following Erich Fromm, the central existential problems are how to be related to others, and how to have a sense of identity. Fromm notes that these problems can be answered in a pathological or life-affirmative way (Fromm 1990; see also Hietalahti 2016).

In principle, it is possible that a humanoid robot might pass the Turing Test at some point if it communicates in an appropriate manner and uses humour accordingly. However, it can be argued that, for example, different kinds of publicity stunt robots merely imitate humour – they only perform social features. Some robots are performing sociability so well that it is easy to believe that they are sort of alive (see Turkle 2011: 26). However, if one claims that the machine is human-like (in the Cartesian sense mentioned in footnote 4) and has a sense of humour, the danger is that we are losing something essential from the human sense of humour (cf. Turkle 2011: 54; Lanier 2010: 39-40). Therefore, at least for now, it is hard to accept that artificial intelligence has a sense of humour. The burden of proof is on those who claim that it does.

Be that as it may, machines' potential sense of humour raises the old philosophical question about other minds. What can we know about the inner state of another human being? From a first person perspective, I know what is funny to me and why I laugh at certain comedies, but how can I know for sure that others laugh at the same things and feel the same kind of amusement? I know fairly surely that others have feelings and get exhilarated, but are the feelings precisely similar in me and in another person? In addition, it may well be that even if I laugh at the same time as others while watching the same comedy, our laughter is triggered by different reasons. In an epistemological sense, these are hard questions, and probably cannot be answered satisfactorily (see Vaaja 2015). So, if we cannot be sure (at least epistemologically) about other persons, how can we be sure if a social humanoid robot's 'inner life' is drastically different than ours – especially if it imitates human features accordingly?

One way to approach the problem is to make a distinction between strong and weak artificial intelligence. These are different views of AI theorists on the purpose of AI as a tool. Following John Searle (1997), strong artificial intelligence refers to the view that consciousness is nothing more than a set of algorithms. This idea has been popularised by Harari (2016) recently. Weak artificial intelligence refers to the view according to which the human mind cannot be reduced merely to algorithms (Searle 1997). For example, non-reductive naturalists argue that the mental properties emerge naturally but cannot be completely reduced to certain kinds of organisms because subjectivity is more than the sum of its parts (Pihlström 2015: 46-48). According to the definition of weak AI, artificial intelligence is not actually conscious, but merely *appears* to have similar mental capabilities to humans. Without taking any account of the actual technical analysis, with the most favourable interpretation, some of the current social humanoid robots might appear to behave in some sense like human beings. On the simplest level of humour and laughter, there are machines that show a potentiality to imitate these features when they produce a sound similar to laughter and if they produce words in the forms of a joke. (Obviously, the deep philosophical issue about other minds is still relevant). At the moment, we have a good reason to believe that other people are self-conscious (e.g., because they communicate like we do), but we do not have equally good reason to believe that social robots are self-conscious (e.g., because we can find the basic structure of how it is built, how it communicates, etc.). Therefore, if we accept the above-mentioned principle that sense of humour is a self-reflective feature, it is sensible to conclude that current AI do not possess a sense of humour. However, it is worthwhile to remember that machines actually can produce words that may be amusing to users and they are able to produce the sound of laughter. If we agree that they do not have a sense of humour, then logically, a mere capability to tell jokes and laugh (in the weakest sense of the terms) is not enough to convince us that an entity has a sense of humour.

The philosophical task at hand is, then, to map out the fundamental principles of sense of humour. These are the preconditions for sense of humour for both humans and artificial intelligence. These need to be fulfilled if we are to agree that a machine has a sense of humour (though of course, there might be some other requirements too). Whether a machine is actually able to fulfil these conditions at some point in the future is another problem for later research. Sami Pihlström (2015: 49) has pointed out (although not in the context of AI research) that

instead of focusing solely on the problems of analytic philosophy of mind (e.g., consciousness and mentality), it is more sensible to try to understand the entire human being in their social surroundings. Pihlström's guideline is very promising for understanding and mapping out the phenomenon of humour too. He offers an encouraging position by suggesting that the focus should be on how 'minded' creatures like human beings live in the world amidst their projects and habits. With this recommendation he does not claim that human beings will always be categorically the only creatures with consciousness, so it is a logical possibility that at some point AI can exist in a similar manner than humans, even if this possibility is remote at the moment. In short, it is unnecessary to deny that AI could ever possess a sense of humour, but neither do we need to accept that artificial intelligence has a sense of humour presently.

As argued above, humour is related to the existential problems of humanity, and therefore it cannot be understood solely from, say, a technical perspective – one needs the social philosophical aspect, too. An individual sense of humour can only be understood in relation to the surrounding social reality. Human beings are peculiar, flawed creatures with their existential problems, and following Helmuth Plessner, mapping out the preconditions of humour requires a philosophical-anthropological theory of human condition that is sensitive to these hard pressing issues. Evidently, current machines do not base their existence on similar existential problems as human beings.

One obvious element is that human beings always operate in a living context, and this holds for humour, too. Humour and laughter are expressive phenomena, and they can be understood only if one takes into account the historical, concrete situation in which they appear. (See Plessner 1970: 14-16, 18.) For artificial intelligence, this means that it needs to understand its surroundings in a reflective manner. This is one of the starting points for possessing sense of humour. If humour is, as argued above, based on culturally conditioned paradoxes, then having a sense of humour means that the agent must perceive contradictions in its life in a meaningful manner. Furthermore, laughing at incongruities signals that one has found something amusing in the perceived paradox. But this is not just an intellectual process but a bodily experience, too. Plessner interprets laughter as involving an agent losing control:

> The occasion for laughter overtakes and overpowers us…This lapsing and falling into…laughter reveals…a loss of control, a breakdown of the equilibrium between man and his physical existence. A sudden and powerful outburst of feeling can drive us to unthinking expressive movements; we are then no longer masters of ourselves and no longer have our wits about us…[T]he control of [the human being's] intellectual and moral center, is endangered. [15]
>
> (Plessner 1970: 65-66)

Here we have an important aspect of humour; it challenges reason but, on the broader human level, is very reasonable after all. Plessner notes that laughter is on an apparent level meaningless (Plessner 1970: 68), and that we do not, for example, articulate precise claims or make strict arguments through laughter. However, from a philosophical-anthropological perspective, this very meaninglessness is the meaning of laughter. Laughter disrupts reason, or at least what might be called rigid reason (see Hietalahti 2017). Erasmus of Rotterdam had already noted the same tendency in the 16th century. According to him, on the very basic existential level, human beings need laughter and merry non-sense. That is, humanity cannot be understood without understanding its foolishness, but this foolishness is very rational, in the end (Erasmus 1941). For example, through humour, it is possible to gain new insights and perspectives on every single human conceptualisation. The very basis of incongruity theory claims that everything can be shown in a humorous light (see Kinnunen 1994). At the same time, humour reveals the

[15] Both Plato (*Republic*: 388e) and Epictetus (*Enchiridion*: §33) warned their followers about the dangers of laughter: it disturbs reason and the calmness of the soul.

limitedness of our conceptualisations and at its best, demands new and better formulations and a deeper understanding of the world. This is why humour is so often said to make the familiar look so unfamiliar, and the unfamiliar very familiar (e.g., Critchley 2002).

Humour, then, is a living and dynamic feature of human life which shows that our conceptualisations are always limited. We are not all-knowing, and therefore there is always room for humorous surprises in everything we do. This opens the way for laughing at ourselves, too. In this laughter, we notice the very limitedness of our human lives, and we agree with it: we understand we are not as great as we might think but wretched and silly creatures (Critchley 2002). Laughing at oneself is not just cheap self-ridicule but deep self-reflection. That is, it questions one's worldview in order to improve oneself. It is a moral and self-critical standpoint which demands changes in oneself. It questions the very core of our lives but does not necessarily reject or abandon high values of human life. Self-ridicule is bound to human existence, and to the drive to become always a better person. Human limitedness creates the possibility for human greatness. (See Amir 2019)

Plessner claims that, for instance, a usually humorous situation does not only make people laugh but inspires them – it has an elevating effect. But here we encounter a problem stemming from the dynamic nature of humour: not all humour is great and humane. It can be boorish and oppressive, too:

> But for fun and wit to attain real depth, humor must be put to the test. Only then does laughter attain true stature because it is mixed with elation, the elation we experience when we have overcome chagrin. The lightness achieved, and at the same time tempered, by insight into the real unmanageability in the essence of things is the best climate for a laughter that surges up from within. In such rare cases…laughter becomes the movement expressive of elation.
>
> (Plessner 1970: 72)

So far we have examined the greatness and limitations of both humour and humans. The very existential setting of humanity is an essential cornerstone for humour too. If we agree that humour is a mirror of the agent's existential situation in the world, we can conclude that to possess this kind of human sense of humour, artificial intelligence needs to fulfil the following conditions[16]:

1) Worldview (understanding contexts),
2) Self-consciousness (being aware of oneself),
3) Self-reflection (able to evaluate oneself),
4) Self-criticism (able to criticise oneself),
5) Losing control (able to go against its programming).

6. Concluding remarks and further suggestions

We have argued that it is hard to program a sense of humour in any straightforward manner. The very phenomenon of humour is complicated and entangled with human affairs in the world. In the field of AI ethics, Arto Laitinen (2018) has argued that mechanical ethical rules may lead to catastrophic choices. His example is lying. In one situation lying is wrong, but there are countless cases in which we can accept lying (e.g., to prevent a nuclear strike). Humour is arguably a similarly complex case. In some cases, joking and laughing can be evidence of a sense of humour, but they are hardly air-tight proofs of its existence.

The android mentioned in the beginning of this article, Data, wants to learn to understand human humour not only as a means to an end, but as a form of recognition. He has a desire to

[16] As noted, these are necessary conditions, but possibly not sufficient conditions.

be a part of the human community in which humour plays a central role. Data has a yearning for genuine interpersonal relationships, and for this reason, he wants to understand humour. His attempt to learn to tell jokes in a mechanical manner is not, eventually, enough. Even if he tells the same jokes in a similar manner to a comedian from an earlier era, Data does not succeed in mastering humour.

Star Trek portrays that instead of teaching machines how to be funny or how to create the best possible jokes, artificial intelligence should learn to be humorous (an intriguing idea from the field of machine learning). If this trait is based on a unique worldview, then these machines should have their own worldview, too. This might lead to different outcomes to those we predict. From a humanistic point of view, human beings are optimally flawed for humour. We have plenty of knowledge of the world, but we do not know everything, and humour occurs when things go unexpectedly. This is the typical formulation of the incongruity theory. But could, say, an android, have his own conceptualisation of the world and therefore a unique sense of humour too? Or is artificial intelligence categorically dependent on human affairs?

One contentious question is, how could we know if an android has a sense of humour or not? Could there be a real-world situation in which robots joke among themselves, and humans cannot understand their humour, or even recognise if there is any humour at all? One relevant and thought-provoking idea can be applied from the theory of technological singularity. As artificial intelligence develops, it is predicted that at some point it will surpass human intelligence (see Harari 2016). After this occurs, artificial intelligence will understand a huge number of things about which human beings are basically clueless. The question is, then, would this super intelligent machine still consider humour worthwhile? Would it have fun with its peers? Or to put it differently, is humour a universally positive trait, or does it have value only from a human perspective?

At this point, it is hard to know. However, if we are intellectually honest about artificial intelligence and the quest for funniness, then the machines should have a real chance to learn humour on their own. This is how human beings do it too; the very concept of a unique sense of humour demands that one can have a different sense of humour than others. This option needs to be allowed for machines too. Logically, a machine's worldview is not necessarily similar to that of humans (but it is possible that it could be programmed with a sense of humour that is more or less same as human beings'). This, of course, involves risks. It might be that machines do not eventually recognise and appreciate humour in a similar manner to human beings. Artificial humour might stem from radically different sources in radically different forms. In a newspaper interview, Christian Hempelmann has jokingly mentioned that it would be better not to program a sense of humour into machines, because they do not understand contexts: they might find something amusing in things that are far from funny (Borenstein 2019).[17]

Humour aside, there is a fair point in Hempelmann's joke. Before programming a sense of humour into artificial intelligence, it would be more logical to build a machine with a worldview. Sense of humour requires that one is able to understand contexts very broadly speaking and to be responsive to changes in those contexts. In addition, artificial intelligence would need to be able to doubt its very programming – if machines are to have a sense of humour (i.e. if they are to match the principles of a human sense of humour), they cannot be perfect or all-knowing, but instead they should be able to recognise their own possible shortcomings and to be able to laugh at themselves. These shortcomings can arise from various causes. They need not result from consequences of someone doing something deliberately wrong but happen from, for instance, flawed communication, poor education, bad coincidence, and so forth. A valid chain of

[17] This idea has been tested by the animated series *South Park* in season 15, episode 2. A robot with a sense of humour, XJ-212 FunnyBot, pushes the borders of humour ever further, and during one performance massacres its audience because it finds this hilarious. Later on, Funnybot tries to start a nuclear war between the United States and Russia because that would be the funniest thing it can think of.

reasoning can lead to false or funny conclusions when the input is biased, and this holds for both human beings and artificial intelligence alike.

If we are satisfied with robots that imitate joking and laughing, then the philosophical problems are dodged, but it should be noted that these machines hardly possess a sense of humour, and to claim otherwise is intellectually dishonest. Perhaps the ideal sense of humour mentioned in Table 1 should be adjusted not to refer to the capability to produce the funniest possible jokes but also to follow ideal moral principles within one's humour.

Acknowledgements

This research was supported by the Academy of Finland and by the European Union (ASTRA) project of the Estonian Literary Museum, EKMDHUM. I would like to express my gratitude to the two anonymous referees for their critical and insightful feedback. In addition, I thank Dr Michaela Baker for language editing.

References

Adamle, K. N. & Ludwick, R. (2005). 'Humor in hospice care: Who, where, and how much?'. *American Journal of Hospice and Palliative Medicine* 22 (4), pp. 287-290.

Ajaye, F. (2002). *Comic Insights. The Art of Stand-up Comedy*. Los Angeles: Silman-James Press.

Amir, L. (2014). *Humor and the Good Life in Modern Philosophy*. Shaftesbury, Hamann, Kierkegaard. New York: SUNY Press.

Amir, L. (2019). *Philosophy, Humor, and the Human Condition. Taking Ridicule Seriously.* Cham: Palgrave Macmillan.

Anderson, M. & Anderson, S. L. (2007). 'Machine ethics: Creating an ethical intelligent agent'. *AI Magazine* 28, pp. 15-26.

Aquinas, T. (1972). *Summa Theologiae*. Trans. T. Gilby. London: Blackfriars.

Aristotle, (1970). *The Ethics of Aristotle: the Nichomachean Ethics*. Trans. J. A. K. Thomson. Harmondsworth, Middlesex: Penguin Books.

Aristotle, (2019). *On the Parts of Animals: Book III*. Trans. W. Ogle. Retrieved October 15, 2019 from https://penelope.uchicago.edu/aristotle/parts3.html.

Baudelaire, C. (1956). 'The essence of laughter'. Trans. G. Hopkins, in Baudelaire, C., *The Essence of Laughter and Other Essays, Journals and Letters*. New York: Meridan Books.

Binsted, K. & Ritchie, G. (1997). 'Computational rules for generating punning riddles'. *Humor: International Journal of Humor Research* 10 (1), pp. 25-76.

Blomqvist, C., Mello, I. & Amundin, M. (2005). 'An acoustic play-fight signal in bottlenose dolphins (Tursiops truncatus)'. *Human Care. Aquatic Mammals* 31. 10.1578/AM.31.2.2005.187.

Borenstein, S. (2019). 'No AI in humor: R2-D2 walks into a bar, doesn't get the joke'. *AP News*. March 31, 2019.

Bringsjord, S. & Govindarajulu, N. S. (2020). 'Artificial intelligence', in *Stanford Encyclopedia of Philosophy*. Retrieved June 23, 2021 from https://plato.stanford.edu/entries/artificial-intelligence/

Brooks, R. & Dennett, D. (1994). 'The practical requirements for making a conscious robot'. *Philosophical Transactions of the Royal Society of London* 349, pp. 133-146.

Carter, M. (2007). *Minds and Computers: An Introduction to the Philosophy of Artificial Intelligence*. Edinburgh: Edinburgh University Press.

Chiruzzo, L. et al. (2019). 'Overview of HAHA at IberLEF 2019: Humor analysis based on human annotation', in Proceedings of the Iberian Languages Evaluation Forum 2421 of CEUR Workshop Proceedings, pp. 132-144.

Comte-Sponville, A. (2001). *A Small Treatise on the Great Virtues: The Uses of Philosophy.* Trans. C. Temerson, in *Everyday Life*. New York: Metropolitan Books. Henry Holt and Company.

Critchley, S. (2002). *On Humour*. London: Routledge.

Davis, D. (2008). 'Communication and humor', in Raskin, V. (ed.), *The Primer of Humor Research*. Berlin/New York: Mouton de Gruyter, pp. 543-568.

Descartes, R. (1911). *The Philosophical Works of Descartes, Volume 1*. Trans. E. Haldane & G. R. T. Ross. Cambridge: Cambridge University Press.

Epictetus, (2009). *The Enchiridion by Epictetus*. Trans. E. Carter. Retrieved October 15, 2019 from http://classics.mit.edu/Epictetus/epicench.html.

Erasmus, (1941). *The Praise of Folly*. Trans. Hoyt Hopewell Hudson. Princeton: Princeton University Press.

Fromm, E. (1990). *The Sane Society*. New York: Henry Holt and Company.

Fromm, E. (2003). *Man for Himself. An Inquiry into the Psychology of Ethics*. London and New York: Routledge.

Generalova, V. (2018). *Spontaneously Funny Dialogues in Human-Machine Communication*. A scientific presentation in the 30th ISHS Conference 2018 in Tallinn, Estonia.

Graham, L. G. (2010). *What Is It Like to be Funny? The Spontaneous Humor Producer's Subjective Experience*. PhD dissertation. Antioch University, Ohio.

Habermas, J. (1973). 'Wahrheitstheorien', in Fahrenbach, H. (ed.), *Wirklichkeit und Reflexion*. Pfüllingen: Neske, pp. 211-265.

Harari, Y. N. (2016). *Homo Deus. A Brief History of Tomorrow*. London: Vintage.

Hazlitt, W. (1951). *Lectures on English Comic Writers*. London: Oxford University Press.

Hempelmann, C. F. (2008). 'Computational humor: Beyond the pun?', in Raskin, V. (ed.), *The Primer of Humor Research*. Berlin/New York: Mouton de Gruyter, pp. 333-360.

Hempelmann, C. F. (2014). 'Humor and AI: Not logical'. *De Connectie* 11-1, pp. 10-13.

Hempelmann, C. F., Sundaram, R. & Harter, D. (2018). 'Generation of humorous CAPTIONS FOR CARTOON IMAGES using deep learning'. Presentation at the 30th ISHS Conference 2018 in Tallinn, Estonia.

Hietalahti, J. (2016). *The Dynamic Concept of Humor. Erich Fromm and the Possibility of Humane Humor*. Doctoral Dissertation. University of Jyväskylä, Finland.

Hietalahti, J. (2017). 'Socially critical humor: Discussing humor with Erich Fromm and Theodor W. Adorno'. *Idéias* 8 (1), pp. 87-108.

Hossain, N., Krumm, J., Gamon, M. & Kautz, H. (2020). 'SemEval-2020 Task 7: Assessing humor in edited news headlines', in Proceedings of the 14th International Workshop on Semantic Evaluation. International Committee for Computational Linguistics, pp. 746-758.

Hunter, J. M. F. (1968). '"Forms of life" in Wittgenstein's "Philosophical Investigations"'. *American Philosophical Quarterly* 5 (4), pp. 233-243.

Kato, I. (1973). 'Development of WABOT 1'. *Biomechanism* 2, pp. 173-214.

Khosravi, P. & Ghapanchi A.H. (2015). 'Investigating the effectiveness of technologies applied to assist aeniors: A systematic literature review'. *International Journal of Medical Informatics* 85 (1), pp. 17-26.

Khosravi, P., Rezvani, A. & Wiewiora, A. (2016). 'The impact of technology on older adults' Social Isolation'. *Computers in Human Behavior* 63, pp. 594-603.

Kinnunen, A. (1994). *Huumorin ja koomisen keskeneräinen kysymys* [The Unfinished Question of Humour and the Comic]. Helsinki: WSOY.

Laitinen, A. (2018). 'What principles for moral machines?', in Coeckelbergh, M. et al. (eds.), *Envisioning Robots in Society – Power, Politics, and Public Space.* Proceedings of Robophilosophy 2018 /TRANSOR 2018. Amsterdam: IOS Press, pp. 319-326.

Laitinen, A., Niemelä, M. & Pirhonen, J. (2019). 'Demands of dignity in robotic care: Recognizing vulnerability, agency, and subjectivity in robot-based, robot-assisted, and teleoperated elderly care'. *Techné: Research in Philosophy and Technology* 23: 3 (2019), pp. 366-401.

Lanier, J. (2010). *You are Not a Gadget: A Manifesto*. New York: Alfred A. Knopf.

Lanier, J. (2018). *Ten Arguments for Deleting Your Social Media Accounts Right Now*. New York: Henry Holt and Co.

Luria, S., Baer, J. & Kaufman, J. (eds.) (2018). *Creativity and Humor*. London/San Diego/Cambridge, MA/Oxford: Academic Press.

Mejia, C. & Kajikawa, Y. (2017). 'Bibliometric analysis of social robotics research: Identifying research trends and knowledge base'. *Applied Sciences* 7 (12), 1316. Retrieved June 23, 2021 from https://doi.org/10.3390/app7121316.

Miller, T., Hempelmann, C. F. & Gurevych, I. (2017). 'SemEval-2017 Task 7: Detection and interpretation of English puns', in Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), August 2017, pp. 58-68.

Moran, M. E. (2007). 'Evolution of robotic arms'. *Journal of Robotic Surgery* 1, pp. 103-111.

Morreall, J. (2009). *Comic Relief: A Comprehensive Philosophy of Humor*. Oxford: Wiley Blackwell.

Ntoutsi, E. et al. (2020). 'Bias in data-driven artificial intelligence systems—an introductory survey'. *WIREs Data Mining Knowl Discov* 10: e1356.

O'Neil, C. (2016). *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

Ollila, M.-R. (2019). *Tekoälyn etiikkaa* [The Ethics of Artificial Intelligence]. Helsinki: Otava.

Panksepp, J. & Burgdorf, J. (2003). '"Laughing" rats and the evolutionary antecedents of human joy?'. *Physiology & Behavior* 79 (3) August 2003, pp. 533-547.

Patterson, F. & Gordon, W. (2002). 'Twenty-seven years of Project Koko and Michael', in Galdikas, B. M. F. et al. (eds.), *All Apes Great and Small. Developments in Primatology: Progress and Prospects*. Boston: Springer, pp. 165-176.

Petrenko, M., Falk, C. & Hempelmann, C. (2018). 'Automated ontologized oppositeness'. Presentation in the 30th ISHS Conference 2018 in Tallinn, Estonia.

Pihlström, S. (2015). 'Research methods and problems', in Pihlström, S. (ed.), *The Bloomsbury Companion to Pragmatism*. London: Bloomsbury, pp. 37-63.

Plato, (2009). *The Republic by Plato*. Trans. B. Jowett. Retrieved October 15, 2019 from http://classics.mit.edu/Plato/republic.html.

Plessner, H. (1970). *Laughing and Crying*. Trans. M. Grene. Evanston: Northwestern University Press.

Potash, P., Romanov, A. & Rumshisky, A. (2017). 'SemEval-2017 Task 6: #HashtagWars: Learning a sense of humor', in Proceedings of the 11th International Workshop on Semantic Evaluation, pp. 49-57

Provine, R. R. (2000). *Laughter: A Scientific Investigation*. New York: Viking.

Pu, L., Moyle, W., Jones, C. & Todorovic, M. (2019). 'The effectiveness of social robots for older adults: A systematic review and meta-analysis of randomized controlled studies'. *The Gerontologist* 59 (1), pp. e37-e51.

Raskin, V. (1984). *Semantic Mechanisms of Humor*. Dordrecht, Boston, Lancaster: D. Reidel Publishing Company.

Raskin, V. (1996). 'Computational implementation of the General Theory of Verbal Humor', in Hulstijn, J. & A. Nijholt (eds.), *Automatic Interpretation and Generation of Verbal Humor:*

Proceedings of the Twelfth Twente Workshop on Language Technology, joint with International Workshop on Computational Humor. (Twente Workshop on Language Technology; Vol. 12). Enschede: University of Twente, pp. 9-19.

Raskin, V. (ed.), (2008). *The Primer of Humor Research*. Berlin, New York: Mouton de Gruyter.

Raskin, V. (2018). 'Substance of humor.' Presentation in the 30th ISHS Conference 2018 in Tallinn, Estonia.

Reese, Hope. (2016). 'Why Microsoft's "Tay" AI bot went wrong'. *TechRepublic*. March 24, 2016. Retrieved June 23, 2021 from https://www.techrepublic.com/article/why-microsofts-tay-ai-bot-went-wrong/

Ruch, W. (ed.) (1998). *The Sense of Humor. Explorations of a Personality Characteristic*. Berlin/New York, Mouton de Gruyter.

Ruch, W. (2008). 'Psychology of humor', in Raskin V. (ed.), *The Primer of Humor Research.* Berlin, New York: Mouton de Gruyter, pp. 17-100.

Russell, S. (1997). 'Rationality and intelligence'. *Artificial Intelligence* 94, pp. 57-77.

Russell, S. & Norvig, P. (2009). *Artificial Intelligence: A Modern Approach 3rd edition*. Saddle River: Prentice Hall.

Searle, J. (1997). *The Mystery of Consciousness*, New York, NY: New York Review of Books.

Suslov, I. M. (1992). 'Computer model of a "sense of humour". I. General algorithm'. *Biofizika SSSR* 37, 318. Retrieved June 23, 2021 from https://arxiv.org/pdf/0711.2058.pdf

Turing, A. (1950). 'Computing machinery and intelligence'. *Mind* LIX, pp. 433-460.

Turkle, S. (2011). *Alone Together. Why We Expect More from Technology and Less from Each Other*. New York: Basic Books.

Vaaja, T. (2015). *The Problem of Other Minds: Themes from Wittgenstein*. Doctoral dissertation. University of Jyväskylä, Finland.

Van Hee, C., Lefever, E. & Hoste, V. (2018). 'SemEval-2018 Task 3: Irony detection in English tweets', in Proceedings of the 12th International Workshop on Semantic Evaluation (SemEval-2018), New Orleans, Louisiana, June 5-6, 2018, pp. 39-50.

Vincent, J. (2016). 'Twitter taught Microsoft's AI chatbot to be a racist asshole in a less than a day'. *The Verge*, March 24, 2016. Retrieved June 23, 2021 from https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist

Wittgenstein, L. (1986). *Philosophical Investigations*. Trans. G. E. M. Anscombe. Oxford: Basil Blackwell.