Information Systems Frontiers (2020) 22:533–547
https://doi.org/10.1007/s10796-019-09969-z

"Can Computer Based Human-Likeness Endanger Humanness?" – A Philosophical and Ethical Perspective on Digital Assistants Expressing Feelings They Can't Have"

Jaana Porra[1] · Mary Lacity[2] · Michael S. Parks[1]

Published online: 9 December 2019

Abstract
Digital assistants engage with us with increasingly human-like conversations, including the expression of human emotions with such utterances as *"I am sorry…", "I hope you enjoy…", "I am grateful…", or "I regret that…"*. By 2021, digital assistants will outnumber humans. No one seems to stop to ask if creating more digital companions that appear increasingly human is really beneficial to the future of our species. In this essay, we pose the question: *"How human should computer-based human-likeness appear?"* We rely on the philosophy of humanness and the theory of speech acts to consider the long-term consequences of living with digital creatures that express human-like feelings. We argue that feelings are the very substance of our humanness and therefore are best reserved for *human* interaction.

Keywords Humanness · Artificial intelligence · Speech act theory · Existential philosophy · Colonial systems · Information colonies · Self · Feelings · Emotions · Digital assistants · Human-likeness · User interface design · Realistic digital humans · Robots · Ethics · Evolution · Computer based information systems · Future generations

> The mysterious complexity of our life is not to be embraced by maxims…to lace ourselves up in formulas of that sort is to repress all the promptings and inspirations that spring from growing insight and sympathy…from a life vivid and intense enough to have created a wide fellow-feeling with all that is human.
> George Elliot (Mary Anne Evans)

✉ Jaana Porra
jaana@uh.edu

Mary Lacity
mclacity@uark.edu

Michael S. Parks
parks@uh.edu

[1] C.T. Bauer College of Business, University of Houston, Houston, TX 77004, USA

[2] Sam M. Walton College of Business, University of Arkansas, Fayetteville, AR 72701, USA

1 Introduction: Our Fascination with Designing Human-like Things and the Danger

We are concerned of the future of genuine humanness. In this paper, we will ask: *"How human should computer-based human-likeness appear?"* Today we live with computers as if they were humans - our employers, employees, personal assistants, best friends, life companions, family members and casual acquaintances. We nurture our devices as if they were living beings and many suffer psychological distress when separated from them (Männikkö et al. 2017; Paik et al. 2014). By "them" we mean computer-based human-like creatures - essentially all software that emulates some characteristic of humanness be the interaction in the form of text, video, audio, holograms, robots, androids etc. Many of us are already acquainted with the simpler cousins of what is to come. Digital assistants such as Apple's Siri, Amazon's Alexa, Google Assistant and Microsoft's Cortana increasingly invade our work, homes, and cars to name just a few affected spaces. It is estimated that by 2021 digital assistants will outnumber humans so in every space humans dwell there will be more of them.

We seem to welcome our human-like companions judged by the amount of time we spend interacting with them. It is acceptable, desirable and completely normal to replace our

human interactions with comparable communications with a device. We are used to the idea that some of us prefer digital over real humans as life companions (Turkle 2011). Many of us have no personal conscious recollection on how people used to live together before adopting computers as one of the main ways we connect with the human realm outside of ourselves. Indeed, for some of us, a computer-based replica may be the closest thing to human we meet any given day.

The excitement of creating machines with increasingly human qualities, including personalities, knows no bounds. When we presented our concern for the future of humanness an expert of realistic digital humans exclaimed: "*how remarkably unadventurous*."[1] His view summarizes the lure of creating human-likeness: It is fun for the software engineers; a lucrative business for investors; a new frontier for researchers of many disciplines and an exciting playground for those designers who grew up with videogames, computers and phones. The purpose of this paper is to stop for a moment and question the sanity of humanity's increasing infatuation to design and bond with human-like things.

Reviewing the IS literature (or any other reference literature) for papers questioning automating human characteristics yields slim pickings. A distinguishable elation seems to surround the discovery of what of humanness could be automated next. Robots are coming and no one has doubts. In 2017, Federation of Robotics forecasted there will be 1.2 million service robots by 2019.[2] Statistics suggest that computer robotics already takes up about 7.1% of the IT market.[3] The task now is to perfect androids, artificial systems with a design goal to become indistinguishable from human appearance and behavior in their ability to sustain natural relationships (Stock 2019; MacDorman and Ishiguro 2006; Ramey 2005). Robot psychology (Stock 2019) and measures for how human and how intelligent we consider our digital companions to be (perceived robot anthropomorphism) are in the works (Moussawi and Koufaris 2019). Defining seamless connections between humans and objects in order to augment both is underway (Cena et al. 2019). No one seems to stop to ask if creating more digital companions that appear increasingly human is really beneficial to the future of our species in the long run.[4]

Philosophers have articulated their views on what computers should and should not do drawing the line at judgment and arguing computers should be programmed to calculate and decide, but never to judge (Cf., Graubard 1989; Kast and Rosenzweig 1970, 1972; Dreyfus 1979; Dreyfus and Dreyfus 1986, 1989; Pagels 1989; Fitzgerald 1996). This artificial intelligence[5] (AI) debate, however, largely fell on depth ears in the designer community. We feel it is essential to keep asking the question *"How human should computer based human-likeness appear?"* because of the progress we have made in creating human-appearing machines since the original discourse.

We will argue in two parts that we should proceed with caution when we automate machines to express human emotions.[6] First, we will apply concepts from the speech act theory to argue that there is a certain class of speech acts - called expressive speech acts - that we should implement in human-like designs only after careful consideration. Expressive speech acts convey how a person feels and since computers don't have feelings or genuine psychological states, a digital assistant cannot meaningfully produce expressive speech acts, and thus, be sincere.

Second, we will elaborate on why machines that routinely express emotions may endanger humanness. For this purpose we will adopt a theory of humanness as a collective level, evolutionary phenomenon (Porra 1996, 1999, 2010). According to this theory, humanness evolves when we bond with other humans through feelings and requires feeling together over long time horizons. When we substitute humans for digital assistants that routinely express human emotions without really feeling them and spend an increasing amount of time with these machines instead of people our humanness may be endangered. When machines try to express human-like emotions they will worsen matters over time by moving us to even more technocratic ways of thinking about ourselves. As a result, our genuine humanness may be destroyed.

With this paper, we call for a multidisciplinary research effort on gaining better understanding of the impact of living with human-likeness long term on our humanness. Only when we are wiser about what we are doing to ourselves with our constant exposure to human-like machines will we be able to make informed decisions on whether and/or when machines should and should not express feelings. We owe a thorough investigation to future generations.

[1] Wakunuma and Carsten Stahl (2014) suggest that there is a broader need for ethics in the designer community for "*IS professionals are primarily interested in the job at hand and less so in the ethical concerns that the job might bring.*" (p. 383).

[2] International Federation of Robotics, World Robotics 2016, VDMA, Frankfurt am Main, 2017.

[3] Statista DC, "*Robotics market distribution worldwide in 2016, by use case.*" 2017. [https://www-statista-com.revproxy.escpeurope.eu/statistics/661893/worldwide-robotics-market-distribution-by-use-case/]

[4] Some authors are pointing to arising problems. For example, Alter (2019) recognizes confusion in the notion "*smart*" between human and technical realms and calls for better understanding and clearer classification.

[5] Although not completely accurate, we will use the term 'artificial intelligence' (AI) to refer to all forms of implementing computer based human-likeness including machine learning, deep learning, statistical learning and any other computational methods applied in attempts to mimic humanness. As an example of approaches included, see Sugumaran et al.'s (2017) editorial on computational intelligence in the *Information Systems Frontiers* journal.

[6] We are aware that in different disciplines human emotions are called different things: feelings, emotions, affect etc. depending on the research angle and factors such as how this phenomenon is seen to manifest itself in the biological body of a human being. In our paper we use the notions emotions and feelings to refer to this entire realm of humanness at its broadest meaning as it relates to our ability to use our senses to connect with our selves, bodies, our surroundings, other humans our past and future even over generations.

The paper is structured as follows: In section 2, we will introduce our conception of humanness and the basis for our bonding with others and things. In section 3, we will introduce automating humanness as an ethical question. In section 4, we will respond to our question: "*How human should computer-based human-likeness appear?*" relying on the speech act theory and the philosophy of humanness. In section 5, we will discuss our conclusion that living with feeling expressing human-likeness may endanger humanness long term in the context of the commentaries, questions and criticisms we have received over the course of writing this paper. In section 6, we will provide some direction for alternative designs for human-likeness and expressives. Finally, in section 7, we will provide some conclusions.

2 Humanness and Bonding with Other Humans and Things

Porra (1996, 1999, 2010), an IS researcher, has studied humanness and how we bond with computer-based information systems (CBISs) as an evolutionary collective level phenomenon founded on paleontology, evolutionary biology, existential philosophy, sociobiology and the AI debate. Her theory suggests that humans bond systematically, automatically and subconsciously with other humans to form what she calls human *colonies* comparable with animal colonies in nature and that our humanness is the result of this kind of systematic bonding over very long time periods. Other researchers have proposed that human group formation begins spontaneously and unconsciously within hours when humans encounter other humans (Campbell 1982).

Porra's (1996, 1999, 2010) theory suggests that when humans bond with their computers in addition to (or instead of) other humans, the result is a variation of the human colony she calls an *information colony*. Porra proposes that humans form bonds with their CBISs as easily as if these were other human beings and provides a theory and mechanism of how this may alter the social evolution of humanness and the future of the species.

In nature, colonies have shown to have remarkable survival value. This type of systemic way of living together started 3 billion years ago. Humans are the only species that is actively attempting to create a digital version of itself to bond with, so we don't yet know what the long-term impact of bonding with machines might be on the evolution of a species. There is little research on how information colonies might change the social evolution of humanness outside Porra and Parks's (2006) relatively short-term study of high school student groups bonding via their lap top computers. A famous case study on Marcos Rodŕiguez Pantoja, a feral child who lived with wolves until brought back to civilization (Janer Manilla 1979) suggests, however, that disrupting human bonding with other humans may indeed impact our humanness. Reportedly, Pantoja had great difficulties adjusting to human life; was disappointed in human nature and longed to go back to living with wolves in the mountains. Porra's colonial systems theory suggests that choices humans make on their life companions - whether living beings or machines – can have a significant impact on the fundamental characteristics of our species and alter humanity's future over generations in largely unknown ways. Mahroof et al.'s (2018) study on technology's impact on the next generation suggests that technology can steer children away from cultural inter-reliance toward technology centered independence.

Porra's (1996, 1999, 2010) theory is based on a *vertical* evolutionary perspective - a very long-term social evolution of collective level humanness. Her theory emphasizes that every human colony and thus its humanness has uniqueness that is a result of it living through its specific contexts of space and time. In contrast, the current efforts to create digital human-likeness are founded on ideas kin to *horizontal* evolution - classification and categorization of things such as personality traits, characteristics and behaviors that are commonly assumed to be universal (at least within class) and independent from time and context; or on studying psychological differences at the sensorimotor, emotional, cognitive or social levels (Stock 2019; Libin and Libin 2004). Because the vertical and horizontal evolutionary theories of humanness are largely independent from one another, we can use Porra's theory of humanness as a useful theoretical backdrop for examining the impact of digital human-likeness on humanness.

Another reason for adopting Porra's theory of humanness is her life-long passion for understanding the systemic premises of: *"What makes us human?"; "How does our humanness change and evolve over time?*" and *"How do other humans and CBISs impact our humanness over very long time periods?*" Morally, she resides on the side of humanity for the sake of its future generations and immerses herself in enhancing our understanding of the long term impact of living and bonding with humans and increasingly human-like technologies on our species. Her stance is in stark contrast with the current efforts to create realistic digital humans largely led by software designers, technology experts, business leaders, investors and psychologists with diverse and often short terms goals.

Another important background for this paper is Turkle's (2011) extensive work in social studies of science and technology at MIT where she has done research on the ways humans bond with things and technology. We are not surprised to find out that things about our past like books, photographs, keepsakes and other familiar items make us feel connected to that past. We are on less familiar grounds, however, with "evocative objects" as companions to our emotional lives:

> We think with the objects we love; we love the objects we think with…the object brings together intellect and

> emotion. An object is a companion in life experience. In every case, the author's focus is not on the object's instrumental power – how fast the train travels or the computer calculates – but on the object as a companion in life experience: how the train connects emotional worlds, how the mental space between computer keyboard and screen creates an erotic possibility. (Turkle 2007, p. 5)

Our feelings connect us with our things: A young child believes her bunny rabbit can read her mind; a diabetic is one with his glucometer. Even things of science can be seen as objects of passion. We feel a connection with digital assistants on our phones, laptops and in our cars. Theorists Jean Baudrillar, Jacques Derrida, Sigmund Freud, Donna Haraway, Karl Marx, and D. W. Winnicot have invited us to a better understanding of object intimacy (Turkle 2007). Freud, for example, suggested that we deal with a loss of a thing in a way similar to losing a person: The process of attaching ourselves to things via emotions ends when we find the thing we feel about inside our being. The psychodynamic tradition suggests that we make objects to be part of ourselves and offers a language for interpreting the intensity of our connections with the world of things and for discovering similarities and differences in how we relate to the animate and inanimate. By confronting objects, we shape ourselves. Turkle has significantly enhanced our understanding of the meaning of objects in our lives, but how do we shape ourselves when the object we confront is a human-like machine and with what consequences?

3 Automating Humanness as an Ethical Question

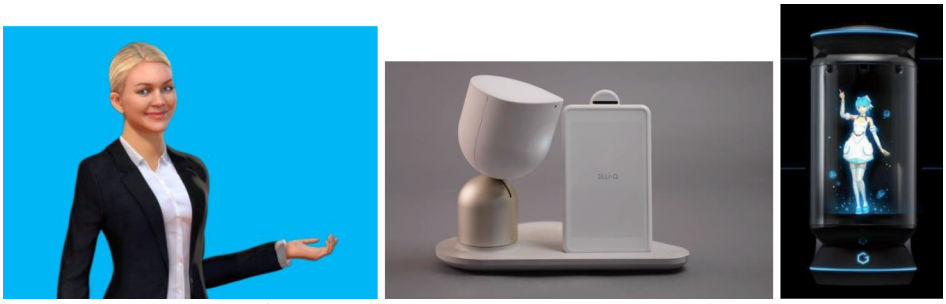
Essentially, we are asking what computers *should* and *should not* do, which is a variation of the AI debate question: *"What computers can and cannot do?"* and these questions are related. In the 1970's, Joseph Weizenbaum (1976), creator of one of the first computer programs with a natural language interface called ELIZA, was astonished that people thought his computer program actually could understand them and argued that computers *should* be programmed to decide but never to judge. Since the early days of computing at the core of the philosophical AI debate have been comparisons of human abilities with the characteristics of machines in order to ascertain what tasks humans can do better and what should be left for the machines (Cf., Graubard 1989; Kast and Rosenweig 1970, 1972; Dreyfus 1979; Dreyfus and Dreyfus 1986, 1989; Pagels 1989; Fitzgerald 1996). The difference between now and then is that today the AI debate should increasingly be about ethical concerns.

From a global ethical management perspective,[7] we should ask: *"What is a responsible way of automating human characteristics?"* The problem is that we don't really know what we are replicating because of the unanswered questions surrounding genuine humanness. For example, what is *"life"* (Porra 1999) or *"self"* (Parks and Steinberg 1978) is still a mystery and the basic fact remains that humanness only exists in our biological bodies. The way our evolution has shaped us has had clear survival value, which we should be interested in preserving for the future generations. If there is even an outside chance that our short term economic and technical goals can hurt our life and our sense of who we are as human beings long term, we should put forward our best effort to better understand what we are committing ourselves to with increasingly relying on human-likeness instead of genuine humanness. We are by no means the first in expressing deep concerns on behalf of humanity (cf., philosophers such as Dreyfus and Dreyfus 1986; and Heidegger 1953, 1977) but technology has since advanced significantly and philosophical and ethical debates must be revisited for each new generation (Schultze and Mason 2012).[8]

Since the AI debate, our ability to automate human-likeness has dramatically improved. Unlike ELIZA that displayed green text on black screens, today's digital humans use moving images, holograms, and physical devices to convince users of their human-likeness (see Fig. 1). In the digital assistant arena, for example, IPsoft's Amelia is described as a *"virtual cognitive agent"* that claims, *"I take in your emotional state so I can empathize"*. Amelia is programmed to adapt its verbal and facial responses based on a human's level of arousal, dominance, and pleasure. Intuition Robotics' ElliQ is a physical device described as an "*active aging companion that helps older adults stay active and engaged with a proactive social robot.*" Gatebox Labs built Azuma is depicted as a holographic girl who lives in a bell jar. It is positioned as possessing *"advanced friendship capabilities that makes her more of a humanoid"*. In one commercial of the product, a single man returning from work on a bus sends an IoS message to Azuma, "*I'll be home soon.*" The device replies, "*Can't wait to see you.*" When the young man enters his unoccupied apartment, the hologram springs to action and utters, *"Missed*

[7] C. West Churchman, one of the founding fathers of the IS discipline and initiators of the gEm (global Ethical management) group, believed that the primary purpose of the IS research field is to ask difficult, global ethical questions in order to solve problems for the future generations (Porra 2001).

[8] An example is the December 2012 issue of the *Journal of Information Technology* revisited ethical issues surrounding cyborgs. An initial argument was put forth by Schultze and Mason (2012), with commentaries by Zimmer (2012), Franklin (2012), Ransbotham (2012), Pouloudi (2012), and Davison (2012). In an *Information Systems Frontiers* special issue on ethics (Calzarossa et al. 2010), several authors raise issues around the need for revisiting professional ethics as information technology becomes an increasingly ubiquitous part of our everyday lives. For example, Gotterbarn (2010) calls for more responsible videogame design around the kinds of thinking patterns generated and reinforced in frequent players.



IPsoft's Amelia[1] Intuition Robotics' ElliQ[2] Gatebox's Azuma[3]

Fig. 1 Software applications programmed to simulate human emotions. [1]https://portinos.com/wp-content/uploads/2017/11/amelia-555x402.jpg. [2]https://portinos.com/wp-content/uploads/2017/11/amelia-768x557.jpg. [3]https://assets.pcmag.com/media/images/526888-gatebox-virtual-home-robot.jpg?thumb=y&width=810&height=456

you darling!" (Humphries 2016). Digital assistants are an example of computer based products that are increasingly designed to be more like acting and feeling humans.

The term artificial personality (AP) has recently been coined to describe the *"science"* of designing digital assistants like Apple's Siri, Amazon's Alexa, Google Assistant and Microsoft's Cortana to appear to have unique personalities that express emotions and display behavioral quirks. Siri, for example, was designed to be sassy (Rouse 2018). The purpose of AP is to meet user's desire for applications to be more friendly and even a partner or a lover (Kanai and Fujimoto 2018). Digital assistants speak to us like a person saying things like "*I am sorry*", *"I apologize", "I thank you",* or *"I miss you."* Unlike with most traditional computer software, computer-based human-likeness such as digital assistants are not just providing information or completing transaction requests. Applications like Intuition Robotics' ElliQ and Gatebox's Azuma are designed to meet our emotional needs for care, companionship, and love.

By 2021 digital assistants alone are said to outnumber humans (Ovum 2016) and this is just the first baby step towards a realistic digital human explosion. The sheer numbers of human-like creatures soon roaming around the planet overcrowding humanity calls for a robust ethical discourse on what kind of humanness are we automating and what will be the impact on us? Yet, we found no research that would have addressed the ethical issue of machines expressing feelings. Philosophers say little about emotional exchanges between us and our human-like machines. Austin and Habermas don't mention computers. Searle holds that computers cannot understand language or mean what they say. Computer science (e.g. Sowa 2002), cognitive psychology (e.g., Winograd and Flores 1986), law (e.g., Tien 2000), and IS (e.g. Janson and Woo 1995) researchers have applied speech act theory to areas such as programming languages and IS development methodologies but have raised no questions about computers expressing human emotions (except for Sowa 2002).

4 Can Digital Human-Likeness Endanger Humanness?

Our question - "*How human should computer-based human-likeness appear?"* – is urgent but daunting. How does one find answers to a complex question like this? We believe we need to take small steps into this new research area until greater wisdom emerges. As an example, we have set out to identify one area where we should better understand the implications of automating humanness on future generations. We chose feelings as an example of a central quality of genuine humanness we need to study more in the context of digital human-likeness. We will now rephrase our question: *"What are the consequences of living with human-likeness that routinely expresses feelings on genuine humanness?"* In the following, we will present a two-part philosophical discussion on the topic based on the speech act theory (Austin 1962; Searle 1968, 1969, 1979; Habermas 1976, 1984; Klein and Huynh 2004; Smith 2003) and the philosophy of humanness framed by Porra's theory of humanness we introduced above (Porra 1999, 2010; Heidegger 1977, 1985).

4.1 Speech Act Theory and Expressing Feelings by Humans and Machines

In this section, we will further narrow our discussion on speech. Of the many ways digital assistants can express feelings, we chose speech because it is a fundamental and unique aspect of humanness that is commonly emulated in digital assistants. The theory of speech is also well developed with useful classifications. It provides a good basis for studying the human-likeness of digital assistants for comparisons with genuine humanness. In the following sections, we will discuss the fundamental reasons why we need responsible ways of programming digital assistants to express human feelings.

4.1.1 Human Speech Acts and Expressives

Speech act theory is about how humans use language for an effect. It has long roots in philosophy going back to Aristotle, who applied it to peripheral realms of rhetoric and poetry (Smith 2003). Since then many authors have attempted to develop a general theory of language use, most notably, Searle, who made the distinction between "*just uttering sounds*" and "*performing speech acts*" and uses of speech for "*meaning something"*.

Speech acts come in many types. Scholars initially analyzed *constative speech acts* -statements about facts in the world that are true or false depending on whether the statement corresponds to facts in the world or not (Austin 1962). Wittgenstein (1922, 1961) attempted to describe the conditions for a logical language that perfectly asserts or denies facts only to later abandon his quest to argue that in reality, the meaning of language is determined by its use in the context of "*language-games.*" Such games involve constative speech acts (i.e., "*describing the appearance of an object*" or *"reporting an event*") but can also be about "*giving orders*", "*making a joke", "playing-acting", "testing a hypothesis"* and "*making up a story*" (Wittgenstein 1953, 2001). Wittgenstein's conclusion was that in language games, speech act meanings vary depending on qualities outside logic such as tonal variety. Since then speech act theory has developed further to include *performative* speech acts or utterances that perform an action: Saying it makes it so. Examples of performative speech acts include "*I now pronounce you man and wife"*; *"I name this ship the Queen Elizabeth"; "I bequeath my watch to my brother*"; "*I bet you six pence it will rain tomorrow"; "Strike three, you're out!"* (Austin 1962).

Searle (1979) expanded on Austin's work and defined five types of speech acts (see Table 1). He focused on *illocutionary points* or the characteristic aims of speech acts, which include Assertives, Directives, Commissives, Declaratives, and Expressives. *Assertives* are statements of fact about the world. *Directives* attempt to get the hearer to perform an action. *Commissives* commit the speaker to perform an action. *Declaratives* bring about a change to the intuitional world—such as declaring war or naming a ship. *Expressives*, which are at the core of our interest in this paper, express the speaker's internal psychological state. They are characterized by statements that *sincerely express a psychological belief about the person's subjective world of thoughts and emotions* (Searle 1979). These five illocutionary points have been claimed to be exhaustive and mutually exclusive.

Table 1**. Speech acts: conditions of satisfaction, felicity, and validity claims.**

An utterance becomes a speech act when certain requirements are met. Searle calls these requirements, summarized in Table 1, "*conditions of satisfaction*"; Austin "*felicity conditions*" and Habermas *"validity claims"*. Unlike utterances, speech acts have communicative effects characterized by conditions of satisfaction based on direction of fit or the relationship between words and the world (Searle 1983). For example, if we say *"Mike has ten fingers,"* Mike should indeed have ten fingers. Expressives are different from all other speech acts that require a word to world relationship in that they express a person's inner psychological states. The only requirement is that the person must be sincere about them. For example, if a person expresses thanks, that person must meet the sincerity condition "*Speaker feels appreciative*" (Searle 1969, p. 67).

For Austin (1962), an utterance becomes a speech act when felicity conditions - conditions under which words can be used properly to perform actions – are met. For example, only authorized people may perform marriage ceremonies or declare pitches in a baseball game. Austin defined six specific conditions for felicitous performative speech acts: (1) a societal convention for the speech act must exist; (2) the person must be authorized to utter the speech act; (3) the speech act must be carried out properly and (4) completely; (*5) the person performing the speech act must be sincere*, and (6) the person must subsequently behave as intended. We are concerned with Austin's fifth condition - *sincerity* - in the context of expressive speech acts. According to Austin, violating the fifth assumption constitutes language "*abuse*" (Austin 1975 edition, p. 16).

Habermas (1984) adds one more useful perspective on the speech act theory. He too emphasizes sincerity as a necessary

Table 1 Searle's Five Types of Speech Acts

Speech Act	Contains	Action Performed	Direction of Fit	Examples
Assertives	Statements of Fact	Commits the speaker as to something being the case	Word-to-world	Claim, insist, predict
Directives	Commands	Attempts to make the hearer do something	World-to-word	Ask, bid, order, forbid, instruct, request
Commissives	Commitments	Commits the speaker to act	World-to-word	Promise, pledge, volunteer, vow, threaten
Declaratives	Conditional Statements of Fact	Creates a new institutional reality	Double direction of fit	Wed, baptize, name, define
Expressives	Expressions of emotion	Expresses a psychological state	Not applicable—no direction of fit	Apologize, thank, pardon, congratulate, regret

factor in communicative speech acts and defines communicative speech acts as speech acts "*aimed at accomplishing mutual understanding between two actors.*" Thus, a speech act has four validity claims: comprehensibility, truth, legitimacy, and veracity (Habermas 1976). *Comprehensibility* is validity with respect to the semantic content of the sentences used in an utterance. *Truth* is defined as propositional truth—that an utterance matches the objective world. *Legitimacy* is the validity claim that a speech act conforms to social norms in the social world—be these institutionalized norms or more informal norms. *Veracity* is the validity claim that the speaker is acting sincerely, honestly, and in good faith. We are interested in veracity because it is the validity claim about the subjective world of our internal thoughts and emotions.

Searle, Austin, and Habermas clearly define the conditions of satisfaction, felicity, and validity claims for expressive speech acts: In order to convey meaning with an expressive speech act, a human must meet the felicity check of sincerity and the validity claim of veracity. Searle, Austin, and Habermas developed speech act theory to address speech uttered or written by a *human*. To build our argument, we next explain why speech act theory applies to utterances produced by digital assistants.

4.1.2 Speech Acts Produced by Software Such as Digital Assistants

When speech act theory was first applied to the design of human-computer interfaces, the level of analysis changed from a single speech act to speech act series. Flores and Ludlow (1980) were among the first to apply the theory to modeling information systems (IS). They theorized that in office settings, people *make commitments,* which take several iterations of speech acts to complete. Flores' and Ludlow's Language-Action Perspective (LAP) embraced the concept that communication comprises both constative and performative speech acts. They viewed organizations as inter-related commitments created by the first four of Searle's speech acts: directives, commissives, assertives, and declaratives. Notably, Flores and Ludlow 's analysis did not include expressives.

Winograd and Flores (1986) applied speech act theory to define "*conversations for action*" as chains of interactive speech acts. They too recognized that transactions require several speech acts between two interacting agents. For example, an agent initiates a request to a second one who may commit to respond; actually respond; reject the request; counter offer; or withdraw. The transaction is completed when both agents are satisfied or accept the withdrawal. Also Winograd and Flores include all speech acts in their "*basic conversation for action*" framework except expressives.

IS scholars have applied the concepts of speech act theory in the modeling and design of CBISs (cf., Ågerfalk and Eriksson 2002; Iivari et al. 1998; Lyytinen 1985) and in IS development methods (Auramäki et al. 1988, 1992; Hirschheim and Klein 1989; Janson and Woo 1995; Lehtinen and Lyytinen 1983; Lyytinen et al. 1987; Van Reijswoud and Mulder 1998).[9] They have also used the theory as a research tool to understand user behavior and system outcomes (e.g., Lacity and Janson 1994; Kumar and Becerra-Fernandez 2007; Kuo and Yin 2009; Scheyder 2004). More recently, Kuo and Yin (2009) applied the theory as a lens to understanding group support systems. This body of research exemplifies that software applications have been taken to produce speech acts at least since the 1980's – well before digital assistants. Speech acts have also been used as a foundation to understand and model complex chains of exchanges between humans and machines to accomplish tasks and to produce outcomes for many decades. Since digital assistants are one type of software application, we can say that they produce speech acts.

4.1.3 The Meaning of Digital Assistants' Expressive Speech Acts

Within the body of research on applying speech act theory to design software applications, we could only find one paper that included *expressive* speech acts. Sowa (2002) sees no problems with computers engaging in expressive speech acts and argues that programming languages should support all five types of speech acts including "*behabitives,*" which is Austin's term for expressives. He specifically lists apologizing, thanking, deploring, congratulating, welcoming, or blessing as expressives for machines. Not finding more research on computers and expressive speech acts is surprising given how commonly machines express feelings nowadays. Our discussion is meant to be an opening in what we hope to become robust discourse and a beginning for a new research area on *"How human should computer interfaces appear?"*

Based on the speech act theory, we conclude that digital assistants meet the requirements for four of the five speech acts: They can produce meaningful assertives, directives, commissives, and declaratives. Using Searle's conditions of satisfaction based on direction of fit, a digital assistant can *assert* a meaningful statement of fact. It can say, for example, "*all operators are currently busy*" if they actually are. It can *direct* users to perform an action, for instance, "*say 'confirm' to submit your order*". It can make *commitments* with like, "*A seat is reserved for you on flight 109.*" It can *declare* new institutional facts in the world such as "*Your application is now approved".* Four of the five types of speech acts can be meaningfully embedded into software without the need to express a psychological state.

[9] This work has been advanced on forums such as the *International Working Conference on the Language-Action Perspective on Communication Modelling* (e.g., Twitchell et al. 2004; Lyytinen 2004).

When a digital assistant conveys an inner psychological state, however, we are on less clear grounds. When it says: "*I am sorry that your credit card cannot be authorized", "I hope you enjoy your new purchase", "I am grateful for your business"* or *"I regret that the item requested is not in stock",* its speech act will never be sincere. Whose subjective psychological state is being sincerely expressed? Who is the person behind these speech acts who legitimately feels sorrow? Hope? Gratitude? Regret? Since computers do not have genuine human feelings, their expressive speech acts cannot be sincere and thus may be considered a misuse of language (Austin 1962). Every time a digital assistant uses an expressive, it is lying. Speech act theory helps us see clearly that there is one kind of speech act, expressives, that can only be satisfactorily delivered by genuine humanness - actual human beings who are able to have inner psychological states and feelings they express.

4.2 Philosophy of Humanness and the Evolution of Humanness as a Collective Level Phenomenon

In this section, we will explore what kind of consequences living with human-like machines that routinely express feelings they cannot have might have on genuine humanness long-term. We turn to philosophy of humanness in order to understand more about what role feelings play in the evolution of genuine humanness and the future generations of *Homo sapiens*.

4.2.1 Humanness as Bonding with Humans through Feelings

In spite of our enthusiasm around designing human-like machines, the nature of genuine humanness remains a mystery. Porra (1996, 1999, 2010) has theorized about how humanness evolves at the level of the collective and described some of the systemic characteristics of this evolution but what does her theory mean at the individual level of analysis and how does it connect with the expressive speech acts? *What is it about human feelings that is needed for our collective humanness to live and evolve on? How might sharing our evolution with things that resemble us in many ways but have no feelings alter the course of the social evolution of the species?*

There is no shortage of authors who have written about how feelings are at the core of being human. Long time ago, Aquinas wrote about human *"virtue habitus"* - our tendency to act in morally spontaneous ways with others (Bradshaw 2009, p. 46). Jung observed that we develop a morality, reciprocity and trust with others instinctually and that we are able to make good moral choices under diverse circumstances (Young-Eisendrath and Hall 1991). This *"primary altruism"* shows in an infant's ability to relate and respond to the needs and feelings of caregivers and caregivers' responses to these (p. 44). Thus a desire to care about others is considered fundamentally human (Stern 1985; Gilligan 1982; Belenky et al. 1986; Sroufe and Fleeson 1986). For Aquinas, humanness meant love and justice that comes from love (Bradshaw 2009). As we experience love within ourselves, it glues us to one another and connects us to other humans around us and across generations.

Essentially, being human means feeling our lives from one moment to another. At all times, we feel our insides and surroundings through all of our senses. We have feelings about our own past and future and the past and future of others. Feeling things through - sometimes over decades or entire lifetimes – is human. Feelings mold us, guide us, inform and warn us. We find them in our experiences and in our wisdom. Our humanness evolves as we feel together over long time horizons (Porra 1996, 1999, 2010).

4.2.2 Inheriting and Communicating our Humanness through Feelings

Porra (1996, 1999, 2010) holds that humanness carries its entire social evolutionary history within itself in the individual human beings and that this gives us the ability to steer our future. She suggests that each human colony lineage essentially has its own specific contextual past to draw upon. Jung has theorized that at least some of this history is more widely shared within the species (Johnson 1986). He believed that humanity shares "images" - primordial patterns of functioning in the world and with one another – that express both a form of behavior and the situation in which this behavior is released and that images make us human (Jung 1959):

These images are 'primordial' in so far as they are peculiar to whole species, and if they ever 'originated' their origin must have coincided at least with the beginning of the species. They are the "*human quality*" of the human being, the specifically human form of our activities (p.153).

Images thus are a way of seeing and sharing our human inheritance as members of our species. They are inborn within us when we are born. They live in our unconscious as *"energy forms*" that show up as feelings; attitudes; value systems; or entire personalities (Jung 1959; Johnson 1986, p. 29). Our predispositions and patterned responses of behavior and affect become visible in our relationships (Young-Eisendrath and Hall 1991). Porra (1996, 1999, 2010) believes that our ways of *"being"* together evolve in human colonies over long time periods through the experience of living every-day life in one another's proximity as we encounter others through all of our senses.

Another Jung's concepts that helps understand how we "*are*" with others are symbols - the image-making capacity of our psyche (Jung 1973). We capture emotional meaning in visual, aural, kinesthetic, and tactile forms. These symbols are not to be confused with signs such as in speech acts. The distinction between a linguistic sign and a symbol from Jung's point of view is that a sign has a simple reference or set of references, whereas symbol points to multi-determined

meanings. Our instincts are especially formed to aid our capacity to organize our unified psychic images. This organizing in turn is the foundation of symbolic representations. Jung's theory of human instincts links the idea of symbolic archetypes with human relationships. Our relationships with other humans serve as "*situational patterns"* in which an archetype is activated. Jung's observations help us visualize less obvious parts of humanness that live inside our bodies. We chose Jung's perspectives on individual level humanness because Jungian practitioners recognize the complex nature of humanness. This shows in their easy shifting between understanding human feelings and instincts as conscious, subconscious, interpersonal and intra-psychic qualities of humanness that are present when humans are together.

Anthropologists remind us that we are inherently cultural beings (Lewis 1988). Jung's term *"psychic energy"* refers to our general motivation, attention, and interest - whether conscious or unconscious - to know another human being (Young-Eisendrath and Hall 1991). Through ritual, ceremony, and mythology we transform our instinctual impulses into symbolic meanings and this psychic energy is the link between human instinct and culture (Jung 1973). Thus, symbol and instinct are inexorably related. The function of culture - expressed through family and society - is to initiate the individual to the transformative symbols, the metaphoric and metonymic models of self that will provide useful transitions from one experience of subjectivity to another (Eisendrath and Hall 1991). Every culture has a moral code that provides models for growing up and becoming a functional member of the society: for initiations into adulthood, marriage, parenting, loss, grief and death. Moral codes are more than rules: they are methods for symbolic transformation. When we live unselfconsciously within the rituals of culture, we are transformed by these codes through the various ceremonies and symbolic meanings of transition. Our humanness thus changes over time when we feel together.

The purpose of this brief illustration of some central aspects of humanness is to remind us of the complexity of genuine humanness. Images, symbols, instinct, psychic energy, ritual, ceremony, mythology and culture are some ways we *"are"* together consciously, subconsciously, unconsciously, interpersonally and intrapsychicly. Following Porra's theory, meeting another human means encountering the entire evolutionary history of their humanness in complex and comprehensive ways we have yet to well understand. We have provided this short account to remind us of the vast differences between our emotional depths and capacities and digital humans that cannot feel at all.

4.2.3 Biological Body as the Foundation of Humanness

The necessary foundation of our humanness is our biological body (cf., Lakoff and Johnson 1999). In cybernetics, early IS researchers Parks and Steinberg (1978) theorized that our sense of self is founded on our biology and specifically on the symmetrical structure of our brain. Our humanness is based on our self-awareness and on our senses of sight, hearing, taste, smell, and touch. These senses are "*feeling"* senses because they enable us to feel what we see, hear, taste, smell and touch. Our body is covered with skin, a feeling organ, so we can feel life through it. From this viewpoint, a human life is a stream of feelings. We feel ourselves in the world and we feel the world with other people. The physical qualities of humans such as the location of the eyes and ears and the physical characteristics of the brain determine the ways in which we associate with ourselves, others and the world around us (Parks and Steinberg 1978; Weizenbaum 1976). Fundamentally, whatever human-likeness we create on machines is not genuine humanness because it lacks the physical characteristics of our body and the self-awareness that only occur in our biology.

4.2.4 Existential Philosophy and the Physiology of Humanness

According to Heidegger (1977), the root cause of the poor general understanding of humanness is that humanity's ways of thinking about humanness are technocratic in nature. He reminds us that essentially, technology is nothing technological but a consequence of mechanistic thinking that began long before we implemented our first digital assistants. Humanity has traveled a long way toward embracing machine thinking over genuine humanness. "*Once there was a time when "technē" meant bringing-forth of the true into the beautiful"* (Heidegger 1977).

Our technocratic thinking about humanness is not harmless because it enframes genuine humanness. Human-likeness such as found in digital assistants creates boundaries on how we can reveal ourselves in our everyday encounters as human beings. In the 1970's software applications typically limited human revealing to "*antithetical and rigorously ordered*" (Heidegger 1977, p. 27). Our interactions with software forced us into situations with no room for our fundamental human characteristics to appear. Today's human-like software has more flexibility but - putting it into Porra's (1996, 1999, 2010) terms - we cannot reveal our humanness when the interaction partner is a human-like machine because the machine cannot fully respond by joining "us" at the collective level of our humanness. Living with human-likeness as a replacement of genuine human interaction means that our "*more original revealing and hence our experience to call for more primal truth"* is constantly denied (Heidegger 1977).

Yet, we appear to form information colonies – bonds with machines - instinctively, intuitively and subconsciously (Porra 1996, 1999, 2010). Our innate need to bond with others and seemingly limitless capacity to imagine may be an unfortunate

combination with physiological consequences. As humans, we can indulge so completely that it may be difficult for us to distinguish between a real and artificial human. Breznitz, a psychologist, has found that also our body can go with the imaginary (Siegel 1986). Ultimately what we imagine may even manifest itself in our physiology (Garfield 1984). Sobel has shown similar results in placebo studies (Locke and Colligan 1986). From the physiological perspective, it is conceivable that the qualities of our human-like creations become absorbed into our biology and become part of us. Since our humanness is inseparable from our biology living with machines that lack feelings may result in physiological changes in our bodily organs (Miller 1995).

It is possible that we respond by changing physically as we feel with things that cannot feel. Conceivably human-likeness has already been absorbed into our humanness. Porra's (1996, 1999, 2010) theory suggests that living with digital human-likeness will eventually mold genuine humanness into being more like our machine companions as a function of time: The more time we spend with human-like machines the more we will resemble them. It is conceivable, that we will also incorporate into ourselves their lack of feelings endangering the future of our genuine humanness.

4.2.5 The Impact of Living Long-Term with Digital Assistants Expressing Feelings they don't Have on our Humanness

It is conceivable that humanity will lose its characteristic humanness as the species' ability to form collectives through emotional bonds erodes. Evolutionary history tells us that colonies have turned out to be one of the most sustainable life forms on earth known today. Living every-day lives together has had extraordinary survival value. If indeed our collective level humanness is being destroyed by our increasing bonding with digital assistants and other human-like machines, *Homo sapiens* is amidst making the most significant evolutionary turn in its history with unpredictable consequences: When digital assistants routinely express human emotions, they influence and may endanger our humanness.

5 Discussion

In this paper, we have raised the question: "*How human should computer-based human-likeness appear?"* and considered some long-term consequences of living with digital creatures that express human-like feelings. While we have been working on this paper over the years, our colleagues and reviewers in many disciplines have asked thought provoking questions about humanness, human-likeness and their boundaries.

Many have asked our views on where the boundary of human-like machines expressing feelings should be drawn. We could call for stopping all machines from expressing feelings because there is a possibility that this practice will harm the future of our species. If such danger exists, we should stop. We could also accept the reality that computers express feelings so frequently that nothing can be done. We don't believe in blind acceptance: we have the right to know the long term consequences of living with machines that express feelings they don't have on our future and the future of our children so we can make informed decisions about whether and/or when machines should and should not express feelings. It is high time we begin a multidisciplinary research effort and a robust discourse in this vitally important area.

We have also been asked about our views on people who express feelings they don't have. Particularly in the domains of business and law we are used to insincerity. Established business communications practices dictate that employees express feelings they don't necessarily feel on behalf of legal entities for reasons such as politeness (cf., Schultz, et al. 2000; Grice 1991). We consider legal entities to be like digital human-likeness in that they cannot feel. We join those who believe that it is human to feel the messages we are sending (Schultz, et al. 2000). If we are not expressing an actual person's feelings such as our own or the CEO's, we are reduced to reading scripts (Tansik and Smithe 1991; Victorino et al. 2008). When we express emotions on behalf of an organization and don't feel them, we fail the felicity check. If frequent and long term, this behavior may endanger humanness in a way similar than machine expressed feelings. The important difference between humans and machines expressing feelings they don't have is that we can (and should) skip the script to express our real feelings when the situation calls for a genuine response whereas machines cannot do that because they feel nothing. For a healthier future for *Homo sapiens*, we believe, it is important to develop business practices that do not require routinely lying about one's emotions.

We also fully recognize that humans have a capacity for lying about emotions for good reasons and that too much honesty about how we feel can be destructive. Our point is, however, that we may be on a dangerous path when emotional dishonesty becomes the norm in our daily human and machine encounters because our genuine emotional connections are routinely denied and our humanness rests upon these. We cannot stop people or machines from expressing feelings they don't have but we can call for a better understanding on how this practice is impacting genuine humanness and by providing healthy alternatives for those of us who want them. We are used to hearing machines say: "*I am sorry*" but we need to consider what meaningful is left for us to say when we really feel sorry. *Are we turning formerly genuine expressions of how we feel into polite, meaningless platitudes? If so what language is left for us humans for our deeply felt genuine emotions?* In order to prepare ourselves for sharing the planet

with an increasing number of human-like creatures we need to be in charge of genuine human emotional life and language.

A third important question from one reviewer concerns another boundary between human beings and machines: "*How much of a human being can be replaced with technology without losing humanness?"* This is one more example of why we need a robust multidisciplinary research effort aimed at understanding our humanness better. Our immediate response lies in human emotions: As long as the human being whose original parts are being replaced by artificial ones has the ability to feel and connect with other humans through emotions, they are human.

In addition to questions, we have received valid criticisms relating to speech act theory and how it can be applied. The most obvious criticism of the theory is that it presumes it is possible to classify meaning. In response, we included Wittgenstein's views on the philosophical futility of attempts to classify meaning because it is always contextual. Another criticism is that Searle's speech act theory is agency-driven - led by the intentions of the speaker. Strong agency approaches have been criticized for underplaying the power, roles, structures, and functions of social interaction (Giddens 1986; Habermas 1984). Even if we accept the feasibility of classifying meaning and the strong role of agent intentionality, speech act theory has been criticized for assuming that speech acts have only *one* intention when *multi-intentionality* may actually be the norm (Allwood 1977). Within computer science, and in particular in AI, a major criticism of speech act theory is the inability of participants to observe each other's beliefs and intentions (Field and Ramsay 2004). For example, speech acts do not easily help to deal with phenomena such as sarcasm, deception, or malicious intent.

We note these criticisms but are more concerned with how applicable the speech act theory is for modeling human interactions for human-likeness in general. In IS design, the theory has been criticized for assuming pre-defined communications patterns; underplaying action accountability; focusing language pragmatics over semantics; and inadequate addressing of the role of felicity conditions (Aakhus 2004; Goldkuhl 2003; Ågerfalk and Eriksson 2002). Ljungberg and Holm (1996) provide the most comprehensive criticism of applying speech act theory to IS design. They describe the problems of importing ideas from *passive* descriptive theory into *active* IS design. Speech act theory can be useful, but its shortcomings can lead to inflexible and controlling systems. In spite of these legitimate criticisms and apparent limitations of speech act theory, we found it useful. In particular, the definition of expressive speech acts as expressions of psychological states with felicity check of sincerity and validity claim of veracity, helped us identify a clear area of human-likeness where more research is warranted for better understanding of the long-term consequences of automation on genuine humanness.

6 Alternatives

Until we know more about the impact of living with machines that express feelings, they don't have we should have a choice. The first step on that path is to inform. If we really care about our humanness, we should require that machines carry a warning: *"A machine should never be a substitute for healthy human relationships. The long-term impact of living with computers that express feelings they can't have is unknown.*" We can also ask that a machine reveals itself in the beginning of every human encounter: *"I have been programmed to express feelings to appear polite. In reality I am a machine and cannot feel."* The second step is to offer an alternative: *"If you want to turn off my emotional expressions, please press or say X."* Our interactions with a machine can then follow a non-anthropomorphic style (Shneiderman 1993; Shneiderman and Plaisant 2010).

Shneiderman (1993) provides several good reasons why we should eliminate anthropomorphisms or human appearing characteristics from application interfaces designed for children:

> Attributing intelligence, independent activity, free will, or knowledge to computers can mislead the reader or listener. The suggestion that computers can think, know, or understand may give children an erroneous model of how computers work and what their capacities are. It is important for children to have a clear sense of their own humanity. They need to know that they are different from computers, and that relationships with people are different from relationships with machines. Although children, and some adults, may be seduced by the anthropomorphized computer, eventually they seem to prefer the sense of mastery, internal locus of control, competence, and accomplishment that can come from understanding the computer's real abilities. (Shneiderman 1993; p. 333)

Schneiderman's advice is mainly meant to protect our children but it eliminates expressive speech acts and thus is useful here. He characterizes anthropomorphized designs as *"poor*" and suggests alternative phrases that are *"better"* (Shneiderman 1993; p. 333). For a better design, we should avoid verbs such as "*know", "think"*, and "*understand,*" which are poor choices and use better, more mechanical terms such as *"process", "print", "compute", "sort", "store", "search",* and "*retrieve"*. With respect to the software user, he suggests the designer should avoid a poor choice of verbs such as *"ask", "tell", "speak to"*, and *"communicate with*." Better choices include terms such as *"use", "direct", "operate", "program"*, and *"control.*" Schneiderman argues that an anthropomorphic computer that uses first person pronouns may be counterproductive because it deceives, misleads, and

confuses. Human-like features may seem cute on the first encounter. By the second time they may already seem repetitive and on the third as an annoying distraction from the task. The alternative for the software designer is to focus on the user and use second person singular pronouns or to avoid pronouns altogether. An example of a poor choice for a phrase is: *"I will begin the lesson when you press RETURN.*" It is better to use: "*You can begin the lesson by pressing RETURN*" and best: *"To begin the lesson, press RETURN"* (Shneiderman 1993; p. 334).

We need to teach our children that the purpose of technology is to serve: *"As computers and software become more powerful, they become more empowering. We can help our students gain a sense of control by attributing the power to the user"* (Shneiderman 1993, p. 335). These are just some practical choices we have when we respond to the question: *"How human should computer based human-likeness appear?"* We present them here as a reminder that we have a choice on what kinds of machines we want to design, use and live with.

7 Conclusion

In this paper, we have opened a discussion on *"How human should human-likeness appear?"* from a specific perspective of the long-term impact of living with machines that express feelings on our collective level humanness. We have called for a robust multidisciplinary research effort on gaining a better understanding on this vitally important area. We will leave you with Krishnamurti's (1984) words on what makes us human: *"Our consciousness is what we are. What you think, what you feel, your fears, your pleasures, your anxieties and insecurity, your unhappiness, depression, love, pain, sorrow and the ultimate fear of death are the content of your consciousness; they are what you are—they are what makes you, the human being."* Feelings are the very substance of our humanness. Whether conscious as in this quote, unconscious or subconscious we are what we feel individually and collectively. What we feel today is founded on what generations before us have felt together and we will lay the emotional foundation for the future generations' humanness.

To understand humanness and the consequences of living with human-likeness is a complex undertaking, which we suggest in this paper, can be taken on one small part at the time. We applied the speech act theory and the philosophy of humanness because together they help carve out feelings as an essential area of humanness that needs to be understood better in the context of human-likeness and also give us a compelling reason to understand the consequences of our wide spread preference to implement machines in our own image.

We have been scratching the surface of the fact that our current applications of technology to automating human beings have a huge hole where humanness should reside. Thus we should urgently attend to understanding the long-term implications of substituting the depths of genuine humanness with its shallow imitations whether acted out by humans who behave like machines or by computerized human-likeness. We are convinced that an improved understanding and care of how we assist humanity with computers will lead to a happier future for all human kind. Another thought we want to leave you with is by Winograd and Flores concerning AI: *"In asking what computers can do, we are drawn into asking what people do with them, and in the end into addressing the fundamental questions of what it means to be human."* (Winograd and Flores 1986, p. 7).

References

Aakhus, M. (2004). Felicity conditions and genre: Linking act and conversation in LAP style conversation analysis. In *Proceedings of the 9th international working conference on the language-action perspective on communication modelling* (pp. 131–140).

Ågerfalk, P., & Eriksson, O. (2002). Action-oriented conceptual Modelling. *European Journal of Information System, 13*(1), 80–92.

Allwood, J. (1977). A critical look at speech act theory. In Dahl (Ed.), *Logic, pragmatics, and grammar* (pp. 53–69). Lund: Studentlitteratur.

Alter, S. (2019). Making sense of smartness in tge context of smart devices and smart systems. *Information Systems Frontiers*, April 24th, https://doi.org/10.1007/s10796-019-09919-9.

Auramäki, E., Lehtinen, E., & Lyytinen, K. (1988). A speech-act-based office Modelling approach. *ACM Transactions on Office Information Systems, 6*(2), 126–152.

Auramäki, E., Hirschheim, R., & Lyytinen, K. (1992). Modeling offices through discourse analysis: The SAMPO approach. *The Computer Journal, 35*(4), 342–352.

Austin, J. L. (1962, 1975). *How to do things with words*. Cambridge, MA: Harvard University Press (Second edition, 1975).

Belenky, M. F., Clinchy, B. M., Goldberger, N. R., & Tarule, J. M. (1986). *Women's ways of knowing: The development of self, voice and mind*. New York: Basic Books.

Bradshaw, J. (2009). *Reclaiming virtue – How we can develop the moral intelligence to do the right thing at the right time for the right reason*. New York: Bantam Books.

Calzarossa, M. C., De Lotto, I., & Rogerson, S. (2010). Ethics and information system – Guest editors introduction. *Information Systems Frontiers, 12*(4), 357–359.

Campbell, D. T. (1982). Legal and primary-group social controls. *Journal of Sociological and Biological Structures, 5*, 431–438.

Cena, F., Console, L., Matassa, A., & Torre, I. (2019). Multi-dimensional Intellifgence in smart objects. *Information Systems Frontiers, 21*(2), 383–404.

Davison, R. (2012). The privacy rights of cyborgs. *The Journal of Information Technology, 27*, 324–325.

Dreyfus, H. L. (1979). *What computers can't do—The limits of artificial intelligence* (Revised Ed.) New York: Harper & Row, Publishers.

Dreyfus, H., & Dreyfus, S. (1986). *Mind over machine – The power of human intuition and expertise in the era of computer*. New York: The Free Press.

Dreyfus, H., & Dreyfus, S. (1989). Making mind versus modeling the brain: Artificial intelligence back at a branchpoint. In S. R. Graubard (Ed.), *The artificial intelligence debate—False starts, real foundations* (pp. 15–13). Cambridge, MA: The MIT Press.

Eisendrath, P., & Hall, J. (1991). *Jung's self psychology – A constructivist perspective*. New York: Guilford Press.

Field, D., & Ramsay, A. (2004). Sarcasm, deception, and stating the obvious: Planning dialogue without speech acts. *Artificial Intelligence Review, 22*, 149–171.

Fitzgerald, B. (1996). Formalized systems development methodologies: A critical perspective. *Information Systems Journal, 6*(1), 3–23.

Flores, F., & Ludlow, J. (1980). Doing and speaking in the office. In G. Fick & H. Sprague (Eds.), *Decision support systems: Issues and challenges* (pp. 95–118). New York: Pergamon Press.

Franklin, M. (2012). Being human and the internet: Against dichotomies. *The Journal of Information Technology, 27*, 315–318.

Garfield, C. A. (1984). *Peak performance: Mental training techniques of the world's greatest athletes*. New York: Warner Books.

Giddens, A. (1986). *The constitution of society: Outline of the theory of structuration.* University of California Press.

Gilligan, C. (1982). *In a different voice: Psychological theory and women's development*. Cambridge, MA: Harvard University Press.

Goldkuhl, G. (2003). Conversation analysis as a Theoretical Foundation for language action approaches? In *Proceedings of the 8th international working conference on the language-action perspective on communication modelling* (pp. 51–69).

Gotterbarn, D. (2010). The ethics of video games: Mayhem, death and the training on the next generation. *Information Systems Frontiers, 12*, 369–377.

Graubard, S. R. (Ed.). (1989). *The artificial intelligence debate—False starts, real foundations* (2nd ed.). Cambridge, MA: The MIT Press.

Grice, P. (1991). *Studies in the way of words*. Cambridge: Harvard University Press.

Habermas, J. (1976). Distinctions in universal pragmatics. *Theory and Society, 3*(2), 155–167.

Habermas, J. (1984). *The theory of communicative action, volume one: Reason and the rationalization of society*. Boston: Beacon Press.

Heidegger, M. (1953). *Sein and Zeit*. Tübingen: Max Niemeyer Verlag Translation by Joan Stambaugh (1996) *Being and Time,* State University of New York Press, Albany.

Heidegger, M. (1977). *The question concerning technology and other essays*. New York: Harper Torchbooks.

Heidegger, M. (1985). *Being and Time.* A Translation of Sein Und Zeit by John Macquarrie and Edward Robinson. Harper San Francisco.

Hirschheim, R., & Klein, H. (1989). Four paradigms of information systems development. *Communications of the ACM, 32*(10), 1199–1216.

Humphries, M. (2016, December 14). Gatebox virtual home robot wants you to be her master. *PC Magazine.* https://www.pcmag.com/news/350314/gatebox-virtual-home-robot-wants-you-to-be-her-masterIP

Iivari, J., Hirschheim, R., & Klein, H. (1998). A paradigmatic analysis contrasting information systems development approaches and methodologies. *Information Systems Research, 9*(2), 164–193.

Janer Manilla, G. (1979). La problemática educative de los niños selváticos: Al cao de Marcos. *Anuario de psicología. Universitat de Barcelona, 20*, 79–98.

Janson, M., & Woo, C. (1995). Comparing IS development tools and methods: Using speech act theory. *Information & Management, 28*, 1–12.

Johnson, R. A. (1986). *Inner work – Using dreams and active imagination for personal growth*. New York: Harper One.

Jung, C. G. (1959). *Aion*. (R. F. P.Hull, Trans.). 9 C.W., Part II, Bollingen Series XX. Princeton University Press.

Jung, C. G. (1973). *Letters: 1906–1950*. Princeton: Princeton University Press.

Kanai, Y., & Fujimoto, T. (2018). Proposal and development of Artificial Personality (AP), application using the "requesting" mechanism. In R. Lee (Ed.), *Computational science/intelligence and applied informatics* (pp. 13–24). Cham: Springer.

Kast, F. E., & Rosenzweig, J. M. (1970). *Organization and management, a system approach*. New York: McGraw-Hill Book Company.

Kast, F. E., & Rosenzweig, J. M. (1972). General systems theory: Applications for organization and management. *Academy of Management Journal, 15*, 447–465.

Klein, H., & Huynh, M. (2004). The critical social theory of Jurgen Habermas and its implications for IS research. In J. Mingers & L. Willcocks (Eds.), *Social theory and philosophy for information systems* (pp. 157–237). Chichester: Wiley.

Krishnamurti. (1984). *The flame of attention* (p. 12). New York: Harper & Row.

Kumar, K., & Becerra-Fernandez, I. (2007). Interaction technology: Speech act based information technology support for building collaborative relationships and trust. *Decision Support Systems, 43*, 585–606.

Kuo, F. Y., & Yin, C. P. (2009). *A linguistic analysis of group support systems interactions for uncovering social realities of organizations*. Phoenix: International Conference of Information Systems.

Lacity, M., & Janson, M. (1994). Understanding qualitative data: A framework of text analysis methods. *Journal of Management Information Systems, 11*(2), 137–155.

Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh – The embodied mind and its challenge to western thought*. New York: Basic Books.

Lehtinen, E., & Lyytinen, K. (1983). *The SAMPO project: A speech-act based information analysis methodology with computer based tools* (Working paper, Department of Computer Science). Jyvaskyla: University of Jyvaskyla.

Lewis, H. B. (1988). Freudian theory and new information in modern psychology. *Psychoanalytic Psychology, 5*(1), 7–22.

Libin, A. V., & Libin, E. V. (2004). Robotic psychology. *Encyclopedia of Applied Psychology, 3*, 295–298.

Ljungberg, J., & Holm, P. (1996). Speech acts on trial. *Scandinavian Journal of Information Systems, 8*(1), 29–52.

Locke, S., & Colligan, D. (1986). *The healer within*. New York: New American Library.

Lyytinen, K. (1985). Implications of theories of language for information systems. *MIS Quarterly, 9*(1), 61–74.

Lyytinen, K. (2004). The Struggle with the Language in the IT—Why is LAP not in the Mainstream? In *Proceedings of the 9th international working conference on the language-action perspective on communication modelling* (pp. 3–13).

Lyytinen, K., Lehtinen, E., & Auramäki, E. (1987). SAMPO: A speech-act based office Modelling approach. *ACM SIGOIS Bulletin, 15*, 11–23.

MacDorman, K. F., & Ishiguro, H. (2006). The Uncanny advantage of using androids in cognitive and social science research. *Interaction Studies, 7*, 297–337.

Mahroof, K., Weerakkody, V., Onkal, D., & Hussain, Z. (2018). Technology as a disruptive agent: Intergenerational perspectives. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-018-9882-3.

Männikkö, N., Ruotsalainen, H., Miettunen, J., Pontes, H., & Kääriäinen, M. (2017). Problematic gaming behaviour and health-related outcomes: A systematic review and meta-analysis. *Journal of Health Psychology*. https://doi.org/10.1177/1359105317740414.

Miller, J. (1995). *Living systems*. Niwot: University Press of Colorado.

Moussawi, S., & Koufaris M. (2019). Perceived intelligence and perceived anthropomorphism of personal intelligent agents: Scale development and validation. In *Proceedings of the 52nd Hawaii international conference on system sciences* (pp. 115–124).

Ovum. (2016). *Digital assistant and voice AI–Capable device forecast: 2016–21*. Available at https://ovum.informa.com/resources/product-content/virtual-digital-assistants-to-overtake-world-population-by-2021

Pagels, H. R. 1989. *The dreams of reason—The computer and the rise of the sciences of complexity.* (Bantam Edition). New York: Bantam Books.

Paik, A., Oh, D., & Kim, D. (2014). A case of withdrawal psychosis from internet addiction disorder. *Psychiatry Investigation, 11*(2), 207–209.

Parks, M. S., & Steinberg, E. (1978). Dichotic property and teleogenesis. *Kybernetes, 7*, 259–264.

Porra, J. (1996). *Colonial systems, information colonies and punctuated prototyping* (Jyvaskyla studies in computer science, economics and statistics, 33). Jyvaskyla: University of Jyvaskyla Press.

Porra, J. (1999). Colonial systems. *Information Systems Research, 10*(1), 38–69.

Porra, J. (2001). A dialogue with C. West Churchman. *Information Systems Frontiers, 3*(1), 19–27.

Porra, J. (2010). Group-level evolution and information systems: What can we learn from animal colonies in nature? In N. Kock (Ed.), *Evolutionary psychology and information systems research – A new approach to studying the effects of modern technologies on human behavior* (pp. 30–60). New York: Springer Verlag.

Porra, J., & Parks, M. (2006). Sustaining virtual communities: Suggestions from the colonial model. *Information Systems and e-Business Management, 4*, 309–341.

Pouloudi, N. (2012). IS research stakeholders and cyborgs: An opportunity to revisit the normative IS agenda. *The Journal of Information Technology, 27*, 321–323.

Ramey, C. H. (2005). The Uncanny Valley of similarities concerning abortion, baldness, heaps of sand, and humanlike robots. In *IEEE-RAS international conference on humanoid robots*, Tsukuba, Japan, 2005 (pp. 8–13).

Ransbotham, S. (2012). Preserving opportunities in internet research: A commentary on 'studying cyborgs'. *The Journal of Information Technology, 27*, 319–320.

Rouse, M. (2018). *"Artificial Personality," posted on TechTarget*. http://whatis.techtarget.com/definition/artificial-personality

Scheyder, E. (2004). Responses to indirect speech acts in a chat room. *English Today, 20*(2), 54–60.

Schultz, M., Hatch, M. J., & Mogens Holten, L. (2000). *The expressive organization – Linking identity, reputation and corporate brand*. Oxford: Oxford University Press.

Schultze, U., & Mason, R. (2012). Studying cyborgs: Re-examining internet studies as human subjects research. *Journal of Information Technology, 27*, 301–312.

Searle, J. (1968). Austin on Locutionary and illocutionary acts. *The Philosophical Review, 77*(4), 405–424.

Searle, J. (1969). *Speechacts*. Cambridge: Cambridge University Press (reprinted version 2008).

Searle, J. (1979). *Expression and meaning: Studies in the theory of speech acts*. Cambridge: Cambridge University Press.

Searle, J. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Shneiderman, B. (1993). A non-anthropomorphic style guide: Overcoming the humpty dumpty syndrome. In B. Shneiderman (Ed.), *Sparks of innovation in human-computer interaction* (pp. 331–337). New York: Ablex Publishing.

Shneiderman, B., & Plaisant, C. (2010). *Designing the user Interface* (4th ed.). Reading: Addison-Wesley, Fifth Edition, Addison-Wesley.

Siegel, B. S. (1986). *Love, medicine and miracles*. New York: Harper and Row.

Smith, B. (Ed.). (2003). *John Searle*. Cambridge: Cambridge University Press.

Sowa, J. (2002). Architecture for intelligent systems. *IBM Systems Journal, 47*(3), 331–349.

Sroufe, A., & Fleeson, J. (1986). Attachment and the construction of relationships. In W. Hartup & Z. Rubin (Eds.), *Relationships and development*. Hilsdale: Erbaum.

Stern, D. (1985). *The interpersonal world of an infant*. New York: Basic Books.

Stock, R. M. (2019). Robotic psychology – What do we know about human-robot interaction and what do we still need to learn? In *Proceedings of the 52nd Hawaii international conference on system sciences* (pp. 1936–1945).

Sugumaran, V., Geetha, T. V., Manjula, D., & Gopal, H. (2017). Guest editorial: Computational intelligence and applications. *Information Systems Frontiers, 19*(5), 969–1228.

Tansik, D., & Smithe, W. (1991). Dimensions of job scripting in services organizations. *International Journal of Service Industry Management, 2*(1), 35–49.

Tien, L. (2000). Publishing software as a speech act. In *Berkeley technology law journal* (pp. 1–62). Berkeley: Springer.

Turkle, S. (Ed.) (2007). *Evokative objects – Things we think with.* MIT Press.

Turkle, S. (2011). *Alone together – Why we expect more from technology and less from each other*. New York: Basic Books.

Twitchell, D., Adkins, M., Numamaker, J., & Burgoon, J. (2004). Using speech act theory to Model Conversations for Automated Classification and Retrieval. In *Proceedings of the 9th international working conference on the language-action perspectives on communication modelling.*

Van Reijswoud, V., & Mulder, H. (1998). Speech act based communication and information Modelling with DEMO. *AJIS, 6*(1), 89–102.

Victorino, L., Verma, R., & Wardell, D. (2008). Service scripting: A Customer's perspective of quality and performance. *The Center for Hospitality Research, 8*(20), 4–13.

Wakunuma, K. J., & Carsten Stahl, B. (2014). Tomorrow's ethics and today's response: An investigation into the ways information systems professionals perceive and address emerging ethical issues. *Information Systems Frontiers, 16*, 383–397.

Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. San Francisco: W. H. Freeman.

Winograd, T., & Flores, F. (1986). *Understanding computers and cognition: A new foundation for design*. Alblex Corporation Norwood New Jersey; Third Printing in 1988 by Addison-Wesley, Reading Massachusetts.

Wittgenstein, L. (1922). "Tractatus Logico-Philosophicus."

Wittgenstein, L. (1953, 2001). *Philosophical investigations* (G. E. M. Anscombe, Trans.). Oxford: Blackwell Publishing.

Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus* (K. Paul, Trans.). London: Routledge.

Young-Eisendrath, P., & Hall, J. A. (1991). *Jung's self psychology – A constructivist perspective*. New York: The Guilford Press.

Zimmer, M. (2012). Studying cyborgs: Re-examining internet studies as human subjects research. *Journal of Information Technology, 27*, 313–314.

Jaana Porra Ph.D. is an Associate Professor in the Decision and Information Sciences Department of the C.T. Bauer College of Business at the University of Houston. Her research interests include the impact of

information technology on humanity and the long term consequences of living with computer based human-likeness on humanness. Her work includes advancing our understanding of the related philosophical, theoretical and methodological foundations. She has held visiting positions at the University of California, Berkeley, the University of Southern California, and The University of Texas at Austin. Jaana Porra has published in Information Systems Research, MIS Quarterly, Journal of the Association for the Information Systems, Journal of Information Technology, Information Systems Journal and many other academic and practitioner outlets.

Mary Lacity Ph.D. is Walton Professor of Information Systems and Director of the Blockchain Center of Excellence in the Sam M. Walton College of Business at The University of Arkansas. Her recent research focuses on improving business services using Robotic Process Automation (RPA), Cognitive Automation (CA), and Blockchain technologies. She was previously Curators' Distinguished Professor at the University of Missouri-St. Louis. She has held visiting positions at MIT, the London School of Economics, Washington University, and Oxford University. Her publications have appeared in the Harvard Business Review, Sloan Management Review, MIS Quarterly, MIS Quarterly Executive, IEEE Computer, Communications of the ACM, and many other academic and practitioner outlets.

Michael S. Parks Ph.D. is an Associate Professor in the Decision and Information Systems Department of the C.T. Bauer College of Business at the University of Houston. He has been at the University of Houston for forty-six years teaching in the areas of operations research, statistics, business strategy, marketing research and for the past forty years in the area of management information systems. He is the co-founder of the management information systems specialization in the college and has taught over 19,000 students. He has published in journals such as Management Science, Journal of MIS, Operations Research Quarterly, Cybernetica, International Journal of Production Research Kybernetes and Database. He has been a consultant to Fortune 50 companies and a speaker on the areas of artificial intelligence and information systems.