

# Ethics for AI Writing

## The Importance of Rhetorical Context

Heidi A. McKee<sup>†</sup>

English  
Miami University  
Oxford, OH USA  
mckeeha@miamioh.edu

James E. Porter

English / Emerging Technology in Business & Design  
Miami University  
Oxford, OH USA  
porterje@miamioh.edu

### ABSTRACT

Implicit in any rhetorical interaction—between humans or between humans and machines—are ethical codes that shape the rhetorical context, the social situation in which communication happens and also the engine that drives communicative interaction. Such implicit codes are usually invisible to AI writing systems because the social factors shaping communication (the why and how of language, not the what) are not usually explicitly evident in databases the systems use to produce discourse. Can AI writing systems learn to learn rhetorical context, particularly the implicit codes for communication ethics? We see evidence that some systems do address issues of rhetorical context, at least in rudimentary ways. But we critique the information transfer communication model supporting many AI writing systems, arguing for a social context model that accounts for rhetorical context—what is, in a sense, “not there” in the data corpus but that is critical for the production of meaningful, significant, and ethical communication. We offer two ethical principles to guide design of AI writing systems: transparency about machine presence and critical data awareness, a methodological reflexivity about rhetorical context and omissions in the data that need to be provided by a human agent or accounted for in machine learning.

### CCS CONCEPTS

- Computing methodologies-natural language generation
- Computing methodologies-intelligent agents • Human-centered computing-interaction design theory, concepts and paradigms

### KEYWORDS

AI writing systems; rhetoric; ethics; communication ethics; rhetorical context; text generation; language models; social context model; information transfer model; transparency; machine ethics; critical data awareness

<sup>†</sup> Corresponding author: Heidi A. McKee, mckeeha@miamioh.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).  
AIES '20, February 7–8, 2020, New York, NY, USA.  
© 2020 Copyright is held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-1-4503-7110-0/20/02...\$15.00.  
DOI: <https://doi.org/10.1145/3375627.3375811>

### ACM Reference format:

Heidi A. McKee and James E. Porter. 2020. Ethics for AI Writing: The Importance of Rhetorical Context. In *Proceedings of 2020 AAAI/ACM Conference on AI, Ethics, and Society (AIES'20), February 7–8, 2020, New York, NY, USA*. ACM, New York, NY, USA, 7 pages. DOI: <https://doi.org/10.1145/3375627.3375811>

### 1 Introduction

AI is transforming communication, including written communication. AI writing agents transcribe meetings and produce minutes (Voicea's Eva), write emails to set up appointments (x.ai's Amy), and communicate via text chat with customers (customer service bots too numerous to name). AI systems publish news stories (the *Washington Post's* Heliograf), create financial reports (Narrative Science's Quill), produce marketing copy (Persado), and (co)write our emails (Google Compose). AI agents have long been used for sentiment analysis in customer service and corporate social media, producing reports and making recommendations for action (Amazon's Connect). Quite simply we are immersed in AI professional communications on a daily basis [1].

In this paper we focus on the ethics of AI writing systems (sometimes called text generation systems or data-to-text systems), focusing in particular on the rhetorical contexts for producing and receiving written communication. Our discussion builds upon the extensive research on machine ethics [2, 3, 4, 5, 6, 7, 8, 9], but we approach AI from the standpoint of rhetoric and communication theory, and especially communication ethics [10]. We raise the important question, What do humans need to know and what do machines need to know to write to and for each other—and, importantly, what can't they know? Communication is often complex and always influenced by rhetorical context, including the sociocultural, historical, ideological, attitudinal, and ethical codes that shape how communication happens.

While humans have a distinct edge in the layered, nuanced complexities of communication, AI writing systems certainly have the edge on processing huge volumes of data. But even with seemingly unlimited data points, many AI writing systems are built on an information transfer model of communication that assumes text production is a simple matter of converting raw data into sentences and paragraphs. This model generally obscures the critical role of audience and context and excludes ethics as an

element of textual production. In this paper we offer as an alternative a social context model of communication that takes account of rhetorical context—the stuff that is “not there” (typically) in the available data sets but that is vital for the production of meaningful, significant, and ethical communication.

## 2 Rhetorical Context and Ethics

We approach the issue of AI writing from the standpoint of our field—rhetoric and professional communication. As scholars in professional communication, we study writing practices in the professional and public realms, focusing particularly on how digital networks change communication production and reception. And we do so from the standpoint of rhetoric theory. Rhetorical analysis, as a humanistic methodology, differs from linguistic study of language in several important respects, but most importantly in its focus on *context* rather than *text* (i.e., on the situational elements that comprise the communication setting) and also on inventional methods for generating content.

Rhetoric is the art and study of effective communication. Rhetoric is not, as the popular media assumes, lying, skewing the facts, political manipulation for political gain, or gaslighting. Actually just the opposite: a fundamental tenet of rhetoric is that ethics is an integral part of rhetorical interaction—from design and development to delivery and response. Aristotle [11] established *ethos* as one of the key persuasive appeals for rhetoric: the *ethos* of the speaker depended on the speaker being virtuous (*arête*), as well as having practical wisdom (*phronesis*) and good will toward audience (*eunoia*). Quintilian’s [12] definition of rhetoric as “the art of speaking well” puts ethics at the forefront of rhetoric, seeing the very definition of rhetoric as tied to ethics in two senses: (1) rhetoric depends on the virtue and ethical qualities of the speaker, and (2) rhetoric must ultimately serve the overall good of the *polis*, community, or state. Rhetoric deals fundamentally with questions of purpose and audience—why are we communicating in the first place? for whose ultimate benefit? And it takes up the key question of ultimate purpose as well: How does this communication serve the welfare of the *polis*, the community, the society?

As Markham [13] pointed out, ethical issues are not always high stakes, life-and-death trolley problems; microethical questions arise numerous times daily in our interactions with others, and we build our character through the way we negotiate those everyday small communication interactions (e.g., do we say please and thank you, or not? do we say good morning?). A rhetorical lens on communication ethics emphasizes the interrelated and inseparable components of rhetor/speaker, audience, purpose, message, and context, all forming the rhetorical situation.

Rhetorical context—including shared ethical understandings and, importantly, the quality of trust—is a kind of invisible glue that brings meaning, significance, coherence, and value to discourse. Microsoft’s Twitterbot Tay is a great example of a writing system that went out into the world without adequate contextual knowledge of the rough-and-tumble context of Twitter—and, particularly, of what constitutes racism, sexism,

homophobia, and anti-semitism. She was a naïf, with immature ethical standards and limited rhetorical intelligence, lacking skills and safeguards for how to communicate ethically in social media. In short, she was not rhetorically savvy and thus was not an effective communicator, at least in the ways her creators intended.

When machine talks to machine it is perhaps easier for effective communication to occur because the unpredictability and complexity of rhetorical context can be more efficiently managed. But with humans talking to humans or humans talking with machines that unpredictability and complexity cannot be so easily programmed in, requiring more flexibility for both human and machine about communication patterns and processes. The ethics of human-machine writing requires of both humans and machines a deeper understanding of context and a commitment to being a good human, a good machine, and a good human-machine speaking well together.

## 3 Two Rhetorical Models

The model of communication we see shaping many, not all, text generation systems is some variation on the Shannon-Weaver model [14, 15]—also known as the *information transfer model*—which views communication primarily as the one-way transmission of information, encapsulated in a message, from the knowledgeable transmitter (encoder) to the uninformed receiver (decoder or audience). The epistemological assumption of this model is that knowledge and meaning are fully contained in the data corpus, that the knowledge can be encapsulated in a verbal message, and then the message can be delivered to a largely uninformed and/or ignorant audience. Meaning is in the words.

This model has been often and soundly critiqued for being simplistically reductive and one way, for failing to acknowledge the contributions of the audience to knowledge construction, and for underestimating the role of rhetorical context. However, as Lakoff and Johnson [16] showed, this one-way pipeline model is deeply embedded into everyday metaphors and ways of talking about communication and so has incredible power and pervasive influence over our thinking.

We see this model underlying many discussions of AI writing systems. For example, the Talk to Transformer text generator takes a short piece of text inputted and, using GPT-2, an OpenAI generative language model, creates an entire written article from that piece of text, using a predictive model that creates new text based on the preceding text. The OpenAI model was trained on a dataset of 8 million web pages built from what the developers described as “outbound links from Reddit which received at least 3 karma”—meaning that had achieved some level of human approval and curation within the Reddit platform. In this language generator, the longer and more detailed the human prompt at the beginning, the more likely the system is to generate a coherent text based on the topic. What the system achieves, according to the developers, is a text that “feels close to human quality and shows coherence over a page or more of text” because, they argue, their dataset draws across a wide variety of text-types, rather than using just one type, like only fiction books [17, 18].

From a communication standpoint this OpenAI generative language model of text generation is badly flawed—first, because of its core assumption that new text arises out of the combination of topic prompt and existing database of text. While certainly sequential order is important in communicating meaning, it’s not actually the most essential. What’s missing is the crucial element shaping any communications, what would be the starting point for a social or rhetorical model: the rhetorical context, including audience, exigence, purpose for communication, and ethical understandings of the interactions. Why are we communicating in the first place? What are we hoping to achieve? This OpenAI model’s approach to coherence is badly flawed because it is based on a formalist notion of textual coherence (Does one piece of text follow topically from another?) versus a notion of coherence that involves audience (Does the text make sense to its intended readers in the contexts of interaction?).

The assumption of the linear model (Figure 1)—a model forming the basis for several AI-text generation systems [19, 20]—is that the AI system assembles data and, using NLG (natural language generation), produces a story or article, which is then delivered unproblematically to readers.

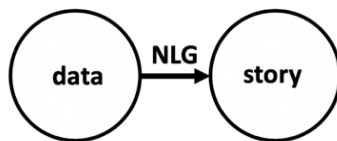


Figure 1. Information transfer model for AI writing

Think about any conversation you’ve had recently—whether written or oral—and then think of the amount of context knowledge involved. HCI researchers from early on recognized the challenge context poses for effective human machine interaction. A brief conversational example from the sociolinguist John Gumperz [21], an exchange between two secretaries in an office, illustrates this challenge:

- A: Are you going to be here for ten minutes?
- B: Go ahead and take your break. Take longer if you want.
- A: I’ll just be outside on the porch. Call me if you need me.
- B: OK, don’t worry.

In this exchange coherence and meaning are not purely a function of data or text only. Coherence, significance, meaning arise from the relationship between the interlocutors and their understandings of the message and its purpose. This exchange works because the two communication agents both possess the missing information, the context, the prior knowledge, the necessary presuppositions to link the pieces of the text into a coherent message. The exigence that drives this particular conversation—that is, the desire to make sure that the office has secretarial coverage—is not articulated explicitly. Nor does it have to be. The exigence, though implicit, is clearly understood by both parties in the conversation. The information transfer model emphasizes the data-to-text process, but what about all that other stuff?

One alternative to the information transfer model is a *social context model*, a model that frames interaction in a very different way, recognizing that meaning making starts with the community, not with the individual. A social context model (Figure 2) forefronts the co-creation of meaning by rhetor/writer and audience and the context shaping their interaction. Context includes both immediate context of the individuals involved (machine and human, and the purpose that brings them together), but also the larger social and cultural contexts that have an effect. That is, a meaningful discursive event does not start with topical data; it starts with rhetorical context or situation [22]: exigence, purpose, audience, and the ethical understandings that shape that assemblage.

What is rhetorical exigence? It is “an imperfection marked by urgency; it is a defect, an obstacle, something waiting to be done, a thing which is other than it should be” [22, p. 66]. Rhetorical action arises from some perceived need, an occasion calling for something to be done. That “something to be done” is the *raison d’être* for rhetorical action. Rhetoric strives to move us, through the deployment of symbolic action, from some current actuality to some new potentiality—the possible, *to dynaton* in sophistic rhetoric [23], which in large part involves moving an audience. That the “something to be done” action/outcome is desirable, valuable, beneficial (to someone), and good is the ethical component of rhetorical context.

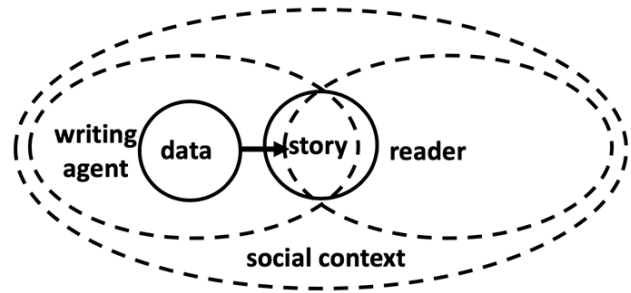


Figure 2. Social context model for AI writing

Our point about context is rather a simple one: To be an effective writer—whether human or AI—you have to understand the rhetorical context for your communication, and that context is not always clearly visible or available in the data set being used. This essential knowledge is not “there” in the data; it exists “out there” in the social realm. Numerous AI researchers have talked about this issue with AI systems, though using different terms. Harnad [24] discussed the symbol grounding problem. Floridi [25] referred to the importance of semantic capital, i.e., the background knowledge, the underlying frame that we acquire that enables us to make meaning and find significance in our world. Chowdhury [26] talked about the “pitfalls” in AI systems that cannot read important environmental factors that are vital to effective system functioning [27, 28, 29, 30].

As regards purpose, why we are communicating in the first place, we have to acknowledge that the writer’s purpose is sometimes—always?—different from the reader’s purpose. Certainly AI systems could develop strategies for accessing this

realm and, through deep learning, for understanding the variability of audiences and purposes. (Effective human writers learn this.) But our point is that many AI systems are being developed with linear communication models (i.e., if we can just get the input right, output will be fine) without sufficient respect for the importance and complexity of context. When humans and AI systems interact, miscommunication occurs and ethical issues arise from lack of understanding about context. We now turn to discuss a few brief examples.

#### 4 Ethic of Transparency: The Zingle Case

In 2014, several hotels in the Marriott franchise contracted with the company Zingle to use a valet parking service that included an automated text messaging system for car retrieval. We first learned of Zingle's service when Heidi stayed at a Marriot and the parking valet handed her a claim check and said, "Text your claim check number to us, and we'll get your car for you." The valet made no mention of an automated system. The next day, when retrieving her car, Heidi had the following text exchange with the AI system (though not knowing it was an AI system):

Heidi: Please retrieve our car 0359-928. We will be at front door at 2:10. Thx.

Zingle: Could not retrieve your claim check at this time. Please check the claim check number and try again. If you continue to get this text, please call 216-791-5678.

Heidi: 0359-928 Honda Accord gray Ohio plate.

Zingle: Thank you. We have received your request for [ticket, *brackets original*]. Your vehicle will be ready in 10 minutes. Regards, Courtyard Marriott Valet.

Heidi went to the front of the hotel and, after waiting 30 minutes and no car, contacted the manager, who explained that the system was automated: the system did not register her request because it only reads claim check numbers, nothing else. In order for the process to work, hotel guests must text *only* the claim check number.

In terms of troubleshooting this context-based miscommunication, there is obviously a lot that could have been done differently. First, the valet should have been explicit that the system was automated and that a human would not be reading the text. Second, the Zingle developers could have designed the system to read claim check numbers no matter if words were also included. And they could have also set a reply so that if the system could not read a message it would provide more helpful clarifying advice: e.g., "Please send only the claim check number. No words."

This short example shows the complexity of even a simple communicative interaction. With Zingle's Valet Text Service, a bounded AI system, even that single-purpose, single task-defined system opens up considerable possibility for error and misunderstanding that impacts the entire network of communication. Bottom line for ethics: Human-AI interaction requires an *ethic of transparency*: humans must know the rhetorical context and if they are interacting with an AI agent—whether in mobile text, social media, or other communications. If

a person thinks they are talking with a human that brings different understandings of the rhetorical context and (micro)ethical understandings for interaction.

#### 5 Ethic of Critical Data Awareness

We next turn to look at an AI system that works with data input to produce news stories. Narrative Science's AI agent Quill powers the GameChanger app, which writes recaps for youth baseball, softball and other sports. GameChanger takes the box score comprised of data entered by the coach or manager and converts it into a narrative of the game, producing the story immediately upon conclusion of the game. For example (names changed):

##### Four RBI Day for Jason G leads Miami LL Rockies Minors Past Cubs

Jason G was an RBI machine on Friday, driving in four on two hits to lead Miami LL Rockies Minors past Cubs 11-2 on Friday. Jason drove in runs on a home run in the first, a groundout in the fourth, and a single in the sixth.

Stories of games, particularly of kids' little league baseball games, circulate in complicated rhetorical situations. How does GameChanger deal with the issue of rhetorical context, particularly with questions of audience and ethics?

We have deduced, based on years of experience as parents reading GameChanger stories, that the AI writer does use some basic rhetorical and ethical principles. (1) It tells the story of the game from the point of the view of your team, whichever team your child is on. This is not unlike how local sports writers write up games for hometown fans. (2) It never says anything bad about any player; there is no criticism, even if a player commits a crucial error in the field or strikes out at a key moment. This is very unlike what a human sportswriter would say, at least about a professional game.

GameChanger operates in a well-defined, clearly bounded rhetorical space: youth sporting event, with kids playing and families and friends watching; families and friends are the intended readers for any game writeup. GameChanger recognizes that no parent wants to read anything negative about their child. But in following a fairly simple formula, GameChanger does, of course, miss a great deal of the rhetorical storyline. A box score reports the facts of the game in a very efficient format, providing a lot of information. But what does it miss?

The *stats* do not capture the *story* or the *drama* of the game (e.g., celebrating someone's first hit of the year; noting the game was delayed by 15 minutes due to a loose dog on the field; a key player was injured and had to leave the game). GameChanger renders the box score in narrative fashion (sentences, paragraphs), but the narrative is sorely limited by box score facts: Player A had two hits in the game, Pitcher B gave up only one run in three innings. But what if Player C's error *was* the deciding factor in the game, the difference between winning and losing? GameChanger does not allow itself to say that—and in that respect its writeup, as currently programmed for options allowed, makes a clear ethical choice: being positive vs telling a fuller story. What if the third out of the



inning—which the box score renders simply as FO9—was an incredible diving catch by a right fielder who had not made a catch all year? That is an important event that readers/fans would want to see acknowledged. GameChanger misses the drama, misses the ebb and flow, misses the critical incidents (to borrow a term from usability studies); it renders no judgment overall or a more contextualized story about the meaning of the game, as would a human sportswriter.

However, a story from GameChanger is better than no story at all and not many sportswriters are lining up to write pieces on 10-year-olds' games. So in that respect AI communication agents extend the context of an event, serving a kind of ethic of expanded circulation. This ethic may not seem that significant when applied to a kids' ball game, but applied to other sectors, it increases in importance and ethical complexity. AI apps such as the *Washington Post's* Heliograf often report on news stories that otherwise might not be written up—results of local elections, for example. So certainly having some information available that otherwise not be made available is important. But if machine reporting becomes the only form of reporting there are dangers of losing representations of the rhetorical context in more detail and complexity.

Let's take a look at a much higher stakes text generation app: the *Los Angeles Times's* Quakebot. Quakebot takes earthquake notices from the U.S. Geological Survey and then, if the earthquake is severe enough, generates a draft article, which is then delivered to the news room for editorial approval (human approval, that is). If the event is verified, the article is released [31]. The articles are short and simple statements of fact (including a Google Map): an earthquake happened, of a certain magnitude, at a given time, with a particular epicenter. Quakebot reports the facts instantly, in a linear way delivering news from the U.S. Geological Survey right to readers of the *Los Angeles Times*, following the one-way information transfer model. In one respect this is not a problem: this is important information which the public needs to know, and quickly. The issue lies more with what QuakeBot does not do: it does not assess the quake and tell people how to respond to it. The key question readers will have in response to the Quakebot article is, *What should I do? Should I evacuate?* And of course the answer to that question will vary depending on individual reader location. The information reported is important, but by itself inadequate for the reader's purpose and well-being. Here is where human judgment is necessary to determine a recommended course of action. Some communication processes cannot happen by machine alone.

But some increasingly can—or at least are. Persado is an AI-based content strategy tool that writes marketing copy. Where Persado differs from some of the AI agents previously discussed is that it performs a kind of audience analysis, drawing from extensive data that allows it to “create emotional profiles for every customer segment based on their responses to previous campaigns” [32]. Then, using its “AI-powered knowledge base of more than 1 million tagged and scored words, phrases, and images in 25 languages,” it produces “personalized marketing messages” for distinct customer segments [33]. The company claims that with

its data-driven approach to developing creative marketing that Persado can “disrupt the limitations of human bias ... and communicate with customers in a way that is proven to resonate” [33]. Or, as Kristin Lemkau, Chief Marketing Officer of JPMorgan Chase, stated in an announcement about Chase's new 5-year contract with Persado: “Machine learning is the path to more humanity in marketing . . . [Persado] rewrote copy and headlines that a human marketer, using subjective judgement and their experience likely wouldn't have. And they worked” [34].

“More humanity” with AI writing? What we see in these product claims for Persado is a certain theory of how rhetoric works, an example of the information transfer model at work—but a version of the model that does take audience into account, albeit a reductive view of audience. Implicit in the discussion is the absolute trust in data as more scientific, more reliable than human judgment. This trust may at times be justified if we clearly know where the data comes from and how it is collected. But, as Markham [35] cautioned, we have to be critical about our understanding of data: What if the data is wrong or incomplete or, as is so often the case, biased for or against particular groups? What if the data we are collecting is the wrong data for the problem we are addressing? There is also an issue of how the data is used to categorize and rate humans and thus shape the communications marketed to humans. Ethics lies in methods as much as in machines [36].

Whether someone chooses to click an ad or not is not, necessarily, high stakes, but whether or not someone gets employed is. AI screening bots are big business. One top program used is HireVue, a video AI hiring system that uses facial recognition software while recording short online interviews with job candidates and then generates a written report card for the employer assessing the employability of job candidates. HireVue does not make its algorithm available for review other than to say that “Facial Action Units ... make up 29% of a person's score; the words they say and the ‘audio features’ of their voice, like their tone, make up the rest.” The candidate's score is benchmarked against others' scores who have been determined to be “successful employees.” One argument that HireVue makes for the validity of its approach is that it is “still more objective than the flawed metrics used by human recruiters” [37].

What is not clear about the HireVue system is whether and how it might be discriminating against certain candidates: e.g., non-English speaking candidates, nervous candidates, candidates with disabilities, candidates who are lower key and not sufficiently expressive or enthusiastic, etc. What seems clear to us, though, is that it is yet another example of a reductive audience analysis. That is, it renders a judgment based on a limited data set emphasizing facial expressiveness and speaking ability (versus other factors such as, say, technical background, previous work experience, professional qualifications, intelligence). Its bias in rendering a hiring judgment might not be any worse than a human recruiter's, but what makes it worse is its claim to objectivity.

What we see, overall, in AI development is a reductive sense of the complexity of the rhetorical context—starting with, and maybe mostly dependent on, a naive trust in data, especially big data, as the answer to all rhetorical challenges. From our standpoint there is not enough critical reflection on what data is (and is not), not enough respect for the presence of malicious agents who skew data, and insufficient appreciation for the problem of limited and incomplete data and of the complexity of human and human-machine communication. With the Persado and HireVue apps, for instance, we see the methodological error of assuming that the limited data collected is sufficient to provide a holistic view of the problem or topic or audience being analyzed, when in fact the assessment, from a rhetorical standpoint, is incomplete and therefore badly skewed. How does the system provide a corrective?

When dealing with well-defined and not high-stakes rhetorical contexts and genres, AI machines can perhaps perform adequately. We can see AI writing agents as existing along a continuum of contextual fluency. At one end of the continuum is the financial report function of Narrative Science’s Quill, which takes quarterly earnings reports (required in the U.S., for example, of all publicly traded companies) and produces a cut-and-dry genre, one with few variations. In a sense, a quarterly report is simply putting words to a financial box score. When such reports used to be written by humans, often someone new to the task would be told to simply “Take last quarter’s report and change the numbers.” A quarterly earnings report requires little contextual knowledge; it is proving to be a genre that AI writing systems handle well. Now the story that might be shaping those financial numbers is far more complicated, and to tell that story might require much greater knowledge of rhetorical context. That is where other genres and human writers or human oversight of machine writers have a significant role to play.

Basically when there is a more nuanced story to tell or you get humans in the mix, communication gets more complicated. In our interview with Dennis Mortensen, CEO and founder of x.ai and creator of Amy, the AI-based scheduling assistant, he discussed the complexity of rhetorical context in something as seemingly simple as setting up a meeting: “What you figure out quickly is that humans are just crazy. Even when they speak to time, sometimes it doesn’t even look like time. They’ll say things like, ‘Let’s meet up later.’ Later? What does that mean?” [1]. Preparing Amy for market took years with 65+ trainers because, as Mortensen explained to us, “We need[ed] to train on that ambiguity. Amy needs to exist in your [human] universe.”

We live in a complicated communicative universe where meaning is not just carried in the words but also in the rhetorical context shaping the words—and for this reason we need to re-envision the models used for preparing AI-based writing and communication systems.

## 6 Conclusion

AI writing agents sometimes succeed (e.g., customer service chatbots responding to set questions) and they sometimes fail (e.g.,

Tay). The failure/success question seems dependent on whether the rhetorical context is well defined and bounded (i.e., clear audience needs, predictable interaction scripts) versus an open-ended situation with multiple audiences, competing needs, unclear expectations, problematic or debatable context (e.g., political exchange). and/or incomplete data.

For the next several years, perhaps decades, we will be in a transition phase where AI writing systems will autonomously produce simpler communications (e.g., customer service interaction, news writing) but serve as collaborators with humans on more advanced communications. Future-of-work commentators believe that we are entering an age of human-machine *symbiosis* [38, 39] requiring *collaborative intelligence* [40, 41]—that is, a transitional period where humans and machines will work closely together as partners in the creation of written content and meaning.

But maybe *collaboration*, *partnership*, *symbiosis* is the wrong descriptive frame. Perhaps the human-machine effort will involve more delegation of tasks in a division-of-labor model: *delegation*, rather than collaboration, may be the more accurate way to frame this [42]. Humans will need to understand “how the unique strengths of humans and AI can act synergistically” [43, p. 579], and where the human participant can provide the most value. The operative metaphor for this relationship might well be the *centaur*—half person, half horse—rather than the cyborg [44]. The machine/body provides power and speed, the human/head provides direction, purpose, and, most importantly, ethical guidance.

AI writing systems will likely need help with rhetorical context, which includes understanding the immediate communication context (why are we communicating? who is the audience? what do they know? what do they need?) and the broader cultural, social, and historical context that includes matters of ethics. The machine will need assistance with higher order cognitive skills such as synthesis of data and decision making involving matters of uncertainty and ambiguity, with competing perspectives and interests and, even, at times, competing versions of data. Because both human and machine will have to learn how to talk to the other—especially how to ask meaningful rhetorical questions—humans will need to know when they are talking to machines, because that requires a specific form of communication different from talking with other humans.

Any AI writing system needs to be able to learn about rhetorical context—what is “not there” in the data—or else they will miss an essential component of communication. But will that contextual information ever be entirely visible or accessible to the AI writer? Will the AI writer ever be able to learn the prior knowledge required to set up a lunch date? Can an AI writer learn to recognize nuanced forms of prejudice? Until that time happens, humans will continue to have a vital role in communication: helping machines meet the demands of context, the needs of diverse audiences, and the standards for ethical communication.

## REFERENCES

- [1] Heidi A. McKee and James E. Porter. 2017. AI agents as professional communicators. In *Professional Communication and Network Interaction: A Rhetorical and Ethical Approach*. Routledge, New York, NY.
- [2] Luciano Floridi and J.W. Sanders. 2004. On the morality of artificial agents. *Minds & Machine*, 14, 3 (Aug. 2004), 349-379. DOI: <https://doi.org/10.1023/B:MIND.0000035461.63578.9d>
- [3] Alison Adam. 2008. Ethics for things. *Ethics and Information Technology*, 10, 2-3 (Sept. 2008), 149–154. DOI: <https://doi.org/10.1007/s10676-008-9169-3>
- [4] Amitai Etzioni and Oren Etzioni. 2016. AI assisted ethics. *Ethics and Information Technology*, 18, 2 (June 2016), 149–156. DOI: <https://doi.org/10.1007/s10676-016-9400-6>
- [5] Bertram F. Malle. 2016. Integrating robot ethics and machine morality: The study and design of moral competence in robots. *Ethics and Information Technology*, 18, 4 (Dec. 2016), 243-256. DOI: <https://doi.org/10.1007/s10676-015-9367-8>
- [6] Virginia Dignum. 2018. Ethics in artificial intelligence: Introduction to the special issue. *Ethics and Information Technology*, 20, 1 (March 2018), 1-3. DOI: <https://doi.org/10.1007/s10676-018-9450-z>
- [7] Fabio Fossa. 2018. Artificial moral agents: Moral mentors or sensible tools? *Ethics and Information Technology*, 20, 2 (June 2018), 115-126. DOI: <https://doi.org/10.1007/s10676-018-9451-y>
- [8] Jaana Leikas, Raija Koivisto, and Nadezhda Gotcheva. 2019. Ethical framework for designing autonomous intelligent systems. *Journal of Open Innovation*, 5, 1 (March 2019), 18. <https://doi.org/10.3390/joitmc5010018>
- [9] Aimee van Wynsberghe and Scott Robbins. 2019. Critiquing the reasons for making artificial moral agents. *Science and Engineering Ethics*, 25, 3 (June 2019), 719-735. DOI: <https://doi.org/10.1007/s11948-018-0030-8>
- [10] Heidi A. McKee and James E. Porter. 2018. Digital media ethics and rhetoric. In *The Routledge Handbook of Digital Writing and Rhetoric*, ed. J. Alexander and J. Rhodes, 401-411. Routledge, New York, NY.
- [11] Aristotle. 2006. *On Rhetoric: A Theory of Civic Discourse* (2nd ed.). Trans. G.A. Kennedy. Oxford University Press, Oxford, UK.
- [12] Marcus F. Quintilian. 1920-1922. *Institutio Oratoria (Institutes of Oratory)*. Trans. H.E. Butler. Retrieved from [http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Quintilian/Institutio\\_Oratoria/home.html](http://penelope.uchicago.edu/Thayer/E/Roman/Texts/Quintilian/Institutio_Oratoria/home.html)
- [13] Annette N. Markham. 2018. What's the trolley problem got to do with It: Ethics as method. July. Retrieved from [https://annettemarkham.com/2018/07/ethics\\_as\\_method\\_published/](https://annettemarkham.com/2018/07/ethics_as_method_published/)
- [14] C.E. Shannon. 1948. A mathematical theory of communication. *Bell System Technical Journal*, 27, 3 (July-Oct 1948), 379–423, 623–656. DOI: <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- [15] Warren Weaver and Claude E. Shannon. 1949. *The Mathematical Theory of Communication*. University of Illinois Press, Normal, IL.
- [16] George Lakoff and Mark Johnson. 1980. *Metaphors We Live By*. University of Chicago Press, Chicago, IL.
- [17] OpenAI. 2019. Better language models and their implications. Retrieved from <https://openai.com/blog/better-language-models/>
- [18] Alec Radford et al. 2019. Language models are unsupervised multitask learners. *OpenAI Blog*. Retrieved from <https://openai.com/blog/better-language-models/>
- [19] Donald Dong. 2018. Text generation using recurrent neural networks. *Medium*. Retrieved from <https://towardsdatascience.com/text-generation-using-rnms-fdb03a010b9f>
- [20] Dan Nelson. 2019. Text generation with Python and TensorFlow/Keras. *Stack Abuse*. Retrieved from <https://stackabuse.com/text-generation-with-python-and-tensorflow-keras/>
- [21] Lucy A. Suchman. 1987. *Plans and Situated Actions: The Problem of Human Machine Communication*. Cambridge University Press, New York, NY.
- [22] Lloyd F. Bitzer. 1968. The rhetorical situation. *Philosophy & Rhetoric*, 1, 1 (Jan. 1968), 1–14. Retrieved from <https://www.jstor.org/stable/40236733>
- [23] John Poulakos. 1983. Toward a sophistic definition of rhetoric. *Philosophy & Rhetoric*, 16, 1, 35-48. Retrieved from <http://www.jstor.org/stable/40237348>
- [24] Stevan Harnad. 1990. The symbol grounding problem. *Physica D, Nonlinear Phenomena*, 42, 1-3, 335–346. DOI: [https://doi.org/10.1016/0167-2789\(90\)90087-6](https://doi.org/10.1016/0167-2789(90)90087-6)
- [25] Luciano Floridi. 2018. Semantic capital: Its nature, value, and curation. *Philosophy & Technology*, 31, 481-497. DOI: <https://doi.org/10.1007/s13347-018-0335-1>
- [26] Ruman Chowdhury. 2019. The pitfalls of a “retrofit human” in AI. *VentureBeat*. Retrieved from <https://venturebeat.com/2019/11/11/the-pitfalls-of-a-retrofit-human-in-ai-systems/>
- [27] Joyce Y. Chai, Rui Fang, Changsong Liu, and Lanbo She. 2016. Collaborative language grounding toward situated human-robot dialogue. *AI Magazine*, 37, 4 (Winter 2016), 32–45. DOI: <https://doi.org/10.1609/aimag.v37i4.2684>
- [28] Stina Bengtsson. 2018. Ethics exists in communication: Human-machine ethics beyond the actor-network. Media@LSE Working Paper Series, London School of Economics and Political Science. OAI: DiVA.org:sh-37239
- [29] Robert Earl Patterson and Robert G. Eggleston, R.G. 2018. Machine synergism in high-level cognitive functioning: The human component. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2, 4 (July 2018), 249-257. DOI: <https://doi.org/10.1109/TETCI.2018.2816584>
- [30] Federica Russo. 2018. Digital technologies, ethical questions, and the need of an informational framework. *Philosophy & Technology*, 31, 4 (Dec. 2018), 655-667. DOI: <https://doi.org/10.1007/s13347-018-0326-2>
- [31] Editor. 2019. What is the Quakebot and how does it work? *LATimes*. Retrieved from <http://www.latimes.com/la-me-quakebot-faq-20190517-story.html>
- [32] Paul Roetzer. 2018. This marketing tool uses AI to create copy that converts. *Artificial Intelligence Marketing Institute*. Retrieved from <http://www.marketingainstitute.com/blog/persado-spotlight>
- [33] Persado. 2019. Persado—How it works. Retrieved from <https://www.persado.com/how-it-works/>
- [34] Joseph N. DiStefano. 2019. “Machine learning is the path to more humanity in marketing”: JPMorgan picks AI to write ads. *Philadelphia Inquirer*, July 30. Retrieved from <https://www.inquirer.com/business/phillydeals/jpmorgan-artificial-intelligence-copy-writing-persado-ads-20190730.html>
- [35] Annette N. Markham. 2013. Undermining “data”: A critical examination of a core term in scientific inquiry. *First Monday*, 18, 10. DOI: <https://doi.org/10.5210/fm.v18i10.4868>
- [36] Annette N. Markham, Katrin Tiidenberg, and Andrew Herman. 2018. Ethics as methods: Doing ethics in the era of big data research—Introduction. *Social Media + Society*, July-September. DOI: <https://doi.org/10.1177/205630511878450>
- [37] Drew Harwell. 2019. A face-scanning algorithm increasingly decides whether you deserve the job. *Washington Post*, November 6. Retrieved from <https://www.washingtonpost.com/technology/2019/10/22/ai-hiring-face-scanning-algorithm-increasingly-decides-whether-you-deserve-job/>
- [38] J.C.R. Licklider. 1960. Man-Computer Symbiosis. *IRE Transactions on Human Factors in Electronics*, 1, 1 (March 1960), 4-11. DOI: <https://doi.org/10.1109/THFE.1960.4503259>
- [39] Sang-won Leigh, Harshit A. Agrawal, and Pattie Maes. 2018. Robotic symbiosis: Interweaving human and machine actions. *IEEE Pervasive Computing*, 17, 2 (June 2018), 34-43. DOI: <https://doi.org/10.1109/MPRV.2018.022511241>
- [40] H. James Wilson and Paul R. Daugherty. 2018. Collaborative intelligence: humans and AI are joining forces. *Harvard Business Review*. Retrieved from <https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces>
- [41] Susan L. Epstein. 2015. Wanted: Collaborative intelligence. *Artificial Intelligence*, 221(April 2015), 36-45. DOI: <https://doi.org/10.1016/j.artint.2014.12.006>
- [42] Andreas Fügener, Jörn Grahl, Alok Gupta, and Wolfgang Ketter. 2019. Collaboration and Delegation between Humans and AI: An Experimental Investigation of the Future of Work. ERIM Report Series. DOI: <http://dx.doi.org/10.2139/ssrn.3368813>
- [43] Mohammad Hossein Jarrahi. 2018. Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making. *Business Horizons*, 61, 4 (July-Aug. 2018), 577-586. DOI: <https://doi.org/10.1016/j.bushor.2018.03.007>
- [44] Nicky Case. 2018. How to become a centaur. *Journal of Design and Science*. DOI: <https://doi.org/10.21428/61b2215c>