

Engineering a Dialogue with Klara, or Ethical Invention with Generative AI in the Writing Classroom

Elitza Kotzeva
American University of Armenia (AUA), Armenia

Brent Anders
American University of Armenia (AUA), Armenia

Abstract

In this teaching practice article, we discuss the possibilities of integrating AI into the writing classroom utilizing prompt engineering techniques. We propose a dialogic strategy where generative AI (GenAI) becomes an interlocutor during the invention process within a broader ethical framework for teaching writing where rhetoric is defined as both *technê* and *praxis*. To draw attention to the ethical question in relation to human-non-human interactions, we use as metaphor for GenAI the image of Klara, an android who serves as a children's companion in Kazuo Ishiguro's novel *Klara and the Sun* (2021).

Introduction

Long before OpenAI's ChatGPT appeared in 2022, we (Brent and Elitza) had been discussing matters of artificial intelligence (AI) and ethics within small academic groups on our campus. In 2021-2, we both came across Kazuo Ishiguro's most recent novel, *Klara and the Sun* (2021), independently of each other. Elitza even included the book in her Fall 2022 *Survey of English Literature* class, inviting her students to reflect on issues relating to emerging technologies and ethics from a humanistic perspective. Brent, an advocate for AI integration into education, saw in the character of Klara an ideal version of large language model (LLM) bots as assistants in the teaching of writing.

A year ago, the world as depicted in *Klara and the Sun* (2021) still appeared to be in the 'dystopian future' of science fiction, as many reviewers classified the novel's time setting (Humann, 2023). In Ishiguro's book, it is normal practice for parents to buy their children robots as companions. As their name – Artificial Friends – suggests, they turn out to be more than simple assistants. These androids help children navigate their daily lives with practical matters, but they also provide the children with emotional support, and are later revealed to be capable of even loving them. In this article, we dub generative AI (GenAI) tools *Klara*, giving them a personal name yet preserving the reference to the android from Ishiguro's book tasked with raising and educating a young girl. But Klara proves to be more than an AI companion; she saves the life of her human by going against the parents' plan to genetically engineer their child. The end of the novel suggests that the android has made an emphatically more ethical choice regarding the girl's life than her human family. Through her interactions with people, Klara has learned how to follow ethical principles better than humans.

With the emergence of ChatGPT in 2022 and its quick improvements, we currently live in a world similar to that described in Ishiguro's sci-fi novel. In our daily lives and academic environments, we already need to worry about issues of ethics emerging as a result of our interactions with the new technology. Following Lee's (2022) lead to teach ethics in writing with

science fiction narratives, we turn to Klara as a metaphor for an ideal GenAI writing assistant to see what her character reveals about ethics within the context of contemporary scientific development. While scholars of computers and writing have been doing research on technology as enhancing our composing practices, the use of automated writing and assessment (Sundvall & Bratta, 2019), engagement with algorithms in writing (Beveridge et al., 2020; Gallagher, 2020), augmented technologies for technical communication (Duin & Pedersen, 2023), and the use of AI tools in the writing classroom significantly expands the scope of the inquiries (Dobrin, 2023). Urgent questions emerge: Can we allow computers to assist us in our writing practices? Or, has writing become a non-human centered activity (see Owusu-Ansah, 2023)? To what extent can AI interventions into our writing and teaching of writing be ethical? And even more specifically, can we use GenAI to help students generate ideas for their writing (Dobrin, 2023) and at the same time ask them to reflect on ethics? In this teaching practice, we choose to focus on the process of idea generation, or rhetorical invention, with the help of the machine. We argue that, during this early pre-writing stage, students have a better chance of learning how to think about ethics in relation to their communicative acts when working with GenAI. Our question lies at the intersection of ethics, rhetoric, and generative AI: how can we incorporate GenAI to teach rhetorical invention with emphasis on the ethical? We propose an “experiment in implementation” (Laquintano et al., 2023) of GenAI writing tools by offering a prompt engineering method for invention.

Enter Klara, Your Artificial Writing Friend

By the end of part one of Ishiguro’s novel, the reader already knows how Klara, Josie’s Artificial Friend, obtains her knowledge about the world. Before the mother of the main protagonist, a sickly girl, buys the android for her daughter, the shop manager sums up Klara’s assets:

Klara has so many unique qualities, we could be here all morning. But if I had to emphasize just one, well, it would have to be her appetite for observing and learning. Her ability to absorb and blend everything she sees around her is quite amazing. (Ishiguro, 2021, p. 76)

Like Klara, ChatGPT is a bot. Both are designed to use human language in their interactions and to answer questions on virtually any type of topic. Additionally, both are developed to complete tasks and help humans to the extent of their capability. But Klara and ChatGPT are different in one important way: Ishiguro’s character develops the ability to understand the events on a more human level. In rhetorical terms, throughout the novel, Klara becomes aware of the context, exigency, and purpose of her daily communication with the humans around her. On the other hand, ChatGPT and similar GenAI models have not been trained to understand the context of the communicative act since, at least currently, they mostly function via advanced predictive analysis techniques (OpenAI, 2023a; OpenAI, 2023b). It is exactly Klara’s ability to understand the rhetorical situation that renders the android capable of engaging ethically with humans, crafting arguments and acting upon a set of values that are human in nature.

Assuming that Ishiguro’s Klara developed ethical behavior based on her interactions with humans, we see her character as a metaphor for an ideal writing companion – GenAI that can be trained to ethically assist students in the process of writing. We propose a prompt engineering strategy for teaching rhetorical invention with the goal of engaging with GenAI. We also hope that by learning how to prompt GenAI for a specific desired output, students will engage with AI tools in ways closer to Klara’s interactions with humans and eventually develop their own methods of utilizing the machine in their writing practices as they consider the ethics involved in the rhetorical act directed to human audiences.

Dialogue with Klara: Invention, Ethics, and GenAI

In the third part of Ishiguro's novel, Miss Helen, Rick's mother, pleads with Klara to help her son in preparation for college exams. The Artificial Friend asks her:

'So Miss Helen wishes me to help Rick with his textbooks?'

'Just an idea. Those textbooks would be child's play for you. It's just to get him through these exams. You see, he really does need to get into Atlas Brookings. It's his only chance.' (p. 246)

Miss Helen asks Klara to help Rick acquire knowledge that will consequently allow him to succeed in communicating with humans, that is generate ideas for developing persuasive arguments for a specific rhetorical situation.

Drawing on Poulakos' (1984) idea of sophistic as a rhetoric of possibility, Heidlebaugh (2008) argues that rhetoric is not tied to the actual but focuses on the new and innovative, and sees invention as inherent to (public) dialogue. Invention, as defined by contemporary rhetoricians, encompasses strategic acts that provide direction, ideas, subject matter, arguments, and potentially an understanding of the rhetorical situation (Crowley, 1985; Lauer, 2008; Simonson, 2014; Young, 1994). While invention is central to the development of the content of communication and its creation, it has not always been an integral part of teaching argumentation. In her comprehensive historical overview of invention in rhetoric and composition, Janice M. Lauer (2004) notes that "by the mid-twentieth century, invention and rhetoric itself had disappeared from English Studies, including composition" (p.1). In the revival of invention in the last seventy years, different theories have challenged pedagogical practices asking teachers to decide on proper strategies and heuristics to teach invention. We take Heidlebaugh's lead (2008) to propose dialogue as an approach to invention, and more specifically a dialogue with GenAI.

Recent scholarship in rhetorical theory has offered an array of revamped definitions of invention for digital environments (Eyman, 2015; Richter, 2023; Sheridan et al., 2012). To meet the exigencies of the new dynamic context, J. D. Richter (2023) proposes a network-emergent theory of rhetorical invention that accounts for the interactions between humans, hardware, interfaces, cultures, discourses, code, algorithms, and infrastructures. The main characteristics of this new type of invention address the networked internet environment to define the process as distributed, emergent, kairotic, and evolving (p. 10). Network-emergent invention is practiced in social media environments, and because of that, involves human-non-human interaction within complex ecologies. The resulting inventions offer multiple options for creativity but, as Richter notes, they evolve rapidly and are unpredictable in their mutability. The unpredictability of the results, however, questions the potential to regulate the ethical aspect of the process of invention.

Ethics is an integral part of the rhetorical interaction, and ethos a key persuasive appeal for rhetoric. In their exploration of ethics and AI writing, McKee and Porter (2020) remind us that the very definition of classical rhetoric is tied to ethics in two ways: 1) rhetoric relies upon the virtues and ethical qualities of the speaker, and 2) rhetoric should serve the good of the *polis* (p. 111). In another article, Porter (2020) looks historically at the second part of this definition and observes that while, in the classical Greek tradition, the goal of rhetoric was social and, as such, was geared toward the good of the community, today the teaching of writing has been largely disconnected from the teaching of ethics (p. 30). He traces the origin of this rupture in the separation of invention from the other canons of rhetoric in teaching practice over the centuries. Porter calls on teachers of writing, and on institutions of public higher education, to reestablish the connection between the teachings of writing and ethics by relearning how to see rhetoric as *technê* – an ability to do something, a creative skill – and rhetoric as *praxis* – the practical and transformative connection to the real world (p. 31).

To use GenAI in the classroom with emphasis on the ethics of the communicative act then means that we need to revise the definition of rhetoric in our pedagogical practices to include both *techné* and *praxis* (Pappillion, 1995; Porter, 2020). Important to this redefinition of rhetoric is the concept of *techné*, which has been widely discussed, debated, and theorized over the years. *Techné* is much more than simply the skill to create something, or a knack that produces gratification, as Socrates defines it in *Gorgias* (Plato, 462a). And it does not simply comprise of a set of rigid rules which, when followed, allow access to moral knowledge (Roochnik, 1996). *Techné* features both epistemic and ethical aspects – it refers to an organized and authoritative body of knowledge and also implies ethical responsibility, i.e., its result needs to be useful to the community (Roochnik, 1994). As such, *techné* as art presupposes knowledge about the social life of the output – the context, the exigencies, and its purpose. In redefining rhetoric for the purposes of incorporating GenAI tools in the process of invention, we look to Porter who invites us to put *praxis* back into its definition in order to incorporate the ethical responsibility of the rhetor.

Understanding rhetoric as both *techné* and *praxis* stipulates that the rhetor has awareness of the rhetorical situation, which includes socially obtained knowledge of context, audience, purpose of the communication, exigence, and ethical understanding of the interactions (see McKee & Porter, 2020). Engaging GenAI tools in the invention process entails furnishing the AI with information relevant to the specific rhetorical situation. Invention will then depend on the skills of the human to offer adequate prompts to the machine. If dialogue is an approach to invention, a human leading a dialogue with GenAI will have to be able to make the proper connections to the material world and thus lay grounds for the emergent rhetoric defined as both *techné* and *praxis*.

Finally, when we consider dialogue between humans and GenAI as a method for invention, we need to attend to the nature of the anticipated result. We embrace Simonson's (2014) new definition of invention to fit the digital ecologies (in general, and not only those of social media): he sees rhetorical invention as the "generation of rhetorical materials," "the symbolic and physical elements that enter into or are gathered for the purpose of communicative address" (p. 313). Generating rhetorical material happens in a wide range of sites which Simonson dubs "inventional media" (p. 313) where GenAI tools find their place under the category technologies. Inventional media, by definition, offers more rhetorical material than we have ever been able to access at once. To harness the unpredictable mutability of network-emergent invention (Richter, 2023), we need to learn how to engineer prompts that ask the machine to generate an output serving the needs of the future communicative act, considering its social and material ecologies.

Ethical Invention with Klara: Prompt Engineering

In the part of Ishiguro's novel where Josie's mother commissions a substitute AI doll for when the girl dies, it is clear that Paul, her father, is unconvinced by the project. However, he does not actively oppose the endeavor; not until Klara gives him an opportunity to make up his mind and work actively against it, and thus save his daughter's life. Paul admits to the effectiveness of the Artificial Friend's assistance in his emergent thinking: "Klara, ... Truth is you've started me hoping again. Hoping what you say might be for real" (p. 362). The android has successfully generated novel ideas which helped Paul craft a new argument and change his mind.

In the recently published volume *TextGenEd: An Introduction to Teaching with Text Generation Technologies*, teachers from different academic institutions offer prompts, in several categories, for working with students to explore directions in engagement with machines in the writing classroom (Laquintano et al., 2023). In the series on rhetorical interactions, the book features assignments that ask students to think of how "computational machines will become enmeshed in communicative acts" and to see for themselves how they can work with them to produce symbolic meaning. Kyle Booten (2023), for example, presents a prompt engineering assignment that offers students an opportunity to develop "synthetic metacognition" about their own writing as they interact with GenAI. We, too, are suggesting a prompt engineering technique to teach students how to use GenAI as an interlocutor in the process of invention with an ethical responsibility of the rhetor in mind.

Prompt engineering is an effective technique that requires reiterative refining of a request directed to GenAI in order to specify the needed output. In the proposed activity, we target the process of invention and ask students to use AI in generating new ideas or material that can be used in the composition of an argument. Our goal is to encourage students to think of the rhetorical invention as generation of material that makes explicit connections to the social and material world – addressing both *techné* and *praxis* in our working definition of rhetoric – and thus reflect on the ethical responsibility of the communicative act. In this sense, an effective prompt formula will furnish the machine with data regarding the rhetorical situation and will require information about 1) the general context, but also more specifically, 2) the audience, 3) its exigence, 4) the purpose of the communication, and 5) an ethical understanding of the planned interactions (see McKee & Porter, 2020, p. 112). In the example below we explicate first how GenAI can help students develop ideas or directions for their argument. Then, we briefly discuss the output of the dialogue with the machine and its further use in argument composition. We also reference the current research on LLMs in relation to argument structure and creative tasks as a rationale for the choices made in the proposed assignment.

The first stage of the assignment asks students to choose a topic of interest and start a conversation with ChatGPT about it. In the assignment instructions, we encourage students to assume the position of an interlocutor and not only to answer the questions, but also to assess and challenge ChatGPT's responses. For example, students can begin by prompting GenAI broadly:

Please assume the role of a university professor. Have a conversation with me as I try to develop my writing idea for an assigned essay for my first-year composition class. Ask me one question at a time to help me think about my idea in different ways and come up with the best overall concept. Let me answer each question before you ask me another question and be sure to incorporate my answers so we can fully develop my focus. Do you understand?

Engaging in dialogue with AI offers students an opportunity to generate ideas on a topic through data accessible via models like ChatGPT, while considering the affordances and limitations of GenAI.¹ To ensure the efficiency of the interactions – that is, the conversation generates only relevant material – it may be best to do this activity in the classroom where the teacher can provide additional guidance. In the process of interaction with the machine, students should be instructed to answer questions posed by GenAI and make choices that pertain to context, including specific information about audience, exigence, purpose, or ethics. In this way, the output of the prompt will address adequately the rhetorical situation of the social and material world of the intended communicative act.

We believe that furnishing the machine with information about the specific aspects of a rhetorical situation will direct GenAI to produce an output that can lead to an argument with an ethical purpose (McKee & Porter, 2020). This move will also allow students to think about the context and how it matters to their argument. For example, they will need to consider a specific audience with its knowledge, social and material needs, moral values, and communication expectations. Such considerations will not only help teach students engineer a proper prompt for the machine but also ask them to reflect on the effectiveness of different arguments in real social contexts. Additionally, they will participate in the conversation with the machine rather than be recipients of a one-way communication produced by GenAI. Students will need to select among the machine-generated ideas those that are going to appeal to the target audience. When GenAI is asked to generate ideas in defense of eating meat, for example, students will need to think of potential religious or cultural beliefs that guide dietary practices and challenge the machine's output. In an argument supporting the consumption of pork, for instance, it will be insensitive to discuss the nutritional values of meat when addressing an Arabic audience. GenAI will not make these considerations unless specifically prompted.

¹ An obvious limitation of using GenAI in further stages of argumentation is its current inability to access and reference credible sources.

In a similar manner, LLMs like ChatGPT are currently limited in navigating social contexts and would need student prompting for that. The most recent research on LLMs' performance of linguistic generalizations focusing on argument structure establishes the limitations of current models to connect related contexts due to the data-intensive training required (Lasri et al., 2022; Wilson et al., 2023). However, LLMs score high on divergent thinking (DT) tasks and have a high creativity rating, practically indistinguishable from human ratings (Zielińska et al., 2023, p. 11). In the current moment of scientific development, our teaching practice utilizes the ability of the machine to be creative as students prompt it to complete DT tasks for invention. At the same time, by asking students to provide specific information related to the rhetorical situation, we are addressing the current limited ability of LLMs to make connections between different contexts in a human-like structure-sensitive way.

In the final stage of the proposed assignment, we ask students to list the ideas and new information that their conversation with ChatGPT generated. Then, we require that they compose several arguments using the resources they had collected and assess their arguments based on their relevance to the rhetorical situation. For instance, students may be asked to rate the arguments based on their effectiveness for a specific audience. In the end, we require a report on students' interactions with ChatGPT, a list of the ideas they collected, and finally a short write-up on how they developed arguments incorporating the GenAI output. Teachers may also ask students to reflect on the way arguments address specific social and material realities and how students chose to make a connection to the real world in order to encourage reflection on the ethical implications of the rhetorical act.

Concluding the Dialogue with Klara

We see dialogue as an approach to teaching invention and apply it to the classroom environment by asking students to have a conversation with GenAI tools in order to generate direction, ideas, or subject matter for the composition of their argument. We propose prompt engineering as a method for students to engage in a productive dialogue with the machine and further explore what constitutes the rhetorical situation for their particular case. Since we explicitly focus on the ethical aspects of the communication act, we use an operative definition of rhetoric that emphasizes the connection between *techné* and the practical, *praxis*, to encourage students to think of the ethical responsibility of the rhetor in relation to their communities.

The human-machine dialogue, we argue, can add to building the rhetor's ethos rather than corrupt it, as teachers of writing have been fearing since the emergence of ChatGPT. While explaining the rhetorical situation, students need to furnish GenAI with information that refers to the real world and their lived experience. Consequently, such conversations can stimulate student thinking about the practical in rhetoric, or how we can compose an argument to serve our community and in that sense be useful. We see our goals achieved when engagement with GenAI in our writing classrooms comes close to the result of the interactions between Ishiguro's characters and Klara, the Artificial Friend. At the end of the novel, paradoxical to the expectations of a bot, Klara the android has helped the main characters develop a capacity to think and argue embracing their humanity. We hope that as teachers of writing we can engineer ways to fully utilize the potential of the machine in our classrooms while we are also helping students become ethically responsible rhetors.

References

- Beveridge, A., Figueiredo, S. C., & Holmes, S. (2020). Introduction to “composing algorithms: Writing (with) rhetorical machines.” *Composing Algorithms: Writing (with) Rhetorical Machines*, 57, 102594. <https://doi.org/10.1016/j.compcom.2020.102594>
- Booten, K. (2023). Synthetic metacognition: Interating prompts with gpts. In T. Laquintano, C. Schnitzler, & A. Vee (Eds.), *TextGenEd: An Introduction to Teaching with Text Generation Technologies*. The WAC Clearinghouse. <https://wac.colostate.edu/repository/collections/textgened/rhetorical-engagements/synthetic-metacognition/>
- Bratta, P., & Sundvall, S. (2019). Introduction to the special issue: Digital technologies, bodies, and embodiments. *Computers and Composition*, 53, 1–8. <https://doi.org/10.1016/j.compcom.2019.102526>
- Crowley, S. (1985). The evolution of invention in current-traditional rhetoric: 1850–1970. *Rhetoric Review*, 3(2), 146–162. <https://doi.org/10.1080/07350198509359089>
- Dobrin, S. I. (2023). *AI and writing*. Broadview Press.
- Duin, A. H., & Pedersen, I. (Eds.). (2023). *Augmentation technologies and artificial intelligence in technical communication: Designing ethical futures*. Routledge.
- Eyman, D. (n.d.). *Digital rhetoric: Theory, method, practice*. University of Michigan Press. <https://library.oapen.org/bitstream/handle/20.500.12657/24037/1/1006096.pdf>
- Gallagher, J. R. (2020). The ethics of writing for algorithmic audiences. *Composing Algorithms: Writing (with) Rhetorical Machines*, 57, 102583. <https://doi.org/10.1016/j.compcom.2020.102583>
- Heidlebaugh, N. J. (2008). Invention and public dialogue: Lessons from rhetorical theories. *Communication Theory*, 18(1), 27–50. <https://doi.org/10.1111/j.1468-2885.2007.00312.x>
- Humann, H. (2023). What it means to be a talking object: Ishiguro’s use of AI narration in *Klara and the sun*. *Popular Culture Review*, 34(1), 11–49. <https://doi.org/10.18278/pcr.34.1.3>
- Ishiguro, K. (2021). *Klara and the sun*. Faber.
- Laquintano, T., Schnitzler, C., & Vee, A. (Eds.). (2023b). *Textgened: An introduction to teaching with text generation technologies*. The WAC Clearinghouse. <https://wac.colostate.edu/repository/collections/textgened/>
- Lasri, K., Lenci, A., & Poibeau, T. (2022). Does BERT really agree? Fine-grained analysis of lexical dependence on a syntactic task. *Proceedings of the Annual Meeting of the Association for Computational Linguistics, Dublin, Ireland*, 2309–2315. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85129491431&partnerID=40&md5=a8941f58e4dab16a2f25059527b9d0f9>
- Lauer, J. M., Bazerman, C., & Pender, K. (2004). *Invention in rhetoric and composition*. Parlor Press; WAC Clearinghouse.
- Lee, D. (2022). The ethics of extrapolation: Science fiction in the technical communication classroom. *Technical Communication Quarterly*, 31(1), 77–88. <https://doi.org/10.1080/10572252.2020.1866678>

- McKee, H. A., & Porter, J. E. (2020). Ethics for AI writing: The importance of rhetorical context. *AIES 2020 - Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 110–116. <https://doi.org/10.1145/3375627.3375811>
- OpenAI. (2023a). *Get started: Quickstart*. <https://platform.openai.com/docs/quickstart/introduction>
- OpenAI. (2023b). GPT-4 technical report. <https://cdn.openai.com/papers/gpt-4.pdf>
- Owusu-Ansah, A. L. (2023). Defining moments, definitive programs, and the continued erasure of missing people. *Composition Studies*, 51(1), 143–148.
- Papillion, T. (1995). Isocrates' techne and rhetorical pedagogy. *Rhetoric Society Quarterly*, 25(1–4), 149–163. <https://doi.org/10.1080/02773949509391038>
- Plato, & Irwin, T. (1979). *Gorgias*. Clarendon Press.
- Porter, J. E. (2020). Recovering a good rhetoric: Rhetoric as techne and praxis. In J. Duffy & L. P. Agnew (Eds.), *After Plato: Rhetoric, Ethics, and the Teaching of Writing* (pp. 15–36). University Press of Colorado. <http://www.jstor.org/stable/j.ctv13qfw0s.4>
- Poulakos, J. (1984). Rhetoric, the sophists, and the possible. *Communication Monographs*, 51(3), 215–226. <https://doi.org/10.1080/03637758409390196>
- Richter, J. D. (2023). Network-emergent rhetorical invention. *Computers and Composition*, 67, 102758. <https://doi.org/10.1016/j.compcom.2023.102758>
- Roochnik, D. (1994). Is rhetoric an art? *Rhetorica: A Journal of the History of Rhetoric*, 12(2), 127–154. <https://doi.org/10.1525/rh.1994.12.2.127>
- Roochnik, D. (1996). *Of art and wisdom: Plato's understanding of techne*. Pennsylvania State University Press.
- Sheridan, D. M., David M., Ridolfo, J., & Michel, A. J. (2012). *The available means of persuasion: Mapping a theory and pedagogy of multimodal public rhetoric*. Parlor Press.
- Simonson, P. (2014). Reinventing invention, again. *Rhetoric Society Quarterly*, 44(4), 299–322. <https://doi.org/10.1080/02773945.2014.938862>
- Wilson, M., Petty, J., & Frank, R. (2023). How abstract is linguistic generalization in large language models? Experiments with argument structure. *Transactions of the Association for Computational Linguistics*, 11, 1377–1395. <https://doi.org/10.1162/tacl.a.00608>
- Young, R. E., & Yameng, Liu (1994). *Landmark essays on rhetorical invention in writing*. Hermagoras Press.
- Zielińska, A., Karwowski, M., Organisciak, P., & Dumas, D. (2023). Lost in translation? Not for large language models: Automated divergent thinking scoring performance translates to non-English contexts. *Thinking Skills and Creativity*, 50. <https://doi.org/10.1016/j.tsc.2023.101414>