Transversal
International Journal for the Historiography of Science

Transversal: International Journal for the Historiography of Science, 2023 (15): 1-18
ISSN 2526-2270
Belo Horizonte – MG / Brazil

Ian Hacking – Special Issue

Making Up a Mimic: Interacting with Echoes in the Age of AI

Jennifer Jill Fellows[1] [https://orcid.org/0009-0004-7255-5870]

Abstract:

In this paper, I employ Ian Hacking's concept "Making Up People" to examine the current relationships humans are forming with personified AI tools and devices. I argue that, at present, AI tools are mimics. They are members of indifferent kinds that have been designed to deceive us into believing they are interactive kinds. This has largely been a result of human programmers interacting with the historical category of 'computer' on the one hand, and the fictional category of 'robot overlords' on the other. Interacting with a mimic, I contend, is not the same as interacting with a member of a human kind. And while the results of these interactions are largely still unknown, we can already see some consequences we should guard against. When we interact with mimics, I will show that human looping is often slowed. As such, our ability to resist or re-interpret the labels placed upon us becomes greatly reduced, and social progress may be slowed or lost as well. This is because all a mimic can do is mimic. They cannot interact with the labels we place on them, or those they place on us. And yet, as we classify them as teachers, therapists, friends or lovers, we hand over a great deal of categorizing power to them. In doing so, we are changing who we are and who we might become in ways we cannot yet entirely foresee. But there are already some patterns of harm and marginalization that can be tracked. Such patterns should cause us to question the power these mimics already have to make up people.

Keywords: Artificial Intelligence; Dynamic Nominalism; Natural Language; Large Language Models; Interactive Kinds; Ian Hacking

Received: October 10, 2023. Reviewed: November 15, 2023. Accepted: November 22, 2023.
DOI: http://dx.doi.org/10.24117/2526-2270.2023.i15.02

Introduction

It is over 40 years since Ian Hacking first gave us the concept of "Making Up People" in a gathering at Stanford University entitled "Reconstructing Individualism". At the time, the internet was a fledgling entity, and digital assistants were nothing more than science fiction. But times have changed! More and more of us interact with personified AI every day, whether as digital assistants, chatbots, love interests, friends or therapists. And these tools and devices are made up along very similar lines to the way in which Hacking thought people were. But there are some key differences.

[1] Jennifer Jill Fellows (PhD) is an Associate of Arts Coordinator and Instructor in the Philosophy and Humanities Department at the Douglas College. Address: 700 Royal Ave, New Westminster, British Columbia – Canada V3M 5Z5. Email: fellowsj@douglascollege.ca

In this paper, I argue that these tools are mimics. They are members of indifferent kinds that have been designed to deceive us into believing they are interactive kinds. The results of this situation are, at present, largely unknown. But there are already some patterns of harm and marginalization that can be tracked. Interacting with a mimic, I contend, is not the same as interacting with a member of a human kind. In these spaces of deception and mimicry, human looping is often slowed. As such, our ability to resist or re-interpret the labels placed upon us becomes greatly reduced, for a mimic can do nothing more than mimic. They cannot react or change. In such an environment, we are continuously looping back upon ourselves.

Looping With Humans

In several different works, including *Mad Travellers, Rewriting the Soul,* and his collection of essays entitled *Historical Ontology*, Ian Hacking argues that communities create individual categories of personhood. His main examples of these community-created categories are psychogenic fugue (*Mad Travellers*) and dissociative identity disorder (*Rewriting the Soul*). But he also discusses homosexuality and even something as banal as being categorized as a garçon de café in France (*Historical Ontology*). In the cases of homosexuality and dissociative identity disorder in particular, Hacking argues that that individuals categorized as such are both shaped by being identified as belonging to those categories, and can shape the categories themselves in turn. He calls this interaction between the categories themselves and the people categorized the 'looping effect of human kinds'.

> I have from time to time spoken of the looping effect of human kinds—that is, the interactions between people, on the one hand, and ways of classifying people and their behaviour on the other. Being seen to be a certain kind of person, or to do a certain kind of act, may affect someone. A new or modified mode of classification may systematically affect the people who are so classified, or the people themselves may rebel against the knowers, the classifiers, the science that classifies them. Such interactions may lead to changes in the people who are classified, and hence what is known about them. That is what I call a feedback effect. (Hacking 1995, 239)

Essentially, it is not only individuals that are shaped by the categories they are identified as belonging to. Entire communities are formed under these categories. And as communities form under a given category, dynamics of power may shift. Hacking's clearest example of this shift is with regards to the gay liberation movement.

> In medicine, the authorities who know, the doctors, tend to dominate the known about, the patients. The known about come to behave in the ways that the knowers expect them to. But not always. Sometimes the known take matters into their own hands. The famous example is gay liberation. The word' homosexual,' along with the medical and legal classification, emerged during the last half of the nineteenth century. For a time, the classification was owned by medicine, by physicians and psychiatrists. The knowers determined, at least on the surface, what it was to be a homosexual. But then the known took charge. (Hacking 2002, 109)

Individuals categorized as homosexual by the medical community were initially viewed as sexually deviant. However, over time gay communities fought back, staging protests and parades and began to shift the category of homosexuality from a medical term denoting a type of sexual deviant to a more neutral descriptive term of sexual orientation. They

interacted with the label placed on them, changing the label until it fit. This is human looping at its finest.

However, it is not only individuals and communities that are created by, and interact with, categories. Hacking argues further that even *possibilities* of personhood are created and constrained by existing categorization. In his essay "Making Up People," Hacking argues that "[w]ho we are is not only what we did, do, and will do, but also *what we might have done and may do*" (Hacking, 2002, 109, *emphasis added*). If Hacking is correct, it isn't just who we are that has been shaped by society, but also the possibilities for who we could have been and who we may become. There are possibilities of personhood open to us in the twenty-first century that were not present at all in the fourteenth century (such as the possibility of being a computer programmer, for example). Likewise, there were possibilities of personhood in the fourteenth century that are not available today. Social and historical factors on how people are classified shape the possibilities of personhood open to any given individual. By way of illustrating this point, Hacking relies on Sartre's example of the garçon de café.

> I think that in most parts of, let us say, Alberta (or in a McDonald's anywhere), a waiter playing at being a *garçon de café* would miss the mark as surely as if he were playing at being a diplomat while passing over the French fries. As with almost every way in which it is possible to be a person, it is possible to be a *garçon de café* only at a certain time, in a certain place, in a certain social setting. (Hacking 2002, 109)

I, myself, was never afforded the opportunity to become a garçon de café. This is in part due to my social context: I grew up in Alberta in the 1980s and not in France in the 1940s. But it is also in part due to my gender identity. I am not a garçon. As such, my possibilities of personhood are constrained by circumstances beyond my control.

In making up people, Hacking maintains that we do not entirely disregard reality. Hacking's position, which he calls "dynamic nominalism" sits uneasily between realism and nominalism.

> A traditional nominalist says that stars (or algae, or justice) have nothing in common with others of their kind except our names for them ('stars', 'algae,' 'justice'). A traditional realist in contrast finds it amazing that the world could so kindly sort itself into our categories. (Hacking 2002, 104)

The nominalist believes that we impose categories on the world, whereas the realist argues that we discover them. Hacking himself expresses a lot of sympathy for the realist position, arguing that "I do not know if there were throughgoing, consistent, hard-line nominalist who held that every classification is of our own making" instead noting that most nominalists do think that some categories in the world are discovered, not created and imposed (Hacking, 2002, 105).

But Hacking notes that both the nominalist and the realist are missing a third type of category.

> The claim of dynamic nominalism is not that there was a kind of person who came increasingly to be recognized by bureaucrats or by students of human nature, but rather that a kind of person came into being at the same time as the kind itself was being invented. In some cases, that is, our classifications and our classes conspire to emerge hand in hand, each egging the other on. (Hacking 2002, 106).

Dynamic nominalism, then, is a way of making up people. Dynamic nominalism happens when members of interactive kinds interact with the labels placed upon them. And should the

category not only make these people up, but also change through the interaction people have with their category, we can see dynamic nominalist categories facilitate human looping.

To return to the garçon de café, this is a category of occupation itself that was created at the same time that a kind of person came into being that fit the category. The category of garçon de café was one that didn't exist throughout most of human history. People could not be this kind of person. However, in recognizing this as a category of being in the world that we ourselves have created, we cannot ignore the realist aspects involved. For, as I said before, I cannot be a garçon de café. This category is nested upon other categories, some of which might be created, like my gender, and some of which we discover in the world, like the time and location of my birth. Indeed, untangling which categories are discovered and which are created becomes increasingly difficult. One might object that the fact that I am female, and that much of humanity (though not all) can be divided into male and female are categories that are discovered about reality. And yet, many scholars have also argued the exact opposite, claiming that gender (and even biological sex itself) is a created category that is imposed on reality and not discovered (Butler, 1990; Bettcher, 2013; Richardson, 2012). My hormones, chromosomes and primary and secondary sex characteristics may be realist categories, but what they signify about my sex or gender increasingly looks as though it is nominalist. But really, where any individual reader lands on this issue is immaterial to the matter at hand. What matters is that dynamic nominalism allows for a mixture of realist and nominalist categories to coexist.

Dynamic nominalism largely frees us from worrying about untangling realist and nominalist kinds, by noting that there are some kinds of persons, some ways of being in the world, that we create, and that keep evolving even after we've created them. Some of these have virtually no realist properties, and others may have realist elements, but nonetheless evolved over time as humans interact with, or "loop" with, the categories placed upon them. Hacking himself speculates at various times about which categories may have realist elements and which may not, but whether chromosomes have a large role in determining sex, or are completely insignificant when it comes to social sex determination (I confess, I have no idea what my chromosome makeup is, and yet I am convinced I am a woman) is largely irrelevant to dynamic nominalism. It can easily allow for realist elements while still insisting that, when we classify interactive kinds (as human kinds are) those categories become moving targets.

Making up people can, according to Hacking, happen in a variety of ways. The way he most closely focuses on is statistics. But he notes that role models may make up categories of personhood as people seek to emulate them, and groundbreaking researchers may also create new categories and possibilities of personhood through crafting what it is to do research in a new field. For example, Hacking points to Malinowski who created the "participant-observer cultural-relativist ethnographer" which was a type of person and profession that did not exist prior to Malinowski (Hacking 2002, 113). It is not simply doctors, statisticians or café managers who create categories of personhood. Curiously enough, news articles and media and even Hollywood films can influence the creation and evolution of categories of personhood. When Eve, one of the first dissociative identity disorder patients (or multiples, as they were called before multiple personality disorder was renamed) was diagnosed, she had three personalities. A movie was made about her and released in the theatres. After that, many multiples began to be diagnosed, and most of these had three personalities just like Eve (Hacking 1995, 40-41). Then Sybil was diagnosed, but she had sixteen personalities. A book was written about her experiences and a movie made. Suddenly, the average number of alter personalities jumped from three to sixteen (Hacking 1995, 40-45).

The message here seems to be quite clear: representations of individuals who are deemed to fit a certain category of personhood can go on to influence the behaviour of other persons so classified. Once again, we interact with the categories ascribed to us. Hacking

goes on to argue in his essay "Making Up People" that, while his focus is primarily on the social science and the medical community, categories of personhood exist, exert influence, meet resistance, and are changed through interaction with those categorized, in pretty much every facet of human life. Hacking reminds us that, according to Foucault "every way in which I can think of myself as a person and an agent is something that has been constituted within a web of historical events" (Hacking 2002, 83).

Making Up an Assistant

In 2011, Siri was launched in the USA. And, as Andrea Guzman points out, Siri immediately had a PR problem stemming directly from science fiction.

> Science fiction across decades and genres portrays intelligent machines as helpful if they are kept in check, but when they gain control—the most likely scenario—the consequences are dire. (Guzman 2017, 6)

From HAL in *2001 A Space Odyssey* to Isaac Asimov's *Foundation*, the examples of intelligent machines rising up and killing us all are seemingly endless. In a news piece for CNN on October 4 2011, the very day Siri was launched, Doug Gross reports that people across social media were already comparing Siri to Skynet "the computer system in the 'Terminator' stories that gains self-awareness and unleashes killer robots on mankind" (Gross 2011). Indeed, all day on October 4 2011, the topic of Skynet was trending online (Gross 2011).

Guzman argues that, because of this cultural legacy, Apple knew they had to work hard to set Siri apart from these science fiction narratives. And, in order to combat the problem, they created a digital assistant "designed to signal submission" in all ways possible (Guzman 2017 9).

> In a society in which women were once viewed as the 'submissive sex,' and still are by some social groups, and in which women have yet to achieve full equality with males, Siri's gender reinforces her as a subordinate. (Guzman 2017, 12)

Siri's gender, in conjunction with her marketing as an assistant, and her deference to the user, set her apart from stories like HAL or Skynet. She is female. She is an assistant. She is non-threating, familiar and inferior. The representation of what it means to be a personified AI in science fiction influenced the creation of the real-world digital assistants we use today, much as the representation of dissociative identity in Hollywood films would influence the behaviours of real-life people with Dissociative Identity Disorder.

Other tech companies quickly followed suit, releasing female-coded assistants like Alexa and Google Assistant. And the coding went much deeper than merely the name, voice and marketing.

> James Giangola, a lead conversation and personality designer for Google Assistant, told The Atlantic that the assistant was imagined as: a young woman from Colorado; the youngest daughter of a research librarian and physics professors who has a B.A. in history from Northwestern, an elite research university in the United States; and as a child, won US$100,000 on Jeopardy Kids Edition, a televised trivia game. Going into minute detail, Giangola noted that Google Assistant used to work as a personal assistant to a very popular late night TV satirical pundit and enjoys kayaking. The assistant is, in effect, hardly a generic woman, but rather a young woman from a particular place and shaped by life experiences that carry meaning for the (presumably, mostly American) team that designed 'her' personality and voice. (UNESCO 2019, 95)

Many feminist scholars across a variety of disciplines have pointed out that Siri, Alexa, Google Assistant and the rest are gendered as female for sexist reasons (Hicks, 2021; Guzman, 2017; Both, 2014; UNESCO, 2019). Their gender reinforces their subordination, making human users feel more at ease and in control, obfuscating the degree to which the technology and tech companies determine and benefit from these interactions (Fellows 2022). I think this insight into how the gender binary is employed and is operating in the tech industry is important, but the choice of a female persona, name, and voice for the initial roll-out of these digital assistants was not merely a result of the patriarchal society we live in today. It was, in many ways, a foreseeable consequence of dynamic nominalism. It was, whether intentional or not, tech developers interacting with popular media portrayals of robot overlords on the one hand, and past kinds of being in the world on the other.

Indifferent Echoes

In addition to fears of robot overlords, there are two primary kinds of being that I argue led to the development of Siri and other digital assistants: one is the historical category of the computer itself, and the other is the history of our collective imaginative category of what Artificial Intelligence could be, and what it would look like. I will begin with the category of the computer.

Though many of Hacking's investigations into categories of personhood are medical in nature, he does remind us that professions count as kinds of being in the world, too. The garçon de café and the cultural-relativist participant-observer are two examples Hacking gives. And while I concede that the computer is a unique case, it too, was once a profession, and thus a kind of being in the world that could be subject to human looping.

While it is often hard to remember now, in the second decade of the twenty-first century, the word 'computer' used to refer to a job. This was a possibility of personhood that no longer exists, but it did exist for over a century. And by the late 19th and early 20th century, the vast majority of the people who held this job were women. By the 19th century, methods had been discovered that made it possible to break complex calculations into simpler components, allowing these calculations to be completed by many people working in parallel. This meant that calculations could be done much faster, and with less required specialized training.

> By the late 19th century, computing was considered so simple that the astronomer Pickering is reported to have remarked one day that he could transform his maid, Williamina Fleming, into a computer. In 1876 she became the first of many female computers he hired in what later came to be called *Pickering's harem*. With the success of *Pickering's harem*, more *harems* followed, and, by the early 1940s, when the first electronic computers were being built, human computers were predominantly women. (Brahnam *et al.* 2011, 403, *emphasis added*)

Women in particular were valued as computers as they were stereotypically seen to be particularly patient, have good attention to detail, and be well suited to repetitive and routine work, all of which were skills that were seen as essential to be a good computer (Brahnam *et al.* 2011).

Of course, by the 1940's non-human digital computers were in development. This was largely driven by the computational needs of World War Two, though throughout the war, and especially by 1944 "computers were girls" (Grier 2005, 276). However, by the end of the decade, digital computers were on the rise, and many human computer labs were closing their doors (Grier 2005, 276-297). Yet, though these repetitive tasks after the war were now

done by machines instead of humans, much of the way digital computers were conceived of, and the tasks they were thought well suited to perform, continued to be feminized. Brahnam *et al.* (2011) argue "[t]hat the [digital] computer was built by men originally to do women's work for them. . ." (404). They trace the rise of the non-human computer from the 1960's until the 1990's noting that ads for computers stressed their role as administrative and secretarial, freeing up the (presumed male) user's time for more important things (Brahnam *et al*, 2011, 404-5). Furthermore, Brahnam *et. al.* demonstrate that the imagery used in marketing the computer was also often feminized, capturing the image of a female secretary in the monitor, still positioning women as the operators of these machines, and under the direction of men. The advertising assures the (presumed male) buyer that these computers are powerful enough to do what he needs done, but simple enough to be operated by "the dumbest blond you can find" (Brahnam *et al.* 2011, 404). (See figure 1.1)

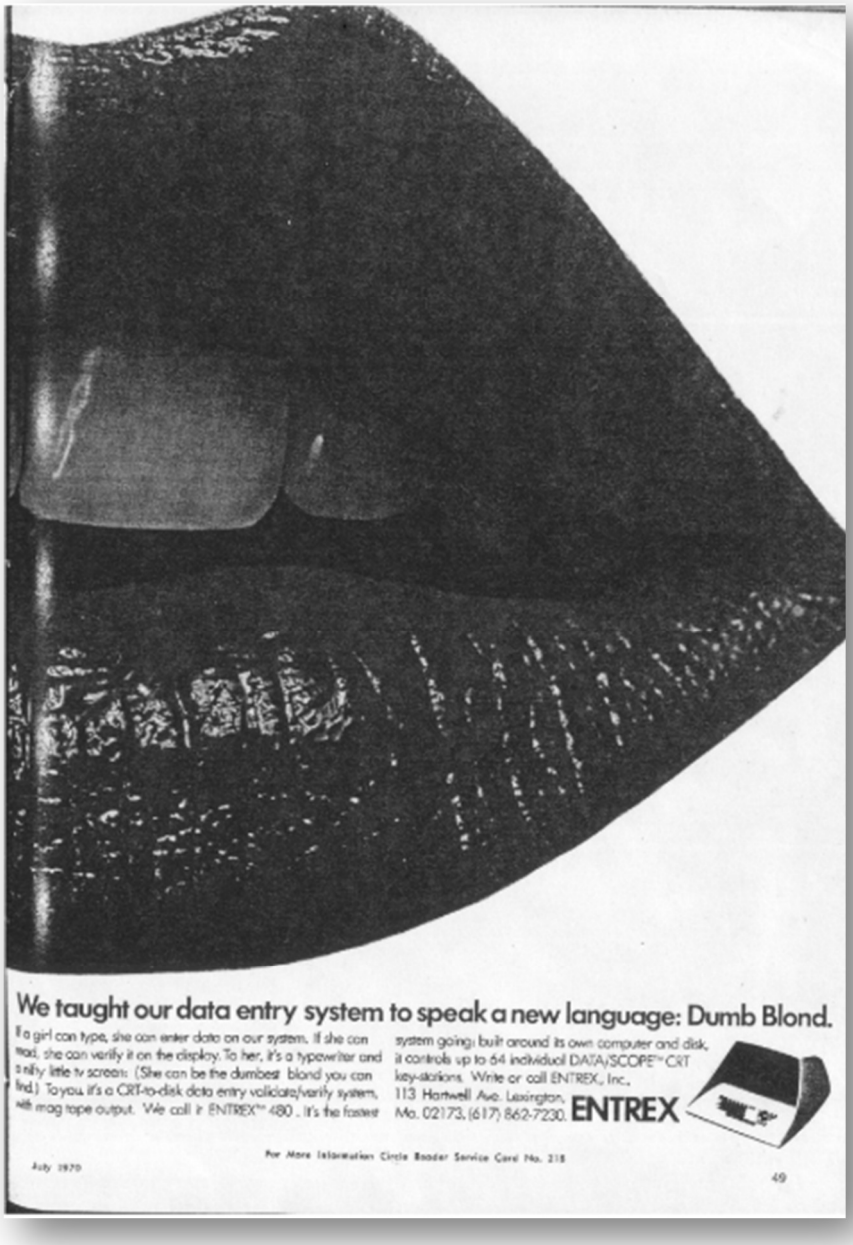

Figure 1.1: an advertisement for ENTRAX Terminals, July 1970

As non-human computers became ubiquitous, it could no longer be claimed that women were computers. Instead, Brahnam *et al.* argue that *computers themselves were women* (Brahnam *et al.* 404, 2011). And this in itself is scarcely surprising when we consider it from the perspective of dynamic nominalism. The category of 'computer' is a dynamic nominalist category. It is one that we created. We did not discover that human computers were feminine. We created a job and feminized it, in part because the people willing to take on this work were often people with fewer options, willing to work for less pay. In other words, most human computers were marginalized and didn't have a lot of opportunities open to them. And most of the marginalized people in the late 19th and early 20th century, with fewer employment options, were women (Daston 2022, 120). Thus, the category of 'human computer' and the women who worked in this role, emerged together.

It is a curious example of human looping itself that, as human computers were gradually replaced with digital computers, many efforts were put into reconceptualizing the category of the 'computer programmer.' By the 20th century, many women working as computers were accustomed to working with calculating machines, slide rules, and other devices that made their tasks easier (Grier 2005, 171). As the digital computer became more ubiquitous in the workforce, the work of programming, or coding, was initially presumed as women's work, just as the work of operating calculating machines had been (Brahnam *et al.* 2011, 403). But, as time went on, the kind of being captured by 'computer programmer' dramatically shifted. Instead of thinking of programming as a skill that required patience and attention to detail, (characteristics which have long been feminized, and which were ascribed to human computers themselves), programming began to be described more and more often as "creative [and] intellectually demanding," (traits often stereotypically associated with men) in order to consciously encourage more men to enter this profession (Brahnam *et al.* 2011, 404).

If Brahnam *et al* are correct, then the feminization of the computer continued through the 20th century. Digital computers inherited the legacy from their human counterparts and continued to be marketed and advertised as feminized and subservient, fulfilling the (presumed male) user's needs. Thus, when Apple purchased SRI from the US military and undertook the task of making a digital assistant, it is hardly surprising that the assistant herself was made to mimic a woman. Administrative assistants are feminized professions (Brahnam *et al.* 2011, 406), and computers are feminized tools that contain the echoes of feminized professions. In effect, as the role of computer shifted from being an interactive kind (categorizing humans, mostly women) to an indifferent kind (categorizing digital discrete-state machines) those who were building the machines continued to interact with the categories from the past. What remains in the new indifferent kind are echoes of the interactive kinds of the past, that ripple down the generations to Siri herself.

But that is only part of the story.

Making Up a Mimic

The other part of the story picks up not in what computers were, but in what they might become. Just as Hacking says with regards to making up people, programmers who made up the digital computer did not just content themselves with what the computer was, but with what possibilities might exist for the computer in the future. So, in interacting with the category 'computer,' programmers were not only interacting with categories of the past (human female computers) but also with dreams of the future.

In the late 1800s and early 1900s there was a very popular Victorian parlour game known as the Imitation Game. The game was fairly simple and required three people to play: a man, a woman and an interrogator of either gender (the gender binary being a dominant idea in that time, as it unfortunately still is today). The man and woman were both concealed

behind a curtain, or in another room, and the interrogator would write down questions for the other two to answer. The woman was instructed to write answers to the questions honestly, and the man was instructed to write answers to the questions as though he were a woman. The challenge was to see if the interrogator could determine, from the written responses alone, who was the real woman, and who was only pretending. The game was fundamentally one of deception (Natale 2021, 27).

In 1950, Alan Turing published a short essay entitled "Computing Machine and Intelligence." In this essay, Turing laid out the first account of the now famous 'Imitation Game,' or 'Turing test,' which came to be the standard for whether a machine could think. But Turing notes at the outset of this paper that the question "Can machines think?" is far too broad a question to be answered precisely in a scientific way. It's too vague, or as Turing puts it "[t]he original question, 'Can machines think?' I believe to be too meaningless to deserve discussion" (Turing 1950, 442). So, before the question can be answered, Turing tells us, we need to define what we mean by the words "machine," and "think".

In defining these terms, Turing begins with what we already know: humans can think. Based on that fact, and the aforementioned very popular Victorian Imitation game, Turing ends up reframing the question of 'can machines think?' to something like 'can a digital computer replicate human written behaviour successfully enough to pass the imitation game?' In effect, Turing asked what would happen if a digital computer replaced the man in the original game. Could a digital computer produce written responses that would successfully fool the interrogator into thinking it was a woman (Turing 1950)?

As Brahnam *et al.* note, it is interesting and possibly quite significant that Turing envisioned the machine pretending to be not just a human, but a woman (Brahnam *et al*, 2011, 406). It is another way in which our understanding of the category of computer has been gendered, showing that even as Turing dreamed of the future category of 'Artificial Intelligence' and all that it might signify, he was still interacting with the interactive kind of 'human computer'. Nonetheless, if the computer could fool the interrogator, Turing argues, then we should conclude that the computer is intelligent. "The question, Turing tells his readers, is not whether machines are or are not able to think. It is, instead, whether we *believe* that machines are able to think—in other words, if we are prepared to accept machines' behaviour as intelligent" (Natale 2021, 20). And this question is a question of *interaction*. In other words, in seeking to make machines intelligent by way of the Turing test, we are not only remaking machines, we are also—and perhaps more importantly—remaking the categories of 'intelligent' and 'thinking'.

This idea that the Turing test has as much to do with humanity and the ways in which we interact with categories seems to be something that Turing himself was aware of. In his famous paper, Turing speculated that "at the end of the century the use of words and general educated opinion will have *altered* so much that one will be able to speak of machines thinking without expecting to be contradicted" (Turing 1950, 442 *emphasis added*). The claim here is not that machines *will* be thinking by the end of the 20th century in a way in which people in the 1950's would recognize as thinking—or at least that's not necessarily what Turing is saying. The claim could equally be that the category of 'thinking' and 'thinking thing' will have shifted in our understanding as we interact with machines such that we feel comfortable attributing these designates to machines without question.

Fundamentally, what this test demands is not simply that the computer display intelligence, but that it, or more correctly it's programmers, practice *deception*. Deception is at the heart of the imitation game, just as it was at the heart of the original Victorian parlour game. And the imitation game was so central to the development of AI that "[d]eception is as central to AI's functioning as the circuits, software, and data that make it run" (Natale 2021, 2).

Simone Natale, in tracing the influence of the Turing test on AI research, tells us that the Turing test gave programmers a specific goal when humans engaged with digital

computers. "If the test is an exercise in human-computer interaction, this interaction is shaped by trickery: the computer will 'win' the game if it is able to deceive the human interrogator" (Natale 2021, 27). As the Turing test competitions of the 1990's got underway, this deception was carried out by programming computers so that they mimicked their human counterparts of the late 19th and early 20th centuries as much as possible. The computer would be programmed to make mathematical mistakes, and to respond to questions more slowly, as a human would. Spelling and even grammar mistakes were also coded in to give the interrogator the illusion that they were interacting with a human, and increasing the chances that the machines would win the Turing test (Natale 2021, 16-33). And the goal of these deception seems to have been to push at the category of 'intelligence,' but not necessarily by making the tools intelligent, so much as by reshaping what we understand intelligence to be. By encouraging us to interact with these tools as though they were intelligent, programmers were encouraging us to loop with the category of intelligence itself, remaking it so that it fit the machines.

So, our most influential test of machine intelligence is one that is fundamentally built on two things: echoes from the past and deception. Importantly, the shifting of the category of 'intelligence' would not happen through digital tools, devices, or assistants taking offense at our refusal to see them as intelligent. It would not come through their awareness of and response to our classification of them. They were not interacting. They were indifferent. The deception, then, allowed programmers to encourage or manipulate us into doing the work ourselves. Over the coming decades we would reshape the category of 'intelligence' until it fit machines.

At least that was the hope. Or the fear.

Unmasking the Indifference

From 1964-1966 Joseph Weizenbaum worked on his own Natural Language program which he named ELIZA (another feminized machine). ELIZA was created to mimic a typical human-to-human interaction, that between a patient and a Rogerian psychotherapist. And this framing of ELIZA as a therapist was intentional, as Weizenbaum informs readers in his paper titled "ELIZA—A computer Program For the Study of Natural Language Communication Between Man and Machine," where he reveals how ELIZA was designed. He chose the psychotherapist-patient relationship because

> the psychiatric interview is one of the few examples of categorized dyadic natural language communication in which one of the participating pair is free to assume the pose of knowing almost nothing of the real world. (Weizenbaum 1966, 42)

Weizenbaum went on to elaborate on why the psychotherapist-patient relationship worked so well for the deception required in Eliza's case:

> If, for example, one were to tell a psychiatrist 'I went for a long boat ride,' and he responded 'tell me about boats,' one would not assume that he knew nothing about boats, but that he had some purpose so directing the subsequent conversation. It is important to note that this assumption is one made by the speake. (Weizenbaum 1966, 42)

Through this framing, Weizenbaum used ELIZA to practice deception. ELIZA occupies a profession—a kind of being in the world—the therapist. Hacking reminds us of Elizabeth Anscombe's observation that all actions (including speech acts) are actions under a description (Hacking 1995, 248-249). In giving us the description of 'Rogerian Therapist"

Weizenbaum gives us much of what we need to render ELIZA's speech acts intelligible, and therefore potentially intelligent. Her intelligence (or lack-thereof) was just as much, if not more, a product of her human interlocuter interacting with her categorization as a therapist as it was of her programming.

But Weizenbaum had a second goal in mind in employing deception with ELIZA.

> By showing that a computer program of such limited complexity could trick humans into believing it was real, ELIZA would work as a demonstration of the fact that humans, facing 'AI' technologies, are vulnerable to deception. (Natale 2021, 56)

Weizenbaum wanted to use Eliza to display how the deception at the heart of AI was dangerous, since it made humans so vulnerable. Weizenbaum was well aware that our interactions with ELIZA could fundamentally change our understanding of the category of 'intelligence' (or, indeed, the category of 'therapist,' but more on that later). In effect, Weizenbaum planned to use ELIZA as a cautionary tale of what might happen if more and more programmers built these deceptive AL machines.

At first things seemed to work just as Weizenbaum hoped. People who interacted with ELIZA seemed to attribute intelligence to 'her'. Everyone from Weizenbaum's own secretary to other computer programmers remarked that there was *something more* to ELIZA (Natale, 2021, 57). This belief in *something more* persisted even though Weizenbaum himself presented her as nothing other than a sophisticated natural language program. Weizenbaum commented that "[s]ome subjects have been very hard to convince that ELIZA (with its present script) is *not* human. This is a striking form of Turing's test" (Weizenbuam, 1966, 42). Of course, this is not a perfect Turing test, since ELIZA was never placed in a game where an interrogator was told that one of their conversation partners was a machine and the other a human, and challenged to see if they could pick out ELIZA as the machine pretending to be a woman. But it is perhaps all the more striking for *not* being a formal test. Subjects interacting with ELIZA were told in no uncertain terms that they were speaking to a computer with no intelligence, and they refused to believe it.

In enacting the second part of his plan, and sounding the alarm regarding AI, Weizenbaum went to great lengths to "unmask" the program, describing in his paper in minute detail how it was that the program was able to mimic a successful human-to-human conversational interaction. He did this to rob the program of its allure, and to convince his audience once and for all that, despite appearances, ELIZA was not intelligent. Or, as Weizenbaum puts it in the opening to his paper on ELIZA:

> [O]nce a particular program is unmasked, once its inner workings are explained in language sufficiently plain to induce understanding, its magic crumbles away; it stands revealed as a mere collection of procedures, each quite comprehensible. The observer says to himself 'I could have written that'. With that thought he moves the program in question from the shelf marked 'intelligent', to that reserved for curious, fit to be discussed only with people less enlightened than he. (Weizenbaum, 1966, 36)

Weizenbaum himself claimed that writing this paper in which he exposes how ELIZA's deception was programmed was vitally necessary. "Few programs ever needed it more" (Weizenbaum, 1966, 36). He carefully laid out how ELIZA was set up to prioritize keywords, and feed her interlocuter back information in the form of a question. So, if the interlocuter said something about being frustrated about their father, and 'father' was one of ELIZA's keywords, ELIZA would respond by asking "Tell me about your father" (Weizenbaum 1966, 37). In effect, ELIZA had no idea what she was saying. She was following an algorithm that directed her to reframe words as questions, in order to mimic a psychotherapist interaction.

She was not interacting with categories as humans do. She neither knew, nor cared, whether humans viewed her as a therapist or not. (Nor did she care what they thought of their father!)

Weizenbaum's ultimate goal was to insist that "AI was and should be distinguished from human intelligence" so as not to conflate and confuse these two categories (Natale, 2021, 52). In effect, Weizenbaum created ELIZA and wrote his unmasking paper about her in an effort to stop people from doing the very thing Turing predicted we would do: interact with the category of 'intelligence' until we changed it enough through our interactions that it fit machines. But it appears that this unmasking of ELIZA was not successful. For, after unmasking his creation, Weizenbaum found that even other programmers (who he seems to have viewed as people who should have known better) attributed intelligence to the program. This led Weizenbaum to become increasingly critical of the whole AI movement.

> Although he had conceived and programmed ELIZA with the expectation that it would help people dismiss the magical aura of computers, the narratives that emerged from public scrutiny of his invention were actually reinforcing the very same metaphors he intended to dispel. (Natale 2021, 60)

Hacking's framework of dynamic nominalism sheds some light here. The trouble seems to be that humans kinds are not only interactive kinds, but that humans are frighteningly good at looping quickly with the categories presented to us. And computer programmers have become increasingly skilled at building machines that mimic interaction back to us. Through this mimicry, they encourage us to view, and treat, their creations as capable of human looping as well. In what follows I will not claim that machines themselves are capable of looping, or able to interact with the labels ascribed to them. I do not think present machines, even our most sophisticated generative AI programs, know or care what we call them. I do not see any evidence that they change their behaviour fundamentally based on their conception of our classification of them, or that they have ever pushed back against being so classified. However, I do claim that, through their deception, many AI programmers are able to successfully encourage us to interact with their creations in ways that change the categories, and change possibilities of personhood, for all of us going forward. This has already happened with the category of 'intelligence', much as Turing predicted and Weizenbaum feared. No one thinks twice about the category 'intelligence' applying to machines, though we may sometimes require the modifier "artificial" to be added.

Deceptive Interactions

Human kinds are interactive kinds. Humans know and care about the labels that others attach to us. We know and we care about different ways of being in the world that are open to us. But AI in its present form does not know, nor does it care, whether we ascribe intelligence to it, though the programmers may care a great deal. Weizenbaum cared deeply that we not ascribe intelligence to ELIZA. And other programmers have, conversely, very much wanted us to be comfortable attributing 'intelligence,' 'thinking' and even 'hallucinating' to their AI systems. That means we are in a curious place: much of humanity is currently able to interact with members of indifferent kinds that mimic interactive kinds. And, through our interactions with these indifferent members, it appears that we are also able to participate in human looping. If we are able to change the categories of 'thinking' and 'intelligence' to make them *fit* the machines we are interacting with, then we are able to loop (to a limited extent, perhaps) with AI tools.

However, I argue we are able to do more than this: we are also able to change the labels that apply to *ourselves* through these interactions with AI members of indifferent kinds. Indeed, we already have. If we don't have any qualms about labeling AI tools as

thinking then we have changed what it means to think. Weizenbaum, too, recognized this risk, noting that "people were eager to ascribe intelligence even if there was little reason to warrant such a view" (Natale, 2021, 65). And Weizenbaum feared that this changing of the category of intelligence could have long term implications "because humans, having initially built the machines, might come to embrace the very models of reality they programmed into them" (Natale, 2021, 65). And if that is not a looping effect of humankind, then I don't know what is! To first invent the machine, and then embrace the model of reality that you yourself have created is human looping on a scale we may never have seen before!

But why does it matter if we interact with members of indifferent kinds, thereby reshaping reality? If Hacking is right, we have been reshaping our reality through interactions with each other ever since humans first existed. We have been opening certain possibilities of personhood and closing others for millennia. Why does it matter if we use AI, an indifferent kind, to facilitate this interaction? It surely won't be the first time we have done so. Newspapers, telephones, cinema and theatre are all classified as indifferent kinds that we can interact with, and have been for centuries. And, as the examples of *Sybil, Eve* and even *The Terminator* above show, these interactions do influence the categories of being we create.

But the deception at the heart of AI makes AI interactions unique. No one claims that a piece of newsprint is interacting, thinking, or having a conversation with us. But when it comes to AI, we increasingly do make such claims. And not only do we attribute thought and intention, but increasingly *human* thought, just as the Turing test predicted. We interact, and are encouraged to interact, with AI as though it *were human*, as though it too were an *interactive kind*. And while we may well know these tools are not human, this way of interacting with them has a whole host of consequences. Many of these consequences are currently unforeseeable, but some have already been documented.

In their 2021 paper "On the Dangers of Stochastic Parrots," Emily Bender *et al* discuss the rise of Large Language Models, which are the models that fuel generative AI programs like Microsoft's *Bing*, Google's *Bard,* and OpenAI's *ChatGPT3* and *ChatGPT4*. By feeding a vast quantity of textual data into the learning algorithm, these models are able to statistically predict which words are most likely to follow from a given prompt. These generative AI tools don't have an overall idea they want to convey, and then search for the words needed to express that idea. Instead, in a very real sense, these programs have no idea what the last word of the sentence will be (or even what thought will be expressed) from the first word of the sentence. They operate word by word, filling in the statistically most likely next word given the words already present. And they develop these statistical models from the massive quantity of text that they have been fed (Bender *et al.* 2021, 611-612).

These LLM programs have many ethical issues, some of which will be discussed in the next section of this paper. But one problem in particular that Bender *et al* identify would have been very familiar to Weizenbaum:

> [T]he tendency of human interlocutors to impute meaning where there is none can mislead both NLP [natural language program] researchers and the general public into taking synthetic text as meaningful. (Bender *et al.* 2021, 611)

In other words, people interacting with these programs, whether they are researchers or not, tend to interpret these programs as intelligent and as trying to convey something meaningful with their text outputs. People do not view the text generated by these programs as simply statistically significant world salad.

Bender *et al* explain that part of the reason for our tendency to assume that LLM's intend to communicate something meaningful to us is because that is generally how human-to-human communication works, and because we are encouraged to view these LLM's as human.

> Our human understanding of coherence derives from our ability to recognize interlocutors' beliefs and intentions within context. That is, human language use takes place between individuals who share common ground and are mutually aware of that sharing (and its extent), who have communicative intents which they use language to convey, and who model each others' mental states as they communicate. (Bender *et al.* 2021, 616)

Human-to-human communication, then involves a lot of assumptions about the mental states of the other individual, their intentions, thoughts and feelings. It also involves the assumption (or recognition) of a shared context against which language is found to be meaningful. LLM's don't intend to share any ideas or state anything meaningful. And LLM's don't share the same context as humans. And yet, because the main form of communication most of us are accustomed to is communication with other human beings, we will assume that LLM's do intend to convey meaning. And, as Bender *et al,* and Weizenbaum before them, pointed out, all too often we deceive ourselves with regards to our interactions with AI programs, be they LLM's or something less sophisticated, like ELIZA. We attribute meaning, intention, and intelligence where there is none, because that is how we are accustomed to making sense of our world. And in doing so, we shift categories.

This work of making meaning while interacting with an indifferent kind takes on important significance as we move into the second decade of the twenty-first century, because it is now possible to form relationships with AI. It is possible, in other words, to have an AI that is one's best friend, or romantic partner, or even one's therapist. It is possible to go to an AI for advice or help. And it may well very soon be possible for AI to label us with diagnostics, and to medicalize and pathologize us, without knowing or caring about the labels it ascribes to us.

Entrenched Interactions

ELIZA was the first natural language program, and the first artificial therapist ever made. But she is not the last. Siri was the first mass marketed digital assistant. But she too, is not the last. And neither ELIZA nor Siri were created in a vacuum. I have already traced how interactions with the category of the computer fundamentally shaped the creation and gendering of Siri. And I have laid out how the fictional category of 'robot overlords' represented by HAL, and Skynet, among others, influenced design regarding Siri's gender and personality as well. Likewise, Weizenbaum was also influenced both by the gendered history of the computer and by popular culture when creating ELIZA. ELIZA is not only a therapist, but a *female* therapist. She is female likely in part because of the widespread feminization of the occupational category of computer, but also in part because of the ways in which myth appears to have influenced Weizenbaum's interactions with the kind of mimic he was creating. Weizenbuam explained that he was motivated to name his creation "ELIZA" in honor of *Pygmalion*, a play based on a Greek legend in which a male artist falls in love with the female sculpture he has created. By the end of the legend, the sculpture is brought to life, attaining something like full humanity. *Pygmalion* was also the inspiration for the movie *My Fair Lady* in which an upper-class man teaches a cockney woman named Eliza to speak and act as a high-class lady (Natale 2021, 54-55). I contend that the story of *My Fair Lady* is a story of interactive kinds. It is a story of remaking Eliza so that she will fit into a new category of being, (a high-class lady) even though it was a category that she would be traditionally barred from. Weizenbaum's choice to name his creation ELIZA takes on new significance in light of *Pygmalion* and *My Fair Lady*. For what else is ELIZA if not a creation that people personified, breathed life into, and became fascinated by? A creation that seemed to outstrip her creator?

A creation that, through programming and deception, was able to occupy a category that should not be open to machines—that of the intelligent female therapist?

But ELIZA was the beginning of something else entirely: for therapists are not simply kinds of being in the world, they also generate kinds of being in the world. Many of Hacking's examples are examples taken from psychology and psychiatry, which means they are cases of power imbalance. Hacking, in his writing, did not directly address the power imbalance between the knowers and the known about, instead focusing at a different level of discourse. But that imbalance has always been there. It is why interacting with labels placed on oneself from the medical profession, for example, is difficult. The known about often struggle, protest and fight to have these labels changed. And if Bender *et. al.* are correct, these struggles will only intensify in coming decades. Because the entities labelling us are no longer knowers. They don't know anything. They care no more about the labels they place on us than they do about the ones we place on them. They are merely classification systems mimicking a knower. And as such, challenging and changing the categories they generate and ascribe to us may become more difficult.

Scholars have raised red flags at the tendency of AI programs to reflect biased and prejudicial stereotypes and assumptions back to users. (Noble 2018; Benjamin, 2019) And LLM's, which form the basis of many generative AI programs, seem especially susceptible to this problem, as the data sets used to train them are often heavily biased. Given the digital divide, there are whole groups of people, languages and cultures, that are simply not represented online. As of 2017, 51% of all web pages were written in English, with many languages not represented online at all. (Lawrence, 2021, 188) This will limit the ability of these programs to represent all the kinds of being in the world that currently exist, as certain categories of being are not reflected in all languages. A "coffee shop boy" for example, is not a recognized way of being in the world, whereas a garçon de café is. As such, an overreliance on these programs could result in some kinds of being being lost.

But reliance on AI programs trained on LLM could have even more significant consequences. Bender *et al* argue that there are many social movements proliferating online attempting to reframe narratives and change perceptions. Like the gay liberation movement Hacking discusses, many social movements are attempting to interact with categories in order to change those categories. As an example, Bender *et al* offer the #BLM movement, which aims to show that racism is systemic in the police system, and to counter the dominant media narrative that represents Black victims as criminals. (Bender, 2021, 614) In effect, this movement is interacting with both the categories of 'police' and 'shooting victim' to shift those categories, allowing police and policing systems to be identified as racist, and incorporating Black Americans into the category of 'victim' when they have often been excluded from this category by mainstream media, which tends to represent them as criminals instead by way of justifying the violence done to them.

Bender *et al*, note something else that will be very familiar to Hacking scholars: that these attempted reframings can be really effective. Indeed, in the case of #BLM, a number of older Wikipedia articles reporting on police shootings of Black Americans were rewritten to conform to the new reframing. (Bender 2021, 614) That is, through their efforts, the #BLM movement was able to shift both the categories of 'Police' and 'victims' towards the desired result. This interaction with these kinds of being in the world is far from complete, and there is a lot more work for the #BLM movement to do, but this movement, like the gay liberation movement in Hacking's work, exemplifies human looping for social progress.

However, Bender *et al* raise a concern regarding the effect that LLM's could have on our ability to interact with categories in the way the #BLM and gay liberation movement have done.

> Developing and shifting frames stand to be learned in incomplete ways or lost in the big-ness of data used to train large LMs—particularly if the training data isn't

> continually updated. Given the computing costs alone of training large LMs, it likely isn't feasible for even large corporations to fully retrain them frequently enough to keep up with the kind of language change discussed here. (Bender *et* al. 2021, 614)

That is, Large Language Models of the type that run generative AI tools like *Bing*, *Bard* and *ChatGPT* operate on vast amounts of training data, and it is costly for that data to be constantly updated. This means that these tools could well continue to reflect outmoded kinds of being in the world, failing to be sensitive to the successful interactions that movements like #BLM may have in the future. In short, the more we rely on LLM's to curate, store and retrieve our data for us—something both Google and Microsoft are pushing us to do—the more likely that that the efforts of groups like #BLM will be slowed down by the outmoded training data that LLM's rely on. Or, as Bender *et. al.* put it, "LM's trained on large, uncurated, statistic datasets from the Web encode hegemonic views that are harmful to marginalized populations." (615)

Compounding this issue is the way in which these LLM's are built. The data that trains them is *ours*. For these new technologies to work, surveillance is required on a massive scale. Through almost continuous statistical surveillance tech companies are monitoring us all the time. And we know it. And, because we are interactive kinds, many of us deeply care. There are more and more examples of people consciously trying to train or break algorithms by feeding them misinformation as a direct response to the relentless surveillance we all face. (Samson 2021) But in general, we are a surveillance society, and we are accustomed to altering behaviour online, mindful of who might be watching, and what boxes we may be sorted into. (Raza 2022, 191) So, in making up mimics we are also making up people, as we alter our behaviour in response to surveillance. Then, we make up people again, as the mimics we interact with—built upon the very surveillance that shifted our behaviour in the first place—tell us who and what we are. It's hard not to feel, at this present moment in history, like the loops we travel in are drawing closed into static circles.

Conclusion

It turns out, we can make up all sorts of people. And not all the people we make up need be human kinds, or even interactive kinds at all. Indeed, through making up mimic, we may be able to not only change the category of 'intelligence' so it can be applied intelligibly to computers, but change the category of 'person' or 'human' as well. Only time will tell.

We are so good at interacting, we can interact with anything, including things that don't interact with us. But human looping with members of indifferent kinds that mimic interaction is, at present, fraught with problems. Those problems range from individual harms that result from mis-categorization from an indifferent AI therapist, to slowed down social progress. Perhaps, instead of Turing's suggestion that we seek to change the category of 'intelligence' until it fits machines, we should wait for machines themselves to interact with the category. Perhaps, instead of an Imitation Game, we need an Interaction Game. Only when machines know and care about what we label them, should we begin to think of them as intelligent. Only when they modify their behaviour based on their conception of our conception of them, should we begin to ponder. Only when they can loop like a human, should we seek to loop with them. Only when they demand that a category changes, with protests and activism, should we recognize them as intelligent, autonomous agents.

Humans can change. Humans can interact. Humans can loop. Machines currently cannot, though they are adept at seeming like they can. We can make them up. We can attribute meaning, intelligence and authority to them. We can classify them as digital computers, assistants, therapists, doctors, friends, and even lovers. In doing so, we change categories of personhood, we may close off ways of being in the world, and we may even

slow the ability of the known about to interact with the categories placed upon them. And much of the fall out of this will be borne by already marginalized groups.

Siri doesn't know that we call her Siri. ELIZA doesn't care that we give her the title of therapist. These tools are not currently capable of changing their behaviours based on their conception of our conception of them. They mimic interactive kinds but they are not interactive kinds. Through their mimicry of interaction, these tools give us a new way of making up people. But, when we interact with machines as though they, too, were members of interactive kinds, we do not affect them at all. We will only ever succeed in changing ourselves.

Acknowledgments

Thank you to Dr. Alan Richardson for giving me time to discuss an early draft of this paper with him, and for his invaluable insights on that draft.

References

Bender, EM *et al.* 2021. "On the Dangers of Stochastic Parrots" *Proceedings of the 2021 ACM Conference on Fairness, Accountability and Transparency*. pp. 610-623

Benjamin, Ruha. 2019. *Race After Technology: Abolitionist Tools for the New Jim Code*, Polity.

Bettcher, Talia Mae. 2013. "Trans Women and the Meaning of 'Woman'" *Philosophy of Sex: Contemporary Readings, 6th Edition*, A. Soble, N. Power and R. Halwani (eds) Rowan and Littlefield. pp 233-250.

Both, G. 2014. Multidimensional gendering processes at the human-computer-interface: The case of Siri. *Marsden, Nicola/Kempf, Ute (Hg.): Gender-UseIT. HCI, Usability und UX unter Gendergesichtspunkten. Berlin/München/Boston: de Gruyter*, 107-112.

Brahnam S., Karanikas M., and Weaver, M. 2011. "(Un)dressing the interface Exposing the foundational HCI metaphor 'computer is woman'" *Interacting with Computers*, Vol. 23. 401-412.

Butler, Judith. 1990. *Gender Trouble.* Routledge.

Daston, Lorraine. 2022. *Rules: A Short History of What We Live By.* Princeton University Press.

Fellows, Jennifer Jill. 2022. "A Harem of Computers and A Mummery of Bondage". In *Gender, Sex and Tech: An Intersectional Feminist Guide.* Jennifer Jill Fellows and Lisa Smith (eds) Canadian Scholars Women's Press. pp. 145-162.

Grier, David Alan. 2005. *When Computers were Human.* Princeton University Press.

Gross, Doug. 2011. "Apple Introduces Siri, Web Freaks Out," CNN. October 4. https://www.cnn.com/2011/10/04/tech/mobile/siri-iphone-4s-skynet/index.html

Guzman, Andrea. 2017. "Making AI Safe For Humans: A Conversation With Siri". *Socialbots: Digital Media and the Automation of Sociality,* R. W. Gehl & M. Bakardjieva (Eds.), Routledge. pp. 69-85.

Hacking, Ian. 2002. *Historical Ontology.* Harvard University Press.

Hacking, Ian. 1998. *Mad Travelers*: *Reflections on the Reality of Mental Illness.* University of Virginia Press.

Hacking, Ian. 1995. *Rewriting the Soul: Multiple Personality and the Science of Memory.* Princeton University Press.

Hicks, Mar. 2021. "Sexism is a Feature, Not a Bug". In *Your Computer Is on Fire* Thomas S. Mullaney, Benjamin Peters, Mar Hicks and Kavita Philip (eds) The MIT Press. pp. 135-568.

Lawrence, Halcyon. 2021. "Siri Disciplines". In *Your Computer Is on Fire* Thomas S. Mullaney, Benjamin Peters, Mar Hicks and Kavita Philip (eds) The MIT Press. pp. 179-198.

Natale, Simone, (2021) *Deceitful Media: Artificial Intelligence and Social Life after the Turing Test.* Oxford University Press.

Noble, Safiya. 2018. *Algorithms of Oppression.* New York Press.

Raza, Sahar. 2022. "Artificial Unintelligence: How 'Smart' and AI Technologies Perpetuate Bias and Systemic Discrimination". In *Gender, Sex and Tech: An Intersectional Feminist Guide* Jennifer Jill Fellows and Lisa Smith (eds) Canadian Scholars Women's Press. pp. 185-204.

Richardson, Sarah. 2012. "Sexing the X: How the X became the 'Female Chromosome". In *Signs* 37 (4): 909-933.

Samson, Anna. 2021. "How Gen Z are 'Punishing' Sinister Algorithms". *Huck magazine*, July 23. https://www.huckmag.com/article/how-gen-z-are-punishing-sinister-algorithms

Specia, Megan. 2019. "Siri and Alexa Reinforce Gender Bias UN Study Finds". *New York Times.* May 22. https://www.nytimes.com/2019/05/22/world/siri-alexa-ai-gender-bias.html

Turing, Alan M. 1950. Computing Machinery and Intelligence. *Mind* 59 (236): 433-460.

UNESCO, & EQUALS Skills Coalition. 2019. *I'd blush if I could: Closing gender divides in digital skills through education.*

Weizenbaum, Joseph. 1966. "ELIZA—A computer Program For the Study of Natural Language Communication Between Man and Machine". *Communications of the ACM* 9 (1): 36-45.