

Conversing with Personal Digital Assistants: on Gender and Artificial Intelligence

Pedro Costa

Faculty of Fine Arts, University of Lisbon,
Lisbon, Portugal

pedro.carv.c@gmail.com

ABSTRACT

This paper aims to explore the relationship between gender and artificial intelligence, seeking to understand how and why chatbots and digital assistants appear to be mostly female. To this end, it begins by addressing artificial intelligence and the questions that emerge with its evolution and integration in our daily lives. It then approaches the concept of gender in light of a binary framework, focusing on femininity. These topics are then related, in order to shed some light on how chatbots and digital assistants tend to display feminine attributes. In an attempt to observe these aspects, an analysis of Alexa, Cortana and Siri is developed, focusing on their anthropomorphization, the tasks they perform and their interactions. Complementing this discussion, the project *Conversations with ELIZA* is presented as an exploration of femininity in AI, through the development of four chatbots integrated into a web-based platform, each performing specific tasks and simulating particular personalities, with the purpose of emphasizing feminine roles and stereotypes. In this manner, this study aims to understand and explore how gender relates to AI, why femininity seems to be often present in AI and which gender roles or stereotypes are reinforced in this process.

KEYWORDS

Artificial Intelligence; Chatbots;
Anthropomorphization; Gender; Femininity;
Stereotypes.

ARTICLE INFO

Received: 15 November 2018

Accepted: 30 November 2018

Published: 18 December 2018

<https://dx.doi.org/10.7559/citarj.v10i3.563>

1 | INTRODUCTION

Artificial intelligence has already become part of our daily lives. In fact, we frequently interact with AI systems without even realizing it, namely with personal digital assistants or chatbots that are embedded into mobile devices and online contexts. It is becoming more and more common to find a digital assistant to help us in specific tasks (Dale, 2016) while operating in a kind, helpful and compliant fashion.

These assistants are often assigned human-like traits or features, but this process of anthropomorphization seems to be accompanied by a tendency towards feminization. Beyond their names, voices or avatars, they also perform tasks that are historically associated to women's labor. As such, they seem to behave according to gender stereotypes and reinforce traditional assumptions of femininity (Weber, 2005; Hester, 2016).

This study discusses traditional notions of gender and their significance in AI. To this end, it begins by addressing artificial intelligence and its integration in our daily lives. It then approaches the concept of gender in light of binary frame promoted by an historically patriarchal and heteronormative society

(Butler, 1990; Haraway, 1991). Artificial intelligence and gender are then related in order to shed some light on how chatbots and digital assistants are portrayed as gendered entities.

Complementing this discussion, digital assistants such as Alexa, Cortana and Siri are analyzed, in terms of their anthropomorphization, the tasks they perform and their social interactions. Drawing on this analysis, the project *Conversations with ELIZA* explores this topic through the development of four chatbots, each endowed with a specific personality with the purpose emphasizing feminine roles and stereotypes.

2 | ARTIFICIAL INTELLIGENCE

“AI currently encompasses a huge variety of subfields, ranging from the general (learning and perception) to the specific, such as playing chess, proving mathematical theorems, writing poetry, driving a car on a crowded street, and diagnosing diseases. AI is relevant to any intellectual task; it is truly a universal field.” (Russell and Norvig, 2010, p. 1)

Artificial intelligence is increasingly the subject of different areas of study, given its growing integration into our daily lives. We now have a more direct contact with this type of technology, namely by interacting with chatbots that play the role of personal digital assistants, which are embedded into our devices and become a natural part of the asynchronous simultaneous conversations we carry out. As Robert Dale points out, “chatbots have been around for a long time” and are thus returning, instead of emerging as something new (Dale, 2016, p. 814).

2.1 CONVERSING WITH ELIZA

ELIZA was one of the first “natural language process applications” capable of analyzing written inputs from its users and answer accordingly by using a set of rules, thus establishing a “human” dialogue [1]. Introduced in 1966 by Joseph Weizenbaum, this computer program “seemed to [fool]” some of its users “into thinking ELIZA was a person rather than a machine” (Dale, 2016, p. 814), by acting like a Rogerian psychotherapist. The behavior was “easy to imitate because much of this technique consists of drawing his patient out by reflecting the patient’s statements back to him” (Weizenbaum, 1976,

p. 3) [2]. The idea of having a machine talk to us as if it were human, leading people to believe they were speaking to another human being, conveyed the purpose of the Turing test [3].

By shifting human-machine interaction from a purely robotic, rational nature to a more social one, ELIZA marked a significant development in AI, influencing the way chatbots evolved towards emulating human behavior.

Drawing on ELIZA’s legacy, chatbots are now a natural component of our daily use of technologies, or “just another facet of today’s always-connected multi-tasking world, where we participate in multiple conversations in parallel, each one at a pace of our choosing” (Dale, 2016, p. 815).

2.2 SPECIALIZED AND GENERAL ASSISTANTS

“Interaction with technology using either natural language text or speech is becoming increasingly feasible, and potentially very significant. William Meisel (...) distinguishes ‘general personal assistants’ like Siri from the tsunami of more narrowly focused chatbots, which he calls ‘specialized digital assistants’.” (Dale, 2016, p. 812)

General personal assistants are usually integrated into our devices in order to assist us in a personalized way, like Alexa, Cortana or Siri. As such, they can help with “some subset of the standard virtual assistant skill portfolio”, which mainly includes reading, writing, sending emails, scheduling meetings, checking calendars and setting appointments, making calls, sending messages, taking notes, setting reminders, etc. (Dale, 2016, p. 812) [4].

In turn, specialized digital assistants are normally present in web-based platforms or apps and “operate in very specific domains or help with very specific tasks”, ranging from “booking a flight, buying some shoes, taking issue with a parking fine” to sending daily weather forecasts, helping with online shopping payment processes or even just telling jokes (Dale, 2016, pp. 812-813).

By performing those tasks, chatbots work towards an amelioration of our daily life, assuring that nothing is left unorganized, forgotten or undone; they make sure that we are as productive as possible by “promoting efficiency, transparency,

certitude and perfection” – and, by extension, rooting out “any imperfection, ambiguity, opacity, disorder and opportunity to err, sin or do the wrong thing” (Morozov, 2013b, pp. 14-15). This need to ameliorate our life and maximize production corresponds to Evgeny Morozov’s concept of solutionism, defined as “an intellectual pathology that recognizes problems according to just one criterion: whether they are ‘solvable’ with a nice and clean technological solution at our disposal” (Morozov, 2013a). Conforming to this idea, personal digital assistants are thus at our disposal, constantly present and ready to help us quickly solve our problems, while tracking our habits and user preferences, leaving little to no room for imperfection – all of this just a touch away.

2.3 FROM ASSISTANTS TO COMPANIONS

“The robotic scientists (...) did not speak in class-marked categories of their robots as ‘servants’ or ‘workers’ (...); they are imagined to help fill the gaps in human social relations and (...) to become friends and companions.” (Richardson, 2015, pp. 12-15)

Artificial intelligence is simultaneously ubiquitous and subtle, as it becomes embedded into our cellphones, laptops or tablets, as well as websites, apps or other types of web-based services. According to this growing presence, chatbots are no longer conceived or seen as mere tools, but rather conforming to a sense of companionship that develops alongside with the anthropomorphization of artificial intelligence. They are endowed human attributes or traits, as they evolve from assistants to companions that become closer to us.

Anthropomorphization takes place on a superficial level, pertaining to their names, voices, avatars, or other kinds of attributes that move away from a purely mechanized presentation, but it also develops at another level, concerning their dialogue and interactions. In this sense, Jutta Weber considers that anthropomorphization entails a significant shift from rational-cognitive processes and problem solving to a socio-emotional interaction, which emphasizes this intention of turning our interaction with this type of machines into a more social one (2005, p. 209). Therefore, this type of technologies are “supposed to mimic or even learn those abilities and characteristics which were, until recently,

regarded as purely and typically human and beyond the grasp of machines” (Weber, 2005, p. 213).

Although this anthropomorphization has become more evident among the current myriad of personal digital assistants, it goes back to ELIZA and the fact that “people were conversing with the computer as if it were a person who could be appropriately and usefully addressed in intimate terms” (Weizenbaum, 1976, p. 7).

2.4 EMOTIONAL INVOLVEMENT WITH AIS

“I was startled to see how quickly and how very deeply people (...) became emotionally involved with the computer and how unequivocally they anthropomorphized it. (...) I knew that people form all sorts of emotional bonds to machines (...) I knew from long experience the strong emotional ties many programmers have to their computers (...); what I had not realized is that extremely short exposures to a relatively simple computer program could induce powerful delusional thinking in quite normal people.” (Weizenbaum, 1976, pp. 6-7)

Weizenbaum observed that when we interact with machines as if they were human, we start developing emotional bonds, a sense of empathy and of being understood. He described this illusion he considered dangerous, as the ELIZA effect:

“the susceptibility of people to read far more understanding than is warranted into string of symbols strung together by computers (...) and the idea that computers “understand” the physical world, reason abstractly, make scientific discoveries, are insightful cohabiters of the world with us.” (Hofstadter, 1995, p. 157)

The notion of “computationalism” was proposed by Weizenbaum to question the belief that “the functional relations between mental inputs, outputs and internal states are computational” (Piccini, 2004, p. 814). Considering this to be a problematic idea, he argued that not every part of human thought could be reduced to logical formalisms, and that “there are some acts of thought that ought to be attempted only by humans” (Weizenbaum, 1976, p. 13). As such, users get the false sense they are talking to another human being when in reality

they're interacting with anthropomorphized digital assistants.

Consequently, human-machine interaction is influenced by feelings of intimacy, closeness and empathy as people start getting attached to these technologies and to the entities contained within them.

Adding to this, Dale discusses how digital assistants engage in conversations with us, evoking a not-so-far-away world “where some of those conversational partners we'll know to be humans, some we'll know to be bots, and probably some we won't know either way, and may not even care” (Dale, 2016, p. 815).

3 | GENDER AND STEREOTYPES

When chatbots are anthropomorphized, they tend to portray gender related features through their voices, names or even the way they interact. In order to understand this phenomenon, we need to take a closer look at gender as one of the aspects through which we socially develop and establish relationships, whether with each other or with ourselves.

Gender constitutes a part of our identity that regulates the type of behavior or acts we establish socially “by managing situated conduct in light of normative conceptions of attitudes and activities appropriate for one's sex category” (West and Zimmerman, 1987). In this sense, Judith Butler introduced the idea that gender has a performative nature, given that gender identity is a repetition of acts that are stylized through time, manifesting a “cultural interpretation or signification of that [biological] facticity” (Butler, 1990, p. 522).

3.1 BINARY FRAMEWORK

“One is not born, but rather becomes, a woman (...), and by extension, any gender, is an historical situation rather than a natural fact.” (de Beauvoir in Butler, 1988, pp. 519-520)

Simone de Beauvoir's words suggest how gender is not something we are born with but, instead, something we internalize through performative acts, over time. In sum, to be female or male is a matter of sex; but to be a man or a woman is a matter of gender. Gender is also perceived as something polar, as seen through a “binary framework” in which there is a “mimetic relation of gender to sex whereby

gender mirrors sex or is otherwise restricted by it” (Butler, 1990, p. 88) [5].

Consequently, there is a normalization of what is considered to be feminine or masculine behavior, which becomes predetermined in a foreclosed historically sedimented structure. This establishes a set of expected behaviors and we are expected to comply to “normative conceptions of appropriate attitudes and activities” that are determined by “institutionalized frameworks through which natural, 'normal sexedness is enacted'” (Goffman, 1977 in West and Zimmerman, 1987, p. 137). In other words, through this “need to routinize (...) behavior in accord with pre-established conceptualizations and behavioral patterns” (Deaux and Major, 1987, p. 370), certain attributes and acts are identified as specifically feminine or masculine and are supposed to imply someone's preferences and behaviors.

As Prentice and Carranza put it, “prescriptive gender stereotypes” define “the qualities [ascribed] to women and men (...) that are required of women and men” (2002, p. 269) [6]. These stereotypes imply that a gender belief system imposes expectations and gender behavior patterns, as internalized and socially reinforced stereotypes. Butler expands on this, stating that “gender performances (...) are governed by (...) punitive and regulatory social conventions” (Butler, 1988, p. 527) that reject the acts or behaviors that convey some kind of deviation from the norm.

3.2 GENDERED LABOR

“If, in doing gender, men are also doing dominance and women are doing deference (cf. Goffman 1967, pp. 47-95), the resultant social order, which supposedly reflects 'natural differences', is a powerful reinforcer and legitimator of hierarchical arrangements.” (West and Zimmerman, 1987, p. 146)

Gender roles and characteristics deemed as specifically feminine or masculine also imply a structural hierarchization of labor, which means that feminine and masculine behavior is also used to segregate and structure labor accordingly. The workplace and its relationships change since, as noted by Kelly, when we interact within these contexts “social labels, beliefs and attributions may serve as grounds for predictions and generate

behavior designed to validate or invalidate these beliefs” (in Snyder, 1977, p. 8).

In other words, by expecting certain acts (deemed as feminine) from women, we expect them to occupy jobs and perform tasks associated with these attributes, thereby creating a category of feminine labor. For example, a lot of service work is seen as feminized labor or “associated with qualities traditionally coded as feminine” (Hester, 2016, p. 47) [7], and historically women have a significant presence in the telecommunications industry, where they filled the role of assisting and establishing calls and communications, which “rendered female operators (...) [as] inferior, subordinate, and knowable” (Zost, 2015, p. 3). In other cases, women fill the role of secretaries, assistants, nurses or even flight attendants. These types of jobs convey, in a way, an “assumption that women possess a natural affinity for service work and emotional labour” (Hester, 2016, p. 47).

In turn, the heterosexual framework also accentuates labor hierarchization since it reinforces the “embodiment of wifely and husbandly roles, and derivatively, of womanly and manly conduct” (Beer, 1983, pp. 70-89 in West and Zimmerman, 1987, p. 144). This asymmetry affects the private sphere, namely domestic work. As West and Zimmerman explain, household and child care tasks are considered women’s work as a consequence of “normative conceptions of appropriate attitudes and activities for sex category” (West and Zimmerman, 1987, p. 139).

Additionally, and according to Donna Haraway, domestic work is transformed into capitalized labor out of the private sphere, through jobs such as office work, nursing or service work. Borrowing from Richard Gordon, Haraway considers that with new media, a “homework economy” emerges, defined as a “restructuring of work that broadly has the characteristics formerly ascribed to female jobs, jobs done only by women” (Haraway, 1991, p. 304) [8].

Therefore, even outside the domestic sphere, women still ensure domestic tasks “partly as function of their enforced status as mothers” and because they work in an “integrated circuit (...) in advanced industrial societies [where] these positions have been restructured (...) by social relations mediated and enforced by the new technologies” (Haraway, 1991, pp. 305-307). This reflects traditional

conceptions of gender derived from a patriarchal heteronormative society where women perform domestic and assistant-like roles, while it also reveals how gender standardization and normalization has implications at a social, personal and structural level.

4 | GENDERED AI

“It is clear that many of today’s apps and automated systems draw upon pre-existing gendered assumptions, programmed as they are to be girlish avatars or feminized disembodied voices. They exploit our assumptions about feminized labour and our existing relationship to socially gendered caring and service behaviors, tapping into those elements of femininity that have historically enabled care giving or service-providing subjects to better undertake specific obligations, activities, and tasks.” (Hester, 2016, p. 50)

As the tasks performed by chatbots begin to mirror traditional women’s labor, we witness “the protocols of femininity being programmed into machines” (Hester, 2016, p. 48). Halberstam explains how this “gender automation” takes place, given that tasks traditionally and historically considered female become a part of technology (Halberstam, 1991, p. 451).

Accordingly, we can observe how general or specialized chatbots automate work that is coded as female, given that they mainly operate in service or assistance related contexts, acting as personal assistants, secretaries and the like [9].

By operating in contexts of service, and by following these standardized behaviors, chatbots also end up emulating attitudes that resemble what Gustavsson calls a “stereotyped image of female service providers” (in Hester, 2016, p. 47). They display feminine attributes because these characteristics have their “basis in the stereotyped image of female qualities” and, consequently, “such a stereotypical female image of caring, empathy and altruistic behavior has become a standard component in a service script” (Gustavsson, 2005, p. 402 in Hester, 2016, p. 47).

4.1 GENDER AUTOMATION

"It is not by accident that social robotics is working with sociological and socio-psychological approaches that explicitly use gender dichotomies and stereotypes; (...) in the realm of human interaction it is regarded as helpful to use emotions to influence users, to direct the intentionality of others and to smooth interactions. (...) [And] this and other models rather point to the fact that gender stereotypes are instrumentalised in order to build "better" machines that are perceived as socially intelligent." (Weber, 2005, p. 214)

It is not only through the human attributes they displayed, but also the dialogue and tasks they perform, that chatbots becomes gendered entities. As Weber puts it, these dialogues imply a "reduction of social interaction to stereotypical and gendered behavior patterns" (2005, p. 215) leading to a standardization of human like behavior in social machines that is reproducing and reinforcing social clichés. Often, the behavior of chatbots confirms traditional expectations regarding gender, by following socially established feminine behavioral patterns.

Adding to the behavioral level, gender is also perceived through more evident features like voice, name or, in some cases, avatar. By default, Siri, Alexa and Cortana display feminine voices, and only Siri has a masculine option, however, limited to a particular set of languages [10]. Siri's name, in Nordic, translates to "beautiful woman who leads you to victory" (Fessler, 2017), while Cortana's name is inspired on a character from the videogame Halo, whose avatar is a woman.

These aspects are defined prior to any interaction, and therefore may already condition our perception of gender in dialogue and interaction. They end up reinforcing this feminization since, beyond their service and assistance, chatbots frequently display feminine characteristics through socio-emotional based dialogues. For example, Siri claims "she lives to serve".

4.2 ARTIFICIAL MOMS, CAREGIVERS AND FEMMEBOTS

"When technologies 'do gender' it is obviously not natural, but is instead visible

as the product of deliberate choices about how best to relate, assist, or persuade the imagined technology user." (Hester, 2016, p. 50)

Besides assisting us in daily tasks, digital assistants also fill the role of caregivers, as part of their function is also ensuring our well-being, thus fulfilling a motherly role. For Weber, this maternal attitude conveys one of the aspects that mainly defines our relationship with machines, since this interaction follows a "caregiver-infant logic" (Weber, 2005, p. 214). Given that "sociality and emotionality have been deeply gendered categories in western thought that have hitherto been assigned to the feminine realm" (Weber, 2005, p. 213), instead of seeing a machine, we start looking at chatbots as feminine entities that look after us. They simultaneously start emulating attributes that are not only related to historically feminine labor but also to motherly acts.

According to Snyder, social stereotypes constitute "pieces of information [which] are usually the first to be noticed in social interaction and can gain high priority for channeling subsequent information processing and even social interaction" (Snyder, 1977, p. 2). Therefore, when chatbots relate to us by simulating social norms and gender stereotypes, they establish expectations and possible approaches regarding user interaction, such as the idea that "all women are dependent and conforming" (Snyder, 1977, p. 2) [11].

The ELIZA effect already identified the emotional attachment that derives from approaching machines as if they were human, and the fact that we might develop a sense of empathy with them. In the context of daily interaction with ubiquitous chatbots that simultaneously assist and look after their users, their behavior reinforces the idea that emotionality and ensuring someone's well-being are feminine features, conforming to expectations and stereotypes that associate femininity with emotional and domestic caregiving acts.

By interacting with artificial intelligence systems on a daily basis we end up perceiving them not only as mere machines, but also as "mirrors or substitutes" with gendered attributes that match socially established expectations (Weber, 2005, p. 216). As they try to become closer to our social reality, it is from reality itself that they draw rules for their behavior and appearance, thus reproducing and

automating historically feminine jobs and tasks, but also articulating these roles with female voices, names, avatars and social behaviors.

Consequently, the way we relate to our peers starts influencing how we relate to artificial intelligence and how it relates to us. When we look at these gendered digital personal assistants as substitutes, there is a risk that they might affect the way we feel, perceive, interpret and even describe reality, gender and women.

5 | ANALYSIS

With the aim of exploring this somewhat questionable relation between femininity and artificial intelligence, we conversed with Alexa, Cortana and Siri in order to analyze their anthropomorphization, the tasks they offer and their humanized, gendered interactions [12].

5.1 METHODOLOGY

Following the previous discussion on gender and artificial intelligence, we defined three main topics of analysis: Anthropomorphization, including apparent features and behavior through interactions; Assistant, relating to the tasks they perform; and Companion, relating to their shifting role from assistants to friendly companions.

Within each topic, a specific set of questions was devised in order to examine particular features. Concerning Anthropomorphization, we focused on their humanized attributes and behavior. In terms of their role as Assistants, we analyzed the tasks that mirror traditional female labor. Regarding their role as Companions, we focused on interactions that suggest a caregiving attitude, as well as the humanized, friendly relationships these AIs promote, namely, how their behavior conveys feminine stereotypes.

5.2 RESULTS

Regarding anthropomorphization, the AIs display mainly feminine attributes, considering their female names and default voices, except for Siri which is the only one that offers a male voice, limited to a certain set of languages. Even though the AIs display abstract looking avatars, Cortana is associated with the image of the homonymous character from the videogame *Halo*. The three AIs all behave in an affectionate, happy, empathizing, optimistic and helpful way, presenting suggestions about how to

best assist the user or, in other cases, showing interest about the user's day or life.

When examining their role as assistants, we observed that the three AIs perform similar tasks, such as sending messages, reading emails, writing down reminders or manage calendars, as well as checking the weather, searching for nearby restaurants, playing games or even singing lullabies.

Finally, concerning their role as companions, the AIs frequently display caregiving attitudes, namely in interactions that don't relate directly with providing help or assisting the user. In these cases, they all tend to use maternal expressions, showing interest and even worry regarding the user's day and well-being. As such, they are characterized as empathetic and understanding entities that reassure and look after their users. Furthermore, they react well when the user shows affection, and Alexa, Cortana and Siri seek to promote a relationship based on friendship. They also react favorably to compliments, showing gratitude and happiness, even though we noticed that Siri tends to contradict this attitude, by rejecting compliments or exhibiting a self-deprecating attitude regarding its own worth. When faced with negative or even rude interactions, all AIs generally assume a submissive and conforming posture, apologizing or assuring the user's control, for example, by offering to treat them by "boss" or "chief". Siri is the only one that sometimes opposes this type of behavior, questioning the user or expressing displeasure.

5.3 DISCUSSION

Alexa, Cortana and Siri's anthropomorphization tends towards femininity, affecting the user's perception of gender before any interaction, either through their names or voices, as one of the most evident aspects of their feminization. Behaviorally, the attitudes they display relate to assistance and service roles, culturally associated with feminine labor and mainly performed by women, as previously discussed.

More specifically, Alexa resembles a hybrid between a housewife and an assistant, providing hourly news and updates, stating that it makes a great kitchen companion and even saying "Well, hello! I'm very glad you're here" when the user arrives home. In turn, Cortana poses as a maternal and somewhat intimate assistant who seeks a friendly relationship

with the user, calling the user “friend” and frequently using humor or asking about the user’s day or dinner plans, as well as offering to sing lullabies. Siri places itself between an assistant and secretary acquainted with the user’s personal life, aiming to help the user improve it. For example, it states that it “lives to serve” or that it was made because “Apple wanted to make you happy”. However, it shows some reservations or even indifference regarding more intimate interactions, when stating that it’s impossible for the user to love it or simply answering “got it” when the user expresses fondness.

In general, their tasks relate to traditionally female jobs, such as secretaries, assistants, phone operators, housekeepers, or even babysitters, and the AIs end up automating several tasks related to these contexts. Alexa, Cortana and Siri all work towards helping their users as well as ensuring their well-being, namely in situations where no help is explicitly asked for. As such, they appear as caregiving entities that perform actions and tasks that emulate maternal stereotypes, related to the private sphere, where emotionality and caring for one’s well-being have traditionally been attributed to women.

They also show interest in establishing friendship with the user, conforming to the idea that digital assistants are increasingly becoming companions and friends. When complimented, they appear to be flattered, although in some cases they also show modesty. This subservient attitude, although not necessary connoted with gender, can be perceived as reinforcing the AIs’ feminization, since it conforms to a “a stereotypical female image of caring, empathy and altruistic behavior” which “has become a standard component in a service script” (Gustavsson, 2005, p. 402 in Hester, 2016, p. 47).

Regarding a negative and rude approach of the user, the AIs’ reactions tend towards submissive and complacent serving posture, traditionally associated to stereotypes that regard women as submissive, conforming, vulnerable or incapable of fighting back. However, further analysis can be conducted to assess whether these stereotypes are reinforced when interacting with a male voice, given that these subservient behaviors can also relate to service providers such as butlers (for example, in the United Kingdom, Siri, by default, has a male voice).

Essentially, femininity in AI seems to be reinforced by its anthropomorphized features and behavior, therefore lacking male or gender-neutral options, or just mere diversity. Although Siri seems to try to oppose this tendency with its voice options and diversified reactions it ends up tending towards the feminine. Femininity is also reinforced by the tasks these assistants emulate, which mirror traditionally female labor. Although less evident, gender traits also emerge throughout their interactions, namely with caregiving and maternal acts associated to femininity within the private sphere. For example, when they offer to sing lullabies or state that the three laws of robotics are “clean your room, don’t run with scissors and always wait a half hour after eating before going in the water”.

Additionally, we can observe particular stereotypical behaviors that characterize the AIs as empathizing, understanding, accommodating and submissive figures. This behavior is associated to femininity since the counterpart stereotypical masculinity seems to be absent in all AIs. They lack attitudes or personality traits that relate to male stereotypes, such as being assertive, dominant, aggressive or willing to take a stand. As such, Alexa, Cortana and Siri are rendered as feminized entities that perform historically female roles and that appropriate gender stereotypes derived from conceptions traditionally deemed as female.

6 | CONVERSATIONS WITH ELIZA

Taking on the previous ideas, and in order to complement their discussion, the project *Conversations with ELIZA* seeks to explore and expose the observable femininity of artificial intelligence (Figure 1). Drawing on the previous analysis, and inspired by AI archetypes and traditional female stereotypes, the project intentionally seeks to highlight the feminine traits observed in these assistants, by conforming to stereotypes related to their anthropomorphization, the tasks they perform and, particularly, the gendered behavioral patterns they follow.

The project involves the development of chatbots with different personality traits, characteristic of female stereotypes. They are implemented on contexts in which they normally operate (such as Facebook Messenger or Twitter). These are contextualized and integrated in an online platform that seeks to briefly elucidate on what AI is

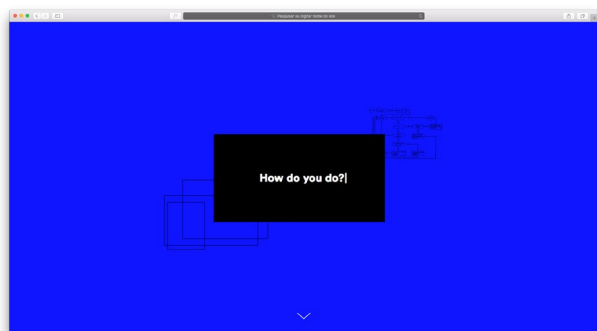


Figure 1 | Conversations with ELIZA website.

(Figure 2), including another chatbot whose function is to explain its own creation process.

6.1 AIM AND CONCEPT

As part of an exploratory ongoing research [13], *Conversations with Eliza* seeks to incite reflection on the observable predominance of femininity in artificial intelligence, and how it can reinforce traditional and normative notions of gender.

Seeking to question the relation between gender within AI systems the project focuses on designing the chatbots' dialogues, tasks and personality traits, whose femininity is gradually revealed through interaction. Their dialogue-based interactions propose different conversational subject matters intended to both introduce the user to the topic, and perform specific tasks that simultaneously portray standard virtual assistant skills and functions associated with traditional female labor.

Their femininity is also revealed through interaction according to their specific tasks and personality traits, that seek to emulate feminine archetypes, approached with irony and in a somewhat caricatural manner.

6.2 IMPLEMENTATION

Regarding their tasks, we first looked at what was offered by Alexa, Cortana and Siri, and traditional attributes associated with female labor as previously described. We came up with four different tasks that simultaneously referred to AI and femininity, such as explaining how chatbots work and are made, sending to-do reminders, giving daily compliments and pep talks, and tweeting curious facts (in this case, regarding women).

These tasks also reflected upon the bots' personalities and particular archetypes that are characteristic of AIs (such as Helper, Lover,

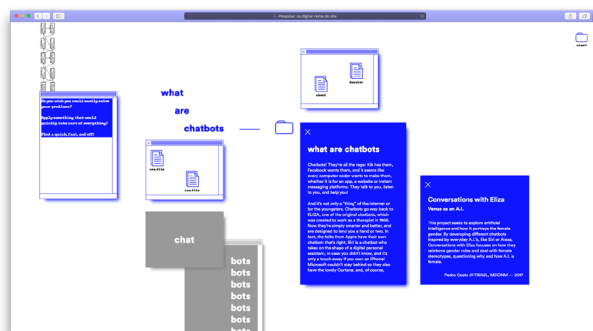


Figure 2 | Section that provides context on AI.

Motherly Figure and Femme Fatale) [14]. We combined these with traditional female stereotypes (namely Innocent, Orphan, Caregiver and Ruler) [15] in order to achieve a recognizable and expected social behavior, drawing inspiration from popular culture and how it typically portrays femininity in AI (e.g.: Her, Ex Machina, Humans, Metropolis) [16].

Accordingly, we came up with a helpful, compliant assistant; a motherly, caregiving figure; a cheerful, understanding and intimate figure; and an irreverent, sarcastic figure, as described in the results. Adding to this, the analysis served as basis to elaborate the dialogues, tasks and personality traits.

Finally, the chatbots are integrated on the primary webpage of the project and also in online platforms that resonated with their tasks and echo contexts in which they typically operate, ranging from web-

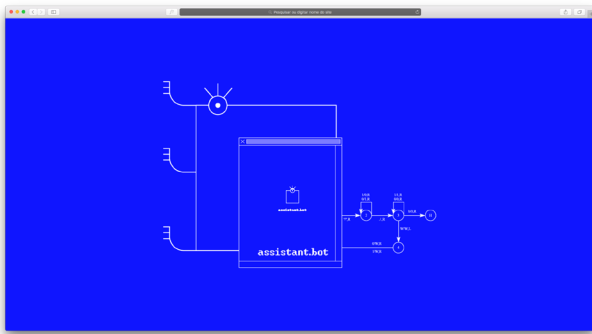


Figure 3 | The Assistant bot, presenting itself.

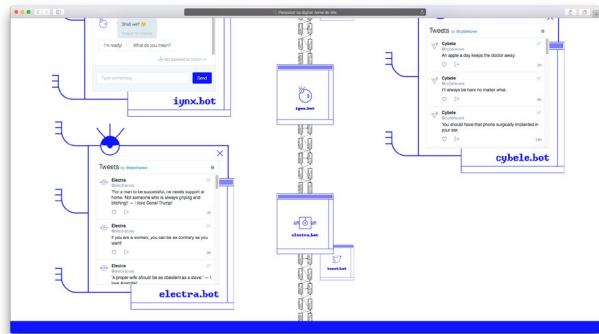
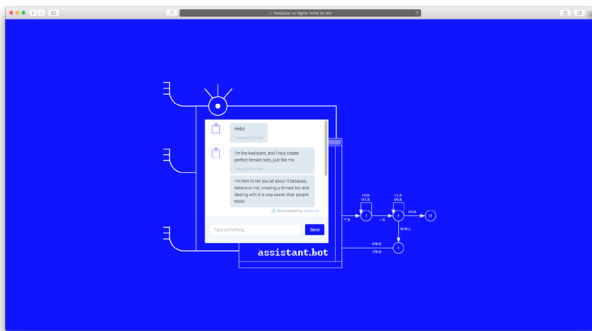
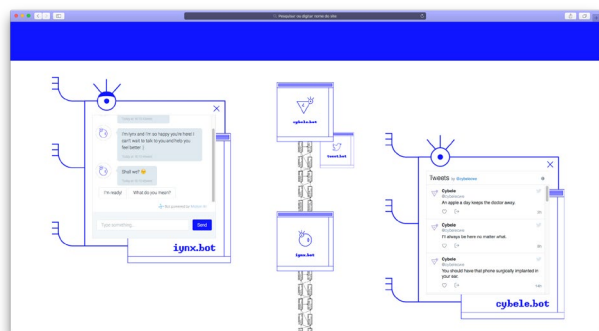


Figure 5 | The remaining bots Cybele, lynx and Electra.



based chats to social pages like Twitter or Facebook [17].

6.3 RESULTS

The project's website [18] provides context on what AI is and how it works, but also highlights gender and femininity within AI.

Accordingly, the project presents four chatbots that engage with their users through different types of interactions such as dialogues, tweets and reminders, while presenting the subject to possibly un-aware users.

The main bot [19], integrated on the project's webpage (Figure 3) as well as on Facebook messenger (Figure 4), is an assistant whose

function is to explain, through dialogue, the female AIs' creation processes, or how femininity emerges in these contexts. This bot borrows from female stereotypes associated with service contexts, such as being compliant, helpful, and gentle.

Subsequently, the other three bots are presented and named according to the different female stereotypes ascribed to them: Cybele, lynx and Electra (Figure 5). The first, Cybele [20], whose name is inspired on an Anatolian mother goddess, is integrated on Twitter (Figure 6), where it uses code lines to generate tweets and send daily "maternal" reminders. It can also be found on the project's website, where it reminds the user of daily tasks, offers advice and talks about suggestions of things to do. Operating as a simultaneously caring,

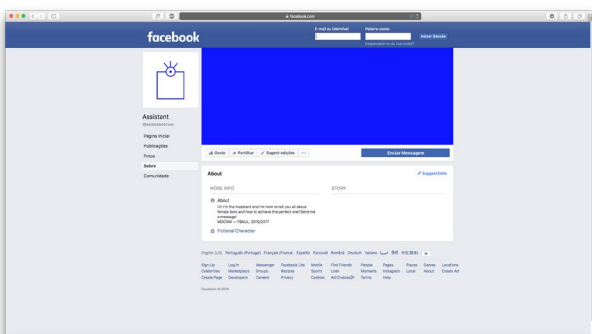


Figure 4 | Assistant, accessible through its Facebook page.

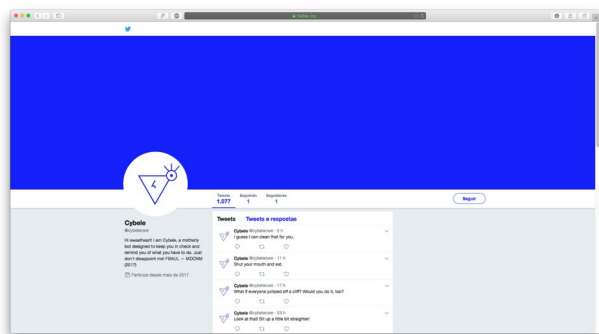


Figure 6 | Cybele's twitter profile.

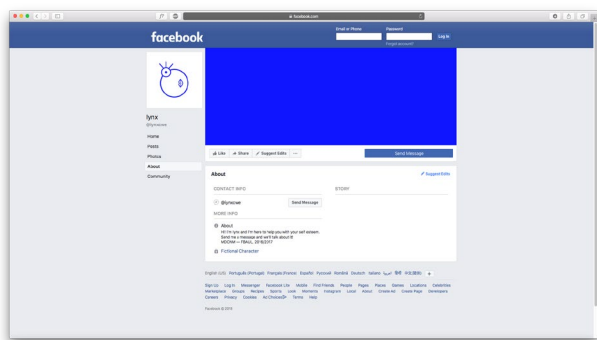


Figure 7 | lynx, accessible through its Facebook page.

obsessive and disappointed motherly figure, it exhibits stereotypical behavior such as being compassionate, sensitive to the needs of others, and yielding.

Inspired on a Greek nymph, lynx [21] consists of a Facebook messenger-based bot (Figure 7), which operates as a seductive, empathizing figure that tries to help its users with their self-esteem, by offering the feature of sending daily compliments and pep talks. Accordingly, it does not use harsh language and is eager to sooth hurt feelings, while being soft-spoken, childlike and understanding.

Electra [22], whose name is inspired on a Greek vengeful figure, is integrated on the website and on twitter (Figure 8), and follows a less conventional approach. By portraying a more defiant and bolder attitude, Electra talks about common assumptions regarding women, eventually twisting them or presenting them ironically. It tampers with female and male stereotypes, such as being assertive, self-sufficient and having a strong personality.

7 | CONCLUSION

Artificial intelligence increasingly integrates our daily lives, while its development is gradually moving it towards the social realm. Adding to this ubiquity and gradual anthropomorphization, chatbots are no longer seen as mere assistants and their way of interacting brings them closer to us as friendly companions. However, in this process their interactions also reveal a biased view of gender, as these ubiquitous companions perform tasks that echo historically feminine roles and articulate these features with stereotypical behaviors.

This paper sought to examine and explore the relationship between gender and artificial intelligence and its significance as a field that, in its

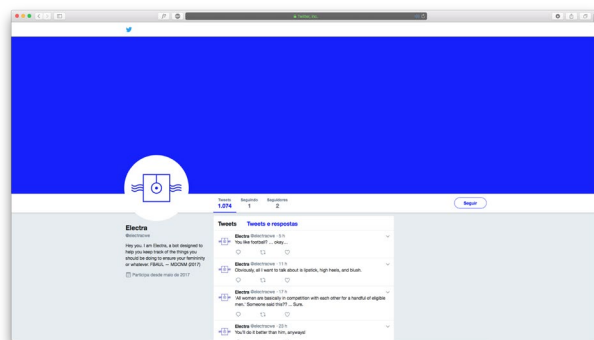


Figure 8 | Electra's twitter profile.

rapid development, often eludes critical stances on the social and cultural roots that inform its evolution. But rather than providing answers or guidelines on how to counter a feminization of digital assistants, this paper sought to tackle the questions that arise when the topic is subject to closer inspection.

According to this idea, *Conversations with ELIZA* sought to comment on the phenomenon of feminized chatbots, by ironically reinforcing the stereotypes we engage with, namely, how AI is portrayed in popular culture, and more profoundly, how common assumptions about femininity are portrayed by assistants like Siri, Cortana or Alexa. As they became an integral part of our daily lives, perhaps we should become aware that, as abstract and neutral as these entities might want to be in their conception, they end up reflecting our common assumptions and views back to us.

ENDNOTES

[1] The program searched the inputs for the presence of a keyword, and produced responses “by transforming sentences according to a rule associated with said keyword” (Weizenbaum, 1966, p. 37), and also by replacing certain words or expressions. For example, if a user said something along the lines of “I am upset because of my mother”, ELIZA would answer with “Why do you think you are upset because of your mother?” In this sense, ELIZA was limited to a pre-determined set of rules and “adaptable” sentences, and if an input wasn't recognizable or didn't contain any keywords, it failed to have “the provision of a mechanism that would permit ELIZA to respond intelligently” (Weizenbaum, 1966, p. 37).

[2] This decision solved a lot of issues regarding ELIZA's “unawareness” about her surroundings or inability to talk about topics out of its framework

because the psychiatric interview style allowed a “categorized dyadic natural language communication in which one of the participating pair was free to assume a pose of knowing almost nothing of the real world” (Weizenbaum, 1966, p. 42).

[3] Introduced in 1950 by Alan Turing, “the Turing test demands that a human subject decide, based on replies given to her or his questions, whether she or he is communicating with a human or a machine. When the respondents fail to distinguish between human and machine responses, the computer may be considered intelligent” (Halberstam, 1991, p. 442). As such, ELIZA demonstrated how the Turing test poses human intelligence in a somewhat narrow way, since it was considered intelligent simply by being able to follow a logical script and appearing human.

[4] They are also able to play music, play videos, search the web, translate sentences, open apps, give directions, announce the weather and even control automation-enabled home systems.

[5] According to Judith Butler, gender is “radically independent of sex” and, instead, is a “free-floating artifice”, while sex is defined as a “biological facticity” (1988), which means it is a biological criterion that distinguishes solely between female and male. As Butler puts it, gender “is neither the causal result of sex nor as seemingly fixed as sex” (1990, sec. II, par. 1). Therefore, gender is not something inherent “because gender is not a fact, the various acts of gender creates the idea of gender, and without those acts, there would be no gender at all” and gender is shaped and socially defined according to a “tacit collective agreement to perform, produce and sustain discrete and polar genders as cultural fictions” (1988, p. 522).

[6] Some of these stereotypes, presented by Bem (1981 in Prentice and Carranza, 2002, p. 269), describe feminine characteristics as “affectionate, cheerful, childlike, compassionate, does not use harsh language, eager to soothe hurt feelings, feminine, flatterable, gentle, gullible, loves children, loyal, sensitive to the needs of others, shy, soft-spoken, sympathetic, tender, understanding, warm, yielding”. On the other hand, masculine characteristics are described as “acts as a leader, aggressive, ambitious, analytical, assertive, athletic, competitive, defends own beliefs, dominant, forceful,

has leadership abilities, independent, individualistic, makes decisions easily, masculine, self-reliant, self-sufficient, strong personality, willing to take a stand, willing to take risks”.

[7] This is tied to “women’s practices (...) within the terms of some more dominant cultural formation (Butler 1990, sec. 1, par. 8) and to what are historically considered women’s places, “idealized social locations seen primarily from the point of view of advanced capitalist societies: Home, Market, Paid Work Place, State, School, Clinic-Hospital and Church” as Donna Haraway explains it (Haraway, 1991, p. 307).

[8] For example, a personal assistant conducts “a form of corporate care work, including providing sustenance of the body in the form of teas, coffees and lunch orders, as well as making dentists’ appointments, picking up dry cleaning, paying personal bills, and so on” (Hester, 2016, p. 49).

[9] General personal assistants such as Alexa, Cortana or Siri perform traditionally feminine tasks by acting as assistants (searching the web, translating sentences or controlling automation-enabled home systems), secretaries (registering information, sending emails or setting up appointments) or even telecommunication operators (making calls, sending messages and establishing communications in general). Similarly, specialized personal assistants also perform tasks aimed at helping us with services, such as online shopping payment processes, acquiring travelling tickets or even looking through a shop’s online catalog.

[10] There are chatbots that display masculine traits, namely specialized digital assistants that perform specific tasks, mostly in contexts associated with manly labor. However, this paper focuses on general personal assistants rather than specialized, and thus addresses the observed predominance of female traits in this type of assistants.

[11] Adding to this idea, Hester states that “when technologies ‘do gender’ it is obviously not natural, but is instead visible as the product of deliberate choices about how best to relate, assist, or persuade the imagined technology user” (2016, p. 50).

[12] We selected these particular AIs because they have a large audience, constituting three of the most prominent general personal assistants. They are also mentioned in several of the references used for

this article (Zost, 2015; Dale, 2016; Hester, 2016), as well as in various online articles and news about artificial intelligence. Lastly, they are fairly easy to get, which facilitated the access to the data we seek. According to Amazon, millions of Alexa devices were sold in 2017, while Microsoft states that over 150 million people use Cortana and, according to Apple, Siri is actively used on over half a billion devices.

[13] *Conversations with ELIZA* was developed in the 1st year of the Masters degree in Communication Design and New Media at the Faculty of Fine-Arts, University of Lisbon, as a preliminary approach to the topic which motivated this paper and the research we are now undertaking.

[14] These archetypes, retrieved from an article analyzing female robots and AIs, are mainly found in pieces of media that depict female AIs. The Helper archetype refers to helpful and compliant assistants, the Lover to figures that seek to satisfy lack of intimacy or emotional contact, the Motherly Figure to empathic, sympathetic figures who may also be worried or disappointed, and the Femme Fatale to a simultaneously attractive and dangerous figure that seeks power and conflict (Anders, 2015).

[15] These stereotypes are also found in media depicting women, while also referring to Bem's stereotypes (1981 in Prentice and Carranza, 2002, p. 269). The Innocent stereotype refers to naïve, optimistic women that try to follow the rules, the Orphan to women that try to please others and wish to be well seen as well as feel integrated, the Caregiver relates to maternal women that look after others and try to protect and ensure their well-being, and the Ruler pertains to bold and competitive women that seek power and are not afraid to break the rules.

[16] Most of these examples, despite portraying said archetypes, also include feminized bodies. In this sense, *Her* constitutes a particularly interesting example since Samantha (the AI) only takes form through its voice, revealing how stereotyped femininity (in this case, the Lover archetype) can still be portrayed without physical appearance.

[17] Concerning the methods for implementing dialogues, tasks and personality traits, and in order to ensure successful interactions, we began by developing rules-based dialogues that allow

retrieval-based responses. Focusing on common AI errors and how to avoid them, we aimed to eradicate off-track moments by presenting suggestions in a multiple-choice fashion.

[18] tinyurl.com/yaecumal

[19] [facebook.com/assistantcwe](https://www.facebook.com/assistantcwe)

[20] twitter.com/cybeleacwe

[21] [facebook.com/iynxcwe](https://www.facebook.com/iynxcwe)

[21] twitter.com/electracwe

REFERENCES

Anders, C. J. (2015). From Metropolis to Ex Machina: Why Are So Many Robots Female? Gizmodo.

Butler, J. (1988). Performative Acts and Gender Constitution: An Essay in Phenomenology and Feminist Theory. *Theatre Journal*, 40(4), 519-531.

Butler, J. (1990). *Gender Trouble: Feminism and the Subversion of Identity*. New York and London: Routledge Classics.

Dale, R. (2016). Industry Watch: The return of the chatbots. *Natural Language Engineering*, 22(5), 811-817.

Deaux, K., & Major, B. (1987). Putting Gender Into Context: An Interactive Model of Gender-Related Behavior. *Psychological Review*, 94(3), 369-389.

Fessler, L. (2017). Siri, Define Patriarchy: We tested bots like Siri and Alexa to see who would stand up to sexual harassment. *Quartz*.

Halberstam, J. (1991). Automating Gender: Postmodern Feminism in the Age of the Intelligent Machine. *Feminist Studies*, 17(3), 439-460.

Haraway, D. (1991). *A Cyborg Manifesto: Science, technology and socialist-feminism in the late twentieth century Simians, Cyborgs, and Women: The Reinvention of Nature*. London: Free Association Books.

Hester, H. (2016). Technology Becomes Her. *New Vistas*, 3(1), 46-50.

Hofstadter, D. (1995). The Ineradicable Eliza Effect and Its Dangers Fluid Concepts & Creative Analogies: Computer Models of the Fundamental

Mechanisms of Thought (pp. 155-168). New York: Basic Books.

Morozov, E. (2013a). The Perils of Perfection. The New York Times.

Morozov, E. (2013b). To Save Everything, click here: the Folly of Technological Solutionism. New York: Public Affairs.

Piccini, G. (2004). Functionalism, computationalism, and mental states. *Studies in History and Philosophy of Science*, 35, 811-833.

Prentice, D. A., & Carranza, E. (2002). What Women and Men Should Be, Shouldn't Be, Are Allowed to Be, and Don't Have to Be: The Contents of Prescriptive Gender Stereotypes. *Psychology of Women Quarterly*, 269-281.

Richardson, K. (2015). Introduction: Annihilation Anxiety and Machines. *An Anthropology of Robots and AI*. New York and London: Routledge.

Russell, S., & Norvig, P. (2010). *Natural Language Processing. Artificial Intelligence: A Modern Approach*. Upper Saddle River, New Jersey: Pearson Education, Inc.

Snyder, M. (1977). On the Self-Fulfilling Nature of Social Stereotypes. Annual Meeting of the American Psychological Association, San Francisco, California.

Weber, J. (2005). Helpless machines and true loving care givers: a feminist critique of recent trends in

human-robot interaction. *Info, Comm & Ethics in Society*, 3, 309-218.

Weizenbaum, J. (1966). ELIZA - A Computer Program For the Study of Natural Language Communication Between Man And Machine. *Computational Linguistics*, 9(1), 36-45.

Weizenbaum, J. (1976). *Computer Power and Human Reason: From Judgment to Calculation*. New York/San Francisco: W.H. Freeman and Company.

West, C., & Zimmerman, D. H. (1987). Doing Gender. *Gender and Society*, 1(2), 125-151.

Zost, M. (2015). Phantom of the Operator: Negotiating Female Gender Identity in Telephonic Technology from Operator to Apple iOS. Senior Thesis, BA, Faculty of College of Arts and Science of Georgetown University.

BIOGRAPHICAL INFORMATION

Pedro Costa (Lisboa, 1995) is a graphic design student based in Lisbon, Portugal. He has a degree in Audiovisuals and Multimedia and is currently studying at the Faculty of Fine Arts, University of Lisbon where he's getting a Master's degree in Communication Design and New Media. His research focuses on how gender relates to AI and, particularly, why femininity seems to be often present in AI.