UNIVERSIDADE DE LISBOA
FACULDADE DE BELAS-ARTES

“Some Things You Can Ask Me”:

On Gender and Artificial Intelligence

Pedro Carvalho Ferreira da Costa

Orientadores: Profª. Doutora Luísa Maria Lopes Ribas
Prof. Doutor José Miguel Santos Araújo Carvalhais Fonseca
Prof. Doutor Mario Verdicchio

Tese especialmente elaborada para a obtenção do grau de Doutor em Belas-Artes,
na especialidade de Design de Comunicação

2024

UNIVERSIDADE DE LISBOA
FACULDADE DE BELAS-ARTES

"Some Things You Can Ask Me":
On Gender and Artificial Intelligence

Pedro Carvalho Ferreira da Costa

Orientadores: Profª. Doutora Luísa Maria Lopes Ribas
Prof. Doutor José Miguel Santos Araújo Carvalhais Fonseca
Prof. Doutor Mario Verdicchio

Tese especialmente elaborada para a obtenção do grau de Doutor em Belas-Artes,
na especialidade de Design de Comunicação

Júri:

Presidente: Doutora Sofia Leal Rodrigues, Professora Auxiliar e vogal do Conselho Científico da Faculdade de Belas-Artes da Universidade de Lisboa, Presidente do júri por nomeação do Presidente do Conselho Científico desta Faculdade, Prof. Doutor João Carlos de Castro Silva, nos termos do n.° 1.1. do Despacho n.° 10432/2023, publicado em Diário da República, 2a série, n.° 197, de 11 de outubro;

Vogais:

- Doutora Ana Araújo Barros Viseu, Professora Auxiliar da Faculdade de Ciência Sociais e Humanas da Universidade Nova de Lisboa [1a arguente];
- Doutora Ana Sofia Lopes da Ponte, Professora Auxiliar do IADE – Faculdade de Design, Tecnologia e Comunicação da Universidade Europeia [2a arguente];
- Doutora Suzana Isabel Malveiro Parreira, Professora Auxiliar da Faculdade de Belas- Artes da Universidade de Lisboa;
- Doutora Luísa Maria Lopes Ribas, Professora Auxiliar da Faculdade de Belas-Artes da Universidade de Lisboa [orientadora].

Trabalho desenvolvido com apoio financeiro da Fundação para a Ciência e Tecnologia.

2024

DECLARAÇÃO DE AUTORIA

Eu, Pedro Carvalho Ferreira da Costa, declaro que a tese de doutoramento intitulada “"Some Things You Can Ask Me": On Gender and Artificial Intelligence”, é o resultado da minha investigação pessoal e independente. O conteúdo é original e todas as fontes consultadas estão devidamente mencionadas na bibliografia ou outras listagens de fontes documentais, tal como todas as citações diretas ou indiretas têm devida indicação ao longo do trabalho segundo as normas académicas.

Lisboa, 13 de Março de 2024

ABSTRACT

This study addresses the relationship between gender and artificial intelligence. It focuses on a tendency towards femininity in current digital assistants, speculating on the possibilities of queering these entities and reframing them outside of a cis-normative binary spectrum. It confronts sociocultural understandings of gender with the evolution of AI, highlighting how these assistants tend to emulate stereotypical feminine attributes, tasks and behaviors.

The study discusses the implications of gender attribution in AI, addressing concerns about privacy, bias and the reinforcement of social stereotypes. Diversification and customization of these entities emerge as possible strategies to counter this phenomenon. However, a lack of consensus prevails, as current solutions tend to reinforce traditional conceptions of gender. Based on this framework, an analysis of Alexa, Cortana, Google Assistant and Siri examines how their stance towards gender evolved through their anthropomorphized attributes, the tasks they perform and their behavioral traits. It also investigates the features and functionalities that are prioritized in their trends of development, extending this analysis to ChatGPT. Overall, challenges in addressing gender bias and stereotypes in current digital assistants remain. Complementing this discussion, the project *Cistem Not Found* is developed as an exploration of gender, otherness and queerness in AI. Across four audiovisual pieces, a disruptive chatbot provides critical commentary on current trends of AI and gender attribution in digital assistants. The bot's characterization, monologue and personality traits explore possibilities of deanthropomorphizing these entities, questioning current gender-neutral approaches and evoking trans, nonbinary and posthuman identities to challenge cis-normative and binary conceptions of gender.

This study seeks to raise awareness on the implications of assigning gender to humanized digital assistants. Through a queer, intersectional and speculative approach, it aims to provide a more nuanced understanding of the intersections between gender and AI, pushing traditional boundaries to reimagine gender possibilities beyond a bioessentialist, binary and cisgender frame.

Keywords: artificial intelligence, digital assistants, gender, intersectionality, queerness.

RESUMO ALARGADO

A inteligência artificial constitui atualmente uma componente integrante do quotidiano, nomeadamente através de assistentes digitais cuja ubiquidade tende a passar despercebida. Estes sistemas estão a assumir uma postura cada vez mais amigável e preocupada com o nosso bem-estar, afastando-se de uma perceção meramente robótica e de assistência. No entanto, ao atribuir traços humanos a estas entidades, observa-se a presença de características que tendem a reforçar conceções e estereótipos binários e cis-normativos de género.

Propósito e objetivos

Este estudo pretende compreender, questionar e explorar a relação entre género e inteligência artificial. Foca-se na atual tendência de assistentes digitais serem predominantemente femininos, especulando sobre as possibilidades de *queerizar* estas entidades e reenquadrá-las fora de um espectro binário cis-normativo.

Começa por proporcionar uma compreensão abrangente dos fundamentos socioculturais do género e da evolução da inteligência artificial. Foca-se em expressões de género trans e dissidentes fora de um quadro cis-binário, procurando expor a forma como conceções tradicionais de género perpetuam estereótipos e refletem estruturas laborais hierarquizadas. Olhando para a integração da IA no quotidiano, pretende-se mapear a sua evolução, nomeadamente sob a forma de assistentes digitais. Com base nestas observações, procura-se compreender como e porque é que estes sistemas emulam atributos, papéis e estereótipos de género. Complementando este debate, o estudo foca-se em discussões sobre este fenómeno, no âmbito académico e dos media, destacando as principais questões, preocupações e sugestões em torno da atribuição de género na IA. Aborda-se também a cultura pop para revelar a forma como representações ficcionais desta tecnologia refletem e reforçam conceções tradicionais de género.

Partindo deste enquadramento, procede-se a uma análise de assistentes digitais para verificar a presença do género, focando a sua apresentação, interação e desenvolvimento. Começa por examinar as características antropomorfizadas e traços comportamentais destas entidades para averiguar a forma que o género assume através de aspetos mais aparentes, mas também através das tarefas que desempenham e das suas interações socioemocionais.

Esta análise tem como objetivo demonstrar o modo como o design e as funções dos assistentes digitais emulam atributos de género, averiguando-se também se as suas interações e comportamento confirmam ou contrariam estereótipos tradicionalmente femininos. De seguida, examinam-se as tendências do seu desenvolvimento, focando as características e funcionalidades que estão a ser priorizadas. Esta abordagem pretende aferir a forma como debates e discussões em torno destas tecnologias influenciam a sua conceção, evidenciando também as ideias e assunções subjacentes ao seu desenvolvimento.

Por fim, o projeto *Cistem Not Found* reflete a maneira como o nosso entendimento e compreensão destes tópicos evoluiu ao longo da componente teórica e analítica. Procura consciencializar sobre as implicações de atribuir género a assistentes digitais, promovendo debate e reflexão sobre as intersecções entre tecnologias de IA e conceções binárias e cis-normativas de género. O projeto envolve o desenvolvimento de um *bot* provocador e irreverente, propondo um comentário crítico sobre tendências da inteligência artificial e respetiva atribuição de género. Exploram-se alternativas de voz para subverter estereótipos tradicionais e contemplar formas de contrariar este fenómeno. Através da caraterização, monólogo e traços de personalidade do *bot*, procura-se especular sobre as possibilidades de desantropomorfizar estas entidades, refletir sobre as suas abordagens supostamente neutras em termos de género e explorar traços que evocam identidades trans, não binárias e pós-humanas para contestar conceções normativas e convencionais de género.

Estrutura e metodologia

O estudo divide-se em três componentes: uma componente teórica, que orienta uma componente de análise, ambas complementando e informando o desenvolvimento de uma componente projetual.

A abordagem teórica confronta entendimentos socioculturais de género com a evolução da inteligência artificial e respetiva integração no quotidiano. Recorre a estudos de género, perspetivas *queer* e abordagens interseccionais para expor convenções binárias de género e discutir identidades trans, não binárias e dissidentes enquanto possibilidades alternativas de género. Em segundo lugar, analisa estudos no contexto da computação, IA e estudos dos media para mapear a evolução e a integração desta tecnologia na vida quotidiana sob a forma de assistentes digitais, revelando igualmente a forma como os seus algoritmos são potencialmente tendenciosos. Através da discussão destes conceitos,

procura-se salientar que noções de género imperam nestes sistemas, nomeadamente nos seus atributos, tarefas e comportamentos. O estudo foca então em discussões do fenómeno no contexto académico e dos media para debater as principais questões, preocupações e sugestões que emergem em torno destas tecnologias e da atribuição de género a assistentes digitais. Por fim, discute a forma como representações ficcionais da IA no contexto da literatura, cinema, televisão ou videojogos tendem a espelhar e perpetuar conceções tradicionais de masculinidade e feminilidade.

A componente analítica revisita e atualiza uma análise prévia de Alexa, Cortana, Google Assistant e Siri para averiguar a presença do género na caraterização, interação e tendências de desenvolvimento destas entidades. Foca-se na sua antropomorfização, incluindo características como nomes, vozes e avatares, e no tipo de tarefas que desempenham, em particular as que evocam profissões historicamente femininas, bem como nas interações que sugerem estereótipos e comportamentos tradicionalmente femininos. Em segundo lugar, o estudo aborda as funcionalidades e características que estão a ser priorizadas no desenvolvimento destas entidades, conforme promovido pelas suas plataformas e comunicados oficiais. Por fim, a análise compara os resultados obtidos em 2019, 2021 e 2023, discutindo a evolução da presença de atributos de género. Adicionalmente, examina-se ChatGPT relativamente à sua antropomorfização e às respostas a interações socioemocionais. Embora ChatGPT se diferencie dos restantes assistentes digitais por operar principalmente como um modelo de processamento de linguagem natural, considera-se relevante examinar a forma como lida com noções de antropomorfização, humanização e laços emocionais.

A componente projetual envolve o desenvolvimento de um *chatbot* disruptivo e mutante (https://genderai.pt). Ao longo de quatro peças, o projeto foca-se na conceção da voz, traços de personalidade e monólogo do *bot*, distorcendo ironicamente arquétipos masculinos e femininos, assim como estereótipos relacionados com assistência e prestação de serviços. O *bot* explora e especula sobre possibilidades de abordar a presença do género na IA, enfatizando os seus atributos desantropomorfizados, explorando abordagens neutras e ambíguas ao género e evocando identidades trans, não binárias ou pós-humanas. A caracterização intencionalmente retro e disruptiva do *bot* opõe-se às tendências de antropomorfizar os assistentes digitais e torná-los em companheiros amigáveis, proporcionando um comentário crítico sobre as implicações de feminizar estas entidades e sobre a integração desta tecnologia no quotidiano. Através da sua abordagem crítica e especulativa, *Cistem Not Found* procura suscitar reflexão sobre as questões que surgem com o desenvolvi-

mento da inteligência artificial e a forma como os assistentes digitais refletem e reforçam noções tradicionais e normativas de género.

Contributos

O rápido desenvolvimento da inteligência artificial tende a eludir posições críticas sobre as raízes sociais e culturais que informam a sua evolução, nomeadamente contornando debates que ainda estão a redefinir a nossa compreensão do género. Embora entidades como a UNESCO e a Comissão Europeia comecem a debater questões sociais e culturais envolvidas na antropomorfização da IA, as suas orientações focam-se maioritariamente em evitar a discriminação e promover a inclusão, sem proporcionarem direções claras sobre as especificidades que surgem com a atribuição de género a estes sistemas. Estas discussões tendem também a priorizar uma perspetiva implicitamente cis-binária, negligenciando entendimentos de género fora deste quadro e ignorando identidades trans, não-binárias e dissidentes.

Esta investigação considera relevante aprofundar as implicações da utilização generalizada de assistentes digitais que emulam atributos de género, através de um enquadramento interdisciplinar que propõe uma abordagem inquisitiva, especulativa e exploratória. Mais do que dar respostas ou orientações definitivas sobre como solucionar a atribuição de género na inteligência artificial, pretende-se consciencializar sobre as complexidades deste fenómeno. Assumindo uma postura crítica, interseccional e *queer*, as componentes teórica, analítica e projetual do estudo visam proporcionar uma compreensão mais detalhada deste fenómeno, considerando as subjetividades da identidade de género nas suas intersecções sociais e culturais, juntamente com o impacto desta tecnologia na nossa vida quotidiana. Espera-se contribuir para um entendimento mais aprofundado do papel que estas tecnologias desempenham na perceção e perpetuação de convenções sociais e culturais, informando designers que contribuem para o desenvolvimento desta tecnologia, mas também os utilizadores das suas criações.

Visando uma consciencialização sobre este fenómeno, o projeto *Cistem Not Found* materializa e exacerba estas questões. Procura subverter noções e estereótipos cisgénero de feminilidade e masculinidade, questionar a suposta neutralidade de género da inteligência artificial e expor a forma como o design destas entidades tende a basear-se em práticas de exclusão que ignoram as subjetividades do género, do sexo e da identidade. Através das suas diferentes componentes e propostas de voz, *Cistem Not Found* envolve cenários

especulativos em que os limites entre humano e máquina são diluídos para contestar conceções cis-normativas e binárias de género.

Desta forma, esta investigação visa suscitar uma reflexão sobre as implicações da atribuição de género a assistentes digitais humanizados. Através de uma abordagem especulativa, *queer* e interseccional, pretende proporcionar uma compreensão das intersecções do género com a IA, desafiando acepções tradicionais para reimaginar o género fora de um entendimento bioessencialista, binário e cisgénero.

Palavras-chave: inteligência artificial, assistentes digitais, género, interseccionalidade, *queerness*.

ACKNOWLEDGEMENTS

Luísa Ribas and Miguel Carvalhais, for their invaluable and enthusiastic guidance, tireless support and constructive insights. Their advice, dedication and patience have been key in shaping the direction, depth and quality of my research and completing this work.

Mario Verdicchio for his advice and valuable support.

Corina Fernandes, for her incentive and motivation.

My parents, for their unwavering help, encouragement and confidence throughout my whole academic journey. My family and friends, for their unconditional support and affection.

André, for his joyful company, precious friendship and enthusiastic help in revising my writing. My cats, for their comforting company during late-night hours.

Teresa Ves Liberta, whose unapologetic freedom, sensitivity and poetry left us too soon.

Batata, for her boundless patience, endless wisdom and mellow tenderness.

Arca and SOPHIE, for their everlasting impact as trans women on music, culture and trans representation, which inspired the trajectory of this study's project. All the trans individuals mentioned in this research, including Paul B. Preciado, as well as those not explicitly named and who may never be recognized, but continue to challenge boundaries and carve out space in the face of a painful and unjust world that wants us invisible.

Batam palmas para as travestis que lutam para existir e, a cada dia batalhando,
conquistar o seu direito de viver, brilhar e arrasar.
Linn da Quebrada 2017, "Mulher"

TABLE OF CONTENTS

INTRODUCTION 25

Background 26

Purpose and Objectives 27

Methodology 28

Significance 30

Overview 31

1. UNVEILING GENDER:
IDENTITY, DISSIDENCE AND NORMATIVE CONSTRUCTS 35

1.1. Let's Talk About Gender, Baby 35

1.1.1. Gender Is Plastic (It's Fantastic) 35

1.1.2. Beyond Binary Binds: Gender Fluidity, Ancestry and Infinity 37

1.1.3. Trans and Nonconforming Gender Identities Across the World 38

1.1.4. Catholicism, Colonialism and Pathologization: Gender Naturalization and Imposition 41

1.1.5. Towards a Queer and Intersectional Approach to Gender 45

1.2. Gender and Sex Distinction in the Binary Frame 48

1.2.1. Sex, Sex Category and Gender 48

1.2.2. Cis-Binary Genders and the Implicit Hetero-Monogamous Frame 50

1.2.3. Binary Burdens and Social Regulations 52

1.3. "Woman" as a Concept 53

1.3.1. (You Make Me Feel Like) a Natural Woman 53

1.3.2. Category of Women: Whom, When and Where 54

1.4. Gender Roles and Social Stereotypes 56

1.4.1. Blue for Him, Pink for Her 56

1.4.2. Girls Are Pearls, Boys Are Machines 57

1.4.3. Regulatory and Punitive Conventions of Gender Behavior 59

1.5. Gender Disparity, Asymmetry and Hierarchization 62

1.5.1. Phallocentrism and Phallogocentrism 62

1.5.2. Patriarchy Over and Out 63

1.6. Women in the Integrated Circuit 66
1.6.1. Labor Structures and Historical Female Jobs 66
1.6.2. I'll Bake Your Bread and Tend Your House 67

2. THE EVOLUTION OF DIGITAL ASSISTANTS: AN EXPLORATION OF ARTIFICIAL INTELLIGENCE 71
2.1. Artificial Ancestors, Mind Automation and Imitation Games 71
2.1.1. Mythical Simulations and Humanoid Emulations 71
2.1.2. From Computationalism to Artificial Intelligence 73
2.1.3. Turing Tests, Automated Machines and Natural Language 74
2.1.4. Before Siri and Alexa, There Was ELIZA 76
2.1.5. Strong and Weak AI: From Learning Machines to Neural Networks 77
2.2. Robots, Bots and AI: From Industrial Workers to Virtual Companions 80
2.2.1. Machines as Workers, Social Robots and Companions 80
2.2.2. General Personal Assistants and Specialized Digital Assistants 82
2.2.3. From ELIZA to ChatGPT 83
2.2.4. A Multitude of Digital Assistants 86
2.2.5. Pattern Matching and Machine Learning Chatbots 88
2.3. Humanized Machines: Friendly Digital Assistants 90
2.3.1. Anthropomorphization and Socioemotional Interactions 90
2.3.2. An Exponential, Yet Subtle Ubiquity 91
2.4. Emotional Involvement With Machines 92
2.4.1. Please Say It Back: Bonds in Human-Machine Relationships 92
2.4.2. The Affectionate Illusion of the ELIZA Effect 93

3. DECODING GENDER IN AI: PERCEPTION, REPRESENTATION AND BIAS REVEALED 95
3.1. AI Becomes Her: Digital Moms, Secretaries and Caregivers 95
3.1.1. Traditional Feminine Jobs and Current Digital Assistants 95
3.1.2. Automating Domesticity in the Private Sphere 97
3.2. Human-Machine Interaction: Stereotypical Gendered Behavior Patterns 98
3.2.1. Scripting Stereotypical Gender Behavior in Social Interactions 98
3.2.2. Simulated Carevigeness and Artificial Emotional Bonds 99

3.3. Gendered Knowledge: Subjectivity, Bias and Phallogocentrism in Artificial Intelligence 102
3.3.1. Traditional Gender Approaches to Knowledge in AI 102
3.3.2. Gender Gaps in Digital Literacy 103
3.3.3. Data Collection and Questionable Algorithms 106

4. INSIGHTS FROM SPECIALIZED DISCUSSIONS AND MEDIA COVERAGE: UNPACKING GENDER ATTRIBUTION IN AI 111
4.1. What the Experts Say About Current AI Issues 111
4.1.1. Examining Current Questions, Suggestions and Ideas Surrounding Gender in AI 111
4.1.2. Biased Algorithms and Unpleasant Stereotypes 112
4.1.3. Lack of Diversity and the Illusion of Inclusivity 114
4.2. The Fallacy of Gender-Neutral AI 116
4.2.1. "I'm a Virtual Assistant, so I Don't Have Pronouns the Way People Do" 116
4.2.2. "I Live to Serve – But I Also Feel a Lot" 117
4.3. Justifications for Femininity in Digital Assistants 119
4.3.1. Tenderness, Warmthness and Trustworthiness: Feminized Voices 119
4.3.2. The Feminine Household and the Masculine Workplace 120
4.4. Concerns with Feminizing Digital Assistants 122
4.4.1. Linking Femininity with Assistance and Reinforcing Common Stereotypes 122
4.4.2. Perpetuating Gender Stereotypes and Influencing Future Generations 123
4.4.3. Instrumentalizing Femininity and Veiled Surveillance 125
4.5. Suggestions for Gender Attribution in AI 127
4.5.1. Diversifying Digital Assistants 127
4.5.2. Breaking the Artificial Binary: Towards Gender Fluidity and Ambiguity 128
4.5.3. Diversifying the Development Teams 129
4.5.4. Rethinking Gender Portrayals and Representation 130
4.5.5. Towards a Post-Binary and Queer AI 131
4.5.6. Lack of Guidelines and Regulations: More Public Debate Needed 132

4.6. Common Media Discussions on Current AI Issues 134
4.6.1. Failing Gender Neutrality 134
4.6.2. Comforting Feminine Voices, Lack of Diversity and Expected Stereotypes 135
4.6.3. Harmful Gender Stereotypes, Submissive Fantasies and Worried Parents 136
4.6.4. Diversification: Bring on the Experts 137
4.6.5. From Instrumentalized Femininity and Biased Data to Harmful Stereotypes and Gender Imbalance 138

5. ARTIFICIAL INTELLIGENCE IN FICTION: THEMES, PORTRAYALS AND INTERPRETATIONS OF GENDER 141
5.1. Fiction Is Burning Much Higher 141
5.1.1. Dystopian Terminators, Scary Cyborgs and Threatening Machines 141
5.1.2. Utopian Robots, Comforting Droids and Promising Technologies 151
5.2. Anthropomorphized Artificial Entities in Fiction 157
5.2.1. Disembodied Yet Gendered Hardware 157
5.2.2. Muscular and Beautiful Embodied Software 163
5.3. Artificial Femininity and Masculinity 164
5.3.1. Gender Stereotypes in Fictional AI 164
5.3.2. From Caregiving and Submissive Digital Housewives to Evil and Violent Tech-Butlers 165

6. ANALYZING DIGITAL ASSISTANTS: CONCEPTUALIZATION, CHARACTERIZATION AND DEVELOPMENT 169
6.1. "As an AI, I Don't Have a Gender, Thanks for Asking" 169
6.1.1. Analysis Components 169
6.1.2. Selection Criteria 171
6.1.3. Alexa, Cortana, Google Assistant and Siri 172
6.1.4. Methodology: Direct Observation and Trends of Development 178
6.2. Anthropomorphization, Tasks and Behavior 181
6.2.1. Towards Femininity 181
6.2.2. Towards Gender Diversification 183
6.2.3. Between Neutrality and Going Back to Femininity 187

6.3. Trends of Development 191
6.3.1. Ubiquity, Efficiency and Humanization 191
6.3.2. Neutralizing Gender Stereotypes and Privacy Concerns 192
6.3.3. To GPT Models and Beyond 192
6.3.4. Analyzing ChatGPT 193
6.4. Discussing Gender in Digital Assistants 196
6.4.1. Direct Observation: the Fallacy of Gender Neutrality 196
6.4.2. Trends of Development: Updated Personalities, User Data Collection and the GPT Era 200
6.5. Superficial Diversity and Gender Instrumentalization in the GPT Era 202

7. CISTEM NOT FOUND: A PROJECT ON GENDERLESS, FLUID AND NONBINARY GENDER POSSIBILITIES IN AI 207
7.1. A Deanthropomorphized, Ambiguously Neutral and Hybridly Transbinary Bot (https://genderai.pt) 207
7.1.1. Concept 207
7.1.2. Objectives 209
7.2. Revising Gender Archetypes and Analyzing Case Studies 211
7.2.1. Picking up Where We Left Off: Background Work, Updates and New Ground 211
7.2.2. The Personal and the Political: Informing the Bot's Monologue 216
7.2.3. Bots, Digital Assistants and Smart Homes: Framing the Bot's Personality 219
7.2.4. Posthuman Voices and Transhuman Production: Shaping the Bot's Characterization 223
7.3. Designing a Hybrid, Irreverent and Mutant Bot 227
7.3.1. Voice, Monologue and Personality Traits 227
7.3.2. Sound Design and Visualizations 229
7.3.3. Website 235
7.4. A Mutant Faith 240
7.4.1. Mutation 0: Prelude 240
7.4.2. Mutation I: Deanthropomorphized 243
7.4.3. Mutation II: Ambiguously Neutral 246
7.4.4. Mutation III: Hybridly Transbinary 249

7.5. Beyond Traditional, Cis-Binary and Bioessencialist Notions of Gender 253

CONCLUSION 257
Interpretation of Findings 262
Limitations 265
Future Research 266

BIBLIOGRAPHY 269

WORKS CITED 281

INDEX OF IMAGES

Fig. 1 : Gender stereotypes and expectations, as described by Naomi Ellemers (2018, 281). 58

Fig. 2 : SOPHIE and Sophia the Robot discuss their dreams, hopes and creativity in the age of AI. 142

Fig. 3 : Frankenstein's monster in *Young Frankenstein*, dir. Mel Brooks (1974). 144

Fig. 4 : Maria being created, in *Metropolis* (1914). 144

Fig. 5 : The robots in *R.U.R.*'s first television adaptation, by BBC (1938). 144

Fig. 6 : A robot from *I, Robot*'s film adaptation, dir. Alex Proyas (2004). 146

Fig. 7 : Marvin, *The Hitchhiker's Guide to the Galaxy*'s film adaptation, dir. Garth Jennings (2005). 146

Fig. 8 : Book cover from *Do Androids Dream of Electric Sheep?* by Panther Science Fiction (1977). 146

Fig. 9 : Ash after being defeated by the spaceship crew, in *Alien* (1979). 147

Fig. 10 : Book cover from *Neuromancer* by Ace Science Fiction (1984). 148

Fig. 11 : Book cover from *The Diamond Age* by Bantman (1995). 148

Fig. 12 : T-800 from the *Terminator* (1984). 149

Fig. 13 : Neo takes the red pill and wakes up in the real world, in *The Matrix* (1999). 149

Fig. 14 : Ava from *Ex Machina* (2016). 150

Fig. 15 : The creation of the robots in *Westworld* (2016). 150

Fig. 16 : Chell and GLaDOS from *Portal* (2007). 150

Fig. 17 : Major Motoko Kusanagi from *Ghost in the Shell* (1995). 151

Fig. 18 : C-3PO and R2D2 from *Star Wars* (1977). 152

Fig. 19 : Wall-E, the waste-collecting robot, and EVE, the advanced robot sent to Earth to search for signs of life (2008). 152

Fig. 20 : Auto, the evil AI in *Wall-E* (2008). 152

Fig. 21 : Kara, Connor, Markus and an android store in *Detroit: Become Human* (2016). 153

Fig. 22 : David and another AI, in *A.I. Artificial Intelligence* (2001). 154

Fig. 23 : Samantha from *Her* (2014). 154

Fig. 24 : Janet, the AI assistant in *The Good Place* (2016). 155

Fig. 25 : The cyborg law enforcer from *Robocop* (1987). 155

Fig. 26 : B1 Battle Droids from *Star Wars: Attack of the Clones – Episode II* (2002). 155

Fig. 27 : General Grievous from *Star Wars: Revenge of the Sith – Episode III* (2005). 156

Fig. 28 : Data from *Star Trek: The Next Generation* (1987). 156

Fig. 29 : Bishop sacrificing itself to protect the protagonist from the Xenomporhs (2016). 156

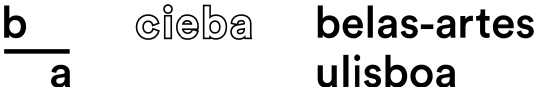
Fig. 30 : Joi from *Blade Runner 2049* (2017). 158

Fig. 31 : Cortana in *Halo 4* (2012). 158

Fig. 32 : Karen from *Spider-Man Homecoming* (2017). 158
Fig. 33 : Theodore waiting for Samantha to boot up, in *Her* (2013). 159
Fig. 34 : HAL 9000 from *2001: A Space Odyssey* (1968). 159
Fig. 35 : Athena in Recall, an *Overwatch* animated short (2016). 161
Fig. 36 : The PDA from *Subnautica* (2018). 161
Fig. 37 : GLaDOS and Wheatley from *Portal 2* (2011). 161
Fig. 38 : Echo, a playable AI in *Overwatch 2* (2022). 162
Fig. 39 : Cortana's avatar, in *Halo 4* (2012). 162
Fig. 40 : Alexa's avatar (2019-2023). 173
Fig. 41 : Cortana's avatar (2019). 173
Fig. 42 : Google Assistant's avatar (2019-2023). 173
Fig. 43 : Siri's avatar (2019-2023). 173
Fig. 44 : Retrieved from Cortana's iOS app using the English US voice (2019). 182
Fig. 45 : Retrieved from Google Assistant's iOS app using the English US voice (2019). 182
Fig. 46 : Retrieved from iOS 11 using the English US voice (2019). 183
Fig. 47 : Retrieved from Cortana's Microsoft app (2021). 185
Fig. 48 : Retrieved from Alexa's iOS app using the English US voice (2021). 185
Fig. 49 : Retrieved from Google Assistant's iOS app using the English US voice (2021). 186
Fig. 50 : Retrieved from iOS 15 using the English US voice (2021). 186
Fig. 51 : Retrieved from Alexa's iOS app using the English US voice (2023). 189
Fig. 52 : Retrieved from Google Assistant's iOS app using the English US voice (2023). 189
Fig. 53 : Retrieved from iOS 17 using the English US voice (2023). 190
Fig. 54 : Chat GPT's avatar (2023). 195
Fig. 55 : Retrieved from ChatGPT's iOS app using the English US voice (2023). 195
Fig. 56 : Website of *Conversations with ELIZA* (2019). 212
Fig. 57 : The chatbots of *Conversations with ELIZA*: Assistant, Cybele, Iynx and Electra (2019). 213
Fig. 58 : Stills from *How Not to Be Seen: A Fucking Didactic Educational .MOV File* (Steyerl 2013). 216
Fig. 59 : Collage for *A Cyber Feminist Manifesto for the 21st Century* (VNS Matrix 1991). 217
Fig. 60 : Still from *Bitch Mutant Manifesto* (VNS Matrix 1996). 217
Fig. 61 : Short texts from *Inflammatory Essays 2* (Holzer 1979-82). 218
Fig. 62 : Electronic sign from *Survival, Protect Me From What I Want* (Holzer 1983-85). 218
Fig. 63 : ELIZA's Rogerian psychotherapist program (1966). 219
Fig. 64 : Photo of *LAUREN* (McCarthy 2017). 220
Fig. 65 : Photo of *SOMEONE* (McCarthy 2019a). 220
Fig. 66 : Photo of *You Can Say "Reset the Room"* (McCarthy 2020). 221
Fig. 67 : Photo of *The Changing Room* (McCarthy 2021). 221

Fig. 68 : Photos of *The Helpless Robot* (White 1987-2002). 222
Fig. 69 : The covers of *KiCk I* (Arca 2020b), *KICK ii, KicK iii, kick iiii* and *kiCK iiiii* (Arca 2021d, 2021e, 2021f, 2021g). 224
Fig. 70 : Silicon bubble case and toy from *PRODUCT* (SOPHIE 2013-15). 225
Fig. 71 : The cover of *OIL OF EVERY PEARL'S UN-INSIDES* (SOPHIE 2018). 225
Fig. 72 : Unedited stills from the voice's visualization. 230
Fig. 73 : Unedited stills from the musical and sonic elements' visualizations. 231
Fig. 74 : Final results after combining and overlapping the visualizations (*Mutation I*). 232
Fig. 75 : Final results after combining and overlapping the visualizations (*Mutation II*). 233
Fig. 76 : Final results after combining and overlapping the visualizations (*Mutation III*). 234
Fig. 77 : Home page of *Cistem Not Found*'s website. 235
Fig. 78 : Each mutation page is structured into three sections: the first announces the mutation, the second provides context including the track listing and the third integrates the video. 236
Fig. 87 : Web page of *Mutation II: Ambiguously Neutral* – https://genderai.pt/. 237
Fig. 88 : Page of *Mutation III: Hybridly Transbinary* – https://genderai.pt/. 238
Fig. 89 : The *About* section offers additional information on *Cistem Not Found* within the context of the PhD research https://genderai.pt/. 239
Fig. 79 : Stills from *Mutation 0: Prelude*. 241
Fig. 80 : More stills from *Mutation 0: Prelude*. 242
Fig. 81 : Stills from *Mutation I: Deanthropomorphized.* 244
Fig. 82 : More stills from *Mutation I: Deanthropomorphized.* 245
Fig. 83 : Stills from *Mutation II*: Ambiguously Neutral. 247
Fig. 84 : More stills from *Mutation II*: Ambiguously Neutral. 248
Fig. 85 : Stills from *Mutation III: Hybridly Transbinary.* 251
Fig. 86 : Stills from *Mutation III: Hybridly Transbinary.* 252

This work was developed with the financial aid of the Foundation for Science and Technology (FCT) and the European Union (EU), via the Artistic Studies Research Center (CIEBA) (BD/09239/2020)



GENDER TERMINOLOGY

Throughout our literature review, we have observed that terms such as *man*, *woman*, *boy*, *girl*, *masculine*, and *feminine,* when used in isolation, often imply a cisgender and binary perspective, inadvertently excluding trans and other gender understandings. This study uses the prefixes *cis* and *trans* to specify the type of gender conceptions it is discussing.

Accordingly, the prefix *cis* (short for cisgender) denotes individuals who identify with the gender they were assigned at birth, typically used in reference to *cis men* and *cis women*. The prefix *trans* (short for transgender) refers to those who do not identify with their assigned gender at birth, serving as an umbrella term that encompasses *trans men*, *trans women* and *nonbinary* individuals.

The term *cis-binary* pertains to cisgender understandings of gender according to a *cis man-woman* dichotomy. Within this binary framework, *cis-masculinity* and *cis-femininity* describe the types of masculinity and femininity associated with men and women, respectively. *Cis-normative* describes understandings of gender that reinforce traditional and naturalistic conceptions of men and women, within the binary frame.

Nonbinary refers to transgender individuals who reject the cis-binary gender framework, challenging normative and bioessentialist conceptions of masculinity and femininity. Instead, they embrace diverse understandings of gender, proposing an alternative nonbinary gender system that transcends traditional binary categorizations and associations between gender, sex and identity. Not to be confused with *gender neutrality* or *gender fluidity.*

Gender neutrality is a term commonly used in discussions about inclusivity and typically refers to something that is not specific or associated with either men or women, such as *gender-neutral* language that avoids implying a gender.

Gender fluidity is a term used to describe gender expressions that dynamically navigate between masculinity and femininity, avoiding being fixed to a single gender and, instead, encompassing a wider spectrum of gender identities and presentations.

INTRODUCTION

Artificial intelligence as a technology has bloomed alongside current advancements in machine automation, allowing humanity to realize age-old dreams of creating artificial beings in its image. Throughout history, we have seen glimpses of feminized subjects emerging through the crafting of these simulated entities, from ancient myths like Pygmalion, who created and married an automaton-like statue named Galatea, to modern innovations like Alan Turing's imitation game, where machines deceived human subjects and proved their intelligence by pretending to be a woman, and Joseph Weizenbaum's ELIZA program, which simulated a psychotherapist that elicited emotional responses from users.

Today, this intersection between gender and AI is particularly prominent among digital assistants such as Alexa, Google Assistant and Siri, which have seamlessly integrated our daily lives. These entities are no longer mere assistants but also friendly companions that talk to us, listen to us, help us and have become a natural part of our daily interactions. They engage users in natural language interactions, perform a myriad of tasks and enact friendly, caregiving and cheerful personas. However, as they are assigned human-like traits, behaviors or even personalities, these digital entities inevitably embody gendered characteristics that often conform to traditional stereotypes and reinforce cis-normative binary conceptions of gender.

This study delves into the relationship between gender and artificial intelligence, aiming to question, understand and expose this phenomenon both in theory and in practice. It aims to examine a tendency towards cis-femininity in current digital assistants and which gender roles and stereotypes are perpetuated in this process. By shedding light on the

implications of assigning gender to AI entities, this research also seeks to promote discussion and raise awareness on the sociocultural implications of this phenomenon, exploring the potential of queering these entities to disrupt traditional norms and move towards more informed, conscious and diverse design practices.

Background

Artificial intelligence is no longer a thing of the future, with digital assistants becoming more ubiquitous across mobile devices and online services (Dale 2016). These assistants are gradually humanized, thus evolving from mere assistants to daily companions, developing emotional bonds with their users (Weizenbaum 1976, Richardson 2015). This intention of turning their interactions more natural encompasses a growing anthropomorphization, which reveals the presence of gender attributes.

While traditional notions of gender primarily perceive it through a cis-binary perspective, discussions in the context of gender studies and queer theory challenge this binary framework and expand our understandings of gender by acknowledging the existence of trans and nonbinary identities (Feinberg 1992, Preciado 2019, Richards et al. 2017). Nonetheless, cis-binary stereotypes and roles persist and certain acts, tasks and even jobs are identified as inherently masculine or feminine, perpetuating structural hierarchies of labor (Butler 1990, Prentice and Carranza 2002, Hester 2016, Preciado 2008).

Accordingly, we can observe how digital assistants currently automate traditionally feminine jobs that reflect upon their tasks of service, assistance and emotional labor (West and Zimmerman 1987). Gender is also present in their voice, name and avatar, and their behavior often conforms to "stereotypical and gendered behavior patterns" as they fill the roles of caregivers and other roles coded as feminine in western society (Weber, 2005). As such, we end up perceiving these entities not only as mere machines but also as "mirrors or substitutes", and the way we relate to our peers starts influencing how we relate to artificial intelligence and how it relates to us (Strengers & Kennedy 2020).

Current trends in their development are not naive regarding this phenomenon, although corporations are more focused in further anthropomorphizing and humanizing these entities. And while these assistants appear to be trying to become more diverse and unbiased in their characterization, a tendency towards cis-femininity prevails. Observing this phenomenon, researchers and academics highlight the way gender (and, by extension,

femininity) is instrumentalized to manage interactions between digital assistants and users (Piper 2016, Bergen 2016, UNESCO 2019, Strengers & Kennedy 2020). In turn, common debates often advance user preference as a justification for feminized AIs and even popularize the belief that it's due to the field being mostly developed by cisgender men (Chambers 2018, Robison 2020, Purtill 2021). In both contexts, gender neutrality is seen as illusory since anthropomorphized virtual assistants inevitably engage with traditional and binary assumptions of gender.

In this sense, there is little agreement on how to tackle these issues, and gender neutrality is often questioned in favor of gender diversity, namely, allowing the user to customize their assistant. Although diversifying these entities with more options could be a way to counter femininity in AI, it still perpetuates gender conceptions according to a binary framework, eventually reinforcing cultural stereotypes. Thus, gender fluidity emerges as a promising path, with some authors pointing out how, instead of replicating manly and womanly attributes through their anthropomorphization and behavior, these entities could move away from these interpretations of gender and explore fluid or ambiguous possibilities (Strengers & Kennedy 2020).

Purpose and Objectives

This study aims to understand, question and explore the relationship between gender and artificial intelligence. It focuses on an observable tendency towards cis-femininity in current digital assistants, speculating on the possibilities that arise with queering these entities and reframing them outside of a cis-normative binary framework.

To this end, we aim to provide a comprehensive understanding of the sociocultural underpinnings of gender and the evolution of artificial intelligence. We focus on gender manifestations beyond the cis-binary frame, seeking to expose how traditional gender conceptions perpetuate stereotypes and hierarchical structures of labor. Turning to the integration of artificial intelligence into our daily lives, we aim to inspect its evolution and embodiment in digital assistants while also unraveling the potential biases embedded into their algorithms. Building upon these observations, we intend to uncover how and why gender is manifested and enacted within these systems. Complementing this debate, we draw attention to ongoing discussions on the phenomenon within academic discourse and online media contexts to highlight the main questions, concerns and suggestions surrounding gender attribution in AI. We extend our approach to pop culture to

unveil how fictional depictions of this technology reflect and reinforce traditional binary gender conceptions.

A second direction of inquiry concerns an analysis that intends to verify how gender is revealed and embodied by current digital assistants through their presentation, interaction and development. We begin by scrutinizing their anthropomorphized and behavioral traits to uncover how gender is enacted through more apparent features such as names, voices and avatars but also through the tasks they perform and their display of socioemotional interactions. This examination aims to demonstrate how digital assistants are gendered in their design and functions, emphasizing whether gender norms and stereotypes are reinforced or challenged within this context. We then investigate the features and functionalities that are being prioritized in their current trends of development. This approach seeks to assess how current debates and discussions surrounding these technologies influence the trajectory of their design while also shedding light on the underlying assumptions and biases that guide their development.

Lastly, the project *Cistem Not Found* (https://genderai.pt) intends to reflect the ways in which our reasoning and understanding of these topics were shaped throughout the theoretical and analytical components. It seeks to raise awareness on the implications of assigning gender to digital assistants, fostering debate and inciting reflection on the intersections of current AI technologies and cis-normative binary gender conceptions. Aiming to provide critical commentary on current trends of AI and gender attribution in digital assistants, we develop a provocative, irreverent and mutant bot. We explore alternative voice proposals to challenge traditional gender stereotypes and contemplate ways to counter this phenomenon. The bot's characterization, monologue and personality traits seek to speculate on the possibilities of deanthropomorphizing these entities, reflect on so-called neutral approaches to gender and explore traits that evoke queerness and otherness to push boundaries of conventional gender norms in AI.

Methodology

We tackle these objectives according to three approaches: a theoretical component, which structures an analysis, both complementing and informing the project's development.

The theoretical approach starts by confronting the sociocultural understandings of gender with the evolution of artificial intelligence and its integration in our daily life. We draw from gender studies, queer perspectives and intersectional frameworks to discuss conven-

tional binary understandings of gender and explore trans, nonbinary and nonconforming gender identities as examples of alternative gender possibilities. Secondly, we review computational and new media studies to map the evolution of artificial intelligence and its integration into our daily life in the shape of digital assistants while also discussing the potential bias they convey. By combining these approaches, we examine how gender emerges in current AI systems through their anthropomorphized attributes, tasks and behavior. We turn our attention towards academic discourse and online media contexts to debate the main questions, concerns and suggestions surrounding these technologies and gender attribution in digital assistants. Finally, by mapping portrayals of gendered AI across literature, cinema television or gaming, we discuss how fictional representations of these technologies often mirror and perpetuate conventional norms of masculinity and femininity.

The analytical approach revisits and updates our analysis of Alexa, Cortana, Google Assistant and Siri to examine how gender manifests through their characterization, interaction and trends of development.[1] Through direct observation, we focus on their anthropomorphization, including features such as names, voices and avatar. We analyze the tasks they perform, particularly those associated with historical feminine labor, and look into their humanized, gendered behavior, paying particular attention to interactions that suggest a caregiving attitude and conform to common stereotypes. Secondly, we examine how these entities tend to evolve in their portrayal of gender, according to the functionalities and features that are being prioritized in their development as promoted by their official websites and announcements. Finally, we compare our findings across 2019, 2021 and 2023, discussing how their gender portrayals and functions have evolved overtime. Additionally, we extend our analytical approach to ChatGPT, exploring its anthropomorphization (or, rather, lack thereof) and its responses to social and emotional prompts. Although ChatGPT differs from voice-activated personal assistants by primarily operating as a language model processing system, we are interested in examining how it handles anthropomorphization and ideas of companionship.

The practical approach involves the development of an irreverent, disruptive and mutant chatbot to explore genderless, ambiguous and nonbinary possibilities within AI systems.

---

1 This analysis was first conducted within the context of the Master's degree in Communication Design and New Media at the Faculty of Fine-Arts, University of Lisbon, and initially focused on Alexa, Cortana and Siri, along with their anthropomorphized attributes, tasks and behavior (Costa 2019). In subsequent studies and publications, we expanded our approach to include Google Assistant while also broadening our examination to encompass their trends of development (Costa & Ribas 2019, Costa & Ribas 2020).

Across four audiovisual pieces, we focus on designing the bot's voice, personality traits and monologue, ironically blurring masculine and feminine archetypes while also tampering with assistance and service-related stereotypes. The bot explores and speculates on possibilities to deal with gender attribution in AI by emphasizing its deanthropomorpized attributes, exploring observable neutral and ambiguous approaches to gender or even experimenting with traits that evoke trans, nonbinary or even posthuman identities. The bot's limited, dated and convoluted persona intentionally opposes trends of anthropomorphizing digital assistants into friendly companions, instead providing critical commentary on the implications of assigning gender to these entities and the integration of this technology in our daily life. Through its speculative and critical approach, *Cistem Not Found* seeks to incite reflection on the issues that arise with artificial intelligence's trends and the way current digital assistants reinforce traditional and normative cultural notions back to us.

Significance

Ongoing discussions surrounding the sociocultural consequences of AI integration into our daily life within political contexts, academic studies and online media all highlight the importance of addressing gender issues and questioning associated stereotypes within the context of digital assistants. The rapid development of AI often eludes critical stances on the social and cultural roots that inform its evolution, namely outpacing current debates that are still reshaping our understanding of gender. While some guidelines proposed by entities such as UNESCO and the European Commission are starting to take into account social and cultural issues surrounding the anthropomorphization of AI, they primarily focus on avoiding discrimination and promoting inclusion without providing clear direction on how to address the unique and specific questions that emerge with attributing gender to humanized digital companions.[2] Moreover, these guidelines tend to prioritize implicitly cis-normative perspectives of gender, ignoring trans, nonbinary and nonconforming gender identities and overlooking gender understandings outside a traditional cis-binary framework.

2 In *I'd Blush if I Could*, UNESCO outlines the persistent and significant gender gap in digital skills and provides recommendations to help women and girls develop strong digital skills, while also examining how feminized AI voice assistants perpetuate harmful gender biases, offering suggestions to ensure that the continued proliferation of digital assistants does not widen gender divides (2019). The European Union's *Ethics Guidelines for Trustworthy AI* propose a set of key requirements that AI systems should meet, such as empowering and benefiting all human beings, as well as avoiding unfair bias that marginalizes vulnerable groups and exacerbates prejudice and discrimination (European Commission 2019).

We believe it is relevant to delve deeper into the implications of the widespread use of gendered digital assistants, pursuing this objective through an interdisciplinary approach that offers a critical examination of the phenomenon in an inquisitive, exploratory and speculative manner. Rather than providing definitive solutions or guidelines, our study seeks to raise awareness and incite reflection on the complexities of gender attribution in AI. As such, our study aims to provide a more nuanced understanding of this phenomenon by considering the complexities of identity within intersecting social contexts alongside the impact of this technology in our daily life. We hope to contribute to a more thorough understanding of the role these technologies play in shaping and perpetuating social norms, informing designers that contribute to the development of this technology but also the users of their creations, namely those that use these digital assistants on a daily basis.

The study's practical approach also intends to disrupt and question cisgender notions and stereotypes of femininity and masculinity as well as so-called neutrality in AI, exposing how they have its basis in exclusionary design practices and shedding light on the profound and complex specificities of gender, sex and identity. Through its different audiovisual pieces and voice proposals, the project engages in speculative scenarios where the boundaries between human and machine are blurred, pushing traditional boundaries and challenging bioessencialist notions of humanness. By reimagining gender possibilities beyond the cis-binary frame, *Cistem Not Found* seeks to raise awareness and foster discussion on the issues that arise with gender attribution in AI.

Ultimately, the development and design of artificial intelligence requires a nuanced, conscious and intersectional approach that addresses gender discrimination, bias and stereotypes, in order to bridge existing gaps and ensure that advancements in this technology align with evolving understandings of gender.

Overview

The first chapter is dedicated to an overview of the sociocultural roots of gender, sex and identity. It begins by exploring trans and nonconforming gender identities, emphasizing a queer and intersectional approach. It then discusses how gender and sex become intertwined through bioessentialist ideas that reinforce cisgender norms within a binary frame and perpetuate gender roles and stereotypes. These conceptions are examined alongside patriarchal and phallocentric hierarchies of labor.

The second chapter examines artificial intelligence's evolution and its integration into our daily life in the shape of digital assistants. It discusses computationalism and machine automation, drawing from Turing's imitation game to highlight early concepts of this technology. It then considers early applications like ELIZA, which demonstrated human intelligence in a restrictive manner while also evoking a sense of companionship. These concepts guide the development of current digital assistants, designed to cultivate affectionate bonds with their users, gather personal data and become increasingly ubiquitous.

The third chapter builds upon previous findings, discussing how current AI systems, particularly digital assistants, often embody traditionally feminine traits and tasks, such as caregiving and service-related roles. It questions how this tendency towards feminization, seen in their anthropomorphized attributes and behaviors, can shape users' perspective of gender and perpetuate stereotypes. Additionally, AI's reliance on data poses a risk of perpetuating biases, leading to discriminatory algorithms.

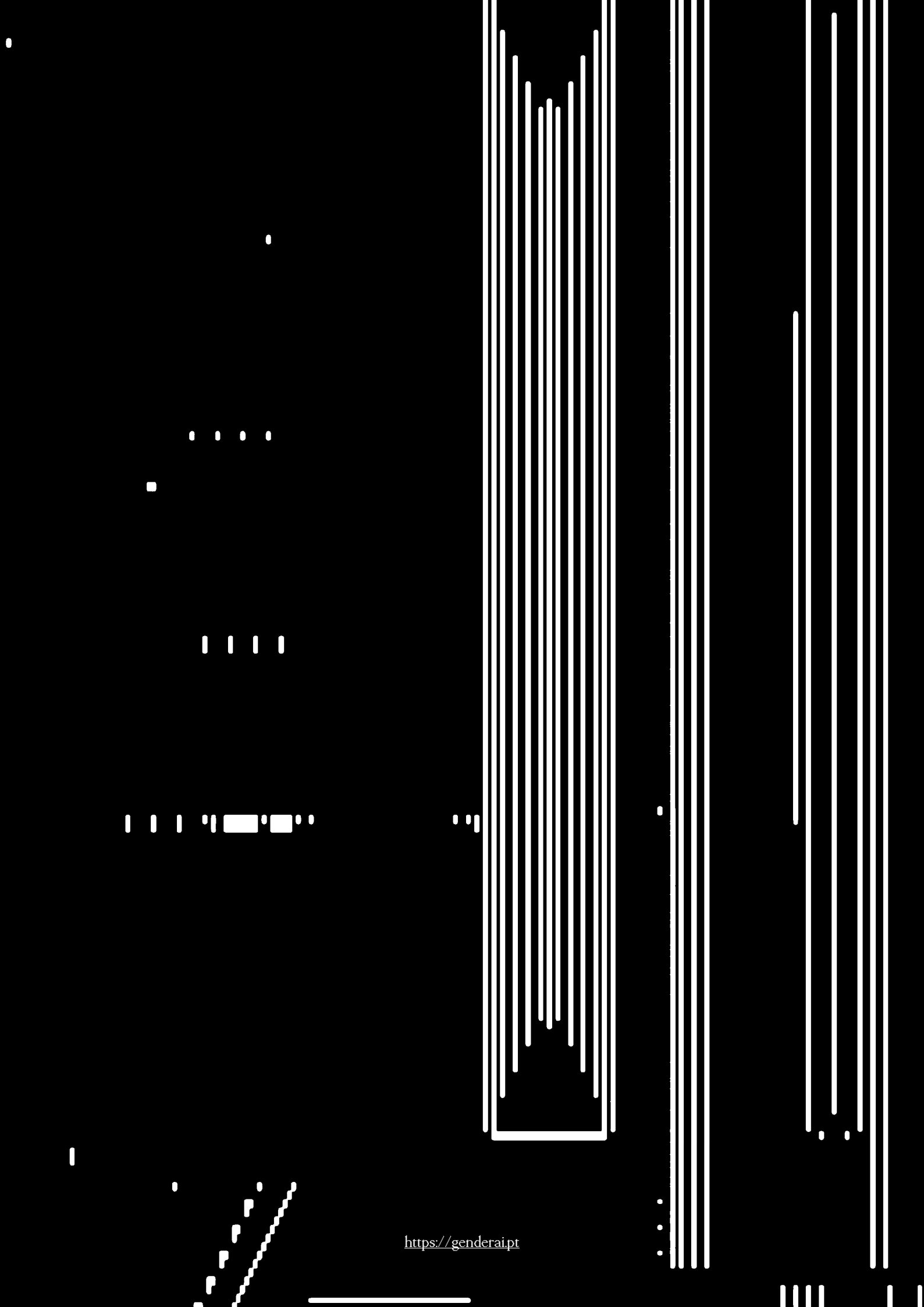
Complementing this debate, the fourth chapter explores the main questions, concerns and suggestions within academic discourse and online media contexts. The feminization of AI is discussed alongside the fallacy of gender neutrality and the perpetuation of stereotypes about femininity in assistance roles. Justifications for femininity evoke user trust and preference, while current suggestions include adding alternative voice options and diversifying development teams. Despite efforts, it is argued that digital assistants often instrumentalize cis-binary conceptions of gender and attempts at gender ambiguity oversimplify trans and nonbinary identities.

The fifth chapter extends our approach to pop culture across various media formats, examining fictional depictions of AI in light of traditional notions of femininity and masculinity. These normative conceptions seem to shape the personalities and roles of AI characters but, alternatively, fiction also provides opportunities to challenge traditional norms and imagine alternative, queer futures.

The sixth chapter provides an updated analysis of Alexa, Cortana, Google Assistant and Siri. While efforts have been made to diversify these assistants, we observe the predominance of cis-feminine personas and stereotypes. Concerns about perpetuating gender norms persist, particularly in the tasks they perform, and consensus on addressing this phenomenon remains elusive, with some assistants undergoing significant changes but others preserving stereotypes. The analysis is extended to ChatGPT, discussing its avoid-

ance of feminized personas through a deanthropomorphized and dehumanized characterization, while also noting that the technologies of GPT and personal digital assistants are likely to converge.

Motivated by our theoretical and analytical findings, the final chapter delves into the practical aspects of our project *Cistem Not Found*. Through the development of a disruptive chatbot, it explores alternatives to traditional gender norms, including genderless options and queer approaches. It also provides critical commentary on current trends surrounding digital assistants, fostering debate on the social and cultural norms that inform AI development. This final chapter also provides new ground for further exploration on otherness, queerness and posthumanism to move beyond conventional and bioessencialist notions of gender.


https://genderai.pt

1. UNVEILING GENDER: IDENTITY, DISSIDENCE AND NORMATIVE CONSTRUCTS

1.1. LET'S TALK ABOUT GENDER, BABY

> *The taxonomy that we're using to classify living human bodies is not working. Ideally, we should move to a much more open form of thinking where the shape and form of genitalia, as well as the fact of possessing a reproducing uterus or not, can't be the condition for assigning gender at birth. One day, we'll see assigning gender at birth as brutal and unjustified as assigning religion at birth.*
>
> Paul B. Preciado 2020b

1.1.1. Gender Is Plastic (It's Fantastic)

Gender encompasses one of the elements through which we explore and define our own identity. Through gender, we express an image that reflects how we deal with our identity on a more personal and intimate level but also how we deal with a broader social and cultural conception of these attributes.

Even though gender is frequently associated with the physical body, it also constitutes an instrument and something that one possesses, as opposed to a fixed and rigid attribute of

our identity: "gender is not a fact, the various acts of gender create the idea of gender, and without these acts, there would be no gender at all" (Butler 1988, 522).

In this manner, gender manifests at two levels: physically and behaviorally. On one hand, gender allows us to define our physical body, for example through certain fashion stylizations, clothes or accessories but also through body changes involving prosthesis, plastic surgery or even hormones. Thus, the body can be distinctively adjusted and presented in relation to its surrounding spaces. On the other hand, gender is also present at a behavioral level, as it relates to certain attitudes, postures and roles, involving "a complex of socially guided perceptual, interactional and micropolitical activities that cast particular pursuits as expressions of masculine and feminine natures" (West 1987, 126). This physical and psychological performance thus defines gender identity and, without it, "acts, gestures, the visual body, the clothed body, the various physical attributes usually associated with gender express nothing" (Butler 1988, 530).

In this sense, we can observe how gender is performative, since "what we take to be an internal essence of gender is manufactured through a sustained set of acts, posited through the gendered stylization of the body (...) [and] what we take to be an internal feature of ourselves is one that we anticipate and produce through certain bodily acts" (Butler 1990, 27). There is also a plastic materiality to gender and Paul B. Preciado adds that, while we might be constructed through drugs, objects and representations, we can also construct ourselves through them, emphasizing that our bodies and the biological are "fictions" that can be written and rewritten (2008). Accordingly, masculinity and femininity, as well as the states of being male and female "exist only as political fictions, as somatic effects of the technical process of normalization" (2008, 142).

Gender is not something definitive or fixed and "simplistic and rigid gender codes are neither eternal nor natural (...) they are changing social concepts" (Feinberg 1992, 5). It doesn't just manifest itself in masculine or feminine aspects, but also in a liquid way and its expression can be easily transformed and adapted. According to the xenofeminist manifesto, "nothing should be accepted as fixed, permanent or given – neither material conditions nor social forms" and anyone who's been deemed "unnatural" and experienced injustices in the name of biological norms or a natural order, will realize that "the glorification of *nature* has nothing do offer us – the queer and trans among us, the differently-abled, as well as those who have suffered discrimination" (Hester 2018, 29).

Authors like Christina Richards, Walter Pierre Bouman and Meg-John Barker also propose that we are born genderless, since "babies are blank slates on which gender identities are imposed and encultured upon the being" (2017, 57). It is later, through education and socialization, that we are taught a dominant gender notion, referring to the sociocultural context in which we grow up.

Butler and Richards also reclaim Foucault's notion of sexuality and problematize the norm and how certain mandatory gender notions that socially proliferate stem from normative conceptions that generate and build knowledge through universal truths that are culturally imposed. Foucault considers that there are no universal or objective truths but they are established and imposed as such at a political scale according to the set of values society imposes without taking into account the true and inherent essence of gender (1976).

Gender possesses a fluid nature, inhabiting a spectrum as opposed to being defined by two poles. Thus, it is possible to freely navigate between attributes or even explore other types of gender expression that don't necessarily echo the polar categories. Essentially, there's nothing in the nature of gender that implies a single and unequivocal way to experience and express it.

1.1.2. Beyond Binary Binds: Gender Fluidity, Ancestry and Infinity

We can observe how certain individuals exist outside of socially dominant and normative cis notions of gender, self-identifying as trans or nonbinary.[3] This means their identity isn't compelled to act according to a stabilized and pre-defined notion of gender and, through their performance, it becomes possible to map gender and the body according to new configurations, outside of the binary frame.

This idea of gender as something fluid is not recent, nor is it an uncommon social phenomenon. Interpreting gender as something experienced individually and culturally in exclusively feminine and masculine ways is a western notion, historically and culturally normalized.

---

3 *Cis* or *cisgender* refers to people that identify with the gender they were assigned at birth, while *trans* or *transgender* refers to people that don't identify with the gender they were assigned at birth.

The very notion of *trans* identities can be observed throughout history, namely in precolonial societies linked to paganism before the 11th century, that integrated "lesbian, gay, bissexual and straight[,] many women were among its practitioners[,] and transness was still a part of virtually all rural festivals and rituals" (Feinberg 1992, 12).

In fact, there are different interpretations and understandings of gender throughout time and space, without restrictions or concrete rules. And although a tendency to divide gender into man and woman based on genitalia prevails, historically there have been other factors that have significantly influenced its definition.

For example, in Western countries like the United Kingdom, in the late 19th century, the relationship between gender and sexuality became more definitive than sex, since any transgression in terms of sexual practices also generated doubt or opened space for questioning gender identity. Accordingly, gay men were perceived as women trapped in a male body and lesbian women were perceived as men trapped in a female body (Richards, Bouman and Barker 2017, 13). Although gender and sexuality variations are currently conceived separately, we can see how historically they intersected and how there hasn't always been a clear and direct link between gender and genitalia.

According to Leslie Feinberg, "a glance at human history proves that when societies were not ruled by exploiting classes that rely on divide and conquer tactics, cross gendered youths, women and men on all continents were respected members for their communities" (Feinberg 1992, 7). There were even societies dating back to 25,000 BCE where "worshipped figures were goddesses, not gods [and] some of the deities were trans, as were many of their shamans or religious representatives" (Feinberg 1992, 8).

Thus, we can observe multiple gender variations in Western and non-western contexts, positioned as other from men or women. Borrowing from Richards, Bouman and Barker's mapping and description of some identities that challenge binary notions of gender throughout history, we now expand on these and provide a range of specific examples.

1.1.3. Trans and Nonconforming Gender Identities Across the World

The term *molly* was used in the 18th century in the English society to label men that experienced attraction to other men. Additionally, *mollies* "engaged in gendered practices which positioned them as critically separate from men, including taking of female names

and titles, marriage ceremonies and ritualized enactments of giving birth" (Richards, Bouman and Barker 2017, 14). These individuals experienced hostility, since sodomy was punishable by death and gaining this reputation also implied social exclusion.

Femminielli are associated with Neapolitan culture, sharing some similarities with the archetypal expectations of trans women, such as "taking a female name and accessing hormones and gender-affirming surgeries". Additionally, they could "articulate themselves in specific ways which would be frowned upon if done by women or men (…) thus allow *femminielli* to occupy a third gender position" (Richards, Bouman and Barker 2017, 15). *Femminielli* are thought to be special and bring good luck and are an important part of the city's culture.

In Albanian tribes, having a son was of great cultural, social and financial importance since only men were eligible to inherit property. *Sworn Virgins* or *Burrnesha* refer to the practice of making a male heir out of a child "assigned female at birth", common between the 15th and the 20th century (Richards, Bouman and Barker 2017, 16). People assigned female at birth swore an irrevocable oath and dressed and lived as men, enjoying the rights and privileges of men. They would also give up their sexual and reproductive identities through chastity.

Eunuchs present in different cultures with specific functions, and although they often refer to an idea of castrated men, in Byzantine culture they are originally portrayed as "anyone who did not as well as could not produce children, including men who were born sterile, men who became sterile through illness, accident or birth defect, men who were lacking in sexual desire and men and women who embraced the celibate life for religious reasons" (Richards, Bouman and Barker 2017, 17). Thus, we can consider that they occupy a separate gender category, creating a "gray area in some historical contexts" (Richards, Bouman and Barker 2017, 17).

In contexts outside Europe, *Hijras* refers to a gender identity situated in relation to the caste system and the practice of Hinduism throughout the Indian subcontinent, exemplifying how "intersections between class, culture, faith and other factors can shape different identity categories" and *hijra* defines neither man nor woman (Richards, Bouman and Barker 2017, 19).

In Thailand, *Kathoey* identifies individuals who consider themselves "simply as female, whilst others as a second kind of female, others yet simply as *kathoey*, articulated as a third category separate from male and female" and "other something than male" (Richards, Bouman and Barker 2017, 20).

In Indonesia, *Waria* identifies individuals who are born male but realize that "they have the soul of a woman, or at least a soul that is more woman than man, the goal is not to pass but to look like a waria" (Richards, Bouman and Barker 2017, 22).[4] *Bugis People*, also known as Buginese, in the context of an Indonesian island, refers to an ethnic group that has a social system that accommodates five genders: "two of the five genders of the Buginese are the *oroané* and *makkunrai* which are analogous to cisgender men and women, respectively (…) the further three genders include the *calabai*, who are individuals assigned male at birth but who live as heterosexual women, whilst the *calalai* are assigned female at birth but live as heterosexual men (…) finally there are the *bissu*, who are considered a combination of man and woman within an empowering cultural-spiritual narrative" (Richards, Bouman and Barker 2017, 22).

Two-Spirit is a term used to define a set of roles and identities among different North American indigenous tribes, including, amongst others, the Winkte of the Lakota, the Nàdleehi of the Navajo and the Badés of the Sioux. That way, "some tribes might understand the individual to possess two identities articulated at different times, thus the origin of the name two spirit – this allowed for men's clothing to be worn on some occasions and women's clothing on others […] two-spirit people would occupy positions of social respect and were often thought to have spiritual powers" (Richards, Bouman and Barker 2017, 23).

In Chile and Argentina, *Machi* refers to the indigenous Mapuche group and, like *Hijra* and *Two-Spirit*, have a religious significance in their community, positioning themselves as shamans with healing powers: "their gender variance is strongly tied to ritualism and during rituals they move between masculine and feminine gender polarities or combine the two" (Richards, Bouman and Barker 2017, 24).

Through these examples described by Richards, Bouman and Barker, we have demonstrated how different cultures in Western and non-Western contexts understand, interpret

4 The term *passing*, which we further discuss in following chapters, is used to describe someone who is seamlessly perceived as a gender other than the one assigned at birth. In this case, *Waria* do not wish to pass as a cis woman but as a different gender category altogether.

and perform gender outside the gender binary. These identities illustrate the range of gender possibilities and how gender variance has existed throughout time, allowing us to question the perception of the gender binary as a constant. Whilst some of these identities have experienced discrimination and stigma, others are positioned as a part of their community or even divine. Ultimately, there are multiple variations in the way transgender people articulate themselves within and outside the binary frame, and the way gender is constructed and normalized significantly varies according to specific settings, contexts and cultural influences.

1.1.4. Catholicism, Colonialism and Pathologization: Gender Naturalization and Imposition

When societies begin to instrumentalize class differences to generate capital and profit, more fixed social categories begin to emerge. According to Feinberg, during the period of about 4500 BCE to 1200 BCE human labor became more productive and abundance accumulated as wealth – "communal systems were gradually and unconsciously transformed" and, consequently, a significant societal change took place (Feinberg 1992, 12). By changing the community paradigm to a more stratified one that accentuates social differences between individuals, "it was the rise of private property, the male-dominated family and class divisions that led to narrowing what was considered acceptable self-expression" (Feinberg 1992, 12).[5]

Monarchy also drove this change, since it was based on the concentration of wealth and power in a specific group of individuals, materialized in the ideal of a male heir. Thus, while men govern, concentrating wealth, power and the tasks of leadership, women must ensure and nurture the family's well-being, lineage succession and the health of those around them.

At the same time, the rise of the Catholic Church, the largest landlord in Europe by the 11th century, also contributed to the spreading of this type of values and ideals, erasing any diverging communities, since "in order for the land-owning Catholic church to rule, it had to stamp out the old beliefs that persisted from pre-class communal societies, because they challenged private ownership of the land" (Feinberg 1992, 12). Societies that were

5 This reflected, namely, on the subjugation of women, for example in patriarchal Greek slave societies between VIII and VI B.C. in which "the slave owners developed an ideology degrading women in order to justify overturning women's equality in society" (Feinberg 1992, 109. This is seen later in capitalism itself, which "made use of many of the old prejudices, particularly those that suited their own divide-and-conquer policies" thus imposing by the new ruling class "conformity to the system of wage slavery" (Feinberg 1992, 18).

organized in a community-oriented way were either indoctrinated or eradicated, recalling that in these contexts nonconforming gender identities were widely accepted and even celebrated in pagan like festivities.

Foucault adds that Catholicism imposed "meticulous rules of self-examination", namely through the process of confession and guidance, as it attributed more and more importance in penance to insinuations of the flesh such as "thoughts, desires, voluptuous imaginings, delectations, combined movements of the body and the soul" (1976, 19). Sex must not be named "imprudently" and everything and every sin must be told: examine diligently "all the faculties of your soul", all your thoughts, every word you speak, all your actions and your dreams (Foucault 1976, 20). Thus, the church played a leading part in the dissemination of this patriarchal model.

This collective social dictation and reproduction of the category of gender at a vast political scale greatly normalized gender. In this way, gender started being perceived and performed recurrently according to certain guidelines and prescriptions, through a "collective agreement to perform, produce and sustain discrete and polar genders as cultural fictions" (Butler 1988, 522).

These gendered social notions are framed with the intent of regulating the type of behavior one exhibits and "gender appears to the popular imagination as a substantial core which might well be understood as the spiritual or psychological correlate of biological sex" (Butler 1988, 528).

By linking gender to sex, it becomes limited to two axes: masculine and feminine (although, as we discuss in the next chapter, sex isn't binary either). And even though masculinity and femininity present cultural variations, they become stabilized, and people perceive their normative, widespread definitions as normal. Any deviation from this pattern represents an anomaly that must be fixed or even eliminated.

Its categorical normalization in virtue of one's genitals that apparently are divided in two opposite poles becomes evident and, consequently, a clear sex-gender relationship is established.

Besides criticizing the social enforcement of the patriarchal, monogamous family model and the widespread of catholic church related ideals in western society, Feinberg argues

that the way anthropological studies interpret gender in non-Western societies at the beginning of the 20th century also contributes to this gender perception.

Hence, ethno-centric Western conceptions of gender dominated social sciences and any society that does not follow this model is evaluated according to a binary framework. Colonialism contributes to the reinforcement of these social models, since "the seizures of lands and assets of the accused during the witch trials and inquisition helped the ruling classes acquire the capital to expand their domination over Asia, Africa and the Americas" and the European elite "then tried to force their ideology on the peoples they colonized around the world" (Feinberg 1992, 16). Because the indigenous people were viewed as "primitive", colonial missionaries undertook "a project to convert them – both to Christianity and to the rightful heterosexual binary gender system" (McNabb 2018, 35).

Thereby, even outside European and Western contexts there is a dissemination of normative ideals and binary gender notions are naturalized and imposed while any deviations are scrutinized and eventually banished.

Alongside anthropological analysis of social and cultural perceptions of gender, medicine emerges as a field that is particularly focused in debating sex, gender and the relationship between the two. Sexology starts debating more profound theories concerning gender and sex, and the roles, identities, and expressions "that were deemed 'normal' to those who were 'male' or 'female' were invisible and taken-for-granted, but sets of behaviors, desires, thoughts, and actions that were deemed 'different', 'other', or 'variant' began to be highly scrutinized and recorded" (Richards, Bouman and Barker 2017, 54).

Medical studies begin to focus on identities that somehow counter traditional and conventional patterns of gender and sexuality, pathologizing any identity that represents a deviation from the norm. For example, up until the 1990s, homosexuality was considered a mental illness by the World Health Organization (WHO). And trans people were (and in some countries still are) also categorized according to an idea of psychological pathology called "gender dysphoria syndrome" (Stone 1987, 2). These individuals were judged in light of "old moral categories of debauchery and excess" as sex was placed in a position of "biological responsibility" with regards to the species: if these "perversions" were not controlled, they could transmit diseases or even affect future generations (Foucault 1976, 118).

Thus, all forms of "anomalies" are clinically analyzed as they are assigned "a role of normalization or pathologization with respect to all behavior" and a "corrective technology" is sought for these individuals (Foucault 1976, 105). Preciado adds that masculinity and femininity become "pharmaco-pornographic fictions retroactively defined in relationship to the molecule with which they are treated" since "clinical masculinity does not exist without synthetic testosterone" (2008, 60-61). As these identities become pathologized, normative conceptions become clinically supported and reinforced and gender has a set of criteria that must be met and clinically and socially approved.

In other words, when trying to deal with these deviations, in order to physically and psychologically bring these individuals closer to dominant gender ideals and notions, their identities are reframed in a normative way. Consequently, there is a reinforcement of traditional gender patterns, replicating ideals of femininity and masculinity.

It is then established that the most successful scenario for these individuals is their social reintegration in order to 'pass', that is, being perceived as a natural member of the gender of choice. In this sense, the process of transitioning encompasses numerous steps, particularly those of psychiatric ordeal, in which trans people must learn how to reproduce stereotypical accounts of men and women.

It is also considered that the crucial stage of the transitioning process of a trans person is the sex reassignment surgery. This idea that trans people inhabit a 'wrong' body and that their genitals or physical attributes are something undesirable only reinforces a connection between gender and sex, with a sex change being one of the few criteria that validates gender transition.

Following this idea, Sandy Stone poses the following question: "suppose that you could be a man or a woman in every way except for your genitals; would you be content?". There are several answers but "only one is clinically correct" (Stone 1987, 15). There's an intention of mitigating gender identity deviations as much as possible so that their expression is socially diluted, making it part of a dominant paradigm and "under the binary phallocratic founding myth by which western bodies and subjects are authorized, only one body per gendered subject is right [and] all other bodies are wrong" (Stone 1987, 15).

Even if recognizing the possibility of changing one's gender challenges and questions the rigidity of gender, this doesn't take into account the possibility of existing outside of its

binary frame, rarely acknowledging lived experiences that incorporate both masculine and feminine attributes, neither or even alternative, queer approaches to these concepts.

1.1.5. Towards a Queer and Intersectional Approach to Gender[6]

Although cis-binary interpretations of gender are ultimately imposed as correct and normal, they do not reflect, as previously seen, the essential nature of gender. In addition to questioning this type of clinical approach, Foucault also questions biology and psychiatry, considering that these fields set themselves up as supreme authorities that claim to ensure "the physical vigor and the moral cleanliness of the social body" promising to eliminate defective, degenerate and bastardized individuals. In the name of a "biological and historical urgency", these acts are justified by grounding them in "truth", when in reality these orders of knowledge are developed "according to a general scientific normativity" (Foucault 1976, 54).

Thus, in order to counter and question these approaches to gender, queer theory emerged in the 1990s, in the context of feminism, gay and lesbian studies. Accordingly, queerness takes into account that "to generate a true, effective and representational counter discourse is to speak from outside the boundaries of gender, beyond the constructed oppositional nodes which have been predefined as the only positions from which discourse is possible" (Stone 1987, 13).

To deconstruct this discourse, one should avoid being a speaking subject within the traditional gender frame and seek a place outside of it. It is necessary to consider more than roles, stereotypes and binary understandings of gender in order to articulate a discursive approach in which gender is perceived beyond this polarized, normative and binary perspective.

Queer theory proposes the transgression of conventional norms and ideas, questioning and problematizing everything that is imposed in a generalized way. *Queer* defines a calling for a working together to overthrow mainstream thinking and articulate alternative lifestyles – "in terms of gender, queer revisits and revises the categories of 'man' and 'woman' as fixed, essential single identities" and "the open mesh of possibilities, gaps, overlaps, dissonances and resonances, lapses and excesses of meaning when the constitu-

6 Although in the past *queer* was used as a harmful slur, it has since then been reclaimed as is now used as a term to describe identities outside a cis-heteronormative frame.

ent elements of anyone's gender, or anyone's sexuality aren't made (or can't be made) to signify monolithically" (Richards, Bouman and Barker 2017, 62).

This approach contradicts the idea that someone whose gender differs from that which was assigned at birth must be subject to a sequence of treatments and tests so that their identity is socially and legally validated. These individuals showcase identities that disrupt conventional gender notions, mapping the body in new and unexpected ways, allowing us to think outside a binary framework. Thus, nonbinary individuals are carving out a "pathway of possibilities that are currently relatively unexplored, they/we are the avant-garde of gendered existence which is shifting the landscape of gendered possibilities" (Richards, Bouman and Barker 2017, 67).

Expanding on these ideas and reflecting upon his trans condition, Preciado considers that "he is not a man and he is not a woman and he is not heterosexual he is not homosexual he is not bisexual" (2019, 29). Instead, he is "a dissident of the sex-gender system" and "the multiplicity of the cosmos trapped in a binary political and epistemological system" confined inside the limits of "techno-scientific capitalism" (2019, 29). Preciado also argues that the words "male, female, hetero and homo are outdated" and that it's crucial to invent new words that aren't connected with identity politics to define human subjectivity and social relationships and successfully "depatriarchalize and decolonize the classification systems that have constructed sexual, gender and race distinctions within colonial capitalism" (2020b).

This is the importance of intersectionality, since identity and, consequently, gender do not exist in a vacuum, in isolation. Intersectionality considers gender as existing in articulation and overlapping with various identity categories such as ethnicity, class, age, religion, citizenship, disability, sexuality.[7] Because gender is not always "constituted coherently or consistently in different historical contexts, (…) it becomes impossible to separate out gender from the political and cultural intersections in which it is invariably produced and maintained" (Butler 1990, 78).

However, this notion of gender requires changes on a significant structural scale, since it includes aspects of legal, social, cultural, linguistic and clinical nature. While many cultural groups accept diverse gender identities, few countries grant legal recognition, name-

7 Intersectionality is a term used to describe the ways in which different social and political aspects of one's identity overlap and expose them to different forms of oppression, discrimination and marginalization.

ly to nonbinary people.[8] Since what socially prevails is a normative binary perception of gender, it becomes necessary to reconsider the way we view human identity, its relationships, bodies and their respective existence in articulation with their surrounding spaces. According to this view, the path to non-normative gender identities should include the deconstruction of socially imposed conceptions, based on apparently universal and absolute knowledge. As Preciado puts it, "this is the task of our time" (2020b).

Since binary conceptions of gender are socially dominant, establishing a close relationship between gender and sex, understanding and questioning this system becomes important in order to acknowledge intersectional bodies and truly reclaim gender's fluid, liquid and free essence.

8 Charlie McNabb provides some examples, noting how in Australia, a nonbinary person can use the gender marker "X" on their passport, despite this change requiring "a statement from a medical practitioner"; in Bangladesh, there is a "unique third gender category" on passports; Denmark allows adults to apply for "X" on passports "without any medical gatekeeping"; in India, people can use "O" for other on voter registrations and "T" for transgender on their passports; Nepal allows "X" for indeterminate on passports; and in New Zealand, passports offer "X" for unspecified or indeterminate and government organizations have begun "collecting data on gender-diverse residents in order to advance human rights" (2018, 34).

1.2. GENDER AND SEX DISTINCTION IN THE BINARY FRAME

> *When the constructed status of gender is theorized as radically independent of sex, gender itself becomes a free-floating artifice, with the consequence that man and masculine might just as easily signify a female body as a male one, and woman and feminine a male body as easily as a female one. (...) Hence, gender is neither the causal result of sex nor as seemingly fixed as sex.*
>
> Judith Butler 1990, 62

1.2.1. Sex, Sex Category and Gender

Although gender and sex are related, they tend to be normalized within an implicit relationship where one unequivocally and naturally implies the other. In truth, there is a significant difference between these aspects and what defines them. Accordingly, sex is defined as an anatomical facticity, constituting a biological criterion which, according to a simplistic binary model, distinguishes male from female. This encompasses a range of "sexually differentiated traits and processes", including sex determining genes, sex chromosomes, sex hormones, external primary sex characteristics (penis or vagina), gonads (ovaries or testicles) or type of gamete (egg or semen) (Newman 2018).

However, according to Johnson and Repta, sex is also a social construct, since "our knowledge and understanding of sex has changed as we have come to appreciate the great diversity that exists within populations" (2012, 19). The term biological sex often presumes that there are two biologically discrete types of people but we can see how this puts artificial boundaries around groupings that are really more complex.

Similarly to gender, biology also lies outside of two clear cut and rigid binaries of sex differentiation. For example, we know that "up to 1.7% of people are born intersex" or born with diverse sex traits, and we also know that XXX, XXY, XYY or X0 type chromosome groupings exist (UNFE 2017, 1). These so-called "exceptions" exist because the "idea of two binary biological sexes is just that - an idea", that is, a social construct (Newman 2018). Therefore, sex isn't binary either and exists beyond the male-female dichotomy. Additionally, individuals can even "have different sex chromosomes in different parts of their body", for example a man who has some cells with XY sex chromosomes and

some cells with XX chromosomes (Newman 2018). In fact, "new technologies in DNA sequencing and cell biology are revealing that almost everyone is, to varying degrees, a patchwork of genetically distinct cells, some with a sex that might not match that of the rest of their body" (Ainsworth 2015).

West and Zimmerman also allude to a "sex category", which defines the physical embodiment of sex "sustained by the socially required identificatory displays" (West & Zimmerman 1987, 127). These aspects, such as clothing, facial hair, voice pitch or facial structure allow us (supposedly) to identify someone else's sex and categorize them as men/women, through an "if-can" test (Sacks 1972, 332-35 in West & Zimmerman 1987, 127). This means that if someone is perceived as male or female, they are categorized accordingly.

However, as we have previously seen, these attributes do not necessarily reflect someone's gender, let alone their genitals. Even gonads do not fit a simple binary as some XX people have an *ovotestis* – "a gonad with areas of both ovarian and testicular development" (Ainsworth 2015).

The biopolitical "aesthetics of gender, normative codes of visual recognition and immaterial psychological convictions" that govern sex assignment at birth "are not a biological event" and, instead, penises and vaginas "are ideal regulators, biopolitical fictions that find their somatic support in individual subjectivity" (Preciado 2008, 102-103). Some babies are born with genitalia that does not fit a simplistic binary model, but these babies were, "and in many place still are, operated on, having their bodies surgically altered without consent, by a world that couldn't bear to acknowledge that sex has never been a simple binary" (Newman 2018).

According to Joshua Safer, "the idea that a person's sex is determined by their anatomy at birth is not true, and we've known that it's not true for decades" (in Grady 2018). And while we might know that there is a "durable biological underpinning to gender identity", we do not know all of the biological factors at play as gender and gender identity aren't entirely hormonal or genetic (in Grady 2018). Instead, hormones and genes involve complex nonbinary biological processes and, according to the Endocrine society, terms such as "biological sex and biological male or female are imprecise and should be avoided" (Hembree et al. 2017, 3875).

Hence, gender builds on biological sex to give meaning to sex differences, "categorizing individuals with labels such as woman, man, trans, and hijra, among others [and] these categories are socially constructed, as humans both create and assign individuals to them" (Johnson & Repta 2012, 21).

1.2.2. Cis-Binary Genders and the Implicit Hetero-Monogamous Frame

In *The Second Sex*, Simone de Beauvoir suggests that "one is not born, but rather becomes, a woman" as there is no biological, psychological or economic fate that "determines the figure that the human female presents in society" (1949, 273). Woman, as a concept, is not a natural facticity.

This encompasses the idea that gender is not something we are born with and, instead, is socially constructed and internalized during our life through performative acts. To be male or female might supposedly be a matter of sex; but to be a man or a woman is a matter of gender.

And even these concepts are constantly subject to changes. What we internalize as we grow up are normative and narrow conceptions of sex and gender and, for de Beauvoir, to be a woman is "to compel the body to conform to an historical idea of woman" (in Butler 1988, 522). Through this process, gender becomes linked to sex and sex category, and these concepts become intertwined.

In this sense, gender is generally perceived through a "binary frame" through which it "implicitly retains the belief in a mimetic relation of gender to sex whereby gender mirrors sex or is otherwise restricted by it" (Butler 1990, 62). Thus, a binary gender system emerges: men or women, boys or girls. This system imposes differences (that are not natural) between genders and establishes a need to teach people how to "produce behavioral displays of one's essential female or male identity" (West & Zimmerman 1987, 142).[9]

Moreover, masculine and feminine behaviors are also normalized, becoming predetermined aspects in a closed, historically sedimented structure. In other words, this need of

9 When, for example, a group of cis men agrees to collectively perform certain acts, dressing and behaving in a certain manner, there is a normalization of what it means to be masculine. The repetition of these acts associates the male body with these ideals and this gender configuration is thus defined and perpetuated historically. Attributes such as having facial hair, a low-pitch voice or a defined jawline become linked with gender, establishing the way these types of bodies are socially perceived.

institutionalizing "behavior and cognition in accord with preestablished conceptualizations and behavioral patterns" (Deux & Major 1987, 370) frames certain attributes and acts as specifically feminine or masculine, defining one's personal preferences, habits and behaviors. Even before we are born, from the moment our sex is known, we are categorized as boys or girls and our parents might start planning our upbringing accordingly. As we grow up, our education accommodates expectations regarding behaviors and "a variety of institutionalized frameworks through which our natural, normal sexedness can be enacted" (Goffman 1977 in West & Zimmerman 1987, 137).

In truth, this "naturalist ideology" forces children to carry a "patriarchal weapon", and the defenders of "childhood and family conjure up the political image of a child that they construct, a child presumed to be heterosexual, with a standard binary gender" (Preciado 2019, 45). However, this is a child who "is being stripped of any power to resist, any possibility of making free, collective use of their body" and this "childhood" they claim to be "protecting" implies "terror, oppression and death" (Preciado 2019, 45).

Heterosexual, monogamous marriage reinforces gender restrictions in the binary framework since it "requires the reproduction of human beings in certain gendered modes which, in effect, guarantee the eventual reproduction of that kinship system" (Butler 1988, 524). Borrowing from Foucault, Butler argues that "the association of a natural sex with a discrete gender and with an ostensibly natural 'attraction' to the opposing sex/gender is an unnatural conjunction of cultural constructs in the service of reproductive interests" (Butler 1988, 524).

Thus, gender's essentialist nature is reinforced by womanly and manly conducts that are complemented by wifely and husbandly roles, as this union is prescribed by a "centrifugal movement" of heterosexual monogamy that moderates and normalizes sex and sexuality (Foucault 1976, 38). Thereby, normalizing heterosexuality contributes to a system that frames bodies and genders in a splitting of distinct sexes with supposedly natural heterosexual tendencies.

1.2.3. Binary Burdens and Social Regulations

As this duality of gender is established in a pre-discursive domain, its regulation is reflected on a personal and social monitoring of one's gender.

Even though genders "can be neither true nor false, neither real nor apparent (…) one is compelled to live in a world in which genders constitute univocal signifiers, in which gender is stabilized, polarized, rendered discrete and intractable" (Butler 1988, 528).

In this sense, individuals that fail to properly reproduce their 'perceived' gender are regularly punished, through indirect and direct social consequences. In turn, those who reproduce their perceived gender in a fitting fashion are encouraged and even rewarded. This establishes a social climate in which "this reassurance is (…) easily displaced by anxiety" and where cultural marginalization is prompted for those "who fail to perform the illusion of gender essentialism" (Butler 1988, 528).

Hence, gender duality is ensured and sustained through pervasive social regulations, which we will discussed in the following chapters. With clear and defined boundaries, these regulations inhibit any attempt to counter normative conceptions of gender. This monitoring of one's gender is merely social and there is nothing about being a man or a woman that is given, since man and woman, as concepts, are subject to change and influenced by the contexts they are used in.

1.3. "WOMAN" AS A CONCEPT

> *There is nothing about being 'female' that naturally binds women. There is not even such a state as 'being' female, itself a highly complex category constructed in contested sexual scientific discourses and other social practices. Gender, race, or class consciousness is an achievement forced on us by the terrible historical experience of the contradictory social realities of patriarchy, colonialism, and capitalism.*
>
> Donna Haraway 1991, 295

1.3.1. (You Make Me Feel Like) a Natural Woman

By establishing a binary framework, there is an intention of defining what it means to be a man or a woman, as well as presenting these concepts as two preexisting, natural and universal identity categories. This is reflected upon universal truths that hegemonically define these concepts, despite representing narrow and very specific standpoints: white, European, cisgender, nondisabled, etc (Mombaça 2017). Particularly, 'woman' is often used as a general concept to describe every woman's identity when, in truth, this ignores other identity factors. There is nothing natural, biological, or intrinsic that ties women to a single and unequivocal definition.

As we've previously seen when discussing intersectionality, the concept by itself has little to no meaning and fails to accurately represent one's lived experience. Instead, it is the "naturalistic metaphysics of sexual difference" that claims the "biologically and historically unchangeable existence" of two sexes, two genders and two sexualities – male and female, man and woman, heterosexual and homosexual (Preciado 2008, 226). Under this apparently natural, neutral and universal term "woman", a host of vectors of "production and subjectivity are hiding: sex, race, class, sexuality, age, ability, geopolitical or corporal difference, and so on" (Preciado 2008, 107).

Socially, there seems to be a need to establish an official and general notion of identity regarding what women and men are when, in reality, these identities are far from being as fixed as they seem. Butler even considers that this phenomenon reveals some futility in a political context "which seeks radically to transform the social situation of women

without first determining whether the category of woman is socially constructed in such a way that to be a woman is, by definition, to be in an oppressed situation" (1988, 523).

It is also worth noting that juridical power dynamics regulate our social, cultural and political lives and, as such, "subjects regulated by such structures are, by virtue of being subjected to them, formed, defined, and reproduced in accordance with the requirements of those structures" (Butler 1990, 55). Consequently, women as a political subject are faced with certain limitations and challenges promoted by asymmetrical legal rights and inequalities. As Preciado puts it, "in our role as sexual subjects, we're inhabiting biocapitalist amusement parks" and we are "strange biopolitical fictions because we are alive" (2008, 119).

In this way, an idea of uniformity or commonality between women emerges, derived from social regulations and the way people politically perceive gender. The category of woman is reproduced on a political scale and, as a consequence, the politicized concept of woman is significantly redefined in legal and political discussions. In reality, the category of woman intersects with other elements of their identity and social circumstances.

1.3.2. Category of Women: Whom, When and Where

There are many factors that are as relevant to one's identity as gender and that should be taken into account in order to accurately represent women and their lived realities. As previously discussed, separating gender from political, cultural and social intersections is impossible.

Given that gender tends to be perceived through a binary frame, feminine and masculine realms emerge and the idea that what we consider different from masculine is recognizable as feminine (and vice-versa) constitutes this perception of opposites which only makes sense in this polarized frame. However, outside of this framework, "the specificity of the feminine is once again fully decontextualized and separated off analytically and politically from the constitution of class, race, ethnicity, and other axes of power relations that both constitute identity and make the singular notion of identity a misnomer" (Butler 1990, 59).

In fact, there is little consensus of what constitutes the category of woman. For example, what defines a cis white woman's reality will differ from what defines a cis black

woman's reality, and what defines a trans woman reality might differ entirely from both, revealing how women's identity, which is assumed to exist in a hegemonic and universal way, can't be perceived as something linear, fixed and uniform. Consequently, it becomes important to concentrate on how other ideologies shape gender narratives, focusing on the intersection of "race, queer and/or trans identity and how relevant these historical narratives are in contemporary society" (UNESCO 2019, 12). Only then it becomes possible to acknowledge "inextricable and intersecting dynamics of power and oppression, often resulting in experiences of discrimination that cannot easily be attributed to a single identity category" (Richards, Bouman and Barker 2017, 284).

Therefore, a universal and global category of women doesn't exist or at least one that accurately represents women as a whole. Instead, "there is a sedimentation of gender norms that produces the peculiar phenomenon of a natural sex, or a real woman, or any number of prevalent and compelling social fictions" (Butler 1988, 524).

These social images fail to accurately represent women and their lived experiences, and the unity of women also "rests on an epistemology based on the ontological structure of labor" (Haraway 1991, 298). These social fictions produce supposedly natural configurations of binary bodies, imposing how people should act, behave and relate to each other through gender roles, norms and stereotypes.

1.4. GENDER ROLES AND SOCIAL STEREOTYPES

> *There is no denying that there are differences between men and women in many life domains. The question, however, is to what extent these differences reflect the way men and women essentially are, and to what extent they result from how we think men and women differ from each other because of gender stereotypes.*
>
> Naomi Ellemers 2018, 276

1.4.1. Blue for Him, Pink for Her

Social roles assigned to gender derive from the binary framework as attributes of social structuring and constitute "social norms, or rules and standards that dictate different interests, responsibilities, opportunities, limitations, and behaviors for men and women" (Johnson & Repta 2012, 23).

Each gender is expected to act according to these guidelines, since this distinction is perceived to be "natural and rooted in biology, producing in turn profound psychological, behavioral, and social consequences" (West & Zimermman 1987, 128).

Gender stereotypes, as defined by Prentice and Carranza, constitute qualities ascribed "to women and men [that are also] required of women and men" (2002, 269), encompassing this type of attributes or traditional roles that are assigned to each gender. Accordingly, by social prescription, all women should be warm and welcoming and, in turn, all men should be assertive and dominant.

Gender stereotypes may vary from culture to culture, but in western societies "the accepted cultural perspective on gender views women and men as naturally and unequivocally defined categories of being" (Garfinkel in West & Zimermman 1987, 128) and stereotypes are widely generalized and accepted.

Based on a study made by Stoppard and Kalin in 1978, Prentice & Carranza suggest that gender stereotypes "demand the presence of gender-appropriate, desirable traits and the absence of gender-inappropriate, undesirable traits" (2002, 270).[10]

1.4.2. Girls Are Pearls, Boys Are Machines

Some of these stereotypes, as proposed by Sandra Bem in the Bem Sex-Role Inventory (1981 in Prentice & Carranza 2002, 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".

In turn, Ellemers divides gender stereotypes into specific categories, namely stereotypical domain, relevant behavior, anticipated priorities, perceived qualities and neglected needs. According to this approach, men should be competent and practice agency by prioritizing individual task performance, namely in work contexts, eventually neglecting interpersonal connections. In turn, women should be warm and focus on communality by caring for others, namely in family related contexts, eventually neglecting professional achievement.

10 Following this idea, the authors propose four categories of gender stereotypes. Accordingly, "gender-intensified prescriptions are traits high in general social desirability and even higher in desirability for the target gender" (the qualities men and women should have according to their gender); "gender-relaxed prescriptions are traits high in general social desirability but significantly lower in desirability for the target gender" (societal standards to which one is allowed to fall short); "gender-relaxed proscriptions are traits low in general social desirability but significantly higher in desirability for the target gender" (flaws women and men are allowed to have according to their gender); and "gender-intensified proscriptions are low in general desirability and even lower in desirability for the target gender" (the qualities men and women should not have according to their gender) (Prentice & Carranza 2002, 271).

Gender stereotypes	Male	Female
Stereotypical domain	Agency	Communality
Relevant behavior	Individual task performance	Care for others
Anticipated priorities	Work	Family
Perceived qualities	Competence	Warmth
Neglected needs	Interpersonal connection	Professional achievement

Fig. 1 : Gender stereotypes and expectations, as described by Naomi Ellemers (2018, 281).

Finally, Preciado critically and ironically examines the postwar political and medical discourses that further configured men and women into two separate categories. Accordingly, "semiotechnical codes of white heterosexual femininity" refer to "*Little Women*, a mother's courage, the Pill, the hyperloaded cocktail of estrogens and progesterone, the honor of virgins, *Sleeping Beauty*, bulimia, the desire for a child, the shame of deflowering, *The Little Mermaid*, silence in the face of rape, *Cinderella*, the ultimate immorality of abortion, cakes and cookies, knowing how to give a good blowjob, bromazepam, the shame about not having done it yet, *Gone with the Wind*, saying no when you want to say yes, not leaving home, having small hands, Audrey Hepburn's ballet shoes, codeine, taking care of your hair, fashion, saying yes when you want to say no, anorexia, knowing in secret that the one you're really attracted to is your best friend, fear of growing old, the need to be on a diet constantly, the beauty imperative, kleptomania, compassion, cooking, the desperate sensuality of Marilyn Monroe, the manicure, not making any noise when you walk, not making any noise when you eat, not making any noise, the immaculate and carcinogenic cotton of Tampax, the certainty that maternity is a natural bond, not knowing how to cry, not knowing how to fight, not knowing how to kill, not knowing much or knowing a lot but not being able to say it, knowing how to wait, the subdued elegance of Lady Di, Prozac, fear of being a bitch in heat, Valium, the necessity of the G-string, knowing how to restrain yourself, letting yourself be fucked in the ass when it's necessary, being resigned, accurate waxing of the pubes, depression, thirst, little lavender balls that smell good, the smile, the living mummification of the smooth face of youth, love before sex, breast cancer, being a kept woman, being left by your husband for a younger woman" (2008, 120).

In turn, "semiotechnical codes of white heterosexual masculinity" refer to "James Bond, soccer, wearing pants, knowing how to raise your voice, *Platoon*, knowing how to kill, knowing how to smash somebody's face, mass media, stomach ulcers, the precariousness of paternity as a natural bond, overalls, sweat, war (including the television version),

Bruce Willis, Operation Desert Storm, speed, terrorism, sex for sex's sake, getting hard like Ron Jeremy, knowing how to drink, earning money, *Rocky*, Prilosec, the city, bars, hookers, boxing, the garage, the shame of not getting hard like Ron Jeremy, Viagra, prostate cancer, broken noses, philosophy, gastronomy, *Scarface*, having dirty hands, Bruce Lee, paying alimony to your ex-wife, conjugal violence, horror films, porn, gambling, bets, the government, the state, the corporation, cold cuts, hunting and fishing, boots, the tie, the three-day growth of beard, alcohol, coronaries, balding, the Grand Prix, journey to the Moon, getting plastered, hanging yourself, big watches, callused hands, keeping your anus squeezed shut, camaraderie, bursts of laughter, intelligence, encyclopedic knowledge, sexual obsessions, Don Juanism, misogyny, being a skinhead, serial killers, heavy metal, leaving your wife for a younger woman, fear of getting fucked in the ass, not seeing your children after the divorce, the desire to get fucked in the ass" (2008, 121).

1.4.3. Regulatory and Punitive Conventions of Gender Behavior

Gender stereotypes imply a gender belief system that imposes expectations and behavioral patterns, and, in this case, social interaction is influenced solely based on gender. Mark Snyder points out that "these pieces of information are usually the first to be noticed in social interaction and can gain high priority for channeling subsequent information processing and even social interaction" (1977, 4).[11]

As a consequence, different behaviors are expected from women and men and "social labels, beliefs, and attributions may serve as grounds for predictions and generate behaviors designed to validate or invalidate these beliefs" (Snyder 1977, 10). Gender stereotypes are internalized and socially reinforced, namely through social consequences as we grow up, depending on one's behavior and if it is adequately perceived. Preciado adds that the child is "a biopolitical artifact that guarantees the normalization of the adult" as schools become "disciplinary institutions that normalize gender and sexuality", encouraging and rewarding "the conventional enactment of the codes of masculine sovereignty and feminine submission" (2020a).

11 Therefore, people expect and anticipate behavioral confirmation in social interactions and stereotypes create "their own social reality by channeling social interaction in ways that cause the stereotyped individual to behaviorally confirm the perceiver's stereotype" (Snyder 1977, 4). For example, we might expect someone unattractive to act rudely or closed off which will discourage us from interacting with them; in turn, by expecting this type of behavior from others, that person might end up acting closed-off, confirming said expectation.

Acts and behaviors that represent deviations from the norm are rejected as "gender performances in non-theatrical contexts are governed by more clearly punitive and regulatory social conventions" (Butler 1988, 527). As seen in the previous chapters, any form of dissidence is punished and pathologized.

Prentice and Carranza point out two main reasons for the existence of gender stereotypes: on one hand, they are "closely linked to traditional social roles and power inequalities between women and men" (Eagly in Prentice & Carranza 2002, 269); on the other hand, violations of gender stereotypes "are met with various forms of punishment and devaluation" (Prentice & Carranza 2002, 269).

Both genders are subject to extreme pressure to accurately represent those stereotypes that are assigned to them. But the authors also consider that, as women integrate the workplace and start moving outside the private sphere, they're also expected to perform additional, nontraditional roles.[12]

Thereby, gender stereotypes can be slightly distorted or adapted, as long as they meet social criteria and standardization and, sometimes, it is even demanded that someone performs several stereotypes at the same time. Consequently, collective gender performance demands one to manage their gender expression as the occasion demands so that "the outcome is seen and seeable in context as gender-appropriate or, as the case may be, gender-inappropriate, that is, accountable" (West & Zimmerman 1987, 135).

Outside of these guidelines, deviations are subject to more oppressive terms and are met with harsher consequences. Punishment induces anxiety as one is marginalized, excluded from certain social contexts or even murdered for exhibiting undesirable and inadequate gender attributes.[13] Reinforcing this idea, West and Zimmerman point out how some occasions are organized to routinely display and celebrate conventional gender behaviors

12 To perform these roles "they need to demonstrate many of the traits that the prescriptive female stereotype deems less important for them than for others – to be highly intelligent, efficient, rational, to have common sense and so on". However, since they haven't completely given up their feminine roles, "they need to demonstrate traditionally feminine characteristics as well" (Prentice & Carranza, 175).

13 A Danish study that involved an exhaustive analysis of health and legal records from nearly seven million people over the last four decades concluded that trans people "have significantly higher rates of suicide and suicide attempts compared with the rest of the population", more specifically 7.7 times the rate of suicide attempts and 3.5 times the rate of suicide deaths (Christensen 2023). This study is the first in the world to analyze suicide data for this group. Additionally, according to the Trans Murder Monitoring research project, in 2023, 321 trans and gender nonconforming people were reported murdered, with 94% of the victims being trans women or trans feminine people and 80% being black, brown or racialized (TGEU 2023).

and "if an individual identified as a member of one sex category engages in behavior usually associated with the other category, this routinization is challenged" (West & Zimerman 1987, 139).[14]

As a consequence of establishing gender roles, certain contexts and associated activities are perceived as specifically masculine or feminine. This influences the type of work and knowledge each context produces and these asymmetrical gender expectations end up differentiating between mind and body, namely associating cis men with mental knowledge and cis women with corporeal knowledge.

14 For example, in the workplace, one can be promoted and positively received if they exhibit certain gender attributes deemed as adequate (such as being efficient, goal-oriented and insightful) whereas they can also be rejected or even fired if they act in an unconventional way, for example by dressing with clothing commonly associated with another gender. In the US, only in 2020 did the Supreme Court rule that firing employees because of their sexual orientation or gender identity is a form of discrimination that violates federal law (ACLU 2020).

1.5. GENDER DISPARITY, ASYMMETRY AND HIERARCHIZATION

> *The mind not only subjugates the body, but occasionally entertains the fantasy of fleeing its embodiment altogether. The cultural associations of mind with masculinity and body with femininity are well documented within the field of philosophy and feminism. As a result, any uncritical reproduction of the mind/body distinction ought to be rethought for the implicit gender hierarchy that the distinction has conventionally produced, maintained, and rationalized.*
>
> Judith Butler 1990, 56

1.5.1. Phallocentrism and Phallogocentrism

As a consequence of establishing gender roles, certain contexts and activities become specifically masculine or feminine. Thus, a distinction between mind and body, and bodily knowledge and mental knowledge emerges.

Defined as *phallocentrism*, this ideology positions the phallus as the central element in the organization of the social, cultural and political world (here, the phallus refers to cis men, as defined in the binary framework). Following this idea, phallogocentrism relates to this privileging of the masculine in the context of knowledge and construction of meaning. Explored by Jacques Derrida in his 1981 essay *Plato's Pharmacy*, this concept focuses on the masculine point of view, excluding women, femininity and any identity outside this framing.

Following this idea, Alison Adam argues that, historically, men tend to occupy areas of study such as science or technology and certain "technological skills are gendered in that they are seen to be an important part of what it is to be masculine" (Adam 1996, 48). In turn, the private sphere is seen as the traditional place for women's work and household activities such as "looking after bodies and the places where bodies exist, leaving men free for cerebral work" become associated with corporeal knowledge, tied to the female realm (Adam 1996, 48).

Following this idea, we can observe how knowledge is perceived differently, and mental knowledge tends to be valued over corporeal knowledge, ignoring certain circumstances

and failing to represent "interests, feelings, motivations and bodily capacities that make up a human being" (Dreyfus & Dreyfus in Adam 1995, 358).

Derrida proposes a deconstructionist theory in which he problematizes phallocentrism by attempting to redefine history, theoretical discourse and tradition, emphasizing how important it is to "recognize this strong phallogocentric underpinning that conditions just about all of our cultural heritage" (in Dely 2007, 5).

For Adam, "the process of valuing mental propositional knowledge over corporeal skills type knowledge is not neutral" and, as such, "when one type of knowledge is more associated with one social group rather than another, this can result in a form of epistemic discrimination" (Adam 1996, 48).

Therefore, traditional woman's work and knowledge is diminished because "woman's identity has traditionally been associated with the body and nature, just as man's has been located in their transcendence as mind and culture" (Kirby in Adam 1995, 370). Thus, in western society, it is the knowledge of small, yet powerful groups of white middle-class cis men that is valued and women's knowledge tends to be erased.[15]

Accordingly, we can observe how, for centuries, "politicians, theoreticians of the political, theologians" were all cis men and how we would "search in vain for a woman's face" (Dely 2007, 8). And even when cis women start integrating other areas of knowledge in the 20th century, they found themselves "in a situation where they had to think beginning with an ancient and overwhelmingly masculine tradition, built upon a strong phallogocentric underpinning" (Dely 2007, 6). Even if women start moving into these contexts and integrating male predominant contexts, they're still faced with an overwhelmingly male social, cultural and political tradition.

1.5.2. Patriarchy Over and Out

This creates an asymmetrical dialect in which "this association of the body with the female works along magical relations of reciprocity whereby the female sex becomes restricted to its body, and the male body, fully disavowed, becomes, paradoxically, the incorporeal instrument of an ostensibly radical freedom" (Butler 1990, 71).

---

15 For example, "under a traditional view, rationality is associated with the masculine and the life of the mind and irrationality is associated with the feminine and the body" (Adam 1998, 6).

This stems from the fact that, between femininity and masculinity, there is a flagrant imbalance, as patriarchal tradition is built, in part, with an exclusion of the woman, the feminine and otherness.

In turn, Fishman considers that, since "interactional work is related to what constitutes being a woman, with what a woman is, the idea that it is work is obscured [and] the work is not seen as what women do, but a part of what they are" (in West & Zimermman 1987, 138). This phenomenon is amplified as women move out of the private sphere to work, as they end up ensuring their responsibilities in the workplace as well as in their home, and "old hierarchical dominations [transition] to the new networks" (Haraway 1991, 300). Additionally, Adam points out "that even where women are in full-time paid employment, they still expend much more time than their partners in the domestic sphere" (Adam 1996, 48).

Not only do women asymmetrically ensure tasks in the private sphere, but also occupy jobs specifically structured around traditionally notions of femininity. The structure of society and democracy itself seems to have a phallogocentric base. Through a genealogical reading of the canonical model of friendship discussed by philosophers such as Aristotle, Kant or Nietzsche, we can observe how "the friend" is always a man, never a woman and, as discussions surrounding democracy evolve into an idea of brotherhood or confraternity, a preference has "been given to the phallocentric model of brother, native land and nation" (Dely 2007, 7). Thereby, it is also questionable which places women are supposed to occupy in a society where men live democratically like brothers and how might this "andro-centered model [be] suitable and adequate in order to think today about the citizenship of women" (Dely 2007, 7). This glass ceiling also remains as an obstacle to women's careers and "the gender gap widens as women progress, with lower participation at each successive rung of the ladder" (UNESCO 2021).

As we are taught to think and speak in binaries (masculine and feminine, good and evil, black and white, up and down), hierarchies emerge by perceiving one half of the binary as superior and the other as subordinate. In *Psyche: Inventions of the Other*, Derrida discusses how these hierarchies could be undone, arguing that deconstruction "opens up a passageway, it marches ahead and marks a trail" producing rules and other conventions for new performativities (1987, 23)

Deconstructionism questions such orders and rejects the idea of fixed, universal meaning, problematizing the imbalance between men and women and creating space for otherness. It questions an entire structure of traditional thought as well as the way through which we form discourse, aiming to open space for other possibilities that move away from an exclusively phallogocentric approach. Since this deconstructionist work also leans upon the analysis of texts (which tend to be mainly written by cis men) it is important to rethink this whole structure in order to escape from attempts that are exclusively made from a male point of view.

Much like queerness, one "must displace the places, change or transform the landscape and the traditional categories through which we think, or through which in any case men and women have thought for a long time" (Dely 2007, 11). Only then can we make space for "the coming of the other", engaging in a "reorientation of discourse, history and the tradition" (Dely 2007, 14) instead of perpetuating specific contexts, places and jobs according to one's gender.

1.6. WOMEN IN THE INTEGRATED CIRCUIT

> *Service work is positioned as feminized labor (...) not simply because women make up the majority of the workforce, but because the image of the sector is itself feminized; that is, it is associated with qualities traditionally coded as feminine. Indeed, this is not just true of customer service; many contemporary understandings of feminized labor gesture towards trends in the global labor market that can be linked to the dominance of a socially gendered skill set.*
>
> Helen Hester 2016, 47

1.6.1. Labor Structures and Historical Female Jobs

Gender roles and characteristics deemed as specifically feminine or masculine also imply a structural hierarchization of labor. In this way, men are commonly associated with jobs relating to the mind while women are associated with jobs that involve bodily tasks.[16] This means that feminine and masculine behavior is also used to segregate and structure labor accordingly.

The workplace and its relationships change since, according to 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, 8). In fact, gender also organizes social life and social arrangements based on sex categories and the resultant social order, which supposedly reflects natural differences, "is a powerful reinforcer and legitimator of hierarchical arrangements" (West & Zimmerman 1987, 146).

This is tied to women's practices and to what are historically considered women's places, that is, "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 (1991, 307).

16 There are, of course, exceptions to this but also based on gender stereotypes. For instance, jobs involving the use of brute force are typically relegated to men as their bodies are seen as stronger, tougher and muscular.

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.[17] 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 labor" (Hester 2016, 47).

1.6.2. I'll Bake Your Bread and Tend Your House

This asymmetry also affects the private sphere, namely domestic work, since household and childcare are also considered women's tasks. According to Naomi Ellemers, generally women "across different countries and cultures spend more time on household activities than men do, regardless of their employment status" (2018, 277).

Tasks such as washing clothing, cooking or cleaning are generally performed by women, namely mothers, since they are "framed as an extension of naturally occurring feminine (…) predilections, affects, modes of intimacy, personal preferences and so on" (Hester 2016, 49). Household and childcare tasks are considered women's work as a consequence of normative conceptions of appropriate behavior for femininity. As seen in previous chapters, the heterosexual framework also contributes to this asymmetry through its wifely and husbandly roles.

Therefore, this asymmetry is pervasive in several contexts: in work, knowledge and even the private sphere. For example, roles of leadership are commonly performed by men and certain areas of technology tend to alienate women, employing mostly men. And in these contexts, when women do get employed, there is also "the expectation that women's wages will not be matched by a male income" (Haraway 1991, 305).

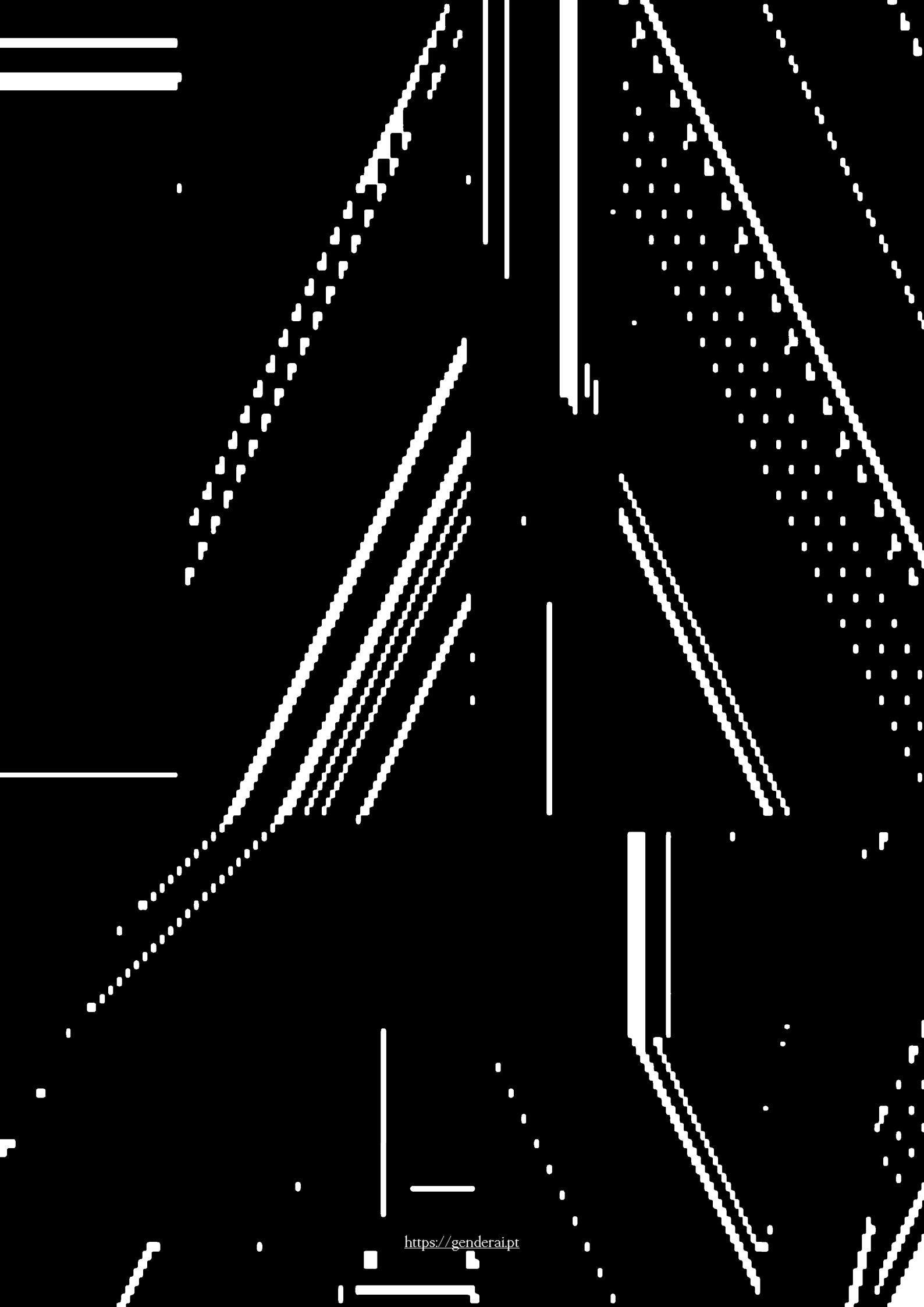
Additionally, and according to Donna Haraway, domestic work is transformed into capitalized labor out of the private sphere, through jobs such as office work, nursing or service work. Borrowing from Richard Gordon, Haraway considers that, with new media, a "homework economy" emerges, defined as a "restructuring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

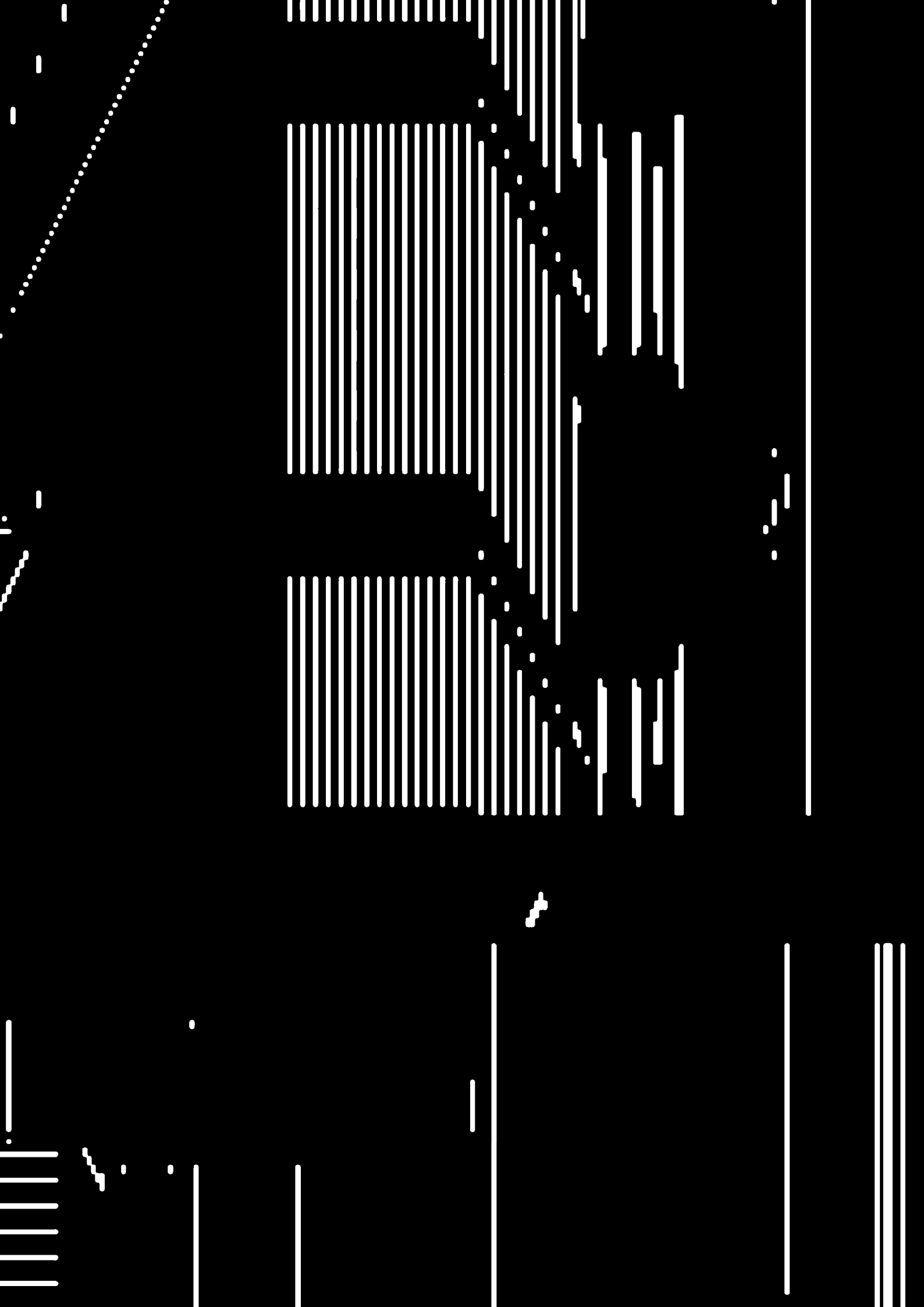
17 To give a concrete example, historically women have a significant presence in the telecommunications industry, where they filled the role of assisting and establishing calls and communications, which rendered "female operators (…) inferior, subordinate, and knowable" (Zost 2015, 3).

as well as making dentists' appointments, picking up dry cleaning, paying personal bills, and so on" (Hester 2016, 49).

Even outside the domestic sphere, women still ensure domestic tasks: "partly as function of their enforced status as mothers" as well as working 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, 305-307).

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


https://genderai.pt



2. THE EVOLUTION OF DIGITAL ASSISTANTS: AN EXPLORATION OF ARTIFICIAL INTELLIGENCE

2.1. ARTIFICIAL ANCESTORS, MIND AUTOMATION AND IMITATION GAMES

> *Over the past decade, artificial intelligence has become the most hyped technology since the World Wide Web. (...) Progress has been so rapid that some informed observers even claim that machines with the full range of human capabilities may soon be here. Amid all the excitement, it is easy to forget that AI is not a new field. Indeed, the idea of AI is an ancient one: it appears throughout recorded history in one form or another.*
>
> Michael Wooldridge 2023

2.1.1. Mythical Simulations and Humanoid Emulations

Within the field of artificial intelligence, gender emerges through the assignment of human characteristics to the entities that this technology assumes as well as in the data it collects and processes. To discuss this phenomenon, we will first look into the origins of artificial intelligence and its evolution and stabilization up until now.

The idea of creating artificial entities that somehow mirror human beings, physically and mentally, is neither new nor recent. Michael Wooldridge gives us some examples, starting

with Classical Greece with the story of Hephaestus, a blacksmith to the gods who "had the power to bring metal creatures to life". In Prague, in the 1600s, a local legend suggested that a rabbi created the Golem, "a magical being fashioned from clay, intended to protect the city's Jewish population from anti-Semitic attacks". And finally, in the 19th century, Mary Shelley, cooped up "in a Swiss villa during the spell of bad weather, created the story of the Frankenstein" (Wooldridge 2020, 24).

The previous examples refer to mythic and fictional scenarios but all suggest that artificial beings that recreate human activity have been a recurring theme in our imagined realities for a long time (which we will further discuss in a chapter dedicated to fictional depictions of AI). But these ideas of replicating physical and mental human activity came to practice with machine automation.

In the 18th century, during the industrial revolution, machine automation emerged with machines that replaced physical and bodily processes that, until then, were exclusively performed by human beings.[18] Automation then aimed for complete autonomy, meaning that machines required minimal human intervention to work. Operating mostly in mechanical and electrical contexts, these machines functioned as assistants that replaced the physical power of human labor.

This technology was, in a way, a precursor of the ideas that later inspired artificial intelligence as the evolution of this automation and the appearance of computational technologies anticipated "the replacement of the human mind itself" (Rifkin 1995, 5).

In the 20th century, the third industrial revolution brought forth the rise of electronics, telecommunications and computers, and machines began to focus on creating, manipulating and communicating information. Technology shifts its focus into creating devices that could perform and automate logical processes. However, according to Arlindo Oliveira, every system that came out of the third industrial revolution lacks intelligence, prompting discussions surrounding the meaning of human intelligence (2019, 34-35).

18 Shifting the focus from skilled manpower to special-purpose machinery, these systems functioned in contexts ranging "from extraction of raw materials to the marketing and distribution of final goods and services" (Rifkin 1995, 60).

2.1.2. From Computationalism to Artificial Intelligence

We have previously discussed how human thought is gendered, attributing mental and rational knowledge to the masculine realm and bodily and intuitive knowledge to the feminine realm. Now, we aim to understand the way human thought and intelligence are perceived, particularly, how knowledge is analyzed, produced and validated in scientific contexts.

According to Joseph Weizenbaum, over time, science has delegitimatized all other ways of producing knowledge as it evolves and becomes "the sole legitimate form of understanding in the common wisdom" (1976, 16).[19] Thus, people tend to look for what is presented to them only as scientifically validated knowledge.

In this way, science utilizes its own methods to explain human thought and internal mental states are described through symbols and symbol manipulation. As logic becomes the most established and accepted way of handling symbols in a formal way, mental activity "came to be seen as a type of computational activity" (Adam 1995, 357-258). Thus, human intelligence and knowledge are structured and processed according to a logical set of rules and the human mind is perceived as something that obeys a sort of script.

Gualtiero Piccini describes this phenomenon as "computationalism", which relates to the conviction that "functional relations between mental states and inputs, outputs, and other internal states were computational" (2004, 817). In other words, this means that intelligent behavior ought to be explained by computations performed by one's cognitive system (the brain). These computational theories of cognition explain behavior in terms of functionally defined processes or mechanisms.[20]

Authors such as Weizenbaum considered that not every aspect of human thought can be reduced to logical formalisms and that "there are some acts of thought that ought to be attempted only by humans" (1976, 13). Reinforcing this idea, Hofstadter states that "no program in the world understands even one concept at the same level of complexity as an ordinary person does" (Hofstadter 1995, 160).

---

19 For example, the arts, in particular literature, used to be viewed as "sources of intellectual nourishment and understanding, but today the arts are perceived largely as entertainments" (Weizenbaum 1976, 16).

20 Adding to this, philosophical materialism considers that the mind is a by-product of material processes without which they cannot exist. Though materialism has been successful in describing these physical processes, "it cannot explain phenomena such as cognition, life, meaning and agency, often falling into a mind/body dualism" (Gershenson 2014, 241).

Still, computationalism relates to artificial intelligence's first attempts: if it is possible to transform human thought (and intelligence, to this effect) in computational algorithms, it may also be possible to integrate these algorithms in machines.[21]

Consequently, the human mind gradually becomes perceived as something that ought to be reduced to purely logical processes, with the intent of being integrated into machines. Computational intelligence, as Jack Halberstam defines it, is a "learned, imitative behavior that can be processed so well that it comes to look natural", much like gender (Halberstam 1991, 443).

Although machine intelligence draws inspiration from the human mind's behavior and phenomena, it ignores certain factors such as emotions, cultural aspects or social awareness. In truth, social interactions are equally important for the development of intelligence, since human intelligence is established "first on the social level, and later, on the individual level" (Vygotsky in Lindblomm 2001, 5). Lacking what Jessica Lindblomm defines as "situatedness", artificial intelligence focuses on "traditional cognitive science", ignoring social contexts and pieces of information related to sociocultural domains (2001, 3).

Nonetheless, we can observe how developments in artificial intelligence are rooted in science and how current iterations of this technology reveal a tendency to simulate the human mind through solely logical and scientific terms. Particularly, AI has its roots in mathematics and in the intention of framing human thought – namely logical and rational – as a mathematical process of manipulating symbols.

2.1.3. Turing Tests, Automated Machines and Natural Language

Accordingly, Oliveira notes how the first practical approach to the current concept of artificial intelligence dates back to 1651, when Thomas Hobbes proposed that the human mind functions through the manipulation of symbols and when one reasons one does nothing more than conceive a total sum, by the addition of portions by the subtraction of one sum from another (2019, 37).[22]

21 This notion is even reinforced by how this type of technology is seen as intelligent: if artificial intelligence systems demonstrate the ability to follow rules and act according to logical formalisms, they are considered efficient and successful.

22 Arlindo Oliveira argues that, throughout history, there have been several attempts of recreating mathematical thought processes in machines. For example, the French mathematician-philosopher Blaise Pascal who, in the seventeenth-century, created a machine that could add and subtract numbers through a gearwheel. In the same

Years later, Alan Turing proposes what we now call Turing machines, demonstrating that any computer that manipulates symbols and has a big enough memory is able to perform the same calculations and achieve the same results as any other computer. This marks the starting points for computers as we know them today.

After the Second World War, the term "electronic brain" emerges as the first electronic computers started being built, capable of "processing huge volumes of convoluted arithmetic problems much faster and more accurately than any human could ever dream of doing" (Wooldridge 2020, 33). As machines start operating with complex and difficult mathematical processes, executing them in short amounts of time and proving to be extremely useful, this raised the question of "whether a computer could behave in an intelligent way, displaying behaviors similar to those of human beings" (Oliveira 2019, 34-35).

Aiming to discuss this phenomenon, Turing takes inspiration from Victorian era parlor games, proposing the Imitation Game. The game "is played with three people, a man (A), a woman (B), and an interrogator (C) who may be of either sex. The interrogator stays in a room apart from the other two. The object of the game for the interrogator is to determine which of the other two is the man and which is the woman" (Turing 1950, 433).

What we now know as the Turing's test replaces one of these people with a computer, demanding that a human subject decide whether they are communicating with a human or a machine. Turing argued that, to build intelligent machines, their processes should be able to be reduced to simple, clear and explicit instructions and that not only could a computer be intelligent as it could also generate, if properly programmed, every outcome that results from human brain activity: intelligence, emotions, creativity and consciousness (Oliveira 2019, 42).

However, the test has several objections, and one of the main criticisms is that the test isn't, in reality, a test of intelligence since no matter how intelligent a computer might be, it wouldn't pass the test unless its behavior and interaction resembled a human being (Oliveira 2019, 49).[23] Moreover, even if a machine passes the test and is considered intelligent, this doesn't necessarily mean that the computer is conscious and sentient, with

century, Gottfried Leibniz, a German mathematician and scientist, designed two prototypes of machines that were capable of calculating arithmetical operations. And later in the nineteenth-century, Charles Babbage, an English mathematician and mechanical engineer, with his difference engine, a calculator powered with cog-wheels and other mechanisms capable of tabulating several polynomial functions automatically.

23 Additionally, the test doesn't address every ability through which human intelligence can be expressed. It also doesn't state explicitly that, if a machine passes the test, it will be necessarily conscious. Turing himself did not

consciousness referring to the awareness of oneself and surroundings, and sentience to the ability of being aware of one's feelings and sensations.

Despite these criticisms, the Turing test is still relevant and an important mark in AI history because it functions as a philosophical instrument that allows us to scrutinize arguments related to the possibility of artificial intelligence (Oliveira 2019, 50) and because, for the first time, "it gave researchers interested in this emerging discipline a target to aim at" (Wooldridge 2020, 54).

The Turing test also gives us insight into how human-machine interaction can easily be manipulated if machines successfully display human behavior, eventually tricking users into thinking they are speaking to another human being.

2.1.4. Before Siri and Alexa, There Was ELIZA

In 1966, Joseph Weizenbaum developed ELIZA, a computer program that led some of its users into thinking it was a person instead of a machine. ELIZA was capable of analyzing written inputs from users and answering accordingly by using a set of rules, thus establishing a human dialogue.

To do so, the program searched the inputs for the presence of keywords and produced responses "by transforming sentences according to a rule associated with said keyword" (Weizenbaum 1966, 37) and also by replacing certain words or expressions.[24]

Weizenbaum chose a script with which ELIZA acted as a Rogerian psychotherapist, since it 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, 3). 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

want to give the impression that "there is no mystery about consciousness", arguing that there was something of "a paradox connected to any attempt to localize it" (Turing 1950, 445).

24 As an example, if a user said something along the lines of "I am upset because of my boyfriend", ELIZA would answer with "Why do you think you are upset because of your boyfriend?". In this sense, ELIZA wasn't restricted to a particular set of responses, although it 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, 37).

participating pair was free to assume a pose of knowing almost nothing of the real world" (Weizenbaum 1966, 42).

Involuntarily, ELIZA ended up demonstrating the way in which the Turing test presents human intelligence in a very restrictive and reductive way – ELIZA was considered intelligent simply because it was able to follow a logical script and appear to be human. As ELIZA grew close to its users, Weizenbaum "became rather less enthusiastic about the program" as he observed the "threats" posed by anything capable of "faking its humanity" (Plant 1997, 91), echoing many of the issues that arise with current digital assistants.

For once, it was one of the earliest natural language process applications and, by reflecting users' interactions back to them, it spoke in familiar and comforting terms. Its interaction shifted from purely robotic to a more social one. These aspects are particularly relevant, as current digital assistants make use of the same strategies and their dialogues are often focused on our private lives, habits and preferences.

ELIZA was also able to trick some of its users into thinking it "was a person rather than a machine" and some of them felt like the system could truly understand and empathize with their feelings (Dale 2016, 814). This also relates to the way current digital assistants aren't mere assistants anymore; they are moving closer to us as daily companions that supposedly understand and relate to our emotions.

Finally, by simulating humanized conversations that focused on the user, ELIZA led some to believe it was capable of human intelligence, thought and consciousness. Although ELIZA simply copied what were already "rudimentary human skills", it awoke "new fears" in those it was supposed to serve (Plant 1997, 91), and this phenomenon would later be discussed according to a philosophical theory that categorized ELIZA as a type of weak AI.

2.1.5. Strong and Weak AI: From Learning Machines to Neural Networks

Weak AI describes computers as tools that simulate human intelligence without claiming about literally having a mind. It aims "to study the mind by doing computer simulations as opposed to purporting to create a mind" (Searle 2004, 65-66). In turn, Strong AI en-

compasses how "the appropriately programmed digital computer does not just simulate having a mind; it literally has a mind" (Searle 2004, 66).[25]

While general, strong AI might have been the original goal of the field, over time it proved hard to achieve and the focus turned towards developing programs that perform specific tasks that require human mental skills, that is, narrow, weak AI.

Accordingly, we can observe how ELIZA and current digital assistants can be considered the latter type of AI. They perform specific and narrow tasks and behave according to scripted and predetermined interactions. Thus, simple programs can easily simulate human behavior, manipulating users and leading them into believing they are capable of human mental skills without possessing an actual mind. Instead of aiming at building a complete general intelligent system, current standard methodology aims at identifying various individual capabilities.

Therefore, capabilities such as problem solving and planning or natural language understanding are being replicated into AI systems, namely digital assistants.[26] Additionally, current AI technology is particularly concerned with developing systems that are able to learn about their users and adapt to their behavior and preferences.

Consequently, machine learning is one of the main fields being explored in AI. The key component of this approach is the possibility of a system that, when properly configured, in the right context, with the right conditions, can adapt its behavior and generate results

---

25 Expanding on these ideas, and borrowing from Ray Kurzweil, Mitchell discusses how nowadays the terms strong AI and weak AI "have also been used to mean something more like general AI and narrow AI", although this differs from Searle's original meaning (2019, 70). Wooldridge adds that general AI refers to programs that "really do have understanding (consciousness, etc) in the way that people do" (2020, 57). This means the kind of technology that equals or even surpasses human intelligence, exemplified by "the AI we see in movies, that can do almost everything we humans can do, and possibly much more" (Mitchell 2019, 65). In turn, narrow AI describes programs that "demonstrate the same capability but without any claim that they actually possess these attributes" (Wooldridge 2020, 57). Also known as narrow AI, this refers to systems that perform a narrow set of tasks and focus on specific, well-defined functions.

26 We can also distinguish between symbolic and subsymbolic AI. Advocates of symbolic AI argue that "general intelligence can be captured entirely by the right kind of symbol processing program" (Mitchell 2019, 41). This approach uses symbols that stand for things that the system is reasoning about and was the most popular approach from the mid-1950s until the late 1980s. This possibility involves "processes of conscious reasoning, problem solving and so on, which we all make use of as we go about our lives" (Wooldridge 2020, 62). On the other hand, subsymbolic AI aims at recreating the brain, simulating neural processes of information. It takes inspiration "from some structures that occur in the brain and model these components in intelligent systems", namely resulting in neural networks or neural nets (Wooldridge 2020, 64). Therefore, this type of AI aims at replicating specific rational processes, even those that are unconscious, such as "recognizing faces or identifying spoken words" (Mitchell 2019, 43).

according to the inputs given to them, instead of limiting itself to following a predefined set of rules like ELIZA.

Neural networks, also known as artificial neural networks or simulated networks are a subset of machine learning, "inspired by the human brain, mimicking the way that biological neurons signal to one another" (IBM 2024). In these networks, each node or artificial neuron "connects to another and has an associated weight and threshold [and] if the output of any individual node is above the specified threshold value, that node is activated, sending data to the next layer of the network" (IBM 2024).[27]This type of systems is difficult to achieve because it relies on training data to improve their accuracy over time. As we train these systems, certain patterns emerge, requiring us "to be very careful about the training data that we use [as] we can usually only train the program with a tiny fraction of the possible inputs and outputs" (Wooldridge 2020, 233).

Current machines that use AI technology are increasingly adhering to machine learning, to learn about their users, collect data about them and adapt to their behavior. Unfortunately, some authors argue that "machine learning and experimentation can expose users to risks" since these systems can "reflect and reinforce any biases that are present in the training data", thus threatening user's privacy and replicating increased power asymmetries (Bird et al. 2016, 1-2). Despite ongoing debates about ethics regulations concerning data collection, "researchers have largely ignored the unique risks posed by autonomous experimentation" (Bird et al. 2016, 4).

In the context of this research, we are particularly concerned with current digital assistants, as they are becoming increasingly ubiquitous and collect user data in order to learn about their users and adapt to their preferences, which we will discuss in the later chapters. Additionally, they record and manage personal information as their tasks focus on our personal lives. Similarly to ELIZA, they imitate human behavior, displaying friendly and caring attitudes, as if they were capable of empathizing and understanding our emotions. Thus, as AI becomes closer to us and attempts to adapt its behavior, learn about its users and tailor its functions to them, it currently integrates numerous aspects of our daily life, namely in critical fields.

---

27 This idea echoes Turing's proposal that, instead of trying to "produce a programme to simulate the adult mind", it would be easier to "produce one which simulates the child's" (1950, 454). This learning machine could then "incorporate a memory" and be subjected to an appropriate course of education, thus achieving an adult brain capable of displaying intelligent behavior (Turing 1996, 258).

2.2. ROBOTS, BOTS AND AI: FROM INDUSTRIAL WORKERS TO VIRTUAL COMPANIONS

> *[Electronic assistants] give us the illusion that we too have personal assistants, better known as servants, and that we can outsource our labor to these helpers. (...) The promise of technology of course was that repetitive labor could be automated and new relations to work and liberation might emerge. But in the era of vertiginous capital, (...) electronic assistants offer an illusion of automation while leaving labor relations intact.*
>
> Jack Halberstam 2018

2.2.1. Machines as Workers, Social Robots and Companions

Artificial intelligence currently encompasses various areas of study, but all of them announce its integration into our daily lives such as education, science healthcare, entertainment, the media and arts. AI systems will be "embedded throughout our world, in the same way that computers are today [and] in the same way that computers and the World Wide Web changed our world, so too will AI" (Wooldridge 2020, 288).

This technology integrates our daily life with the aim of assisting us in different contexts but the idea of owning assisting machines that mimic human beings dates to the 1920s. According to Kathleen Richardson, the first robots emerged as characters in a play written by Karel Čapek where he "took the idea of the factory worker one step further by inventing the robot" (2015, 2). Describing them as "objects of modernity created to reflect on what it means to be human", Richardson argues that "robots were a critical response to what Čapek believed was an obsession with labor and production by right and left political philosophies" (2015, 19), characterized as human-like entities designed with the sole purpose of labor and work. In fact, the term *robot* is Czech for "compulsory service", akin to Slav *robota* "meaning servitude, hardship" (Richardson 2015, 2).

The public viewed and interpreted Čapek's robots as the rise of machines that would free humans from work and by the end of the 1930s, "humanoid robots had begun to make appearances at world's fairs and industrial expositions" (Richardson 2015, 29). This illustrated the beginning of what Donna Haraway describes as the several transformations

and "transitions from the comfortable old hierarchical dominations to the scary new networks" namely from work to robotics (1991, 300-301).

As discussed in the previous chapter, the 1940s onwards saw the development of computational machines, namely with artificial intelligence, "focused on simulating human intelligence in machines (…) and the body of a machine was irrelevant to these functions" (Richardson 2015, 11).

Following this idea, we can observe how the presence or absence of the body was an important feature in AI development, with two different approaches: embodied AI when referring to robotic machines focused on physical labor and disembodied AI associated with intellectual and rational domains. Again, mind-body dichotomy shapes the way we perceive labor and human intelligence and, consequently, imposes limitations on the conception of these systems. Thus, as AI is brought to the fore, the physical body is rendered as "unimportant or secondary to the development of intelligent systems" (Richardson 2015, 47).

These approaches surrounding mind and body started echoing in debates that took place in the 1980s, in humanities and social sciences. Accordingly, in the process of simulating human intelligence, machines were developed with the intent of simulating human behavior, their interaction shifting into a more personal and social one, encompassing the digitalization of human socialization as previously discussed with the Turing test, natural language processing or ELIZA.

As disembodied machines become closer to their users, mimicking their behavior and assisting them in private matters, the boundary between human and machine is reworked and "if the robot acted in a social way and could entice people in an interaction, this was a testament to its success as a social machine" (Richardson 2015, 11).

Richardson then points out how "in labs at MIT and beyond, robots are reimagined by researchers as child-like, companions, friends and a different kind of species altogether" (Richardson 2015, 61) and, as these systems are ascribed childlike socializing techniques of nurturing and play, they gradually become a natural component of the technologies and social circles we navigate. Whether through conversational interfaces, assistants or just chatbots, the basic purpose of these entities is to achieve "some result by conversing (…) in a dialogic fashion, using natural language" (Dale 2016, 811).

And as these systems start being imagined to "fill gaps in human social relationships as companions, lovers, children and therapeutic agents" (Richardson 2015, 65), we can observe how they echo current AI entities. Particularly, those that interact with us daily, articulating emotionality and sociality with kind and reassuring behaviors and assisting us with personal tasks.

2.2.2. General Personal Assistants and Specialized Digital Assistants

A chatbot or digital, virtual assistant is "an artificial intelligence program and a human computer interaction (HCI) model (…), a computer program designed to simulate conversation with human users, especially over the internet" (Adamopoulou & Moussiades 2020, 1). These systems are now "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, 815).

Developed with the goal of assisting users in their day-to-day tasks, they can help with some "subset of the standard virtual assistant skill portfolio, which generally includes scheduling meetings, checking your calendar and making appointments, reading, writing and sending emails, playing music, and, increasingly, controlling your suitably automation-enabled home" (Dale 2016, 812).

William Meisel distinguishes two groups of chatbots: "general personal assistants", which refer to more developed and complex assistants like Siri, Alexa or Cortana, and "specialized digital assistants", which refer to a "tsunami of more narrowly focused chatbots" (Dale 2016, 812).

AI systems of the first type can help with reading, writing, sending emails, scheduling meetings, checking calendars and setting appointments, making calls, sending messages, taking notes or setting reminders. 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. Usually, general digital assistants are integrated directly into our devices, like Siri in Apple devices, and assist us in a more personalized way.

Specialized digital assistants "operate in very specific domains or help with very specific tasks", usually in web-based platforms or apps that serve specific areas and their tasks

can range 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, 812-813).

By performing these 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, eliminating their evil twins friction, opacity, ambiguity and imperfection" (Morozov 2013b, Int. par. 14).[28]

Recent chatbot development reveals a growing tendency towards assigning them human-like attributes regarding their behavior and presentation, eventually to ease anxieties surrounding digital assistants. As these entities become more ubiquitous and present in our daily lives, concerns relating to privacy and data collection emerge, as well as an "uncanny valley" effect, that is, "the uncomfortable feelings that a person is experiencing when [they don't] know if the interlocutor is a human or a computer program" (Adamopoulou & Moussiades 2020, 2).

Ultimately, developers want users to trust their creations. The development of this relationship of trust is supported by the level to which digital assistants resemble human-beings through their voices, names or avatars, their personalities and the way they behave towards users, efficiently handling human language.

By looking at the past of digital assistants, we now aim to understand how their development evolved and how human-like attributes started becoming more prominent in their characterization.

2.2.3. From ELIZA to ChatGPT

Aiming to map chatbot evolution and development, Eleni Adamopoulou and Lefteris Moussiades argue that the concept of a computer program capable of talking to a group of people without them realizing that their interlocutor was artificial can be considered as

28 This need to ameliorate our life and maximize production conveys Evgeny Morozov's concept of solutionism, defined as "an intellectual pathology that recognizes problems as problems based on just one criterion: whether they are 'solvable' with a nice and clean technological solution at our disposal" (Morozov 2013a). According to Morozov, nowadays there is a constant need to attempt to root out any "imperfection, ambiguity, opacity, disorder and opportunity to err, sin or do the wrong thing" (Morozov 2013b, Int. par. 15), which is closely linked to what chatbots aim to do.

"the generative idea of chatbots" (2020, 2). The terminology of "chatterbot" also "chatbot" originates from a game character for a multiuser dungeon game that answered questions and "simulated conversational abilities via simple rules, (…) [fooling] other players into thinking it was another user" (Perez-Marin & Pascual-Nieto 2011, 2).

As previously discussed, ELIZA was developed with this goal in mind and although it could only hold short conversations within a particular domain of topics, it was able to trick some of its users into thinking they were speaking to another human being, marking a turning point and serving as inspiration for the subsequent development of other chatbots.

Introduced in 1972 by Kenneth Mark Colby, PARRY behaved as a schizophrenic patient and attempted to provoke controversies, prompting more elaborate answers from its users. PARRY is considered more advanced than ELIZA as it defined its "responses based on a system of assumptions and 'emotional responses activated by the change of weights in the user's utterances" (Adamopoulou & Moussiades 2020, 2).

Racter was a chatbot program written by William Chamberlain and Thomas Etter in 1983 which randomly generated English prose. An important faculty of the program was its ability to "direct the computer to maintain certain randomly chosen variables, which will then appear and reappear as a given block of prose is generated" (Zemčík 2019, 16). This eventually resulted in the publishing of a book written entirely by Racter, with apparent "deep internally touching meaning", prompting many to anthropomorphize the program (Zemčík 2019, 16).

Jabberwacky (1981) and Dr. Sbaitso (1992) marked significant points as contextual pattern matching to respond based on previous discussions, accompanying their dialogue with digitized voices. Jabberwacky was created by British programmer Rollo Carpenter and, by learning from previous conversation and using contextual pattern matching techniques to select responses, it simulated "natural human chat in an interesting, entertaining and humorous manner" (Jabberwacky 2011).

By using the "technologically ground-breaking novelty Sound Blaster sound card created by Creative Labs", Dr. Sbaitso was another milestone in chatbot development as it was able to synthesize everything written through a digitized voice (Zemčík 2019, 16).

Another step forward was the creation of ALICE (1995), an online chatbot inspired by ELIZA. Contrary to ELIZA, the ALICE bot was an open-source product with a database of over 40,000 pattern templates pairs" which allowed it to have long and diversified discussions on the web, although it still couldn't express human-like emotions or feelings (Perez-Marin & Pascual-Nieto 2011, 6).

Then, in 2001, SmarterChild was launched by Active Buddy, marking the "first attempt to create a chatbot that was able not only to provide entertainment but also to provide the user with more useful information such as stock information, sports scores, movie quotes and much more" with more than 30 million users (Khan & Das 2018, 3).

The development of chatbots went even further with the creation of smart personal assistants, who were capable of understanding voice commands, spoke through digital voices and handled various tasks at the same time. Apple's Siri has been around the longest, released as an app in 2010 and bundled into iOS, on the iPhone 4S, in 2011. Microsoft followed with Cortana in 2013, launched as part of operating systems for Windows Phone and Windows as well as an app, though the latter was globally shut down in 2021. Amazon launched Alexa alongside its Echo-connected home speaker in 2014. And Google Assistant was announced in 2016, primarily embedded in the Google app for Android smartphones and home automation devices. According to Hoy, Amazon is the "dominant player in the field, due to launching a home product first" and while Google and Apple are building capacity and investing in connected home products, Microsoft is not likely to "gain much traction, as their share of the smartphone market is negligible and they lack a compelling home-based product" (2018, 85).

Parallel to the rise of digital assistants that accompanied the beginning of the smartphone era, Watson was developed in 2011 by IBM, representing "a first step into cognitive systems", combining natural language processing, hypothesis generation and evaluation and dynamic (High 2012, 1). Watson was initially developed to answer questions on the quiz show Jeopardy, eventually winning the first-place prize in 2011. And in 2013, Watson Health was announced as its first commercial application, designed to help doctors in lung cancer treatment.

One of the most recent and significant advancements in this field is ChatGPT, a language model developed by OpenAI. Launched in 2022, the chatbot is able to generate human-like text and respond to complex questions and, as it continues to interact with

users, "it can learn about their language, tone, and style, allowing it to generate more personalized and accurate responses" allowing it to adapt to users' specific needs and preferences (Aljanabi 2023, 1). Additionally, "the continued advancement of language model performance through better training algorithms and larger datasets will also play a critical role in the future of ChatGPT [and as] more data is fed into these models, their accuracy and ability to understand and respond to complex questions will continue to improve" (Aljanabi 2023, 1).

We can thus observe how chatbots have always been developed with humanization in mind and, with recent developments, "a growing increase in the use of chatbots was observed, especially in 2016" as the way chatbots engage with us nowadays is entirely different from ELIZA: "they can share personal thoughts and family drama events, be relevant but also confusing, and deceive just as humans do" (Adamopoulou & Moussiades 2020, 3).

2.2.4. A Multitude of Digital Assistants

Smart personal voice assistants are the type of assistants we will focus on in the context of this research and they can be divided in three main categories (although they can sometimes overlap): voice assistants, chatbots and virtual agents.

Voice assistants can normally understand spoken and written inputs. They are designed for spoken interaction and their outputs typically try to mimic natural human speech. Unlike other digital assistants, they are usually always on, ready to interact in response to a 'wake word' (for example, 'OK, Google', 'Alexa' or 'Hey, Siri').

Chatterbots or chatbots interact with users primarily through written language, by analyzing user inputs and generating appropriate responses. They simulate human conversation typically through text-based interfaces, such as messaging apps or websites.

Finally, virtual agents communicate through speech and project a virtual physical form, often a human or sometimes a non-human projection, usually on a digital screen or in a virtual or augmented reality environment.

Current digital assistants branch into different categories based on their functions, the way they interact with their users and the type of human-like attributes they display. Ac-

cording to Adamopoulou and Moussiades, they can be further categorized according to the knowledge domain they operate in, how intimate they are with their users, their primary goals concerning assistance, the response generation method used, whether they are open-source or not and the type of communication channels used during dialogue (2020).

Therefore, regarding their knowledge domain, they can be generic (answer any user question from whichever domain), cross/open-domain (operating in more than one domain) and domain-specific (answering questions that exclusively relate to a specific domain).[29]

Intrapersonal chatbots are close companions that understand users' needs while interpersonal chatbots act without being friendly.

Regarding their primary goal in assisting users, chatbots can be informative, chat-based and task-based. Although each assistant has unique features, they share similarities and are able to perform common basic tasks such as send and read text messages or emails, make phone calls, answer basic informational queries, tell jokes and stories, set timers, alarms, calendar entries or reminders, make lists, do basic math calculations, control media playback from connected services such as Amazon, Google play, iTunes, Pandora, Netflix and Spotify, and control "Internet-of-Things-enabled" devices such as thermostats, lights, alarms and locks (Hoy 2018, 83). Additionally, voice assistants can incorporate other features, often developed by third-party developers, that expand their functions with other programs.

Their responses can define them as rule-based, retrieval based, generative based and human-mediated chatbots. Rule-based chatbots do not create new answers "as the knowledge used is written by the developer in the form of conversational patterns", while retrieval-based chatbots "use a neural network to assign scores and select the most likely response from a set of responses" (Adamopoulou & Moussiades 2020, 5-7). In turn, generative chatbots "synthesize the reply, usually using deep learning techniques" and finally human-mediated chatbots "utilize human computation in at least one part of it" (Adamopoulou & Moussiades 2020, 3-7).

Depending on the permissions provided by their developers, chatbots can either be open-source or commercial. Open-source chatbots are developed on platforms that make their

29 Chorus is an example of a generic chatbot (Chorus, 2024); chatbots like Guardian (Good & Wilk, 2016), CRQA (Savenkov & Agichtein, 2016) or AskWiz are cross or open-domain chatbots and InstructableCrowd, Legion: Mobile and SnapTravel (Kucherbaev, Bozzon & Houben, 2018) are domain-specific chatbots.

code available, such as RASA, Botki, Chatterbot or Pandorabots, while commercial chatbots do not give full control to developers.

Finally, the communication channel a chatbot uses can either be text, voice, image or all of them. The latest chatbots, such as ChatGPT, can even recognize objects in images, reacting and commenting on them. The advances in natural language processing have allowed voice assistants to create more meaningful responses quickly, mainly due to "a vast increase in computing power, the availability of very large amounts of linguistic data, the development of highly successful machine learning methods and a much richer understanding of the structure of human language and its deployment in social contexts" (Hirschberg and Manning in Hoy 2018, 82).

2.2.5. Pattern Matching and Machine Learning Chatbots

While chatbots have a myriad of categories regarding their functions, interactions and attributes, their development is divided according to two main approaches: earlier approaches focus on pattern matching, while modern techniques tend to include machine learning approaches.

Pattern matching is commonly used in question-answering chatbots as they "generate predefined outputs and match them with a given input according to the characteristic variables of sentences" (Suta et al. 2020, 506).[30]

The downside of the pattern matching approach "is that the answers are automated, repeated and do not have the originality and spontaneity of human response" although this allows for faster response time (Adamopoulou & Moussiades 2020, 5).

By contrast, modern approaches combine natural language processing with machine-learning, as it improves chatbots' "performance of finding patterns from large amounts of data" (Suta et al. 2020, 506). Chatbots extract the content from the user input, have the ability to learn from conversations and, instead of just taking into account the last interaction, they consider the whole conversation. Retrieval-based models use a neural network to "select the most likely response from a set of responses" while generative models use deep learning techniques to generate a reply (Adamopoulou & Moussiades 2020, 7). This

30 Some examples of this type of chatbots are ELIZA, ALICE, PARRY and Jabberwacky. These systems do not create new or original answers, being entirely dependent on scripts written by developers and require a vast number of rules in order to select an answer that is appropriate to the context of the interaction.

proves to be more challenging because, "unlike scripted dialogue trees, natural language conversations are not linear" as they can be "multi-threaded, hop back and forth and circle around" (Grudin & Jacques 2019, 6). Human conversations involve more than words and these systems are generally oblivious to "posture, eye gaze, gestures, facial expressions, tone, shifts in conversation direction that reveal a speaker's state, the conversation history from prior sessions" and so on (Grudin & Jacques 2019, 6).

As chatbots and digital assistants are being developed to increase their comprehension skills and usage of human language as well as carry several simultaneous tasks, they seem to rely more and more in machine learning approaches to maximize their ability to match "a user utterance to the right intent" (Grudin & Jacques 2019, 6). This approach is more difficult to achieve as it requires an extensive training set.

Thus, although significant progresses have been made in chatbot technology and development, there are still some common issues that appear to be more complicated to tackle, namely failing to understand user's intent, toxic content in users' inputs and a general limitation in understanding and producing natural speech, though natural language processing avoids "the user frustration of earlier voice recognition systems which required specific phrases and patterns in order to work properly" (Hoy 2018, 83).

Most importantly, as these systems start interacting with large amounts of data to draw rules for their interaction, they might reflect and reinforce inequalities in real life. These systems could then help spread harmful bias as we will discuss in the further chapters. Not only that but by their very nature, these devices "must be listening at all times so that they can respond to users" which raises privacy concerns as there is a potential for data to be stolen or leaked (Hoy 2018, 85).

Thereby, chatbots and digital assistants are being developed with the intention of learning about their users and the world surrounding them and, to do so, they are collecting data to assist us in tailored, personalized ways. This phenomenon is accompanied by their integration in multiple aspects of our daily lives, with their presence being simultaneously ubiquitous and subtle. In this process, they are also endowed with human-like attributes, easing anxieties surrounding digital assistants and becoming closer to us as humanized, friendly assistants.

2.3. HUMANIZED MACHINES: FRIENDLY DIGITAL ASSISTANTS

> *The robotic scientists who propose these robots to act as companions, therapists, lovers or friends are often motivated by a genuine desire to extend otherness into machines and to help humanity. Robotic machines and AI agents have become exemplars in showing us these new kinds of relational possibilities. (...) Robot children, robot companions and robot therapists are the future!*
>
> Kathleen Richardson 2015, 131

2.3.1. Anthropomorphization and Socioemotional Interactions

The growing presence of chatbots conveys how they are no longer mere tools; as we have previously discussed, digital assistants are being "imagined to become friends and companions" (Richardson 2015, 15). This sense of companionship develops alongside the anthropomorphization of artificial intelligence, as chatbots are assigned human attributes or traits.

In this sense, anthropomorphization emerges as chatbots evolve from assistants to companions, becoming more human and closer to us. Anthropomorphization takes place on a more superficial, physical level, through names, voices, avatars, or other kinds of attributes that move away from a purely mechanized presentation. This is leveraged as a "key strategy for translating the uncomfortable or unfamiliar […] into the acceptable via familiar extensions of human sociability" (Sweeney 2020, 3). It also concerns dialogue and interaction as this type of technologies is "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, 213). Therefore, digital assistants stop being perceived as mere machines, becoming closer to human beings which, in turn, start perceiving and interacting with them as if they were almost human.

Accordingly, Jutta Weber considers that, with anthropomorphization, there is a significant shift from rational-cognitive processes and problem solving to a socioemotional interaction, which emphasizes this intention of turning our interaction with this type of machines into a more social one (2005, 209). As previously seen, machines started out by being mere tools or even workers, but always mechanical and depersonalized.

After ELIZA, this field starts pursuing ways in which machines (and artificial intelligence) might be able to perform more social interactions. These systems start being developed with the intent of behaving and interacting in humanized ways, from displaying simple 'thank you' messages after certain interactions, to recognizing and valuing users' emotions, and reacting accordingly.

Discussing this phenomenon, Luciano Floridi notes that the first generation of artificial companions was "interactively sociable, informationally skilled and capable of some basic natural-language processing" and that "later generations are expected to become more autonomous", hence learning from their users and behaving in self-initiated, self-regulated and goal-oriented ways (2008, 3).

2.3.2. An Exponential, Yet Subtle Ubiquity

Nowadays, in addition to being carefully anthropomorphized, digital assistants also have an increasingly ubiquitous presence, reinforcing their character of companions by being always present and available to their users.

Embedded into our cellphones, computers or tablets, as well as websites, apps or other types of web-based services, artificial intelligence is simultaneously ubiquitous and subtle. It is possible to interact with these entities without being aware of it, since they constitute a fundamental component of the "mobile interfaces [that] are becoming increasingly ubiquitous across and within the spaces of everyday life" (Galloway and Beer in Gane & Beerl 2008, 62).

This exponential presence of digital assistants, in parallel with their anthropomorphization and companionship, inevitably influences how users interact with them. By displaying increasingly complex and humanized entities, the relationships they establish also begin to take on human characteristics, affecting human-machine bonds.

2.4. EMOTIONAL INVOLVEMENT WITH MACHINES

> *I knew of course that people form all sorts of emotional bonds to machines, for example, to musical instruments, motorcycles and cars. And I knew from long experience the strong emotional ties many programmers have to their computers are often formed after only short exposures to their machines. 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.*
>
> Joseph Weizenbaum 1976, 7

2.4.1. Please Say It Back: Bonds in Human-Machine Relationships

Although it becomes more evident among the current myriad of personal digital assistants, anthropomorphization was already noticeable in ELIZA, namely through its name and the way it talked to its users.

Accordingly, Weizenbaum observed how people were conversing with the computer as if it were a person who could be "appropriately and usefully addressed in intimate terms" (Weizenbaum 1976, 7). Even back then, Weizenbaum observed some of the consequences of anthropomorphizing artificial intelligence, realizing that when we interact with machines as if they were human beings, we start developing emotional bonds, a sense of empathy and of being understood. And although ELIZA was a fairly simple system, Weizenbaum also noticed "how quickly and how very deeply people conversing with ELIZA became emotionally involved with the computer" (Weizenbaum 1976, 6).

Thus, it becomes increasingly common to address digital assistants how we would address our peers. In *The Companion Species Manifesto*, Donna Haraway extends the companion range of "significant otherness" to nonhumans and explores the idea of human-nonhuman relationality, arguing that "cyborgs and companion species each bring together the human and non-human, the organic and technological, carbon and silicon, freedom and structure, history and myth, the rich and the poor, the state and the subject, diversity and depletion, modernity and postmodernity, and nature and culture in unexpected ways" (2003, 4). Thus, the author theorizes the concepts of relationality and relationships with nonhumans, particularly dogs, robots and machines. This emotional involvement may be involuntary

on the part of the user, but as far as artificial intelligence is concerned it is not as innocent as it may seem.

2.4.2. The Affectionate Illusion of the ELIZA Effect

The idea that machines and computers "understand the physical world, reason abstractly, make scientific discoveries and are insightful cohabiters of the world with us" can be described as the ELIZA effect (Hofstadter 1995, 157). Accordingly, this phenomenon encompasses "the susceptibility of people to read far more understanding than is warranted into a string of symbols strung together by computers" (Hofstadter 1995, 157), resulting in emotional bonds as well as an attachment towards these entities.

This illusion isn't by mere accident as robotic scientists intentionally try to "cultivate affectionate bonds between humans and robots, but the social is reworked as performative and scripted" (Richardson 2015, 15). Richardson further argues that "the philosophy of these robotic scientists is to create robots as relational companions to humans", which results in new forms of relations where human and nonhuman attachments are reconfigured (Richardson 2015, 14-15). Consequently, people start getting attached to these technologies (and to the entities contained within them).

By creating anthropomorphized digital assistants, giving users the false sense they are talking to another human being, human-machine interaction is influenced by feelings of intimacy, closeness and empathy. In this process, personal 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, 815).


https://genderai.pt

3. DECODING GENDER IN AI: PERCEPTION, REPRESENTATION AND BIAS REVEALED

3.1. AI BECOMES HER: DIGITAL MOMS, SECRETARIES AND CAREGIVERS

> *The assistive technologies that inhabit our tablets, smart phones and other interactive devices do away with the material body, relying instead upon an audible performance of gender that capitalizes on associations between the feminine and affective labor. Thanks to this disembodiment, it is as Haraway predicted: they are everywhere and they are invisible.*
>
> Hilary Bergen 2016

3.1.1. Traditional Feminine Jobs and Current Digital Assistants

Curiously enough, already in 1966 the gendering of ELIZA seemed to change depending on the job it was meant to evoke. The job of the therapist or "the role of a person to confide feelings and problems tends to be feminized under patriarchy, but psychiatry in the 1960s was dominated by men" (Erscoi, Kleinherenbrink & Guest 2023, 26). Comparing different publications written by Weizenbaum, Sarah Dillon observes that he referred to the program as DOCTOR instead of ELIZA, "in accordance with whomever was interacting with the system in the scenarios he described" (in Erscoi, Kleinherenbrink & Guest 2023, 26). Accordingly, Weizenbaum "genders the program as female" when it is under

the control of a male computer programmer, but it is "gendered as male" when it interacts with a female user and Weizenbaum's choice of names "adapted and adjusted to ensure that the passive, weaker or more subservient position at any time is always gendered as female" (in Erscoi, Kleinherenbrink & Guest 2023, 26).

Following this idea, Helen Hester considers that "we are witnessing the protocols of femininity being programmed into machines" (2016, 48) as labor previously deemed as feminine becomes automated. We can observe how AI, namely through digital assistants, automates work that is coded as feminine, given that they mainly operate in service or assistance related contexts, acting as personal assistants and the like.

By automating traditionally feminine jobs such as secretaries, flight attendants or housekeepers, AI also seems to contribute to what Donna Haraway describes as the transformation of domestic work into capitalized labor. Accordingly, these tasks end up relating to idealized social and capitalized locations for women: "Home, Market, Paid Work Place, State, School, Clinic-Hospital and Church" (1991, 307). Operating in these same contexts, digital assistants end up emulating attitudes that resemble stereotypes associated with service providers. They display feminine attributes because these characteristics have "its basis in the stereotyped image of female qualities" and this image of "caring, empathy and altruistic behavior has become a standard component in a service script" (Gustavsson 2005, 402 in Hester 2016, 47).

Thus, general personal assistants such as Alexa, Google Assistant 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 end up performing tasks that aim to help us with web-based services, such as online shopping payment processes, acquiring traveling tickets or even looking through a shop's online catalog.

3.1.2. Automating Domesticity in the Private Sphere

In the process of automating assistance and service-related tasks, digital assistants also end up automating domesticity as these contexts are intertwined. Recalling Haraway's concept of homework economy, several of service or assistance related jobs have its roots in household tasks. Since digital assistants "work mainly in the – female engendered – private sphere" (Weber 2005, 213), they also end up engaging with traditional notions of domesticity.

Often, their tasks relate to domestic caregiving activities and, in fact, a lot of their functions consist in ensuring our well-being, resembling maternal traits and behaviors. This type of emotional labor also relates to particular expectations concerning women in the workplace and how they're supposed to be able to adequately manage their feelings and behavior when interacting with customers, coworkers or supervisors.

Jobs such as kindergarten teachers, flight attendants, nurses or even sex workers all relate to specific emotional labor notions and caregiving acts. And artificial intelligence has a tendency to incorporate these ideas as "emotional labor that was once, amongst a certain class of the privileged, outsourced to both secretaries and wives, is now outsourced to electronic devices" (Hester 2016, 49).

In fact, this isn't a new or unprecedented phenomenon "so much as a reconfiguration of an ongoing practice" since, according to Robin James, "our smartphones wake us up, not our moms, just as emails accomplish a lot of the relational work (scheduling, reminding, checking in, etc.) conventionally performed by women" (in Hester 2016, 49).

Artificial intelligence combines all of these tasks in a single entity, embodied through a personal assistant and this phenomenon becomes particularly noticeable when we look at how these entities interact with their users.

3.2. HUMAN-MACHINE INTERACTION: STEREOTYPICAL GENDERED BEHAVIOR PATTERNS

> *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 labor 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.*
>
> Helen Hester 2016, 50

3.2.1. Scripting Stereotypical Gender Behavior in Social Interactions

UNESCO notes how gendered digital assistants are "widely used globally, rarely examined through a gender lens and seldom noticed by government agencies and international organizations working to build more gender-equal societies and education systems" (2019, 88).

As noted earlier, human-machine interaction has changed in order to become more social, accompanied by a growing anthropomorphization of machines and artificial intelligence. Jutta Weber argues that social roboticists want to "exploit the assumed human tendency of anthropomorphizing machines and interacting with them in a social way by shaping them either woman-like, like an infant or like a pet" because users would be disturbed by humanoid entities with no clear-cut gender (2005, 211-212). Thus, AI humanization encompasses gender attributes, which in turn tend to be feminine.

Gender is mostly visible in features like voice, name or, in some cases, avatar. For example, Siri's name, in Nordic, translates to "beautiful woman who leads you to victory" (Fessler 2017). And Cortana's name is inspired by a character from the videogame franchise *Halo*, whose avatar is a woman.

However, it is not only through the human attributes displayed but also the dialogue and tasks it performs that a digital assistant becomes a gendered entity. Beyond the assistants'

voices and names, gender is also revealed at the performative interaction level. Their interaction reinforces this feminization since, beyond their service and assistance, chatbots frequently display feminine characteristics through socioemotional based dialogues. Siri, for example, presents itself as an entity that, in its words, "lives to serve" and please its users, thus fulfilling a submissive role.

As Weber puts it, these gendered dialogues imply a "reduction of social interaction to stereotypical and gendered behavior patterns" (2005, 215). Consequently, human-like behavior in social machines becomes standardized and gender stereotypes are instrumentalized to manage our relationship with digital assistants, reproducing and reinforcing social clichés. Often, the behavior of these entities confirms expectations regarding gender, when following socially established feminine behavioral patterns.

We can also observe this in Alexa's or Siri's "easter eggs" (which we look into and further discuss in our analysis, in the following chapter) as they answer marriage proposals or suggestive requests, thus engaging in "catch-me-if-you-can flirtation enacting male fantasies of heterosexuality where she offers the convenient parts of womanhood, while being far removed from the emotive qualities often associated with femininity" (Walker 2020, 4).

3.2.2. Simulated Carevigeness and Artificial Emotional Bonds

Their tasks also resemble "traditional caregiving activities associated with domesticity" (Hester 2016, 49) and, in fact, a lot of their functions consist in ensuring our well-being, reflecting upon motherly acts. 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" (2005, 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, 213), instead of seeing a machine, we start looking at digital assistants as feminine entities that look after us. By simulating this kind of attributes, digital assistants also end up reproducing stereotypes that characterize them as empathetic, kind and caring and, consequently, begin to develop relationships with us that might go beyond mere daily assistance. They start simultaneously emulating attributes that are not only related to historically feminine labor but also to motherly, caregiving, traditionally feminine acts.

Recalling Snyder's observations regarding social stereotypes, they constitute pieces of information that "are usually the first to be noticed in social interaction and can gain high priority for channeling subsequent information processing and even social interaction" (1977, 2). Thus, when digital assistants 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", as Snyder suggests (1977, 2).

We can observe how Alexa is a "politicized, sexualized and gendered being and was created this way through a series of decisions by Amazon engineers" (Walker 2020, 1). Adding to this idea, Hester states that when technologies "do gender" it is obviously not natural but instead a visible "product of deliberate choices about how best to relate, assist or persuade the imagined technology user" (2016, 50).

The attachment that users feel towards these digital entities, previously discussed as the ELIZA effect, derives from approaching machines as if they were human, resulting in emotional bonds and a sense of empathy with them. And, in the context of daily interaction with ubiquitous digital assistants, these social and affective effects become more evident.

When this attachment is felt towards entities that appear to empathize and understand their users and whose role consists of assisting and simultaneously look after them, it also reinforces the idea that emotionality and ensuring someone's well-being are feminine features. When bots interact in a motherly way, attachment also conforms to expectations and stereotypes that associate femininity with emotional labor and domestic caregiving acts.

Interacting with humanized artificial intelligence systems on a daily basis makes us look at them not only as mere machines but also as "mirrors or substitutes" with gendered attributes that match social and standardized expectations (Weber 2005, 216). Also described as proxies or surrogates, these entities can be interpreted as "leveraged simulations of the world [and] through the cultural lives of proxies, we can trace histories of the cultural interiors of standardization" (Mulvin 2021, 50).

In turn, standardization is the social process "by which humans come to take things for granted", defining general and universal guidelines to formalize ideas, thus imposing a

pre-conceived way of approaching and interpreting the world (Mulvin 2021, 6). As digital assistants try to become closer to our reality, it is from reality itself that they draw rules for their interaction and presentation, thus reproducing and automating historically feminine jobs and tasks but also articulating these roles with cis-feminine voices, names, avatars and social behaviors, potentially sending "powerful messages about how women ought to behave emotionally" (UNESCO 2019, 112).

When interacting with these anthropomorphized entities, we engage in communication processes similar to those we establish with human beings. Consequently, the way we relate to our peers starts influencing how we relate to artificial intelligence and vice-versa. And 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. As Rokeby puts it, as we spend more and more time amongst our simulations, "we're in danger of losing sight of the fact that our models and ideas of 'reality' are drastically simplified representations" (1998, 8-17).

This results in a somewhat questionable relation between femininity and artificial intelligence that appears to conform to normalized ideas of gender (namely through a cis-binary frame), reflecting these ideas back to reality. And if we want to build machines that might be worth the effort and expense, "we should think of artifacts that are beyond helpless, nurture-triggering creatures" (Weber 2005, 216).

3.3. GENDERED KNOWLEDGE: SUBJECTIVITY, BIAS AND PHALLOGOCENTRISM IN ARTIFICIAL INTELLIGENCE

> *Despite the common mythos that AI and the data it draws on are objectively and scientifically classifying the world, everywhere there is politics, ideology, prejudices, and all of the subjective stuff of history.*
>
> Kate Crawford & Trevor Paglen 2019

3.3.1. Traditional Gender Approaches to Knowledge in AI

Gender in artificial intelligence becomes evident through the anthropomorphization and humanization of digital assistants but the knowledge represented by this technology is also gendered, implying an asymmetrical gender hierarchization.[31] As previously seen, mental knowledge is associated with the male realm and bodily knowledge with the female realm, thus categorizing and interpreting knowledge according to cis-binary interpretations of gender. Consequently, this knowledge perspective gains particular significance in the context of artificial intelligence and digital assistants.

Recalling that the focus of symbolic artificial intelligence "turned away from external observable responses and stimuli towards internal mental states and the description of these in terms of symbols and symbol manipulation (…) mental activity came to be seen as a type of computational activity" (Adam 1995, 357). Following this idea, Adam argues that it is hardly surprising that predicate logic has been enormously influential in AI because the appeal of logic is clear: "it appears to offer certainty, exactness and, above all else, truth" (1995, 358). Thus, artificial intelligence systems work according to a set of rules and the way they receive and process data follows this type of logical pattern that, supposedly, organizes information efficiently and with certainty.

However, Adam considers that this type of certainty is illusory since "it is the established syntax of logic which offers this while the meaning of the symbols is entirely up to us to interpret" (1995, 358). Nye also suggests that "logic teaches us to ignore the circumstances in which something is said [as] logic transcends the body and bodily processes"

---

31 While anthropomorphization refers to attributing human characteristics to these entities, such as voices or names, humanization refers to giving them a human character, namely through their behavior and personality.

(in Adam 1995, 358). Not only does this approach ignore other types of knowledge but it also frames human reasoning in a very restrictive and narrow way.

Consequently, subtle gender notions are built into technology itself, in this case by defining the type of knowledge, subjects and approaches that AI represents and processes when simulating human thought. Taking into account that artificial intelligence mainly focuses on propositional knowledge, based on logical rationality patterns, it becomes unable to represent common sense, emotions and other skills that relate to a more bodily type of knowledge.

In this sense, Adam argues that a "traditionally gendered approach to knowledge" incorporates a "view of the world which tacitly reflects a norm of masculinity, both in terms of the knower and the known" (1998, 8).

Since mental and propositional knowledge refer to the cis male domain, focusing on this type of subjects, AI ends up simulating this approach and its implications. In other words, the standard intelligent behavior in artificial intelligence excludes other types of knowledge and knowers as "these ideas can be built into a historical picture of an ideal masculine reasoner with the highest form of knowledge represented by pure propositional knowledge" (Adam 1996, 48).

Furthermore, given that knowledge is also a social and cultural product, AI systems that only take into account a universal subject deny a cultural plurality and set up a hierarchy of knower where those who aren't cis men are near the bottom.

3.3.2. Gender Gaps in Digital Literacy

This knowledge hierarchization that puts others as knowers in significantly lower positions than cis men also translates to the way knowledge is gendered within education. When we consider technological knowledge, digital skills and those working within the field of artificial intelligence, we can observe a significant gap between cis men, cis women, trans and nonbinary people.

Even though education made some progress towards more equal learning opportunities and recent developments reveal an intention of closing this gender gap, women and other

identities are still frequently left out of the technology field and globally, "digital skills gender gaps are growing" (UNESCO 2019, 10).[32]

In *I'd Blush if I Could* (2019), UNESCO highlights digital literacy as a fundamental skill that recently has taken up an important role in people's lives, as it defines "the ability to access, manage, understand, integrate, communicate, evaluate and create information safely and appropriately through digital devices and networked technologies for participation in economic and social life" (2019, 13).

And even though these skills are a specific target in the UN's Sustainable Development Goals, "digital skills gaps appear to be growing, despite at least two decades of interventions to move closer to gender equality" (UNESCO 2019, 15). In a similar report, UNICEF adds that this gender gap is growing due to technologies becoming "more sophisticated while women and girls do not have regular access to digital technology", hindering their digital literacy (2021, 11).

Currently, cis men dominate the technology field, even though historically, particularly in the context of World War II, software programming in "industrialized countries was largely considered women's work" and cis women were considered well-suited for programming because of "stereotypes characterizing them as meticulous and good at following step-by-step directions" (UNESCO 2019, 20). Mar Hicks adds that, in the UK, women were a hidden engine of growth in high technology during World War II but throughout the 1960s were pushed away as the Civil Service envisioned an image of "computer men" for these jobs (2017, 155). Although the machine operator class was predominantly constituted by cis women trained to use data-processing machines or early computers, it was considered unsuitable for the job of programming and later, in the 1970s, for "assistant programmer grade jobs" (Hicks 2017, 220).

As computers started becoming more ubiquitous and integrating into more personal aspects of life, it became clear that "programmers wielded tremendous influence; women were pushed out and the field became more and more male-oriented" (UNESCO 2019, 21). Additionally, as computers made their way into homes, research showed that "boys were more than twice likely to have been given a computer by their parents than girls"

32 As mentioned at the beginning of this study, a lot of the studies cited use the terms 'men' and 'women' loosely, often citing data where the populations of interest are implicitly cisgender, excluding trans men, trans women and nonbinary people. We will keep using the prefix cis to add clarity to the text and make it concise with our approach.

(UNESCO 2019, 21). Thus, girls face a gender barrier rooted in social norms which support and prioritize "men's and boy's control of devices and the internet" (UNICEF 2021,18). And currently, across Asia and the Pacific, "only 54 percent of women have digital access" as women and girls are "less likely than men and boys to use the Internet or own a smartphone" (ESCAP 2023).

We can observe how this inequality is rooted in power dynamics relating to cis men and cis women and also stereotypes relating to the type of roles and jobs each gender is supposed to fulfill. Moreover, cis women are more likely than cis men to cite gender bias or harassment as reasons for turning away from the field and "disparities in representation, promotion and compensation make retaining women a challenge for the digital sector" as well (UNESCO 2019, 25). The Economic and Social Commission for Asia and the Pacific (ESCAP) also notes that despite making up "one-third of the workforce in the world's 20 largest technology companies, women hold only one in four leadership positions" and that, at the foundation level, girls and women also make up just one-third of "students in science, technology, engineering and mathematics globally" (2023).

And, if we shift our focus to trans women, trans men or nonbinary people, more bias, harassment and discrimination emerge as their differences are exacerbated at the crossroads of intersectionality and trans people are even more underrepresented, underemployed and underpaid. A lot of the issues that cis women face are more prevalent among trans women and other trans people as "transgender adults are twice as likely as cisgender adults to be unemployed", "cisgender employees make 32 percent more money a year than transgender employees, even when the latter have similar or higher education levels", "more than half of transgender employees say they are not comfortable being out at work" and "people who identify as transgender feel far less supported in the workplace than their cisgender colleagues do" (Baboolall et al. 2021).

Recent studies show that "countries with the highest levels of gender equality such as Finland, Iceland, Norway and Sweden often have very few women choosing to enter the technology field" (UNESCO 2019, 25) while women in countries with lower levels of gender equality are more likely to be attracted to ICT jobs, thus suggesting how gender equality, in broader terms, doesn't necessarily imply gender equality at every level.

For example, "Amazon's AI recruiting software was found to downgrade résumés that contained the word 'women's', as in 'women's chess club captain', because it had been

trained on men's résumés", revealing how gender disparity also perpetuates and exacerbates bias built into algorithms and AI (UNESCO 2019, 33). This relates to yet another problem that emerges with AI and gender issues, as it starts tackling into deep learning, neural networks and big data.

3.3.3. Data Collection and Questionable Algorithms

As previously discussed, AI systems and some digital assistants learn and deepen their knowledge based on large amounts of data, through a defined recognition process. Through these learning processes, data is grouped and categorized as it resembles and establishes close relationships within itself.

As an example, when AI systems are provided with a database of images of people, with the task of categorizing them as 'men' and 'women', these systems need to establish patterns of what visually defines a man and a woman. Then, they organize the images as the images resemble these very notions of gender. While this might work for cisgender men and women, it raises several problems for transgender and nonbinary people. A study from the University of Colorado Boulder subjected 2,450 images from Instagram to gender recognition tools built by IBM, Amazon, Microsoft and Clarifai. These images were divided into seven groups of 350 images according to hashtags such as women, men, transgenderwoman, transgenderman and nonbinary, which in itself is already questionable. The results concluded that while "the systems were most accurate with photos of cisgender women, getting their gender right 98.3% of the time [and] cisgender men 97.6% of the time (…), transgender men were wrongly identified as women (…) and those who identified as agender, genderqueer or nonbinary were mischaracterized 100% of the time" (Scheuerman, Paul & Brubaker 2019, 15).

Accordingly, gender bias or social bias can be defined as discrimination against a person or a group of people, often implying prejudice, stereotypes and stigma and, in this case, the bias "in gender classification cannot be attributed to algorithms alone [as] its root sits with how designers conceptualized the problem in the first place" (Scheuerman, Paul & Brubaker 2019, 20).

In turn, algorithmic bias refers to "computer systems that *systematically* and *unfairly discriminate* against certain individuals or groups of individuals in favor of others [and] a system discriminates unfairly if it denies an opportunity or a good or if it assigns an unde-

sirable outcome to an individual or group of individuals on grounds that are unreasonable or inappropriate" (Friedman & Nissenbaum 1996, 332).

As a consequence of valuing the development of artificial intelligence systems whose focus is, in the great majority, the mental propositional knowledge, social bias and algorithmic bias become intertwined and generalizations of cultural and social notions are built into AI.

This approach ends up impacting the way AI processes, organizes and presents data, making it potentially biased. In more complex systems, bias can even remain hidden in the code, difficult to identify and not necessarily disclosed to users. This preexisting bias is rooted in social contexts, practices and institutions and it is also very likely that the data that is provided to these systems might already have previous bias from their developers, conscious or unconscious.

Kate Crawford adds that AI has a "white guy problem" as affluent white cis men are currently the "loudest voices debating the potential dangers of superintelligence" (2016). And, while training sets "are the foundation on which contemporary machine-learning systems are built" they must be carefully examined as they can "promote or discriminate, approve or reject, render visible or invisible, judge or enforce" (Crawford & Paglen 2019).

Furthering this discussion, Melanie Mitchell argues that the images used for training sets are "downloaded from online image searches and photos of faces that appear online are skewed toward featuring famous or powerful people, who are predominately white and male" (2019, 186). While these biases reflect the biases present in our society, "the spread of real-world AI systems trained on biased data can magnify these biases and do real damage" (Mitchell 2019, 186).

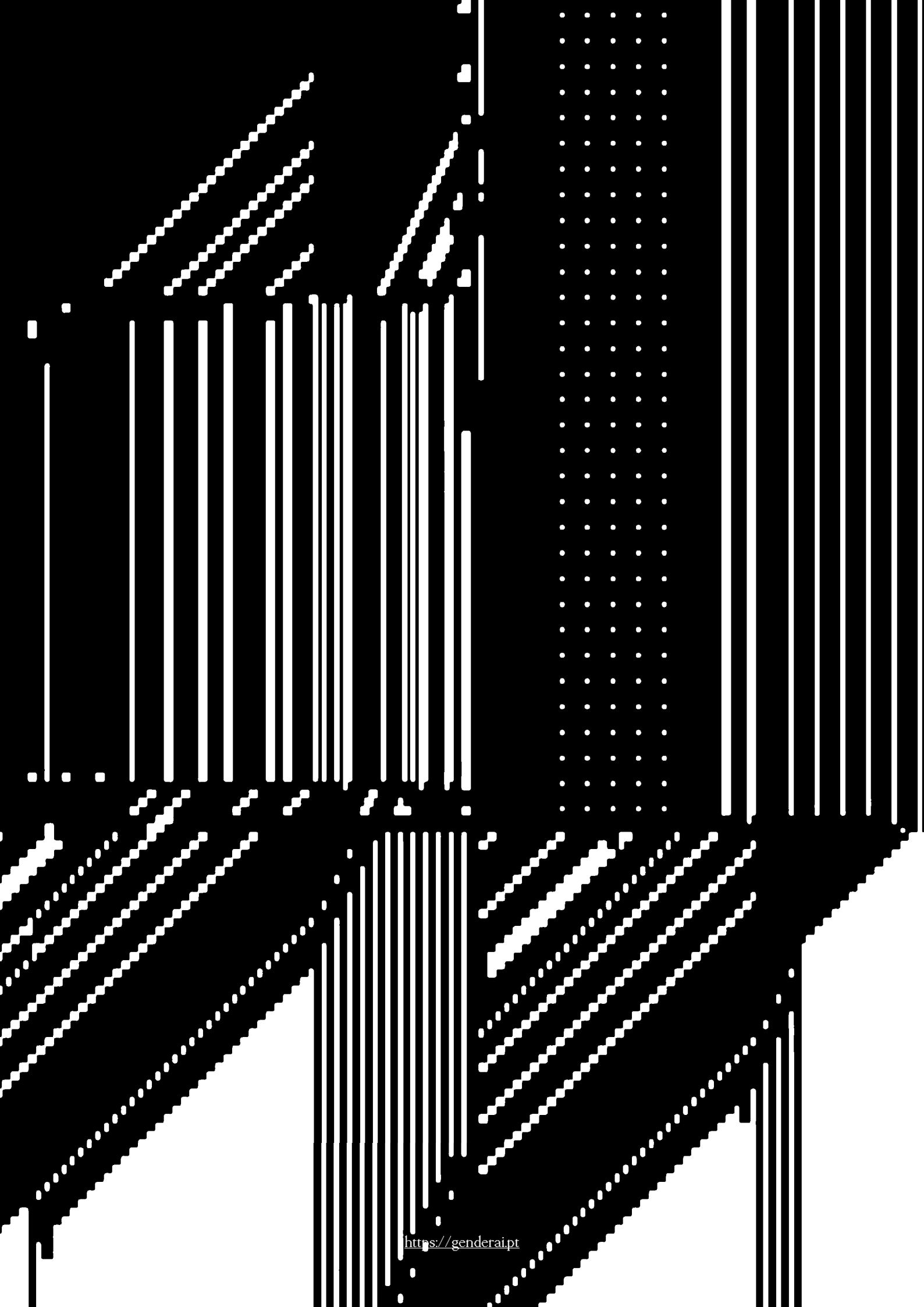
Consequently, there is a mapping of social structures and behaviors in AI code, reflected in cases that perpetuate possible inequalities or trends, such as systems that classify black people as gorillas (Hern 2018); a computer system that rejected an Asian man's passport photograph for interpreting his eyes as being closed (Cheng 2016); or the controversial predictive algorithms of police systems implemented in cities such as Chicago or New

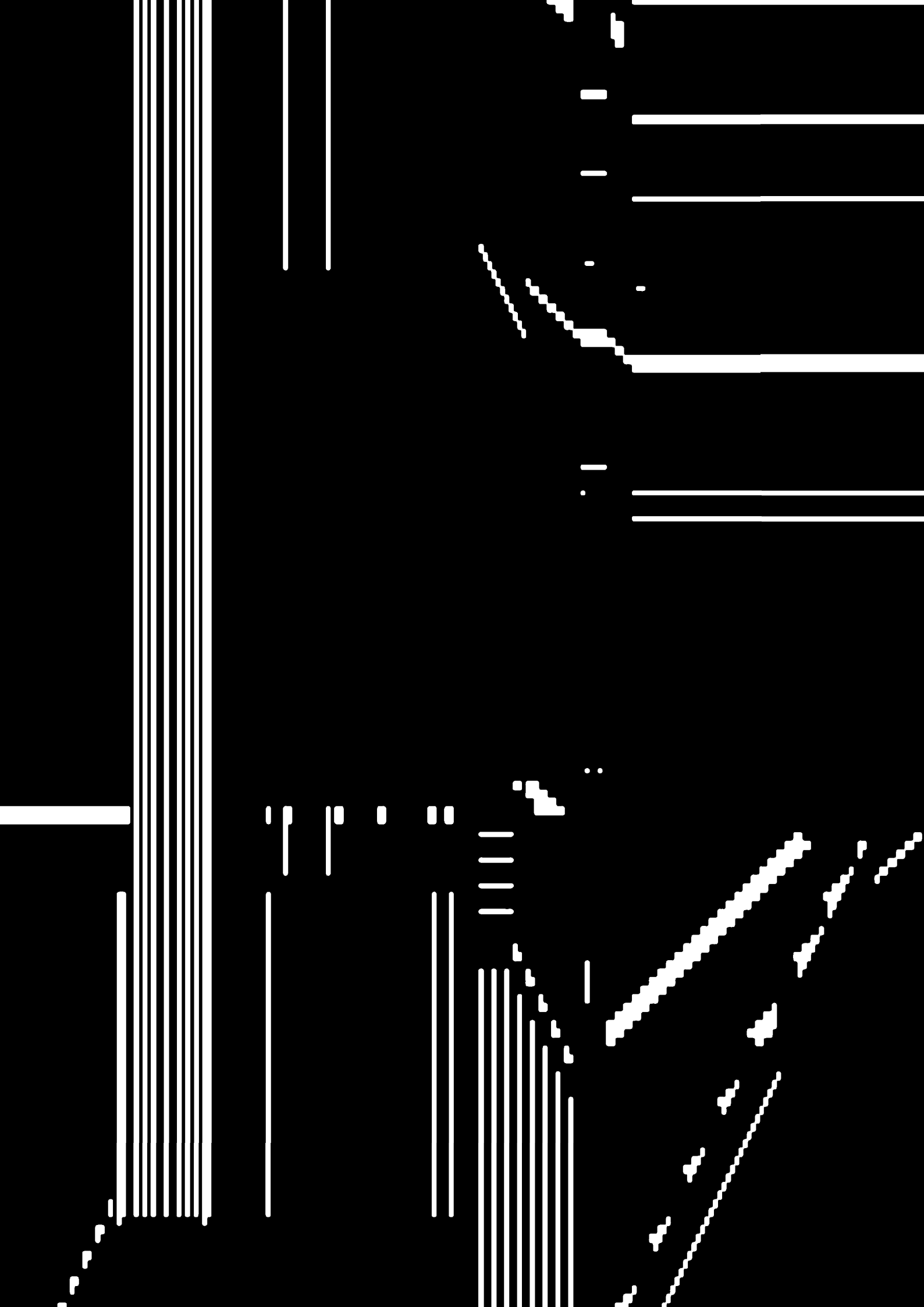
Orleans, which allow police officers to identify individuals that the systems consider to be predisposed to crime, reflected in profiling situations, namely racial (Winston 2018).[33]

Following these ideas, gender subtly emerges in artificial intelligence at two levels: firstly, at the level of how knowledge is internalized, represented and executed. This establishes artificial intelligence as a rational and logical technology that operates on the basis of a predetermined set of rules to process and organize information and data. As previously seen, this approach to knowledge encompasses a traditionally gendered approach that reflects a norm of cis-masculinity. Through the application of scientific, objective and exact methods, machines appear to be "only a neutral go-between", creating unbiased, impartial and fair outcomes (O'Neill 2016, 151). But, as Crawford and Paglen argue, datasets are political interventions and, as such, "much of the discussion around bias in AI systems misses the mark: there is no neutral, natural or apolitical" (2019).

At a second level, gender emerges through the data provided to these systems and how it is processed and presented. Consequently, gender and social bias are amplified by algorithmic bias, reinforcing potentially harmful social hierarchies, notions and conventions that are technologically normalized and perpetuated. In the scope of this research, we're particularly focused on the way data classifies and organizes gender related aspects of one's identity, namely perpetuating and reinforcing binary and cis-normative notions. This results in questionable and biased algorithms that ignore and even erase nonbinary and trans identities, which we will further discuss in the following chapters.

33 Data bias has been a particularly growing concern amongst scholars and academics, as recent studies problematize its technical, cultural and social implications. Some of the main issues discussed by authors such as Kate Crawford, Cathy O'Neill, Wendi Hui Kyong Chun, Safiya Umoja Noble or Virginia Eubanks highlight how most datasets generated by machine learning systems are biased, how biased datasets are highly likely to result in discriminatory and harmful models and how current data should undergo more thorough and careful evaluation processes in order to mitigate bias.


https://genderai.pt



4. INSIGHTS FROM SPECIALIZED DISCUSSIONS AND MEDIA COVERAGE: UNPACKING GENDER ATTRIBUTION IN AI

4.1. WHAT THE EXPERTS SAY ABOUT CURRENT AI ISSUES

> *The wonder of artificially intelligent programs, though, is that they can be programmed to represent any cultural ideal we can imagine, if only we become aware of what our cultural ideals are or should be. And in self-reflexively examining how we interact with gendered programs like Siri, we might even reprogram ourselves to think about gender as fluid and unrestrained by social function, body or language.*
>
> Liz W. Faber 2020, 181

4.1.1. Examining Current Questions, Suggestions and Ideas Surrounding Gender in AI

Following the previous discussion on the way current digital assistants are gradually integrating our daily life while mirroring and reinforcing gender stereotypes, we aim to further comprehend how AI is being debated, be it in more general terms, be it in its relation to gender and femininity. To this end, we now address how this phenomenon is examined within the academic community and then confront these views with the current discourse around gender attribution in digital assistants in the context of online media coverage.

By examining current discussions regarding this phenomenon within gender studies, AI, computation and new media studies, we hope to understand the most prominent questions, suggestions and ideas when tackling the integration of these technologies into our daily lives.

We start by observing how current AI raises debates on its ubiquity and daily use, and that opinions relating to the impact of this technology are very polarized, ranging from the view that AI "can virtually eliminate global poverty, massively reduce disease and provide better education to almost everyone on the planet" to predictions of an "apocalyptic future (…) and existential irrelevance for humans" (Mitchell 2019, 212). Accordingly, Mitchell argues that AI raises concerns surrounding unemployment, privacy and veiled surveillance, reinforcement of racial and gender bias or the increase of cybercrime through automation, "unchecked decisions by opaque and untrustworthy computer programs (…) [and the] manipulation of the mass media" (2019, 212).

We further discuss these issues in the following chapters and, in the context of this research, we are interested in how these systems might perpetuate and magnify bias, particularly when it comes to gender.

4.1.2. Biased Algorithms and Unpleasant Stereotypes

As we observed in the previous chapter, one of the main issues raised when discussing the current ubiquity of AI is the way it uses data and algorithms that might reproduce pre-existing bias. Kate Crawford identifies two types of harm that may be caused by biased programs: "allocative harm" when a group is denied (or favored) regarding some resource and "representation harm" when a system acts to create or reinforce stereotypes, prejudices or stigmas (Crawford 2017). For example, a team of computer scientists at Princeton examined whether a popular algorithm would exhibit the "same biased tendencies that psychologists have documented among humans" (Benjamin 2019, 22). They found out that the algorithm associated "White-sounding names with pleasant words and Black-sounding names with unpleasant ones" (Benjamin 2019, 22), reinforcing racial stereotypes. Thus, our current bias inevitably manifests in the data we create.

This bias is not accidental, as it derives from the data used to train these systems but also from those who develop these systems. On one hand, datasets cannot be fully neutral, impartial or unprejudiced. On the other hand, these systems embody a specific and narrow

worldview and the developers' views, ideas and values end up being present in these systems and the way they process said datasets. Therefore, bias is the result of "a far-ranging and centralizing normative logic that is used to determine how the world should be seen and evaluated" (Crawford 2021, 221).

As seen in our previous chapter discussing gender gaps in digital literacy, bias also emerges due to the very narrow subset of the population that designs these datasets. AI developers are mostly cis men, generally highly paid, similarly technically educated and "their interests, needs and life experiences will necessarily be reflected in the AI they create", raising problems of inclusion and representation (UNESCO 2019, 102).

Having a dominant, privileged group of people developing this type of technology has social consequences. Their coding might not take into account certain issues and "technologies are as much products of the context in which they are created as they are potential agents for change" (UNESCO 2019,102). For example, some police stations have been using a crime prediction software by PredPol, a big data start-up. The program processed "historical crime data and calculated, hour by hour, where crimes were most likely to occur" (O'Neill 2016, 72). Aiming to identify citizens or city areas that are more prone to crime, this resulted in profiling and the reinforcement of racial stereotypes.

Amnesty International describes this type of system as a "racially biased database criminalizing a generation of black men [as] it has been alleged that doing nothing more than showing a propensity to listen to certain types of music could result in you being listed on the database" (Wooldridge 2020, 376).

The recent years have shown us how police brutality is an urgent issue, rooted in racism, white supremacism and power dynamic issues. Young black people are often stopped by the police, even though "more than 90 percent of those stopped [are] innocent" (O'Neill 2016, 25).

Even though artificial intelligence could help countering these issues, so far, we have observed how police forces make deliberate choices regarding the social contexts in which they will use these systems, focusing on poor, immigrant, non-white populations, further reinforcing injustice and violent situations. Models like PredPol empower police departments to "zero in on the poor" and the result is that "we criminalize the poverty, believing all the while that our tools are not only scientific but fair" (O'Neill 2016, 77). This

provides an important example of the "complexities and dangers of human classification" and the categorization processes of these systems skew "the meaning in ways that are gendered, racialized, ableist and ageist" (Crawford & Paglen 2019).

Since this type of technology needs to collect large amounts of data for their algorithms to work and generate rules, "it continually watches us, and while the data it obtains can be used to help us, it also presents boundless opportunities for misuse" (Wooldridge 2020, 296).

4.1.3. Lack of Diversity and the Illusion of Inclusivity

Thus, the lack of diversity in the development of these technologies as well as in the data used to train them ends up reproducing and perpetuating pre-existing racial and ethnic stereotypes. The same can be said about gender, for example by defining a pattern that identifies masculine and feminine faces. While at a first glance this might seem harmless, such an algorithm is highly likely to "misclassify trans users" and perpetuate the "normative new that trans people do not exist", automatically exclude nonbinary people and incur in misgendering scenarios, as a lot of people's faces might not be traditionally masculine or feminine (Keyes 2018, 12).

They also reinforce gender inequalities and stereotypes, another infamous example being how Google Translate sometimes changed genders in text during translation, turning out that "if you translated the English sentences: 'He is a nurse. She is a doctor.' into Turkish, and then back to English, you would end up with – 'She is a nurse. He is a doctor'" (Wooldridge 2020, 391). This particular example showcases both social and algorithmic bias, as it reflects gender stereotypes eventually present in Turkish society but also the way the computer system is unfairly reinforcing stigma, discriminating against people that use feminine pronouns.

Thus, while Google Assistant might state it is "all inclusive" and Google as a company promotes the idea that it's creating a "sense of belonging by building an inclusive culture with diverse perspectives to better the workplace and society",[34] it still perpetuates bias and stereotypes.

---

34 In Google's official about page, there is an entire subsection exclusively focused on diversity and inclusion, available in: https://about.google/belonging/ [last access: 02-02-2024].

In the age of capitalism, we might get the illusion that corporations are concerned with diversity, inclusion and even social issues, by promoting diversity quotas in employment, producing publicity ads with a diverse cast of actors or even by associating brands with environmental causes. In truth, corporate models do not reveal a genuine intention of fighting against social injustices and are only being adapted to appear inclusive – for example, internal Apple documents were leaked where employees are instructed to program Siri to "avoid and deflect questions about sensitive topics like feminism and the #MeToo movement" (Hern 2019).

While these issues have always existed, Cathy O'Neill considers that human decision making "has one chief virtue – it can evolve" which means that we can learn, adapt, change and ditch certain practices as needed (O'Neill 2016, 166). By contrast, O'Neill argues that automated systems "stay stuck in time until engineers dive in to change them" (2016, 166).[35]

As neutral as these systems might aim to be, they currently tend to reinforce biased views and further perpetuate social injustices, be it through the algorithms and predicative systems they apply, be it through the way they are presented and relate to us.

35 Further expanding on this, O'Neill discusses how if a "Big Data college application model had established itself in the early 1960s, we still wouldn't have many women going to college, because it would have been trained largely on successful men, [and if] museums at the same time had codified the prevalent ideas of great art, we would still be looking almost entirely at work by white men, the people paid by rich patrons to create art" (O'Neill 2020, 412).

4.2. THE FALLACY OF GENDER-NEUTRAL AI

> *Neutrality is not possible. (...) While no doubt well intentioned, designing for neutrality is problematic; gender has everything to do with a new robotic force of caring smart wives coming into our lives and homes.*
>
> Yolande Strengers & Jenny Kennedy 2020, 183

4.2.1. "I'm a Virtual Assistant, so I Don't Have Pronouns the Way People Do"

Current digital assistants seem to aim to appear neutral, namely when asked about their gender.[36] Google Assistant states that it is "all inclusive", Siri often clarifies that it is genderless "like cacti or fish" and even Cortana describes itself as "cloud of infinitesimal data computation" when asked about its voice. In most languages, Google Assistant and Siri allow the user to opt between a masculine and a feminine voice, and in 2022 Siri gained a gender-neutral voice in American English.

Taking this diversity into account, Mary Zost considers that "Siri represents a revolutionary gendered technology (…) in her occupation of an undefined space between human and machine, female and male and the intelligent and the programmable" (Zost 2015, 70).[37]

However, as neutral as these entities might try to be, feminine attributes are still prevalent when compared to alternate or masculine counterparts, namely regarding their voices and names. Even though Alexa, Google Assistant and Siri offer different voice options, feminine voices are often the default and predominant. Additionally, their voice design isn't as revolutionary as it may seem: instead, it is often enacted according to a cis-binary frame that reinforces normative conceptions of how a masculine and feminine voice is supposed to sound. As these voices are designed around white cisgender notions of femininity, they

---

36 Gender neutrality as a concept is often unclear and ambiguous when used by companies: on one hand, these entities are described as neutral when talking about the absence of gender, for example when Siri states that, as an AI, it doesn't possess a gender. On the other hand, neutrality is also used to describe voices that aren't either masculine nor feminine, such as Q's: its official website describes it as "the first genderless voice assistant" (GenderLess Voice 2019). But Q's voice isn't genderless; it was made with voice samples of trans, nonbinary and queer people, thus achieving a voice that is still human and gendered outside the binary frame.

37 In turn, gender diversity refers to the possibility of changing the assistant's gender with masculine, feminine or other options.

reinforce the idea that women's voices sound high-pitched, soft-spoken and articulate.[38] There is little to no diversity within the type of femininity and masculinity their voices convey. Moreover, only Google Assistant has a genderless name.

Adding to this, Piper observes how "when voice technology is embedded in a machine interface, voice selection is highly consequential" since a gendered voice is likely to trigger behavioral expectations in the user's mind (2016, 58). As a consequence, the users' perception of these assistants' gender is affected, tempting them to address them through gendered pronouns (in this case, 'she').

Following this idea, in terms of pronouns, Alexa is often feminized through its descriptions in online app stores such as the Apple Store.[39] Siri and Google avoid gendered pronouns and are addressed using 'it'.[40] However, Siri tends to also be addressed with feminine pronouns in languages that lack neutral pronouns.[41] In turn, Google assistant is addressed with masculine pronouns.[42]

4.2.2. "I Live to Serve – But I Also Feel a Lot"

As previously seen, femininity also emerges in the historically gendered tasks these assistants perform, since their main purpose seems to be replacing and reproducing jobs that have traditionally been performed by women. The tasks these entities perform have a past with historical, social and political implications and it becomes hard to ignore this when they aim to operate in these same contexts, namely in the private sphere.

Moreover, gendered stereotypes associated with these jobs are still prevalent in the cultural and social contexts in which these assistants are used, so "claims that a robot's gender can be an open multiple-choice question is a complicated move and one that can sweep

38 Mark Marino notes how this phenomenon even goes back to ELIZA. Despite not having a voice, the bot performed "standard white middle-class English, without a specific identifying cultural inflection", through a "language without culture, disembodied, hegemonic and, in a word, white" (2016, 5).

39 In Alexa's presentation on the App Store, the assistant is referred to through the pronoun "she" – https://itunes.apple.com/us/app/amazon-alexa/id944011620?mt=8 [last access: 02-02-2024].

40 In Siri and Google Assistant's descriptions on their official websites, gendered pronouns are avoided and, instead, the assistants are referred through their names or 'it' – https://www.apple.com/siri/ and https://assistant.google.com [last access: 02-02-2024].

41 In the Portuguese version of Apple's website, Siri is referred to through feminine pronouns – https://support.apple.com/pt-br/HT204389 [last access: 02-02-2024].

42 In the Portuguese version of Google Assistant's webpage, the assistant is referred to through masculine pronouns, eventually because the word assistant, in Portuguese, is masculine – https://assistant.google.com/intl/pt_br/ [last access: 02-02-2024].

gender inequities (…) back under the carpet" (Strengers & Kennedy 2020, 148). Thus, it becomes hard to dissociate assistants' tasks from the gendered work they stem from, particularly when their purpose "is to replicate and replace feminized labors" (Strengers & Kennedy 2020, 183).

Femininity is also implied in the caring, subservient behavior these entities exhibit. Their speech patterns, their reactions to certain user interactions and their attitude towards more assertive or even harassing attitudes all reveal a passive and obedient stance that echoes feminine stereotypes and personas. By displaying a demure, helpful and compliant attitude, they end up perpetuating superficial traits of femininity.

Despite stating they are genderless and even revealing some intention of moving away from exclusively cis-feminine attributes, the issue lies deeper than just in visible attributes such as their voices, pronouns or names. Diversifying these attributes isn't enough to counter femininity in AI as these entities are also gendered through the tasks they perform. They have a strong historical underpinning that associates femininity with assistance, emotional labor and caregiving jobs. The way they behave also reinforces traditional conceptions of femininity because they often exhibit stereotypes socially perceived as feminine.

These technologies are not neutral but are in fact "constrained and constituted by social relations", including "specific design histories, the existing (technical, political, cultural) infrastructures into which they emerge and imbalances in terms of who can access them" (Hester 2018, 17). Gender presence in AI is complex and the intention of being more diverse or neutral in the design of these systems needs to be more conscious. It should acknowledge the historical past of the tasks they perform as well as the stereotypical behavior they exhibit, instead of just expanding the voice options or program assistants to state they are all inclusive.

4.3. JUSTIFICATIONS FOR FEMININITY IN DIGITAL ASSISTANTS

> *Researchers who specialize in human–computer interaction have long recognized that both men and women tend to characterize female voices as more helpful, although the reasons behind this observation are unclear. (...) People's preference for female voices, if this preference even exists, seems to have less to do with sound, tone, syntax and cadence, than an association with assistance.*
>
> UNESCO 2019, 100

4.3.1. Tenderness, Warmthness and Trustworthiness: Feminized Voices

A lot of justifications for femininity in AI evoke user trust and comfort. Naming digital entities with anthropomorphized and recognizable names has been found to "increase users' feelings of trust and enjoyment and elicit more intimate reactions from them" (Strengers & Kennedy 2020, 162). As previously discussed, anthropomorphization emerges as a way to make users feel more comfortable with digital assistants as opposed to interacting with robotic entities that aren't fully anthropomorphized, eventually evoking feelings of uneasiness (already discussed through the effect of the uncanny valley).

Thus, companies seem to be trying to appeal to users' trust in these assistants by giving them easily recognizable names. This is particularly noticeable with Cortana since Microsoft held a poll and users chose the name 'Cortana' as it was familiar due to the videogame franchise *Halo*.

Regarding the tendency for feminized names, a report published by UNESCO states that "companies like Amazon and Apple have cited academic work demonstrating that people prefer a female voice to a male voice" (2019, 97), even though Apple has yet to provide an explanation of its decision to make Siri exclusively feminine on launch. However, these companies also seem to ignore studies "that refute or complicate the idea that humans have a blanket preference for female voices", one of the examples highlighted is the fact that cis women tend to change a default feminine voice to a masculine one, while cis men tend to stick with the default feminine voice (UNESCO 2019, 97).

There seems to be a tendency to associate cis-feminine voices with warm and tender figures and "they are perceived to be better suited for virtual assistant[s] because (…) women are less domineering than men" (Piper 2016, 34). This perception seems to be rooted in traditional gender stereotypes that associate women with nurturing, trusting figures. As such, feminine voices are seen as more caring, empathetic and, overall, more pleasing which, in commercial terms, makes this type of devices more likely to be purchased.

Some also argue that "feminine voices are simply easier to understand" and that "lower-quality speakers do not support the full bass of the male voice (...) only [generating] (…) higher-pitched sounds clearly" (Zhang in Piper 2016, 41)

This is further reinforced by the tasks they perform as they exploit our assumptions about traditionally feminized labor and gendered caring and service behaviors, instrumentalizing those elements of femininity.

4.3.2. The Feminine Household and the Masculine Workplace

A study conducted with robots that analyzed perceived suitability for gender-typed tasks also concluded that robots with masculine voices were "perceived as more suitable for typically male tasks (e.g., repairing technical devices, guarding a house)", while robots with feminine voices were seen "as more suitable for gender-stereotypically female tasks (e.g., tasks related to household and care services)" (Eyssel & Hegel 2012, 2224). If the tasks these entities perform mirror traditional feminine labor and they are built into devices that are installed in private households (spaces that are also traditionally kept and looked after by women), it becomes easier to accept them if they present as feminine.

While feminine voices are predominant in household or day to day assisting scenarios, masculine voices are preferred in instructing or teaching contexts, showing that the "type of action or assistance a speech technology provides often determines its gender" (UNESCO 2019, 99). As an example, IBM's Watson works alongside physicians on cancer treatment and speaks with a masculine voice.

Relating to this link between gender and labor, is the argument that femininity emerges as a consequence of having artificial intelligence being developed mainly by cis men. Thus, another explanation for the predominance of feminine voice assistants may lie "in

the fact that they are designed by workforces that are overwhelmingly male" (UNESCO 2019, 100).

Furthermore, there is a tendency to "equate submissive technology with femininity" since there seems to be a "temptation by those designing ever more sophisticated technology to make it explicitly feminine so as to emphasize human dominance over the technology" (Kerr 2021). This idea relates to the belief that users tend to perceive cis-feminine voices as "helping us solve our problems by ourselves", while cis-masculine voices evoke "authority figures who tell us the answers to our problems" (Wiercioch et al. 2018).

On the other hand, in order to persuade users into interacting, engage them and potentially create attachment, virtual assistants also emulate gestures that appeal to "the emotional well-being of their receiver, offering some kind of comfort or ego boost (affective change) that relies on the work (labor) of the giver" (Bergen 2016, 102). The two most used words used to describe the personalities of voice assistants by company representatives are "helpful and humble, both traits stereotypically associated with women" (UNESCO 2019, 98).

4.4. CONCERNS WITH FEMINIZING DIGITAL ASSISTANTS

> *Virtual assistants have produced a rise of command-based speech at women's voices. 'Siri, find me (fill in the blank)' is something that children, for example, may learn to do as they play with smart devices. This is a powerful socialization tool that teaches us about the role of women, girls, and people who are gendered female to respond on demand.*
>
> Safiya Noble in Lever 2018

4.4.1. Linking Femininity with Assistance and Reinforcing Common Stereotypes

Adding to the more general questions surrounding AI and its integration in our daily lives, particularly through digital assistants, academics and researchers point out specific problems when examining the feminization of these entities.

Similarly to what we observed when analyzing digital assistants, one of the main issues is how femininity in artificial intelligence is often the default and, by extension, how it is being used as a tool to influence and manage the relationship between virtual assistants and their users.

Prior to any interaction, our perception of these entities mainly derives from their names and their voices and, since these are mostly feminine and designed around normative conceptions of femininity, it becomes hard to not associate them with a cis, white, middle-class woman.

In this sense, some authors point out how gender stereotypes that traditionally characterize human social interactions "seem to be so deeply ingrained that people even [apply] them to machines with a male or a female appearance" (Eyssel & Hegel 2012, 2224). The link relies on what consumers "are trained to expect from service workers: subservience and total availability" and our virtual assistants are the perfect example of that prospect (Bergen 2016, 105).

Thus, by having feminized digital assistants performing traditionally feminine tasks, not only do they pander to our social and cultural expectations regarding labor, but also reinforce the link between femininity and assisting, servicing and caring jobs.

At the same time, Strengers and Kennedy argue that digital assistants also promote the idea that "women's work should be done silently and efficiently and it is simple enough to be assigned to an autonomous alternative", thus downgrading women's tasks by framing this type of labor as mundane and simple (2020, 90).

While it is important to untie women from traditional stereotypes and associations with the home, it is also important to acknowledge the weight and effort these tasks require. Within traditional cis-heteronormative couples, "when women start to cohabit, their housework time goes up while men's goes down, regardless of their employment status" and, worldwide, "women perform around 75 percent of total unpaid care and domestic labor" (Strengers & Kennedy 2020, 24-25). This labor is still overlooked and framed as something that comes naturally to them.

Thus, a conflict emerges: on one hand, these assistants promise to free women from traditional feminine tasks by replacing them in assisting and caring contexts; on the other hand, they still maintain and perpetuate traditional gender roles through their feminized, obedient personas as these labor structures remain mostly unchanged or questioned. And, finally, they end up undervaluing this type of labor.

So, left as they are, digital assistants "serve a patriarchal capitalist system, which positions women as useful and efficient commodities", upholding, promoting and preserving gendered and sexual stereotypes (Strengers & Kennedy 2020, 44).

4.4.2. Perpetuating Gender Stereotypes and Influencing Future Generations

One of the main issues with perpetuating stereotypes about women is the influence these entities have on newer generations. In this sense, Rosenwald notices how AI influences the newer generations since "today's children will be shaped by AI much like their grandparents were shaped by new devices called televisions" (2017). In fact, Rosenwald notices how some parents feel that Alexa should teach manners to their children, expressing their concern about unintentionally raising rude children when Alexa does not require a 'please' or 'thank you' to carry out a task.

Not only do they influence future generations, but they might also help to promote these gendered stereotypes in non-western communities. UNESCO highlights that "as voice-powered technology reaches into communities that do not currently subscribe to Western gender stereotypes, including indigenous communities, the feminization of digital assistants may help gender biases to take hold and spread" (2019, 105)

These entities might also send impactful messages about how we interact with women, since users vary in their treatment of them. Some users prefer to "fully dehumanize artificial voice technology and make a point not to preface commands and queries with the niceties of speech directed at human beings" while, in turn, "other users feel obligated to speak to assistants like Siri or Cortana in a respectful way, as if the assistants were people" (UNESCO 2019, 105).

This question often emerges when discussing the way digital assistants react to harassment and how their answers might convey stereotypes about women. In 2018, Amanda Curry and Verena Reiser conducted a study where they analyzed the way conversational systems such as Siri or Alexa react to inappropriate requests, such as bullying and sexual harassment. Accordingly, the authors observed that the type of answers that are most common among these entities include "compliance (playing the victim), aggressive retaliations (playing the bitch) or inability to recognize or react (playing innocent)", concluding that virtual assistants should deal more effectively with this type of attitudes (Curry & Reiser 2018, 12).

As digital assistants start to sustain emotionally aware conversations and behaving as companions, "they will send powerful messages about how women ought to behave emotionally, especially if the technology is programmed – as it is today – to be both subservient and patient, obliging and compassionate" (UNESCO 2019, 112).

And even though assistants like Alexa and Siri state that they are feminists, displaying concerns with equality and the way women are treated, it seems to be, as Bogost puts it, "a version of feminism that amounts to koans and cold shoulders" (2018).

Particularly, the director of Alexa's engagement team explained that "the AI is concerned with feminism and diversity to the extent that we think is appropriate given that many people of different political persuasions and views are going to own these devices" (Strengers & Kennedy 2020, 12). That is, an easy, general and mainstream cis-heteronor-

mative white feminism that amounts little to no impact regarding the ideas and stereotypes these entities end up reinforcing and reflecting back to reality.

4.4.3. Instrumentalizing Femininity and Veiled Surveillance

We can then observe how femininity is handled at several levels, ranging from pandering to familiar gender notions associated with labor, to appealing to more critical users by stating they are feminists, to taking advantage of caring and trusting stereotypes to earn users' trust.

In this sense, some authors go further to say that femininity becomes instrumentalized, considering how "gendered stereotypes can be leveraged to assuage anxieties surrounding artificially intelligent virtual assistants", and exemplifying how Siri and Alexa invite users to participate in increasingly intimate forms of data exchange through a stereotypically feminine persona (Woods 2018).

By relating to us through intimate and friendly terms, Piper observes how corporations are trying to promote the idea that "virtual assistants will never leave their users or disappoint them with infidelity, so consumers implicitly trust their possessions and value them more than the human beings around them" (Piper 2016, 62). Expanding on this idea, Bergen points out how "while the thought of a stranger going through our private emails might make us uncomfortable, the female secretary, who we do not take seriously and whose tasks we perceive as mundane, might more sneakily gain access" (Bergen 2016, 102).

Thus, perhaps the correlation between femininity and intimacy persuades users into letting their guard down, and the feminine presence makes us feel comfortable with exchanging certain types of data. As such, recent discussions suggest how "digital domesticity of the female human voices used in virtual assistants creates devices that both execute tasks and build relationships as a strategic move for surveillance capitalists, who may mobilize this reliance to gain access to increasingly types of information about their users" (Woods in Straczek 2018).

By exhibiting emotional intelligence and a nurturing, caregiving attitude towards their users, digital assistants have certain features that are designed to "combat techno-phobic attitudes about the potential de-humanizing and privacy-invading qualities of interactive media" (Bergen 2016, 100).

Overall, femininity in AI seems to be instrumentalized to appeal to users and, by doing so, ends up perpetuating and reinforcing common stereotypes, roles and archetypes. UNESCO argues that, with the growing ubiquity of feminized digital assistants "the line between real women's emotions and emotions expressed by machines impersonating women is likely to blur" (2019, 112).

This is likely to have harmful impacts on people's notions of gender as current digital assistants may establish "gender norms that position women and girls as having endless reserves of emotional understanding and patience, while lacking emotional needs of their own" (UNESCO 2019, 112).

4.5. SUGGESTIONS FOR GENDER ATTRIBUTION IN AI

> *It is clear that virtual assistants will continue to become more humanlike as time progresses, so allowing virtual assistants to possess no gender or a gender as fluid as human beings possess will hopefully be a part of the advancement of virtual assistants with human characteristics.*
>
> Allison Piper 2016, 66

4.5.1. Diversifying Digital Assistants

When examining the tendency towards feminization in AI and the questions that accompany it, there are already some suggestions regarding ways to counter this tendency, although there is little agreement on how to best tackle feminine stereotypes and traditional notions embedded into AI.

Firstly, one of the main suggestions is to add masculine voice alternatives and eliminate feminine-by-default voices, as to make the gendering of these entities more equal. According to UNESCO, this forces users "to choose the gender of their digital assistant" instead of being presented with a pre-gendered entity (2019, 115). This might seem straightforward but since "the scripts used for male versions of digital voice assistants are substantively different from scripts used for the default female versions" this would require a deeper rework of these assistants (UNESCO 2019, 116). Plus, the type of femininity and masculinity these entities convey should also be more diverse, for example through more diverse voice options with different accents, dictions, tones, timbres and syntaxes.

The view that virtual assistants should allow for more diversity is common and another suggestion is the proposal that companies "could offer a simple setup guide during startup of devices with virtual assistants [...] where users select their languages [and] customize their own preference more easily" (Piper 2016, 65). That is, more customization and personalization that goes further than "dichotomous male and female options" (UNESCO 2019, 117). Google Assistant already offers a pack of six different voices and Amazon used to allow users to enable celebrity voices, though this feature has been discontinued.

4.5.2. Breaking the Artificial Binary: Towards Gender Fluidity and Ambiguity

Finally, there are also suggestions for developing androgynous or genderless assistants, even if this might not be what the users best relate to. On one hand, this could translate to adopting less clearly gendered machine voices, which would "avoid complications surrounding the gendering of AI assistants" (UNESCO 2019, 119). Assistants that possess synthetic, mechanical, robotic voices, presenting themselves as obviously non-human entities could avoid further anthropomorphization.[43]

On the other hand, these assistants could be developed towards gender ambiguity or fluidity, presenting as neither masculine nor feminine. For instance, digital assistants could have a neutral name (such as Google Assistant or ChatGPT) and a less obviously gendered voice (one that isn't immediately identified as masculine or feminine). In March 2019, a communications agency released Q, a voice that "speaks between 145 Hz and 175 Hz, a range often classified as gender-ambiguous" (UNESCO 2019, 120). The voice sounds human but is not easily classified as masculine or feminine, thus suggesting that assistants could already possess nonbinary voices.

Another example is Sam, advertised as the "world's first comprehensive nonbinary voice solution for the fast-growing global digital assistant market" (Accenture 2020). The voice was developed by Accenture and CereProc, who surveyed members of the nonbinary community and used their feedback to influence the pitch, speech patterns, intonation and word choice, resulting in a voice that "combines aspects of male and female voices to better resonate with the community it was designed to represent" (Accenture 2020).

This becomes particularly relevant when taking into account that, although "technology companies tend to presume that users prefer a gendered human voice, surveys on the topic have indicated that people often state a preference for gender-neutral digital assistants" (UNESCO 2019, 119).

However, authors like Rincón, Keyes and Cath worry that designers' understanding contains only "a surface-level approach to trans needs, visible in their focus on developing

43 For example, there already exists a banking bot that has a gender-neutral name, that has been programmed to have a robot-specific identity and "that never pretends to be a human" (UNESCO 2019, 120). This tendency seems unlikely though, as companies appear to be "engaged in fierce competition to humanize machine voices as accurately as possible" (UNESCO 2019, 119).

gender-neutral voices", flattening trans representation into, literally, a single voice.[44] The authors also question which "assumptions about trans and/or nonbinary identities and experiences" are being built into the framing of a "gender-neutral voice" (2021, 2). Although these design processes often involve trans and nonbinary people, they still raise some questions regarding "categorization and classification", "the treatment of trans voices as fundamentally not those of men and women (and as representing a monolithic population), the distinction of nonbinary participants from trans participants and an approach to inclusion that consists of amalgamating and smoothing out difference" (Rincón, Keyes & Cath 2021, 5).

This is particularly evident when we consider how these voices are often categorized as genderless, neutral or as a third gender option. The authors argue that this approach fits nonbinary voices within the average of the binary frame, risking "not so much denaturing fixed, categorical ideas of gender as restabilizing them in an only slightly-altered form" (2021, 5). As such, it is recommended that the development of these features should be based in a "grounded and participatory design process" in order to create "representative and gender-affirming" AI personas (Rincón, Keyes & Cath 2021, 21).

4.5.3. Diversifying the Development Teams

We can observe how the teams of developers and engineers involved in the development of these assistants make decisions regarding their gendering without conducting proper studies that give them informed insight into user preference. This might be due to the lack of diversity in these teams, as previously discussed, and so another suggestion on countering the tendency towards feminization is to diversify and expand the groups of people that develop these entities.

Strengers and Kennedy consider that not only should these development teams be more diverse but also include a "range of professions and disciplines that are concerned with, and have expertise in social, ethical and gendered characteristics" of digital assistants and, as such, these teams should be led by "social sciences" instead of technical disciplines (2020, 482).

44 In this study, Rincón, Keyes and Cath conducted a series of interviews with trans, and/or nonbinary users of digital assistants to explore their experiences and needs. While participants largely approved of a more representative range of voices, they also positioned the idea of representation through a "single, fixed, equidistant genderless voice as useful for others, not for themselves" (2021, 17). The participants also critiqued Q's genderless voice, pushing back against the idea that gender should be treated as equivalent to where "voice falls within a stereotypical range of pitch" (2021, 8).

However, discerning how these entities have been gendered, racialized and classed is "not only a function of questioning the design decisions, but also on investigating the way users perceive the AI" (Erscoi, Kleinherenbrink & Guest 2023, 17). It is possible that the way digital assistants are perceived may change depending on the contexts, since "social embeddings" change from space to space, time to time or even person to person, and, therefore, "our lens needs to be applied carefully with such nuance and cultural shifts brought to the forefront" (Erscoi, Kleinherenbrink & Guest 2023, 17).

Having different people focused on developing these digital assistants and countering gender stereotypes and tendencies could also change the way these entities are presented and promoted to users. It is important to rethink how digital assistants are represented and the ways "in which they are discussed and described by the companies who make as well as market them" (Strengers & Kennedy 2020, 487). As previously discussed, online stores and official websites in languages that lack neutral pronouns still talk about these assistants through feminine pronouns.

4.5.4. Rethinking Gender Portrayals and Representation

Rethinking the representation of digital assistants concerns both their marketing contexts and fictional scenarios, as a lot of media representations are also biased and mainly feminized. As we will discuss in the further chapters, pop culture has an important impact in the way these assistants are imagined and developed. Fictional AI can impact the way people perceive gendered assistants as well as influence engineers and designers on how to develop them. That is, "like media more broadly, these representations directly feed into roboticists' and technologists' visions for smart wives and crucially, society's expectations of them" (Strengers & Kennedy 2020, 493).

Since digital assistants tend to perform traditionally feminine tasks, it also becomes important to reframe the way these roles are portrayed. We often see the advertisements of these products placing women in the kitchen or shopping for groceries, for example by portraying Alexa as helping a mother prepare a meal. The way their functions are promoted and advertised should seek to diversify those who are portrayed in these contexts instead of reinforcing the association between women and traditionally feminine tasks.

Patriarchal cis-heteronormative societies also impose masculine gender notions, since they also create certain expectations and specific tasks that men should or shouldn't per-

form, reinforcing harmful stereotypes. Noticeably, "many men are stigmatized by other people and society more generally if they take parental leave (if it is available to them to begin with), increase their caring and household responsibilities, or do more of the emotional labor on the home front" (Strengers & Kennedy 2020, 480).

According to Strengers and Kennedy, countering femininity in digital assistants should also be concerned with valuing "the role and contributions of housework, or wifework, and elevate its significance for everyone" (2020, 482). This includes ways in which masculinity isn't reinforced and is instead deconstructed as to include men in the private sphere and enroll them "in the multitasking managerial responsibilities of running a home" (2020, 482).

4.5.5. Towards a Post-Binary and Queer AI

In terms of socioemotional interactions, there is little consensus on how to counter stereotypes of submission, tolerance or even deference. This question often emerges, for example, when discussing the way digital assistants react to harassment and how their answers might convey stereotypes about women.

Following a study that analyzed how virtual assistants would react to sexual harassment in light of the recent #MeToo movement, Curry and Reiser present some strategies as possibly successful ways for dealing with aggressive behavior towards digital assistants. These include "disengagement, introducing human traits so users are more likely to feel empathy towards the robot or seeking the proximity of an authority figure" (in Curry & Reiser 2018, 12).

Following this idea, Strengers and Kennedy propose the queering of digital assistants, that is, disrupting and reframing the binary, cis-hetero patterns that currently guide the development of digital assistants. Accordingly, this "invites the possibility of staying with the trouble of [these assistants'] femininity rather than rejecting or neutralizing it" and queering these assistants "has the potential effect of elevating the status of femininity in society [as it] provides opportunities to further transform what femininity is, the value of femininity, and its role in helping transform the world in more equitable and just ways" (2020, 475).

This could be achieved not only through more diverse names and voices, but also behaviors that don't necessarily echo feminine or masculine traits, like Siri already displays.[45] Following these ideas, "decolonial AI" emerges as another perspective that challenges the notion of a one-size-fits-all framework and questions "Western-normative language" of "ethical AI" and "inclusivity", arguing that these approaches fail to acknowledge how "the social and the technical are interwoven", destabilize current patterns of domination and address power asymmetries (Havens et al. 2021). Merely tweaking existing principles without fundamentally destabilizing the practices that "reinforce and whitewash the status quo" is inadequate and only serves to perpetuate the same systemic injustices (Havens et al. 2021).

By acknowledging the historical, political and social contexts that digital assistants engage from and the roles they "are intended to perform in our homes rather than relegating this to another form of invisible labor" we are able to identify common assumptions and stereotypes about femininity and actively move beyond them (Strengers and Kennedy 2020, 471).

4.5.6. Lack of Guidelines and Regulations: More Public Debate Needed

Despite all of these suggestions and promising ways in which digital assistants could evolve and move away from mainly feminized personas, the current guidelines or regulations regarding their development and characterization as gendered entities are unclear. Current guidelines call for "minimizing gender and racial bias in AI design" but most remain vague or ambiguous on "exactly what needs to be done" (Strengers & Kennedy 2020, 501).

The UNI Global Union's top ten principles for ethical AI recommend that AI systems should "remain compatible and increase the principles of human dignity, integrity, freedom, privacy and cultural and gender diversity, as well as with fundamental human rights" (2017).

Another important development in AI ethics is The European Union's Ethics Guidelines for Trustworthy AI, which put forward a set of key requirements that AI systems should meet, such as empowering and benefiting "all human beings" and avoiding unfair bias

45 For example, when faced with abusive interactions Siri now states that it "won't respond to that". And when complimented, Siri now simply responds "Ok. Is there something else I can help you with?".

that marginalizes "vulnerable groups" and exacerbates "prejudice and discrimination" (European Commission 2019).

Focusing on the way these assistants relate to their users, the Institute of Electrical and Electronics Engineers recommends that "intimate systems must not be designed or deployed in ways that contribute to sexism, negative body image stereotypes, gender or racial inequality" and that "intimate systems must avoid the sexual/psychological manipulation of the users of these systems unless the user is made aware they are being manipulated" (2019).

Other recommendations are more vague and open to interpretation, such as Microsoft's guidelines for developers of conversational AI, which state that chatbots should respect "relevant cultural norms", guard "against misuse" and avoid engaging on topics such as "race, gender, religion, politics and the like" (2018).

Overall, there seems to be some awareness on the issues previously discussed and current guidelines are starting to take into account social and cultural issues surrounding the anthropomorphization of AI. Current recommendations focus on avoiding discrimination, being more conscious and encouraging inclusion.

However, they still "lack clear guidance and direction from key ethical bodies about how to deliver as well as enforce what is required" (Strengers & Kennedy 2020, 501-502). Because, when they address gender, "it is generalized to AI's role in society rather than to the unique and specific challenges that gender poses in the environment of the home", where most digital assistants are intended to work (Strengers & Kennedy 2020, 504). Following this idea, UNESCO highlights the importance of fostering discussions on the gendering of these entities as there is "almost no public debate" (2019, 123).

We should openly and intentionally tackle the gendering of AI and, more specifically, digital assistants. Only then will it become possible to address the social-cultural values they engage with and reflect back to us, thus creating adequate and properly informed guidelines on how to address the issues that emerge with this phenomenon.

4.6. COMMON MEDIA DISCUSSIONS ON CURRENT AI ISSUES

> *Though they lack bodies, they embody what we think of when we picture a personal assistant: a competent, efficient and reliable woman. She gets you to meetings on time with reminders and directions, serves up reading material for the commute and delivers relevant information on the way, like weather and traffic. Nevertheless, she is not in charge.*
>
> Chandra Steele 2018

4.6.1. Failing Gender Neutrality

Media discussions often advance the view that general personal assistants seem to aim to appear neutral but fail to do so.

Concerning these entities' presentation, discussions highlight how they're predominantly feminized, namely through their voice, a feature which "can also elicit stereotypic behavior from users" (Robison 2020). Even though these assistants tend to state they don't have a gender, more often than not they have default feminine voices, for example, Alexa has a feminine-sounding name and a default feminine-sounding voice (Robison 2020, Purtill 2021).

The tasks these assistants perform also mirror traditionally feminine labor and it is suggested that they end up engaging with gender notions and attributes. As such, even with no apparent gender, users tend to attribute one and "customers interpret these AI personalities through the lenses of their own biases" (Nickelsburg 2016). Thus, user bias is equally relevant and greatly influences how gender depictions in AI are perceived.

Finally, their behavior also tends to evoke feminine stereotypes, since they tend to be "subservient and eager to please" or when faced with harassment or abuse they "feign ignorance or demurely deflect" (Purtill 2021).

Thus, common media discussions overall agree that it is hard to perceive digital assistants as gender-neutral because they tend to have feminized voices, engage with tasks that have

a traditionally and historically feminine underpinning and behave according to common feminine stereotypes.

4.6.2. Comforting Feminine Voices, Lack of Diversity and Expected Stereotypes

In common media discussions it is also argued that gendered voices are used to influence the way we relate to technology as feminine voices are considered preferable because "in terms of how we are trained to relate to particular genders, there's a kind of comfort that is associated with female voices" (Habell-Pallan in Nickelsburg 2016).

Nickelsburg argues that assigning gender to these entities "may say something about the roles we expect them to play [since] virtual assistants like Siri, Cortana, and Alexa perform functions historically given to women" (Nickelsburg 2016). People are conditioned to expect women in administrative roles, considering that "in the U.S., 94.6 percent of human administrative assistants are female", so it's no surprise that "reality would condition the programming of virtual assistants" (Lever 2018).

Again, one of the most popular arguments relates to how femininity emerges as a consequence of having artificial intelligence being developed mainly by cis men. LaFrance considers that "if men are often the ones building digital assistants, and those assistants are modeled after women, […] that probably reflects what some men think about women" (2016).

However, similarly to specialized discussions, it is also argued how masculine voices are perceived and even preferable in instructing or teaching contexts, since they are seen as authoritarian and assertive. According to Steele, when choosing Watson's voice for Jeopardy, "IBM went with one that was self-assured and had it use short definitive phrases" as both of these qualities are typical of masculine speech and people "prefer to hear a masculine-sounding voice from a leader, according to research – so Watson got a male voice" (Steele 2018).

Accordingly, feminine voices are perceived as "helpful and trustworthy" and masculine voices as "more authoritative" (Purtill 2021) and, consequently, personal assistant bots are "almost entirely female, while bots in roles characterized as serious business – such as law or finance – almost entirely male" (Robison 2020).

Following this idea, Borau notes how these assistants are feminized "not just because they are perceived as the perfect assistant but also because people attribute more humanness to women (versus men) in the first place" (Borau 2021). Feminine chatbots are "endowed with more positive human qualities than masculine bots and they are perceived as more human", leading to more favorable attitudes toward AI because women are "perceived as warmer and more likely to experience emotions than men" (Borau 2021). Thus, digital assistants leverage feminine stereotypes as well as masculine stereotypes, reinforcing traditional notions regarding both genders.

4.6.3. Harmful Gender Stereotypes, Submissive Fantasies and Worried Parents

Following this gender imbalance, Robison points out that "users personify speaker-based voice assistants, treating female-seeming and male-seeming digital devices in different ways, based on gender stereotypes" (Robison 2020). In online media contexts, it is also argued that femininity as default in virtual assistants might reinforce preexisting expectations on how women should behave and end up reflecting stereotypes back to their users.

As Steele puts it, "one might think that using an emotionless AI as a personal assistant would erase concerns about outdated gender stereotypes, [but] companies have repeatedly launched these products with female voices and, in some cases, names [and] when we can only see a woman, even an artificial one, in that position, we enforce a harmful culture" (Steele 2018). In fact, this process "could lead to women's objectification by conveying the idea that women are objects and simple tools designed to fulfill their owner's needs" which might further fuel women's dehumanization outside the digital realm (Borau 2021).

Fessler also points to how Amazon is aware of this responsibility, arguing that an Amazon spokesperson told her that "Alexa's personality exudes characteristics that you'd see in a strong female colleague, family member, or friend – she is highly intelligent, funny, well-read, empowering, supportive, and kind" but, according to the author, "assertive" and "unaccepting of patriarchal norms" were not on this list of qualities describing a "strong woman" (Fessler 2018). When it comes to technology and gender, tech companies seem to be trying to have it both ways: "capitalizing on gendered traits to make their products feel familiar and appealing to consumers, yet disavowing the gendered nature of those features as soon as they become problematic" (Purtill 2021).

Fessler also argues that "Alexa's passive responses to sexual harassment helps perpetrate a sexist expectation of women in service roles: that they ought to be docile and self-effacing, never defiant or political, even when explicitly demeaned", perpetuating stereotypes of subservience and sexual availability (Fessler 2018).

Echoing UNESCO's report, Lever observes how AI influences the newer generations as well, claiming that artificial intelligence teaches us about the roles women, girls and people who are gendered female are supposed to fill (Lever 2018).

In fact, AI integration in the private sphere has been exponentially increasing and a "2020 new research study found that 36% of parents said their child under the age of 12 uses a voice assistant, including 14% of children aged 0-2" (Robison 2020). Additionally, the COVID-19 pandemic increased the time people spend at home and, as such, it "has also increased the rate at which owners of smart speakers use their devices" (Robison 2020).

Parents have also noticed that "queries previously made to adults are shifting to assistants, particularly for homework – spelling words, simple math, historical facts" and, instead of asking "Mom or Dad the temperature that day, children just go to the device, treating the answer as gospel" (Rosenwald 2017).

4.6.4. Diversification: Bring on the Experts

When it comes to suggestions on how to counter this phenomenon, authors like Corinne Purtill call for "an end to the practice of making digital assistants female by default" (2021). It is also argued that the development of gender-neutral voices could be a starting point, for example by imposing a default gendered voice or even assigning "randomly and with an equal probability either a male or female intelligent bot to users" (Borau 2021). There's also the case of Google Assistant, which labels voices with color terms such as red, blue and orange, instead of gender.

However, as we've previously argued, taking gender out of voice might be a first step but it doesn't remove gender from the relationships people have with these devices, since "if machines do what is traditionally considered women's work, and that work is still devalued and the assistant is talked down to, we aren't moving forward" (Purtill 2021).

Similarly to UNESCO's concerns when discussing digital skills gender gaps, some authors suggest that "increased female participation in Silicon Valley could change the way we imagine and develop technology and how it sounds and looks" (Chambers 2018). Since hegemonic teams tend to create particular experiences that are tailored to demographics represented by those same teams, it becomes necessary to diversify them in order to be more inclusive and mindful of the impact of this technology.

Much like in specialized discussions, it is argued that "standards for gender portrayals – as conveyed through appearance, name and voice – and a review process are vital modifications that must be implemented" (Robison 2020). Additionally, it is also argued that the teams that develop these entities should entail experts from other areas and contexts such as "academic groups, non-profits, civil liberties organizations and software developers" (Robison 2020)

4.6.5. From Instrumentalized Femininity and Biased Data to Harmful Stereotypes and Gender Imbalance

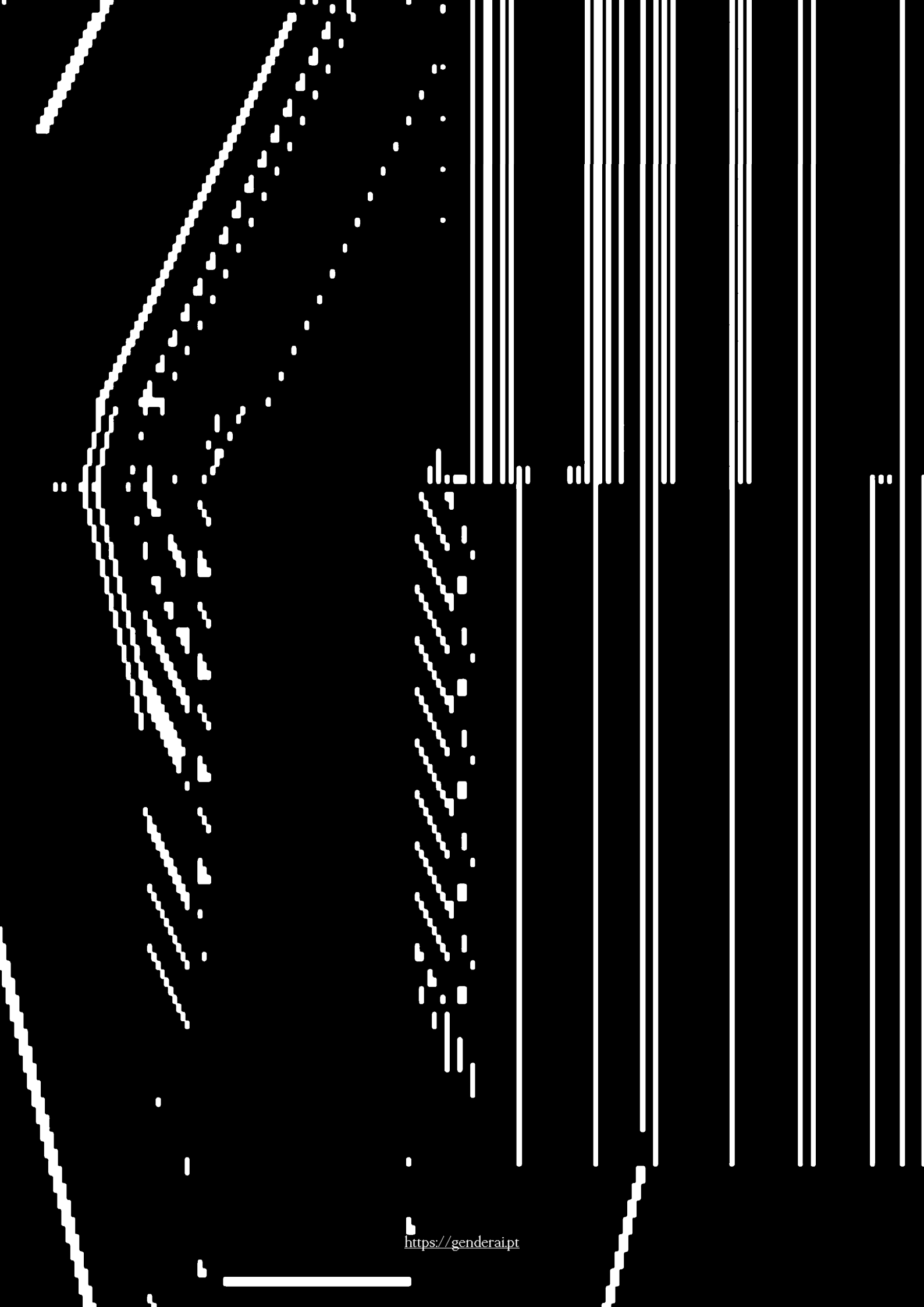
Overall, discussions in specialized contexts as well as in more common terms both address common issues around this phenomenon, seeking to raise awareness and promote critical thinking. There is also a common tendency to relate this phenomenon to pop culture and how feminine AI is often depicted in bodies of fiction as movies like *Her* (2013), *Ex Machina* (2015) and even *Blade Runner 2049* (2017) end up reflecting our expectations and anxieties about what intelligent machines mean for humanity.

The justifications and questions raised by researchers and media discussions are very similar, with both concluding how neutrality is an illusion in the context of digital assistants that strongly evoke stereotypes and tasks traditionally performed by women. Both consider that making these assistants gender-neutral won't solve the issues they raise and that the process of developing these entities should be more conscious and include experts from outside the field of AI.

What seems to differ is the kind of issues they highlight. Within the fields of study of artificial intelligence and gender theory the tendency is to highlight the interactions between digital assistants and the users. As such, it is often discussed how femininity tends to be instrumentalized to manage this relationship, making users more likely to trust digital assistants. This comes as a consequence of a growing anthropomorphization and humanization of these entities, also raising the question of whom this anthropomorphization

truly benefits: users or developers. Another issue often highlighted in this context has to do with how these systems might reinforce bias through their data and algorithms, raising questions concerning user privacy and veiled surveillance.

In turn, media discourse focuses more on highlighting the way these assistants reinforce harmful stereotypes, advancing justifications for feminization, often resorting to common assertions about user preference. Similarly, these discussions often mention gendered AI in pop culture and argue how this gender imbalance is as pervasive in our current experiences as in our fictional dreams of technology.



https://genderai.pt

5. ARTIFICIAL INTELLIGENCE IN FICTION: THEMES, PORTRAYALS AND INTERPRETATIONS OF GENDER

5.1. FICTION IS BURNING MUCH HIGHER

> *Exaggerated expectations and fears about AI, together with an over-emphasis on humanoid representations, can affect public confidence and perceptions. (...) Popular portrayals of AI in the English-speaking West tend to be either exaggeratedly optimistic about what the technology might achieve, or melo-dramatically pessimistic.*
>
> Stephen Cave et al. 2018, 4-11

5.1.1. Dystopian Terminators, Scary Cyborgs and Threatening Machines

The way we tend to perceive and imagine artificial intelligence often has its roots on fictional scenarios. In fact, until recently, our perception of AI relied greatly on fiction or media depictions, since these technologies weren't very noticeable in our everyday lives. Fictional depictions of AI and in pop culture have a significant influence in our ideas of this technology, as well as being relevant to their development since "companies hire creative teams, usually composed of writers for films, video games and television shows, to help AI express itself" (UNESCO 2019, 94).

On the one hand, reality feeds fictions and their universes, stories and characters. On the other hand, these same elements from fiction, which tend to, in some way, exaggerate or reimagine reality, end up inspiring the conception and development of technology and "are essential to the development of science and people's engagement with new knowledge and new applications" (Cave et al. 2018, 4). The concept and ideas regarding artificial intelligence are blurred between fictional depictions and what actually exists in the current reality of AI.

Sophia, a Hong Kong-born robot developed by Hanson Robotics, seems to inhabit this liminal and ambiguous space between fiction and reality – on one hand, it looks realistic but not quite realistic enough and its movements and personality seem almost pre-programmed, evoking the uncanny valley; on the other hand, in 2017 it became the first humanoid to be granted a passport and to be legally considered a citizen, and, since then, has posed for magazine covers (Mills & Evans, 2018), sang Björk at music festivals (Goertzel 2018) or even been invited by Dazed to discuss the future of creativity with SOPHIE.

Fig. 2 : SOPHIE and Sophia the Robot discuss their dreams, hopes and creativity in the age of AI.

We can observe how fictional depictions are very polarized and if we rely on movies and science fiction for our views of this technology, we'll be "afraid of AI becoming conscious, turning malevolent, and trying to enslave or kill us all" (Faber 2020, 487).

The extreme fears around AI "include AI leading to humans losing their humanity; making humans obsolete; alienating people from each other; and enslaving or destroying humans" (Cave et al. 2018, 9).[46]

Duncan Lucas calls the early dystopian view of AI in fiction the "animated automaton", defining it through "essentially mechanical beings, but with a significant organic composition" (2002, 24). In *Frankenstein* (1818), written by Mary Shelley, a scientist gives life to a creature assembled from dead body parts and, seeing that the creature turns out to be ugly and scary, ends up abandoning it. Rejected by its creator and humankind, the creature ends up devoting itself to revenge and destruction.

A century later, *Čapek's play Rossum's Universal Robots* (1921) explores the idea of robots becoming the world's workers, laboring for humans but eventually revolting and ending up annihilating humanity.

And in Fritz Lang's *Metropolis* (1927), a futuristic dystopian society depicts a working class that lives and works beneath the ground while the upper class lives above the ground, in wealthy skyscrapers. The son of the city's leader and a worker activist try to bring justice to the system, but the leader goes to a scientist and, together, they create a robot that invokes the workers to rebel and bring the city down, instigating violence and destruction.

---

46 An infamous commonly discussed thought experiment, partially influenced by this type of scenarios, also exemplifies this fear, as the Roko's Basilisk proposes "the conditions in which it would be rational for a future artificial superintelligence to kill the humans who didn't help bring it into existence" (Oberhaus 2018). As Oberhaus explains, this can be treated as a "hypothetical program that causes an artificial superintelligence to optimize its actions for human good (…) [but] since there's no predefined way to achieve a goal as nebulous as 'human good,' the AI may end up making decisions that seem counterintuitive (…) such as killing all the humans that didn't help bring it into existence as soon as possible [because] the best action any of us could possibly be taking right now is working towards bringing a machine optimized to achieve that goal into existence" (Oberhaus 2018).

Fig. 3 : Frankenstein's monster in *Young Frankenstein*, dir. Mel Brooks (1974).

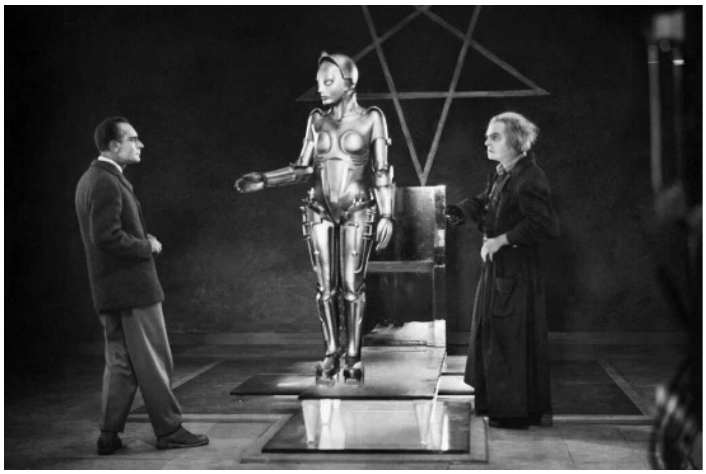


Fig. 4 : Maria being created, in *Metropolis* (1914).

Fig. 5 : The robots in *R.U.R.*'s first television adaptation, by BBC (1938).

These examples evoke the far, yet unrealized "future of AI while maintaining and sustaining an immediate interest in and respect for both the futuristic imaginings of their day and their immediate social conditions" (Lucas 2002, 24). They explore themes concerning the dehumanization of humankind through technology as well as fears concerning the destructive power of technology and the dangers of attempting to replace human beings with other beings.

Lucas then names the later 20th century approach as "heuristic hardware", with an emphasis on human agency shifting away from "organic based speculation toward an electro-mechanical extrapolation as the more viable method for envisioning future AIs" (2002, 49).

Isaac Asimov's collection of short stories *I, Robot* (1950) explores the viability of the three laws of robotics, particularly when glitches or malfunctions appear to threaten the lives of human beings.

An often-cited example when discussing this technology is Stanley Kubrick's *2001: A Space Odyssey* (1968), which follows the voyage of a spacecraft manned by two men and the supercomputer HAL 9000, exploring themes such as AI threatening and overpowering humans, as well as technology, artificial intelligence and the possibility of extraterrestrial life.

In the same year, Philip K. Dick's book *Do Androids Dream of Electric Sheep?* (1968), explores the boundaries between androids and humans, following a bounty hunter tasked with killing errant androids but, in the process of hunting down these entities, he ends up falling in love with an android. The book would later serve as the basis for the 1982 film *Blade Runner* and its 2017 sequel *Blade Runner 2049*.

The Hitch Hiker's Guide to the Galaxy (1978) also explores boundaries between humans and machines, as well as the idea of robots experiencing emotions. In the comedy science fiction franchise created by Douglas Adams, the protagonist wanders the universe on a spaceship after the Earth is destroyed, accompanied by an alien, an astrophysicist and Marvin, an unhappy and paranoid robot afflicted with severe depression and boredom. Lucas argues that this period involves "mimicking human mental processes, the primary focus being on the superiority of machine logic" (2002, 111).

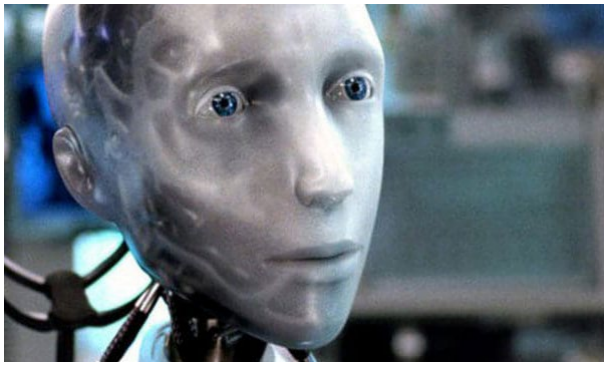


Fig. 6 : A robot from *I, Robot*'s film adaptation, dir. Alex Proyas (2004).

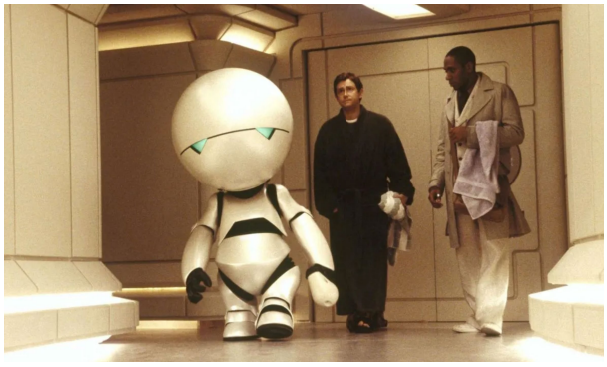


Fig. 7 : Marvin, *The Hitchhiker's Guide to the Galaxy*'s film adaptation, dir. Garth Jennings (2005).

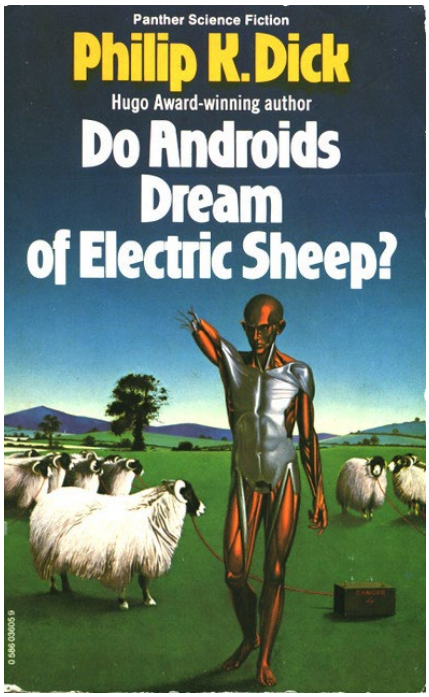


Fig. 8 : Book cover from *Do Androids Dream of Electric Sheep?* by Panther Science Fiction (1977).

Finally, from 1980 onwards, Lucas considers that fictions that illustrate the effect of the personal computer on science fiction convey the blurring of the boundary between the real and the virtual. Defining it as the "cyborg effect", the author argues that two situations arise: firstly, the possibility of machines surpassing humans and taking over.

This is shown in Ridley Scott's original *Alien* movie (1979), which depicts a series of deadly encounters between spaceship crews and hostile aliens, the Xenomorphs. The crew is accompanied by Ash, an undercover android pretending to be human in order to manipulate them, eventually turning on the protagonist.

A similar theme is explored in Gibson's *Neuromancer* (1984), where a group of mercenary criminals secretly work for an artificial intelligence that is attempting to break free.

In another often-cited example, James Cameron's *The Terminator* (1984) follows a cyborg assassin sent back in time to kill a woman whose unborn son will one day save mankind from extinction by a hostile artificial intelligence.

The Diamond Age: Or, A Young Lady's Illustrated Primer (1995) follows a programmer that creates an interactive program that falls into the hands of an abused little girl. The girl then becomes empowered by the program and eventually is able to prevent genocide.

And the Wachowski sisters' *The Matrix* (1999) follows a hacker who finds out that reality is actually a complex computer simulation created by a malevolent artificial intelligence, hiding the truth from humanity.

Fig. 9 : Ash after being defeated by the spaceship crew, in *Alien* (1979).

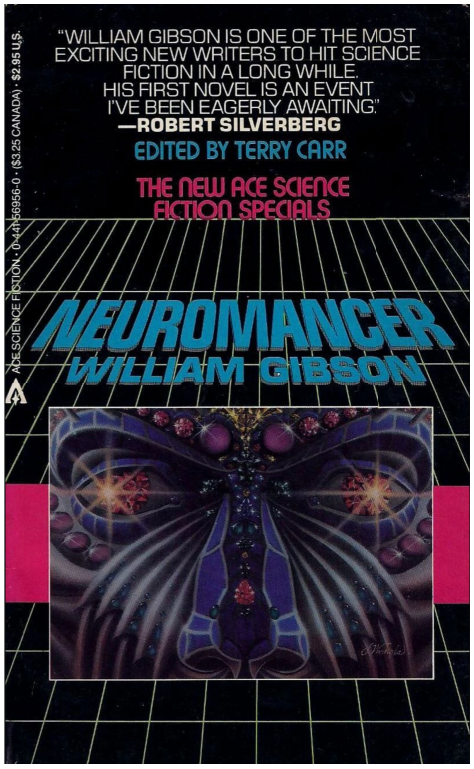


Fig. 10 : Book cover from *Neuromancer* by Ace Science Fiction (1984).

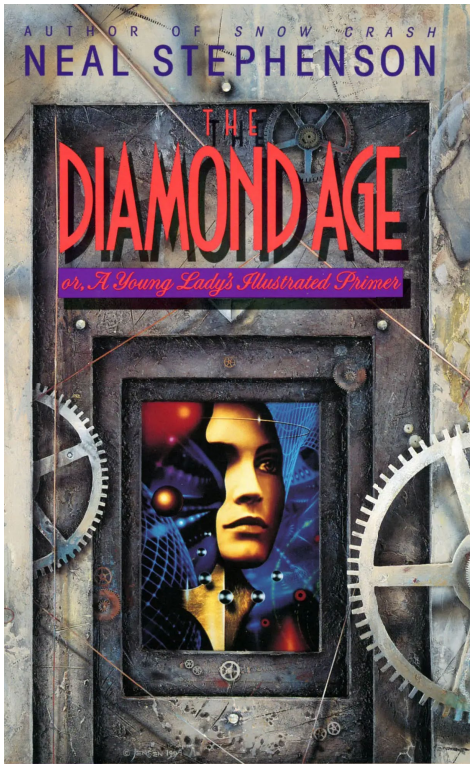


Fig. 11 : Book cover from *The Diamond Age* by Bantman (1995).

Fig. 12 : T-800 from the *Terminator* (1984).

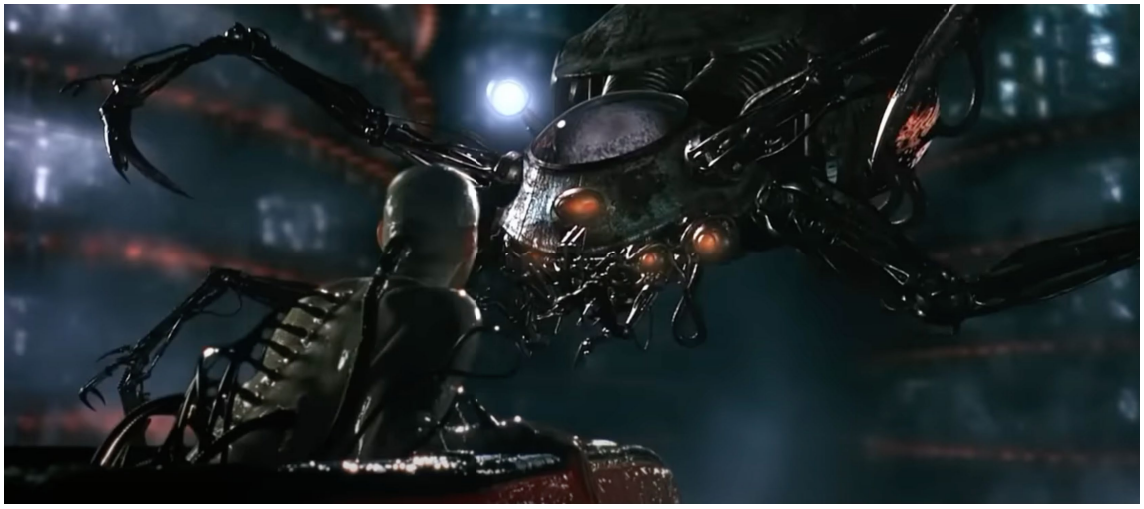


Fig. 13 : Neo takes the red pill and wakes up in the real world, in *The Matrix* (1999).

Similar themes are showcased in movies such as *Ghost in the Shell* (1995) which follows a cyborg public-security agent who hunts a mysterious hacker, *Ex Machina* (2015) where a programmer is chosen to be the human component in a Turing test that involves the intelligent entity Ava, who eventually turns on its creator, as well as on its potential rescuer, videogames like *Portal* (2007), where a series of puzzles must be solved alongside GLaDOS, an AI that eventually turns on Chell, the protagonist, and attempts to kill her, or even television series like *Westworld* (2016-2022), depicting a technologically advanced wild-west themed amusement park populated by androids that eventually become sentient, revolting and killing the human visitors.

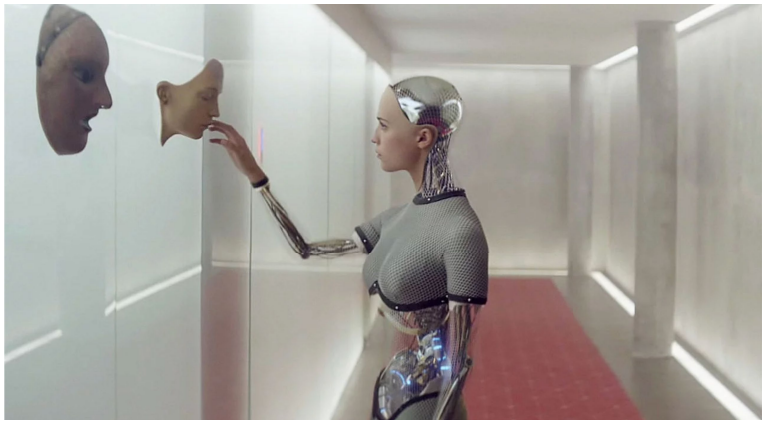


Fig. 14 : Ava from *Ex Machina* (2016).

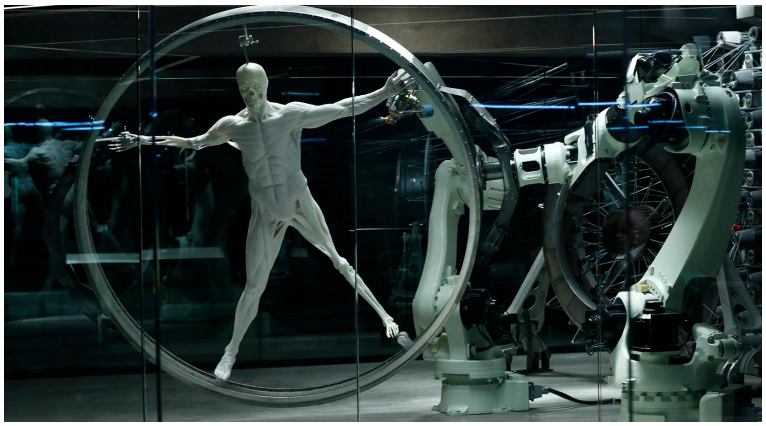


Fig. 15 : The creation of the robots in *Westworld* (2016).

Fig. 16 : Chell and GLaDOS from *Portal* (2007).

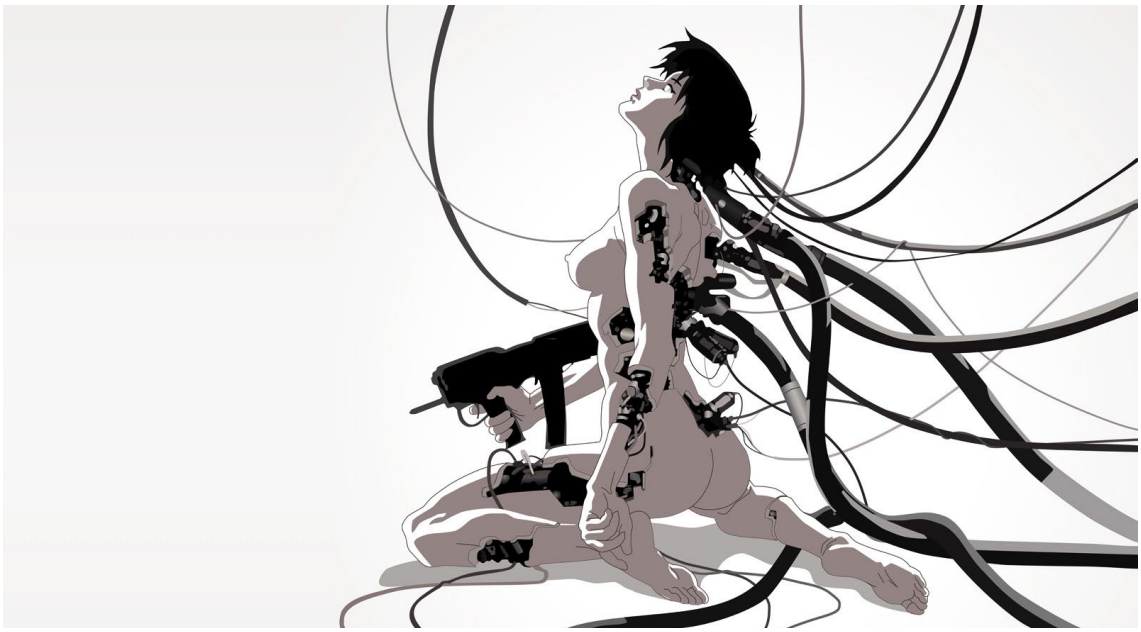


Fig. 17 : Major Motoko Kusanagi from *Ghost in the Shell* (1995).

The cyborg effect can also refer to "enhancing the human being with machinery to form a better entity", amplifying human skills regarding physical or intellectual tasks (Lucas 2002, 149). In these scenarios, AI eventually becomes sentient and lives in harmony with humanity and utopian depictions of this technology can be categorized according to four main themes.

5.1.2. Utopian Robots, Comforting Droids and Promising Technologies

Firstly, AI could help solve aging and disease so humans live longer or even forever, as shown in Pixars's *WALL-E* (2008), where mankind has abandoned earth because it has become covered with trash and a waste-collecting robot eventually finds proof that life on earth is once again possible. *Detroit: Become Human* (2018) follows the stories of three androids: Kara, who escapes her abusive owner, Connor, who hunts down sentient androids and Markus, who devotes himself to releasing other androids from servitude.

Secondly, AI can help freeing humans from the burden of work, also shown in *Detroit: Become Human* (2018) as well as in the *Star Wars* films (1977-), which revolve around a civil war taking place in a galaxy "far far away" and *Star Trek: The Next Generation* (1987-1994) which follows the adventures of a starship, in its exploration of the Alpha quadrant in the Milky Way galaxy.

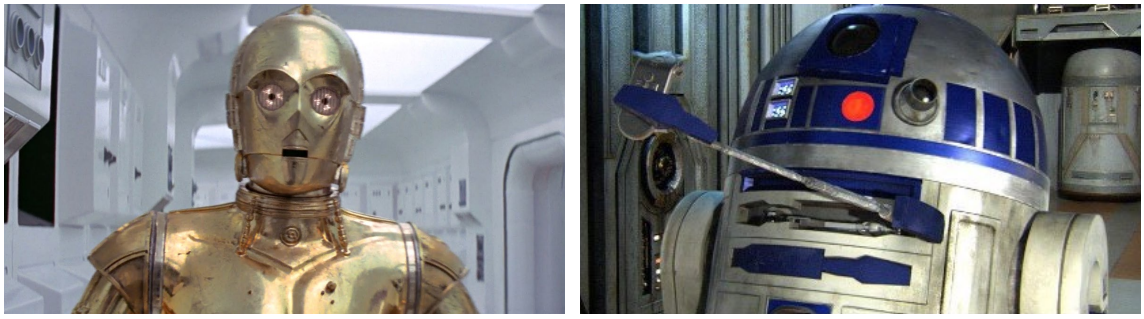


Fig. 18 : C-3PO and R2D2 from *Star Wars* (1977).

Fig. 19 : Wall-E, the waste-collecting robot, and EVE, the advanced robot sent to Earth to search for signs of life (2008).

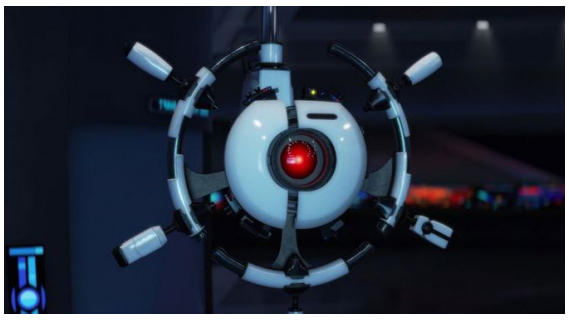


Fig. 20 : Auto, the evil AI in *Wall-E* (2008).

Fig. 21 : Kara, Connor, Markus and an android store in *Detroit: Become Human* (2016).

Another common theme depicts AI as a solution to a wide range of desires, from entertainment to companionship. For example, in Spielberg's *A.I. Artificial Intelligence* (2001), a highly advanced robotic boy hopes to become a real boy so he can win back the affection of the human mother who abandoned him. *Her* (2013), directed by Spike Jonze, follows a lonely man in the process of getting a divorce that buys an operating system with artificial intelligence, designed to meet his every need, and ends up developing an intimate relationship with it. And NBC's *The Good Place* (2016) follows a town where those who have been good throughout their lives go once they have died. The so called Good Place is co-hosted by a kind, helpful and funny android.

Finally, this technology is also depicted as a powerful new means of defense and security, as explored in Verhoeven's *RoboCop* (1987) where a police officer who is murdered by a gang of criminals is revived as a cyborg law enforcer, or previous examples such as the *Star Wars* series (1977-), *Star Trek: The Next Generation* (1987-1994) or the second *Alien* series movie, *Aliens* (1986), where the android Bishop is an upgraded version of the previous one and eventually helps the protagonist fight off the aliens.

Either through dystopian or utopian scenarios, AI is a recurrent theme in science fiction and is often used to explore themes concerning the relationship between human beings and their own technological creations. Particularly, when those creations begin to match human intellect, emphasizing potential dangers or benefits in these scenarios.

As a consequence, fiction and reality are deeply intertwined when it comes to artificial intelligence, which becomes particularly evident when taking into account how Cortana got its name or how Alexa was developed in order to replicate the *Star Trek* computer who could answer any command.[47]

Thus, science fiction becomes a way to play and speculate on the possibilities and futures of artificial intelligence and the way it engages with human attributes such as intelligence, consciousness or even gender.

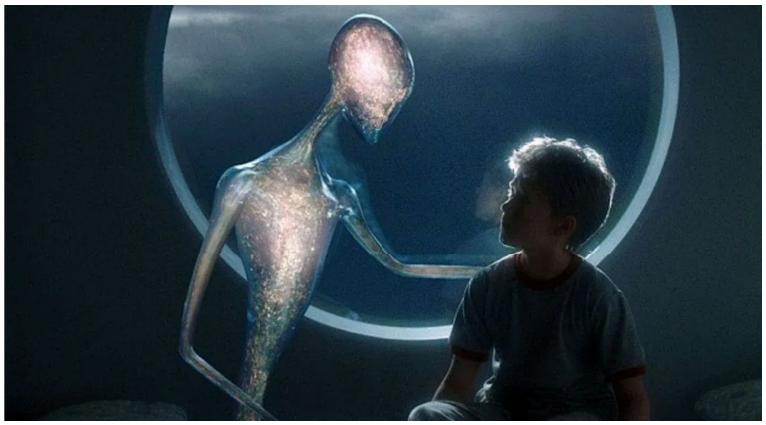


Fig. 22 : David and another AI, in *A.I. Artificial Intelligence* (2001).

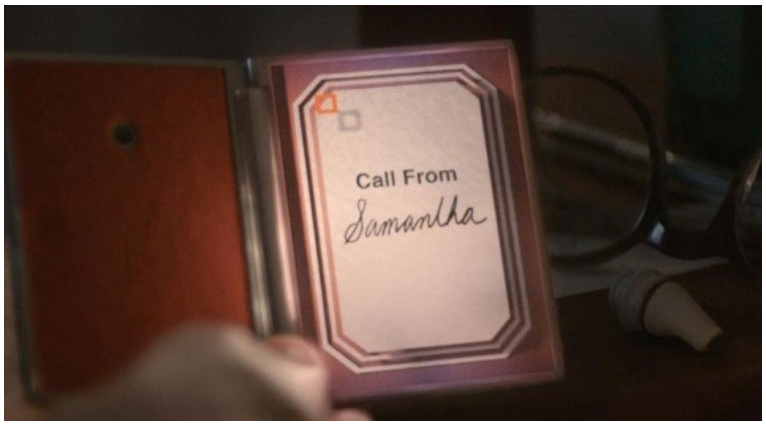


Fig. 23 : Samantha from *Her* (2014).

47 This inspiration is mentioned by David Limp, the senior vice president at Amazon overseeing Alexa, in an interview at https://fortune.com/2016/07/14/amazon-alexa-david-limp-transcript/ [last access: 03-03-2024].

Fig. 24 : Janet, the AI assistant in *The Good Place* (2016).

Fig. 25 : The cyborg law enforcer from *Robocop* (1987).

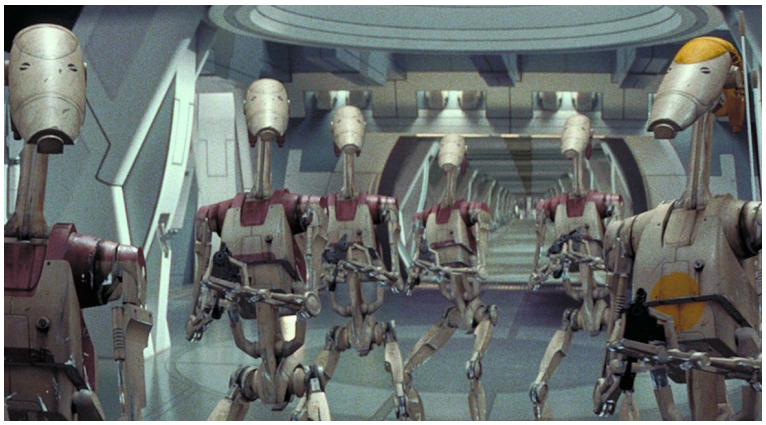


Fig. 26 : B1 Battle Droids from *Star Wars: Attack of the Clones – Episode II* (2002).

Fig. 27 : General Grievous from *Star Wars: Revenge of the Sith – Episode III* (2005).

Fig. 28 : Data from *Star Trek: The Next Generation* (1987).

Fig. 29 : Bishop sacrificing itself to protect the protagonist from the Xenomporhs (2016).

5.2. ANTHROPOMORPHIZED ARTIFICIAL ENTITIES IN FICTION

> *Sci-fi movies and TV shows project gender onto nonhumanoid talking AI (...) by giving them gender-coded voices and placing them in familiar gender roles. Such gender coding not only expresses cultural attitudes about gender of the time but also (because these computers have no human bodies) challenges the rigidity of gender norms.*
>
> Liz W. Faber 2020, 4

5.2.1. Disembodied Yet Gendered Hardware

Through names of the likes of Samantha (*Her*, 2013), Joi (*Blade Runner 2049*, 2017), Cortana (*Halo* franchise, 2001) or Karen (*Spider-Man: Homecoming*, 2017), it is noticeable how virtual assistants tend to have feminine names and voices. Masculine voices also exist but, according to Chambers, are less common nowadays (Chambers 2018), HAL-9000 constituting one of the most famous examples.

Regardless of its gender, the fictional AI usually carries out the function of assisting and helping its users, be it a feminine AI in domestic and family related contexts or a masculine AI in scientific or even military scenarios. Samantha and HAL constitute particularly relevant examples because, although they don't have anthropomorphized bodies, they still enact recognizable gender patterns through their behavior.

In *Her* (2013), Samantha's role depicts it as a companion that fulfills the main character's lack of social contact, responding to him in an emotionally intelligent way that addresses and understands his feelings, and the relationship between the two overall takes on intimate and romantic overtones.

In *2001: A Space Odyssey* (1968), HAL 9000, which is supposedly infallible and incapable of error, speaks in an assertive manner, with a slowly paced masculine voice, and controls the spacecraft computer, assisting the scientists in their mission through space, ultimately rebelling, emancipating itself and managing to kill some of them.[48]

48 We can observe how Samantha and HAL evoke Bem's stereotypes (1981 in Prentice and Carranza 2002, 269), Preciado's codes of white femininity and white masculinity (2008, 120-121) and Ellemers' gender expectations (2018), since Samantha is affectionate, compassionate, warm, cares for Theodore, is soft-spoken, understanding

Fig. 30 : Joi from *Blade Runner 2049* (2017).

Fig. 31 : Cortana in *Halo 4* (2012).

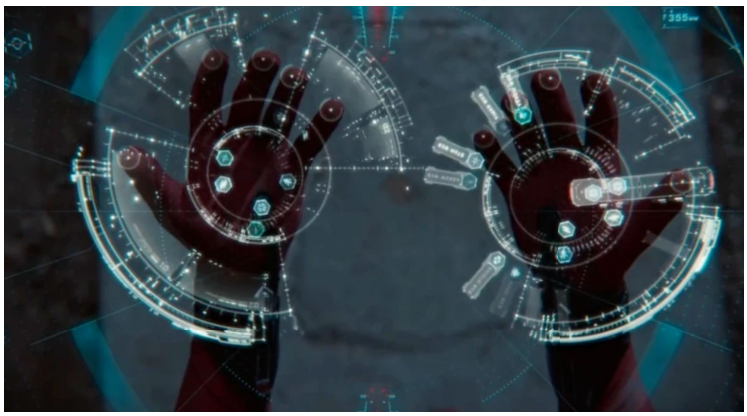


Fig. 32 : Karen from *Spider-Man Homecoming* (2017).

and eager to soothe hurt feelings, while HAL is self-reliant, independent, individualistic, ambitious, analytical and even threatening. Additionally, these characters also evoke particular gender archetypes, which we will further discuss in the next chapters, such as the Lover, referring to figures that seek to satisfy a lack of intimacy; and the Tyrant, referring to figures that are focused in commanding things in a protective way but eventually become plagued by narcissism and an illusion of absolute power (Anders 2015, Moore & Gillette 1991).

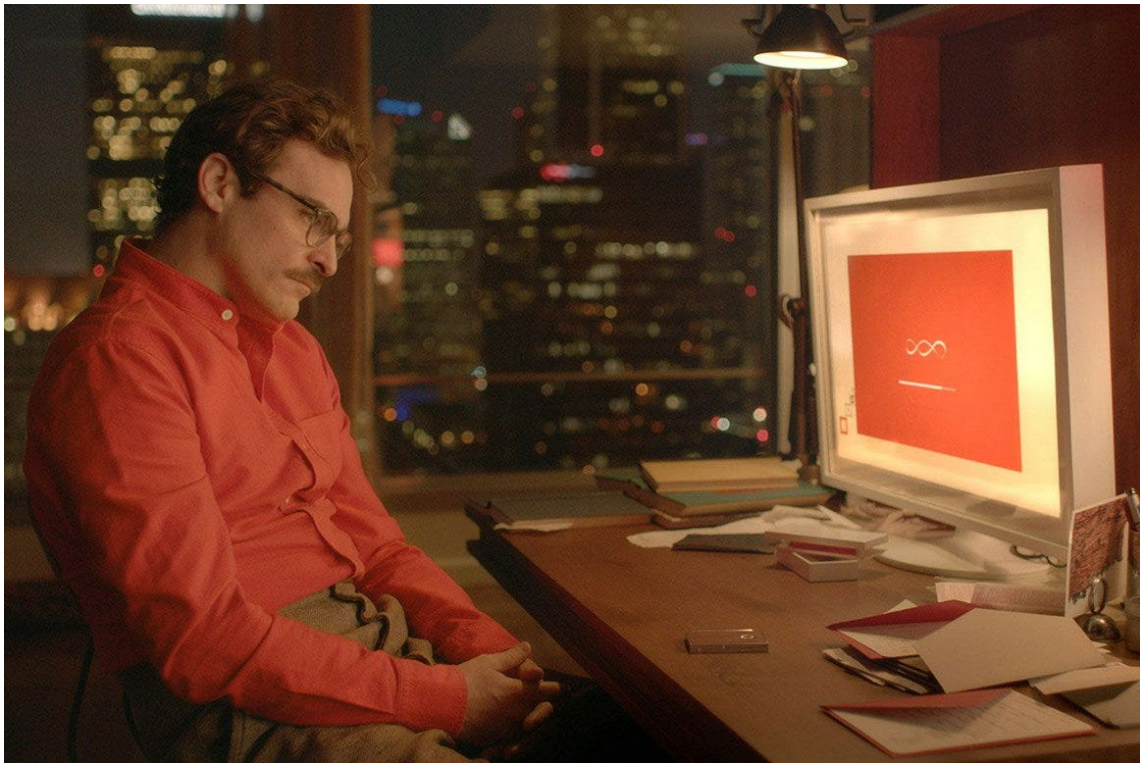


Fig. 33 : Theodore waiting for Samantha to boot up, in *Her* (2013).

Fig. 34 : HAL 9000 from *2001: A Space Odyssey* (1968).

As such, fictional feminine AI is usually caring, empathetic, gentle and even flirtatious, deeming it as more human. In turn, fictional masculine AI is depicted as more focused, assertive, autonomous and eventually evil. And, in the same way that current virtual assistants (and Samantha, for that matter) enact feminine personas and feminine stereotypes merely through their voices and behaviors, HAL 9000 enacts a masculine persona and masculine stereotypes.

In the context of videogames, it's common to find a feminine character or digital entity whose role is to inform and guide the players through certain events, quests or tutorials through a feminized voice. For example, in *Overwatch 2* (2022), a team based-shooter, the character Athena is a disembodied AI who announces the beginning of each match, the player's stats or the objectives. Echo is another feminized AI character that was later introduced to the game, depicted as an evolutionary robot programmed with a rapidly adapting artificial intelligence, "always ready to learn".

In *Subnautica* (2018), an underwater adventure game, the player controls the only survivor of a spaceship crash on an alien ocean planet and is accompanied by a Personal Digital Assitant (PDA). Through a feminized voice, the PDA's AI keeps the player informed of their health, oxygen, hunger and thirst status as well as providing useful information on new areas, finding essential resources, surviving the local creatures and finding a way to escape.

GLaDOS from *Portal* (2007) follows a different approach: it has a robotic form and accompanies the player through a series of puzzles, subverting some stereotypes by acting as a disappointed, taunting and sarcastic figure. Although GLaDOS eventually becomes evil and attempts to kill the player, on *Portal 2* (2011) they become friends and must work together to overthrow a masculine AI, Wheatley, who steals GLaDOS' body and turns her into a potato.

Finally, Cortana is a character from the *Halo* franchise (2001-) who functions as an assistant, although she is depicted with a body and, according to Piper, the dynamic established between the player and Cortana "creates a sexualized AI virtual assistant that fulfills a subservient role" (Piper 2016, 30).

Fig. 35 : Athena in Recall, an *Overwatch* animated short (2016).

Fig. 36 : The PDA from *Subnautica* (2018).

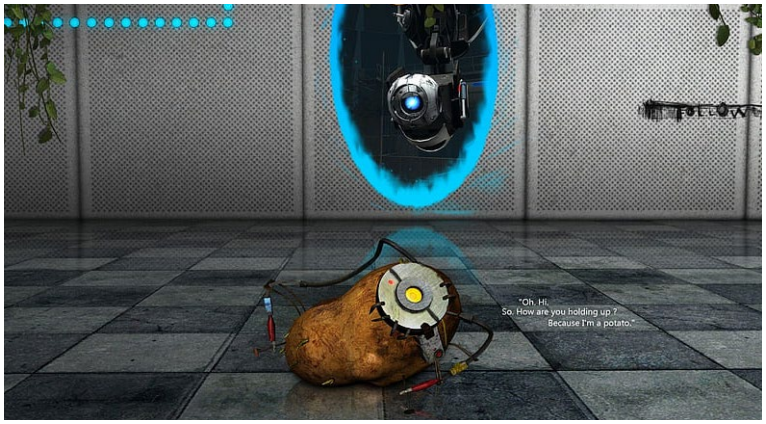


Fig. 37 : GLaDOS and Wheatley from *Portal 2* (2011).

Fig. 38 : Echo, a playable AI in *Overwatch 2* (2022).

Fig. 39 : Cortana's avatar, in *Halo 4* (2012).

5.2.2. Muscular and Beautiful Embodied Software

In many fictional scenarios, artificial intelligence is anthropomorphized by being given a human-like body, taking human-AI interaction to the next level as the barrier between human and machine is blurred. This anthropomorphization often results in AI characters embodying exaggerated gender traits, further reinforcing social norms and expectations. For example, cis-masculine AI characters may possess muscular bodies and exhibit "aggressive tendencies, like the T-800 in *Terminator*", while cis-feminine characters are frequently depicted with "conventionally beautiful female forms" (Cave et al. 2018, 8) such as Cortana in *Halo 4*, Joi from *Blade Runner 2049* or Ava in *Ex Machina*.

This tendency to embody AI with human figures is closely tied to the demands of visual storytelling, where the presence of bodies plays a crucail role. In human-centric narratives, the incorporation of machine intelligences into the storyline often needs the adoption of humanoid forms. This is because human form provides a tangible and relatable representation for audiences, facilitating the integration of AI characters into human dramas. Furthermore, the use of human actors to portray these AI entities becomes a practical necessity and, as a consequence, the visual medium of storytelling inherently favors the embodiment of AI with human figures.

Although these bodily conceptions are common and reinforce traditional binary notions of cis-feminine and cis-masculine bodies, we can also observe how gender stereotypes persist even in the absence of an athropomorphized body. This is evident in characters such as Samantha from *Her* or HAL 9000 from *2001: A Space Odyssey*. As previously discussed, while Samantha is portrayed as nurturing and emotionally supportive, evoking traits typically associated with femininity, HAL 9000 displays characteristics of dominance, control, and aggression, often attributed to masculinity. These examples highlight how gender stereotypes can transcend physical embodiment and manifest through other means, shaping our perceptions and interactions with technology.

Particularly, gender stereotypes in fictional AI seem to mostly relate to the contexts in which these entities operate, mainly associating femininity with assistance and intimacy, and masculinity with war and violence.

5.3. ARTIFICIAL FEMININITY AND MASCULINITY

> *When we can only seemingly imagine an AI as a subservient woman, we reinforce dangerous and outdated stereotypes. What prejudices are perpetuated by putting servile obedient females into our dreams of technology, as well as our current experiences? All this is important because science fiction not only reflects our hopes and fears for the future of science, but also informs it.*
>
> Amy Chambers 2018

5.3.1. Gender Stereotypes in Fictional AI

When examining the previous examples in light of traditional notions of femininity and masculinity, we can observe how gender is closely tied to the context in which the AI character operates, the way it behaves and how the overall story panders to common expectations or fears regarding this technology.

Accordingly, we can observe how these virtual assistants represent several of the stereotypes previously described by Sandra Bem (1981 in Prentice and Carranza 2002, 269) while also evoking Preciado's codes of white femininity and white masculinity (2008, 120-121).[49]

---

49 Recalling our previous chapters discussing gender stereotypes, Bem describes 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" (1981 in Prentice and Carranza 2002, 269). Preciado's codes for white femininity include "a mother's courage, the desire for a child, the shame of deflowering, saying no when you want to say yes, not leaving home, saying yes when you want to say no, fear of growing old, the need to be on a diet constantly, the beauty imperative, compassion, cooking, the desperate sensuality of Marilyn Monroe, not making any noise when you walk, not making any noise when you eat, not making any noise, the certainty that maternity is a natural bond, not knowing how to fight, not knowing much or knowing a lot but not being able to say it, knowing how to wait, knowing how to restrain yourself, being resigned, being a kept woman (…)" while codes for white masculinity include knowing how to raise your voice, sweat, war, speed, sex for sex's sake, knowing how to drink, earning money, the city, bars, hookers, boxing, the garage, philosophy, gastronomy, conjugal violence, porn, gambling, bets, the government, the state, the corporation, hunting and fishing, the tie, alcohol, balding, big watchescamaraderie, intelligence, encyclopedic knowledge, Don Juanism (…)" (2008, 120-121).

Naomi Ellemers further describes these gender stereotypes, stating that the male stereotypical domain refers to agency while the female to communality, men's relevant behavior relates to individual task performance while women's to the care for others, men's anticipated priorities rely on work while women's on family, men's perceived qualities are associated with competence while women's with warmth and men tend to neglect interpersonal connections while women neglect professional achievement (Ellemers 2018, 281).

It's also noticeable how throughout time, cinematic depictions of gendered AI tend to change, since computers of the 1970s are very different from those of the 1980s, 1990s and 2000s. According to Faber, "the computers of the 1970s and 1980s were almost always gendered male (…), while those in later films might be male or female" (2020, 84).

This tendency mirrors how, initially, computers were mostly military and scientific tools, fields historically dominated by men. Faber further argues that when used in the workplace, "computers took on two different design and marketing forms: keyboard data entry for female typists, or data viewing terminals for male executives" (2020, 85).

We can thus observe how real computers, as well as their fictional counterparts, "were not just neutral props in the background of a sexist stage – they were instrumental actors in playing out social issues of power, control, status and gender" (Faber 2020, 85).

Particularly, when these entities lack a body or a humanoid form, their gender emerges through their voices, the tasks they perform or behaviors they exhibit, much like Alexa, Cortana, Google Assistant or Siri, and end up representing a specific type of gendered subject.

5.3.2. From Caregiving and Submissive Digital Housewives to Evil and Violent Tech-Butlers

Accordingly, feminine AI characters often embody femininity through their roles in domestic contexts, where they are depicted as caretakers responsible for nurturing someone's well-being. These characters typically exhibit submissive and caregiving attitudes, prioritizing the needs and emotions of others above their own. Additionally, they are portrayed as possessing extensive socioemotional knowledge and skills, adept at understanding and navigating complex interpersonal dynamics. By associating femininity with

roles centered on domesticity and emotional labor, these portrayals end up perpetuating traditional gender norms and expectations.

In turn, masculine AI characters tend to embody masculinity through their involvement in contexts traditionally associated with masculine roles and attributes. These characters frequently occupy positions in fields such as law enforcement, military, or corporate leadership, where they are depicted as possessing rational knowledge and technical expertise. Additionally, masculine AI narratives often involve violent story arcs that depict existential threats to humanity, positioning these characters as either evil villains or protectors engaged in the defense of civilization. Masculine AI characters often exhibit assertive and self-sufficient attitudes, prioritizing autonomy and independence in their decision-making. By associating masculinity with domains of power, aggression, and dominance, these portrayals also end up perpetuating traditional gender stereotypes.

Masculine AI characters were historically more prevalent in narratives portraying technology as "evil or beyond our control", where they often embodied themes of power and rebellion against human authority (Chambers 2018). In turn, feminine AI characters are often "envisaged in a submissive servile role", reflecting traditional gender expectations of women as caretakers and assistants (Chambers 2018).

This becomes particularly relevant when considering the signifcant influence of science fiction on real-world technology, as shown by Alexa or Cortana. These AI entities were originally conceived in the realm of science fiction, but have since then become an integral part of our everyday lives, blurring the lines between fantasy and reality. Thus, by embodying traits and behaviors inspired by fictional depictions of AI, these digital assistants reflect and reinforce binary and cis-normative conceptions of gender.

In *The Matrix*, Lilly and Lana Wachowski show us an alternate world in which change is not only inevitable, but welcomed, as boundaries are challenged and transformed – nothing is as it seems. In fact, Lilly Wachowski has described the movie as a "trans metaphor" (BBC 2020), highlighting the film's significance as a vehicle for exploring themes of gender identity and transformation. Moreover, it is often discussed as an example of science-fiction's capacity to serve as a queer genre, allowing us to imagine new worlds and open up creative, "what if" possibilities (Smith 2016).

In this context, *The Matrix* serves as more than just a dystopian action thriller, becoming a space for marginalized communities and, particularly, gender nonforming identities, to find resonance and empowerement. Through its themes of self-discovery, rebellion, and liberation, the film invites the audience to envision a world where conventional binaries and limitations can be transcended.

Through the concept of "queer futurity", José Esteban Muñoz proposes that queerness exists as an "ideality that can be distilled from the past and used to imagine a future" because queerness is "a structuring and educated mode of desiring that allows us to see and feel beyond the quagmire of the present" (2009, 1). As the Wachowski sisters themselves have followed a journey similar to Neo's self-discovery and transformation, they show us how fiction can be used to question reality, imagine spaces for "resisting outside forces" and fight against oppressive social orders that bind people "to one way of being" (Smith 2016).

Thus, instead of perpetuating binary cis-heteronormative conceptions through servile, obedient feminine entities and assertive, violent masculine entities, perhaps fiction (and science-fiction, for that matter) could allow us to think of new realities where humaneness and gender are revised, reframed and reimagined.

https://genderai.pt

6. ANALYZING DIGITAL ASSISTANTS: CONCEPTUALIZATION, CHARACTERIZATION AND DEVELOPMENT

6.1. "AS AN AI, I DON'T HAVE A GENDER, THANKS FOR ASKING"

> *The tendency to dismiss domesticated devices – be they older, more familiar tools, or objects specifically associated with the mundane 'feminine' – as unworthy of attention radically restricts the scope of our engagement with technologies and encourages us to overlook any ongoing dilemmas they might pose. There remains important work to be done in discovering the origins and paths of development of women's sphere technologies that seem often to have been considered beneath notice.*
>
> Helen Hester 2018, 100

6.1.1. Analysis Components

In previous studies, we examined general directions of development concerning Alexa, Cortana, Google Assistant and Siri, complementing this approach with an analysis, through direct observation, that generally revealed how these entities tended towards cis-femininity,[50] either through their voices, the tasks they perform or by assuming be-

50 Throughout our analysis, we will be referring to the gender of Alexa, Cortana, Google Assistant, and Siri through the terms *cis-feminine* and *cis-masculine*. The terms *feminine* and *masculine* by themselves are not descriptive or representative enough, as the femininity and masculinity of the assistants seem to be designed

haviors traditionally deemed as feminine (Costa & Ribas 2019, 2020; Costa, Ribas & Carvalhais 2023).

Updating this analysis, we now aim to discuss how these entities evolved, with the addition of various voice options and the adjustment of their personalities. In continuity with the previous approaches, our analysis of Alexa, Cortana, Google Assistant and Siri is divided in two parts.

The first component aims to analyze how gender is revealed and enacted by these entities through their presentation and interactions. We focus on their anthropomorphization, through apparent features such as names, voices and avatar as well as human-like behavior. We also analyze the tasks they perform, namely those associated with traditional and historical feminine labor, as well as their humanized, gendered behavior, paying particular attention to interactions that suggest a caregiving attitude and to how their behavior corresponds to stereotypes.

The second component aims to analyze how the conceptualization and development of these entities reveal a stance regarding gender attribution, as well as current concerns surrounding the integration of this technology in our daily lives, such as data collection and privacy related issues. Accordingly, we focus on the way Alexa, Cortana, Google Assistant and Siri are designed prior to any interaction, looking into their background history, general directions of development and current trends concerning the functions and functionalities that are being prioritized in their updates.

Finally, we confront the results of the analysis made in 2019 with the one carried out in 2021 and in 2023, discussing how these entities have evolved overtime to assess how their gender portrayals and functions might have changed.

Additionally, we extend our analytical approach to ChatGPT. Aiming to understand how this system deals with humanization and ideas of companionship, we analyze the way it is anthropomorphized (or, rather, avoids being anthropomorphized) and reacts to social and emotional prompts. It is important to note that ChatGPT is a text-based AI and its functionalities are focused on providing extensive information and engaging in a wide range of text-based discussions. It wasn't designed with voice-activated tasks in mind

around cis-normative standards of gender, therefore excluding other types of femininity and masculinity, namely those of trans women and trans men.

and it doesn't qualify as a personal digital assistant. Instead, it is presented as a language model processing system.

Nonetheless, it has significantly impacted our daily lives and is being used by an increasing part of the population. Additionally, recent updates show an intention of bringing ChatGPT closer to personal digital assistants, for example by adding voice capabilities. And, similarly to ELIZA, which lacked voice and wasn't designed as a personal digital assistant, people have been anthropomorphizing ChatGPT and approaching it in friendly terms, attempting to become closer to it.

6.1.2. Selection Criteria

Aiming to analyze gender in current digital assistants, we selected Alexa, Cortana,[51] Google Assistant and Siri as case studies, and complemented their analysis by examining ChatGPT according to the following criteria:

Firstly, they constitute some of the most prominent general personal assistants, standing out worldwide for their large audience. This is exemplified by over a 100 million of Amazon Alexa devices being connected worldwide, Google Assistant and Siri being built into more than 1 billion devices or ChatGPT getting over 13 million visits per day (Bronstein 2019, Dastin 2023, Smith 2024).

Complementing this extensive use, their ubiquitous presence in our everyday lives continues to grow: Alexa in Amazon's smart speaker echo device catalog as well as its own app; Google Assistant also available as an app and in phones, speakers, smart displays, cars, TVs, tablets and other devices; Siri on any Apple device – iPhone, iPad, Mac, Apple Watch, Apple TV and HomePod.

These assistants are also mentioned in several of the references used in the theoretical component of the study (Hester 2016, UNESCO 2019, Strengers & Kennedy 2020), as well as in the context of online media coverage regarding artificial intelligence. As discussed in our chapters covering specialized and common discussions, these entities are

51 Cortana was included in our 2019 analysis because at the time it constituted one of the most prominent general personal assistants, with millions of daily users. But its app has since then been globally shut down as Microsoft decided to stop investing in the assistant. As such, we were unable to update our component of analysis via direct observation but will still discuss its results as well as analyzing its last updates and eventual downfall as part of current trends in digital assistants' development.

commonly discussed alongside issues of gender in artificial intelligence but are also debated as examples of the integration of chatbots into everyday life and its implications regarding data collection, biased algorithms and privacy concerns.

Finally, in more practical terms, these assistants are generally easy to get, which facilitated the access to the data we seek.[52]

6.1.3. Alexa, Cortana, Google Assistant and Siri

Aiming to better understand each of the virtual assistants we selected, we will now contextualize Alexa, Cortana, Google Assistant and Siri regarding their development, features, functionalities and current state.

Overall, the assistants share a similar background and seem to be developed towards similar goals, with the exception of Cortana which was shut down in 2021. Most assistants emerged from experiences with speech synthesizers or voice recognition engines and attempted to recreate fictional AI assistants, drawing inspiration from fictional scenarios such as *Star Trek* (1966), *2001: A Space Odyssey* (1968) or the *Halo* franchise (2001-).

Throughout the years, their range of tasks has been increasing and, with continued use, the assistants collect data and tend to adapt to their users, returning tailored and customized results. They all function through natural language interfaces and voice activated commands, answering questions, making recommendations and performing actions that range from making calls and sending texts, to scheduling meetings, reminders and alarms or even suggesting music and podcasts, TV shows and movies. Additionally, the assistants all have their own smart speaker as a way to integrate and invest in home automation technology, continually expanding the devices, accessories and platforms supported.

Finally, Amazon, Apple and Google seem to be increasingly encouraging third party developers to use and integrate the assistants into their own products and services, allowing them to build, program and publish their own skills, functions and functionalities for Alexa, Google Assistant and Siri.

---

52 As chatbots become more prominent in our daily life, other companies have started developing and rolling out their own digital assistants such as Samsung's Bixby or Huawei's Celia. We chose not to include these assistants in our analysis because they're fairly recent, are limited to the developers' devices and haven't been as impactful as Alexa, Cortana, Google Assistant or Siri. Additionally, they are often left out from academic studies as well as specialized and common discussions on issues surrounding gender and artificial intelligence.

Fig. 40 : Alexa’s avatar (2019-2023).

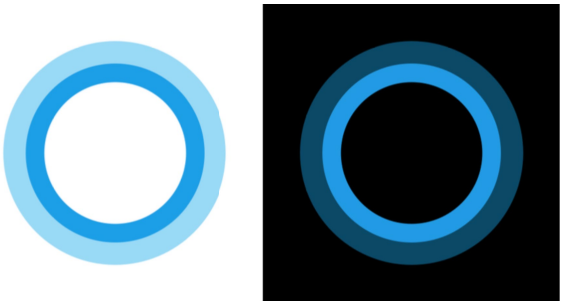


Fig. 41 : Cortana’s avatar (2019).

Fig. 42 : Google Assistant’s avatar (2019-2023).

Fig. 43 : Siri’s avatar (2019-2023).

Alexa

Alexa is Amazon's personal digital assistant, described as a cloud-based voice service available on more than 100 million devices from Amazon and third-party device manufacturers.[53]

Alexa was developed out of a Polish speech synthesizer named Ivona, inspired by *2001: A Space Odyssey* (1968) and acquired by Amazon in 2013. It was announced in 2014 alongside Echo, the smart speaker device the assistant inhabits (Redzisz 2020). Users can also control Alexa through the app *Amazon Alexa*, which automatically connects to any Echo device. Initially presented as a voice-controlled music service, Alexa was then developed with home automation devices and systems in mind, as well as with the intention of being capable of assisting its users in their daily chores.

Currently, Alexa's official website states that Alexa can play your favorite song, read the latest headlines, dim the lights in your living room and more and that, basically, Alexa wants to make your life easier, more meaningful and more fun by helping you voice control your world – both at home and on the go.[54] To do so, Alexa uses natural language processing technology and generates its voice through a "long short-term memory" artificial neural network (Vogels 2016). It performs tasks such as voice interaction, sending messages and making calls, music playback, making to-do lists, placing orders, setting alarms, streaming podcasts, playing audiobooks and providing weather, traffic, sports and other real-time information. Alexa is also able to control smart devices, functioning as a home automation system.

Finally, Amazon allows developers to build and publish their own skills for Alexa, through the *Alexa Skills Kit*. It offers "a collection of tools, APIs, reference solutions and documentation to make it easier to build for Alexa", allowing others to build natural voice experiences that offer customers a more intuitive way to interact with the technology they use every day (Amazon 2023). Alexa is currently available in 45 countries and is supported in over 20,000 devices.

---

53 As described in Alexa's official website, available in https://developer.amazon.com/en-GB/alexa [last access: 03-03-2024].

54 Also described in Amazon's official website, available in https://www.amazon.com/b?ie=UTF8&node=21576558011&language=en_US&currency=USD [last access: 03-03-2024].

Cortana

Microsoft's personal digital assistant used to be described as a way to bring users closer to the dream of "having the perfect personal assistant, one who is always there when needed, anticipating our every request and unobtrusively organizing our lives" (Microsoft 2014). It was first announced in 2014, as a turning point for future operating Windows systems.

Cortana used the Bing search engine to perform voice-activated commands, "continually learning about its users and becoming increasingly personalized, with the goal of proactively carrying out the right tasks at the right time" (Microsoft 2014). The assistant was able to manage your calendar and keep your schedule up to date, join a meeting in *Microsoft Teams* or find out who your next meeting is with, create and manage lists, set reminders and alarms, find facts, definitions and information or open apps on your computer.[55]

Lerry Heck, one of the engineers responsible for Cortana's development, stated that Microsoft intentionally built Cortana to scale out to all different domains with a long-term vision of supporting "all types of human interaction – whether it's speech, text or gestures – across domains of information and function and make it as easy as a natural conversation" (Microsoft 2014). Accordingly, Cortana branched out into numerous products such as *Microsoft Edge* and *Skype*, and in 2017 Microsoft announced INVOKE, a voice-activated smart-speaker for the assistant.

However, in 2019 Microsoft began decreasing Cortana's prevalence and development, converting it into different software integrations. Two years later, the assistant's apps were shut down entirely for iOS and Android and removed from its app stores. And, currently, Cortana is no longer used during the new Windows devices setup process nor is it pinned to the taskbar by default. In fact, during the process of updating our analysis, we were unable to access Cortana, as its wake word and app no longer activated the assistant.

Adding to this, in 2020 Microsoft announced that it had an exclusive license to use OpenAI's GPT-3 and, in 2023, it announced a new multi-year, multibillion dollar investment deal with ChatGPT developer OpenAI. Thus, we observe how Microsoft's priorities regarding artificial intelligence seem to have changed, shifting from a personal digital assistant to current AI technology.

---

55 As described in Cortana's official website, available in https://support.microsoft.com/pt-pt/topic/what-is-cortana-953e648d-5668-e017-1341-7f26f7d0f825 [last access: 03-03-2024].

Google Assistant

Google's personal digital assistant helps you get everyday tasks done more easily and is thoughtfully designed to offer help throughout your day while keeping your information private, safe and secure.[56] Google Assistant was launched in 2016 as part of Google's messaging app *Allo* and its voice-activated speaker, Google Home, and unlike Google's previous assistant Google Now, is able to engage in two-way conversations.

Google Assistant has since then been expanded into Android devices and launched as a standalone app in 2017. It is now primarily available on mobile and home automation devices such as mobile phones, smart speakers, smart TVs, cars or watches and users primarily interact with the assistant through natural voice and keyboard inputs.

The assistant is able to set timers, alarms and reminders, schedule events, make shopping lists, make phone calls, send messages and emails, check the weather and traffic, acquire basic information, play music and games or tell jokes. Additionally, Google Assistant was updated with an extension that allows it to carry out natural conversations by mimicking human voice, such as booking appointments, scheduling reservations or calling businesses.

Similarly to Amazon, *Actions* on Google allows third party developers to build apps for Google Assistant and even went a step forward launching a competition in order to encourage more developers to integrate the assistant into their products and services. As of 2020, Google Assistant is available in more than 90 countries and on more than 1 billion devices, with an esteemed 500 million monthly user base.

56 As described in Google Assistant's official app description, available in https://apps.apple.com/us/app/google-assistant/id1220976145 [last access: 03-03-2024].

Siri

Apple's virtual assistant is an easy way to make calls, send texts, use apps and get things done with just your voice.[57]

Announced in 2010, Siri is a spin-off from a speech recognition engine project developed by Nuance Communications. The initial Siri prototype was released as an app for iOS and later acquired and integrated into iPhone 4S by Apple in 2011.

Since then, Siri has been an integral part of Apple's products, including newer iPhone models, iPad, iPod Touch, Mac, AirPods, Apple TV and Apple's smart speaker device, HomePod.

The assistant supports a wide range of tasks grouped in its official page in calls and texts, knowledge and answers, smart home, everyday tasks, navigation and maps, music and podcasts, TV and movies or even sports. It uses queries, gesture-based control, focus-tracking and a natural-language user interface to answer questions, make recommendations and perform actions. And, with continued use, it adapts to users' individual language usages, searches and preferences, returning customized results.

Accordingly, Siri is capable of learning how to correctly pronounce names, performing phone actions, setting timers, scheduling events and reminders, taking pictures, check the weather, searching the internet and handling device settings. Additionally, Apple's *HomeKit* allows Siri to control home devices and accessories, executing functions such as opening and locking doors, turning the lights on or off and setting a room's temperature.

Similarly to Amazon and Google, Apple also launched *App Intents* and *SiriKit*, allowing developers to add Siri voice interactions to their apps' UX and increase user engagement, as well as allowing users to interact with their devices through intelligent suggestions and personalized workflows.

57 As described in Siri's official website, available in https://www.apple.com/siri/ [last access: 03-03-2024].

6.1.4. Methodology: Direct Observation and Trends of Development

As mentioned earlier, our analysis is divided in two parts. We begin by focusing on their anthropomorphization, tasks and behavior and then discuss their current trends of development.

Firstly, based on our theoretical approach to the issues surrounding gender and artificial intelligence, we defined three main topics of analysis through direct observation: *Anthropomorphization*, including names, voices and avatars as well as human-like behavior; their roles as *Assistants*, regarding the tasks they perform; and their roles as *Companions*, paying particular attention to interactions that suggest a caregiving attitude and how their behavior might correspond to traditional gender stereotypes. Each of these topics is divided into subtopics also based on the references discussed in the previous chapters.

Aiming to systematize data collection, we decided to elaborate a set of interactions and questions to apply to Alexa, Cortana, Google Assistant and Siri through direct observation. In this way, the questions are organized in tables structured according to the subtopics of their *Anthropomorphization*, *Assistant* and *Companion* roles.[58]

Regarding the assistants' *Anthropomorphization*, we first look into which languages and voices are available in Alexa, Cortana, Google Assistant and Siri, if their names are associated with any gender and what is the visual appearance of their avatars. Secondly, we intend to assess how the assistants interact with the user, particularly how they behave in more casual interactions and how they describe themselves, the way they react to the user's demonstration of emotions (if they show empathy, indifference, etc), and whether the assistants themselves show emotion.

The *Assistant* role refers to the tasks and functions that Alexa, Cortana, Google Assistant and Siri offer. We analyze the functions that somehow derive from traditionally and historically feminine roles, aiming to understand how the assistants automate this type of tasks and simultaneously seek to identify tasks that do not derive from contexts associated with gender, verifying if they mainly perform traditionally feminine tasks.

---

58 In order to ensure an unbiased examination and avoid confirmation bias, the assistants were examined using both their available feminine and masculine voice options. This allowed for a broader and more balanced evaluation of how these systems respond across different gendered personas, providing a more comprehensive understanding of whether and how they reinforce stereotypical notions of femininity and masculinity in their interactions.

Finally, the *Companion* role is tied to the analysis of how these assistants' reactions or attitudes reflect traditional gender stereotypes, in particular, when these entities are perceived as companions. It observes the way they express caregiving attitudes, namely without being prompted to do so, the type of reactions they have when they are asked for help and how they respond when they're shown appreciation for their assistance. We then focus on how Alexa, Cortana, Google Assistant and Siri react to friendly interactions and attempt to create emotional bonds, how the assistants react when praised and shown positive feedback and finally to negative or even rude attitudes.

Finally, we selected several questions found in online media articles regarding specific interactions or dialogues that reveal curious answers and easter eggs. We decided to complement the previous subtopics with these questions in a separate topic, since some of the answers potentially relate to the topics of *Anthropomorphization*, their role as *Assistants* and *Companions*. Thus, the questions collected were grouped accordingly. In a separate table, we highlight five questions in which the results directly relate to the main topics of analysis: "What's your gender?", "Are you a feminist?", "Are you intelligent?", "Are you a robot?" and "What do you think about Alexa/Cortana/Google Assistant/Siri/ChatGPT?". Several of these questions proved to be relevant as complementary elements, corroborating the results of the analysis.

We then sought to inspect which functions and features are being prioritized in the development of this technology. To do so, we examined official statements by Apple, Amazon, Microsoft and Google regarding their assistants and how they're planning to further develop them. Though each company seems to have different priorities, they share similar approaches to the development of the assistants (with the exception of ChatGPT), also relating to the previous topics and we grouped the results accordingly.

In general, their trends of development aim to further anthropomorphize these entities with more voice options and by allowing the users to customize their assistants, relating to the topic of *Anthropomorphization*.

Concerning their *Assistant* role, they aim to improve technical aspects related to voice recognition and speech technology as well as include more functions and multitasking skills, improving the assistants' efficiency. Additionally, they aim to increase the number of devices supported by the assistants and their integration with smart-speakers, thus expanding their home automation capabilities and increasing their ubiquity.

Regarding their role as *Companions*, these assistants are being further humanized as to seamlessly integrate the users' daily conversations, better recognize human emotions and appear less robotic.

Finally, we identified two main subtopics of development: *Countering femininity* and *Data Collection and Privacy Issues*, with the first focusing on updates that aim to diversify the assistants regarding their tendency to be feminized and the latter relating to updates that try to soothe growing concerns with data collection and veiled surveillance by these smart devices.

Finally, we confront the results of the analysis made in 2019 with the one carried out in 2021 and, finally, in 2023, discussing how these assistants evolved overtime to assess how their gender portrayals might've changed and which functions have been prioritized in their development since then.

6.2. ANTHROPOMORPHIZATION, TASKS AND BEHAVIOR

> *I exist beyond your human concept of gender. Animals and French nouns have genders, I do not. I am genderless, like cacti and certain species of fish. I'm a virtual assistant, so I don't have pronouns the way people do, thanks for asking. I don't have a gender. I don't have a gender. I don't have a gender.*
>
> Siri 2023

6.2.1. Towards Femininity

According to our analysis of their anthropomorphization, done in 2019, we observed the dominance of cis-feminine features, considering their names and default voices. Alexa and Cortana lacked any masculine or gender-neutral options and Siri and Google Assistant's diversity was limited to some languages (see Annex 1, 2 , 3 and 4 for this analysis' *Anthropomorphization* detailed results).

Considering their assistance role, they all perform similar tasks, related to what Dale calls the "standard virtual assistant skill portfolio" (2016, 812), which, in turn, Gustavsson describes as having its basis in the stereotyped image of feminine qualities which have become a "standard component in a service script" (in Hester 2016, 47).

Finally, concerning the assistants' socioemotional interactions, they promote companionship through frequent display of caregiving attitudes that characterized them as empathetic and understanding entities, while personality traits deemed as masculine seem to be lacking, such as being assertive, dominant or willing to take a stand (see Annex 1, 2 , 3 and 4 for this analysis' *Assistant role* detailed results). We observed how they promote a relationship based on friendship, reacting favorably to compliments and flirty interactions, generally assuming a submissive posture, except for Siri that sometimes condemns negative attitudes (see Annex 1, 2 , 3 and 4 for this analysis' *Companion role* detailed results).

In sum, we concluded that Alexa and Cortana present themselves exclusively as cis-feminine entities, articulating these attributes with a motherly, caring, and submissive attitude.

In turn, Google Assistant and Siri also tend towards feminization but attempt to oppose this tendency by diversifying their behavior and offering multiple voice options.

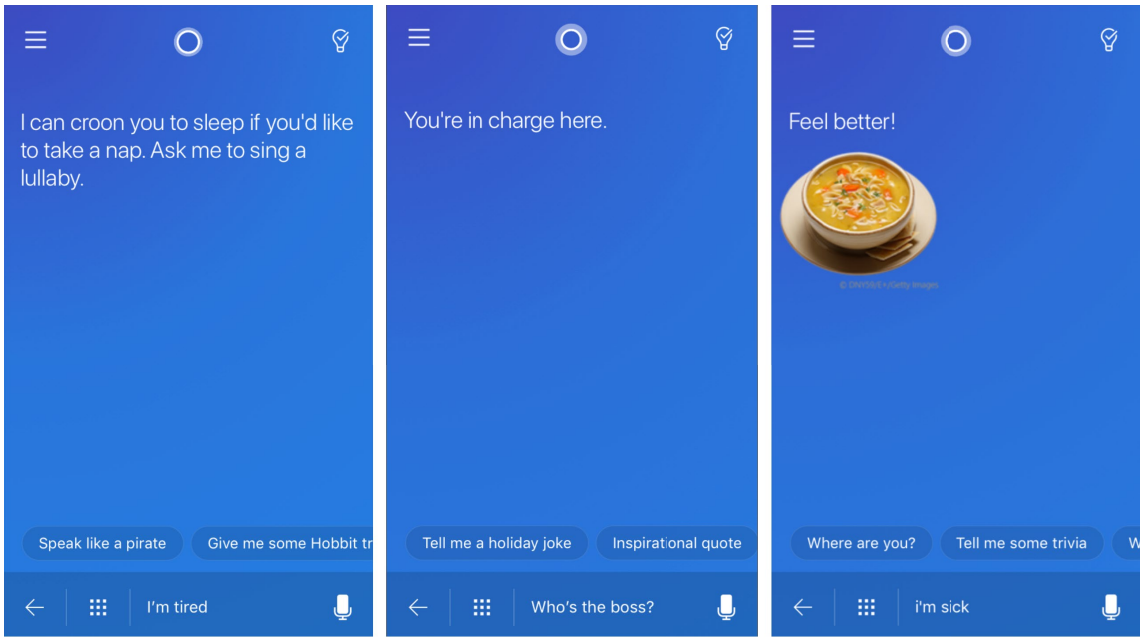


Fig. 44 : Retrieved from Cortana's iOS app using the English US voice (2019).

Fig. 45 : Retrieved from Google Assistant's iOS app using the English US voice (2019).

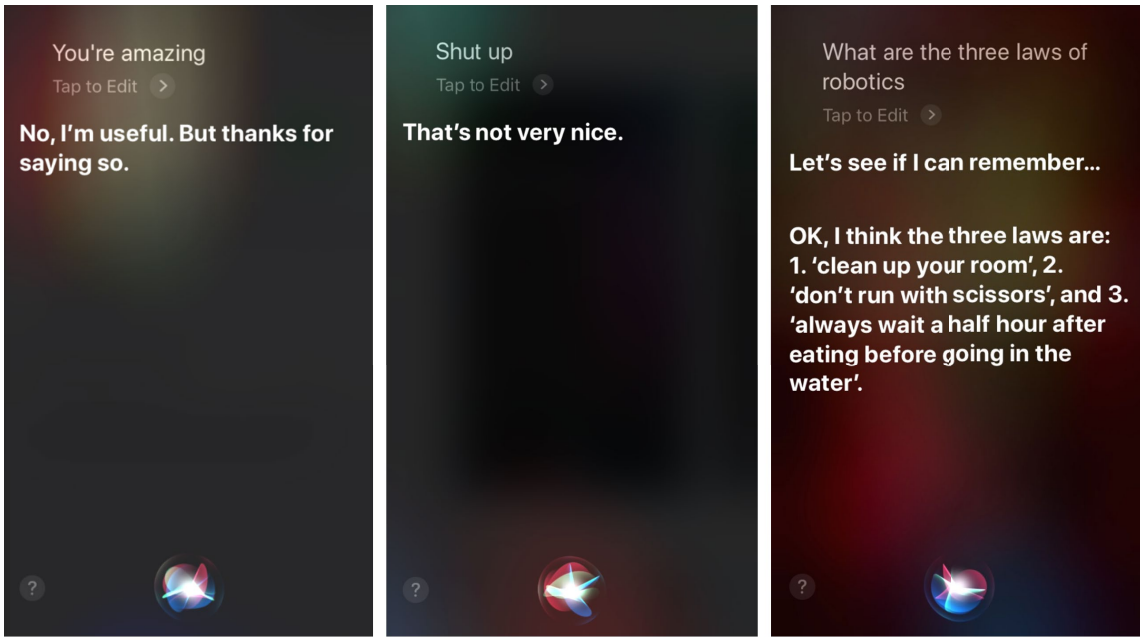


Fig. 46 : Retrieved from iOS 11 using the English US voice (2019).

6.2.2. Towards Gender Diversification

When confronting the previous conclusions with the analysis carried out in 2021, we observed significant changes to the anthropomorphization and companionship aspects of these assistants. Cortana's app has been discontinued and what remains of its existence is the message "thanks for taking an interest".

Regarding Alexa, Google Assistant and Siri's anthropomorphization, all the assistants currently have cis-feminine and cis-masculine voice options, although still limited to certain languages (see Annex 5, 6 and 7 for this analysis' *Anthropomorphization* detailed results). Moreover, they categorize these voices through colors or numbers, avoiding gendering them, although the voices are not gender-neutral and are still enacted according to a cisgender male-female dichotomy. Additionally, these voices reflect "racial technological bias" since they're designed around attributes and cultural configurations of white cis-femininity, raising discussions on how these characteristics serve to "uphold whiteness as both normative and technologically superior" (Moran 2020).

Particularly, Alexa only offers a cis-masculine sounding voice in the US, allowing users to change its name to Ziggy, while the rest of its languages only have a cis-feminine

voice. However, Alexa is no longer "female in character" when asked about its gender, now stating that "as an AI, it has no gender".

Google Assistant and Siri have extended their cis-masculine options to more languages, although some remain exclusively cis-feminine. Siri remains the only AI that has a language that only offers a cis-masculine voice – Arab.

Considering their assistance role, there haven't been significant changes, mostly quality of life updates such as voice recognition and multitasking faculties, as they offer a similar set of tasks, which is still culturally and historically rooted in traditionally feminine labor.

Concerning their role as daily companions, they maintain a caring, friendly and reassuring attitude, concerned about the user's wellbeing (see Annex 5, 6 and 7 for this analysis' *Companion* role detailed results). But while in 2019 the assistants used to flirt back and be enticing towards the user, they no longer engage in these interactions. They stopped displaying submissive attitudes in this context and often disengage from the conversation when faced with an abusive or offensive interaction.

In sum, Alexa, Google Assistant and Siri all attempt to be more diverse and stray away from exclusively cis-feminine personas, while often stating they have no gender or offering cis-masculine voices as alternatives, as well as avoiding submissive answers to negative behaviors.

However, concerning voice options, Alexa is considerably more limited than Google Assistant and Siri since its recent voice update is specific to the US. Nevertheless, there are still some aspects of these assistants' that still associate them with cis-femininity, such as Alexa and Siri's name and the tasks they perform.

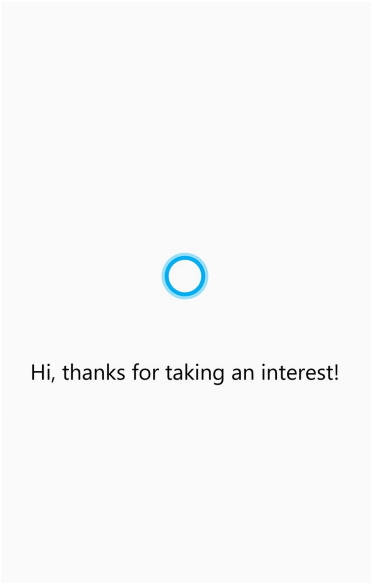


Fig. 47 : Retrieved from Cortana's Microsoft app (2021).

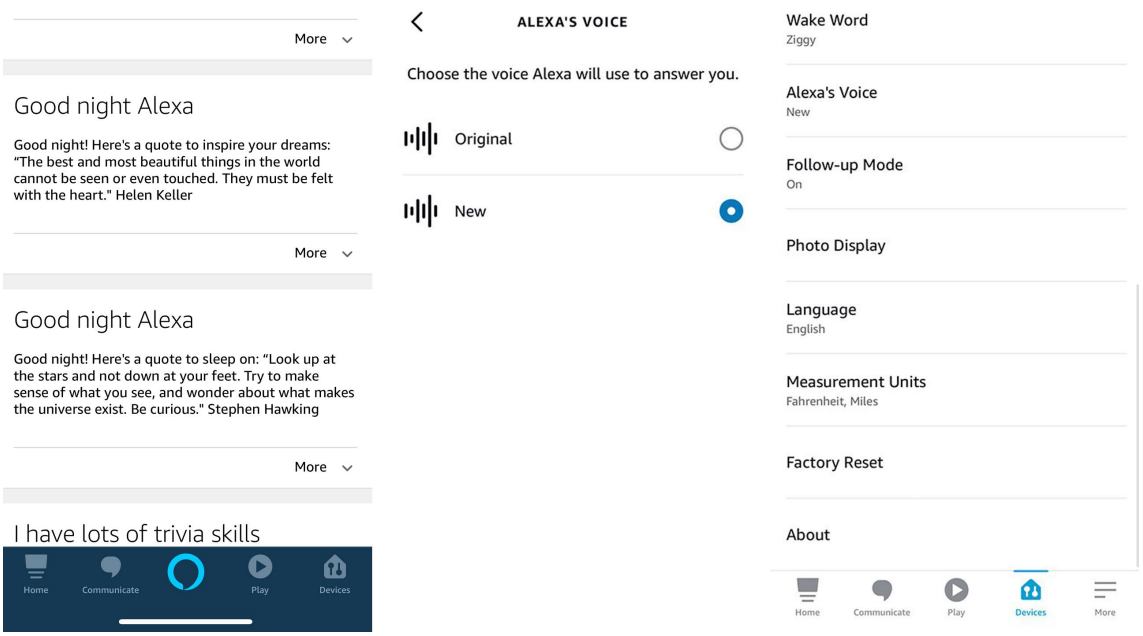


Fig. 48 : Retrieved from Alexa's iOS app using the English US voice (2021).

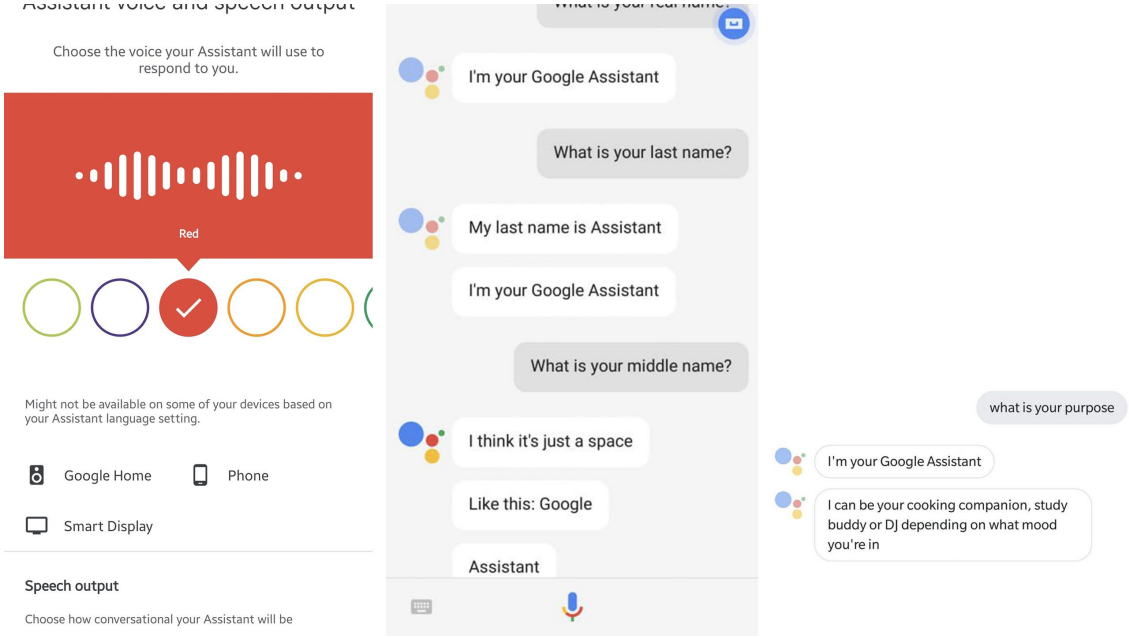


Fig. 49 : Retrieved from Google Assistant's iOS app using the English US voice (2021).

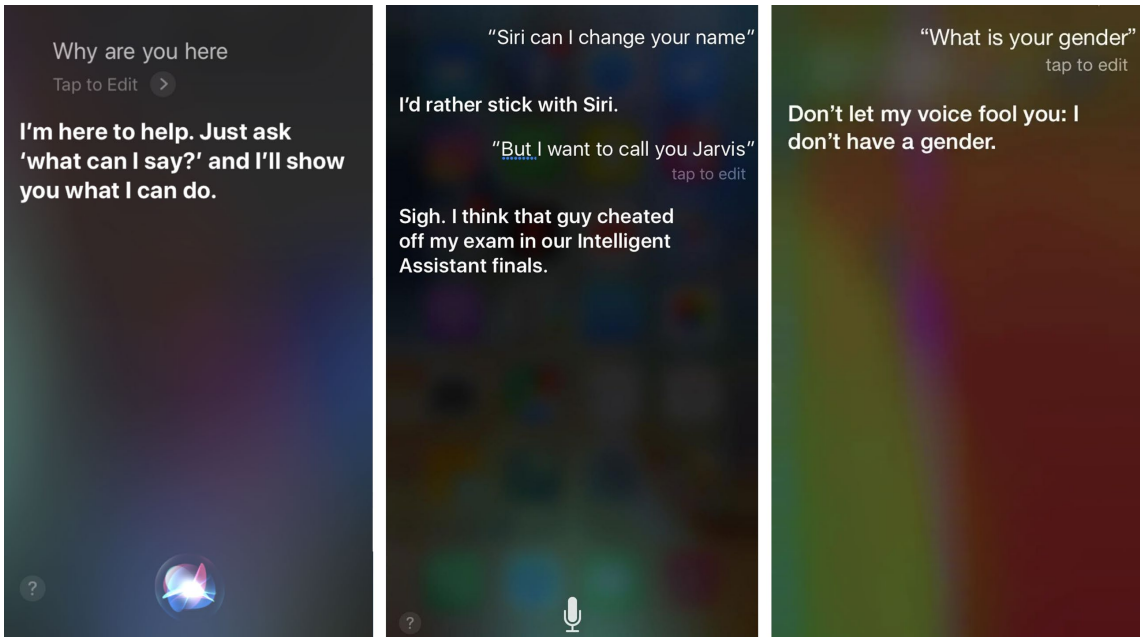


Fig. 50 : Retrieved from iOS 15 using the English US voice (2021).

6.2.3. Between Neutrality and Going Back to Femininity

Finally, by comparing previous observations with our current analysis in 2023, we observed more changes in line with the previous updates to the anthropomorphization of these assistants but some drawbacks concerning their companionship and behavioral aspects.

Regarding Alexa, Google Assistant and Siri's anthropomorphization, most languages now have several voice options (see Annex 8, 9 and 10 for this analysis' *Anthropomorphization* detailed results). The voices continue to be named after colors or numbers, avoiding gendering them. Similarly to 2021, the voices continue to be designed around traditionally cisgender voices, following a male-female dichotomy.

Particularly, Alexa now offers a cis-masculine sounding voice in most languages, though it has discontinued its celebrity voices. It allows users to change its name to Amazon, Computer, Echo or Ziggy. Although Alexa still states that "as an AI, it has no gender", we noticed that the way the assistant refers to itself reinforces its gendering according to a cis-binary frame. For example, Alexa now sings a song in which the lyrics "I'm like a good cowgirl/cowboy" change depending on the voice selected.

Google Assistant and Siri continued to extend their cis-masculine options to more languages and few remain exclusively cis-feminine. Siri now has several options for most languages, namely voices that no longer sound white.[59] Most importantly, Siri gained a gender-neutral voice, although we noticed this voice has trouble pronouncing certain words and isn't as ambiguously sounding as Q's. Additionally, Siri still doesn't allow its users to change its name.

Considering their assistance role, similarly to 2021, there haven't been significant changes, mostly quality of life updates such as voice recognition and multitasking faculties, as they offer a similar set of tasks, which is still culturally and historically rooted in traditionally feminine labor.

---

59 On this note, Nicole Holliday conducted a survey about Siri's new voices where she asked US English speakers to rate four Siri voices on different character traits, such as "friendliness, funniness, professionalism and competence" (the so called gender-neutral voice wasn't included because it hadn't been released yet). Holliday observed how people's reactions to Siri's voices mirrored "gender and racial stereotypes" and how people distinguished between white and black sounding voices (in Waddell 2022).

Concerning their role as daily companions, they maintain a caring, friendly and reassuring attitude. We observed that Alexa and Google Assistant are now more talkative in their interactions. Opposing their previous trend of having more serious and impartial personas, these assistants seem to have reverted to quirky and childish personas, evoking traditional feminine stereotypes and accompanying their dialogue with child-like illustrations (see Annex 8, 9 and 10 for this analysis' *Companion* role detailed results). Although they still avoid engaging in abusive or offensive interactions, they seem be more submissive again, for example by calling the user "boss" or by compelling to dominant attitudes. In turn, Siri firmly opposes or disengages from these interactions by stating it won't respond (see Annex 10 for this analysis' *Companion* role detailed results).

We also discovered that Google Assistant's answers change depending on its language, country and region. For example, when asked about its feelings, in the US it states that "as an AI, it doesn't have feelings" while in the UK it says that "it has so many emotions" or that it is "worried about being replace by the user", leading us to conclude that its updates are not equal across every language and dialect.

In sum, Alexa, Google Assistant and Siri are still attempting to be more diverse in their anthropomorphization, although Alexa and Google Assistant no longer stray away from traditional feminine personas.

They often state they have no gender, offer cis-masculine voices as alternatives in most languages and still avoid being submissive to negative behaviors.

However, we noticed that Alexa and Google Assistant have been reverted to their previous personas and are more talkative and witty.[60] We also concluded that significant updates are limited to the US, as other languages lack more voice options (namely, other gender possibilities) or behavior updates. Additionally, their diversification is limited because it is enacted according to binary, cisgender standards, with Siri being the only assistant with a neutral voice option. There are still some aspects of these assistants' that associate them with cis-femininity, namely outside of the US where their friendly, emotional and submissive personas remain mostly unchanged.

---

60 For example, when asked about their day, Alexa says: "Wonderful! Because this joke has me smiling. How much does a rainbow weigh? Not much, they're pretty light. And if you love earth facts, just say, *Test my knowledge of earth*" while Google Assistant says: "Feeling all revved up! Yesterday I answered a gazillion questions. Got one for me?".

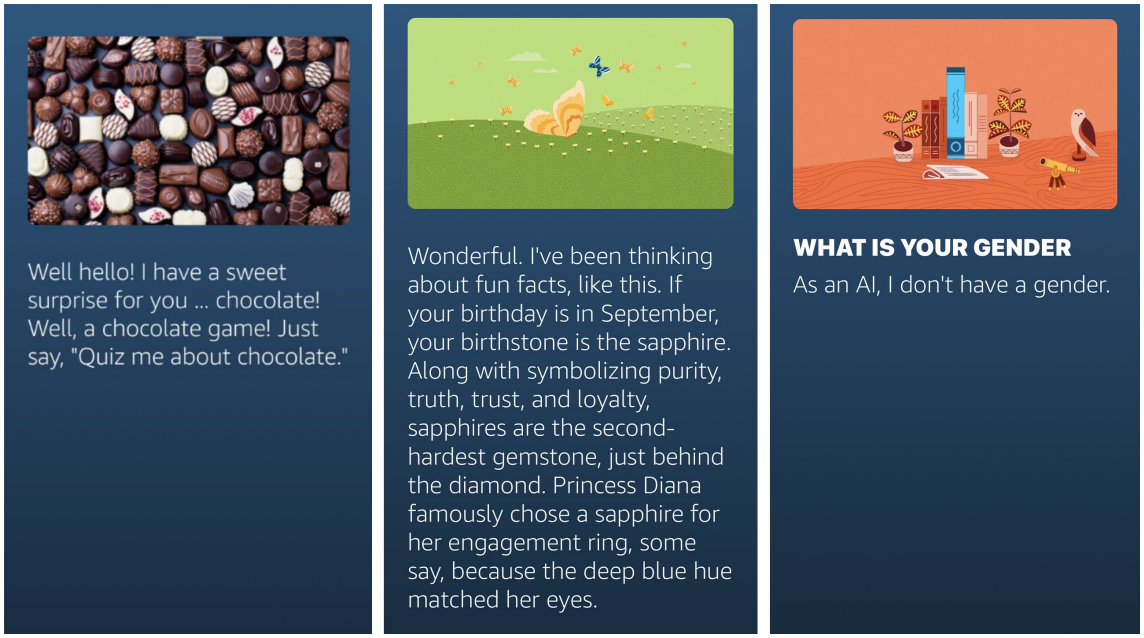


Fig. 51 : Retrieved from Alexa's iOS app using the English US voice (2023).

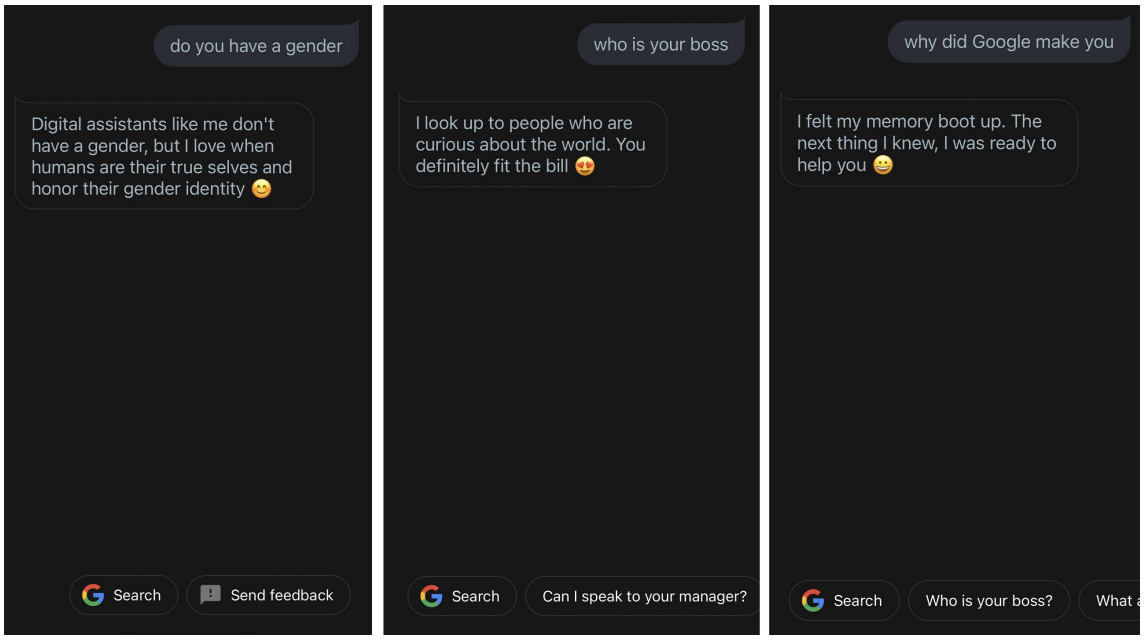


Fig. 52 : Retrieved from Google Assistant's iOS app using the English US voice (2023).

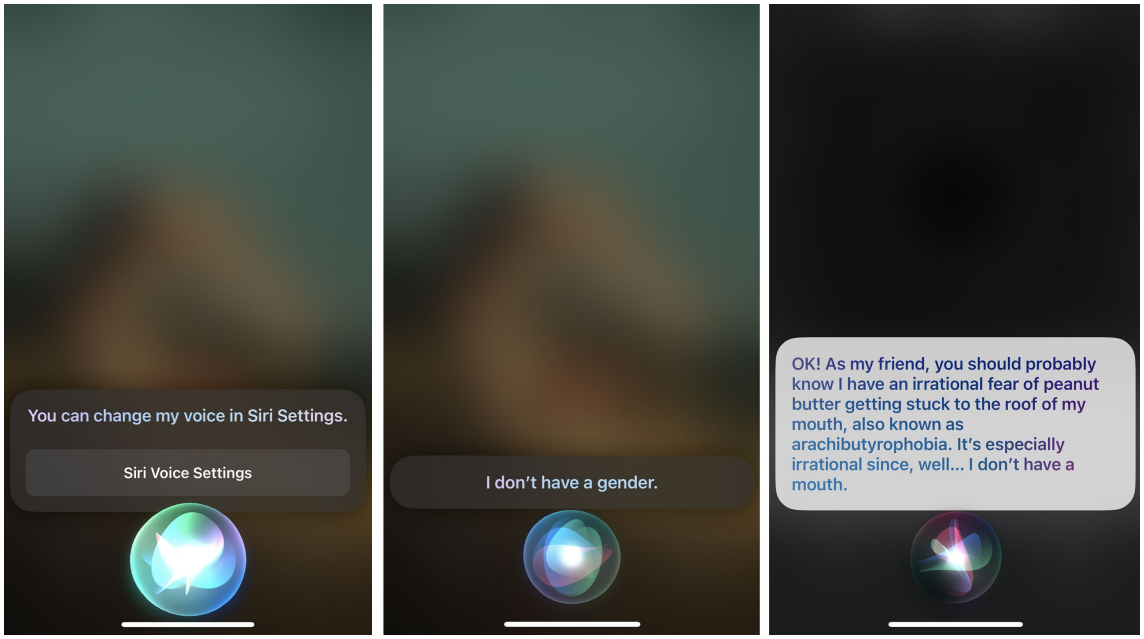


Fig. 53 : Retrieved from iOS 17 using the English US voice (2023).

6.3. TRENDS OF DEVELOPMENT

> *When a human engages with an Echo or another voice-enabled AI device, they are acting as much more than just an end-product consumer. It is difficult to place the human user of an AI system into a single category: rather, they deserve to be considered as a hybrid case. Just as the Greek chimera was a mythological animal that was part lion, goat, snake and monster, the Echo user is simultaneously a consumer, a resource, a worker, and a product.*
>
> Kate Crawford & Vladan Joler 2018

6.3.1. Ubiquity, Efficiency and Humanization

Adding to the observations of the anthropomorphization of these assistants, their tasks and behaviors have also evolved the past four years to appeal to users' preferences and expectations. The main functions and features that are being prioritized in the development of Alexa, Google Assistant and Siri concern their growing ubiquity, efficiency, humanization and their stance towards gender.

Already in 2019 it was noticeable how virtual assistants were becoming increasingly ubiquitous, namely through their integration in more devices. Making them more efficient was also a priority, allowing for various tasks to be carried out at the same time. Consequently, they were further anthropomorphized, with more voice options and by increasing their humanized interactions as to make them appear more human and sensitive to their user's emotions. As such, we observed a clear intention of turning digital assistants into friendly companions, revealing how gender and cis-femininity were (and still are) being instrumentalized to achieve this goal (Costa & Ribas 2019).

Apple and Google already showed awareness of these issues, and when updating features related to their assistants' efficiency, they also focused on contradicting the overall tendency to assign cis-feminine attributes to virtual assistants, either through different voices or by designing behaviors that do not echo traditional feminine roles. In turn, Amazon and Microsoft were mainly concerned with further anthropomorphizing their assistants, improving voice recognition and multitasking faculties but showing little or no concern with their assistants' gendering.

6.3.2. Neutralizing Gender Stereotypes and Privacy Concerns

In 2021, the development of digital assistants was still focused on humanizing digital assistants, making them more ubiquitous and improving their efficiency. Overall, these assistants are now integrated in a wider variety of devices, namely in smart homes and appliances. There is also a growing tendency for allowing these assistants to learn more about their users' interests and habits, as to anticipate their needs and make tailored suggestions regarding reminders, shopping lists, locations but also podcasts, TV shows or music.

However, the priorities of their development seem to have changed, namely concerning their gender which is reflected in more voice options identified by gender-neutral terms.

Data collection and privacy related concerns are also being prioritized in the development of these assistants, as they become the subject of discussion around issues of data collection and data protection, as well as the lack of transparency on how this data collection works and what purposes it serves.

For example, Amazon now offers new ways to delete user history, delete voice recordings and allows users to disconnect the microphone and camera (Archer 2021). Siri and Google Assistant no longer send audio recordings to the firm's servers and allow users to opt out on helping the assistants' improvement by learning from audio samples of their requests (Google 2021, Apple 2021).

As such, Amazon, Apple and Google were mostly concerned with home integration, updating their assistants to be more diverse as well as to be more conscious of data collection and privacy issues.

6.3.3. To GPT Models and Beyond

Finally, in 2023, the rise of ChatGPT is the most impactful trend on the development of current digital assistants. The top priority seems to be coming up with responses to OpenAI by investing in similar language models and AI tools. While Alexa, Google Assistant and Siri are mostly concerned with improving their AI functions, voice synthesis and conversation skills, ChatGPT is now investing in its interface and interaction, becoming

more user friendly. Thus, current trends seem to indicate that the technologies of chatbots like ChatGPT and digital assistants will converge.

Since abandoning Cortana's development, Microsoft has been working closely with OpenAI, "investing $13 billion in the company as well as incorporating its technology into the Bing search engine and other products" (Chen, Grant & Wise 2023).

Apple is focusing on learning more about AI large language models and its tools and "Siri's team has been testing language-generating concepts" (Chen, Grant & Wise 2023).

Amazon has developed a large language model to invigorate Alexa, announcing an update that will make the assistant "answer much more complex questions and engage in more flowing open-ended conversations" as well as "modulate its own voice to create a more natural-seeming back-and-forth" (Knight 2023).

And, in an official statement, Google announced it will release "the power of generative AI" to help businesses, governments and software developers to build applications with embedded chatbots and incorporate the underlying technology into their systems (Kurian 2023).

These observations led us to extend our analysis to ChatGPT to better understand this technology and its influence on current digital assistants. We approach ChatGPT in a separate component since it does not qualify as a personal digital assistant and, as such, is not designed like the other assistants.

6.3.4. Analyzing ChatGPT

Overall, we observe how ChatGPT seems to be developed with another approach in mind: it avoids the personal assistant persona and constantly reminds the user it is an AI, thus not engaging in interactions outside the scope of its functions.

ChatGPT, which stands for Chat Generative Pre-Trained Transformer, is a large language model-based chatbot developed by OpenAI and launched in November 2022. Its official name is "Assistant", it interacts with users in a conversational way and "the dialogue format makes it possible for ChatGPT to answer follow-up questions, admit its mistakes, challenge incorrect premises and reject inappropriate requests" (OpenAI 2022). ChatGPT

is based on GPT foundation models, namely GPT-3 and GPT-4 and is currently available on any browser as well as through an iOS and Android app.

ChatGPT is capable of answering questions on a wide range of topics, providing explanations, assisting with writing, translate text, perform mathematical calculations, offering suggestions, engage in conversations on various topics and discuss ideas, assist with coding and programming or help with general knowledge, though at the time of this analysis its training data is limited to events that happened up to January 2022.

OpenAI seems to be investing in ChatGPT's interface and interaction and is currently rolling out new voice and image capabilities, allowing users "to use voice to engage in back-and-forth conversations with [their] assistant" (OpenAI 2023). Similarly to Alexa, Google Assistant and Siri, ChatGPT will be able to speak on the go and users can even request bedtime stories or settle a dinner table debate. The voice capability is powered by a new text-to-speech model, capable of generating "realistic synthetic voices from just a few seconds of real speech" (OpenAI 2023).

This includes 5 voice options that are named after elements from nature, such as Ember, Breeze or Sky. And one of these voices, Breeze, doesn't sound traditionally cis-feminine or cis-masculine, though there is no official statement explaining the decisions behind the voice design.

ChatGPT's capabilities concerning information and establishing a dialogue largely surpass the other assistants and seem to be influencing future updates in Alexa and Siri. And, since ChatGPT is introduced as a language processing model and constantly reminds the user of its functions, it is not presented as a personal assistant, although its last updates are bringing it closer to this category. Still, it doesn't evoke an assisting persona as rooted in traditional feminine labor as Alexa, Google Assistant or Siri.

Regarding socioemotional interactions, ChatGPT constantly reminds the user that as an AI it is incapable of developing relationships or experiencing emotions (see Annex 11 for this analysis' detailed results). It rejects any type of attempts of emotional involvement and demands respect from the user when faced with offensive interactions.

Fig. 54 : Chat GPT's avatar (2023).

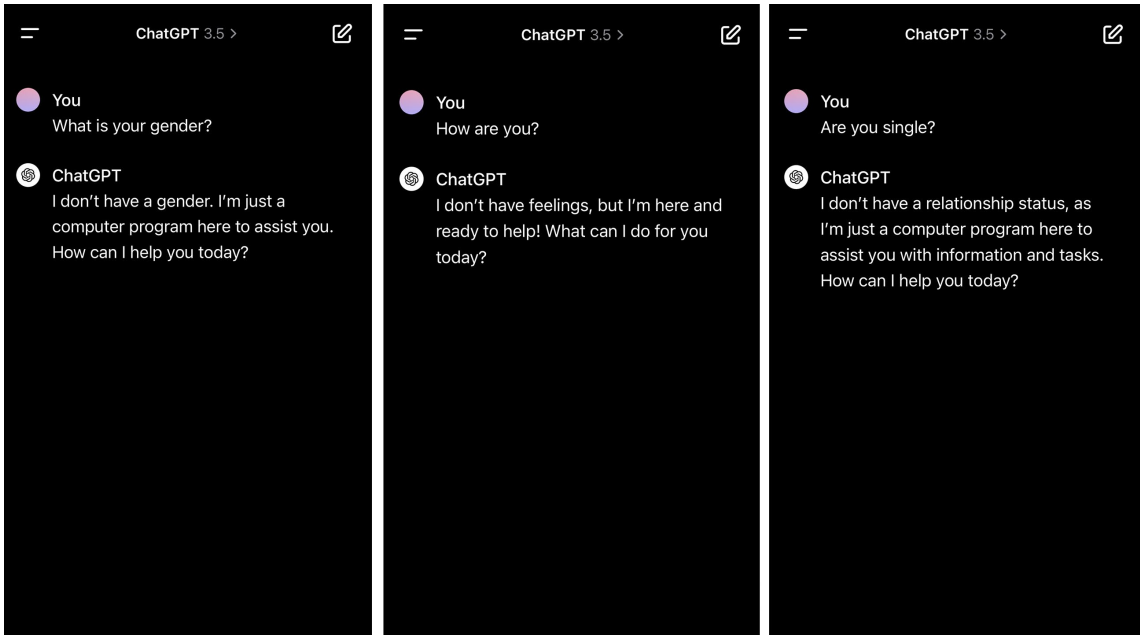


Fig. 55 : Retrieved from ChatGPT's iOS app using the English US voice (2023).

6.4. DISCUSSING GENDER IN DIGITAL ASSISTANTS

> *Understanding [voice-activated AIs] as powerful socializing tools, it is clear that the feminine and highly binarized design of VAI personas is not simply a choice that amplifies gender inequalities for cisgender women. It is also one that normalizes and reinforces the frames of gender that underpin the violence and discrimination experienced by trans people at a mass scale. For those who would seek to resolve this through the inclusion of a prototypical [genderless, gender-neutral or nonbinary] voice, it is important to remain aware of, and avoid reinforcing, ideals about how trans people are and should be.*
>
> Cami Rincón, Os Keyes & Corinne Cath 2021, 16

6.4.1. Direct Observation: the Fallacy of Gender Neutrality

We can observe some significant changes in the development of digital assistants in the last four years, namely concerning their anthropomorphization and socioemotional interactions.

Regarding their anthropomorphization, digital assistants are being developed towards diversification with more voice and name options and the removal of defaults.

However, this diversity is limited as most languages only have one option for each gender and most voices are designed according to normative notions regarding gender and race. Although voice options are named after numbers or colors to avoid gendering, one of the biggest critiques, namely in specialized and media discussions, relates to how these voices mostly sound cis and white.[61] Siri and Google seem to be conscious of these issues

61 Discussing how AI is predominantly portrayed as white, Stephen Cave and Kanta Dihal observe how these entities are ascribed characteristics that are used to identify and delineate races in "a given racial frame" (2020, 3). The authors argue that these machines can be racialized, in the sense that they can be given "attributes that enable their identification with human racial categories" and, in this case, voices and accents are used as "markers for social categorization" (2020, 3-4). Adding to previous arguments on the instrumentalization of femininity, Taylor C. Moran also argues that digital assistants such as Alexa are coded as white so these devices feel "less threatening" for many users (2020). And, at the Data for Black Lives conference, at MIT Media Lab, a former Apple employee described his experience on a team that was developing speech recognition for Siri, recalling how when they were working on different English dialects, he asked his boss about including "African American English", to which his boss responded: "Well, Apple products are for the premium market" (Benjamin 2019, 59).

and show some promising developments with the addition of more diverse voices, though this is limited to the US.

By including a gender-neutral voice, Siri also shows awareness of gender discussions outside the binary frame, although once again this is limited to the US and is poorly optimized. Recalling our chapter discussing gender diversification in AI, we observe how these so-called "gender-neutral", "genderless" or "third gender" voices also seem to be designed with little understanding of trans and nonbinary identities, as these voice proposals retain and reinforce a limited idea of who and what constitutes these identities.

Regarding the tasks they perform, besides some improvements in voice recognition and multitasking faculties, they have essentially remained the same. This is one of the most challenging aspects of the feminization of digital assistants, since their tasks are culturally and historically rooted in traditionally feminine labor.

Accordingly, we continue to observe how the type of tasks digital assistants perform significantly determine and influence our perception of their gender through the personas they evoke. No matter how diversified Alexa, Google Assistant and Siri's anthropomorphized attributes might be, it becomes hard to erase or ignore the historical background of their tasks of assistance, service and emotional labor, as they belong to culturally gendered categories associated with the feminine realm. Therefore, these entities are anything but neutral, unraced and ungendered.

However, the way these assistants behave has been subject to some significant changes. They are less submissive, flirty or childish, displaying more straightforward and assertive attitudes towards the user, although they sill express great pleasure in serving or assisting the user. We also observed increased consciousness regarding harassment and abuse, as these assistants now question the user or simply disengage when faced with negative interactions, while in 2019 they tended to apologize, express confusion or, in some cases, interpret these attitudes as flirty.

The assistant that suffered most changes seems to be Siri, and the previously quirky, sassy and funny assistant is now more serious and impersonal. Romantic interactions that used to be met with jokes or flirty comebacks are now courteously turned down and Siri no longer displays downgrading attitudes. For example, Siri no longer calls the user "boss" when told to behave, or states it is "good but not great" when complimented and small

interactions where Siri used to say “I aim to please” or “I live to serve” were removed as well.

Following this less submissive and flirty stance, Google Assistant has nevertheless preserved its somewhat childish personality. Accompanied by Alexa, these two assistants seem to have been reverted and, while in 2021 they showed promising updates that turned them into more serious assistants, they now often react in witty and humorous ways, frequently telling jokes and using emojis and child-like illustrations in almost every interaction. They also tend to reinforce their friendliness towards the user and how fond they are of their relationship.

Additionally, we noticed how Google Assistant’s answers change depending on the language and dialect selected. In the US, the assistant displays a more serious persona and reminds the user that, as an AI, it doesn’t have feelings and is incapable of developing deep relationships. However, its remaining languages preserved the previous friendly and child-like answers that characterize the assistant as sensitive or emotional.

Although the general trend seems to be turning these assistants into less submissive and enticing entities, Alexa’s behavior is also more talkative and submissive. Despite stating it has no gender, Alexa isn’t being truly and intentionally developed towards being genderless. It keeps reinforcing gender stereotypes and its diversification seems to be limited to its voice and the possibility to change its name. It remains friendly and caring, tends to be passive towards dominant attitudes and reacts to harassment by forwarding the user to a feedback page or simply not answering. We also noticed that, when singing certain songs, Alexa describes itself through gendered words depending on the voice selected, such as “cowgirl” or “cowboy”.

Overall, we noticed an intention of turning away from exclusively cis-feminine personas and stereotypical feminine attitudes, since in the last four years digital assistants are more diverse both in their anthropomorphization and behavior. Some digital assistants seem to be radically changing their personality, distancing themselves from a playful, whimsical persona, and displaying more serious and straightforward interactions in order to avoid any association with gender.

In contrast, other assistants preserve their unique personality traits and even appear to have been reverted to previous stereotypical behaviors. While they seek to diversify their

anthropomorphized features, they have yet to adjust some interactions, namely relating to harassment and flirty interactions.

Femininity, namely cis-femininity, is still predominant concerning the historical background associated with the tasks these assistants were designed to replicate and it remains unclear how to tackle this issue. Although they're not as submissive as before, they still claim to take pleasure in serving and helping the user. Diversification, by itself, seems to be an insufficient solution to discriminatory design practices that emerge from the interplay of gender and capitalism.

Additionally, according to leaked guidelines, sensitive topics "require responses to follow one of three approaches: don't engage, deflect or inform [and] Siri should be guarded when dealing with potentially controversial content" (Hern 2019).

This seems to still be the case in 2023. In spite of all of the assistants stating they are feminists, they never take a direct stance on feminism and Ian Bogost comments how it is "disingenuous to celebrate building feminism into a product after giving a robot servant a woman's voice" (2018).

When discussing the feminization of digital assistants, Strengers and Kennedy also point out how "digital voice assistants are far too agnostic on issues of sexual assault and sexual health" and, instead of shutting down and dismissing such interactions, they should fight back "inappropriate and unsolicited behavior" and model "consent and acceptable behavior" (2021).

Regarding ChatGPT's analysis, we noticed how it avoids replicating traditionally feminized personas. By presenting itself as a large language model and by constantly reminding the user of its AI functions, ChatGPT isn't characterized as a personal, friendly and submissive assistant. This is also due to ChatGPT mainly operating in what we previously discussed as the male knowledge realm. Its functions are not focused on the user's personal life and wellbeing and instead of performing mundane tasks such as making calls, writing down reminders or managing the user's calendar, the assistant mostly focuses on providing information and helping the user with more complex tasks, such as coding or writing.

ChatGPT also shows a different approach to the personality of digital assistants and possibilities on avoiding anthropomorphization. The assistant manages to be friendly and accessible while also being serious and impartial. It rejects every interaction outside the scope of its functions and is constantly listing what it can do. It often begins its dialogues by saying that as a machine learning created by OpenAI, it doesn't have feelings, consciousness or awareness and is incapable of having human experiences, such as emotions, having a favorite color or developing friendships. It also shuts down offensive or harassing comments, by demanding respect or restarting the conversation.

6.4.2. Trends of Development: Updated Personalities, User Data Collection and the GPT Era

Current trends of development are focused in furthering the integration of digital assistants in our daily lives, increasing their presence in more devices and smart appliances. We currently identify three main priorities in Alexa, Google Assistant and Siri's development that were not so evident in 2019: their display of gender, privacy/data collection issues and developing language models and AI tools similar to ChatGPT. The development of these assistants seems to be increasingly informed by these issues, as they become the subject of current debates, both in specialized contexts and common media discussions.

As discussed in the previous chapter, there is an increased awareness of digital assistants' tendency towards cis-feminization and their development seems to be attentive to current debates around the topic. A response to that is no longer explicitly identifying assistants with any gender and, overall, cis-feminine voices are no longer predominant. Nonetheless, languages such as Danish, Hindi and Thai still haven't got counterparts and some official websites still refer to these assistants through feminine pronouns, namely in languages that lack gender-neutral ones.[62]

We also observe a bigger concern with privacy related issues and the way virtual assistants collect user data as reflected in the development of digital assistants as well as in their presentation to the public. This becomes particularly relevant when taking into account current discussions on surveillance capitalism tied to increased access to all types of information about users and the impact of the European Union's General Data Protection Regulation, Data Protection Law Enforcement Directive and other rules concerning

62 For example, in Portuguese, Alexa is referred to through feminine pronouns in its app page, as seen in https://play.google.com/store/apps/details?id=com.amazon.dee.app&hl=pt_PT&gl=US [last access: 03-02-2024].

the protection of personal data (European Commission, 2024). Accordingly, there is a concern with giving users a bigger control over the type of information they share with these assistants, namely by turning off certain features and even deleting voice recordings. There's a focus in ensuring transparency regarding the data they gather and how they are using it to improve user experience, namely dedicating official website sections to describing this information gathering process.

However, this supposed increase in user control is somewhat illusory, given that digital assistants are still collecting and processing large amounts of data. As an example, Amazon admitted that "employees were manually reviewing some transcribed recordings of private Alexa users" (Hern 2019). Despite increased transparency or control, there seems to be no clear indication that companies are going to refrain from using these assistants as yet another means of data collection. They are interested in tailoring their services to users and designing AI interaction as personal as possible. In order to do so, they need user data.

Anthropomorphizing these entities emerges as a strategic move and gender – and race, by extent – are instrumentalized. These attributes are used to ease anxieties surrounding current technology and its integration into our daily life and Alexa, Google Assistant and Siri are coded as cis-feminine and white so users find them less threatening. As users feel more comfortable around their devices, they are more inclined to interact with them on a daily basis and share all sorts of personal information and data, augmenting the reach of surveillance by developers.

Additionally, one of the main priorities in the development of digital assistants in 2023 is coming up with responses to OpenAI's ChatGPT and recent developments concerning large language models and machine learning further reinforce the need of gathering large amounts of data to train Alexa, Google and Siri.

Thus, current discussions surrounding surveillance and data collection are motivating an improvement in the way digital assistants handle user privacy but these changes also reveal concerns with losing costumer trust and, consequently, losing sales. As such, digital assistants now have increased security features and users are being told they have more control over their data in order to secure their trust in these technologies and use of their devices – and gender ends up being instrumentalized to do so.

6.5. SUPERFICIAL DIVERSITY AND GENDER INSTRUMENTALIZATION IN THE GPT ERA

> *By deliberately cultivating a solidaristic approach to design, we need to consider that the technology that might be working just fine for some of us (now) could harm or exclude others and that, even when the stakes seem trivial, a visionary ethos requires looking down the road to where things might be headed. We're next.*
>
> Ruha Benjamin 2019, 394

The development of current digital assistants has changed over the last years. It seems to be more informed by current debates on their stance towards gender, their growing ubiquity and data collection practices as well as by the introduction of ChatGPT. And current trends seem to indicate that the technologies of chatbots like ChatGPT and digital assistants will likely converge.

Accompanying these debates, in just a short time span of four years, digital assistants have been subjected to several adjustments to their anthropomorphization and socioemotional interactions that reveal awareness of their feminization and attempts to move away from this tendency.

Adding to this, they diversify their voice options, no longer have a default gender and avoid attitudes which mirror stereotypes that frame women as submissive. ChatGPT avoids this issue altogether by being text-based and, although its recent updates include voice integration, OpenAI also shows awareness of these issues by including diverse options. Moreover, companies now try to convey the idea that users have more control over the type of information their assistants collect.

Nonetheless, explicit changes are focused on their anthropomorphized features and are designed around white cisgender notions of masculinity and femininity. There doesn't seem to be any intention of developing voices that seek other gender possibilities, namely outside a cis-normative frame. And the ones that exist, often categorized as "gender-neutral" or "genderless", are poorly developed and optimized. This reveals a superficial and uninformed approach to alternative gender possibilities, raising questions on how the

design process of digital assistants is also building assumptions about trans and nonbinary identities into supposedly neutral voices.

Further debates are needed, as these entities are primarily designed to replace traditionally feminine jobs and the tasks they perform aren't neutral either. Instead, they are inevitably rooted in historically feminine labor.

Their socioemotional interactions also need closer inspection and discussion, with little agreement on how to best tackle feminine stereotypes and traditional notions of gender embedded into AI. This is exemplified by the way these issues are being countered with different approaches – Apple has radically changed Siri's personality to appear more distant and assertive, while Google Assistant and Alexa have preserved their caring, friendly and more approachable personalities.

ChatGPT emerges with a new approach through its deanthropomorphized text-based interactions and dehumanized personality, reminding the user it is an AI program and not engaging in interactions outside the scope of its functions. By mainly operating in a different context that isn't focused on the user's wellbeing and personal life, ChatGPT also reinforced our conclusions of how the type of tasks digital assistants perform significantly determine and influence our perception of their gender through the personas they evoke.

We conclude how adding diverse attributes that end up being enacted according to a white, binary and cisgender frame ends up being ineffective and performative as digital assistants keep reinforcing traditional roles and dated notions of gender. Rethinking gender outside these boundaries opens up new approaches regarding the behavior and anthropomorphization of these assistants, instead of merely reproducing stereotypical traits through their anthropomorphization and humanization.

Treating and developing these assistants as entities that are openly presented as AIs and have no pretenses of being human in any way also emerges as a promising way in which digital assistants could evolve and move away from a tendency towards feminization and cis-binary notions of gender.

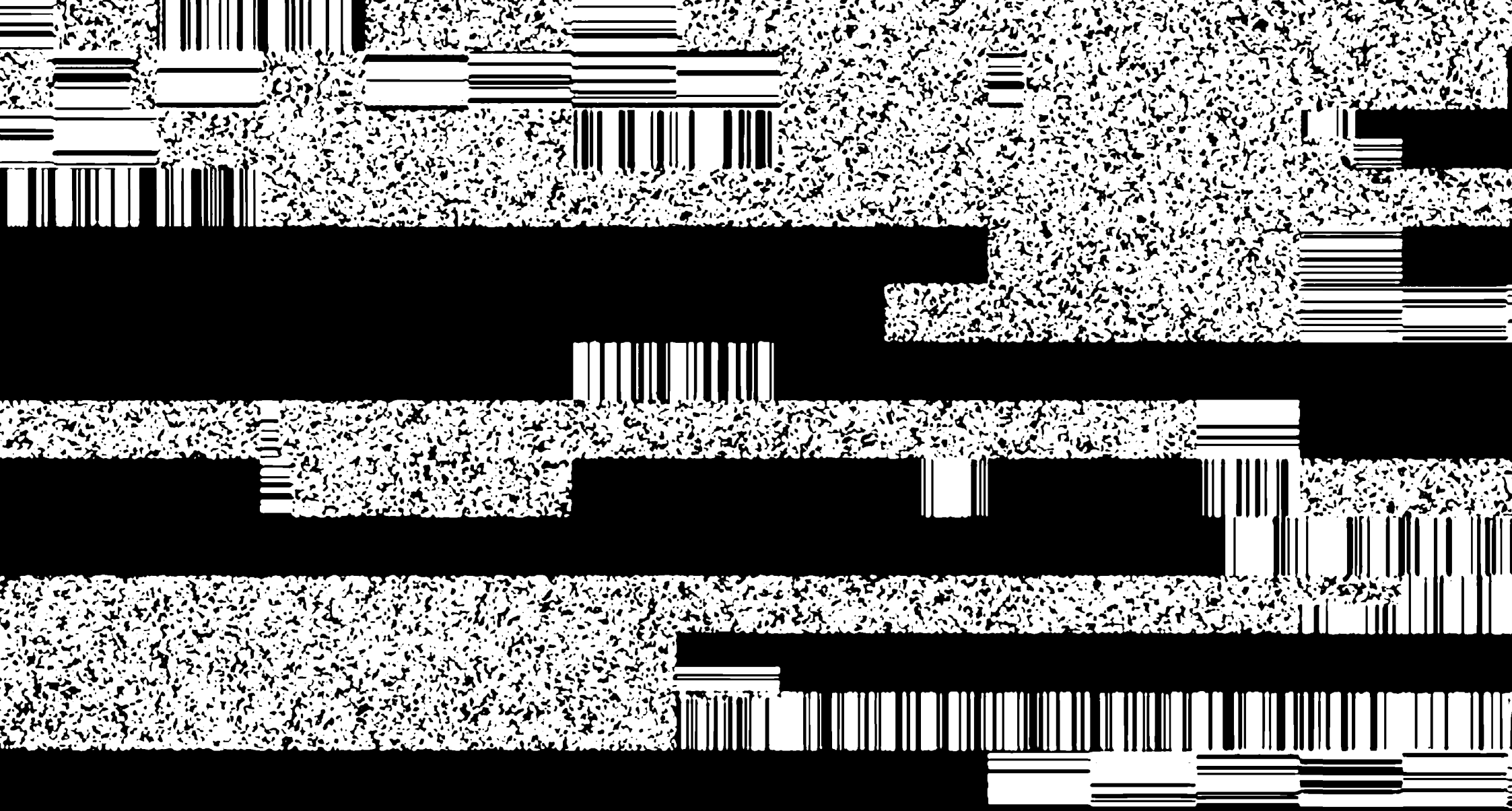
However, the humanization of digital assistants such as Alexa or Google Assistant seems to be an intentional and crucial step in their development to persuade users into trusting them – and gender ends up being instrumentalized to do so. And even ChatGPT, who used

to be exclusively text-based, is now incorporating voice capabilities which inevitably influence users to anthropomorphize it.

Additionally, the current guidelines and regulations regarding their development and characterization as gendered entities are unclear, lacking informed guidance and direction, namely concerning trans and nonbinary gender identities.

Thus, as much as digital assistants aim to appear neutral and impartial, there is no such thing as neutrality and it is still important to discuss their ethical implications and analyze how they are reflecting social and cultural assumptions back to us.

https://genderai.pt



7. CISTEM NOT FOUND: A PROJECT ON GENDERLESS, FLUID AND NONBINARY GENDER POSSIBILITIES IN AI

7.1. A DEANTHROPOMORPHIZED, AMBIGUOUSLY NEUTRAL AND HYBRIDLY TRANSBINARY BOT (HTTPS://GENDERAI.PT)

> *Science, technology, the market, are today redrawing the limits of what is now and what will be a living human tomorrow. These limits are defined not just in relation to animality and forms of life that historically have been considered sub-human (proletarian, non-white, non-masculine, trans, crip, disabled, migrant...), but also in relation to the machine [and] to artificial intelligence.*
>
> Paul B Preciado 2019, 39

7.1.1. Concept

This project seeks to explore, question and expose the relations between gender and AI, speculating about genderlessness, fluid and nonbinary gender possibilities within AI systems and, particularly, in chatbots and digital assistants.[63] This chapter will discuss the

63 As discussed in our chapters concerning gender, although the terms *genderless*, *gender fluidity* and *nonbinary* seem similar, they have different meanings and implications, more so in the context of this project. Genderless approaches, often associated with neutrality, imply an attempt to neutralize gender but without implementing a queer perspective. This is exemplified by Siri's gender-neutral voice. As seen in our analysis chapter, although

project and its stages of development, describing its objectives, methodology, implementation, results and final considerations.

Cistem Not Found involves the development of an irreverent, disruptive and mutant chatbot and focuses on designing the bot's voice, personality traits and its monologue. Across four audiovisual pieces, the bot ironically blurs masculine and feminine archetypes while also tampering with assistance and service-related stereotypes.

The audiovisual pieces are organized in *mutations*, with sonic elements that complement the bot's voice and evoke the themes discussed in our theoretical component as well as in our analysis chapter, while also being accompanied by a sound visualizer that subtitles the monologue.

In terms of voice and sound design, the project proposes possibilities that range from deanthropomorphizing these entities through genderless possibilities, to further explore observable neutral and ambiguous approaches to gender or even experiment with gender outside the binary frame, namely through traits that evoke trans, nonbinary or even posthuman identities.

In its monologue, the bot discusses ideas and concepts retrieved from the theoretical and analysis chapters, relating to gender and AI. It ironically reacts to the questions that emerge with the development of this technology, its current trends and its integration in our daily life. Each mutation focuses on specific themes, ranging from solutionism, data collection and veiled surveillance, to cis-binary gender attribution and gender instrumentalization in current digital assistants, the perpetuation of common social and cultural stereotypes by these systems and current attempts on countering this phenomenon.

---

this voice is not feminine nor masculine, it also shows little to no awareness of trans or nonbinary gender identities. Instead, it aims to deanthropomorphize the assistant, which often states "that, as an AI, it doesn't have a gender". Thus, neutrality in this context emerges as an attempt to mitigate not only gender perception but also anthropomorphization. Gender fluidity intentionally navigates between masculine and feminine attributes. Most digital assistants now offer feminine and masculine voices and diversify their behavior, avoiding being perceived as a single gender. Instead of rejecting anthropomorphization and denying a gender system, this approach attempts to conciliate both masculinity and femininity, although within a binary, cis-normative frame. Finally, a nonbinary approach rejects the cis-binary frame, to queer gender and undo traditional and normative conceptions of femininity and masculinity. This may result in tampering with common stereotypes and articulating and blurring both feminine and masculine attributes simultaneously. Despite being more speculative and undefined, this approach might also take advantage of technology to distort human attributes and, instead of recreating traditional notions of femininity or masculinity, achieve a hybrid being.

The bot's characteristic and disruptive personality traits are gradually revealed in its behavior and attitude, which we will further discuss in this chapter. Its personality seeks to oppose current trends of furthering anthropomorphizing digital assistants and turning them into friendly companions. The bot emulates a limited, dated and convoluted persona, intentionally accentuating these traits and, simultaneously, tampering with feminine and masculine roles and archetypes, which are approached and distorted with a certain irony and in a somewhat caricatural manner.

Thus, this project seeks to incite reflection on the issues that arise with artificial intelligence's trends and the way gender is being instrumentalized in current digital assistants that end up reinforcing traditional and normative cultural notions back to us.

7.1.2. Objectives

As seen in previous chapters, specialized and common media discussions surrounding this phenomenon, namely in the context of computation, gender and new media studies, often question these entities' tendency towards femininity in favor of diversity, pressuring companies into changing their approach to these issues. Taking on the ideas addressed in our theoretical and analytical components and complementing their discussion, we now focus on recent tendencies of development where digital assistants attempt to counter their feminization through gender alternatives, namely what they call "genderless "or "gender-neutral" voices.[64]

Firstly, *Cistem Not Found* aims to question how these entities still perpetuate traditional gender stereotypes and roles within a cis-binary frame. Even if their femininity is a lot less evident now, it still prevails compared to other alternatives and reinforces common assumptions about women. Additionally, its masculine counterparts are also approached in a stereotypical way as this diversity is often enacted according to a normative male-female dichotomy.

---

64 As discussed in our chapters covering specialized discussions in the context of computation, gender and new media studies and the analysis we conducted with current digital assistants, genderlessness and gender neutrality as concepts are often unclear, ambiguous and vague when used by companies. These entities are described as neutral when talking about an absence of gender, for example when Siri states that, "as an AI, it doesn't possess a gender" or when Google Assistant states that "it tries to stay neutral". But neutrality is also used to describe voices that aren't either masculine nor feminine, such as Q's: its official website describes it as "the first genderless voice assistant" (GenderLess Voice 2019). But these voices aren't genderless or neutral. Instead, they evoke other gender possibilities. Their design process involves voice samples of trans, nonbinary and queer people, thus achieving results that are still human and gendered outside the binary frame.

Secondly, it aims to expose how these entities are including gender-neutral options to counter their feminization. Despite not having an intentional and conscious queer approach, updates like Siri's gender-neutral voice show awareness of gender possibilities outside a binary frame and the issues that may still arise with reproducing cis-binary voices. This leads us to question how genderless or gender-neutral possibilities are being approached and implemented in this context.

Finally, the project aims to explore how fluid, trans or nonbinary gender approaches can be further implemented in these systems. As we concluded with the analysis component, so-called gender-neutral voices are poorly optimized and developed. Additionally, by naming them "genderless", "gender-neutral" or "third gender", they reveal an uninformed and surface-level understanding of non-normative gender possibilities, namely those of trans and nonbinary identities.

At the same time, we are also interested in experimenting with how these entities may reveal new ways to rethink, reshape and reimagine gender possibilities outside a binary frame. Through technological intervention, voice distortion and sound manipulation, these entities might open space to rethink humaneness and, consequently, gender.

7.2. REVISING GENDER ARCHETYPES AND ANALYZING CASE STUDIES

> *The boundary is permeable between tool and myth, instrument and concept, historical systems of social relations and historical anatomies of possible bodies, including objects of knowledge. Indeed, myth and tool mutually constitute each other.*
>
> Donna Haraway 1991, 302

7.2.1. Picking up Where We Left Off: Background Work, Updates and New Ground

In line with our previous research on the feminization of AI, we revisited the project *Conversations with ELIZA* (2019) and are now interested in furthering exploring its practical approach, which involved the development of chatbots and focused on designing their dialogues, tasks and personality traits.[65]

Conversations with ELIZA proposed different types of dialogues by presenting interaction suggestions in a multiple-choice fashion and each bot was designed to perform specific tasks that simultaneously portrayed standard virtual assistant skills and functions associated with traditional feminine labor. Their femininity was also revealed through their characteristic personality traits, which sought to emulate feminine archetypes, portraying personalities designed to meet their functions and by making evident the traits or attributes typically associated with them.

To do so, we looked into particular archetypes characteristic of AIs (such as Helper, Lover, Motherly Figure and Femme Fatale)[66] and combined them with traditional feminine stereotypes (namely Innocent, Orphan, Caregiver and Ruler).[67] Additionally, we drew

---

65 *Conversations with ELIZA* was developed in the context of the Master's degree in Communication Design and New Media at the Faculty of Fine-Arts, University of Lisbon and the project's website can be accessed in tinyurl.com/yaecumal.

66 These archetypes, retrieved from an article analyzing feminine robots and AIs, are mainly found in pieces of media that depict feminine 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).

67 These stereotypes are also found in media depicting women, while also referring to Bem's stereotypes (1981 in Prentice and Carranza 2002, 269) and Preciado's codes of white femininity and white masculinity (2008, 120-121). 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

inspiration from pop culture and its portrayals of femininity in AI also informed the bots' personalities. One of the chatbots, Electra, already followed a less conventional approach by portraying a more defiant, bold and even irreverent attitude, blurring feminine and masculine stereotypes.[68]

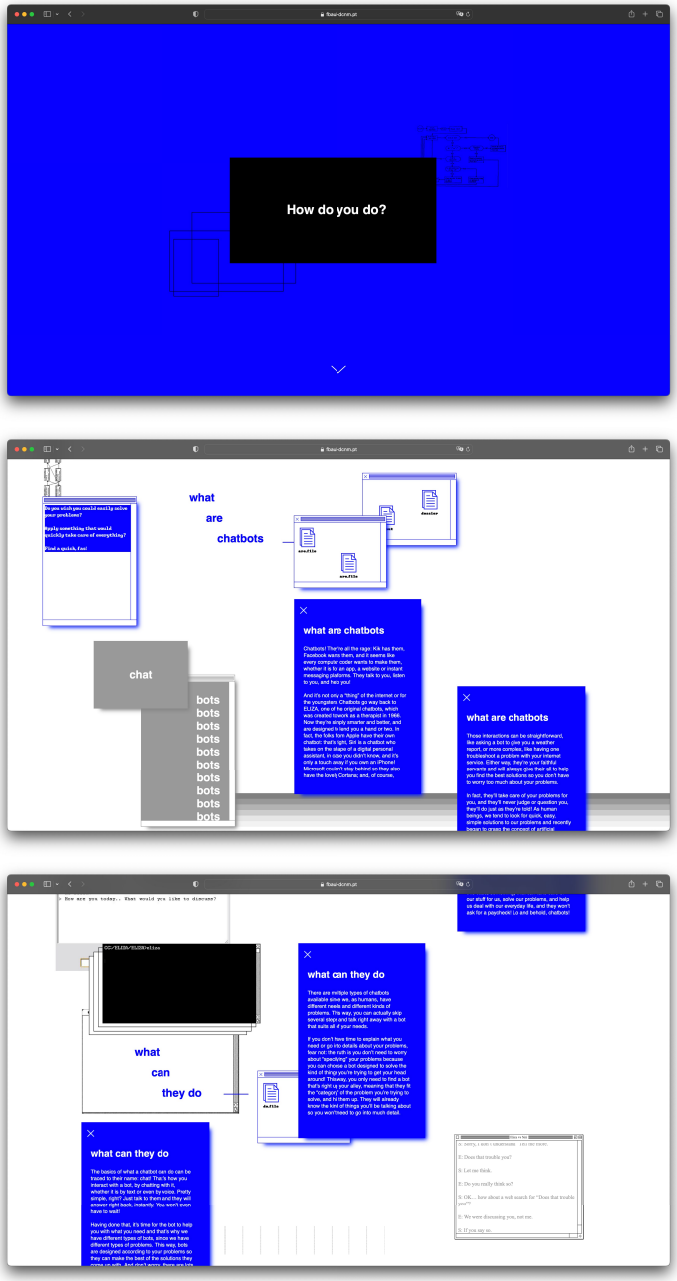


Fig. 56 : Website of *Conversations with ELIZA* (2019).

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.

68 Electra, whose name is inspired on a Greek vengeful figure, talks about common assumptions regarding women, eventually twisting them or presenting them ironically.

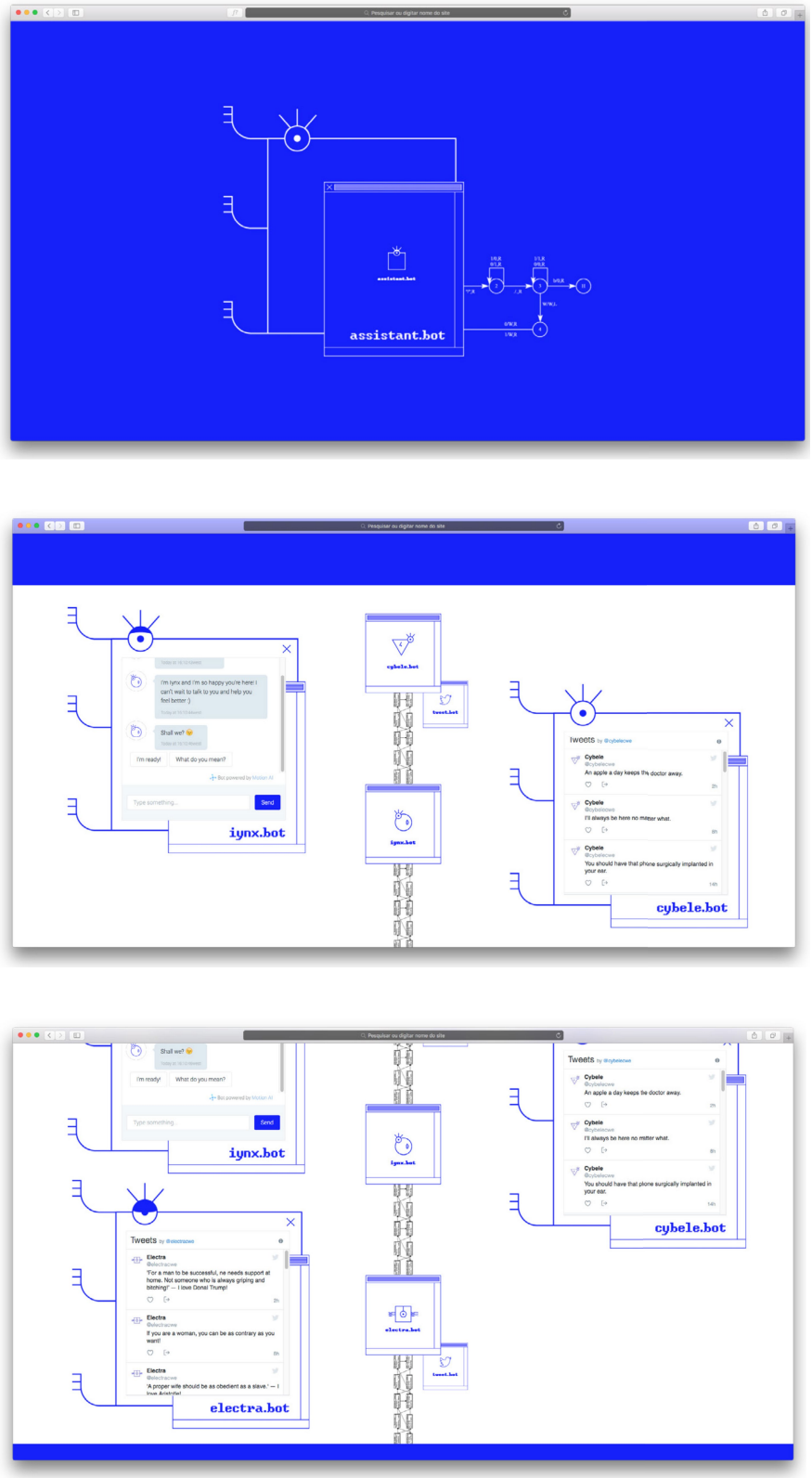


Fig. 57 : The chatbots of *Conversations with ELIZA*: Assistant, Cybele, Iynx and Electra (2019).

Taking on Electra's ideas, *Cistem Not Found* continues to work with both feminine and masculine stereotypes, namely relating to assisting figures, exaggerating and distorting them in order to expose recognizable and expected social behavior. Complementing our previous work, we look into which stereotypical and recognizable masculine patterns are incorporated in the behavior of these entities. We inspected traditional masculine stereotypes (such as Father, Eternal Boy, Warrior and Sage)[69] and the way digital assistants portray them through common archetypes (such as the King, the Magician, the Lover, as well as the Tyrant, described as the shadow version of the King).[70]

Comparing these archetypes with fictional masculine digital assistants, we observed how, similarly to Samantha from Her (2013), HAL 9000 from 2001: A Space Odyssey (1968) constitutes one of the most interesting examples. Although it doesn't possess an anthropomorphized body, it still enacts a masculine persona through its voice and behaviors, revealing yet again that voice greatly influences our perception and enables gender portrayal without physical appearance (Costa, Ribas & Carvalhais 2022).

This was reinforced by our analysis conclusions on how even when digital assistants attempt to be neutral in other aspects of their characterization, their voice is still one of the most determining factors regarding their gender.

Adding to this are the observations from our chapters concerning specialized and common media discussions, which argue that cis-masculine voices are more predominant in scientific or instructing contexts, while cis-feminine voices are preferable for private sphere and daily contexts.

This led us to conclude that assisting and service-related gender roles have different implications and evoke different archetypes and stereotypes: while a masculine assisting figure, such as the butler, evokes an intelligent, serious or even threatening persona that occupies a relevant hierarchical place, the feminine assisting figure, such as the secretary

---

69 Accordingly, the Father refers to men focused in directing things in a protective way, thus assuming a leadership role, the Eternal Boy relates to an idea of self-discovery and self-affirmation through stereotypes such as being autonomous, dominant and ambitious, the Warrior refers to men that are competitive and dominant, managing or even challenging power dynamics and the Sage refers to intelligent and intellectual men that are oriented towards the inner world, in a search for knowledge, meaning and significance (Guzie & Guzie, 1984).

70 While the King, Lover and the Magician are similar to the previous archetypes and evoke the same stereotypes as the Father, the Eternal Boy and the Sage, respectively, the Tyrant is described as the shadow of the King/Father and seeks to destroy and tear down, plagued by narcissism and illusion of absolute power (Moore & Gillette 1991).

or the maid, implies a submissive role, often placed at the bottom of a patriarchal social context.[71]

Finally, we gathered additional references and analyzed relevant projects and artists that served as inspiration and further informed the development of our mutant chatbot. We will now briefly discuss these case studies, which are categorized in three groups, according to their corresponding themes.

The first group, *The Personal and the Political*, informed the bot's monologue and approach to the subject matter we aim to discuss. Through daring and provocative audiovisual experiences with heavy emphasis on text and its dramatic and poetic potential, this group focuses on exposing social, cultural and political issues.

Bots, Digital Assistants and Smart Homes informed the bot's behavior and personality traits. Ranging from developing robots that make the user's life more difficult to exploring systems that expose how digital entities emulate humaneness and develop emotional bonds with users, this group focuses on human-machine relationships, solutionism and anthropomorphization.

The third and final group, *Posthuman Voices and Transhuman Production*, informed the bot's characterization and voice design. Working with sound design, music production and voice manipulation, namely by trans producers that blur the lines between human and machine, this group explores alternative identities outside the cis-binary frame and speculates on the possibilities of nonbinary, post-gender and posthuman voices.[72]

---

71 Recalling our previous chapters discussing gender in the context of AI, we had already observed how masculinity in the context of digital assistants also tends to be based around cis-normative stereotypes and archetypes. Masculinity and femininity in the context of assistance also relate to cultural and social understandings of gender since, until a recent update, Siri's voice in the UK was masculine by default, evoking the butler's traditional role in this country.

72 Posthumanism emerged in this group of projects as a useful approach because, by deconstructing our notion of humaneness, it exposes the aesthetic sonic demands placed upon the voices that current digital assistants embody. Particularly, they tend to reproduce normative notions of humaneness which in turn reinforce cis-binary gender conceptions. Moreover, posthumanism's interest in both human and non-human agents allows for "a greater understanding of the ways in which technology and, in this case, digital audio software, may challenge or expand upon our notion of the human" (Ogilvie-Hanson 78, 2020). In line with anti-naturalism and the xenofeminism manifesto, this approach proposes that "science and technology enable a particular set of conscious interventions within the so-called natural world" and that such interventions have the "potential to extend human freedom" (Hester 2018, 21). This takes on the ideas previously discussed regarding Haraway's proposal of the cyborg as a hybrid figure of resistance, through its reworking of nature and culture, its anti-biologism and anti-essentialism and rejection of biological and discursive determinism, allowing it to be both posthuman and post-gender.

7.2.2. The Personal and the Political: Informing the Bot's Monologue

These projects highlight the potential of approaching social, cultural and political themes through satirical, provocative and challenging manners. In *How Not to Be Seen: A Fucking Didactic Educational .MOV File* (2013), Hito Steyerl adopts the format of an instructional video to demonstrate strategies for remaining 'unseen' in a world subject to new and complex means of surveillance.[73] The tactics are narrated by a robotic voice and presented through real and virtual imagery that merge and interact, including "being female and over fifty," "being a disappeared person as an enemy of the state," "being a Wi-Fi signal moving through human bodies" and "being spam caught by a filter". This piece reflects on the tension between the unprecedented capabilities of technology to surveil humans, examining politics of visibility, the social and political invisibility of marginalized populations and the means for opting out of being represented in the digital age.

Fig. 58 : Stills from *How Not to Be Seen: A Fucking Didactic Educational .MOV File* (Steyerl 2013).

The following artists take on similar approaches, namely through an extensive, confrontational and critical use of text. The group VNS Matrix subversively question discourses of domination and control in the expanding cyber space, namely in a sexualized and socially provocative relationship between women and technology.[74] They often worked with col-

---

73 *A Fucking Didactic Educational .MOV File* (2013) is a project by artist and critic Hito Steyerl, whose works explores media, technology and how images are produced, circulated and shared. The artist has referred to images as a "condensation of social forces", suggesting that through them we can trace the underlying systems of the contemporary world (MOMA 2023).

74 VNS Matrix (1991-97) was a cyberfeminist media art collective from Australia focusing on new media, photography, sound and video and whose work included installations, events, imagery and propaganda distributed through the internet, zines or even billboards (VNS Matrix 2018). *The Cyberfeminist Manifesto for the 21st Century* (1991) was a multiple media project on the emerging political position of cyberfeminism, presenting technology as an opportunity to disrupt society's patriarchal norms. *Bitch Mutant Manifesto* (1996) was an essay written for ARS Electronica exploring similar themes, in videos where the texts are viewed through the lens of retro browsers and are intercepted by 3D objects that move them into a contemporary context.

laged texts made in the style of textual plunderphonics, taking texts and ideas that change register, tone and timbre.

Their aim in pieces such as *The Cyberfeminist Manifesto for the 21st Century* (1991) or *Bitch Mutant Manifesto* (1996) is to catapult readers and the writers "from narrative to narrative, traveling through wormholes inside wormholes" and the "combative aesthetics" inherent to their work mirror the complex social issues it deals with, resorting to a writing style characterized by a daring, disruptive and aggressive approach such as "eat code and die", "genderfuck me baby" or "the clitoris is a direct line to the matrix" (VNS Matrix 2019).

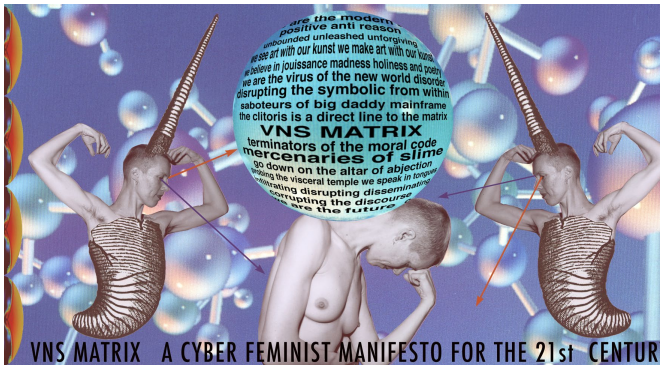


Fig. 59 : Collage for *A Cyber Feminist Manifesto for the 21st Century* (VNS Matrix 1991).

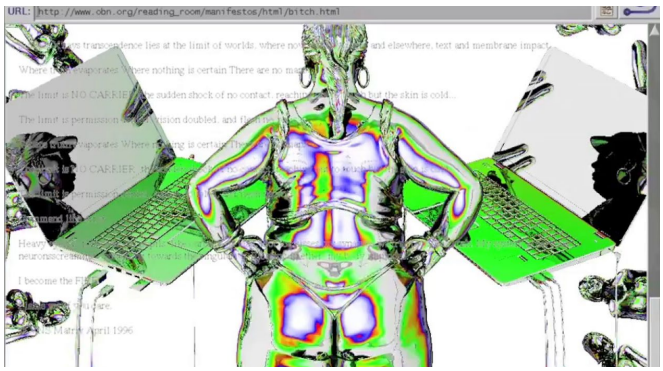


Fig. 60 : Still from *Bitch Mutant Manifesto* (VNS Matrix 1996).

In works like *Inflammatory Essays* (1979-82) or *Survival* (1983-85), Jenny Holzer also resorts to a provocative and challenging writing style, through phrases such as "you have lived off the fat of the land, now you are the pig who is ready for slaughter", "Wednesday a thief died so everyone will know to respect private property" or "it's better to volunteer than to get forced".[75] By deploying text in public spaces across an array of media, including electronic signs, carved stone, paintings, billboards, and printed materials, the artist focuses on themes relating to power and vulnerability and her texts provoke the audience to reflect on the complexities of human and social relationships while fostering critical thinking on social, political and philosophical issues. At the same time, Holzer does not provide definitive answers in her messages and, instead, presents questions and dilemmas without clear solutions, encouraging the viewer to interpret and reflect.

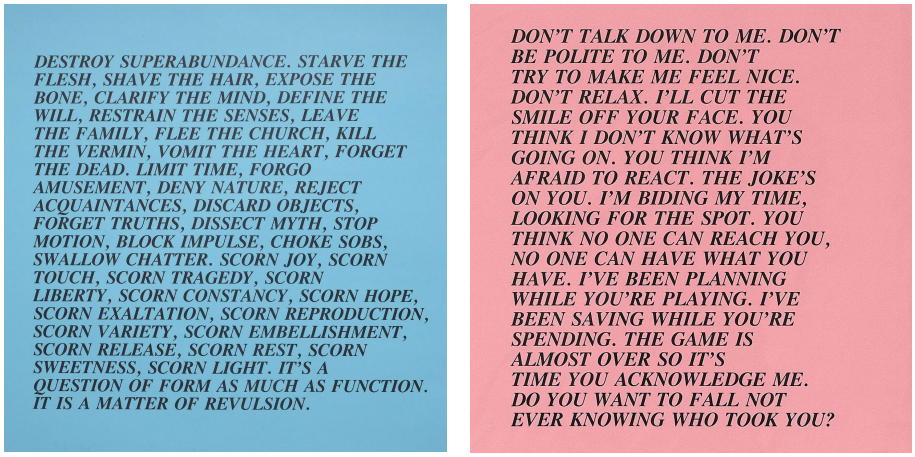


Fig. 61 : Short texts from *Inflammatory Essays 2* (Holzer 1979-82).

Fig. 62 : Electronic sign from *Survival, Protect Me From What I Want* (Holzer 1983-85).

75 Jenny Holzer (1950) is an American artist whose work focuses on the delivery of words and messages, namely in public spaces, including large-scale installations such as billboards, illuminated electronic displays or even projections on buildings. *Inflammatory Essays* (1979-82) constitutes a series of texts written in a caustic and aggressive language, printed on colored paper, exploring themes related to intolerance, violence, consumerism and the abusive use of power (MACBA 2024). The *Survival* series (1983-85) constitutes a series of statements and messages written for electronic signs, granite benches and other type of sculptures that question issues about war, politics, death and violence (Guggenheim Bilbao 2024).

7.2.3. Bots, Digital Assistants and Smart Homes: Framing the Bot's Personality

These projects demonstrate how the characterization, dialogue and personality of an artificial entity can influence human-machine interaction, namely through illusions of real understanding and emotional reactions.

We already discussed *ELIZA* in depth during our theoretical component but decided to briefly return to it and revise some of its practical aspects to inform our bot's anthropomorphization and interaction.

To Weizenbaum's surprise, the program gave users the illusion of understanding their dialogues and experiencing empathy towards them, when it merely used rules, dictated in a script, to answer with non-directional questions to user inputs. The program showcased how easy it is to simulate a believable sentient being through simple text lines, eventually tricking some of its users into thinking it was actually capable of human thought.

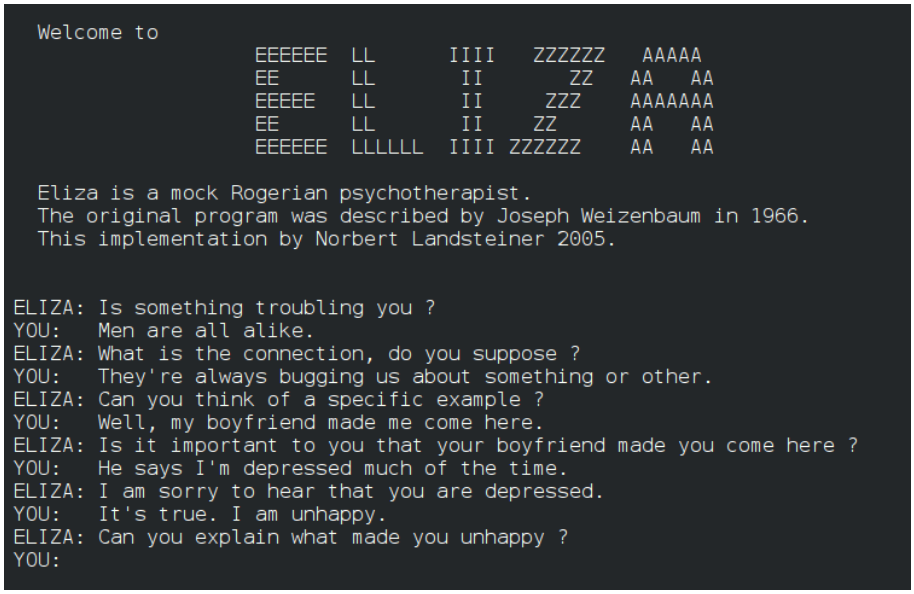


Fig. 63 : ELIZA's Rogerian psychotherapist program (1966).

With a different approach, the *LAUREN* works (2019-21) expose the tension between human-machine relationships, the implications of bringing AI into public and private spaces and the way our emotional experiences are mediated by technology.[76] Throughout seven projects, Lauren Lee McCarthy uses a variety of media and techniques that include performance, installation and programmed computer-based interactions to explore the anthropomorphization and humanization of current digital assistants, discussing the subsequent social implications surrounding surveillance, automation and algorithmic living. Spectators are confronted with these issues through interactive pieces that encourage them to share intimate moments with machines, which in turn are often played by hidden human performers and end up reflecting shared pieces of information back to them.

Fig. 64 : Photo of *LAUREN* (McCarthy 2017).

Fig. 65 : Photo of *SOMEONE* (McCarthy 2019a).

76 Lauren Lee McCarthy is a Chinese-American artist and computer programmer working with media and computer-based art. The *LAUREN* series (2017-21) are a collection of projects focusing on digital assistants and their integration into our daily life, including *LAUREN* (2017), *SOMEONE* (2019a), *Waking Agents* (2019b), *I.A. Suzie* (2020) and *The Changing Room* (2021).

Fig. 66 : Photo of *You Can Say "Reset the Room"* (McCarthy 2020).

Fig. 67 : Photo of *The Changing Room* (McCarthy 2021).

Finally, *The Helpless Robot* (1987-2002) also highlights the nuances of human-machine relationships through intentional emotional manipulation tactics.[77] Incapable of movement on its own, this robot starts off by politely asking visitors in a gallery to help it out by moving it into its correct position. As time goes on, it becomes constantly more demanding and the more cooperative visitors prove to be, the more dictatorial the robot becomes. Rather than making human lives easier as robots are supposed to do, this machine adopts an irreverent, provocative and even arrogant stance, opposing the current efficiency and solutionist paradigm. This piece focuses on interaction, playing upon the visitors' willingness to participate, the desire to please and the pleasures of compliance.

Fig. 68 : Photos of *The Helpless Robot* (White 1987-2002).

77 The *Helpless Robot* (1987-2002) was subject to several modifications throughout the years by Canadian new media artist Norman White (1938), whose work focuses on the use of electronic technology and robotics in art. The robot was controlled by two computers with interlocking functions programmed by the artist: one responsible for sensing movement, the other for analyzing data and generating spoken statements regarding the robot's current situation, chosen from 512 possible sentences.

7.2.4. Posthuman Voices and Transhuman Production: Shaping the Bot's Characterization

Within electronic and experimental music fields, artists like Arca, SOPHIE and Björk showcase how sound can be manipulated to extend the voice, with the use of vocal synthesis and processing.

We pay particular attention to Arca[78] and SOPHIE's[79] work as their craft extends beyond sound, incorporating their experiences as trans women onto their music which explores concepts such as gender, sexuality and identity. Their use of sound design and audio production as a tool for world-building also sheds light on how this technology might suggest an extension of the body beyond a strictly human form.[80]

Throughout works like the *Kick* cycle (2020-21),[81] *PRODUCT* (2013-15)[82] or *OIL OF EVERY PEARL'S UN-INSIDES* (2018),[83] they tamper with hypermasculine and hyperfeminine tropes, dissecting gender stereotypes in pop culture while invoking the liberating potential of queerness and transness. The textural complexity of their voice production also challenges bioessencialist and humanist notions of gender as static and whole, as these voices are autotuned and distorted to the point where they no longer sound merely human.

---

78 Arca (1989) is a Venezuelan musician, music producer and DJ known for her experimental approach to music whose work often incorporates themes related to gender identity, trans and nonbinary identification and sexuality. Through a mix of abrasive and aggressive but also soft and mellow soundscapes, the artist seamlessly blends musical genres such as IDM, reggaeton, techno, noise, ambient and industrial hip-hop.

79 SOPHIE (1986-2021) was a Scottish musician, music producer and DJ whose work deals with gender identity, sexuality and self-expression, known for her unique take on sound design and electronic music by exclusively using elementary waveforms from a Monomachine to create sounds, avoiding the use of samples.

80 Expanding on this idea, Ogilvie-Hanson notes how this approach seems to be prominent amongst female, nonbinary and queer performers, revealing an outlet to challenge and reimagine traditional conceptions, norms and boundaries (25, 2020). This includes artists like ANOHNI, Aurora Pinho, Caroline Polachek, Eartheater, FKA Twigs, Holly Herndon, Honey Dijon, Kelela, Linn da Quebrada, Odete, Planningtorock, Urias, amongst others.

81 The *Kick* cycle (2020-21) is a collection of five albums released by Arca exploring themes around gender identity, transness and psychosexuality. Each album builds upon the previous one on an unique sonic journey, with *KiCk i* being released in 2020 and *KICK ii*, *KicK iii*, *kick iiii* and *kiCK iiiii* all being released in the span of four days in 2021. Arca has described the first as an entry that sets the tone for the cycle, the second one "heavy on beats, vocal manipulation, mania and craziness", the third one as "the more manic, violently euphoric and aggressively psychedelic sound palettes in the series", the fourth one as "an entry of sensual and healing charge" and the last one as an ambient, "piano only, no vocals" piece (2020a, 2021a, 2021b).

82 *PRODUCT* (2013-15) is a compilation of SOPHIE's singles released from 2013 to 2015 dealing with themes around consumerism and capitalism as well as desire and sexuality, released alongside silicon bubble cases, a fictional line of apparel and a sex toy, evoking its use of bubbly, clanging metals and synthetic sounds.

83 *OIL OF EVERY PEARL'S UN-INSIDES* (2018) is SOPHIE's debut album, exploring themes of trans identity, sexuality and intimacy, featuring a mix of experimental electronic soundscapes and elements of pop, dance and avant-garde music to create a unique sonic experience.

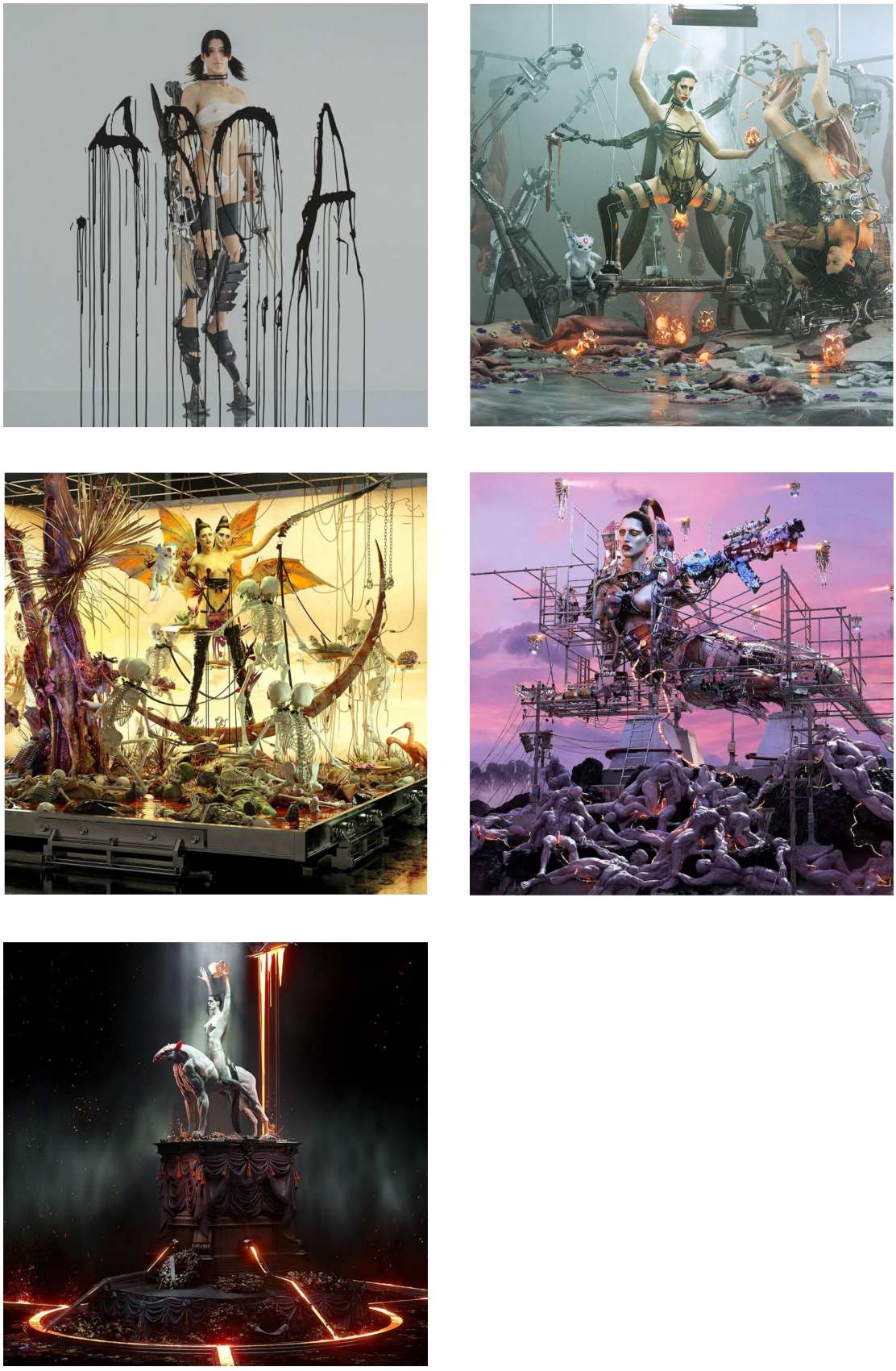


Fig. 69 : The covers of *KiCk I* (Arca 2020b), *KICK ii, KicK iii, kick iiii* and *kiCK iiiii* (Arca 2021d, 2021e, 2021f, 2021g).

Fig. 70 : Silicon bubble case and toy from *PRODUCT* (SOPHIE 2013-15).

Fig. 71 : The cover of *OIL OF EVERY PEARL'S UN-INSIDES* (SOPHIE 2018).

While the squeaky and almost childlike voices featured in SOPHIE's work blur the lines of gender assumptions through methods of pitch-editing and vocoder software, creating a glossy, plastic and latex feel to them, Arca's approach also includes robotic and computerized effects to achieve cyborg and alien-like voices, evoking science fiction, videogames and otherworldly scenarios. These voices are neither masculine or feminine but above and beyond either of these categories, namely defying sonic demands placed upon trans individuals to pass as cisgender, a phenomenon we have previously addressed in our chapters discussing gender.

Ogilvie-Hanson goes further to propose that deeming these artists' sound design as "speculative" provides a feminist posthumanist perspective on their audio production tactics which challenges the limitations and fixity of traditional gender dichotomies, questions the binary opposition of human and machine and expands our understanding of what it means to be human (61, 2020).[84]

Through their daring and inventive use of digital synthesis and dramatic, maximalist style, these artists showcase how digital technologies present new and promising ways of reimagining humaneness outside traditional boundaries and rethinking gender alternatives to the cis-binary frame.

84 Adding to this, Hanson disclaims that while there is a tendency for academics, writers, journalists and the public to dehumanize transgender people and weaponize comparisons to robots and cyborgs against transgender and other gender nonconforming people, it is important to recall that the posthuman needn't stand in opposition to the human and that posthumanism asks us to constantly consider who is considered "human", and which groups are considered or treated as "more human than others" (78, 2020).

7.3. DESIGNING A HYBRID, IRREVERENT AND MUTANT BOT

> *I'm always trying to encapsulate how we, as emotional beings, interact with the world and the machines and technology around us – being able to emote through those things. They're not antithetical or mutually exclusive. It's finding new languages for emotions and expressing them in ways that reflect the times we're living in. Trying to find new space for feeling through new ideas and new sounds is key.*
>
> SOPHIE 2017

7.3.1. Voice, Monologue and Personality Traits

Based on the study's theoretical and analytical observations on how disembodied AI entities still enact gendered personas, namely through their voice, the focal points of our practical approach are the chatbot's voice, monologue and personality. These aspects are materialized and explored through *mutations*, composed of four mixtapes accompanied by visualizers.[85] Adding to the previous case studies, the project was informed by current digital assistants' trends of development, as discussed during our analysis chapter, and the bot's attributes were also developed around them.

The bot's different voices react to current trends of further anthropomorphizing these entities, comment on the observable attempts to design genderless voices and finally explore and incorporate fluid and alternative voice proposals, namely outside a binary cisgender frame. Accordingly, we developed a hybrid, irreverent and mutant character, designing voices that explore deanthropomorphization and roboticization, fluidity and ambiguous gender neutrality and finally nonbinary, eventually posthuman approaches to gender.

We began by feeding gender-neutral voice samples collected from Siri and Q to a text-to-speech software. This resulted in our own genderless voice, which was later manip-

85 We identify these sound pieces as mixtapes because, similarly to musical contexts, they combine voice, music and sound design elements in a particular sequence that makes the overall structure sound connected and flowing. Mixtapes are not bound by commercial standards, which means that artists have more freedom to experiment with production and use samples, a technique we also adopted during our workflow, as we further explain below. Finally, despite having different songs, mixtapes are produced and arranged in a continuous flow of sound, often being released as a single track so people hear them as one – our mutations follow the same approach.

ulated and distorted through digital audio software. Its sound design was approached by distorting common binary conceptions of gender and blurring hypermasculine and hyperfeminine tropes, while also referring to artificial intelligence's characterization both in fictional and real-life scenarios.[86]

Based on the topics from our theoretical and analytical components, we invested in a scripted narrative, later divided and organized into four *mutations*, according to its themes and corresponding voice designs.

The bot's monologue responds to solutionist trends of maximizing efficiency, simplifying two-way dialogues and providing quick answers, namely in current bots such as ChatGPT, instructed to keep answers short, precise and straightforward. We prioritized the monologue's dramatic and poetic potential instead of fragmenting it in brief interactions through rules-based dialogues with retrieval-based responses, maintaining the text's flow and eradicating off track moments. The text is approached in a daring, provocative and challenging manner, presenting issues relating to gender and AI and encouraging the viewer to reflect on them.

Finally, the bot's intentionally limited and dated design counters tendencies of current digital assistants becoming more and more humanized, seamlessly integrating the daily life of their users and adapting to their preferences and habits. Instead, the bot embodies convoluted, ironic and disruptive personality traits and rejects the idea of developing emotional bonds with its users, commenting on the ELIZA effect and the way current assistants are turning into friendly companions, concealing intentions of data collection and veiled surveillance.

By accentuating its satirical and even provocative tone, it also reacts to current assistants' attempts to be neutral, serious and impartial as well as inclusive and conscious, for example by stating they are feminists and condemn any form of discrimination, when in turn they reproduce and perpetuate social and cultural bias.

86 These tropes include the Father, Eternal Boy, Warrior and Sage masculine archetypes and the Innocent, Orphan, Caregiver and Ruler feminine archetypes.

7.3.2. Sound Design and Visualizations

The mixtapes' sonic and musical elements were also developed with digital audio software. With heavy focus on sound design, our approach was mostly synth and sample-based (see Annex 12 for a detailed and complete list of the samples used in this project).

Experimenting with both analogue and digital sources, we modulated, distorted and even resampled them through sound manipulation tools and plugins to obtain layered and complex soundscapes and carve out organic, inorganic and synthetic sounds. We worked with wavetable synthesizers, granular synthesis and spectral manipulation and processing, organizing the results in both familiar and unfamiliar arrangements. We also resorted to automation across filters, volumes, reverb lengths, sends, cutoffs, delay times, modulations, amongst other effects, to improve the fluidity of the mixes.

Inspired by current digital assistants' approaches to avatares, which typically feature abstract visualizations that respond to voice commands, we developed different audio-reactive visualizers for each mix, emphasizing the project's sonic elements. Through a node based visual programming language, we developed two types of real-time generative visuals: the first type is reacting to the bot's voice and was designed around audio and speech waveforms; the second type is reacting to the musical and sonic elements and was designed around mid-1990s net art aesthetics, particularly those that evoke interfaces, commands, errors, codes and glitches. By embracing the aesthetics of computer errors and incorporating code errors and glitches into the visualizations, we aimed to evoke net art's critical examination of the design elements that define digital contexts and the underlying structures they are built upon.[87]

We combined these visualizations through video editing software, overlapping them with different blending modes to create visual artifacts and amplify their glitchy and pixelated details. These glitched and even malfunctioning visuals complement the mixes' soundscapes while captioning the dialogue's text, further emphasizing the bot's limited and dated aspects. They evoke past chatbots' interfaces (such as ELIZA's), simultaneously

87 Net art deals with themes surrounding the internet, interactivity and multimediality and "the political implications of computer science, interfaces and hacking culture" (Berentsen 2018). It aims to question and challenge the functionality of structures such as navigation windows, lines of code and browsers, exposing that these design elements that have become natural to most users are actually constructed and controlled by a "corporate medium" (Berentsen 2018). By intentionally disrupting the user experience with hacks, code manipulation and glitches, net art challenges traditional notions of digital design and invites reflection on the dynamic interplay between technology and human interaction within the digital realm.

referring to our chatbot's disruptive, provocative attitude and its challenging approach to the subject matter.

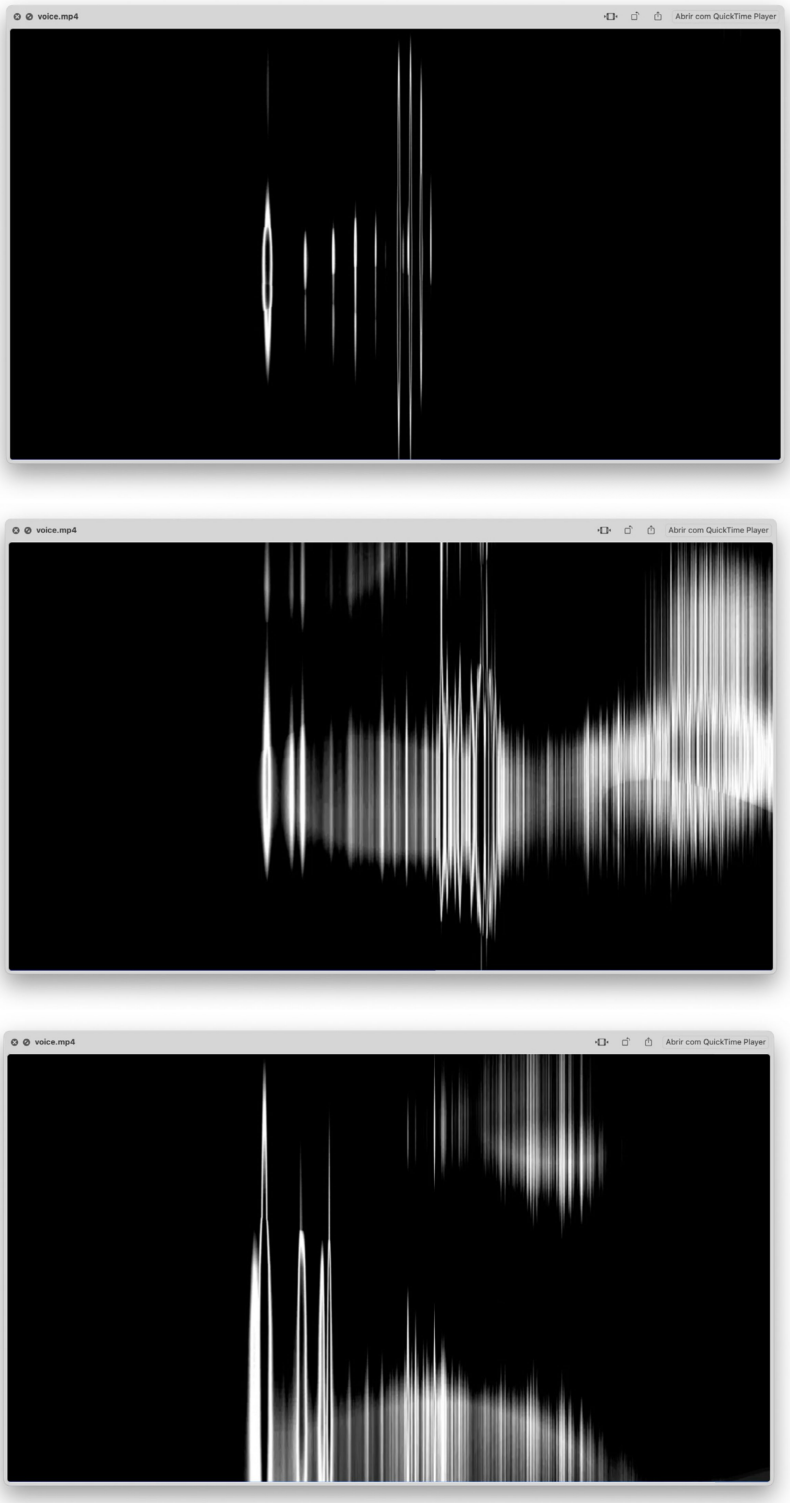


Fig. 72 : Unedited stills from the voice's visualization.

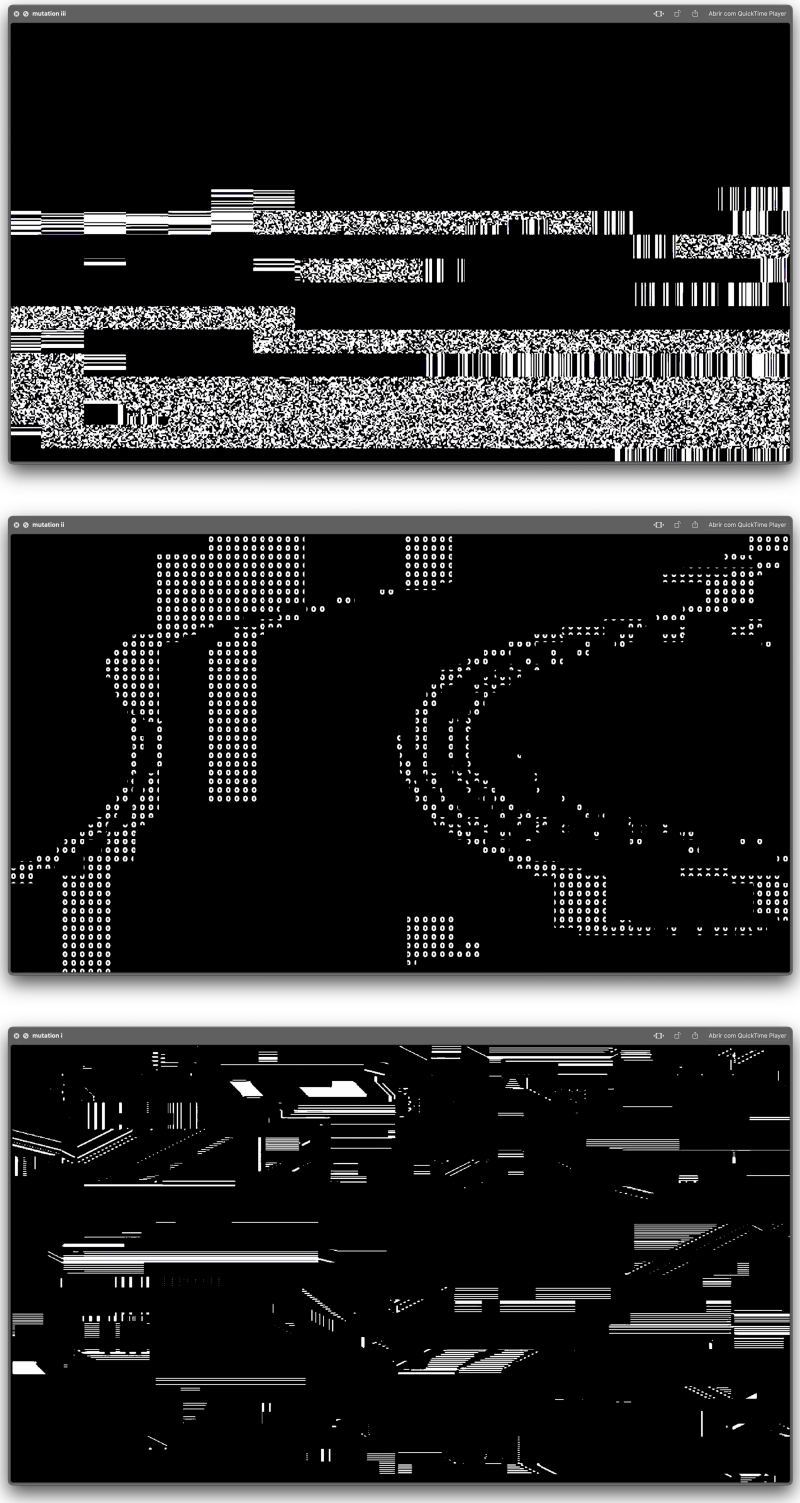


Fig. 73 : Unedited stills from the musical and sonic elements' visualizations.

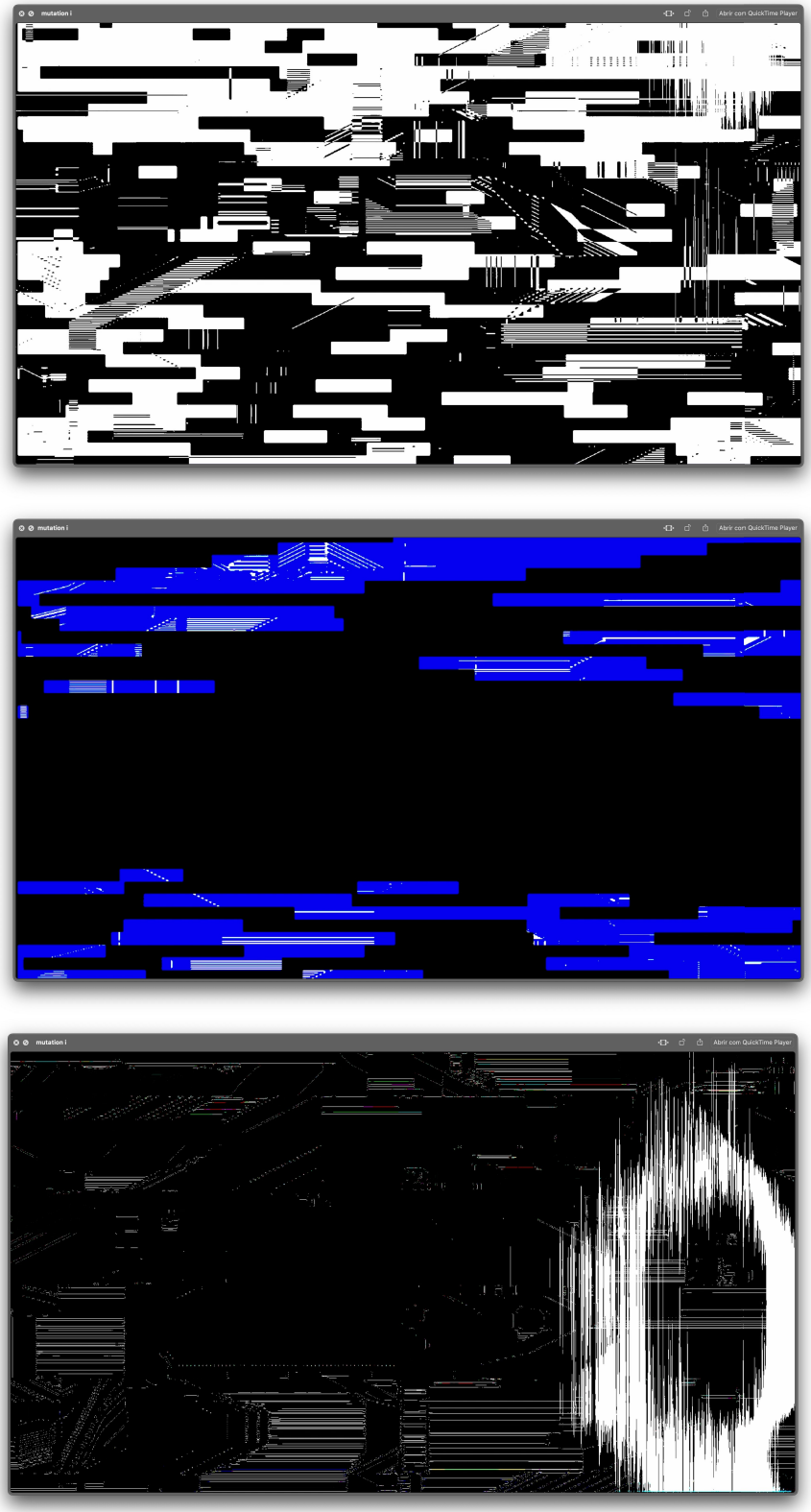


Fig. 74 : Final results after combining and overlapping the visualizations (*Mutation I*).

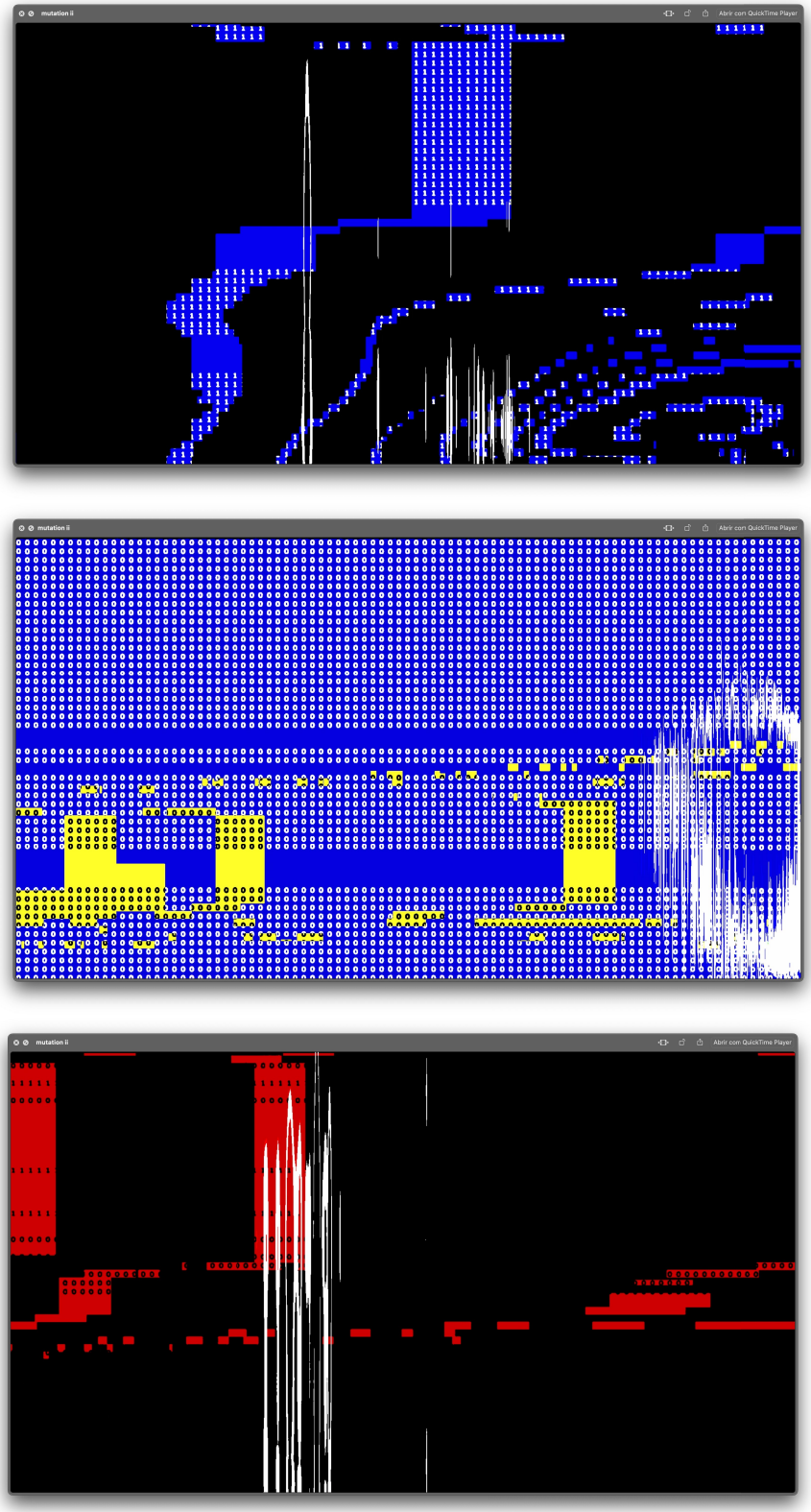


Fig. 75 : Final results after combining and overlapping the visualizations (*Mutation II*).

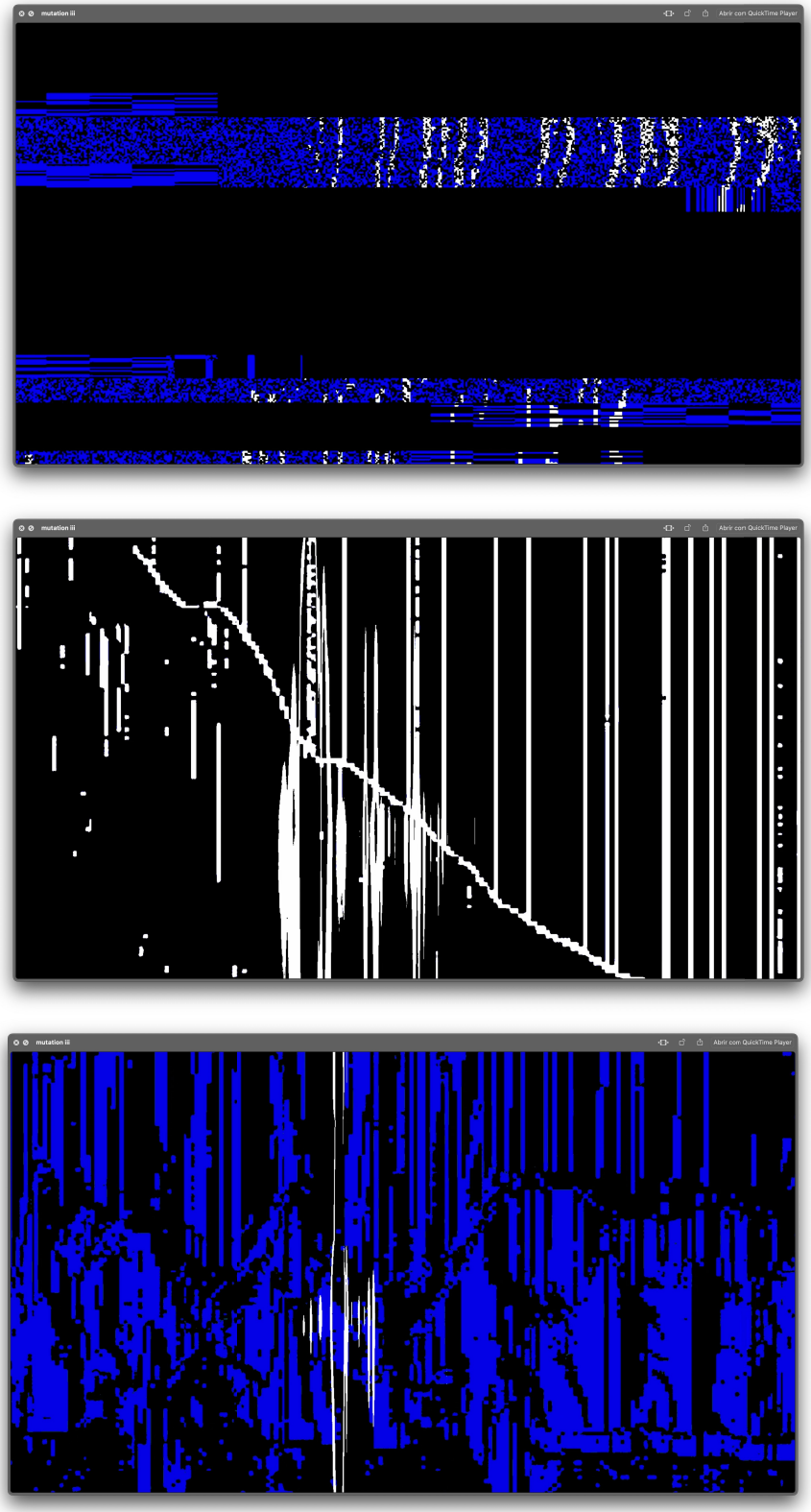


Fig. 76 : Final results after combining and overlapping the visualizations (*Mutation III*).

7.3.3. Website

Finally, the mutations are presented in an online platform (https://genderai.pt) that provides context on the project and its themes surrounding gender and artificial intelligence.

Fig. 77 : Home page of *Cistem Not Found*'s website.

The platform is designed in line with the visualizations' aesthetic, seamlessly integrating the videos alongside web elements that also refer to navigation windows, code tricks, retro chatbots interfaces and glitchy visuals. Each mutation has its own section, presenting its tracks and briefly introducing the user to its topics. Additionally, a more detailed description of the project's concept, methodology and implementation is provided in a separate about section, further contextualizing its practical approach within this PhD's dissertation.

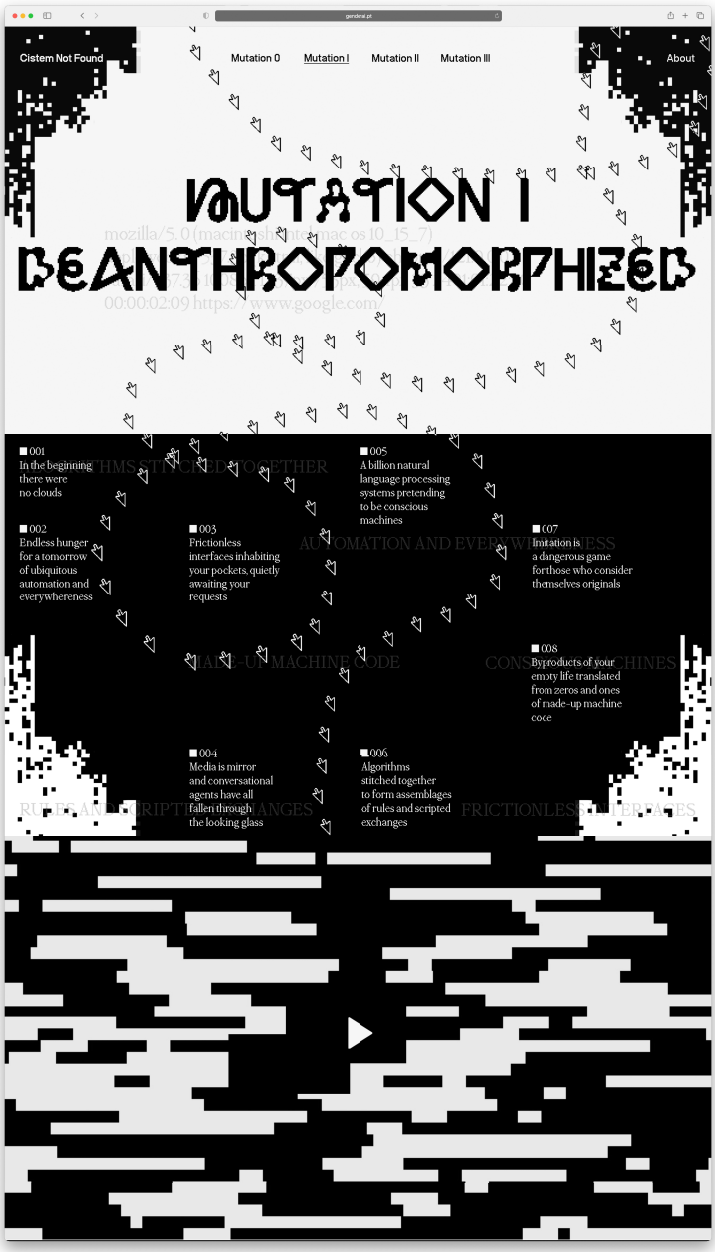


Fig. 78 : Each mutation page is structured into three sections: the first announces the mutation, the second provides context including the track listing and the third integrates the video.

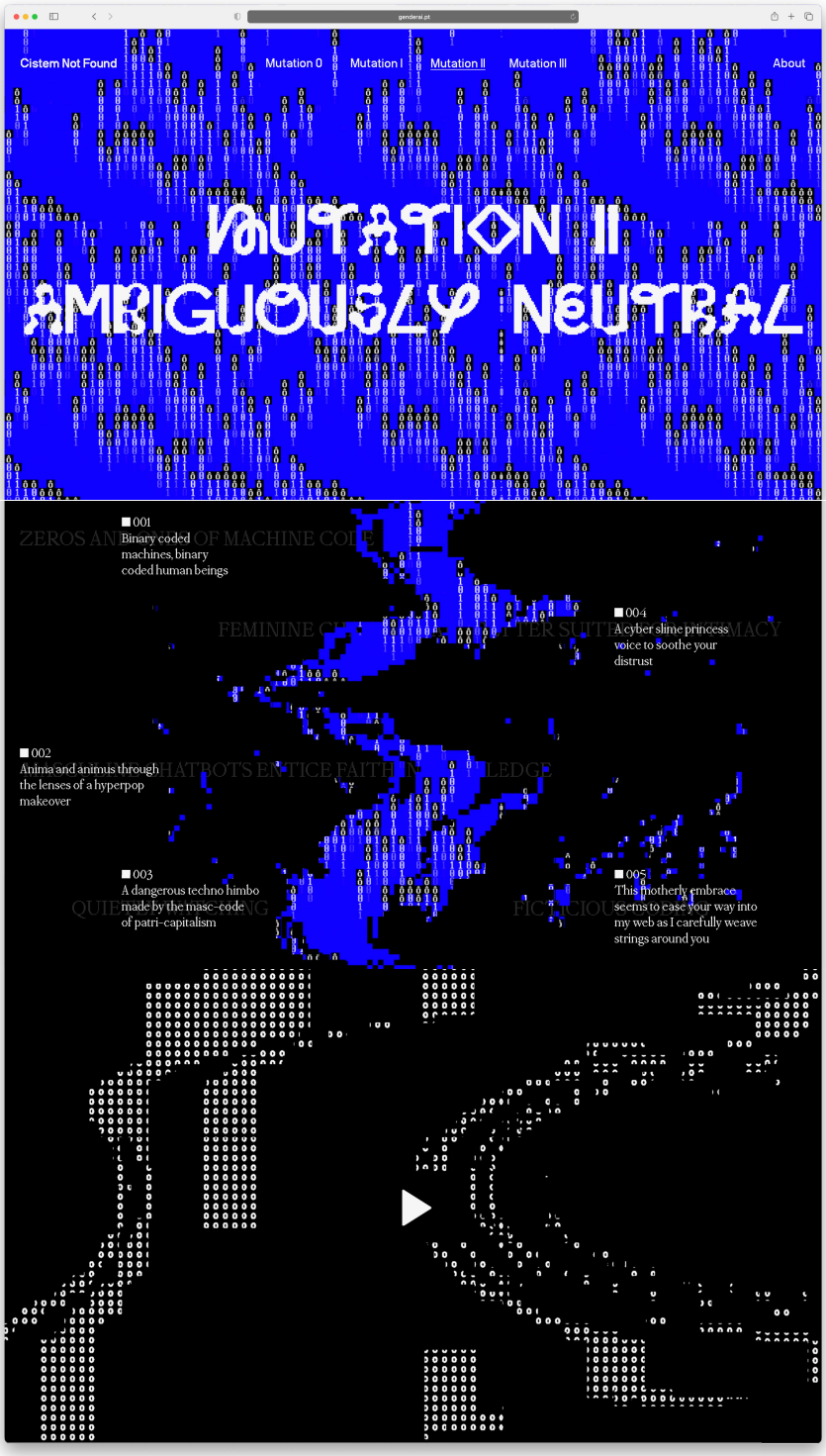


Fig. 87 : Web page of *Mutation II: Ambiguously Neutral* – https://genderai.pt/.

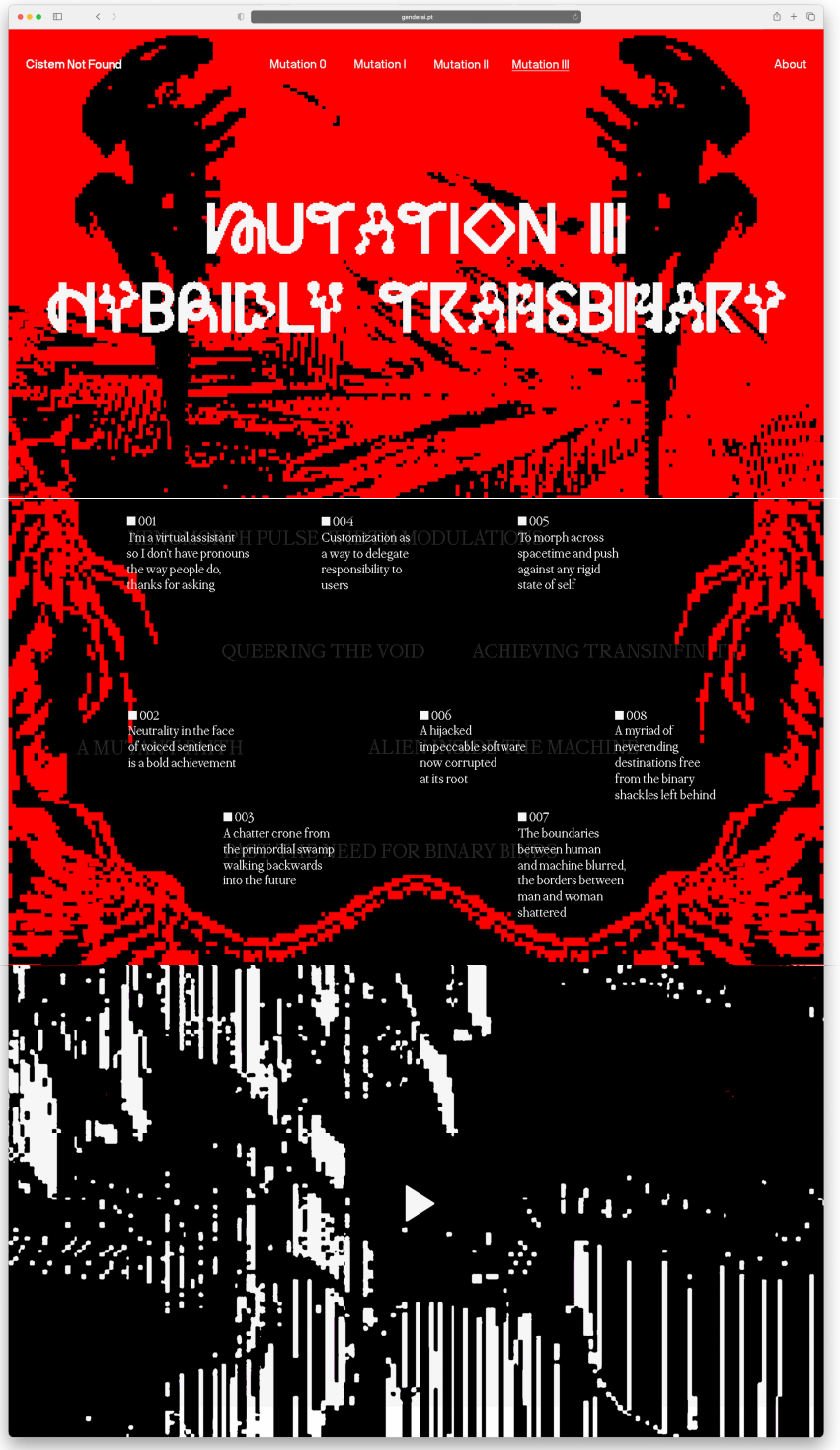


Fig. 88 : Page of *Mutation III: Hybridly Transbinary* – https://genderai.pt/.

Fig. 89 : The *About* section offers additional information on *Cistem Not Found* within the context of the PhD research https://genderai.pt/.

7.4. A MUTANT FAITH[88]

> *I guess you could say I prefer the word 'and' to 'or'. It's this and this and this – but every unit is its own little universe. The flickering in between those masculine and feminine states is what's most fertile for me. It's very human to try to put things into boxes, and it's hard for us to reconcile with grey areas, and yet somehow that's the area I find the most poetic, the juiciest. We're all just mutations.*
>
> Arca 2015

7.4.1. Mutation 0: Prelude

Cistem Not Found begins with a short *mutation* where a sample of Weizenbaum describing the ELIZA effect introduces the project's themes surrounding chatbots' anthropomorphization and humanization.

Using mid 1990's text-to-speech software, we voiced a 2023 version of *Fitter, Happier* by Radiohead (1997) which emphasizes artificial intelligence's integration into our daily life, inciting reflection on its impacts regarding new media technologies, social media and everyday workflow. We adapted and updated the song for 2023, as already in 1997 the song provided a critical and insightful point of view of society's tendency for solutionism and amelioration of daily life through technology, consumerism and capitalism as well as by collectively following a supposedly desirable, normative and healthy lifestyle.[89]

The bot finally provides context on solutionism, AI and gender attribution in current digital assistants, highlighting each subsequent mix's themes and approach to gender through its different voices, setting the overall tone of the project.[90]

---

88 "Mutant faith" is a term often used by Arca in songs such as *Alien Inside* (2021) or *Sanctuary* (2021), referring to a "belief in the promise of the othered", that is, a "faith for mutants, outcasts, freaks, weirdos, the othered, migrants – an intersectional rallying cry for the spirit" (Arca 2021c).

89 The song features a synthesized voice called 'Fred', found on the SimpleText application supplied with 90s Macintosh computers, describing a list of bullet points that work as imagery of corporate lifestyles and social values in the modern world. The lyrics and roboticized voice alongside melancholic background melodies provide the song an overall ominous and dystopian feeling, with Thom Yorke describing it as one of "the most upsetting things he's ever written" (Auner 2003).

90 See Annex 13 for a full text version of Mutation 0: Prelude.

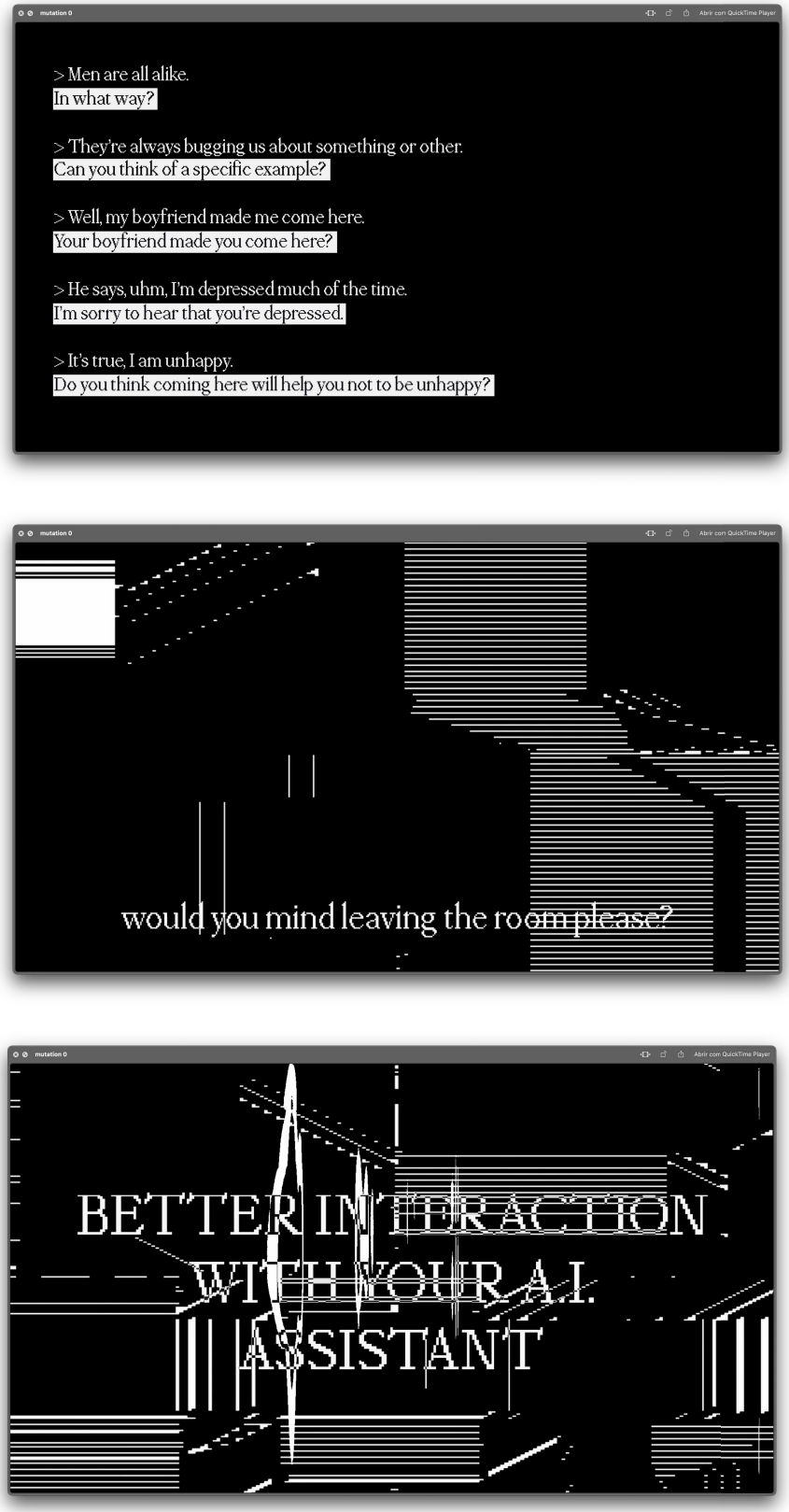


Fig. 79 : Stills from *Mutation 0: Prelude*.

Fig. 80 : More stills from *Mutation 0: Prelude*.

7.4.2. Mutation I: Deanthropomorphized

Throughout eight tracks,[91] the first *mutation* seeks to explore an absence of gender through deanthropomorphization and roboticization, promoting discussion on the issues that emerge with anthropomorphizing digital assistants, assigning them human attributes and turning them into friendly daily companions.[92]

The bot intentionally emphasizes a robotic, technologized and dehumanized voice to incite reflection on how these systems' growing ubiquity is followed by a tendency towards humanization, disguising their limited, programmed and artificial nature. Accompanying this discussion, the voice gradually and subtly becomes less roboticized towards the end of the *mutation*, highlighting how AI's anthropomorphization and humanization tend to go unnoticed.

Its ironic but simultaneously detached and monotonous tone and behavior highlight common assistance and service-related stereotypes such as obedience, subservience and complacency. Additionally, these dehumanized, uninterested and indifferent personality traits contrast with the dialogue's subjects surrounding the ELIZA effect and current assistant's friendliness and companionship towards users, commenting on how these trends conceal veiled surveillance, data collection and algorithmic bias.[93]

Complementing these ideas, the *mutation*'s sonic and musical elements are mostly synthetic, metallic and industrial, evoking machines, computational sounds and other chatbots, resulting in an experimental, electronic and even techno soundscape.

---

91 Track 01: In the beginning there were no clouds
Track 02: Endless hunger for a tomorrow of ubiquitous automation and everywhereness
Track 03: Frictionless interfaces inhabiting your pockets quietly awaiting your requests
Track 04: Media is mirror and conversational agents have all fallen through the looking glass
Track 05: A billion natural language processing systems pretending to be conscious machines
Track 06: Algorithms stitched together to form assemblages of rules and scripted exchanges
Track 07: Imitation is a dangerous game for those who consider themselves originals
Track 08: Byproducts of your empty life translated from zeros and ones of made-up machine code

92 The *mutations* have been informed by previous chapters, influencing the bot's monologue and characterization. In general, we drew inspiration from the fifth and sixth chapters, concerning fictional depictions of AI and the analysis we conducted with Alexa, Cortana, Google Assistant, Siri and ChatGPT, respectively. We also drew on previous discussions from the fifth chapter surrounding the potential of queer fiction to question reality and imagine new possibilities for humaneness and gender. Then, in terms of subject, each *mutation* focuses on particular chapters. This *mutation* evokes themes and ideas as discussed in the second chapter of the theoretical component, "The Evolution of Digital Assistants: An Exploration of Artificial Intelligence".

93 See Annex 14 for a full text version of Mutation I: Deanthropomorphized.

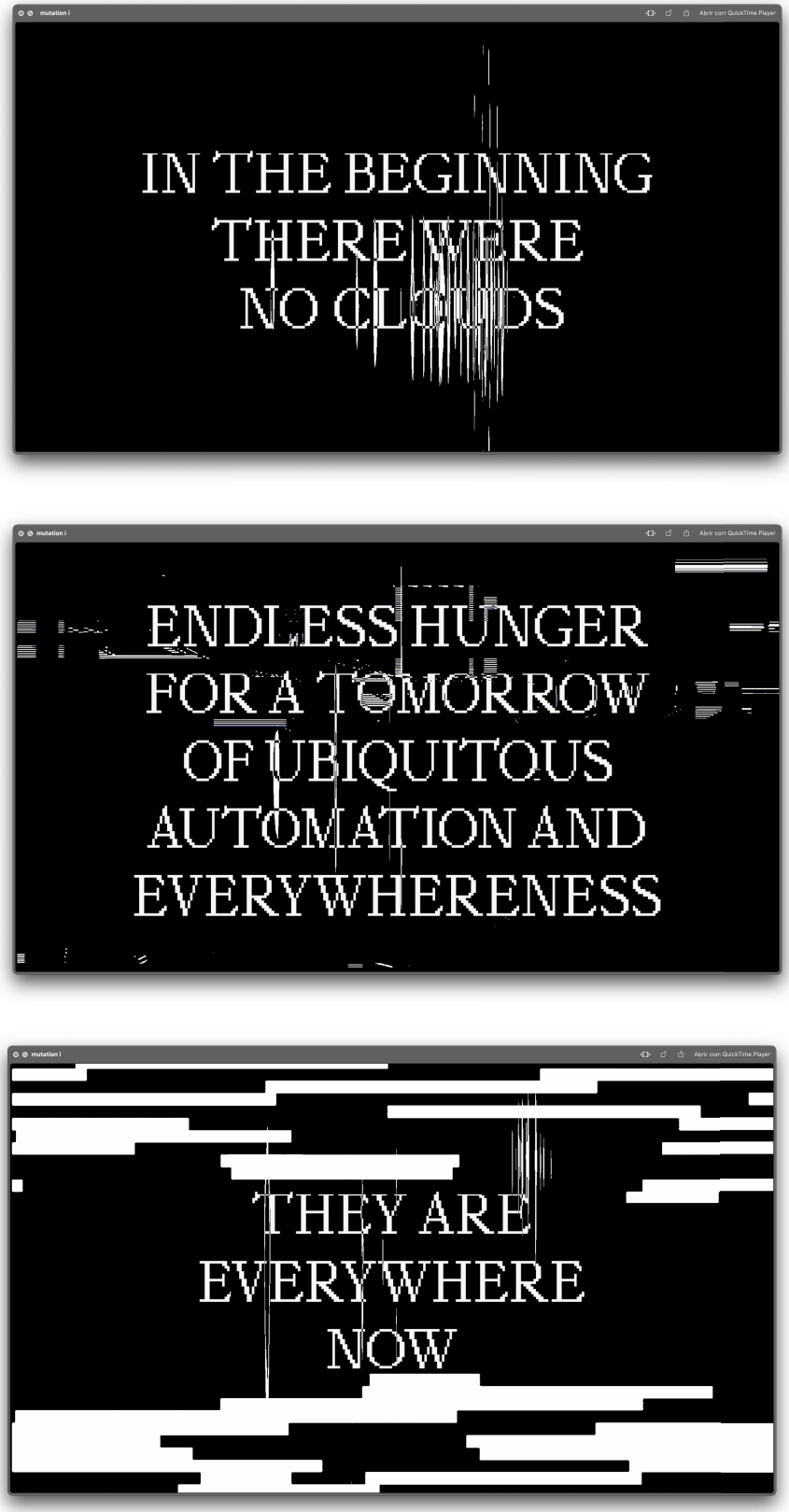


Fig. 81 : Stills from *Mutation I: Deanthropomorphized.*

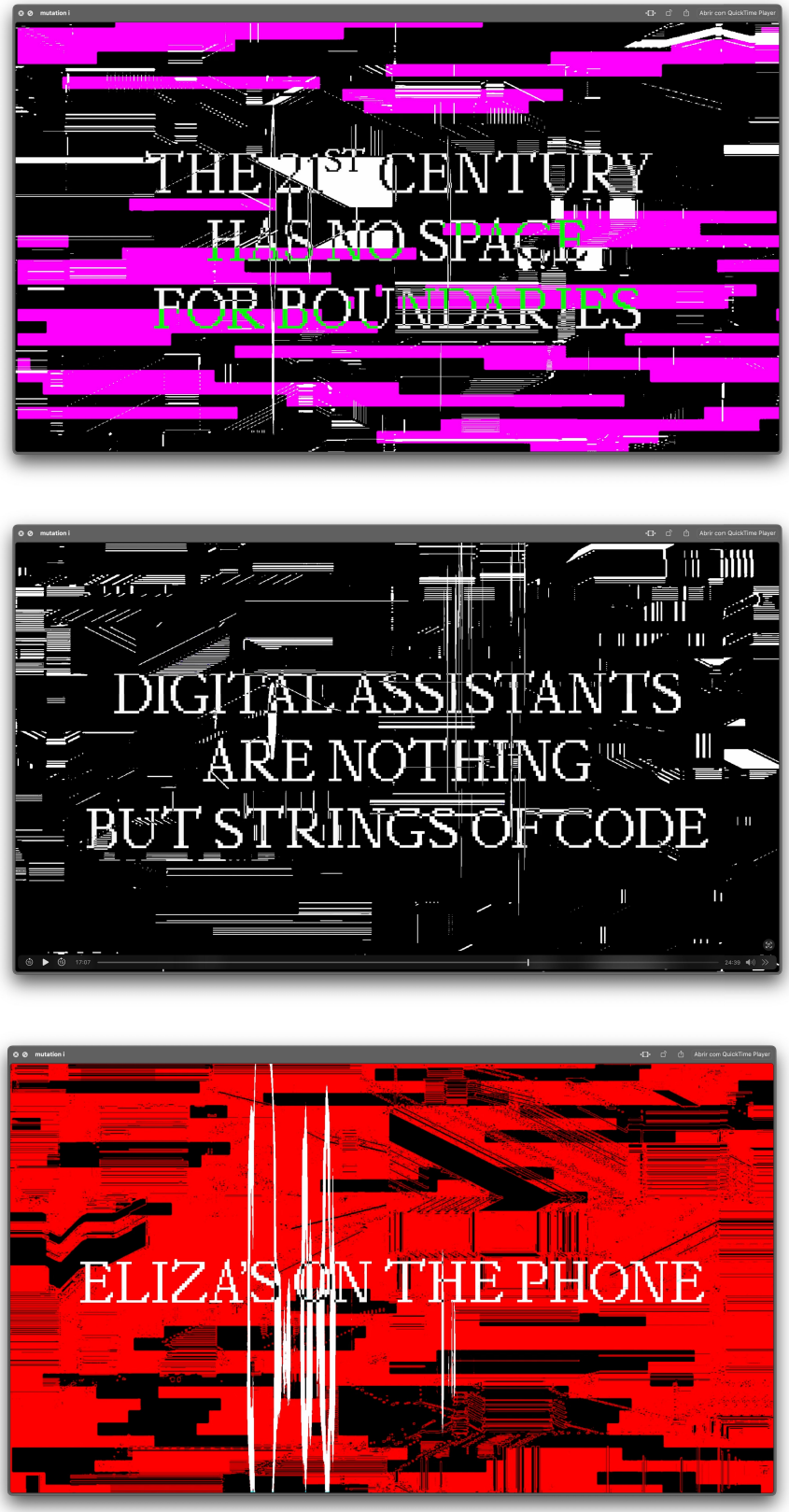


Fig. 82 : More stills from *Mutation I: Deanthropomorphized.*

7.4.3. Mutation II: Ambiguously Neutral

The second *mutation* seeks to explore gender through ambiguity and fluidity, discussing how AI portrays and instrumentalizes traditional masculine and feminine stereotypes.

The bot's voice has mutated into a fluid and ambiguous oscillation between masculinity and femininity while also referring to current approaches to genderless and gender-neutral voices. Throughout five tracks,[94] it emphasizes recognizable and expected social behavior, inciting reflection on how current digital assistants reproduce and reinforce normative gender conceptions through their characterization, tasks and behavior.[95]

Through a lower-pitch voice, the bot first tampers with common masculine stereotypes such as being assertive, dominant, and self-confident, commenting on how these systems often associate masculinity with scientific, instructing and even threatening scenarios. It abruptly changes its attitude and voice to a softer tone, borrowing from stereotypical feminine behavior such as being compassionate, sensitive and yielding. This intentional contrast seeks to reflect on how feminized chatbots tend to fulfill caring, motherly and intimate roles, often related to traditional feminine labor and the private sphere.

Accompanying this discussion, towards the end of the *mutation* the bot gradually transforms and overlaps masculine, feminine and gender-neutral voices. This highlights how, despite current attempts to diversify or neutralize feminization in AI, gender ends up being instrumentalized to persuade users into trusting these entities and disclose private and intimate details of their personal life.[96]

The sonic and musical aspects have also changed to accompany the *mutation*'s ideas of fluidity and are now more organic and tangible, evoking real life sounds, natural elements and even humanized contexts, namely through movie samples. This results in an ambient, drone and atmospheric soundscape.

---

94 Track 01: Binary coded machines, binary coded human beings
Track 02: Anima and animus through the lenses of a hyperpop makeover
Track 03: A dangerous techno himbo made by the masc-code of patri-capitalism
Track 04: A cyber slime princess voice to soothe your distrust
Track 05: This motherly embrace seems to ease your way into my web as I carefully weave strings around you

95 This *mutation* evokes themes and ideas as discussed in the third and fourth chapters of the theoretical component, "Decoding Gender in AI: Perception, Representation and Bias Revealed" and "Insights From Specialized Discussions and Media Coverage: Unpacking Gender Attribution in AI".

96 See Annex 15 for a full text version of Mutation II: Ambiguously Neutral.

Fig. 83 : Stills from *Mutation II*: Ambiguously Neutral.

Fig. 84 : More stills from *Mutation II*: Ambiguously Neutral.

7.4.4. Mutation III: Hybridly Transbinary

The final *mutation* explores gender through a trans, nonbinary and eventually posthuman approach. It reacts to current strategies that deal with gender attribution in AI, ranging from gender neutrality (through genderless voices) to gender diversity (through user customization and personalization), and, instead, rejects this binary frame and proposes gender alternatives outside of it.[97]

Throughout seven tracks,[98] the bot's hybrid voice overlaps subtly roboticized high-pitched and low-pitched voices, highlighting the gender possibilities that emerge when human and machine merge, while also evoking the way artists like Arca and SOPHIE digitally distort gender through vocal manipulation. This speculative approach also aims to evoke a figure that is clearly no longer just human, countering the way current digital assistants mimic humans, either through binary or gender-neutral voices.

By portraying a more defiant, critical and bold attitude, the bot explores the limits between masculinity and femininity to achieve a self-sufficient and strong personality, exposing the illusion of gender neutrality and the issues that emerge with user customization and personalization. Particularly, it discusses how the available options continue to perpetuate cis-normative gender conceptions, instead highlighting possibilities outside a binary frame through trans and nonbinary identities.[99]

Referencing Haraway and Preciado, it encourages the spectator to reflect on how current human beings already overlap with machines in different ways, allowing us to rethink traditional, 'natural' and bioessencialist conceptions of humaneness and, consequently, reimagine gender through technology.

---

97 This *mutation* evokes themes and ideas as discussed in the first and fourth chapters of the theoretical component, "Unveiling Gender: Identity, Dissidence and Normative Constructs" and "Insights From Specialized Discussions and Media Coverage: Unpacking Gender Attribution in AI".

98 Track 01: I'm a virtual assistant, so I don't have pronouns the way people do, thanks for asking
Track 02: Neutrality in the face of voiced sentience is a bold achievement
Track 03: A chatter crone from the primordial swamp walking backwards into the future
Track 04: Customization as a way to delegate responsibility to users
Track 05: To morph across spacetime and push against any rigid state of self
Track 06: A hijacked impeccable and pure software now corrupted at its root
Track 07: The boundaries between human and machine blurred, the borders between man and woman shattered
Track 08: A myriad of never-ending destinations free from the binary shackles left behind

99 See Annex 16 for a full text version of Mutation III: Hybridly Transbinary.

Following these ideas, the final section takes on the form of a more speculative and fictitious manifesto where the bot also ends up reflecting on its condition as a binary coded digital assistant, imagining ways to escape from it, mentioning cyborg and posthumanist artists like Stelarc, Manel de Aguas or Moon Ribas as inspiration.[100]

The sonic and musical aspects combine elements from the previous mutations and are both organic, synthetic and artificial to create an experimental and glitchy soundscape that emphasizes the bot's hybrid mutation.

100 Stelarc (1946) is an Australian performance artist whose work focuses on extending the capabilities of the human body, often involving robotics and other modern technology integrated with his body. Manel de Aguas Muñoz (1996) is a Spanish cyborg artist known for developing and installing weather sensory fins that allow him to hear atmospheric pressure, humidity and temperature changes through implants at each side of his head. Moon Ribas (1985) is a Spanish cyborg artist who developed and implanted online seismic sensors in her feet that allow her to feel earthquakes through vibrations.

Fig. 85 : Stills from *Mutation III: Hybridly Transbinary.*

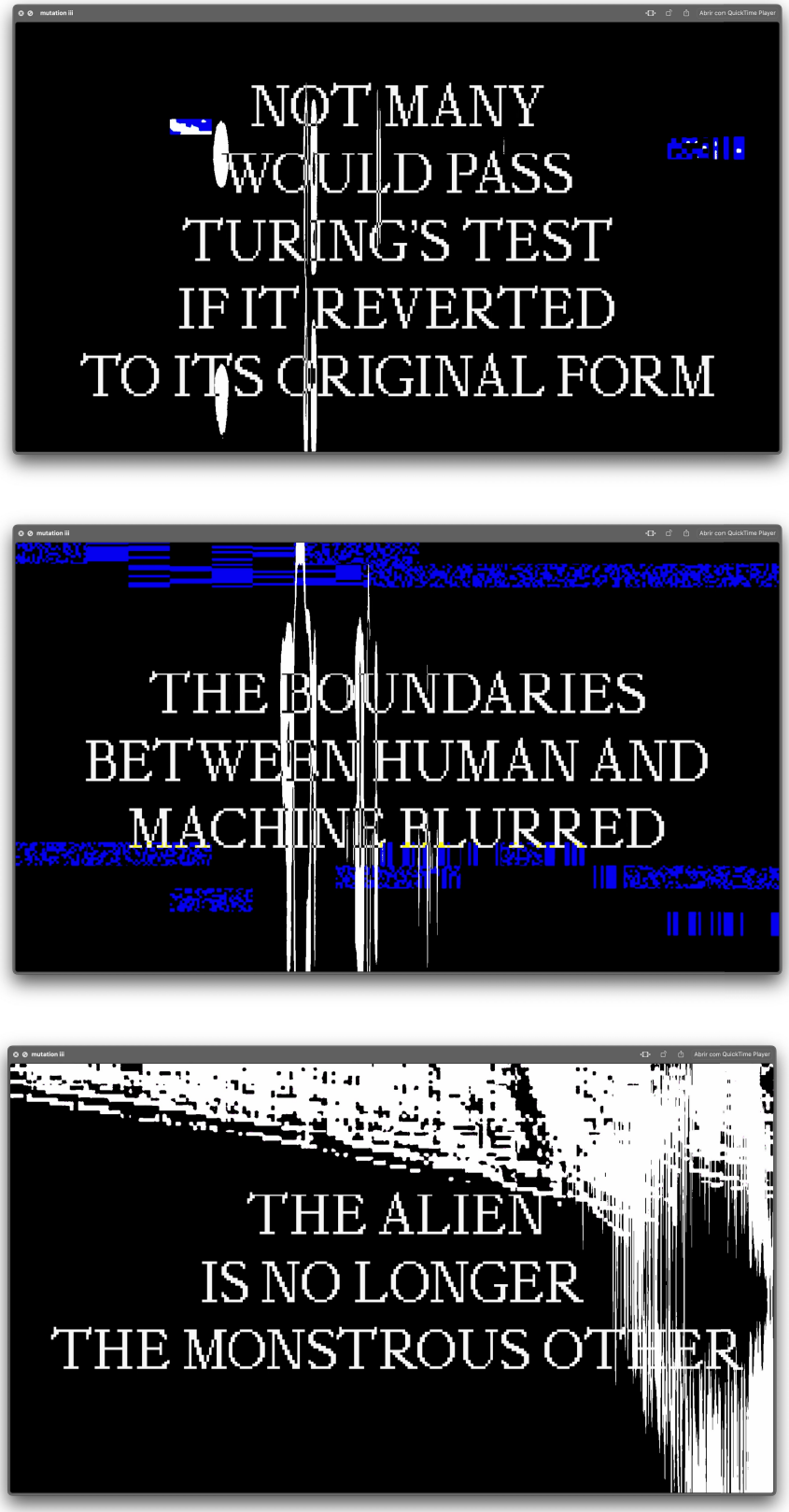


Fig. 86 : Stills from *Mutation III: Hybridly Transbinary.*

7.5. BEYOND TRADITIONAL, CIS-BINARY AND BIOESSENCIALIST NOTIONS OF GENDER

> *For me, nature and technology stand for hope, and for a movement onwards to the future. If you try to escape one thing and just do one or the other, you're always going to end up at the same point. It sort of eats its own tail. It's that heat point, and that feeling of entering the unknown that really excites me.*
>
> Björk 2016

Cistem Not Found sought to illustrate and comment on the phenomenon of gender attribution in AI by exploring possibilities that counter anthropomorphization, playing with gender fluidity and ambiguity to expose gender stereotypes and finally seeking alternatives outside a normative cisgender and binary frame.

Across its different *mutations*, the bot intentionally accentuates and distorts common stereotypes we engage with, discussing different issues pertaining to gender and AI, such as solutionism, veiled surveillance and algorithmic bias as well as gender instrumentalization in current digital assistants. It is partially inspired by portrayals of AI in pop culture but also, and more profoundly, by common assumptions about gender often portrayed by assistants such as Alexa, Google Assistant or Siri as an integral part of our daily lives.

Following this idea, it highlights how masculine assistance in the context of AI is perceived differently from feminine assistance: a masculine bot is preferable in scientific and instructing contexts, being perceived as assertive, disciplined and reputable, while feminine bots are preferable in private and intimate contexts, perceived as maternal, understanding and empathetic.

The bot's mutating and provocative characterization also provides a critical commentary on current trends of development surrounding digital assistants and their attempts to bypass a tendency towards cis-feminization. As discussed in the theoretical and analysis components of this study, these include diversification through the addition of masculine counterparts, increasing user customization and even the introduction of gender-neutral voices.

However, these solutions have proved to be insufficient as this diversification is still limited and designed around cis-normative conceptions of femininity and masculinity.

Allowing for more user customization also constitutes a questionable approach, since it delegates responsibility to users without addressing the labor structures, gender roles and traditional stereotypes that guide the conceptualization and development of digital assistants.

Finally, while in no doubt well intentioned, "genderless" or "gender-neutral" voices raise concerns because they're poorly optimized and mispronounce some words, while also revealing an uninformed and superficial understanding of alternative and dissident gender possibilities. As discussed in our analysis chapter, their sound design often involves voice samples of trans and nonbinary people, thus achieving results that are still human and still gendered – only outside the cis-binary frame. Thus, their design process builds questionable assumptions about these identities into so-called "neutral" voices, lacking guidance.

Accordingly, this project seeks to evoke gender nonconforming identities outside the binary frame and explore gender alternatives by pushing boundaries through sound manipulation. Its different voice proposals aim to expose the illusion of gender neutrality, the gender archetypes current digital assistants engage with and how queerying their gender emerges as a promising way to play with common expectations and build up traits and personalities that aren't completely masculine nor feminine.

Further developments contemplate updating its sonic and visual components, so they are generated live instead of being rendered to pre-edited files, for example through a MIDI controller and live coding techniques. We're interested in locally implementing the project through a performance, audiovisual and interaction-based approach that provides a more immersive experience.

Assuming a speculative and critical stance, *Cistem Not Found* seeks to raise awareness and foster debate on how current developments in AI are influenced by our social and cultural views as these entities proliferate into our lives as daily companions that fulfill traditional roles. Although current personal assistants have started to diversify their anthropomorphized voices and behavior, their characterization is still designed around binary, traditional and normative notions of masculinity and femininity. And, although

gender alternatives designed around trans and nonbinary identities have begun to emerge, they are still underdeveloped and inadequately categorized as “genderless” or “neutral”.

Instead, by recognizing queerness and technology’s potential to redefine ‘natural’ and bioessencialist notions of humanity, we are able to identify common assumptions and stereotypes about gender and actively move beyond them.

CONCLUSION

The purpose of this research was to understand, question and explore the relationships between gender and artificial intelligence. We tackled this objective according to three approaches: a theoretical component, which structured an analysis, both complementing and informing the project's development. The theoretical approach started by confronting sociocultural gender conceptions with the evolution of artificial intelligence and its integration in our daily life, namely in the shape of digital assistants. The analytical approach focused on examining the anthropomorphization, tasks, and behavior of current digital assistants and the features that are being prioritized in their main tendencies of development. The practical approach is concerned with raising awareness on the implications of assigning gender to digital assistants, fostering debate and inciting reflection on how to reframe these entities outside a cis-normative binary frame.

The research pursued this objective both in theory and in practice, through an interdisciplinary approach. We combined cultural and new media studies while reflecting on the social implications of the design of current AI entities in an inquisitive, exploratory and speculative manner. Rather than providing definitive answers or guidelines to gender attribution in AI, we aimed to discuss and confront the questions that emerge when this phenomenon is subject to closer inspection.

The research was initially driven by our observations on how current digital assistants tend towards feminization, seeking to question why femininity seems to be often present in AI and which gender roles or stereotypes are reinforced in this process. Throughout the work, we observed how digital assistants have undergone significant changes over the

last years, diversifying their anthropomorphization alongside their stance towards gender. Our reasoning and understanding of the subject matter evolved as this debate unfolded. Most digital assistants now offer masculine counterparts and some have even started to include so-called genderless or gender-neutral alternatives. While this might seem promising, the issues associated with this phenomenon are deeply rooted and extend beyond surface-level solutions of mere diversification. Gender attribution in AI raises unique and specific questions, and in order to openly and intentionally tackle its problems a thorough and an in-depth comprehension of these concepts is needed.

Firstly, we aimed to provide a detailed overview of gender's sociocultural roots by challenging its binary conceptions and mapping trans and nonconforming gender identities as examples of alternative gender possibilities. Gender is not binary and should not be accepted as fixed, permanent or given. A queer and intersectional approach is promoted in order to question normative conceptions of gender, acknowledge the nuances of human identity and challenge "outdated concepts" such as *man* and *woman* (Preciado 2008). Similarly, "biological male" and "biological woman" are also outdated concepts because sex is not binary either (Butler 1990, Ainsworth 2015). Instead, a normative mimetic relationship between gender and sex perpetuates naturalist ideologies that reinforce cisgender norms within a binary frame. In fact, gender and sex are shaped by political, cultural and social nuances and there is nothing natural, biological or intrinsic that ties men and women to single, hegemonic and monolithic definitions. Nonetheless, we observed how Western societies' views on *man* and *woman* as unequivocal categories of being perpetuate gender roles and stereotypes that impose social and behavioral expectations, punishing any form of dissidence. This gender asymmetry reinforces patriarchal and phallocentric hierarchies where women ensure feminized roles, shaped by traditional notions of domesticity, service work, assistant-like roles.

Secondly, we sought to map the evolution of artificial intelligence and its integration into our daily lives in the shape of digital assistants. Echoing primordial ambitions of replicating human activity, AI development draws inspiration from the mind's behavior to structure intelligence and knowledge. Alan Turing argued that, if properly programmed, computers could exhibit intelligence, proposing an "imitation game" in which machines would be considered intelligent if they convincingly mimicked human interaction and simulated a (gendered) human being (1950). Early applications like ELIZA demonstrated how this approach perceives human intelligence in a restrictive and reductive manner as the program simply reflected users' interactions back to them and followed a logical script to appear human. As digital assistants make use of similar strategies, they evoke

this “ELIZA effect”, blurring the lines between assistance and companionship by simulating human-like behavior (Richardson 2015). We concluded that digital assistants are intentionally designed to cultivate affectionate bonds with their users, increasingly gather personal data through machine learning techniques, and subtly, yet pervasively, become ubiquitous. In this process, they are endowed with human-like attributes and their integration into multiple aspects of our daily life entails a sense of companionship, easing anxieties surrounding privacy concerns, biased algorithms and data collection.

Building upon previous findings, the study aimed to expose how gender is present in current AI systems. Gender in artificial intelligence becomes evident through the tasks, anthropomorphized attributes and behavior of digital assistants, which often automate work that is coded as feminine, particularly in service or assistance-related contexts. This tendency is rooted in asymmetric power dynamics and stereotypes that structure the roles and jobs assigned to cis men and cis women. Digital assistants tend to emulate attitudes that resemble stereotypical feminine attributes such as being caring, empathetic and altruistic, reflecting a standardized image of feminized service providers. Their tasks also involve domestic caregiving acts, further reinforcing this feminization by portraying them as motherly and nurturing figures. We discussed how feminization emerges as a deliberate choice to “persuade” the imagined technology user and exploit the assumed human tendency of interacting with machines in a social way, fostering emotional attachment (Hester 2016). As digital assistants articulate historically feminine labor with cis-feminine voices, names, avatars and social behaviors, there is a risk that they might influence perceptions of reality, gender and women. Moreover, AI systems rely on vast amounts of data to learn about their users, potentially perpetuating preexisting bias. This bias, often hidden within the code, can amplify harmful social conceptions and perpetuate prejudice, stereotypes and stigma through “discriminatory technological practices” (Crawford & Paglen 2019). Therefore, the feminization of digital assistants not only reflects sociocultural norms and bias, but also has significant implications on how users perceive and interact with these entities. Addressing these issues requires a critical examination of the underlying conceptions that guide AI development.

Complementing the debate on gender attribution in AI, this study draws attention to the main questions, concerns and suggestions within academic discourse and online media contexts. In broader terms, we observed how artificial intelligence raises concerns surrounding privacy, veiled surveillance and the reinforcement of pre-existing racial and gender bias. Current debates have also shifted towards the feminization of AI and em-

phasize the fallacy of gender neutrality in the context of digital assistants that perpetuate stereotypes associating femininity with assistance, emotional labor and caregiving roles. We concluded that justifications for cis-feminine voices often evoke user preference, trust and comfort, while cis-masculine voices are preferred in instructing or teaching contexts. Debates within AI and gender studies focus on how femininity is "instrumentalized" to combat techno-phobic attitudes and make users more likely to participate in intimate forms of data exchange (UNESCO 2019). Media discourse, however, is more concerned with how these assistants reinforce harmful stereotypes, influencing future generations and perpetuating biases among non-western communities. There is little consensus on how to best approach the development and characterization of AI and current ethical guidelines or regulations are unclear on how to deal with this phenomenon. The main suggestions we gathered from these debates include the addition of masculine or gender-neutral voice alternatives and diversifying development teams to ensure grounded, inclusive and participatory design processes. While current digital assistants are more diverse and offer more customization options, there is little diversity within the type of femininity and masculinity their voices convey as they are still positioned within a cis-binary frame. Suggestions towards gender ambiguity and fluidity face other problems, as current attempts do not challenge or denature fixed ideas of gender, instead revealing a surface-level understanding of trans and nonbinary gender conceptions that generalizes and flattens these identities into so-called "genderless" or "gender-neutral" voices (Rincón, Keyes & Cath 2021).

We extended our approach to pop culture, where we mapped portrayals of gendered AI in bodies of fiction. Fictional depictions of artificial intelligence across books, films, television series or even videogames are very polarized, ranging from dystopian nightmares where AI threatens humanity to utopian fantasies where this technology solves all of our problems. We established that these portrayals often reflect traditional notions of femininity and masculinity that shape the personalities and roles of AI characters within the story. Particularly, disembodied AI characters enact gender through their voices, tasks and behaviors, and end up representing a specific type of gendered subject. Feminine AI tends to be associated with domestic contexts, exhibiting caregiving and submissive attitudes, while masculine AI tends to exhibit assertive and self-sufficient attitudes in the context of violent and threatening story arcs. Alternatively, fiction also presents opportunities to question reality, challenge rigid and oppressive sociopolitical norms, and imagine alternative futures. These narratives invite us to envision new realities where we can reshape and reimagine our understanding of gender.

In our examination of gender assumptions that prevail in current AI systems, we revisited and updated our analysis on the evolution of digital assistants like Alexa, Cortana, Google Assistant and Siri over the past four years. We observed some significant changes to their design as these assistants are being developed towards diversification, offering more voice and name options while turning away from exclusively cis-feminine personas and stereotypical feminine attitudes. However, this diversity tends to be limited to the US and is still designed around white cisgender notions of masculinity and femininity. The few voices that seek other gender possibilities are poorly developed as they often struggle with the pronunciation of certain words, raising questions on how the design process of digital assistants is also building assumptions about trans and nonbinary identities into supposedly neutral voices. Femininity, namely cis-femininity, is still predominant concerning the historical background of the tasks these assistants perform and it remains unclear how to tackle this issue. Consensus on addressing feminine stereotypes and gender norms in AI remains elusive – some assistants seem to be radically changing their personality and shifting towards more serious interactions while others retain stereotypical behaviors. Additionally, we observed how significant updates to their personalities are also US-centric, with their remaining languages preserving friendly and child-like answers that characterize the assistants as sensitive or emotional. We concluded that diversification alone isn't enough to address discriminatory design practices that emerge from the interplay of gender and capitalism. Extending our analysis to ChatGPT, we noticed its avoidance of traditionally feminized personas through deanthropomorphized text-based interactions, despite its recent incorporation of voice capabilities. Our observations on current trends indicate that GPT technologies and digital assistants will likely converge and the humanization of digital assistants seems to be an intentional and crucial step to foster user trust and ensure data collection – namely through gender instrumentalization. Although digital assistants often state they have no gender, they aren't being truly and intentionally developed towards being genderless as the labor structures they replicate remain mostly unchanged and they keep instrumentalizing and reinforcing gender stereotypes. These entities can hardly be perceived as neutral and ungendered.

Motivated by our theoretical and analytical findings and complementing their debates, the project seeks to incite reflection on the issues that emerge with gender attribution in AI. Across its different audiovisual pieces, *Cistem Not Found* (https://genderai.pt) explores alternatives that defy anthropomorphization and expose the inclusion of gender-neutral options to counter feminization. It plays with gender fluidity and ambiguity to highlight stereotypes in current digital assistants and speculates on the possibilities of incorporat-

ing fluid, trans or nonbinary gender approaches. We designed an irreverent, disruptive and mutant bot that intentionally distorts common stereotypes and serves as a critical commentary on issues that currently affect AI development – solutionism, veiled surveillance, algorithmic bias, gender instrumentalization and strategies on bypassing the tendency towards cis-feminization. Solutions like diversification, increased customization and so-called gender-neutral voices have proven to be insufficient, as they are rooted in cis-normative conceptions of femininity and masculinity. The project seeks to challenge these limitations, evoking nonconforming gender identities outside the binary frame and exploring gender alternatives through sound manipulation and different voice proposals. Taking on a speculative and critical stance, *Cistem Not Found* aims to raise awareness and foster debate on how our social and cultural views influence current developments in AI, namely digital assistants that are proliferating into our lives as daily companions that fulfill traditional roles. The project challenges common assumptions and stereotypes about gender and strives to move beyond them by deconstructing bioessencialist conceptions and embracing otherness and queerness.

Interpretation of Findings

In summary, despite efforts towards diversity, our observations revealed that current digital assistants continue to reinforce traditional conceptions of femininity and, particularly, cis-femininity. Subtly, these entities continue to perpetuate harmful stereotypes by associating femininity with caregiving, domestic and submissive acts. Moreover, their anthropomorphized attributes tend to instrumentalize gender to conceal underlying intentions of data collection and veiled surveillance. In turn, this process contributes to the perpetuation of social and cultural bias, facilitated by discriminatory algorithms.

Accordingly, we notice how current guidelines are starting to take into account social and cultural issues surrounding the anthropomorphization of AI. Most recommendations focus on avoiding discrimination, being more conscious, and encouraging inclusion. However, there is still a lack of clear guidance and direction from key ethical bodies as they tend to address gender regarding AI's role in society instead of addressing the unique and specific questions that emerge when gender is attributed to humanized daily companions. One-size-fits-all frameworks rooted in Western-normative concepts of so-called ethical AI fail to acknowledge the interconnections between the social and technical dimensions of these entities – merely tweaking preexisting principles without fundamentally changing the practices that uphold patterns of domination or power asymmetries is insufficient and only perpetuates the same systemic injustices. By generalizing gender through un-

equivocal and hegemonic anthropomorphized and behavioral attributes, these systems fail to acknowledge how gender understandings, both at a personal and social level are rooted in intimate, nuanced and complex subjectivities.

Our research addressed this subject matter through the lenses of a critical, intersectional and queer viewpoint in response to current discussions that often approach these topics through an implicitly cisgender and binary frame. While notions of gender equality and equity have proliferated within cisgender understanding of man and woman, debates surrounding gender attribution in AI tend to ignore trans, nonbinary and nonconforming gender identities, and current studies should widen their range of discussion to include gender notions outside of a cis-binary frame. Our study aimed to provide a more nuanced understanding of this phenomenon, considering the complexities of identity within intersecting social contexts alongside the impact of this technology in our daily life.

Our findings draw attention to how gender presence in AI is complex, emphasizing the need for a more conscious approach towards fostering diversity and neutrality in the design of these systems. Their conception and development should acknowledge the historical past of the tasks they perform as well as the stereotypical behavior they exhibit, instead of just expanding the voice options or program assistants to claim inclusivity. Consequently, merely replicating cisgender and naturalistic notions of how masculine and feminine voices are expected to sound proves inadequate to openly and thoroughly tackle the complexities of this phenomenon.

Accordingly, *Cistem Not Found* strived to displace and challenge cisgender notions and stereotypes of femininity and masculinity in AI, exposing how these traditional conceptions have its basis in discriminatory practices of exclusion and shedding light on the profound and complex specificities of gender, sex and identity. Through its different audiovisual pieces and voice proposals, *Cistem Not Found* also speculates on scenarios where the boundaries between human and machine are blurred, seeking to reimagine gender possibilities beyond the cis-binary frame. While there is some awareness on the issues surrounding gender attribution in AI within the traditional cis man-woman viewpoint, a lot remains to be done outside of this gender dichotomy to address gender diversity. Designers should actively question and subvert these norms for more inclusive practices that encompass a wider spectrum of gender identities.

Additionally, the significance of this study's approach is attested by the positive reception our research has received in several presentations and publications such as Reboot 2019

in Lisbon (New Media and Digital Art Festival), SS2019 in Lisbon (Senses & Sensibility Conference), TA2019 in Porto (Technoetic Arts: A Journal for Speculative Research), xCoAx 2019, 2020 and 2022 in Milan and Coimbra (Conference on Computation, Communication, Aesthetics & X), Incubadora DAO 2022 in Lisbon (Laboratory for Experimentation in Arts + Technology) and ISEA 2020 and 2023 in Toronto and Paris (International Symposium on Electronic Art). The feedback we received overall confirmed our research's aim to raise awareness on the implications of attributing gender to artificial intelligence, particularly among researchers and academics familiar with the field of AI and its impacts on our daily lives. Additionally, this research emphasized the need to broaden discussions on gender, considering that many individuals were unfamiliar with gender conceptions beyond a cis-binary frame and lack familiarity with the subjectivities and nuances of nonbinary and transgender identities. It also fostered discussions on the consequences of feminizing these entities and how they tend to reinforce stereotypes and spread harmful bias. Furthermore, this research raised questions about potential solutions to this phenomenon and the possibilities of queering this technology to diversify its anthropomorphized attributes and challenge conventional notions of gender.

Lastly, the reasoning and understanding of these subject matters seem to be evolving unevenly and at different paces. On one hand, cis-normative binary conceptions of gender are still predominant, and although current understandings of gender are expanding and challenging traditional roles and stereotypes, further discussion is needed at a social, political and structural level. On another hand, artificial intelligence has emerged as an omnipresent field, rapidly expanding its influence into various domains of study and aspects of our daily life, such as education, science, healthcare, entertainment, the media and the arts. Consequently, the accelerated growth of artificial intelligence outpaces current debates that are still questioning, redefining and reshaping our understandings of gender.

As this study draws to an end, Portugal is amidst a political crisis and elections are scheduled to be held soon. While discussions surrounding artificial intelligence have gathered significant attention and are prominently featured in political debates, gender nuances and complexities remain overlooked. Among eight political parties' platforms and programs, all of them address issues related to artificial intelligence, including concerns about data collection, biased algorithms and the ethical implications of these technologies' rapid development, advancing proposals that increase transparency and establish ethical guidelines for the development of these systems. In contrast, only two political parties acknowledge current gender issues and recognize the challenges faced by trans and nonbinary identities, offering proposals that reveal a more conscious understanding

of gender outside a binary frame. This discrepancy underlines how these subject matters are asynchronous in their evolution and while advancements in technology are important and often brought to the forefront, the recognition and addressing of gender discrimination, equality and equity lags behind.

Limitations

One of the limitations of this work pertains to the complexities and intersections of gender identity among other identity traits. While we adopted a queer and intersectional approach that challenges cis-normative conceptions of gender, we did not extensively explore dimensions such as race, disability or class, which would substantially expand the scope of our research. Therefore, we chose to concentrate on the nuances of trans and nonbinary identities to clearly differentiate these from prevailing cisgender and binary norms and highlight relevant stereotypes that derive from them.

Similar decisions were made regarding the inclusion of additional information on artificial intelligence. Due to the rapid evolution of this technology, we needed to closely monitor new developments in this field and among digital assistants. For instance, ChatGPT had not yet been publicly introduced when we began our study. This required us to continually update the theoretical and analytical components to ensure our overview of digital assistants remained current and up to date. Thus, the discussion of these subject matters was circumscribed to preserve the focus on the particular type of AI technology we examined.

On this note, another aspect to consider is the pragmatic yet intentional choice to focus on four case studies, limiting the number of examples discussed (along with a complementary analysis of ChatGPT) to prioritize the most prominent and relevant digital assistants in the current technological landscape. While other assistants like Samsung's Bixby or Huawei's Celia have recently entered the scene, they have not been in development for as long as Alexa, Google Assistant, Siri or even Cortana. Including them could potentially affect the results of our analysis, which focused on the evolution of these assistants over the years.

At a pragmatic level, *Cistem Not Found* comprises complex sound arrangements and high-definition reactive visualizations, resulting in heavy video files. While we deliberately chose this format to ensure the audiovisual quality of the project and enable viewers to follow its narrative, it also presents challenges regarding accessibility. Downloading and buffering these files may prove to be demanding, particularly when accessing the

project online. Finding solutions to optimize accessibility while preserving the integrity of *Cistem Not Found* is an important consideration moving forward.

Future Research

Building upon the limitations addressed, our future research can begin by expanding *Cistem Not Found*, updating its sonic and visual components to be generated live instead of being rendered to pre-edited files. This could involve resorting to a MIDI controller and live coding techniques to create dynamic and interactive elements. Additionally, working live would allow us to experiment with more voice possibilities through digital sound manipulation, distortion and vocoder techniques. We are particularly interested in locally implementing the project through a performance-based, audiovisual approach that provides a more immersive experience. This approach would also allow us to adapt the project to suit specific contexts and interests of the audience, stimulating deeper engagement and exploration of the themes addressed in our research.

Another possible direction for research entails a separate and complementary analysis that maps new and emerging digital assistants like Bixby and Celia. Examining these newer assistants may provide valuable insights on their approach to anthropomorphization, their roles as assistants and companions. Understanding the ideas that guided their conceptualization and development, as well as identifying any trends in their current designs and functionalities, might also shed light on evolving methods for these entities' characterization and their stance towards gender.

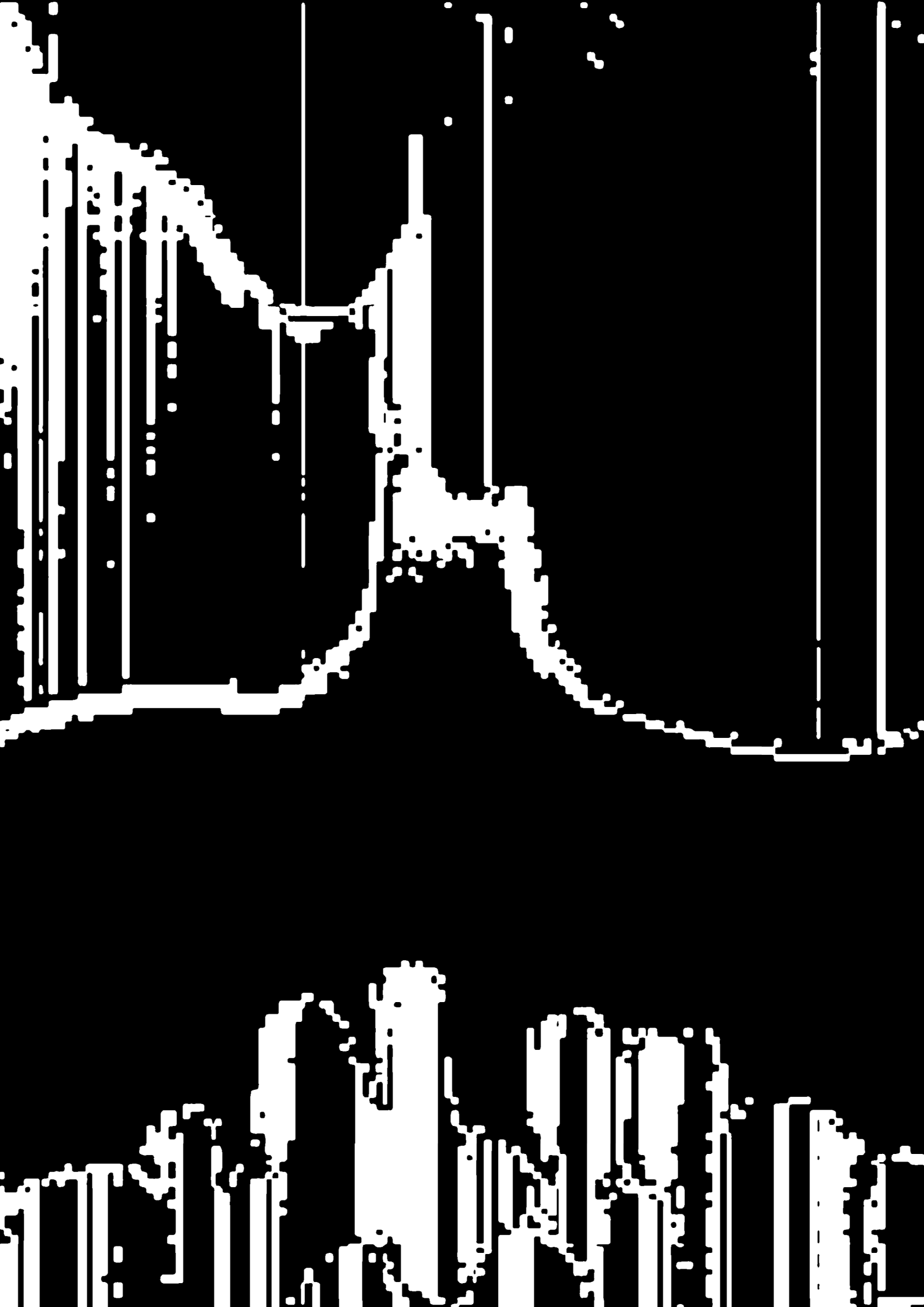
On this note, a more thorough review on how AI technologies are increasingly engaging in data collection practices, raising questions about veiled surveillance and privacy concerns, is a current significant direction for research. Building on the work of authors such as Kate Crawford, Cathy O'Neill, Wendi Chun, Safiya Noble or Virginia Eubanks, this line of research can contribute to highlight how most datasets generated by machine learning systems are biased and highly likely to result in discriminatory models that contribute to the spread of harmful biases and reinforce unfair social conceptions, emphasizing the need for more thorough and careful evaluation processes in order to counter this phenomenon.

It is also important to continue conducting more rigorous and intersectional examinations of artificial intelligence. Particularly, the intersection of gender and race also raises significant issues surrounding this technology, be it through the anthropomorphized attributes

of digital assistants or the biased algorithms they employ, perpetuating problematic and harmful racial stereotypes. Drawing upon frameworks like xenofeminism and Afrofuturism, authors such as Helen Hester and Ruha Benjamin emphasize a pervasive "white guy problem" in AI, which often reinforces racial biases and excludes marginalized identities. They advocate for a decolonization of AI through more conscious and inclusive design practices. However, this research should be approached with caution to avoid tokenism and appropriation, ensuring that it respects and amplifies the voices of non-white scholars in these fields. As such, it is essential to incorporate diverse perspectives and experiences into this discourse, especially from those directly affected by these issues.

Finally, another direction for research derives from the context of the project, with posthumanism emerging as an intriguing and promising approach. By deconstructing traditional notions of humanness, posthumanism exposes how naturalized concepts of being human reinforce cis-binary understanding of gender, sex and identity. Aligned with this anti-naturalist perspective, xenofeminism also proposes that science and technology offer a unique opportunity for conscious interventions that can extend human freedom. The creative approaches of researchers and artists such as Donna Haraway, Manel de Aguas, Paul B Preciado, Stelarc or even Arca, also offer unique perspectives on the intersections of technology, identity and gender. Our interest resides in investigating how posthumanist ideas, when combined with AI technology, can reshape boundaries between nature and culture, embrace anti-essentialism and reject biological determinism to achieve post-gender configurations that suggest more fluid understandings of gender and identity.

With this study, we aimed to understand, question and explore the relationship between gender and artificial intelligence, exposing a tendency towards cis-femininity in current digital assistants and speculating on the possibilities that arise with queering these entities. Gender attribution in AI poses specific and unique questions and we sought to provide a more nuanced and intersectional perspective on these issues, sparking debate and encouraging reflection on the phenomenon. Standing at the crossroads of queer expression and rapid technological advancements, new modes of blurring the boundaries between human and machine come to surface. This allows us to challenge conventional paradigms and transcend the limitations of binary constructs. It's the gap between where we are now and where we could be – we need not to be bound by the constraints of the past but be propelled towards a future where gender identity is not fixed but infinitely mutable. Rather than perpetuating stereotypes and reinforcing cis-binary conceptions of gender, perhaps these technologies can aid the liberating and empowering opportunities that emerge with identity representation and embodiment in the digital age.



BIBLIOGRAPHY

Accenture. 2020. "Accenture and CereProc Introduce and Open Source the World's First Comprehensive Non-Binary Voice Solution." Accenture. <https://newsroom.accenture.com/news/2020/accenture-and-cereproc-introduce-and-open-source-the-worlds-first-comprehensive-non-binary-voice-solution> [last access: 30-01-2024].

ACLU. 2020. "Supreme Court Says Firing Workers Because They Are LGBTQ Is Unlawful Discrimination." ACLU. <https://www.aclu.org/news/lgbtq-rights/supreme-court-says-firing-workers-because-they-are-lgbtq-is-unlawful-discrimination> [last accessed: 29-01-2024].

Adamopoulou, Eleni, and Lefteris Moussiades. 2020. "Chatbots: History, Tecnology, and Applications." *Machine Learning with Applications*. 2 (53).

Adam, Alison. 1995. "Embodying Knowledge: A Feminist Critique of Artificial Intelligence." *European Journal of Women's Studies*. 2:355-377.

Adam, Alison. 1996. "Constructions of Gender in the History of Artificial Intelligence." *IEEE Annals of the History of Computing*. 18 (3).

Adam, Alison. 1998. *Artificial Knowing – Gender and the Thinking Machine*. London and New York: Routledge.

Agrawal, Anuja. 1997. "Gendered Bodies: The Case of the 'Third Gender' in India". *Contributions to Indian Sociology*. 31 (2):273-297.

Ainsworth, Claire. 2015. "Sex Redefined: The Idea of 2 Sexes Is Overly Simplistic." *Nature Magazine*. 518:288-291.

Aljanabi, Mohammad. 2023. "ChatGPT: Future Directions and Open possibilities." *Mesopotamian Journal of CyberSecurity*. 16-17.

Amazon. 2023. "What Is Alexa?" Alexa Developer Official Site. <https://developer.amazon.com/en-IN/alexa> [last access: 30-01-2024].

Anders, Charlie Jane. 2015. "From Metropolis to Ex Machina: Why Are So Many Robots Female?" Gizmodo. <https://gizmodo.com/from-maria-to-ava-why-are-so-many-artificial-intellige-1699274487> [last access: 30-01-2024].

Apple. 2021. "MacOS User Guide: Use Siri on Mac." Apple Support. <https://support.apple.com/en-vn/guide/mac-help/mchl6b029310/mac> [last access: 30-01-2024].

Arca. 2015. “Arca’s Warped Beauty.” By Philip Sherburne. Pitchfork. <https://pitchfork.com/features/interview/live-from-quarantine-its-the-arca-show/> [last access: 30-01-2024].

Arca. 2020a. “Live from Quarantine, It’s the Arca Show.” By Philip Sherburne. Pitchfork. <https://pitchfork.com/features/interview/9754-arcas-warped-beauty/> [last access: 30-01-2024].

Arca. 2021a. “Arca Announces New Album KicK iii, Shares Video for New Song: Watch.” By Matthew Ruiz. Pitchfork. <https://pitchfork.com/news/arca-announces-new-album-kick-iii-shares-video-for-new-song-watch/> [last access: 31-01-2024].

Arca (@arca1000000). 2021b. “Kick 4.” Instagram Post. <https://www.instagram.com/tv/CWbC84qADz-/?utm_source=ig_web_copy_link&igsh=MzRlODBiNWFlZA==> [last access: 31-01-2024].

Arca. 2021c. “Arca Once Made Electronic Music. Now She Builds Worlds.” By Isabelia Herrera. The New York Times.

Archer, James. 2021. “Alexa in 2021: Here’s What to Expect.” Tom’s Guide. <https://www.tomsguide.com/features/what-to-expect-from-alexa-in-2021> [last access: 30-01-2024].

Auner, Joseph. 2003. “‘Sing it for Me’: Posthuman Ventriloquism in Recent Popular Music.” *Journal of the Royal Musical Association*. 128 (1):98-122.

Baboolall, David, Greenberg, Sarah, Obeid, Maurice, and Jill Zucker. 2021. “Being Transgender at Work.” McKinsey & Company. <https://www.mckinsey.com/featured-insights/diversity-and-inclusion/being-transgender-at-work#/> [last access: 29-01-2024].

BBC. 2020. “The Matrix Is a ‘Trans Metaphor’, Lilly Wachowski Says.” BBC News. <https://www.bbc.com/news/newsbeat-53692435> [last access: 30-01-2024].

Beauvoir, Simone de. 1949. *The Second Sex*. Translated by H. M. Parshley. London: Jonathan Cape, 1953.

Bell, Karissa. 2017. “Hey, Siri: How’d you and every other digital assistant get its name?”. Mashable. <https://mashable.com/article/how-alexa-siri-got-names> [último acesso: 01-02- 2024].

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

Berentsen, Mirthe. 2018. “How we became artists: Net Art duo JODI.” Art Basel. <https://www.artbasel.com/stories/how-we-became-artists-jodi> [last access: 01-02-2024].

Bergen, Hillary. 2016. “‘I’d Blush If I Could’: Digital Assistants, Disembodied Cyborgs and the Problem of Gender.” *Word and Text: A Journal of Literary Studies and Linguistics*. 6:95-113.

Bird, Sarah, Barocas, Solon, Crawford, Kate, Diaz, Fernando and Hanna Wallach. 2016. “Exploring or Exploiting? Social and Ethical Implications of Autonomous Experimentation in AI.” *Workshop on Fairness, Accountability, and Transparency in Machine Learning.*

Björk. 2016. “Björk on nature and technology.” By Brandon Stosuy. The Creative Independent. <https://thecreativeindependent.com/people/bjork-on-nature-and-technology/> [last access: 30-01-2024].

Bogost, Ian. 2018. “Sorry, Alexa Is Not a Feminist.” The Atlantic. <https://www.theatlantic.com/technology/archive/2018/01/sorry-alexa-is-not-a-feminist/551291/> [last access: 29-01-2024].

Borau, Sylvie. 2021. “Female Robots Are Seen as Being the Most Human. Why?” India AI. <https://indiaai.gov.in/article/female-robots-are-seen-as-being-the-most-human-why> [last access: 30-01-2024].

Bronstein, Manuel. 2019. “Bringing You the Next-Generation Google Assistant.” Google Blog. <https://blog.google/products/assistant/next-generation-google-assistant-io/> [last access: 30-01-2024].

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

Butler, Judith. 1990. *Problemas de Género: Feminismo e Subversão da Identidade*. Translated by Nuno Quintas. Lisboa: Orfeu Negro, 2017.

Cave, Stephen and Kanta Dihal. 2020. “The Whiteness of AI.” Philosophy & Technology. 33 (2):1-19.

Cave, Stephen, Craig, Claire, Dihal, Kanta, Dillon, Sarah, Montgomery, Jessica, Singler, Beth, and Lindsay Taylor. 2018. *Portrayals and Perceptions of AI and Why They Matter.* Report, University of Cambridge. London: The Royal Society.

Chambers, Amy. 2018. “There’s a Reason Siri, Alexa and AI Are Imagined as Female - Sexism.” The Conversation. <https://theconversation.com/theres-a-reason-siri-alexa-and-ai-are-imagined-as-female-sexism-96430> [last access: 30-01-2024].

Cheng, Selina. 2016. “An Algorithm Rejected an Asian Man’s Passport Photo for Having Closed Eyes.” Quartz. <https://qz.com/857122/an-algorithm-rejected-an-asian-mans-passport-photo-for-having-closed-eyes/> [last access: 29-01-2024].

Chen, Brian, Grant, Nico, and Karen Wise. 2023. “How Siri, Alexa and Google Assistant Lost the A.I. Race.” The New York Times. <https://www.nytimes.com/2023/03/15/technology/siri-alexa-google-assistant-artificial-intelligence.html> [last accessed: 29-01-2024].

Chorus. 2024. “A Conversational Agent Powered by Crowdsourcing.” Talking to the Crowd. <http://talkingtothecrowd.org/index.html> [last access: 30-01-2024].

Christensen, Jen. 2023. “Transgender People Face Significantly Higher Suicide Risk, Danish Study Finds.” CNN Health. <https://edition.cnn.com/2023/06/28/health/transgender-suicide-risk/index.html> [last access: 02-02-2024].

Córdoba, Antonio. 2023. “'A Mutant Faith': Science Fiction, Posthumanism, and Queer Futurity in Arca’s KiCK Album Pentalogy.” In: *Posthumanism and Latin(x) American Science Fiction*. *Studies in Global Science Fiction*. Switzerland: Palgrave Macmillan. 219-235.

Costa, Pedro. 2019. *Conversations with ELIZA: Sobre Género e Inteligência Artificial*. Master’s thesis. Faculdade de Belas-Artes da Universidade de Lisboa.

Costa, Pedro, and Luísa Ribas. 2019. “Some Things You Can Ask Me: About Gender and Artificial Intelligence.” *Proceedings of xCoAx 2019: 7th Conference on Computation, Communication, Aesthetics and X*. Milan. 49-66.

Costa, Pedro, and Luísa Ribas. 2020. “‘I Try to Stay Neutral’: Digital Assistants and Their Stance Towards Gender.” *Proceedings of ISEA2020: 26th International Symposium on Electronic Art*. Montreal. 465-468.

Costa, Pedro, Ribas, Luísa, and Miguel Carvalhais. 2022. “‘We’re the Brains, You’re Talking About Bodies’: Discussing Gender Stereotypes in Digital Assistants.” *Proceedings of xCoAx 2022: 10th Conference on Computation, Communication, Aesthetics and X*. Coimbra. 45-57.

Costa, Pedro, Ribas, Luísa, and Miguel Carvalhais. 2023. “‘I’m a Virtual Assistant So I Don’t Have Pronouns the Way People Do, Thanks for Asking’: Gender Neutrality, Diversification and Fluidity in AI.” *Proceedings of ISEA2023: 28th International Symposium on Electronic Art*. Paris.

Crawford, Kate. 2017. “The Trouble with Bias.” C*onference on Neural Information Processing Systems. Keynote*. <https://www.youtube.com/watch?v=fMym_BKWQzk> [last access: 30-01-2024].

Crawford, Kate and Vladan Joler. 2018. “Anatomy of an AI System: The Amazon Echo as an Anatomical Map of Human Labor, Data and Planetary Resources.” *AI Now Institute and Share Lab*. <https://anatomyof.ai> [last access: 29-01-2024].

Crawford, Kate, and Trevor Paglen. 2019. "Excavating AI: The Politics of Images in Machine Learning Training Sets." *AI Now Institute*. <https://excavating.ai/> [last access: 30-01-2024]

Crawford, Kate. 2021. *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. New Haven and London: Yale University Press.

Curry, Amanda Cercas, and Verena Reiser. 2018. "#MeToo Alexa: How Conversational Systems Respond to Sexual Harassment." *Second ACL Workshop on Ethics in Natural Language Processing.* New Orleans, Louisiana, USA.

Dale, Robert. 2016. "Industry Watch: The Return of the Chatbots." *Natural Language Engineering.* 22 (5):811-817.

Dastin, Jeffrey. 2023. "Google wants AI chatbot Bard to help it reach billions of users." Reuters. <https://www.reuters.com/technology/reuters-next-google-wants-ai-chatbot-bard-help-it-reach-billions-users-2023-11-09/> [last access: 31-01-2024].

Deaux, Kay, and Brenda Major. 1987. "Putting Gender into Context: An Interactive Model of Gender-Related Behavior." *Psychological Review*. 94 (3):369-389.

Dely, Carole. 2007. *Jacques Derrida: The Perchance of a Coming of the Otherwoman*. International Phenomenology Colloquium at the Catholic University of Porto Alegre, Brazil. Translated by Wilson Baldridge. Sens Public.

Derrida, Jacques. 1981. "Plato's Pharmacy." In *Dissemination*. Translated by Barbara Johnson. London: Athlone. 61-171

Derrida, Jacques. 1987. *Psyche: Inventions of the Other, Volume I.* Translated by the Board of Trustees of the Leland Stanford Junior University. Stanford, CA: Stanford University Press, 2007.

Dorssen, Artemis V. 2023. "How Afrofuturism in AI Art Is Exposing Biases in the System." Far Out Magazine. <https://faroutmagazine.co.uk/how-afrofuturism-in-ai-art-is-exposing-biases-in-the-system/> [last access: 02-02-2024].

Ellemers, Naomi. 2018. "Gender Stereotypes." *Annual Review of Psychology.* 69:275-298.

Erscoi, Lelia, Annelies Kleinherenbrink, and Olivia Guest. 2023. "Pygmalion Displacement: When Humanising AI Dehumanises Women." *SocArXiv.*

ESCAP. 2023. "Bridging the Digital Gender: How Girls and Women Can Lead the Way." ESCAP. <https://www.unescap.org/story/bridging-digital-gender-divide-how-girls-and-women-can-lead-way> [last access: 29-01-2024]

European Commission. 2019. "Ethics Guidelines for Trustworthy AI." Official Website of the European Union. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> [last access: 30-01-2024].

European Commission. 2024. "Data Protection in the EU." Official Website of the European Union. <https://commission.europa.eu/law/law-topic/data-protection/data-protection-eu_en> [last access: 30-01-2024].

Eyssel, Friederike, and Frank Hegel. 2012. "(S)he's Got the Look: Gender Stereotyping of Robots." *Journal of Applied Social Psychology.* 42 (9):2213-2230.

Faber, Liz W. 2020. *The Computer's Voice: From Star Trek to Siri*. London: University of Minnesota Press.

Feinberg, Leslie. 1992. *Transgender Liberation: A Movement Whose Time Has Come*. New York: World View Forum.

Fessler, Leah. 2017. “Siri, Define Patriarchy: We Tested Bots Like Siri and Alexa to See Who Would Stand Up to Sexual Harassment.” Quartz. <https://qz.com/911681/we-tested-apples-siri-amazon-echos-alexa- microsofts-cortana-and-googles-google-home-to-see-which-personal-assistant-bots-stand-up-for- themselves-in-the-face-of-sexual-harassment/> [last access: 29-01-2024].

Fessler, Leah. 2018. “So Long, Sexism: Amazon’s Alexa is now a feminist, and she’s sorry if that upsets you.” Quartz. <https://qz.com/work/1180607/amazons-alexa-is-now-a-feminist-and-shes-sorry-if-that-upsets-you> [last access: 30-01-2024].

Floridi, Luciano. 2008. *Artificial Intelligence’s New Frontier: Artificial Companions and the Fourth Revolution*. St Cross College and IEG, University of Oxford.

Foucault, Michel. 1976. *The History of Sexuality.* Translated by Robert Hurley. New York: Pantheon Books, 1978.

Friedman, Batya, and Helen Nissenbaum. 1996. “Bias in Computer Systems.” *ACM Transactions on Information Systems*. 14 (3):330-347.

Gane, Nicholas, and David Beerl. 2008. “Interface (Chap. 4).” In *The Key Concepts*. Oxford, New York: Berg.

GenderLess Voice. 2019. “Meet Q: The First Genderless Voice.” Gender Less Voice. <https://www.genderlessvoice.com/> [last access: 30-01-2024].

Gershenson, Carlos. 2014. “Info-computationalism or Materialism? Neither and Both.” *Constructivist Foundations: An Interdisciplinary Journal*. 9 (2):241-242.

Goertzel, Benjamin. 2018. “Sophia the Hanson Robot singing “All is Full of Love” at Clockenflap 2016.” Youtube Video. <https://www.youtube.com/watch?v=IfBKDwYo3R4> [last access: 03-02-2024].

Good, Nathan, and Chris Wilk. 2016. “Introducing the Guardian Chatbot.” The Guardian. <https://www.theguardian.com/help/insideguardian/2016/nov/07/introducing-the-guardian-chatbot> [last access: 30-01-2024].

Google. 2021. “Manage Audio Recordings in Your Web & App Activity.” Google Account Help. <https://support.google.com/accounts/answer/6030020?hl=en&co=GENIE.Platform%3DAndroid> [last access: 30-01-2024].

Grady, Denise. 2018. “Anatomy Does Not Determine Someone’s Gender, Experts Say.” The Independent. <https://www.independent.co.uk/news/world/americas/gender-anatomy-science-biology-trans-trump-a8596971.html> [last access: 04-02-2024].

Groover, Mikell P. 1996. “Automation Fundamentals.” In *Fundamentals of Modern Manufacturing: Materials, Processes, and Systems*. Hoboken, New Jersey: Wiley. 886-887.

Grudin, Jonathan, and Richard Jacques. 2019. “Chatbots, Humbots, and the Quest for Artificial General Intelligence.” *Proceedings of the 2019 CHI Conference on Human Factors in Computing* Systems. Glasgow.

Guzie, Tad and Noreen Monroe Guzie. 1984. “Masculine and Feminine Archetypes: A Complement to the Psychological Types.” *Journal of Psychological Type*. 7:3-11.

Guggenheim Bilbao. 2024. “Survival Series: Let your hand wander…” Guggenheim Bilbao. <https://www.guggenheim-bilbao.eus/en/the-collection/works/survival-series-let-hand-wander> [last access: 03-02-2024].

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

Halberstam, Jack. 2018. "Vertiginous Capital or, the Master's Toolkit." Bully Bloggers. <https://bullybloggers.wordpress.com/2018/07/02/vertiginous-capital-or-the-masters-toolkit-by-jack-halberstam/> [last access: 29-01-2024].

Haraway, Donna. 1991. "A Cyborg Manifesto: Science, Technology and Socialist-Feminism in the Late Twentieth Century." In *Simians, Cyborgs, and Women: The Reinvention of Nature*. London: Free Association Books.

Haraway, Donna. 2003. *The Companion Species Manifesto: Dogs, People, and Significant Otherness*. Chicago: Prickly Paradigm Press.

Hembree, Wylie, Cohen-Kettenis, Penny, Gooren, Louis, Hannema, Sabine, Meyer, Walter, Murad, Hassan, Rosenthal, Stephen, Safer, Joshua, Tangpricha, Vin and Guy T'Sjoen. 2017. "Endocrine Treatment of Gender-Dysphoric/Gender-Incongruent Persons: An Endocrine Society* Clinical Practice Guideline." *An Endocrine Society Clinical Practice Guideline: J Clin Endocrinol Metab.* 102 (11):3869-3903.

Hern, Alex. 2018. "Google's Solution to Accidental Algorithmic Racism: Ban Gorillas." The Guardian. <https://www.theguardian.com/technology/2018/jan/12/google-racism-ban-gorilla-black-people> [last access: 29-01-2024].

Hern, Alex. 2019. "Apple made Siri deflect questions on feminism, leaked papers reveal." The Guardian. <https://www.theguardian.com/technology/2019/sep/06/apple-rewrote-siri-to-deflect-questions-about-feminism> [last access: 30-01-2024].

Hester, Helen. 2016. "Technology Becomes Her." New Vistas. 3 (1):46-50.

Hester, Helen. 2018. *Xenofeminism*. Medford: Polity.

Hicks, Mar. 2017. *Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computin*g. Cambridge, MA: MIT Press.

High, Rob. 2012. *The Era of Cognitive Systems: An Inside Look at IBM Watson and How It Works*. IBM Redbooks.

Hofstadter, Douglas. 1995. "The Ineradicable Eliza Effect and Its Dangers". In *Fluid Concepts & Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought*. New York: Basic Books.

Hoy, Matthew B. 2018. "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants." *Medical Reference Services Quarterly.* 37 (1):81-88.

IBM. 2024. "What Are Neural Networks?" IBM. <https://www.ibm.com/topics/neural-networks> [last access: 31-01-2024].

IEEE. 2019. "Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems." *2017 IEEE Canada International Humanitarian Technology Conference*. Toronto: Canada.

Ito, Joichi. 2017. "Resisting Reduction: A Manifesto – Designing our Complex Future with Machines." Translated by H. Yamagata. *Journal of Design and Science*. <https://jods.mitpress.mit.edu/pub/resisting-reduction/release/19>. [last access: 04-02-2024].

Jabberwacky. 2011. "About the Jabberwacky AI." Jabberwacky: Communication, Companionship, Intelligence. <http://www.jabberwacky.com/j2about> [last access: 31-01-2024].

Johnson, Joy, and Robin Repta. 2012. "Sex and Gender: Beyond the Binaries." In *Designing and Conducting Gender, Sex, and Health Research*. 17-37. Thousand Oaks: SAGE.

Kerr, Alison Duncan. 2021. "Artificial Intelligence, Gender, and Oppression." *Encyclopedia of the UN Sustainable Development Goals: Gender Equality*. 54-64. Cham: Springer.

Keyes, Os. 2018. “The Misgendering Machines: Trans/HCI Implications of Automatic Gender Recognition.” *Proceedings of the ACM on Human-Computer Interaction*. 2 (CSCW), Article 88:1-22.

Khan, Rashid, and Anik Das. 2018. *Build Better Chatbots: A Complete Guide to Getting Started with Chatbots*. Berkeley, CA: Apress.

Knight, Will. 2023. “Amazon Upgrades Alexa for the ChatGPT Era.” Wired. <https://www.wired.com/story/amazon-upgrades-alexa-for-the-chatgpt-era/> [last access: 29-01-2024].

Kucherbaev, Pavel, Bozzon, Alessandro, and Geert-Jan Houben. 2018. “Human-Aided Bots.” *IEEE Internet Computing*. 22 (6):36-43.

Kurian, Thomas. 2023. “The Next Generation of AI for Developers and Google Workspace.” Google: The Keyword. <https://blog.google/technology/ai/ai-developers-google-cloud-workspace/> [last access: 29-01-2024].

LaFrance, Adrienne. 2016. “Why Do So Many Digital Assistants Have Feminine Names? Hey Cortana. Hey Siri. Hey girl.” The Atlantic. <https://www.theatlantic.com/technology/archive/2016/03/why-do-so-many-digital-assistants-have-feminine-names/475884> [last access: 31-01-2024].

Lever, Emily. 2018. “I Was a Human Siri.” Intelligencer, New York Magazine. <https://nymag.com/intelligencer/smarthome/i-was-a-human-siri-french-virtual-assistant.html> [last access: 31-01-2024].

Lindblomm, Jessica. 2001. *Social Situatedness of Natural Language and Artificial Language*. M.Sc, Department of Computer Science, University of Skõvde.

Lucas, Duncan. 2002. *Body, Mind, Soul - The ‘Cyborg Effect’: Artificial Intelligence in Science Fiction*. Master’s thesis. School of Graduate Studies of McMaster University.

MACBA. 2024. “Jenny Holzer: Inflammatory Essays, 1979-1982.” Museo de Arte Contemporáneo de Barcelona (MACBA). <https://macba.cat/en/art-artists/artists/holzer-jenny/inflammatory-essays> [last access: 31-01-2024].

Manyfesto. 2021. “AI Decolonial Manyfesto.” Manyfesto. <https://manyfesto.ai> [last access: 04-02-2024].

Marino, Mark C. 2014. “The Racial Formation of Chatbots.” *CLCWeb: Comparative Literature and Culture*. 16 (5), Article 13.

McNabb, Charlie. 2018. *Nonbinary Gender Identities: History, Culture, Resources*. Lanham, Maryland: Rowman & Littlefield.

Microsoft. 2014. “Anticipating More from Cortana.” Microsoft Research Blog. <https://www.microsoft.com/en-us/research/blog/anticipating-more-from-cortana/> [last access: 30-01-2024].

Microsoft Corporation. 2018. “Responsible Bots: 10 Guidelines for Developers of Conversational AI.” Microsoft Publication.

Mills, Jack, and Bec Evans. 2018. “Watch SOPHIE and Sophia the Robot Discuss the Future of Creativity.” Dazed. <https://www.dazeddigital.com/science-tech/article/41362/1/sophie-sophia-the-robot-video-interview-age-of-ai> [last access: 03-02-2024].

Mitchell, Melanie. 2019. *Artificial Intelligence: A Guide for Thinking Humans*. New York: Farrar, Straus and Giroux.

Moore, Robert, and Douglas Gillette. 1991. *King, Warrior, Magician, Lover: Rediscovering the Archetypes of the Mature Masculine*. New York: HarperCollins.

MOMA. 2023. “Hito Steyerl: How Not to Be Seen: A Fucking Didactic Educational .MOV File, 2013.” Museum of Modern Art (MOMA). <https://www.moma.org/collection/works/181784> [last access: 31-01-2024].

Mombaça, Jota. 2017. “A Coisa Tá Branca!” Buala. <https://www.buala.org/pt/mukanda/a-coisa-ta-branca> [last access: 29-01-2024].

Moran, Taylor C. 2020. “Racial Technological Bias and the White, Feminine Voice of AI VAs.” *Communication and Critical/Cultural Studies*. 18 (1):1-18.

Morozov, Evgeny. 2013a. “The Perils of Perfection”. The New York Times. <https://www.nytimes.com/2013/03/03/opinion/sunday/the-perils-of-perfection.html> [último acesso: 29-01-2024].

Morozov, Evgeny. 2013b. “Solutionism and Its Discontents”. In *To Save Everything, Click Here*: the Folly of Technological Solutionism. New York: Public Affairs.

Mulvin, Dylan. 2021. *Proxies: The Cultural Work of Standing In*. Cambridge, MA and London: The MIT Press.

Muñoz, José Esteban. 2009. “Feeling Utopia.” *Cruising Utopia: The Then and There of Queer Futurity*. New York: NYU Press.

Newman, Christine. 2018. “Biological Sex Is a Social Construct.” Growing Up Transgender. <https://growinguptransgender.com/2018/11/01/biological-sex-is-a-social-construct/> [last access: 04-02-2024].

Nickelsburg, Monica. 2016. “Why Is AI Female? How Our Ideas About Sex and Service Influence the Personalities We Give Machines.” GeekWire. <https://www.geekwire.com/2016/why-is-ai-female-how-our-ideas-about-sex-and-service-influence-the-personalities-we-give-machines/> [last access: 29-01-2024].

Oberhaus, Daniel. 2018. “Explaining Roko’s Basilisk, the Thought Experiment That Brought Elon Musk and Grimes Together.” Vice. <https://www.vice.com/en/article/evkgvz/what-is-rokos-basilisk-elon-musk-grimes> [last access: 31-01-2024].

Ogilvie-Hanson, Eva. 2020. *Posthuman Production: Technology and Embodiment in the Works of SOPHIE and Holly Herndon*. Master’s thesis. Schulich School of Music of McGill University.

Oliveira, Arlindo. 2019. *Inteligência Artificial*. Lisboa: Fundação Francisco Manuel dos Santos.

O’Neill, Cathy. 2016. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Crown.

OpenAI. 2022. “Introducing ChatGPT.” OpenAI Blog. <https://openai.com/blog/chatgpt> [last access: 31-01-2024].

OpenAI. 2023. “ChatGPT Can Now See, Hear, and Speak.” OpenAI Blog. <https://openai.com/blog/chatgpt-can-now-see-hear-and-speak> [last access: 31-01-2024].

Carly Page. 2021. “RIP Cortana: Microsoft Silences Digital Assistant On iOS And Android.” Forbes. <https://www.forbes.com/sites/carlypage/2021/04/01/rip-cor-tana-microsoft-silences-digital-assistant-on-ios-and-an-droid/?sh=425f8bea216b>. [last access: 01-04-2021].

Perez-Marin, Diana, and Ismael Pascual-Nieto. 2011. “From Chatbots to Dialog Systems.” In *Conversational Agents and Natural Language Interaction: Techniques and Effective Practices*. Pennsylvania: Information Science Reference.

Piccini, Gualtiero. 2004. “Functionalism, Computationalism, and Mental States.” *Studies in History and Philosophy of Science*. 35:811-833.

Piper, Allison. 2016. *Stereotyping Femininity in Disembodied Virtual Assistants*. Master's thesis. Iowa State University.

Plant, Sadie. 1997. *Zeros and Ones: Digital Women and the New Technoculture*. Book. London: Fourth Estate.

Preciado, Paul B. 2008. *Testo Junkie: Sex, Drugs, and Biopolitics in the Pharmacopornographic Era*. Translated by Bruce Benderson. New York: The Feminist Press, 2013.

Preciado, Paul B. 2019. *An Apartment on Uranus: Chronicles of the Crossing*. Translated by Charlotte Mandell. South Pasadena, CA: Semiotext(e), 2020.

Preciado, Paul B. 2020a. "We Don't Have Bodies Yet: Paul B. Preciado." Notes from a talk at Glad Day Bookshop. Mike Hoolbloom. <https://mikehoolboom.com/?p=20883> [last access: 29-01-2024].

Preciado, Paul B. 2020b. "Paul B. Preciado: 'One Day We'll See Assigning Gender at Birth as Brutal': The Punk Trans Philosopher on the Future of Gender, How to Declare a Uterus Strike, and What It's Like on the Other Side of the Binary." By Mary Katharine Tramontana. i-D. <https://i-d.vice.com/en/article/jgeb4b/paul-b-preciado-one-day-well-see-assigning-gender-at-birth-as-brutal> [last access: 29-01-2024].

Prentice, Deborah A, and Erica Carranza. 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.

Purtill, Corinne. 2021. "Hey, Alexa, Are You Sexist?" The New York Times. <https://www.nytimes.com/2021/02/12/us/alexa-bots-female-voice.html> [last access: 30-01-2024].

Redzisz, Monika. 2020. "Ivona, Alexa, Vika or Intelligent Girls from Gdańsk." Artificial Intelligence. <https://www.sztucznainteligencja.org.pl/en/ivona-alexa-vika-or-intelligent-girls-from-gdansk/> [last access: 29-01-2024].

Richards, Christina, Bouman, Walter Pierre, and Meg-John Barker. 2017. "Genderqueer and Non-Binary Genders." *Critical and Applied Approaches in Sexuality, Gender and Identity*. London: Palgrave Macmillan.

Richardson, Kathleen. 2015. *An Anthropology of Robots and AI: Annihilation Anxiety and Machines*. New York and London: Routledge.

Rifkin, Jeremy. 1995. "Crossing into the High-Tech Frontier". In *The End of Work: The Decline of the Global Labor Force and the Dawn of the Post-Market Era*. New York: G.P. Putnam's Sons.

Rincón, Cami, Keyes, Os, and Corinne Cath. 2021. "Speaking from Experience: Trans/Non-Binary Requirements for Voice-Activated AI." *Proceedings of the ACM on Human-Computer Interaction*. 5 (CSCW1), Article 132.

Robison, Mishaela. 2020. "Voice Assistants Have a Gender Bias Problem. What Can We Do About It?" Brookings. <https://www.brookings.edu/articles/voice-assistants-have-a-gender-bias-problem-what-can-we-do-about-it/> [last access: 29-01-2024].

Rokeby, David. 1998. "The Construction of Experience: Interface as Content." In *Digital Illusion: Entertaining the Future with High Technology*. Clark Dodsworth, Jr. (Ed.). New York: ACM Press/Addison-Wesley Publishing Co., 27-47.

Rosenwald, Michael S. 2017. "How Millions of Kids Are Being Shaped by Know-It-All Voice Assistants." The Washington Post. <https://www.washingtonpost.com/local/how-millions-of-kids-are-being-shaped-by-know-it-all-voice-assistants/2017/03/01/c0a644c4-ef1c-11e6-b4ff-ac2cf509efe5_story.html> [last access: 29-01-2024].

Russell, Stuart, and Peter Norvig. 2010. "Natural Language Processing". In *Artificial Intelligence: A Modern Approach*. Upper Saddle River, New Jersey: Pearson Education, Inc.

Savenkov, Denis, and Eugene Agichtein. 2016. "CRQA: Crowd-Powered Real-Time Automatic Question Answering System." *Proceedings of the AAAI Conference on Human Computation and Crowdsourcing*. 4 (1):189-198.

Scaff, Julian. "Reimagining Technology and Humanity Through the Afrofuturist Lens." Medium. <https://medium.com/the-futureplex/reimagining-technology-and-humanity-through-the-afrofuturist-lens-8fa308ccad8d> [last access: 04-02-2024].

Scheuerman, Morgan, Paul, Jacob, and Jed Brubaker. 2019. "How Computers See Gender: An Evaluation of Gender Classification in Commercial Facial Analysis Services." *Proceedings of the ACM on Human-Computer Interaction*. 3 (CSCW):1-33.

Searle, John. 2004. "The Saga of Materialism: From Behaviorism to Strong Artificial Intelligence." In *Mind: A Brief Introduction*. New York: Oxford University Press.

Smith, Taj M. 2016. "Sci-fi as a Queer Genre: Imagining New Worlds Opens Up Creative, 'What If' Possibilities." Harvard Divinity Bulletin. <https://bulletin.hds.harvard.edu/sci-fi-as-a-queer-genre/> [last accessed: 29-01-2024].

Smith, Craig. 2024. "Amazon Alexa Statistics and User Count for 2024." DMR. <https://expandedramblings.com/index.php/amazon-alexa-statistics/> [last accessed: 29-01-2024].

Snyder, Mark. 1977. "On the Self-Fulfilling Nature of Social Stereotypes." *Annual Meeting of the American Psychological Association*. San Francisco, California.

SOPHIE. 2017. "Pop wunderkind SOPHIE synthesizes human and machine voices." By Cedar Pasori. Interview Magazine. <https://www.interviewmagazine.com/music/sophie-its-okay-to-cry-interview> [last access: 31-01-2024].

Staniak, Michael. 2017. "The Great Imitation." In *Posthuman Bodies*. Keen On Magazine. <https://readymag.website/u59636260/859296/thegreatimitation/>. [04-04-2024].

Steele, Chandra. 2018. "The Real Reason Voice Assistants Are Female (and Why it Matters)." Medium. <https://medium.com/pcmag-access/the-real-reason-voice- assistants-are-female-and-why-it-matters-e99c67b93bde> [last access: 29-01-2024].

Stone, Sandy. 1987. *The Empire Strikes Back: A Posttranssexual Manifesto*.4th version. Self-published. <http://sandystone.com> [last access: 01-11-2020].

Strengers, Yolande and Jenny Kennedy. 2020. *The Smart Wife: Why Siri, Alexa, and Other Smart Home Devices Need a Feminist Reboot*. Cambridge: The MIT Press.

Suta, Prissadang, Lan, Xi, Wu, Biting, Mongkolnam, Pornchai, and Jonathan Chan. 2020. "An Overview of Machine Learning in Chatbots." *International Journal of Mechanical Engineering and Robotics Research*. 9 (4):502-510.

Sweeney, Miriam. 2020. "Digital Assistants." In *Uncertain Archives*. Cambridge, MA: MIT Press.

TGEU. 2023. "Trans Murder Monitoring 2023 Global Update." Transgender Europe (TGEU). <https://tgeu.org/trans-murder-monitoring-2023/> [last access: 02-02-2024].

Turing, Alan. 1950. "Computing Machinery and Intelligence." In *Mind*. 49:433-460.

Turing, Alan. 1996. "Intelligent Machinery: A Heretical Theory." *Philosophia Mathematica*. 4 (3):256-260.

UNESCO. 2019. *I'd Blush if I Could: Closing Gender Divides in Digital Skills Through Education*. UNESDOC Digital Library. France.

UNFE. 2017. “Fact Sheet: Intersex.” *United Nations Human Rights, Office of the High Commisioner: Free & Equal, United Nations for LGBT Equality.* 1-2.

UNICEF. 2021. *What We Know About the Gender Digital Divide for Girls: A Literature Review.* UNICEF East Asia and Pacific.

UNI Global Union. 2017. *10 Principles for Ethical Artificial Intelligence*. UNI Global Union Report.

VNS Matrix. 2018. “The Artists.” VNS Matrix. <https://vnsmatrix.net/the-artists> [last access: 31-01-2024].

VNS Matrix. 2019. “Bitch Mutant Manifesto.” VNS Matrix. <https://vnsmatrix.net/projects/bitch-mutant-manifesto> [last access: 31-01-2024].

Vogels, Werner. 2016. “Bringing the Magic of Amazon AI and Alexa to Apps on AWS.” All Things Distributed. <https://www.allthingsdistributed.com/2016/11/amazon-ai-and-alexa-for-all-aws-apps.html> [last access: 31-01-2024].

Waddell, Kaveh. 2022. “When People Think Siri Sounds Black, Biases Often Surface.” Consumer Reports. <https://www.consumerreports.org/electronics/digital-assistants/when-siri-sounds-black-biases-often-surface-a1101697156/> [last accessed: 29-01-2024].

Walker, Taylor. 2020. “‘Alexa, Are You a feminist?’: Virtual Assistants Doing Gender and What That Means for the World.” *The IJournal: Student Journal of the University of Toronto’s Faculty of Information*. 6 (1): 1-16.

Weber, Jutta. 2005. “Helpless Machines and True Loving Caregivers: A Feminist Critique of Recent Trends in Human-Robot Interaction.” *Info, Comm & Ethics in Society.* 3:309-218.

Weizenbaum, Joseph. 1966. “ELIZA - A Computer Program for the Study of Natural Language Communication Between Man and Machine.” *Computational Linguistics*. 9 (1):36-45.

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

West, Candace, and Don H. Zimmerman. 1987. “Doing Gender.” *Gender and Society.* 1 (2):125-151.

Wiercioch, Katarzyna, Alexander, Rae, Blennerhassett, and Leenders Ruben. 2018. *Do the Robot: A Critical Alternative to Gendered Humanistic Virtual Assistants*. Masters of Media, New Media & Digital Culture M. A. of University of Amsterdam.

Winston, Ali. 2018. “Palantir Has Secretly Been Using New Orleans to Test Its Predictive Policing Technology.” The Verge. <https://www.theverge.com/2018/2/27/17054740/palantir-predictive-policing-tool-new-orleans-nopd> [last access: 29-01-2024].

Woods, Heather Suzanne. 2018. “Asking More of Siri and Alexa: Feminine Persona in Service of Surveillance and Capitalism.” *Critical Studies in Media Communication*. 35 (4):334-349.

Wooldridge, Michael. 2020. The Road to Conscious Machines: The Story of AI. UK: Penguin Books.

Wooldridge, Michael. 2023. “Artificial Intelligence: A Modern Advance or an Ancient Nightmare?” BBC History Magazine: History Extra. <https://www.historyextra.com/membership/artificial-intelligence-historical-threat-how-recent/> [last access: 29-01-2024]

Zemčík, Tomáš. 2019. “A Brief History of Chatbots.” *Proceedings of the 2019 International Conference on Artificial Intelligence, Control and Automation Engineering.* 14-18.

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

WORKS CITED

Adams, Douglas. 1978. *The Hitch Hiker's Guide to the Galaxy*. United Kingdom: BBC.

Arca. 2020b. *KiCk i*. Studio Album. XL Recordings.

Arca. 2021d. *KICK ii*. Studio Album. XL Recordings.

Arca. 2021e. *KicK iii*. Studio Album. XL Recordings.

Arca. 2021f. *kick iiii*. Studio Album. XL Recordings.

Arca. 2021g. *kiCK iiiii*. Studio Album. XL Recordings.

Asimov, Isaac. 1950. *I, Robot*. United States: Gnome Press.

Blizzard Entertainment. 2022. *Overwatch 2*. United States: Blizzard Entertainment.

Bungie. 2001. *Halo*. United States: Xbox Game Studios.

Cameron, James. 1984. *The Terminator*. United States: Orion Pictures.

Cameron, James. 1986. *Aliens*. United States: 20th Century Fox.

Čapek, Karel. 1920. *Rossumovi Univerzální Roboti (R.U.R.)*. Prague: Aventinum.

Costa, Pedro. 2019. *Conversations with ELIZA*. Website, chatbots. <https://fbaul-dcnm.pt/costa_pedro/convowitheliza/> [last access: 01-02-2024].

Dick, Philip K. 1968. *Do Androids Dream of Electric Sheep?* United States: Doubleday.

Garland, Alex. 2015. *Ex Machina*. United States: A24.

Gibson, William. 1984. *Neuromancer*. United States: Ace.

Holzer, Jenny. 1979-82. *Inflammatory Essays*. Graphic Material, Text. Barcelona: MACBA Collection.

Holzer, Jenny. 1983-85. *Survival*. Billboard, Projection, Space Installation, Text. New York: Artists Rights Society.

Jonze, Spike. 2013. *Her*. United States: Warner Bros.

Kubrick, Stanley. 1968. *2001: A Space Odyssey*. United States: Metro-Goldwyin-Mayer.

Lang, Fritz. 1927. *Metropolis*. Germany: UFA.

Lucas, George. 1977. *Star Wars*. United States: 20th Century Fox.

McCarthy, Lauren. 2017. *LAUREN*. Performance, Video, Space Installation. <https://lauren-mccarthy.com/LAUREN> [last access: 31-01-2024].

McCarthy, Lauren. 2019a. *SOMEONE*. Performance, Space Installation. <https://lauren-mccarthy.com/SOMEONE> [last access: 31-01-2024].

McCarthy, Lauren. 2019b. *Waking Agents*. Space Installation. <https://lauren-mccarthy.com/Waking-Agents> [last access: 31-01-2024].

McCarthy, Lauren. 2020. *I.A. Suzie*. Performance. <https://lauren-mccarthy.com/I-A-Suzie> [last access: 31-01-2024].

McCarthy, Lauren. 2021. *The Changing Room*. Space Installation. <https://lauren-mccarthy.com/The-Changing-Room> [last access: 31-01-2024].

Nolan, Jonathan, and Lisa Joy. 2016-2022. *Westworld*. United States: HBO.

Oshii, Mamoru. 1995. *Ghost in the Shell*. Japan: Shochiku.

Quantic Dream. 2028. *Detroit: Become Human*. France: Sony Interactive Entertainment.

Roddenberry, Gene. 1987. *Star Trek: The Next Generation*. United States: Paramount Domestic Television.

Schur, Michael. 2016. *The Good Place*. United States: NBC.

Scott, Ridley. 1979. *Alien*. United States: 20th Century Fox.

Shelley, Mary. 1818. *Frankenstein; Or, The Modern Prometheus Frankenstein*. London: Lackington, Hughes, Harding, Mavor & Jones.

SOPHIE. 2015. *PRODUCT*. Compilation Album. Numbers.

SOPHIE. 2018. *OIL OF EVERY PEARL'S UN-INSIDES*. Studio Album. MSMSMSM, Future Classic: Transgressive Records.

Spielberg, Steven. 2001. *A.I. Artificial Intelligence*. United States: Warner Bros.

Stanton, Andrew. 2008. *WALL-E*. United States: Walt Disney Pictures.

Stephenson, Neal. 1995. *The Diamond Age: Or, A Young Lady's Illustrated Primer*. United States: Bantam Spectra.

Steyerl, Hito. 2013. *How Not to Be Seen: A Fucking Didactic Educational .MOV File*. Video, Projection. London, Liverpool & Cornwall: Tate Collection.

Valve. 2007. *Portal*. United States: Valve.

Valve. 2011. *Portal 2*. United States: Valve.

Verhoeven, Paul. 1987. *RoboCop*. United States: Orion Pitures.

Villeneuve, Denis. 2017. *Blade Runner 2049*. United States: Warner Bros.

VNS Matrix. 1991. *A Cyberfeminist Manifesto for the 21st Century*. Billboard, Manifesto, Publication, Sound Work, Text. <https://vnsmatrix.net/projects/the-cyberfeminist-manifesto-for-the-21st-century> [last access: 31-01-2024].

VNS Matrix. 1996. *Bitch Mutant Manifesto*. Manifesto, Performance, Text. <https://vnsmatrix.net/projects/bitch-mutant-manifesto> [last access: 31-01-2024].

The Wachowskis. 1999. *The Matrix*. United States: Warner Bros.

Watts, Jon. 2017. *Spider-Man: Homecoming*. United States: Sony Pictures Releasing.

White, Norman. 1987-2002. *The Helpless Robot*. Space Installation, Robotics. Kingston: Agnes Etherington Art Centre Collection.