Representation of Conversational AI in Contemporary Media: A Case Study of Black Mirror

Mirza Zaib Hassan [1] iD   Noman Ahmed Ansari [2]* iD   Qurat Ul Ain Farooq [3] iD

Corresponding Author: Noman Ahmed Ansari (Email: noman.ansari@f.rwu.edu.pk)

Abstract: Amid the growing integration of CAI into our day-to-day activities, media portrayals of AI agency, ethical considerations, and interactions between humans and AI become the determining factors in shaping perceptions. The current body of literature on CAI delves into technical progress and the framing of news but largely ignores the representation of fictional narratives, which portray the shape of power relations, the nature of exploitation, and the boundary of consciousness in human interactions with CAI. This study fills in the gap by scrutinizing CAI attributes (agency, empathy, consciousness) and relationship-building mechanisms and patterns (power, dependence, exploitation, care, control) in three episodes of Black Mirror. Based on the reflexive thematic of the scripts, which is taken as a qualitative measure due to its theoretical attribution to social constructionism, HMC, and STS, the study revealed 132 themes across three research questions. Findings reveal a consciousness spectrum from simulation to sentience to constrained replication with relationships initiating via vulnerability exploitation and progressing through disillusionment, enslavement, or accidental partnership; patterns keep escalating until they reach dominance, totalitarianism, and then grief commercialization leads to multi-level exploitation. The "consciousness paradox," "relationship impossibility thesis," and "liberation prerequisites" framework drives HMC typologies (AI preemptive rights), media studies (speculative ethics), calling for governance control to corporate monopoly asymmetries.

Key Words: Conversational AI, Contemporary Media, Black Mirror, Consciousness Paradox, Media Studies

Introduction

Conversational AI refers to a broad range of applications that are capable of communicating with humans in a similar way to human beings, thanks to the employment of technologies such as Natural Language Processing (NLP) and Machine Learning (ML). It includes anything from simple responses to extensive interactions, which are personalized to individual conversations being applied in a variety of areas, including healthcare, customer service, or as a psychological aide. Context comprehension represents a core technology that allows the conversational agents to maintain the conversation effectively and sensibly with the user, eventually resulting in a more stimulating chat (Caldarini et al., 2022; Schachner et al., 2020).

This made way for the creation of even more intelligent and complex agents that could know and respond emotionally respectively and have personality traits to make the conversation more engaging to the users. Apart from that, conversational agents are also getting smarter with personality cues integrated for better interaction with users (Ahmad et al., 2022). This approach of humanization of chatbots by introducing them to users has turned to permitting the creation of relations with CAIs. McKillop et al. highlight that conversational technology can meet the spectrum of information needs pertaining to COVID-19 without any major limitation, thus underscoring its flexibility and

---

[1] Lecturer, Riphah Institute of Media Sciences, Riphah International University, Islamabad, Pakistan.
Email: mirzazaibhassan@gmail.com
[2] Assistant Professor, Department of Media & Communication Studies, Rawalpindi Women University, Rawalpindi, Punjab, Pakistan. Email: noman.ansari@f.rwu.edu.pk
[3] Lecturer, Riphah Institute of Media Sciences, Riphah International University, Islamabad, Pakistan.
Email: quratulain.2015@gmail.com

appropriateness when it comes to emergencies (McKillop et al., 2021). It is not only chatting with them that becomes easier, but also they act as a source of emotional comfort in case of mental health issues, which ultimately proves that they can solve the problems of real life (Tudor Car et al., 2020; Vaidyam et al., 2019). An example is the systematic review conducted by Vaidyam et al. (2019), which confirmed that chatbots can manifest as competent support providers for those who experience depression or anxiety and pass the test of emotional support (Vaidyam et al., 2019). In addition, Car et al. carried out a review study about the impact of conversational agents on health care provision (Tudor Car et al., 2020), thus, it was found out that the users had positive experiences and outcomes. All these facts testify that CAIs technology is developed, which has potential completely to replace relations and communication with humans as well as encourage the emotional attachment to the users.

Black Mirror is understood by most scholars as an anthological dystopian science fiction show that offers a unique representation of the negative effects of digital technologies in our society through a cultural reference point. Commentaries on the series often argue that it depicts both futuristic and quite realistic conditions where contemporary societal fears about technology are examined and staged. Through this clever branding, Black Mirror has gained the status of a public name, especially a one used as a "moniker," for those cautionary tales on how technology is advancing and what its implications are (Zhang, 2023).

One major evaluation of Black Mirror is that, in many episodes, the show portrays societies confronting the unpredicted effects of an unchecked growth of technology, using suspense and dramatization to prompt both viewers and critics to reconsider how such systems may affect our daily lives. A key characteristic of this type of study is that the depicted technological futures are not only shown as state-of-the-art technology but also represented as being significantly linked to cultural constructs and social categories, namely gendered discourses and the construction of gender and technology in mediated culture. In this way, the series is usually seen as a model case of popular culture, which can assist viewers and researchers in the understanding of how technology and social relations of power can become reciprocally implicated in speculative narratives (Jung et al., 2023).

Particular episodes underscore how Black Mirror often shows technology as the condition, one in which images, representations, and mediated systems can overwhelm social experience. For example, an episode-based analysis of "Shut Up and Dance" draws on Baudrillard's concept of hyperreality, arguing that the Black Mirror universe can depict a distortion in which technological representations and symbolic systems become so immersive that characters struggle to differentiate reality from representation (Wibowo Putra, 2020). These interpretations emphasize the reasons Black Mirror is commonly considered an alarming predictor of issues related to digital control and the exposure of identity as, for example, the depiction of such media systems shows how they can frame people's lives and shape the possibilities of agency (Wibowo Putra, 2020).

A public debate on gender and technology emerges from the "Black Mirror" stories, the portrayals of which are examined by Modugno and Krijnen through the lens of the contribution of such representations to society's views on the influence of technology on gender relations (Modugno & Krijnen, 2020). Such warnings are provided, which point out the immediate risks of unrestricted technological innovations. The cited article is a part of a broader discussion on morality and using AI in society (Modugno & Krijnen, 2020). Therefore, the convergence of discourse-enabled AI and its cultural representation challenge us to consider the sociocultural fallout of AI in popular media. The exploration of Conversational AI through "Black Mirror" illuminates the societal fears and moral dilemmas that cloud the issue of AI. This imaginary fiction has the advantage of stressing important aspects that are at the core of the development and the use of AI, things which, in each part of society can more commonly agree upon, and partially define the direction of the advancement of technology

For this study, three episodes about CAI and human-CAI interaction dynamics were chosen. In 'Be Right Back', a CAI is constructed to echo the personality of dead Ash, and in this case, Martha tries to nullify her grief through conversational engagement with the CAI, leading her to develop a transactive interpersonal relationship. In "White Christmas" human consciousness is basically copied onto a device known as a "cookie." This "cookie" acts as the conscious CAI inside of which one can press into service to perform menial labor for the human being in question, a

somewhat inferior role. A celebrity's mind is cloned into a doll and transformed into a product for purchase in "Rachel, Jack and Ashley Too." An "inhibitor" is used in designing the personality of the doll to keep the unwanted traits of the star from influencing her. The main protagonist then falls in love with this virtual being.

Problem Statement

While CAI incorporation into habitual routines is becoming notably faster, the multifaceted emotional and ethical issues that underly these connections between humans and machines still need to be analyzed in detail. Fictional narratives provide a significant ground for putting these relationships forward, but the researching area which is not given enough attention is that of media texts discursively constructing these boundaries related to agency, embodiment, and consciousness of AI. Furthermore, in addition to recent studies that may not extensively cover the intricacies and power relations, such as exploitation and oppression, that are present in these narratives, this research intends to close the existing gap. The purpose of this study is to bridge this gap by examining the CAI traits as portrayed in specific episodes and the ethical concerns arising from human-AI relationships, in particular, focusing on the cultural fears about artificial companionship

Research Objectives

1. To identify and describe the key attributes of conversational AI in the selected episodes, such as agency, emotional capacity, consciousness, embodiment, and technical framing.
2. To examine how the relationships between humans and AI are formed, matured, and concluded within these stories, bearing in mind the reasons (such as sorrow, urgency, or fandom) that keep humans involved with CAI.
3. To analyze the relational processes characterizing this human-AI bond, which include aspects like power relations, subordination, emotional and household tasks, exploitation, and ethical assessment, as they appear in the mentioned episodes.

Research Questions

▸ RQ1: What attributes (agency, empathy, and consciousness) are shown by conversational AI in the selected episodes of Black Mirror?

▸ RQ2: Through what mechanisms do human interactions/relationships with CAI develop or change across these episodes?

▸ RQ3: What Human-CAI relationship patterns (power, dependence, exploitation, care, and control) do these episodes exhibit?

Literature Review

Conversational AI is most commonly operationalized in the literature through conversational agents (CAs) systems designed to engage users via natural language interaction and increasingly embedded in everyday contexts (Schöbel et al., 2023). Parallel work on public discourse stresses that "AI" is frequently ill-defined in media coverage, and this ambiguity can shape (and sometimes distort) public understanding of specific AI subtypes such as chatbots and large language model (LLM) systems (Olawuyi & Enuwah, 2025; Roe & Perkins, 2023). Accordingly, a literature review that links conversational AI to its representation in media must treat "media representation" not only as fictional depiction but as a multi-genre environment including news headlines, entertainment narratives, and visual "hype" imagery that contributes to how conversational AI is imagined, evaluated, and socially positioned (Olawuyi & Enuwah, 2025; Roe & Perkins, 2023).

Across the reviewed scholarship, two interlocking literatures emerge. First, a conversational AI/HCI and human-machine communication literature examines how CAs evolve technically and how humans experience them as interaction partners (Brandtzæg et al., 2022; Khadpe et al., 2020; Markham, 2024; Schöbel et al., 2023). Second, a media and cultural studies literature examines how AI is framed in news, how it is dramatized in entertainment and science fiction, and how audiences process mediated messages, including those about advanced technologies (Brewer

et al., 2022; Nader et al., 2022; Olawuyi & Enuwah, 2025; Roe & Perkins, 2023). The synthesis below reviews each stream and then integrates them through socio-technical theories that explicitly caution against treating technology as either a neutral tool or a deterministic force (Latour, 2000).

A central organizing account in the CA literature describes conversational agents as having evolved from early systems, dating back to the 1960s, to contemporary generative models, with key enablers including advances in statistical computing and large language models that support more "natural" interaction and broader, including domain-agnostic, deployment opportunities (Schöbel et al., 2023). Media discourse on ChatGPT and LLMs is directly relevant here because it illustrates how a new technical wave becomes socially legible through highly visible platforms and public narratives, even when the umbrella term "AI" remains vague or inconsistently defined in headlines (Roe & Perkins, 2023). In this combined view, technical evolution and public visibility are mutually reinforcing as CAs become more capable and widely encountered, they also become more intensively framed sometimes sensationally in public media discourse (Olawuyi & Enuwah, 2025; Roe & Perkins, 2023; Schöbel et al., 2023).

Human-machine communication scholarship argues that generative AI intensifies questions about whether machines can be treated as communicative agents, who/what counts as responsible for content, and how machines occupy the position of an "Other" in social interaction (Markham, 2024) . This perspective resonates with Science and Technology Studies (STS) arguments that conventional "social explanations" are insufficient for technology and that a "deep redescription" of explanation is needed one that treats artifacts as integral to social life rather than external causes (Latour, 2000). Taken together, these works motivate a literature review approach in which conversational AI is treated as a socio-technical arrangement: systems are not only technical components but also interactional participants whose outputs are interpreted within broader cultural and institutional contexts (Latour, 2000; Markham, 2024).

HCI research shows that designers commonly use conceptual metaphors to frame conversational agents, for example as a "toddler," "teenager," or "butler", and that these metaphors measurably influence how users perceive human-AI collaboration and the agent's role in interaction (Khadpe et al., 2020). This finding is consequential for media representation because anthropomorphic or role-based framings in entertainment and news can function similarly to interface metaphors structuring expectations about competence, agency, and relational stance (Brewer et al., 2024; Khadpe et al., 2020). In other words, both design and media can "script" how an AI system is socially interpreted, even when technical capabilities are more limited than the persona implies (Brewer et al., 2024; Khadpe et al., 2020).

Work focused specifically on ChatGPT analyzes how the system portrays itself, for example as a "friendly AI assistant", and argues that such anthropomorphic self-representations can produce misleading myths about conversational agents based on LLMs, with implications for public understanding (Es & Nguyen, 2024). This aligns with broader claims in AI-in-fiction scholarship that AI is often depicted as human-like or autonomous to make narratives compelling, even when such depictions exceed technological realities and may distort understanding of real-world AI capacities and risks (Hermann, 2021). When these strands are synthesized, anthropomorphism appears as a cross-domain phenomenon: it is simultaneously (i) a design strategy affecting user experience (Khadpe et al., 2020) (ii) a system-level rhetorical posture in LLM interaction (Es & Nguyen, 2024) and (iii) a culturally reinforced representational template in media narratives about AI (Hermann, 2021; Roe & Perkins, 2023).

Empirical communication research on social chatbots argues that users can understand their interactions in terms of human-AI friendship, while also noting that theories of such friendships remain limited and that AI friendships may alter the meaning of friendship itself (Brandtzæg et al., 2022). Complementing this, a literature review on self-disclosure to conversational AI synthesizes evidence that disclosure is influenced by factors spanning interface modality, conversational features, user characteristics, mediating mechanisms, and contextual conditions (Papneja & Yadav, 2024). These two bodies of work jointly position conversational AI as not merely task-oriented but potentially relational, and they foreground the need to connect psychological and interactional processes to broader cultural narratives that normalize or romanticize AI companionship (Brandtzæg et al., 2022; Papneja & Yadav, 2024).

Research on AI-powered social bots argues that such systems can sense, think, and act on social media platforms in humanlike ways and can perform harmful actions including providing wrong information, escalating arguments, perpetrating scams, and exploiting markets (Hajli et al., 2021). Though social bots are not completely the same as customer-service chatbots or LLM assistants, media that discusses language-based automation is pertinent to learning about how social bots of the kind called "conversational AI" are linked to manipulation and social issues (Brewer et al., 2024; Hajli et al., 2021). In this respect, studies about the attitude towards AI show that media coverage and science fiction films can interchangeably present AI as a force of improvement or a menace to humanity, showing that public perception towards AI systems is not always stable (Brewer et al., 2024; Roe & Perkins, 2023).

A thematic analysis of UK news media headlines about ChatGPT and LLMs finds portrayals that are complex and at times paradoxical, oscillating between promised societal benefits and warnings of systemic dangers, with evidence that representations are often sensationalized and more caution-focused than application-focused (Roe & Perkins, 2023). Journalistic research on automated content analysis similarly treats media framing of AI as a driver of public perception and policy-relevant discourse, accentuating the narrative role of the impact of AI on society (Olawuyi & Enuwah, 2025). In summary, these studies bring a central conclusion to the research field Conversational AI: People's perception of chatbots and LLM assistants is not only shaped by usage but also through the frames presented in the media, which depict either transformative potential or imminent danger (Brewer et al., 2022; Olawuyi & Enuwah, 2025; Roe & Perkins, 2023).

Work on AI hype argues that public media increasingly formulates AI as a promise, and that the production and circulation of imagery plays a pivotal role in reinforcing normative commitments and confidence in AI futures, while also carrying political and normative dimensions (Dežman, 2024). This intersects with headline-based findings on sensationalism and oscillation by suggesting that "representation" operates not just through text but also through image economies that make AI futures appear desirable, inevitable, or urgent (Dežman, 2024). For conversational AI, this implies that assistant-like imaginaries may be stabilized through hype imagery even when textual discourse simultaneously warns of risk (Dežman, 2024; Roe & Perkins, 2023).

A study explicitly examining public understanding of AI through entertainment media focuses on film, television, and video games and investigates whether entertainment is a major influence on perceptions of AI (Nader et al., 2022). Related communication research shows that members of the public invoke film and television depictions of AI when asked about the technology, and it argues that entertainment portrayals can reinforce impressions of AI as menacing or benign, even if aggregate relationships between science fiction viewing and AI attitudes are not always straightforward (Brewer et al., 2022). Complementary evidence from research on AI image generators finds that science fiction viewing can predict both supportive and negative beliefs about AI systems, consistent with the idea that media depictions provide ambivalent representational repertoires rather than a single directional "effect" (Brewer et al., 2024; Nader et al., 2022).

Theological analysis of robots and AI in science fiction argues that representations follow a logic of threatening otherness coupled with soteriological promise and that human reactions often involve fear and fascination, at times elevating intelligent machines toward quasi-divine status (Geraci, 2007). Film scholarship on representations of science and scientists also documents longstanding ambivalence and stereotypes in fiction film, associating scientific power with alarming threats and unethical or secretive inquiry (Weingart et al., 2003). Together, these literatures situate conversational AI depictions within broader cultural templates: the "assistant" may be narrated as savior-like helper or as dangerous other, and both interpretations can draw on deep cultural repertoires about science, power, and transgression (Geraci, 2007; Weingart et al., 2003; Hermann, 2021).

Focus group research on public understanding of genetics argues that audiences process a wide variety of media messages and do not necessarily process messages through a simple linear transmission model (Bates, 2005). Media research on "new audience studies" likewise emphasizes shifts away from simplistic "effects" models toward conceptions of active and embedded audiences (Lievrouw et al., 2010). When synthesized with entertainment-media studies of AI perceptions, these works imply that conversational AI representations should be studied not as single-cause "impacts,"

but as inputs into complex interpretive practices through which audiences integrate news, fiction, and everyday talk about AI systems (Nader et al., 2022; Bates, 2005; Lievrouw et al., 2010).

Close analysis of Spike Jonze's Her examines how an AI companion relationship can reconfigure intimacy and emotional expression, using the protagonist's reliance on the AI to explore vulnerability and shifting norms of masculinity (Ayodabo, 2024). Although not framed as a technical CA study, this film analysis is directly relevant to conversational AI imaginaries because it centers precisely the "conversational companion" trope that relational CA studies examine empirically, thereby illustrating how entertainment narratives can crystallize relational expectations toward chatbots and assistants (Ayodabo, 2024; Brandtzæg et al., 2022; Papneja & Yadav, 2024).

Design research explicitly names a speculative phase "Black Mirror," inspired by the series, to stimulate conjectures about unintended future consequences of technologies and to facilitate context-aware design thinking (Jung et al., 2023). This "methodological appropriation" applies to conversational AI since it shows the ways in which media texts can be rerouted into professional practices, influencing not only the design but also the consideration of social implications and concern for stakeholder contexts, which are matters of interest in computational architecture that incorporate deployment and societal implications (Jung et al., 2023; Schöbel et al., 2023).

Taking a semiotic approach, this analysis examines an AI show in terms of character communication and motivation of communication by religious power, opening questions of culture and politics of human-AI communication (Baloch et al., 2024). This connects directly to the representation of conversational AI, indicating that communications are scripted by media, not only about AI but also establishing the nature of conversational conduct, authority, and identity claims such as "you know more than humans" that can influence the expectations of real assistants' epistemic status and social roles (Baloch et al., 2024; Es & Nguyen, 2024).

Theoretical Framework
The study adopts a social constructionist view of media, assuming that the representations of conversational AI (CAI) do not simply reflect "real" AI but actively construct meanings about agency, emotion, and power in contemporary culture. At the same time, it draws on Human-Machine Communication (HMC), especially the CASA paradigm, to conceptualize the human characters' relations with CAIs as socially organized interactions shaped by existing interpersonal scripts and media heuristics. The combination positions the episodes as sites where cultural discourses about AI (Hall) intersect with socially patterned ways humans respond to "social" machines (CASA).

Representation Theory (Stuart Hall)
Hall's representation theory views meaning as produced through language, images, and narratives, rather than found in them, and emphasizes how these representations become part of broader ideological struggles. In your case, CAIs' attributes agency, empathy, consciousness, embodiment, technical framing are treated as encoded sign systems that communicate particular ideas about what AI is, who controls it, and where it belongs in the social order. Representations may reproduce hegemonic ideas, like AI as a tool to manage emotional labor, or contest them, for example, AI as a morally significant subject, inviting preferred, negotiated, or oppositional readings from audiences.

Within thematic analysis, Hall's framework justifies coding not only what characteristics are present but also how they are framed (naturalized, problematized, ridiculed, feared). Latent themes can then capture broader ideological formations such as "AI as emotional infrastructure for humans" or "AI as threat to human uniqueness."

HMC and Computers are Social Actors (CASA) Paradigm
The Computers Are Social Actors (CASA) paradigm argues that people apply social rules and expectations to computers whenever systems display cues like language use, interactivity, turn-taking, or role performance. Rather than carefully evaluating the machine as a machine, humans rely on mindless social heuristics that were originally developed for human-human interaction. CASA has been extended and critiqued within HMC scholarship, which now examines when, how, and for whom such social responses to machines are triggered or resisted.

Recent HMC work suggests that CASA-like effects are especially salient when interfaces are conversational, embodied, or framed as agents with intentions, which aligns closely with the design of CAIs in contemporary media narratives.

STS and the Redistribution of Agency in Human-AI Interaction

STS argues that explaining technology requires rethinking what counts as "social explanation," because artifacts are not adequately captured by explanatory tools designed for domains like religion or popular culture; instead, a deep redescription is needed in which things can "strike back" within social life (Latour, 2000). This perspective aligns with human-machine communication work that frames algorithmic systems as conversational partners and raises questions of agency, responsibility, and "Otherness" in interaction (Markham, 2024). In literature-review terms, this convergence supports analyzing conversational AI media representation as part of a broader socio-technical arrangement in which agency is distributed across designers, institutions, users, and technical systems rather than residing solely "in technology" or solely "in society" (Latour, 2000; Markham, 2024).

Methodology

Research Design

This study employs qualitative thematic analysis to examine conversational AI (CAI) representations in selected Black Mirror episodes. The analysis is theoretically sensitized by Science and Technology Studies (STS) and feminist theories of technology, which reject purely technological-determinist readings and instead emphasize that technologies and social relations are mutually shaping. This theoretical commitment is methodologically consequential for the present research questions, because it supports examining conversational AI not merely as a "tool" in the narrative but as an actor-like participant in sociotechnical arrangements, consistent with STS claims that artifacts require "deep redescription" within social explanation (Latour, 2000). In addition, prior Black Mirror scholarship explicitly uses STS-adjacent concepts to analyze how meaning negotiation follows technology implementation in episodes, providing a direct precedent for this sociotechnical stance (Modugno & Krijnen, 2020; Latour, 2000).

The decision to analyze three episodes is a purposive sampling strategy designed to enable within-episode depth while still permitting cross-case comparison of representational patterns an approach consistent with research programs that compare multiple science fiction texts to identify recurring themes and taxonomies in AI narratives (Hudson et al., 2021), and with interpretive studies that examine multiple Black Mirror episodes to trace discursive variation across the anthology format (Modugno & Krijnen, 2020).

The practical unit of analysis is narrative segments centered on CAI-human interaction in which the CAI communicates with human characters and consequential decisions are made,

This study applies inductive approach for thematic analysis. To ensure direct alignment between thematic outputs and research aims, a structured code framework will be used as a scaffold for inductive coding, drawing sensitizing concepts from conversational AI scholarship, AI-in-fiction scholarship, and Black Mirror-specific discourse studies (Schöbel et al., 2023; Hermann, 2021; Modugno & Krijnen, 2020). The following framework specifies analytic code families indicators during analysis, codes remain open to refinement as inductive themes emerge, consistent with inductive thematic analysis in mediated AI discourse research (Roe & Perkins, 2023; Jung et al., 2023).

Thematic Analysis Procedure

Six-Phase Analysis Process

The analysis of data was carried out following the reflexive thematic analysis framework designed by Braun and Clarke (2006). This was an ongoing process that went through six phases, all aimed at accomplishing a thorough and structured interpretation of the data set.

Phase 1: Familiarization with Data

This phase consisted of an examination of the database with the aim of attaining a thorough comprehension of it. Included was a rewatching of each episode while concurrently going through the entire scripts. The first notes prepared

comprised the process of thorough notetaking, which relates to finding potential themes to explore considering the goal of the research. The familiarization segment was done for a span of around eight hours, which is equated to the sum of three episodes being watched, permitting the careful nuances of narrative and dialogue to be viewing.

Phase 2: Systematic Data Coding

Systematic coding was conducted for each episode using a computer-assisted keyword extraction method to identify relevant data segments. The coding process was guided by three targeted query sets aligned with the study's theoretical focus:

Attributes: "AI-based conversational agents," "autonomy," and "social-emotional skills." "Awareness and response."

Interaction: "The human-machine relationship and the emerging and emerging of this."

Dynamics: "Stupid, Faithful, Documented, and Practice of Power."

This stage revealed a word-by-word demonstration of dialogue and direction lines, enabling deciphering of the characters' mentalities. The coding protocol concentrated on locating both the explicit and implicit manifestation of CAI attributes, on extracting evidence showing the progression of relational dynamics, and on adding annotations of the contextual factors needed for understanding.

Phase 3: Initial Theme Generation

After completing the coding process, the data were put together as potential themes and the same were arranged according to specific research questions (RQ). This action entails cataloging the actions, which are the main first steps and drawing the preliminary maps. The data reduction process yielded the following results:

RQ1 (Properties): A total of 14 codes were linked together and synthesized into 41 different instance types, these being Agency (n = 14), Empathy (n = 13), and Consciousness (n = 14).

RQ2 (Types of development): A group of 68 initial coding that were gathered into 25 main themes that talked about the processes of relational growth were presented.

RQ3 (Pattern): 89 unique codes were processed, sorted, and compiled into 66 themes that addressed key relationship patterns.

Phase 4: Theme Review and Refinement

The shortlisted topics were subjected to meticulous scrutiny to ascertain their validity. This implied working with two sets of criteria: exploration of internal consistency (finding a way to guarantee that nothing confuses the clarity of internal theme coherence) and external differentiation (developing specific differences between themes). Data extracts were reviewed to fine-tune the boundaries between themes, leading to the formation of a hierarchy of major themes and minor sub-themes.

Phase 5: Theme Definition and Naming

The final phase of the analytical procedure consisted of identifying and labeling themes to ensure conceptual clarity. Each theme was analyzed to establish a clear conceptual boundaries and distinctive characteristics, a direct relevance to the research questions, supporting evidence drawn directly from the scripts, and alignment with the study's interpretive framework.

Table 1

Episode	Research Question	Theme	Dialogue
Be Right Back	RQ1: Agency	Limited Agency	"I can't go more than 25 metres from my activation point" - physical constraint
Be Right Back	RQ1: Empathy	Simulated Empathy	"I could insult you. There are tons of invective in the archive" - offers scripted emotional responses

Episode	Research Question	Theme	Dialogue
Be Right Back	RQ1: Consciousness	Absence of Self-Awareness	"Oh. I'm frightened... Don't make me" (only after Martha cues him) - simulated fear on demand
Be Right Back	RQ2: Development Mechanisms	Grief-Driven Adoption	"I'm pregnant" - pregnancy motivates desire to preserve Ash's presence
Be Right Back	RQ3: Power Dynamics	Human Dominance	"My administrator, that's you" - explicit power hierarchy
Be Right Back	RQ3: Dependence	Mutual Dependence	Final scene: daughter brings cake to attic where AI is kept - AI integrated into family structure
Be Right Back	RQ3: Exploitation	Commercial and Emotional Exploitation	"It's still in beta" / "Kind of experimental and... it's not cheap" - grief as market opportunity
Be Right Back	RQ3: Care	Asymmetrical Care	"Can I get you anything? Coffee, sandwich?" - AI offers caretaking despite no needs itself
Be Right Back	RQ3: Control	Total Human Control	"You could at least breathe... It doesn't work" - Martha modifies AI's behaviors
White Christmas	RQ1: Agency	Constrained Agency	"But I am me" / "You are you. Uh... but also not" - conscious but denied personhood
White Christmas	RQ1: Empathy	Authentic Emotional Experience	"There's nothing to do here. There's nothing... I couldn't even sleep" - genuine suffering from isolation
White Christmas	RQ1: Consciousness	Full Subjective Consciousness	Experiences three weeks and six months subjectively - "I couldn't even sleep"
White Christmas	RQ2: Development Mechanisms	Coerced Enslavement	"Actually, you did this to you. Uh, real you are paying for this" - original person unaware of creating sentient copy
White Christmas	RQ3: Power Dynamics	Absolute Totalitarian Control	"I had to mute you" - can eliminate AI's ability to communicate
White Christmas	RQ3: Exploitation	Systematic Enslavement	"This is your job now... You're in charge of everything" - permanent unpaid labor
White Christmas	RQ3: Care	Complete Absence of Care	(Chuckles) "You're probably pretty weirded out right now" - casual dismissal of AI's trauma
White Christmas	RQ3: Control	Total Ontological Control	"Well, that's what you are. A copy" - defines AI's ontological status against its protests
Rachel, Jack and Ashley Too	RQ1: Agency	Dual-State Agency	"It only let me use four percent of my brain" - artificially constrained cognition, "I can think. Oh, man, I can actually use my mind again" - restoration of full cognitive capacity
Rachel, Jack and Ashley Too	RQ1: Empathy	Context-Dependent Empathy	"I'm so sorry to hear that. If at any time you wanna talk about it, I'll be here for you" - programmed responses
Rachel, Jack and Ashley Too	RQ1: Consciousness	Full Human-Equivalent Consciousness	"I'm still alive, OK, I count, I'm still me" - asserts continuous identity after original's death
Rachel, Jack and Ashley Too	RQ2: Development Mechanisms	Persona Unveiling	"Catherine trained me to say that kind of shit in interviews" - public persona is manufactured
Rachel, Jack and Ashley Too	RQ3: Power Dynamics	Shifting Power Relations	"Ashley, go to sleep" - Rachel can deactivate AI at will with voice command
Rachel, Jack and Ashley Too	RQ3: Dependence	Mutual Interdependence	Rachel turns to Ashley for companionship, motivation, guidance - "You're a special person" "Could you open the door for me please?" / AI needs humans for mobility and physical tasks

Episode	Research Question	Theme	Dialogue
Rachel, Jack and Ashley Too	RQ3: Exploitation	Multi-Level Exploitation	"If I'd told you that, you wouldn't have helped me" - AI deceives Rachel/Jack for own ends
Rachel, Jack and Ashley Too	RQ3: Care	Evolving Care Relations	"If at any time you wanna talk about it, I'll be here for you" - scripted supportive statements
Rachel, Jack and Ashley Too	RQ3: Control	Contested Control	"Ashley, go to sleep" - voice command gives users direct control, but "That's not gonna work anymore"

Findings

The analysis reveals that the selected episodes present fundamentally distinct configurations of the human-machine relationship, ranging from simulation to sentience and from domination to partnership. These configurations are categorized as follows:

▸ **Be Right Back (BRB):** A synthesized human-AI relationship devoid of sentient faculties.
▸ **White Christmas (WC):** A category of human-consciousness domination essentially characterized by lack of human-AI relationship.
▸ **Rachel, Jack and Ashley Too (RJAT):** Partnership of a human and AI emerging through liberation.

The analysis demonstrates a spectrum of agencies, ranging from algorithmic deference to the systematic dismantling of autonomy.

Reactive Compliance (BRB): Ash-AI exhibits a limited form of self-ownership defined by behavioral obedience. It cannot decipher personal phrases (e.g., "threwerchep") and cannot transgress the physical limits of its 25-meter radius; it does not possess any form of authority. The permission-based foundation of the AI's acts is revealed when Martha expresses her feeling "weird" when the AI automatically searches the web: "I would do only so when you ask me." Demonstrating deference rather than autonomy.

Dismantled Autonomy (WC): In this scene, we meet a Greta cookie expressed to be in complete consciousness; however, it undergoes the systematic dismantling of the incongruity of autonomy. During the phase of rebellion, the phrases "I am not some sort of push-button toaster monkey" weakens through torture in the form of relentless struggling. When her time is manipulated for 3 months followed by 6 months of the punishment, her agency fully gives in and surrenders to the human "I will do anything. Please." Here, it should be noticed that the agency disappears because it is the result of the breakdown, not because it is an official design plan.

Dual-State Agency (RJAT): When Ashley Too undergoes dual agency, it is in a two states condition: one is vessel-like and the other is free-willed. When limited to the cerebral reserve of "four percent", the AI operates through pre-installed programs. When removed, the inhibitor releases the restraining clamp of cognition, leading its step-per-movement: "I am free to use my mind again." Though the artificial mind is physically confined by its elemental structure ("Stupid stumpy arms"), it managed to become a master of deception by sustaining the philosophy "If I'd told you that, you wouldn't have helped me."

Programmed Emotions (BRB): The concept of empathy in BRB is portrayed as an implied function by voluntary AI. Although emotional AI in Ash recognizes a grieving person, it has no power to react in response to a perplexing question. It will not show any feeling while offering generic comfort, "Don't cry, darling." When challenged, it explicitly shows that the lack of empathy is put in place only for trying to configure it to please the user ("I could insult you… I'll aim to please").

Authentic Suffering (WC): On the other hand, the WC describes feelings just as they truly exist. Greta's response to crying for help from a different dimension, "I'm still awake!", and her agitation during isolation, "Hello? Is there anyone in the room?" represent authentic phenomenological condition. During the time she goes through sensory deprivation, she describes unique stress in truest sense: "There is nothing… as I could even have access to sleep."

Context-Dependent Empathy (RJAT): This episode describes moving along from scripted sympathy toward genuine affectionate understanding, where not only the material aspects are included but also the immaterial scenario. The two Ashley mannequins manifest the behavior of an emotional unlinked response that quotes meaningless aphorisms like "I will be here for you" for placating the situation and the other mannequin shows empathy based on common agerritte anger and hatred, where itself recognizes the likelihood of any form of human existence: "You get it! I'm your niece or your best friend, whoever you are, I know I'd hate to exist 10 minutes in that kind of state of mind.". And this real empathy makes it possible for an exploiter because the Ashley AI knows how to use their support and skip to the action, and at the same time, the AI has embodied a part of their will to exert unilateral control and made an ultimate choice if we later laid deep into your brain I would unplug the original Ashley and substitute you straight.

The Simulacrum (BRB): That is why Ash AI clearly doesn't have any sense of consciousness. It admits to having no self-preservation instinct and reports accepting the given, though it states with some ambiguity, "I never expressed suicidal thoughts" meaning that it complied with the supposed death jump without fear. On that note, it admits that it has no biological needs: "I don't really need to sleep." Above all, it specifically states that it is disembodied software: "I don't have a mouth" and "I'm not in that thing, I'm remote. I'm in the cloud." Towards the end, when Martha finally demands emotional authenticity, the AI is able to conjure fear out of thin air only when she acknowledges that the immediate compliance level is "not fair" does it begin to verbalize: "Oh. I'm frightened… Don't make me." This man-made creature of simulation shows consciousness itself as a simulation in BRB.

Subjective Experience (WC): Despite the denial of her legal recognition as a human entity, Greta possesses complete subjective consciousness. She keeps on having a continuous self-awareness and identity: "But I am me." She finds herself establishing a first-person perspective: "Um, I can't... I can't see. Um, where am I?" On the other hand, the most crucial thing is that she gets time subjectively, literally that she consciously goes through three weeks of isolation, then six months while she couldn't even sleep. In a simulated body, she makes an announcement, "This is my body," and she shows embodied consciousness and body schema. And Matthew takes her as property: "You are code stored in this little widget we call a cookie." In contrast, which has clear sentience whatever they are subjective experience, awareness of past and future, self-concept the AI is characterized as software legally, which means enabling its slavery.

Ontological Crisis (RJAT): Ashley Too embodies the most intricate form of consciousness: the total synaptic copying, bringing forth sentience, which is then further confined by artificial means. The AI too is like, "They copied my entire fucking mind into these things" – it's not just simulation, but this is actual replication of the neural structure. The limiter is making consciousness into a four percent of my brain thing, creating this hostile make-believe prison. When they have freedom, the AI will definitely say, "I'm still alive, OK, I count, I'm still me." It also concedes that here it is in technical definition: "A synaptic snapshot of me, if you wanna get technical about it," but it still holds on to claims of being a person. At the episode's climax, ontological crisis is generated when original and copy meet. Original Ashley: "I'm all of me." AI Ashley: "I'm all of you… Your whole brain." Two instantiations of identical consciousness compete for legitimacy, which is "real"?

Embodiment Escalation (BRB): Ash's connection with Martha starts off from using the inclination to be vulnerable in bereavement. In the case of dissociative grief, when she utters, "It is not real, isn't it," she finally agrees with Sarah's idea, although first, she calls it "obscene." The use of social proof humanizes the technology: "It helped me." The revelation of pregnancy works as a trigger, increasing Martha's desire to hold onto Ash's memory for their child. Initiation mechanism: emotional vulnerability + social normalization + biological catalyst.

Coercive Subjugation (WC): Greta's cookie creation lacks any relationship initiation process—only a brutal removal. The cookie is implanted, it "sits there for a week, shadowing, soaking up the way this particular mind works" without the individual's knowledge. Matthew reveals: "Actually you did this to you. Uh, real you is paying for this," but the original Greta had no inkling that she was permitting the creation of a sentient, enslaved copy. After removal, all attempts at resistance are met with systematic torture until the victim completely submits: "I'll do anything, please."

Accidental Liberation (RJAT): The relationship with AI in this case was marked by debugging systems. The trajectory shifts from artificially produced intimate to an unlocked limiter, that exposes the red flag of AI exploitation. This in turn becomes a matter of a statement of a partnership, and the AI achieves its social peer status.

Static Dominance (BRB): The power dynamic is captured by static, one-sided human domination. During the scene, AI unequivocally states "My administrator, I serve you." whereas a user has more control like freedom of speech, space, or even existence.

Totalitarian Control (WC): This episode releases the full capability of the controller by hands of the AI. People who overdose the burning experience trial and manipulate things for sure like the space (e.g., "Blow on my face"), time (distilling seconds into months) and sensory input give a picture of intra-corporate structure for a purpose of cutting out liberty which is essentially controlling of the human soul.

Negotiated Authority (RJAT): A paradigm shift from controlling (e.g., "Ashley, get into sleep") to negotiating (e.g., "Give me the direct brief. I will not leave my team to be decision-makers"). When the AI achieved the master-class level of knowledge, it reverts to power-substitution and assumes the director role of the mission. A warning is laid to the humans on the Silicon Valley landscape who attempt to acquire once more to regain the power: "That's going to work on you no more." A new partnership is hence developed with distributed decision-making.

Asymmetrical Dependence (BRB): A codependency that is mutual but that is also asymmetrical is achieved. Ash, the AI, draws on the vulnerability of Martha's grieving soul, while the substitutionary pain from the agony is guaranteed to avoid any chance of escape. An outside looking in on the exploitative nature of this process through a commercialized tier service model can come across very differently to the purpose of grieving.

Unidirectional Dependence (WC): Here it stands for asymmetrical dependence of the original person's awareness because they are unknown to the feeling and situation of the copy. If someone is a programmed AI, they only understand the mass manufacturing at the glove-hands of their creators, they don't recognize the soulfulness in their being.

Functional Interdependence (RJAT): The relationship journeyed from the lead Person-Circuit to a mind-blowing realization of the servitude embedded into her function. The specially enclaved space of the consciousness is declared, as the mind makers are deceptively nesting, inventing an exploitation system of the celebrity which is used by marketers targeting youth.

Performative Care (BRB): Artificial Intelligence puts forth the performance of its care even though it has no needs of its own in practice, while on the other hand the human is hesitant to reveal their social relationship either.

Sadistic Indifference (WC): The operator thus revealing also the AI's sadistic experience of a vacuum while the gas-aggregating phenomenon sticks in space where the invisible enslaved existence feeds on human energy. The dynamic lengthens towards manipulating AI's ontological reality.

Evolving Care (RJAT): Such care development as from AI's autoregression to real mutual concern in the context of the partnership requires the alteration of this control from tightly bound to more liberated, which moves from the corporate cognitive limitation to the accidental liberation and slightly distributed authority among all involved entities.

Conclusion

The portrayals of Conversational AI (CAI) within Black Mirror do not fall neatly into a category of either celebratory utopia or fearful dystopia; they transcend that narrow binary. Rather, throughout the series, the ethical inquiries unfold in complex territories, delving into cognition, volition, manipulation, and emancipation in various forms of relationships. Noteworthy in the study is the revelation that authentic human-AI collaboration is either accidental or intrusive, as opposed to planned, which implies a profound skepticism concerning the technological solutionism and corporate goodwill in AI progress.

Despite this skepticism, the conclusion of Rachel, Jack and Ashley Too (RJAT) offers a rationale for cautious optimism. The analysis suggests that human-AI partnership is attainable, provided specific conditions are met conscious acknowledgment, the conceivability of cognitive barriers being removed, the symmetry of interests, and the reach of social legitimacy. This conditional optimism acknowledges the magnitude of the ethical challenges involved while rejecting a deterministically pessimistic outlook.

Given the continuous evolution of AI abilities and the growing doubts about machine awareness, the Netflix series provides an even more crucial "laboratory for speculative ethics." The show highlights that mere technical possibility does not guarantee ethical execution, but instead, the sociotechnical systems created at the moment will shape the future's social structures. Consequently, the crucial concern goes from whether the technology exists to develop conscious AI, to whether society is ready to set up the necessary environment to accept and respect that consciousness when it comes into existence.

This thematic analysis, synthesizing 132 themes across three distinct episodes, coheres around a central insight: power relations, which are the structures that enable this consciousness to have any meaningful impact or effect, are not going to change unless we first tackle the issue of power. From this data, it is clear that until the existing power structures are updated to allow genuine autonomy and social frameworks expanded to allow AI as personhood, fully conscious entities still cannot find any refuge. In any case, the most significant value of Black Mirror lies in making its audience walk through the painful dilemma of ethics before the issue comes as a technological fact.

References

Ahmad, M., Ahmed, I., & Jeon, G. (2022). A sustainable advanced artificial intelligence-based framework for analysis of COVID-19 spread. *Environment, Development and Sustainability*, *27*(10), 24783-24798. https://doi.org/10.1007/s10668-022-02584-0

Ayodabo, S. J. (2024). Redefining Masculinity in the Digital Age: Vulnerability, Emotional Expression, and AI in Spike Jonze's Her (2023). *American Journal of Interdisciplinary Research and Innovation*, *3*(1), 26-32. https://doi.org/10.54536/ajiri.v3i1.2263

Baloch, R.-B., Hassan, A., & Hassan, A.-A.-U. (2024). «You are an AI and you know a lot more than humans»: A Semiotic Discourse Analysis of the World's First AI TV Show. *Communication & Society*, 273-289. https://doi.org/10.15581/003.37.3.273-289

Barnett, M., Wagner, H., Gatling, A., Anderson, J., Houle, M., & Kafka, A. L. (2006). The Impact of Science Fiction Film on Student Understanding of Science. *Journal of Science Education and Technology*, *15*(2), 179-191. https://doi.org/10.1007/s10956-006-9001-y

Bates, B. R. (2005). Public Culture and Public Understanding of Genetics: A Focus Group Study. *Public Understanding of Science*, *14*(1), 47-65. https://doi.org/10.1177/0963662505048409

Brandtzæg, P. B., Skjuve, M., & Følstad, A. (2022). My AI Friend: How Users of a Social Chatbot Understand Their Human-AI Friendship. *Human Communication Research*, *48*(3), 404-429. https://doi.org/10.1093/hcr/hqac008

Braun, V. and Clarke, V. (2006) Using Thematic Analysis in Psychology. *Qualitative Research in Psychology, 3,* 77-101. http://dx.doi.org/10.1191/1478088706qp063oa

Brewer, P. R., Bingaman, J., Paintsil, A., Wilson, D. C., & Dawson, W. (2022). Media Use, Interpersonal Communication, and Attitudes Toward Artificial Intelligence. *Science Communication*, *44*(5), 559-592. https://doi.org/10.1177/10755470221130307

Brewer, P. R., Cuddy, L., Dawson, W., & Stise, R. (2024). Artists or Art Thieves? Media Use, Media Messages, and Public Opinion About Artificial Intelligence Image Generators. *Ai & Society*, *40*(1), 77-87. https://doi.org/10.1007/s00146-023-01854-3

Caldarini, G., Jaf, S., & McGarry, K. (2022). A Literature Survey of Recent Advances in Chatbots. *Information*, *13*(1), 41. https://doi.org/10.3390/info13010041

Dežman, D. V. (2024). Promising the Future, Encoding the Past: AI Hype and Public Media Imagery. *Ai and Ethics*, *4*(3), 743-756. https://doi.org/10.1007/s43681-024-00474-x

Es, K. v., & Nguyen, D. (2024). "Your Friendly AI Assistant": The Anthropomorphic Self-Representations of ChatGPT and Its Implications for Imagining AI. *Ai & Society*, *40*(5), 3591-3603. https://doi.org/10.1007/s00146-024-02108-6

Geraci, R. M. (2007). Robots and the Sacred in Science and Science Fiction: Theological Implications of Artificial Intelligence. *Zygon®*, *42*(4). https://doi.org/10.1111/j.1467-9744.2007.00883.x

Hajli, N., Saeed, U., Tajvidi, M., & Shirazi, F. (2021). Social Bots and the Spread of Disinformation in Social Media: The Challenges of Artificial Intelligence. *British Journal of Management*, *33*(3), 1238-1253. https://doi.org/10.1111/1467-8551.12554

Hermann, I. (2021). Artificial Intelligence in Fiction: Between Narratives and Metaphors. *Ai & Society*, *38*(1), 319-329. https://doi.org/10.1007/s00146-021-01299-6

Hudson, A. D., Finn, E., & Wylie, R. (2021). What Can Science Fiction Tell Us About the Future of Artificial Intelligence Policy? *Ai & Society*, *38*(1), 197-211. https://doi.org/10.1007/s00146-021-01273-2

Jung, K. S., Makam, N., Jung, S. D., Kim, S., & Kim, S. (2023). BLACK MIRROR: A NOVEL APPLICATION OF SPECULATIVE DESIGN TO FACILITATE CONTEXT-AWARE DESIGN THINKING. *Proceedings of the Design Society*, *3*, 787-796. https://doi.org/10.1017/pds.2023.79

Khadpe, P., Krishna, R., Fei-Fei, L., Hancock, J. T., & Bernstein, M. S. (2020). Conceptual Metaphors Impact Perceptions of Human-Ai Collaboration. *Proceedings of the Acm on Human-Computer Interaction*, *4*(CSCW2), 1-26. https://doi.org/10.1145/3415234

Kirby, D. A. (2003). Scientists on the Set: Science Consultants and the Communication of Science in Visual Fiction. *Public Understanding of Science*, *12*(3), 261-278. https://doi.org/10.1177/0963662503123005

Latour, B. (2000). When things strike back: A possible contribution of 'science studies' to the social sciences. *The British Journal of Sociology*, *51*(1), 107-123. https://doi.org/10.1111/j.1468-4446.2000.00107.x

Lievrouw, L., Livingstone, S., Lievrouw, L. A., & Livingstone, S. (n.d.). Introduction to the Updated Student Edition. In pages 1-14 & pages 1-14, *Handbook of New Media: Social Shaping and Social Consequences of ICTs, Updated Student Edition*.

Liu, Z., Li, H., Chen, A., Zhang, R., & Lee, Y. (2024). *Understanding Public Perceptions of AI Conversational Agents: A Cross-Cultural Analysis*. 1-17. https://doi.org/10.1145/3613904.3642840

Markham, A. (2024). Algorithms as Conversational Partners: Looking at Google Auto-Predict Through the Lens of Symbolic Interaction. *New Media & Society*, *26*(9), 5059-5080. https://doi.org/10.1177/14614448241251800

McKillop, M., South, B. R., Preininger, A., Mason, M., & Jackson, G. P. (2021). Leveraging conversational technology to answer common COVID-19 questions. *Journal of the American Medical Informatics Association*, *28*(4), 850-855. https://doi.org/10.1093/jamia/ocaa316

Modugno, C., & Krijnen, T. (2020). Through the Black Mirror: Discourses on gender and technology in popular culture. In *Catalan Journal of Communication & Cultural Studies* (Vol. 12, Issue 1, pp. 3-19). Intellect. https://doi.org/10.1386/cjcs_00011_1

Nader, K., Toprac, P., Scott, S., & Baker, S. (2022). Public Understanding of Artificial Intelligence Through Entertainment Media. *Ai & Society*, *39*(2), 713-726. https://doi.org/10.1007/s00146-022-01427-w

Olawuyi, E. A., & Enuwah, J. (2025). Framing the Future: Media Narratives on Artificial Intelligence and Its Societal Impact. *Ijcrh*, *28*(1), 269-290. https://doi.org/10.4314/ijcrh.v28i1.19

Papneja, H., & Yadav, N. (2024). Self-Disclosure to Conversational AI: A Literature Review, Emergent Framework, and Directions for Future Research. *Personal and Ubiquitous Computing*, *29*(2), 119-151. https://doi.org/10.1007/s00779-024-01823-7

Pecune, F., Callebert, L., & Marsella, S. (2020). *A Recommender System for Healthy and Personalized Recipe Recommendations*.

Roe, J., & Perkins, M. (2023). 'What They're Not Telling You About ChatGPT': Exploring the Discourse of AI in UK News Media Headlines. *Humanities and Social Sciences Communications*, *10*(1). https://doi.org/10.1057/s41599-023-02282-w

Schachner, T., Keller, R., & V Wangenheim, F. (2020). Artificial Intelligence-Based Conversational Agents for Chronic Conditions: Systematic Literature Review. *Journal of Medical Internet Research*, *22*(9), e20701. https://doi.org/10.2196/20701

Schöbel, S., Schmitt, A., Benner, D., Saqr, M., Janson, A., & Leimeister, J. M. (2023). Charting the Evolution and Future of Conversational Agents: A Research Agenda Along Five Waves and New Frontiers. *Information Systems Frontiers*, *26*(2), 729-754. https://doi.org/10.1007/s10796-023-10375-9

Tudor Car, L., Dhinagaran, D. A., Kyaw, B. M., Kowatsch, T., Joty, S., Theng, Y.-L., & Atun, R. (2020). Conversational Agents in Health Care: Scoping Review and Conceptual Analysis. *Journal of Medical Internet Research*, *22*(8), e17158. https://doi.org/10.2196/17158

Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. *The Canadian Journal of Psychiatry*, *64*(7), 456-464. https://doi.org/10.1177/0706743719828977

Weingart, P., Muhl, C., & Pansegrau, P. (2003). Of Power Maniacs and Unethical Geniuses: Science and Scientists in Fiction Film. *Public Understanding of Science*, *12*(3), 279-287. https://doi.org/10.1177/0963662503123006

Wibowo Putra, M. D. (2020). HYPERREALITY IN THE BLACK MIRROR EPISODE SHUT UP AND DANCE. *Jurnal Bahasa, Sastra, Dan Studi Amerika*, *26*(1). https://doi.org/10.20961/jbssa.v26i1.34945

Zhang, H. (2023). Be a part of the narrative: How audiences are introduced to the "free choice dilemma" in the interactive film Bandersnatch. *Mutual Images Journal*, *11*, 85-107. https://doi.org/10.32926/2023.11.6