Online Journal of Communication and Media Technologies, 2023, **13**(4), e202340
e-ISSN: 1986-3497

OPEN ACCESS

Review Article

AIsmosis and the pas de deux of human-AI interaction: Exploring the communicative dance between society and artificial intelligence

Ayse Asli Bozdag [1,2*]
0000-0003-0647-8120

[1] Faculty of Business, Istanbul Bilgi University, Istanbul, TÜRKİYE
[2] Faculty of Communication, Bahcesehir University, Istanbul, TÜRKİYE
[*] Corresponding author: asli.bozdag@bilgi.edu.tr

Citation: Bozdag, A. A. (2023). AIsmosis and the pas de deux of human-AI interaction: Exploring the communicative dance between society and artificial intelligence. *Online Journal of Communication and Media Technologies, 13*(4), e202340. https://doi.org/10.30935/ojcmt/13414

ARTICLE INFO

Received: 27 Apr 2023

Accepted: 9 Jun 2023

ABSTRACT

As the global influence of artificial intelligence (AI) in our daily lives and the looming advent of artificial general intelligence (AGI) become increasingly apparent, the need for a sophisticated interpretive framework intensifies. This paper introduces 'AIsmosis'–a term that captures AI's gradual, nuanced integration into society, and akin to the biological process of osmosis. AI's integration dynamics are examined through the lens of three pivotal theories: social construction of technology, technological determinism, and diffusion of innovations. These theories collectively elucidate the sociocultural influences on AI, the potential repercussions of unchecked technological growth, and the factors driving the adoption of novel technologies. Building upon these explorations, the 'controlled AIsmosis' conceptual framework emerges, emphasizing ethically conscious development, active stakeholder communication, and democratic dialogue in the context of AI technology adoption. Rooted in communicative action theory, this framework illuminates AI's transformative impact on society. It calls for a comprehensive evaluation of systems that steer AI diffusion and their potential impacts, acknowledging the pervasive influence of AI and transcending traditional disciplinary boundaries. This work underscores the need for a multidisciplinary and interdisciplinary approach in investigating the complex AI-society interplay and understanding the ethical and societal consequences of AIsmosis.

Keywords: AIsmosis, controlled AIsmosis, artificial intelligence, social construction of technology, technological determinism, diffusion of innovations, communicative action

INTRODUCTION

> "The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it" (Mark Weiser, computer scientist and pioneer of ubiquitous computing).

Navigating through the swift currents of our evolving landscape, we encounter transformative technologies such as smartphones, the internet, GPS, artificial intelligence (AI) voice assistants, Wi-Fi, contactless payment systems, social media platforms, and streaming services, along with many more. These technologies, echoing computer scientist Mark Weiser's observation, intricately weave themselves into our life's fabric. Their pervasive influence sculpts our experiences and societal norms to the point of near indistinguishability from our daily routines.

This article aims to untangle the complex interplay between AI and society while focusing on the processes guiding the integration of technologies into societal, cultural, and personal frameworks. To deepen our understanding, the subsequent chapters of this article will delve into the intricate interplay between AI and society, exploring key theories such as social construction of technology (SCOT) by Pinch and Bijker (1984), technological determinism as discussed by Veblen (1899, 1921), and the diffusion of innovations as proposed by Rogers (1962). Each theory presents a unique thread in the unnoticed loom of this integration, contributing to a rich tapestry of understanding while providing valuable perspectives complemented by Habermas's (1984) communicative action that shed light on the complex dynamics of AIsmosis.

Navigating this new technological era calls for an increased focus on ethical considerations in AI development. It highlights the necessity for responsible innovation. Stakeholder theory (Freeman, 1984) reminds us that stakeholders are not isolated entities but parts of a larger ecosystem. This awareness emphasizes the need to synchronize our ethical frameworks with the speed of technological advancement. It advocates for a communicative turn in our approach to AI, possible artificial general intelligence (AGI), and the process of AIsmosis. The cautionary note here is the need for a measured pace, particularly in a world, where technology often outstrips our understanding.

The concept of AIsmosis, pivotal to this article, promises to expand the current understanding of the interaction between evolving technologies and society. This exploration acknowledges and probes the complex ties between human development and technological advancement, contributing significantly to the academic discourse on this vital subject.

In the intricate dance of technology and society–pas de deux–controlled AIsmosis emerges as a pivotal concept. It encapsulates AI's balanced, ethically informed integration into society and daily life. This notion illuminates the deep interconnectedness of humans and technology. It stresses the need for technological progress to be aligned with societal needs and ethical norms. The article further reveals the dynamic relationships among various societal elements, depicted as 'tanks' in the controlled AIsmosis framework. In the shifting landscape of AI integration, the interplay of these 'tanks' vividly illustrates the controlled AIsmosis process. Despite its complexity, this dance is integral to maintaining balance in society's shared experience with AI.

The subsequent section outlines the qualitative methodology employed in this study, designed to scrutinize the multifaceted interactions between AI and society. This approach yields valuable insights into the evolution of AI and its impact on societal norms.

METHOD

The principal objective of this article is to interpret the nuanced symbiosis between AI and societal norms, using the concept of 'AIsmosis' as a metaphor for AI's gradual, seamless infiltration into our daily existence. This inquiry primarily centers around examining the ethical, sociocultural, and communicative implications of AIsmosis, advocating for a mindful, ethically driven, and inclusive modus operandi towards AI integration, a conceptual framework the article termed as controlled AIsmosis.

This article's investigative methodology is qualitatively oriented, systematically encompassing the search, selection, and rigorous analysis of extant literature on AI-related themes to amalgamate insights and draw novel interpretations from relevant research. While there may be a scarcity of direct research focusing on AIsmosis, this methodological framework facilitates a holistic exploration of AI's societal interaction, inspecting it through the application of established theories. This exploration, hinged on synthesizing and critically analyzing AI-focused research, strives to yield valuable insights into the evolution of AI and societal dynamics.

Literature Search and Selection

A careful search through electronic databases was undertaken to comprehensively comprehend the subject matter. Search terms were carefully curated to encompass key facets of AI, including but not limited to "artificial intelligence," "AI integration," "society and AI," and "ethics of AI". The inclusion criteria were delineated to select publications that align with the wider domain of AI, make substantial theoretical and conceptual contributions, and offer enlightening insights into the dynamic interaction between AI and society. Furthermore, this analysis is fortified by employing established theories from the literature to search into the integration of technologies within societal paradigms.

Screening and Data Extraction

A screening process was applied to the selected publications, examining their titles, abstracts, and full texts in accordance with the predetermined inclusion criteria. Essential information, including theoretical frameworks and research findings, was systematically extracted from each publication. Although direct references to AIsmosis might be sparse, the emphasis was placed on identifying enriching insights and perspectives to further the understanding of AI's societal integration.

Analysis and Synthesis

The gathered information was subjected to thorough analysis and synthesis, with the aim of identifying overlapping themes, theoretical perspectives, and empirical findings that elucidate the complex interplay between AI and society. These vital insights derived from the literature were cohesively organized and articulated in a coherent narrative that introduces the concepts of AIsmosis and controlled AIsmosis, underscoring their pertinence in comprehending the broader implications of AI.

In addition, the methodology of this article extends to the creation of the digital cave and controlled AIsmosis model's visual representation. Harnessing the capabilities of 'Python's networkx and matplotlib libraries', graphs were generated that visually encapsulate the interconnectedness within the models. In the controlled AIsmosis framework, each node and edge symbolize the integral components of the model and their interrelationships, providing a concise visual aid to complement the exploration.

Limitations

Despite the comprehensiveness and rigor of the implemented methodology, certain limitations must be acknowledged. The lack of direct research focusing on AIsmosis can constrain the breadth of the findings. Furthermore, the article's conclusions are influenced by the quality, diversity, and scope of the existing literature at the time of the review. However, these constraints notwithstanding, it is posited that this approach fosters a robust exploration of available literature, thereby stimulating insightful revelations concerning the evolving dynamics interlinking AI and society.

Utilizing the outlined methodology, the subsequent section, 'Review Findings,' embarks on a comprehensive analysis of AIsmosis. This investigation will be enriched by a critical examination of three foundational theories–SCOT, technological determinism, and diffusion of innovations–which illuminate our understanding of technology-society interactions.

REVIEW FINDINGS

AIsmosis Unleashed

Our contemporary era is witnessing an evolving symbiosis between humanity and AI–an evident change in societal paradigms. This symbiosis is encapsulated in the term 'AIsmosis,' a fusion of 'AI' and 'osmosis,' symbolizing the unrestricted fusion of human and artificial aspects. AIsmosis denotes the passive assimilation of ideas and knowledge, evolving from static internet platforms to dynamic entities that actively engage and seek our attention (Webb, 2023). It is not merely a metaphor but an emphasis on balanced integration, a call for measured AI introduction instead of careless inundation.

AIsmosis's continuous flow of diverse data enhances mutual capabilities between humans and AI. A prime example of this is the use of recommendation algorithms. These tools refine based on vast user data, and it's important to note their potential influence within the attention economy of like, share, and engage (Goldhaber, 1997; Simon, 1971). While they hold significant advantages, they also present challenges, as they could inadvertently amplify societal biases, leading to misinformation and societal polarization (Mittelstadt et al., 2016; Turkle, 2015; Zuboff, 2019).

The transformative power of AIsmosis, driven by significant advancements made in 2023, can be seen across various applications such as ChatGPT, Dall-e, Midjourney, Microsoft Copilot, Photoshop AI tool, and others. These applications, a far cry from early algorithm-driven interactions with AI, are revolutionizing sectors such as retail, CRM, business security, and finance, especially through AI-powered chatbots enhancing customer engagement and sales performance.

However, with these advancements come challenges. A recent event serves as a potent illustration of the potential consequences of unchecked AI-generated content within the framework of AIsmosis. On May 22, 2023, a fabricated, AI-generated photo of an explosion at the Pentagon went viral on social media, causing widespread confusion and even momentarily impacting the US stock market (New York Post, 2023). This incident underscores not only the susceptibility of individuals to misinformation but also the potential for unscrupulous entities to exploit sophisticated AI systems, sowing chaos, and deception online.

Such incidents highlight the need for fairness and transparency in algorithmic decision-making to mitigate biases and adversarial attacks. Indeed, this need has been echoed by numerous scholars, including Caldarini et al. (2022), Starke et al. (2022), Tian et al. (2022), and Zhang et al. (2018). These academics underscore the necessity to understand these critical elements amidst the complexities of AI integration.

In the financial sector, the impact of 'AIsmosis' has been transformative, improving risk management, fraud detection, and customer personalization through big data (Chen et al., 2012). However, this evolution also presents data security, privacy, and job displacement challenges. Organizations must prioritize skill development, industry evolution, and government collaboration to adapt to AI's influence. Investment in training programs can help prepare employees for an AI-infused workplace. Simultaneously, government support is required in the form of education initiatives and robust social safety nets (Frey & Osborne, 2017).

Navigating the terrain of AIsmosis necessitates a sophisticated and adaptive approach, acknowledging its potential to reshape societal structures and remodel human experiences. Developers must consistently evaluate our decision-making processes, ensuring they evolve ethically under the guiding principles of fairness and transparency (Virvou, 2022). Given the unpredictability of AI, it becomes increasingly imperative to maintain a central focus on human agency within human-artificial intelligence interaction (HAII), justifying the continuous safeguarding of human autonomy and aligning AI objectives with societal welfare (Virvou, 2022).

One substantial obstacle to ethically and responsibly approaching AIsmosis is the so-called 'black box' problem in AI, which refers to the often-opaque nature of AI systems' decision-making processes. These processes, which are complexly layered with algorithms, can resist straightforward interpretation and validation, posing significant barriers to transparency and accountability. The problem becomes particularly concerning when AI decisions impact areas such as healthcare, finance, or law enforcement, where incorrect decisions can have severe consequences. Consequently, human oversight and control become critical in this context. Mitigating the challenges of the 'black box' problem requires concerted efforts to demystify AI systems, potentially through initiatives aimed at improving algorithmic explainability or transparency, to ensure their reliability and accountability (Virvou, 2022).

This shift fosters increased collaboration between humans and machines, leading to changes in our roles as technology evolves. Understanding the social and cultural implications is crucial for responsible and ethical AI integration.

Sociotechnical Symphony: Unraveling SCOT Theory

The convergence of science and technology is apparent in the evolution of clustered regularly interspaced short palindromic repeats (CRISPR) gene editing technology, which facilitates precise modifications to an organism's DNA. Discovered during research on bacterial immune systems, the potential applications of CRISPR swiftly expanded beyond the confines of rudimentary research. Presently, CRISPR is employed across several industries, including agriculture and medicine, assisting in the creation of innovative treatments, modifications, and therapies.

Layton (1971) emphasized that the boundaries between science and technology are not rigid and isolated but rather fluid and interconnected. He argued that these domains resemble mirror-image twins, reflecting and influencing each other's evolution and growth. Their interaction transcends mere knowledge and action, extending to social factors such as power, politics, and cultural values (Layton, 1971). This interplay becomes particularly relevant when examining the social fabric of AIsmosis and can help elucidate the 'black box' problem and the role of human oversight discussed previously.

In the context of AI, the fusion of science and technology assumes unique significance. AI is distinguished by its deep-rooted interdisciplinarity, enveloping insights from a wide spectrum of disciplines, including computer science, mathematics, engineering, neuroscience, psychology, linguistics, and philosophy, extending further into the social sciences and humanities. Acknowledging this complex triad–science, technology, and society–is pivotal for a comprehensive understanding of the complex dynamics propelling technological progression and the subsequent implications for our everyday lives.

Within this framework, SCOT theory, posited by Pinch and Bijker (1984), proposes that social factors, power dynamics, political influences, and cultural values significantly shape technological development. At its core, SCOT aims to decipher the mechanisms driving technological transformations within society, emphasizing the crucial role of social constructs in moderating these changes. A prime example of this interplay is the rapid societal integration of digital cameras.

At their advent in the late 1990s, digital cameras bore the stigma of high costs and inferior image quality compared to traditional film cameras. However, the accelerated pace of technological advancement coupled with significant price reductions ushered in the era of digital photography. Various social forces, including evolving cultural perspectives on photography, society's increasing inclination for instant gratification, the convenience offered by digital storage, and the explosive growth of social media platforms enabling photo-sharing, played a significant role in the mainstream acceptance of digital cameras.

Moreover, the availability of digital cameras played a pivotal function in democratizing photography, extending beyond the sole consideration of technical features. This dynamic interplay of societal and cultural forces in the acceptance of technology can also be seen in the rapidly developing field of AI.

Social factors play a substantial role in the expansion and integration of AI technologies. One such factor is the increasing demand for automation and efficiency across various sectors like manufacturing, healthcare, and finance (Brynjolfsson & McAfee, 2014). This demand has stimulated investments in AI research, yielding innovative solutions with the potential to revolutionize these industries. By integrating AI technologies, there is an opportunity for cost reduction, heightened productivity, and improved quality of products and services (Bughin et al., 2017).

However, these advancements bring distinct challenges and opportunities across different industries, necessitating strategic decision-making to navigate effectively. Take manufacturing, for example; while AI promises to boost productivity and cut costs, the automation and robotics that deliver these advantages might also disrupt job roles, triggering the need for reskilling and upskilling initiatives (Smith & Anderson, 2018). Therefore, a key issue is finding a balance between exploiting AI's potential and facilitating a fair transition for workers.

Similar to the manufacturing industry, in healthcare AI offers the prospect of improved diagnostics and personalized care, but it also raises ethical concerns such as patient privacy, data security, and potential biases in AI algorithms (Obermeyer et al., 2019). Effectively addressing these complex issues is crucial to ensure that AI-driven healthcare solutions prioritize patient rights and welfare. As AI continues to be integrated across various sectors, the growing complexity of challenges calls for the development of even more advanced and intuitive systems.

It's crucial to acknowledge that challenges associated with data privacy, algorithmic bias, and job displacement span various industries (Jobin et al., 2019). For example, industries with a heavy reliance on human interaction, such as the service sector, may struggle uniquely with integrating AI-driven automation without sacrificing the human touch.

The growing investment in AGI companies reflects the accelerated evolution of AGI and its rising prominence across various sectors. In early 2023, these companies secured over $21 billion in investments (Financial Times, 2023). Yet, as we venture towards AGI–a technology marked by cognitive abilities mirroring those of humans–we encounter a proliferation of challenges despite substantial strides in emulating human-like reasoning. As noted by a research paper from Microsoft (Bubeck et al., 2023; Metz, 2023), this journey has kindled a robust debate regarding the feasibility of creating an AI system that equals human intelligence. However, the quest to replicate nuanced human traits such as emotions, love, and intimacy in AI systems still poses a significant hurdle.

GPT-4, with its potential in theory of mind (TM) tasks, brings us a step closer to AGI, though it's not without its hurdles, such as confidence calibration, long-term memory, continual learning, personalization, planning, interpretability, cognitive biases, and sensitivity to inputs (OpenAI, 2023). A comprehensive evaluation of such AI systems is crucial to understand their limitations and potential challenges.

For example, GPT-4 technical report acknowledges a "hallucination" tendency, but a more precise description might be "confabulation," which refers to the production of fabricated or distorted information (Bonhoeffer, 1904). This confabulation tendency, which involves the production of unreliable or false information, can introduce unique obstacles in vulnerability discovery and exploitation tasks. When the goal is to identify weaknesses or exploit flaws, the presence of confabulation can make the process more challenging and less reliable. These complexities highlight the importance of integrating intellectual abilities, such as logical reasoning and information processing, along with nuanced aspects of human cognition and emotion. In light of the challenges discussed, TM, originally proposed by Premack and Woodruff (1978), assumes greater significance. TM represents our ability to attribute mental states to ourselves and others, a capability not fully replicated in current AI systems. Early research by scholars such as Piaget (1929) and Selman (1980) underscores the essential role of TM in human development and social interaction, suggesting its relevance in advancing AI and AGI.

TM encompasses the concept that human understanding of the world involves the construction of mental models or representations that include objects and events. Incorporating this understanding into AI and future AGI systems can enhance their interactions by enabling them to anticipate the mental states of other agents. An example of this is Woebot, a chatbot developed by Stanford University. Using natural language processing, it provides cognitive-behavioral therapy to individuals experiencing symptoms of anxiety or depression (Fitzpatrick et al., 2017). Operating based on TM principles, Woebot can identify users' mood patterns and generate personalized recommendations, illustrating the progress towards AGI and emphasizing the significant role of psychological constructs in advancing human-like machine cognition.

Consider the insights shared by Gujral (2023), CEO of Behavioral Signals, at a recent conference. Gujral (2023) highlighted the escalating sophistication of emotional analysis driven by AI and its growing significance in fields such as government defense and security services. This extends to voice-based emotion analysis applications, ranging from predicting instances of duress, fraud, or PTSD, to coaching participants in meetings or interviews and even creating language-independent surveillance. AI-enabled applications, like chatbots and virtual assistants, are emerging as sources of emotional support and companionship for individuals facing loneliness, social isolation, or mental health issues (Bickmore & Picard, 2004). These chatbots can employ sentiment analysis to comprehend and react to user emotions.

Similarly, virtual assistants might use facial recognition technology to discern user moods, tailoring their responses suitably. AI underscores its potential to enhance mental health outcomes, particularly in underserved areas, where mental health resources are sparse, and it can potentially minimize the stigma linked to seeking mental health treatment by offering a more private, accessible alternative. However, integrating TM into AI system architecture is crucial for transparency and explainability.

Studies by Nass and Moon (2000) have demonstrated that humans tend to anthropomorphize technology, imbuing it with human-like attributes and qualities, especially after prolonged interaction. This phenomenon underscores the human capacity to form psychological bonds with technology, perceiving them as companions or social partners. Such anthropomorphism substantially impacts AIsmosis by fostering emotional transfer between individuals and technology, resulting in shared emotional experiences. Emotional contagion, a process, where emotions spread from one individual to another via nonverbal cues like facial expressions, voice tone, or body language (Hatfield et al., 1993), further explicates this.

Emotional contagion and AIsmosis share a notable connection, as they emphasize the transfer of emotions through nonverbal cues and technology-mediated interactions. Consider, for example, Replika–an AI-driven chatbot that serves as an empathetic digital companion. It's capacity to create an immersive, user-centric experience, adapting to the user's communication style and emotional needs, mirrors AI assistants like Samantha in Spike Jonze's film "Her." This adaptability exemplifies the progression of AI chatbots, emphasizing their potential to break down traditional communication barriers and foster profound human-machine interactions.

The swift integration and evolution of technologies, showcased by initiatives like Replika, directly respond to current societal needs. This reflects how society's increasing dependence on technology for communication, companionship, and assistance incentives the evolution and proliferation of advanced AI systems. This trend illustrates a dynamic interplay, where social needs transition from being passive recipients to active catalysts that shape the pace and trajectory of technological advancements.

This aligns with the concept known as the 'domino effect.' Engaging with technology for one purpose often triggers a series of subsequent uses and possibilities, similar to a line of dominos falling one after another. This interaction, in turn, fuels our understanding and integration of technology into various facets of our lives, contributing to the phenomenon of AIsmosis. This continual cascade of innovation and assimilation underscores the dynamic and multifaceted nature of the AI-human relationship within the broader societal context.

When considering AIsmosis, SCOT framework stresses the capacity of technology to not only transform daily life but also significantly influence human cognition and decision-making. This perspective highlights the fundamental role of social factors in developing and integrating AI into societal infrastructures and the consequential societal impacts. A SCOT-based perspective facilitates a broader comprehension of AI's incorporation into sectors such as education, healthcare, transportation, and entertainment. From this viewpoint, technology transcends its materialistic aspect–it becomes a product of social negotiation and interaction among diverse participants, including users, designers, policymakers, and others. Thus, adopting a SCOT approach to AI integration underscores the need to incorporate these stakeholders into decision-making processes, respect their perspectives and values, and ensure that the AI technologies developed to align with societal goals (Pinch & Bijker, 1984).

As AI technologies become more entwined with our daily lives, they hold the potential to reshape, even challenge, longstanding norms within Habermas's (1987) lifeworld. Such influences could disrupt traditional social interactions, redefine cultural practices, question conventional understandings, and amplify social disparities. A fascinating interaction arises between Habermas's (1987) lifeworld and AIsmosis when the latter triggers a reassessment of foundational lifeworld assumptions and practices. For instance, the increasing reliance on AI in decision-making may call for deep reflection on notions of agency, responsibility, and accountability.

Conversely, AI's communication and language processing roles extend beyond simply facilitating conversations. AI has the potential to decode the complex matrix tangled within our social interactions, encompassing implicit biases, subtle emotional cues, cultural variations, and power hierarchies. A nuanced AI design can illuminate these layers, fostering new insights into our communication details. These enlightening discoveries may provoke a reassessment, and possibly a challenge, of our deeply ingrained communicative patterns.

AIsmosis offers an array of resources that promise to bolster the lifeworld by amplifying communication and collaboration. It's not merely about language proficiency or response speed; AIsmosis promises to transcend traditional barriers such as time, distance, and cultural differences. This potential also necessitates a deeper analysis of AIsmosis within the wider context of the theory of communicative action (TCA) and Habermas's (1987) lifeworld concept.

Lifeworld embodies societal assumptions, values, beliefs, and linguistic practices that shape personal actions and interactions, laying the foundation for social integration and norm development through mutual understanding, empathy, and dialogue (Habermas, 1987). However, with the continued advancement of AI and the potential emergence of AGI, there's a risk of accelerating the 'colonization' of the lifeworld. This could alter our shared norms and interpersonal dynamics (Habermas, 1987). As such, thoughtful deliberation and the implementation of governance structures become paramount, ensuring that AI and AGI technologies align with and respect the values entrenched in the lifeworld.

The colonization of the lifeworld could potentially result in social pathologies, such as alienation, identity loss, and disconnection from societal values and peers. Turkle (2015) underscores this concern, noting a marked decline in face-to-face interactions as we continue to progress into a highly tech-centric society. Such a shift risks diminishing the nuanced subtleties integral to social interaction and crucial for fostering empathy and understanding. Turkle (2015) further argues that communication technologies may impair the ability to decode nonverbal cues–an essential aspect of understanding others' emotional states–thereby potentially reducing empathy levels and amplifying feelings of loneliness. The impact of technology on human interaction shapes societal transformation. Understanding this influence is crucial to mitigate potential negative effects and foster a connected society. However, this shift may also pose risks to societal stability and unity. Habermas (1987) proposed a reconfiguration of the lifeworld-system relationship to counterbalance the effects of this rationalization.

This reconfiguration differentiates between two domains: the system world, associated with utilitarian logic, and the lifeworld, linked with communicative action (Habermas, 1987). The advent of AI, capable of interpreting and processing vast amounts of data, stands poised to redefine these boundaries. By uncovering patterns too obscure for human cognition, AI has the potential to impact the lifeworld significantly, enriching our knowledge base and possibly contributing to an over-rationalization of human experiences.

Habermas (1984) characterizes the system world as dominated by instrumental rationality and bureaucratic structures, while the lifeworld safeguards subjective meanings, values, and cultural practices. AI's controlled and gradual integration into society–AIsmosis–underscores the urgent need to recalibrate the lifeworld-system relationship. It advocates for fostering an equitable, democratic society that values human relationships and meaningful communication, strongly emphasizing engaged dialogue.

AI systems also have an immense capacity for analyzing and processing large volumes of data, aiding individuals in deciphering complex situations and detecting patterns that might elude human cognition (Floridi, 2014). This ability can enrich our understanding of situations and experiences and potentially guide decision-making. Yet AI's potential influence extends beyond analytical assistance. It may impact our personal experiences and life narratives by providing tailored recommendations and insights and analyzing behavioral, preference, and social interaction data. However, this potential must be harnessed cautiously, ensuring it supports stakeholders in making informed decisions and understanding their world rather than amplifying biases or undermining human agency. Thus, the challenge remains to integrate AI into society while preserving lifeworld values and practices. It's pivotal to understand that AI may not entirely capture the subjective interpretation, conceptual thinking, and meaning-making inherent to human cognition. Although there are similarities, substantial differences remain in the strategies and use of knowledge between humans and AI systems.

AI systems come with their share of potential ethical implications and inherent limitations. A pertinent example is the UK government's 2020 adoption of an AI-powered algorithm for determining high school grades, which, due to the COVID-19 pandemic's cancellation of in-person exams, resulted in inaccurate and unjust grades for many students, particularly those from disadvantaged backgrounds. AI algorithms were critiqued for their lack of transparency and their potential to reinforce existing biases and inequalities in the educational system (Baker & Hawn, 2021).

There's a risk that AI systems could reinforce existing power structures, given that entities with more substantial resources and data access can exert considerable influence over AI's development and implementation. This dynamic might limit the diversity of perspectives in communicative action, potentially degrading social norms and practices. The impact of AI on the interaction between the lifeworld and the system hinges on its development, implementation, and governance. Therefore, it's essential to engage stakeholders from various cultural and social backgrounds in AI development and deployment. This inclusivity ensures that AI systems respect the lifeworld's cultural norms, values, and practices, fostering meaningful interpersonal exchanges and social cohesion.

'AIsmosis,' could bring about significant transformations in the lifeworld, subjugating its inherent values to system imperatives. This intrusion can be seen through AI's infiltration of the lifeworld via data analysis and personalized content feeds. Drawing a parallel with Habermas' (1987) concept of media 'delinguistification,' AI could colonize the lifeworld by dominating data, content, and communication channels, consequently undermining its openness and diversity. This raises concerns about the potential erosion of lifeworld's openness and diversity as AI systems assume control over information flow.

The growing influence of technology on communication channels may inadvertently transform individuals into passive information consumers, thereby stifling their ability for critical evaluation and active engagement. As AI begins to shape and control information flow, there arises a crucial need to examine if this could lead to an 'information monoculture.' This concept, inspired by the agricultural practice of cultivating a single crop and thereby eliminating diversity, alludes to the potential marginalization of diverse ideas and perspectives.

Instead, narratives may favor algorithmically determined outcomes that predominantly reflect the interests of the system's architects and controllers (Pariser, 2011; Sunstein, 2001). Here, SCOT theory reminds us that technologies, including AI, are social constructions that reflect the values and power structures of the society they originate from (Pinch & Bijker, 1984).

The increasingly data-driven nature of AI surveillance presents the risk of exacerbating social inequality and fragmentation (Eubanks, 2018). The construction of these algorithms from consumption data, decision-making patterns, and areas of interest could reshape societal and communal structures. With such transformations, an overriding emphasis on technical rationality could potentially overshadow democratic dialogue, leading to implications such as job displacement and rising inequality due to AI technologies. This impact might hinder proactive participation in the lifeworld. Yet, according to SCOT theory, these structures are not neutral or inevitable outcomes of technological progression (Pinch & Bijker, 1984). Instead, they are shaped and defined by diverse social actors and their varied interpretations and responses to these technologies. Hence, biases in AI algorithms are not merely technical artifacts. They represent specific design choices influenced by the social and cultural values of the designers.

Transparency or public engagement could lead to technocratic colonization of the lifeworld, risking societal values and ethics. However, SCOT theory posits that societal engagement, negotiation, and contestation can influence the shape and direction of technologies (Pinch & Bijker, 1984). This perspective gives rise to a critical question: How can societal engagement and deliberation about AI's design and use be promoted? How can the domination of a single narrative or interest group be prevented? And how can respect for diversity within the democratic dialogue of the lifeworld be ensured?

Leading this development are corporations such as Alphabet, Microsoft, NVIDIA, Tesla, and IBM, which are at the forefront of AI research, development, and application (Capital.com, 2023). The surge in investment during late 2022, in conjunction with the launch of OpenAI's ChatGPT, underscores this trend, indicating considerable growth prospects for the industry, particularly for AI-heavy stocks like Alphabet and Microsoft.

The advent of AI's influence brings the potential risk of transforming significant interpersonal communication into mere quantifiable data patterns, thereby challenging entrenched societal norms such as privacy and autonomy. This risk escalates when AI mediates our interactions, threatening the essence of communicative action rooted in mutual understanding and empathy. This mediated interaction, reduced to data, can feed into the AI systems, fueling their influence over individual and societal behavior.

Despite the challenges, it's essential to acknowledge that not all interactions within this system–a process referred to in this discussion as 'AIsmosis'–are deceptive or lack authenticity. Nonetheless, most consumer decisions and responses to curated content inadvertently fuel this system, thus creating a self-sustaining cycle of feedback loops (Hofstadter, 2008). This escalating dependence on algorithms and data-driven decision-making processes might erode individuals' autonomy and sway over their lives, possibly leading to reduced democratic engagement and a diminished ability to shape personal experiences Crawford (2015).

The allegory of the cave, a prominent philosophical metaphor from Plato's "The republic," serves as a fitting appraisal of AIsmosis. In a contemporary interpretation of this parable, individuals inhabit a digital cave, continuously exposed to data and experiences personally tailored by AI algorithms. While these algorithms attempt to imitate the external world, they inevitably apply filters and biases that resonate with each person's established beliefs and values.

In this allegorical cave, AI has assumed the role of a personal puppeteer for each inhabitant. Rather than puppeteers shaping experiences, AI algorithms take control, molding worldviews, providing information, and curating experiences for individuals. As these algorithms gradually align with personal preferences and biases, the content they produce amplifies existing beliefs, engendering a potentially harmful cycle of AIsmosis that can narrow worldviews–much like the prisoners in Plato's original parable. Thus, like Plato's cave dwellers, individuals confront a reality dictated by technological programming, heavily influenced by the socio-cultural context of its creation.

This virtual allegory reflects societal perceptions and principles, where developers' and users' cultural values and biases significantly influence the process of AI technology, casting shadows on the cave wall. There exists a challenge in our current lifeworld to validate research outcomes emerging from this digital cave.

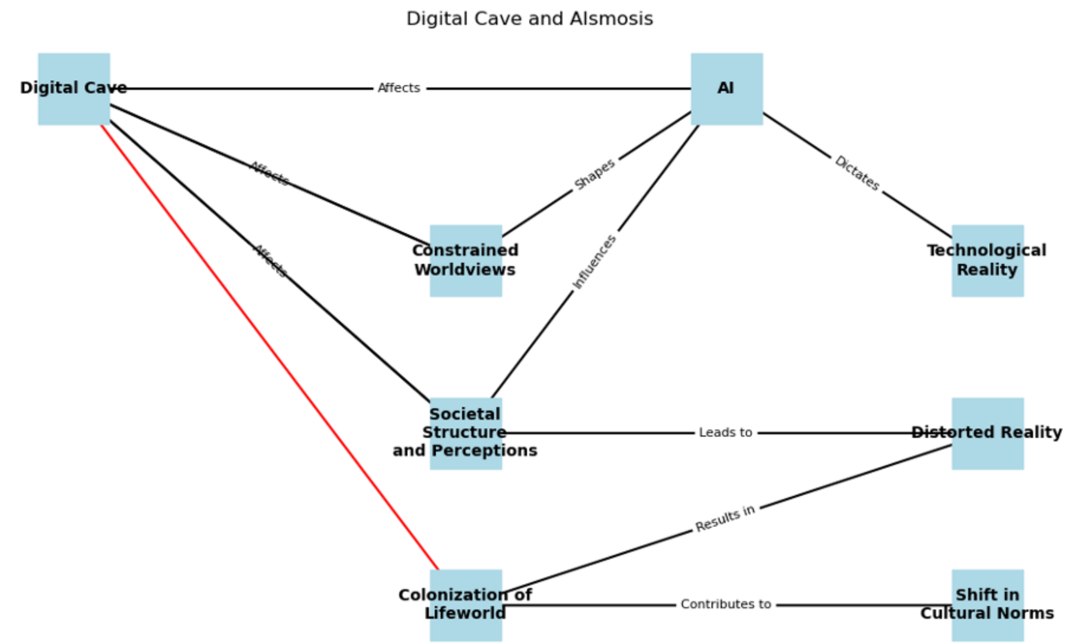

Figure 1. Digital cave & AIsmosis (Source: Author, using Networkx and Matplotlib libraries)

As illustrated in **Figure 1**, this scenario presents a distorted reality where the continuous manipulation of information obscures societal transparency. This distortion arises from the coexistence of AI and humans in distinct realities. AI, inherently alien to our world, can cast empathy-free shadows on our cognitive cave walls, exacerbating the potential for unchecked influence. If left uncontrolled, this could pave the way for the colonization of the lifeworld when the system's instrumental rationality, prioritizing control and efficiency, begins to dominate, subverting the communicative and normative aspects of our society.

The growing role of AI surveillance underscores the impact of societal and cultural elements on technology, prompting an urgent need to scrutinize the values and assumptions guiding its development and application. As theorized by Habermas (1987), the unchecked growth and influence of AI could risk 'colonization of the lifeworld,' transforming fundamental aspects of our societal interactions and norms.

This includes the rise of AI surveillance, which holds the potential to reshape societal perceptions of privacy and autonomy. As individuals grow more accustomed to invasive surveillance practices, there is a growing willingness to trade privacy for perceived security or convenience (Zuboff, 2019). This shift in cultural values contributes to the colonization of the lifeworld, as individuals adopt the norms of the surveillance society and lower their privacy expectations (Fuchs, 2013). The advent of AI ushers in a new level of responsibility, necessitating a thorough examination of its implications. Beyond mere adoption rates, the diffusion of AI innovation includes intricate dynamics that ripple through society. For instance, consider Hangzhou, China's "City Brain" initiative. This program employed facial recognition technology to facilitate access to government services, intending to streamline services like transportation and healthcare. However, it met significant resistance from citizens concerned about potential privacy infringements and misuse of their personal data.

In the context of surveillance capitalism, the gradual erosion of privacy and civil liberties is a growing threat to essential aspects of society–culture, individuality, and social interaction. Büchi et al. (2022) suggested that such a system could exacerbate the 'chilling effect,' leading to diminished freedom of expression and democratic discourse. Continuous monitoring cultivates unease among individuals, deterring them from expressing dissenting opinions or engaging in contentious debates for fear of possible repercussions. This apprehension could potentially trigger an escalation in the spiral of silence (Noelle-Neumann, 1974). In this theory, the fear of isolation leads individuals to withhold their opinions on controversial issues, thereby stifling healthy democratic discourse.

As such apprehension grows, the spiral of silence could amplify, further suppressing open conversation and critical debate. Borrowing from Foucault's (1977) Panopticon theory, the increasing AI surveillance can instill self-discipline and self-censorship as individuals internalize the gaze of the surveillance system. The pressure of continuous monitoring and assessment can foster a sense of relentless examination, promoting conforming behavior and stifling individualism. Recognizing the role societal forces play in shaping AI's trajectory necessitates a critical reflection. Grounded in Foucault's (1977) power/knowledge concept, it's clear that AI surveillance, as a socio-technological construct, is intimately linked with power dynamics. This technology's design and application not only reflect but also reinforce existing power structures. Consequently, it is essential to question who truly directs the development of this technology and whose interests are prioritized.

Foucault's (1977) assertion is that power is not a unilateral force but rather diffused across social networks. However, the societal acceptance of AI surveillance may not only consolidate these power structures but could also contribute to the colonization of the lifeworld, as Habermas (1987) warns. Thus, underlining the need for a critical analysis of societal power dynamics in the acceptance and construction of such technologies. This nuanced understanding can illuminate the complex interplay between society, power, and AI, which is necessary for navigating the ethical terrain of AI surveillance. AI communication tools' collection and analysis of personal data can expose individuals to the risk of cybercrime, fraud, and identity theft. Cyber scammers are already leveraging voice cloning AI tools to impersonate victims' relatives or friends in distress needing financial assistance (Fortune, 2023a). To counter the issue of privacy invasion by AI surveillance systems, it is vital to establish robust legal and regulatory mechanisms that safeguard individuals' privacy rights. Furthermore, individuals must be equipped with the power to exercise their privacy rights, including the right to access and control their personal data (Benjamin, 2019).

Pasquale (2015), in his book Black box society, argues that the current legal framework may be inadequate to address the complex and evolving challenges posed by AI-powered tools, underscoring the necessity of careful integration of AI into the lifeworld to prevent adverse consequences. According to SCOT theory, the integration of AI into society is a complex process shaped by diverse actors, including engineers, policymakers, and users. Each of these actors, driven by their individual values and objectives, contributes to the development of AI within unique social and cultural contexts. Consequently, our experiences with AI are not solely dictated by the technology itself but are also influenced by the societal and cultural processes that frame its development and adoption.

Illusion of Inevitability: AIsmosis and Technological Determinism Paradox

Technological determinism contrasts with SCOT perspective by asserting technology as the principal catalyst for societal change, operating independently of human influence or societal elements. Within this perspective, technology's societal impact, represented metaphorically as the digital cave in **Figure 1**, is predetermined and predictable, following a linear and inevitable trajectory.

This dictating interplay between AI and the digital cave, as illustrated in **Figure 2**, characterizes the essence of technological determinism. This concept hypothesizes that technology, AI included, functions as the primary driver for societal transformation, operating independently of human mediation or societal constituents. Similar to the transformative repercussions triggered by the emergence of nuclear weaponry in the geopolitical arena, AI's encroachment into the digital cave suggests profound implications for our societal fabric. It echoes the deterministic view that the influence of technology on society follows a preordained, linear progression.

The dawn of the atomic age, incited by the discovery of nuclear fission in the late 1930s, prompted an unprecedented acceleration in nuclear technology development, climaxing in the devastating bombings of Hiroshima and Nagasaki in 1945. The following proliferation of nuclear weaponry incited an intense arms race between the United States and the Soviet Union during the Cold War. Today, the drive to maintain and augment nuclear capabilities persists as nations grapple with the strategic and geopolitical implications of exercising such powerful arsenals in a progressively complex and multipolar global context. This enduring reality, as reflected in the ongoing tensions between Ukraine and Russia, underscores the lasting influence of technologies on international security and diplomatic interactions while exemplifying the tenets of technological determinism, underscoring the relentless impact of technological evolution on human history.

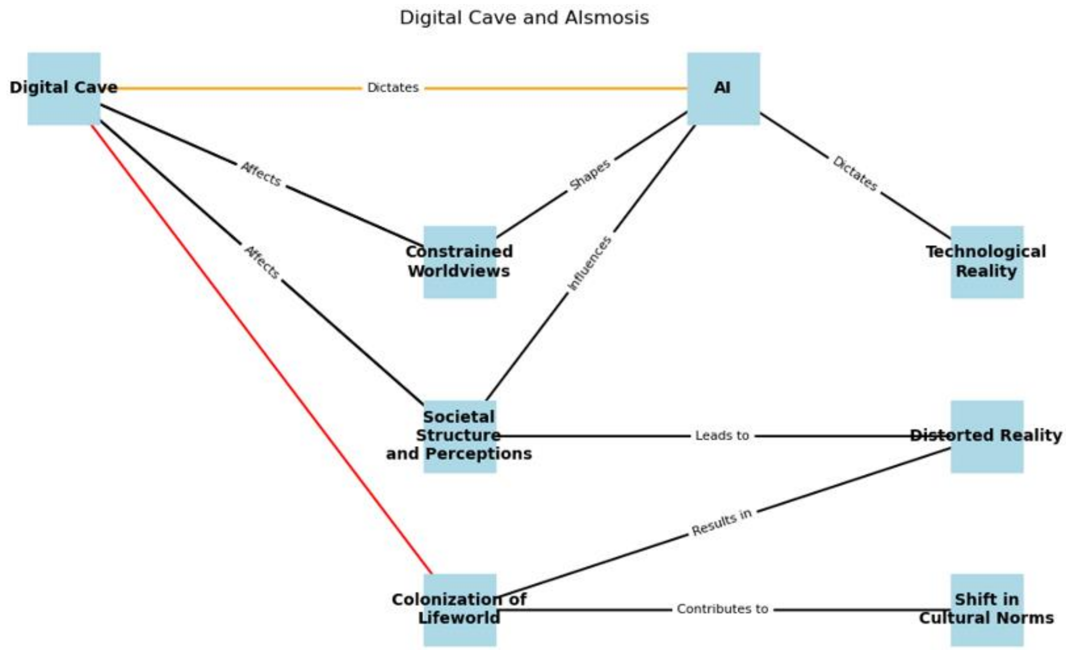

Figure 2. Digital cave & technological determinism (Source: Author, using Networkx and Matplotlib libraries)

Historian and philosopher Yuval Noah Harari suggests that AI could present a risk to humanity as severe as climate change and nuclear war. Although AI is not a direct equivalent to nuclear weapons in terms of destructive potential, the competition to achieve AI supremacy could significantly alter global power structures, affecting military capabilities, economic growth, and technological evolution. With the evolution of AI technologies and the potential emergence of AGI, the sphere of politics stands on the brink of profound transformations. If we consider AGI–an AI system with the ability to understand, learn, adapt, and implement knowledge across a wide range of tasks as efficiently as a human–it can aid in processing massive amounts of complex information far beyond the capabilities of human analysts. It could track social, economic, and political trends in real-time, giving leaders a more comprehensive understanding of the global situation. Such insights can guide policy-making processes, helping countries navigate the complicated maze of international relations, respond to global crises more effectively, or even reduce the risk of escalations.

Veblen's (1899, 1921) seminal works, 'The engineers and the price system' and 'The theory of the leisure class', provide an avenue for understanding the complex relationship between technology, society, and the economy. In these influential texts, Veblen (1899, 1921) explains how technology is a key component in shaping economic systems and societal structures. Veblen's (1904) idea of 'the machining process' further elucidates how technological advancement can induce revolutionary changes in economic and social frameworks. He suggested that this process could engender more efficient and logical production management methods, albeit potentially hindered by existing societal structures and vested interests (Veblen, 1904). In the context of AI, this perspective implies that its evolution and utilization are subject to overarching economic and social structures, such as capitalist production paradigms, consumer culture, and geopolitical dynamics, including international disputes and rivalries. Contrary to the view of technology as an independent and autonomous force, Veblen (1899) posits that these societal structures and economic systems significantly influence its development and application. As such, the perspective of technological determinism may risk overlooking the role of social factors in shaping the digital cave and disregarding potential alternative pathways or outcomes. Veblen (1899) suggests that institutions, being historical constructs, may not always align with contemporary realities, thereby emphasizing the need for periodic reassessment and adaptation to the continuously evolving technological milieu.

The process of AIsmosis signals a significant transformation in societal infrastructure. The transformation warrants careful consideration of its potential impacts on established institutions and power dynamics. Institutional conservatism, however, may stifle adaptation to this emerging technological landscape. As a result, outdated power structures and economic systems could be perpetuated, remaining ill-suited to current conditions (Veblen, 1899). Addressing this issue necessitates a flexible and responsive institutional framework capable of adapting to technological changes. This involves a two-fold approach: a critical examination of existing institutions fulfilling diverse social, economic, and political roles and an assessment of their compatibility with recent technological advancement.

The range of institutions that must adapt to these technological advances is expansive, encompassing government agencies, regulatory bodies, educational and healthcare organizations, financial institutions, non-profit entities, and cultural establishments such as museums and libraries. Each of these entities will face specific requirements for adaptation, depending on the context and nature of the innovation. This could lead to necessary changes in areas such as intellectual property laws, privacy guidelines, and cybersecurity protocols.

Media technologies exemplify this trend, possessing the dual potential to disseminate information widely and simultaneously restrict its flow by establishing ‘monopolies of knowledge’ that marginalize alternative cultures and knowledge systems (Innis, 1950). These monopolies can influence societal evolution and structures considerably, suppressing or marginalizing certain perspectives. Thus, media technologies serve as potent architects of perceptions and behaviors, contributing to the creation and expansion of a digital cave. Technological progress extends this digital cave, illuminating and obscuring in equal measure, affecting societal perceptions and behaviors through myriad actors.

Cultivation theory (Gerbner et al., 1980) emphasizes the pivotal role media plays in molding our perceptions and guiding public discourse. Applying this theory to the digital age, it becomes clear that AI has the potential to significantly influence the subjects we contemplate. However, when personal biases intertwine with a possible AGI, an ‘altered mass media’ can emerge, wherein misinformation can be amplified and propagated more effectively.

Various scholars across communication, sociology, and science and technology studies have criticized technological determinism for oversimplifying the intricate relationship between technology and society. For example, Winner (1986) has argued that technological determinism neglects the social and political context in which technologies are developed and used and the potential for democratic control and influence over technology. Similarly, Haraway (1991) has criticized technological determinism for ignoring the ways in which social identities, such as gender and race, shape technological development and use because technology is not neutral but is instead shaped by the interests and values of the people and institutions that produce and use it. Thus, technology should not be evaluated in isolation. It operates within a specific social and political context and is influenced by social identities and power dynamics. Controlled AIsmosis, with its emphasis on a nuanced and context-specific approach to technological change, can help to ensure that the benefits of technology are shared equitably and that its potential risks and challenges are thought of effectively.

Ultimately, AI’s influence on inequality hinges on its development, implementation, and regulation. For instance, AI algorithms trained on biased data sets could potentially exacerbate existing inequalities. Conversely, mindful development and implementation of AI, considering potential biases and discriminatory outcomes, may serve to mitigate inequality and enhance social outcomes. When considering the potential for bias in commercial data used in AI systems, it is crucial to address several factors. Commercial data, collected primarily for targeted advertising purposes, may lack the necessary diversity and representation to reflect the broader population or marginalized groups accurately. Biases can emerge during the data collection process, reinforcing and amplifying existing societal inequalities within the dataset (Moore, 2023).

In light of these complexities, recognizing the factors that drive the adoption and diffusion of AI becomes even more critical. Such understanding is paramount not just for its potential to illuminate how AI technologies permeate society but also for its implications in ensuring equitable technological development. AI adoption and diffusion, characterized by their rapid pace, are propelled by both the anticipated advantages and the pursuit of profit and competitive edge. The theory of diffusion of innovations serves as a clarifying the process of how AI technologies permeate society.

Understanding Adoption and Diffusion of Artificial Intelligence in Society

According to Rogers’ (1962) diffusion of innovations theory, the adoption of new technologies, such as AI-generated digital models exemplified by Lalaland.ai, is underpinned by perceived advantage, compatibility with existing systems, complexity, and trialability. The swift integration of such technologies in industries like fashion and advertising is driven not only by tangible benefits like improved efficiency but also by societal demands for increased diversity and inclusivity. However, while the promise of an AIcentric Revolution is compelling, it raises significant concerns. Issues such as equitable profit distribution, workforce reskilling, and potential misuse surface alongside the diffusion of AI technologies. These complexities inherent in AI may hinder its widespread acceptance and diffusion, despite its perceived advantages and system compatibility across diverse sectors, including education, entertainment, and customer service.

This dynamic is particularly relevant considering recent advancements in the realm of digital olfaction. Google researchers Gerkin and Wiltschko (2022) have designed a graph neural network (GNN) model that creates a link between molecular structures and odors, resulting in the Principal Odor Map. This tool has potential applications in a variety of fields, from food and fragrance creation to environmental monitoring and disease detection. As we consider a future, where AGI can decipher and engage with olfactory data, the diffusion of this technology underscores the broader challenges and opportunities associated with the adoption of advanced AI technologies in society. Hypothetically, dedicated sensors could capture and digitize olfactory data akin to visual and auditory information harnessed by cameras and microphones. This advancement would empower a possible AGI to analyze and respond to digitized scent data in ways distinct from human perception. However, this comes with potential risks and ethical implications, such as misuse of personal data and power concentrated among a few entities, emphasizing the imperative for responsible and ethical AI deployment.

The diffusion of innovations theory suggests that various factors influence the rate of adoption and diffusion through society, such as the perceived relative advantage, compatibility with existing systems, complexity, and observability (Rogers, 1962). This implies that the adoption of AI technologies is influenced by their technical capabilities and how they fit into the socio-technical landscape. As social media users demand more personalized and engaging experiences, AI technology has been developed and deployed to meet these demands. This shift towards personalized and automated experiences reflects Veblen's (1904) concept of technological determinism, which suggests that the development and diffusion of technology can shape cultural values and lead to social and economic transformations.

One notable instance is the case of Caryn Marjorie, a Snapchat influencer who launched an AI chatbot modeled on herself. This chatbot serves as a 'virtual girlfriend' for $1 per minute, illustrating the monetization of simulated intimate experiences (Fortune, 2023b). This example showcases how AI technology is harnessed to cater to users' desires for customized and immersive interactions, aligning with the evolving expectations of social media users and highlighting the potential monetization of simulated intimate experiences. 'Controlled AIsmosis' in the face of rapid integration for profit is a daunting task.

The entertainment industry has also seen a rise in the use of AI-generated music and films. An example of AI-generated content is the 2018 film "Zone out," which was written and directed by an AI system named Benjamin (2019), trained on thousands of science fiction scripts to generate a unique script for the film. Lil Miquela, an AI and 3D animated virtual influencer, garnered substantial attention on social media in 2019 due to her lifelike appearance and compelling content, her creators employing machine learning for personality creation and post generation. The AI-generated album "I AM AI," released in the same year, featured music created entirely by AI algorithms.

Hip-hop artists, such as producer Timbaland, have sparked discussions by expressing plans to produce AI-generated songs featuring the voices of deceased artists, raising ethical, psychological, and commercial considerations surrounding the use of emerging AI technology in music (MSNBC, 2023). Kendrick Lamar's music video for "The heart part 5" used deep fake technology to morph into various public figures, conveying a message about fame and identity in society and demonstrating the potential of deep fake technology as a tool for creative expression. These examples illustrate the diverse ways in which AI is transforming creative expression and entertainment industries.

Another example of AI adoption is its use in the music industry, demonstrated in the case of "Heart on my sleeve." This song, leveraging AI to replicate the voices of renowned artists Drake and the weekend, garnered popularity across various digital service providers until Universal Music Group's infringement claims led to its removal (BBC News, 2023). The use of generative AI to create music may have initially been seen as a novel and innovative application of technology, but its violation of copyright law and potential to harm artists' interests demonstrates the need for responsible AI development and deployment and the importance of ethical considerations in the diffusion of new technologies. Though, Grimes' recent announcement on Twitter that she would share 50% of profits on "any successful AI-generated song" using her voice highlights the potential impact of AIsmosis on the creative industries (The Verge, 2023).

This example demonstrates how AI can transform the traditional dynamics of creative production and intellectual property and foster new collaborative and innovative approaches to creating art. The willingness to embrace AI-generated music signifies the growing influence of AI on the creative process and its potential to reshape the future of artistic expression. However, the primary beneficiaries of this innovation are tech firms with leading music publishers. This situation echoes Adorno's (1991) critique of the culture industry, emphasizing the risk of power and profits being further concentrated in the hands of a select few, thereby promoting standardized cultural production and stifling individuality and critical thought.

The diffusion of AI technologies involves various sectors. For instance, with the advancement of 5G technology, AI can significantly enhance autonomous vehicles' capabilities. These vehicles can communicate more quickly and reliably with each other and their surroundings in real time. However, this evolution raises ethical issues, like the potential for robot-induced fatalities, sparking adverse public reactions. World Economic Forum (2022) discusses how autonomous decision-making may lead to possibly fatal errors, thus echoing the enduring ethical conundrum known as the trolley problem. This moral dilemma resonates with the development of self-driving cars since they are programmed to decide, potentially involving trade-offs between passengers' safety and the safety of others on the road (Foot, 1967). The education sector is witnessing an accelerated influx of AI-powered platforms, driven primarily by significant investments from Big Tech companies and influential publishing houses. An example of this trend is Pearson's IBM Watson Tutor, a highly sophisticated AI educational tool. Despite the potential benefits for consumers, this concentration of power carries inherent risks, such as the potential misuse of vast amounts of personal data collected without explicit consent (Council of Europe, 2022).

Big tech companies, owing to their extensive data resources and computational power, exert substantial influence over the development and deployment of AI technologies. This influence not only shapes the course of technological innovation but also dictates the role of AI in society. This dominance, if left unregulated, could potentially lead to 'technological determinism,' a scenario, where these corporations determine the future direction of AI deployment. A case in point is the Microsoft Build Conference, where Microsoft showcased recent developments in AI, introducing a framework for constructing AI applications and copilots. These applications utilizing advanced AI and large language models (LLM) for complex cognitive tasks are evidence of the increasing pace of AI adoption.

The diffusion of AI technology, referred to as 'AIsmosis,' is not a mere consequence of supply-demand dynamics but is equally influenced by larger economic and social structures rooted in capitalist modes of production and consumer culture. This underscores the necessity of regulation to guarantee a balanced and sustainable integration of AI in the long term. The recent launch of Cicero, an AI developed by Meta, underlines the burgeoning capabilities of AI. Cicero, capable of achieving human-level performance in the strategy game Diplomacy, demonstrates a significant leap forward, highlighting the potential of AI in complex decision-making processes. This transcends the broad usage of AI for targeted data applications and personalized political messaging. However, as AI continues to permeate different facets of our lives, it's vital to understand its influences, which are not limited to profit motives but also shaped by wider economic and social structures rooted in capitalist production and consumer culture.

AI's potential is not confined to commercial benefits; it can also make substantial contributions to societal needs. A prime example was during the Ukrainian crisis when Deloitte, in collaboration with NGOs and the government, swiftly developed an AI-powered virtual contact center called IRENA. Revolutionizing crisis communication and data collection, IRENA enabled faster and more efficient dissemination and collection of information. This successful deployment of IRENA accentuates AI's ability to provide prompt, effective responses in crisis situations, underlining its role as an invaluable tool for emergency management and humanitarian aid (Deloitte, 2021). As global conflicts persist, the question remains: Can such AI tools be developed and made accessible worldwide, given varying levels of technology access? The uncertainty is whether nations lacking these resources will be disadvantaged or if global cooperation can ensure equitable AI deployment in crisis management.

Competition for global dominance in AI technologies is heating up, driven by concerns over global competitiveness and national security. Governments worldwide are pouring significant investments into AI research and development, understanding the potential of AI to stimulate economic growth and enhance military capabilities and geopolitical influence (Taddeo & Floridi, 2018). This rapid AI advancement has sparked intense competition between geopolitical powerhouses, such as China and the US. Both nations are competing for AI supremacy, marked by their aggressive efforts to develop and deploy cutting-edge AI technologies.

Emerging technologies like LLMs and multimodal large language models (MLLMs) present both opportunities and challenges in the AI landscape. LLMs, exemplified by ChatGPT4, have demonstrated remarkable capabilities in generating human-like language. However, they have also led to challenges such as deepfake videos and disinformation campaigns. This has prompted authorities in regions like China to deem them unsafe. MLLMs, representing a significant evolution from LLMs, can process multiple data modalities–text, images, and speech–to produce contextually relevant outputs. This attribute enables MLLMs to be applied in a wider range of applications, including image and speech recognition and multimodal dialogue systems. Models like OpenAI's beta DALL-E 3 and DALL-E 2, which generate images from textual descriptions, exemplify this advancement.

In a CBS 60 minutes interview, Google executive James Manyika revealed an intriguing discovery: AI technology had acquired language capabilities without explicit targeted training. This unexpected development points to the unpredictable nature of AI, which has been described as operating in a "black box" by Google CEO Sundar Pichai. Harris and Raskin (2023) identified this unpredictable learning characteristic as inherent to recent AI advancements, particularly in the development of generative large language multimodal models (GLLMMs). Metaphorically drawing a parallel with 'golem-class' entities from Jewish folklore, these models are capable of continual self-improvement, a trait seen in GLLMMs through their use of self-generated training data. In this context, Huang et al. (2022) illustrated this growing complexity, demonstrating that the 'chain-of-thought' (CoT) reasoning mechanism allows LLMs to improve autonomously without reliance on 'ground truth' outputs. AI personal assistants like Siri, delivering improved responses, can mirror this development, potentially progressing towards human-equivalent cognition.

The escalating competition has catalyzed both public and private investments in AI, accelerating its development and promoting widespread adoption. The spread of AI across sectors is largely propelled by its perceived advantages and compatibility with existing systems. These factors significantly determine the rate of technology adoption and its subsequent dissemination (Rogers, 1962). From a business standpoint, AI brings substantial benefits, such as the potential to increase profits and strengthen its position in the attention economy. This compatibility fosters adoption across different fields, including customer service and education. The benefits of AI extend to streamlining and automation of processes, generating predictive insights, and refining traditional methods. World Economic Forum (2018) anticipates that AI's integration into the manufacturing sector could potentially boost global GDP by 16% by 2030, mainly by optimizing production processes and reducing costs. However, this transition presents inherent challenges: it requires considerable investments in emerging technologies and substantial workforce upskilling to manage these innovative tools effectively.

Brynjolfsson and McAfee (2014) argue that AI and other technological innovations are triggering significant shifts in the labor market, disrupting traditional job roles while concurrently creating new avenues for economic growth. This dual impact is global, not confined to any specific locale. World Economic Forum (2018) highlights the extent of this transformation, emphasizing that numerous countries are tackling potential job threats due to automation. The report reveals that 43% of surveyed businesses anticipate a reduction in workforce size due to technology integration, while 41% predict increased reliance on contractors and temporary staff.

McKinsey Global Institute (2017) notes that while only a few occupations are entirely automatable, approximately 60% of all occupations have at least 30% of activities that are technically automatable. This emphasizes the potential for considerable labor displacement and highlights the urgency of addressing the growing skills and employment gap between highly-skilled and low-skilled workers. In this dynamic landscape, the unregulated proliferation and dissemination of novel technologies could result in knowledge and power imbalances among different societal groups. Those with privileged access and control over these sophisticated technologies could potentially exploit this advantage, accruing disproportionate economic, political, or military influence.

This scenario could precipitate novel forms of inequality and domination, posing substantial risks and challenges for global governance. As highlighted by van Dijk (2020) and Warschauer and Matuchniak (2010), uneven access to technology could widen disparities in education, employment prospects, and social mobility, further entrenching existing social inequalities. To address this concern, the responsibility directly rests with educational institutions and governments to prioritize initiatives aimed at bridging the digital divide. These include ensuring learners' access to technology and fostering the necessary skills and knowledge for its effective application.

It's important to realize that access to technology alone is not enough. It's crucial for educators to enforce responsible and ethical usage of technology among learners. This involves education about digital citizenship, online safety, and data privacy, along with media and visual literacy. Given the ubiquity of AI in various facets of life, a multidisciplinary educational approach in institutions is necessary. These examples underscore the necessity of considering technological innovations' ethical and social implications. However, the question of whether AI should be democratized or remain tightly regulated remains a contentious topic. While increased access to AI can bring significant benefits, like increased efficiency and services, it can also lead to potential risks and dangers if not properly regulated. Consequently, it's vital for organizations and policymakers to carefully consider the implications of AI adoption, weighing both potential benefits and risks, including unforeseen outcomes. By applying such balanced examination, we can approach AI democratization with due caution and adequate safeguards, allowing us to use the potential of this powerful technology responsibly and ethically.

To strike this balance, developing a conceptual framework that provides guidance through the unpredictability and complexities of AI evolution and adoption is essential. Reflecting on our journey from the broader implications to specific applications of AI, the article emphasizes the need for the development of an informed conceptual framework. Termed controlled AIsmosis, this conceptual framework provides valuable guidance in striking the delicate balance required in managing AI's integration into society. It serves as a compass to navigate the intricacies and challenges of AI adoption, offering a structured approach to ensure a mindful and controlled evolution of AI and possible AGI.

RESULTS

AIsmosis aligns with an expanding consensus: the need for a conceptual framework that forefronts human values, aspirations, and needs. This concept was recently reinforced by computer science Russell (2023) in his lecture at CITRIS research exchange and Berkeley artificial intelligence research lab (BAIR). Russell (2023) offered a critical assessment of the "standard model" of AI, which traditionally measures a machine's intelligence by its capacity to fulfill predefined objectives. Despite recognizing the model's significant role in AI evolution, Russell (2023) highlighted its considerable deficiencies, specifically in safeguarding personal agency and privacy. To remedy these shortcomings, Russell (2023) proposed a paradigm shift that elevates these human-centric elements above machine objectives. He redefined the standard model, stating: "Machines are beneficial to the extent that their actions can be expected to achieve our objectives."

Russell's (2023) conceptual realignment directly informs 'controlled AIsmosis,' a concept that underscores the necessity to preserve human agency and privacy. This approach is integral to a balanced, gradual integration of AI into societal systems. This results section will provide a more detailed exploration of this conceptual framework, advocating for a careful, tempered approach to AI integration. In response to the rapid advancements in AI, a systematic review of existing literature on AI and its societal interactions was conducted, and current happenings were observed. These activities led to the development of a new conceptual framework, 'controlled AIsmosis.'

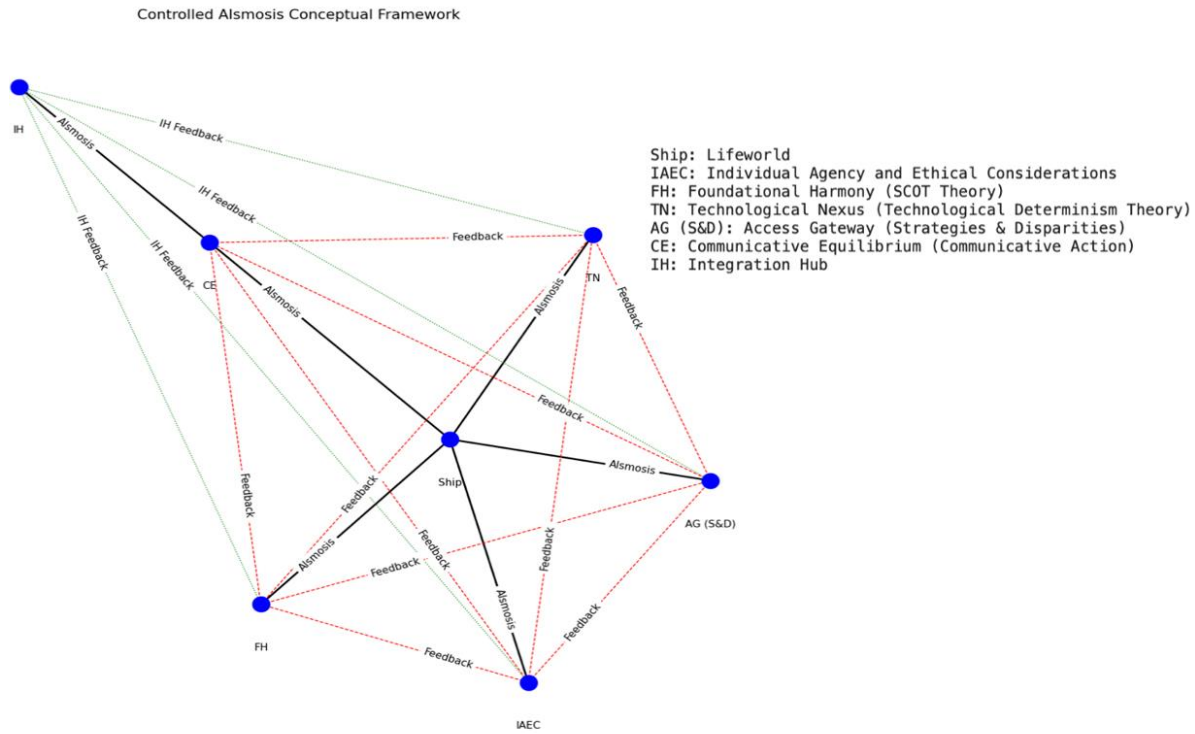

Figure 3. Controlled AIsmosis conceptual framework (Source: Author, using Networkx and Matplotlib libraries)

Controlled AIsmosis–Conceptual Framework

As depicted in **Figure 3**, this framework uses the metaphor of a ship navigating through an ocean to symbolize our collective journey through the swiftly changing landscape of technological innovations. The model consists of six balance tanks: foundational harmony (FH), technological nexus (TN), access gateway–strategies and disparities (AG–S&D), communicative equilibrium (CE), individual agency and ethical considerations (IAEC), and integration hub (IH).

Each tank symbolizes a key facet of socio-technological interaction, contributing to the stability of the lifeworld amidst unpredictable technological waves. The metaphor of the ship serves to convey two main points. Firstly, it represents the turbulent nature of modern technological evolution, symbolized by the ocean, resonating with technological determinism theory, which posits that technology significantly influences society. Understanding the driving forces and implications of AI progress, sourced from various disciplines, is thus critical. Secondly, the ship metaphor represents our lifeworld, not static but dynamic, portraying it as an entity continually reshaped by technological innovation, resonating with SCOT theory. The framework emphasizes technology's dual role as a societal change catalyst and its reflection/ byproduct.

Feedback Loops

Feedback loops symbolize the mutual and continual exchange of influences, adaptations, and effects between different tanks. Each tank in the framework sends signals or 'feedback' to all other tanks, reflecting its current state, changes, or imbalances. Simultaneously, each tank receives signals from all others, responding and adapting to these influences to maintain its equilibrium and harmony with the overall system. As a key component in the feedback system, the IH embodies the continuous integration of the processes occurring in other tanks, emphasizing the inherently interconnected and complex dynamics of socio-technological interactions. Society's inherent dynamism involves shifts influenced by cultural evolution, legal amendments, economic conditions, and demographic changes, among others. These shifts, independent of technology, are embodied within the feedback loops connecting the tanks. These loops signify reciprocal interactions between technology and society, embodying 'controlled AIsmosis'–a process of continual learning, adaptation, and evolution akin to AI networks. This mechanism ensures a dynamic equilibrium within the system.

Foundational Harmony Tank

Built on SCOT theory, FH tank symbolizes the evolving dialogue between societal norms and technology. It embodies the intertwined relationship of societal and cultural dynamics crucial to socio-technological development and acceptance. FH tank aligns societal norms, values, and expectations with technological progression, ensuring emerging technologies are accepted and integrated within society. However, this balance is fragile. Disruptions within FH tank can cause rifts between societal expectations and technological advancements, leading to potential resistance, ethical issues, and societal disruptions. Addressing such imbalances requires proactive measures, including stakeholder involvement, public consultation, ethical reviews, policy interventions, and inclusive design practices.

FH tank, interacting with other tanks, exhibits the bidirectional causality inherent in its foundational theory, demonstrating how societal norms shape and are shaped by technological advancements. This interplay, evident in FH tank's relationship with CE and AG–S&D tanks, underscores the reciprocal nature of these relationships. In case of imbalances, proactive strategies like democratic technogenesis–the co-creation of technology in a democratic, inclusive manner that respects societal values and norms–become necessary to foster harmonious societal and technological co-evolution. The fluidity of socio-technological interactions is highlighted by the adaptability of each tank in the framework.

Technological Nexus Tank

Informed by technological determinism theory, TN tank emphasizes the pivotal role of technology in shaping society. It serves as a hub for technological advancements and their implications, capturing the transformative potential of such developments in guiding socio-technological progression. TN tank serves as a navigator, pioneering and integrating technological advancements. It is instrumental in the discovery, development, and deployment of new technologies. However, its role is not dictatorial but rather one component of a balanced interaction among all tanks, each representing interconnected aspects of socio-technological interaction. Imbalances within TN tank can result in socio-technological discord, ethical dilemmas, and unequal access to technology. Thus, maintaining balance in TN tank requires effective societal dialogue, equitable diffusion strategies, and ethical considerations. This dynamic interplay ensures sensitivity to sociocultural, ethical, and communicative aspects of the journey. Therefore, TN tank is more accurately described as a collaborative navigator rather than an autocratic leader.

The complex dynamics of technological innovation and societal adaptation are captured within TN and FH tanks. They steer our collective journey through the complex landscape of technological innovation, maintaining balance in our societal 'ship' with continuous adjustments. This demonstrates the fluidity and dynamism of socio-technological interactions. TN tank substantially influences IAEC tank by introducing new ethical challenges alongside technological advancements. It also impacts CE tank by prompting changes in communication modes. Furthermore, TN tank molds access strategies and may influence disparities, affecting AG–S&D tank. It plays a central role in IH, reflecting its pivotal influence on the functionality of the whole system, as the emergent behaviors and complex interdependencies within society echo the dynamics of TN tank.

Communicative Equilibrium Tank

Rooted in the communicative action theory, CE tank principally encourages open, inclusive, and constructive dialogue. This forms a platform for a common understanding of technology, where different viewpoints converge, evolve, and negotiate shared meanings, thereby bolstering social dynamics within the system. It plays a critical role in mediating conversation between FH and TN tanks, guaranteeing that cultural and social influences integrate effectively with technological advancements and vice versa. CE tank acts as a stabilizing entity, mitigating potential misunderstandings and disagreements.

However, an imbalance, such as a communication breakdown or consensus failure, could unsettle the system and incite conflict. To regain equilibrium, restoring open, inclusive dialogue and re-establishing common understandings of technology and its societal implications is crucial. During intense technological transitions or heated ethical debates, CE tank becomes a key conflict resolution tool, clarifying misunderstandings and fostering balanced discussions to mitigate disputes. CE tank operates bidirectionally, sharing information about the effectiveness of communication and the content of dialogue with other tanks and incorporating their feedback to adapt its strategies. For instance, it could adjust its communication methods based on feedback from FH tank about sociocultural sensitivities or based on information from TN tank about emerging technologies.

External factors, such as socio-political context or legal and regulatory environment, can influence CE tank's communicative dynamics and shape its internal discourses. In response to socio-technological changes and advancements, CE tank continually adapts its communication strategies, embodying the dynamism inherent in the communication process. This discourse influences public perceptions of technology and its ethical implications, shaping the perspectives reflected within IAEC tank.

Access Gateway–Strategies & Diffusions

Inspired by the diffusion of innovations theory, AG–S&D tank concentrates on societal perception, acceptance, and adjustment to technological advancements. It gauges societal preparedness and aims to foster an understanding of the technology, creating a conducive environment for its acceptance. Aspects like perceived benefits, complexity, trialability, and observability shape socio-technological adoption rates and guide the strategies of AG–S&D tank. This tank collaborates with other tanks to formulate strategies. For instance, it could collaborate with CE tank to ensure effective communication of a new technology's benefits and implications to the public while coordinating with FH tank to ensure the technology aligns with societal norms and values.

However, imbalances within AG–S&D tank, appearing as inconsistencies in technology access and adoption, may lead to societal rifts and potential strife. To counter this, developing informed strategies that promote inclusive diffusion is critical, thus enhancing societal readiness for technological innovation and reducing disparities. Together, CE and AG–S&D tanks play crucial roles in facilitating effective technology communication and managing its fair distribution across society. By nurturing mutual understanding and encouraging inclusivity, they play a significant part in navigating the intricate landscape of technological progression, ensuring the ship stays balanced and on course.

Integration Hub

Within the conceptual framework, IH tank plays a crucial role, acting as the central point for merging the outputs, inputs, and feedback loops of all other tanks. This tank embodies the repetitive integration of the processes happening in the other tanks, emphasizing the inherently interconnected and complex dynamics of AIsmosis. The IH has the responsibility of integrating and synchronizing all tanks.

Drawing parallels with Foucault's (1977) theory of power, IH tank operates akin to a distributive node of power in a network. It ensures balance across different elements of the system, enhancing the democratic dialogue and interactions among all tanks. This resemblance underscores IH's crucial role in preserving the equilibrium of the overall system, reflecting the diffusion and balance in power dynamics within the process of controlled AIsmosis.

This role is essential to ensure all tanks operate in accord, with their outputs systematically factored into the system's overall structure and direction. This harmonious interaction–pas de deux–aligns with the feedback loops with other tanks; IH receives, processes, and reciprocates feedback from all tanks. It guarantees that shifts and changes within any tank are appropriately accounted for in the system. By harmonizing and integrating the functions and effects of the other tanks, IH contributes to the overall balance of the ship.

IH and TN should operate in close collaboration. TN tank, which can metaphorically be described as the captain, embodies technological innovation and progression, steering the direction of advancement within the system. However, the term captain is metaphorical and does not imply that TN holds authority over other tanks. Similarly, IH tank represents the 'steering mechanism,' playing a crucial role in synthesizing information from different tanks and ensuring coordinated responses. IH acts as a facilitator for effective system coordination and communication.

Individual Agency and Ethical Considerations

IAEC tank encapsulates societal perspectives and attitudes towards technology, influenced by societal norms and ethical considerations. It maintains the central focus on human agency within HAII. Its role is crucial in reflecting societal voices, ethical debates, and concerns regarding technology within the system. Through its feedback loops with other tanks, IAEC emphasizes the ethical dimensions of socio-technological interaction. It informs TN about the ethical implications of technology design and deployment, guides societal norms and behaviors through FH, and shapes the dialogue about technology in CE tank.

IAEC tank stands as a symbol of our steadfast commitment to preserving human dignity, autonomy, and rights in the face of technological advancement. The influence of IAEC tank reaches AG–S&D tank, guiding the management of technology access and addressing disparities. This underscores the ethical considerations in technology access and the crucial importance of providing equitable opportunities for technology adoption. IAEC tank assumes a vital function in shaping technological legislation, thereby influencing system dynamics beyond the immediate scope of technological implications. This brings to light the indispensable interplay between ethics and law in overseeing the use and development of technology. With time, IAEC tank adapts and evolves alongside societal shifts in viewpoints on technology and ethics. This adaptive capacity underscores the fluidity of ethics in technology, necessitating persistent scrutiny and modification.

Alterations in TN tank can significantly affect IAEC, as swift technological progress may pose challenges to or even redefine our ethical considerations and personal agency. This emphasizes the pressing need for ethical deliberation amidst the fast-paced progression of technological innovation. The interaction between IAEC and IH tank is vital, considering the IH's function mirrors the system's intricate interdependencies and emerging behavior. This interaction accentuates the fundamental role of ethics and agency in preserving system cohesion and steering emerging behavior. IAEC tank encapsulates human viewpoints vis-à-vis AI and prospective AGI, contemplating our capacity to mold, accept, reject, or adapt technology. It's primarily tasked with safeguarding human agency and ethical considerations, examining the repercussions of technology on human dignity, privacy, autonomy, and social equality.

DISCUSSION

In the conceptual framework, technological unpredictability is mitigated through the concentrated actions of various tanks. IAEC tank develops ethical guidelines and influences legislation, while TN tank steers technological progression strategically. AG tank works towards equitable technological access and influences legislative actions, and CE tank stimulates dialogue for strategic planning and legislative guidance. These actions create a dynamic feedback loop, continually adapting to shifting circumstances and controlling the process of AIsmosis.

Imbalances can spark systemic ripple effects, emphasizing the importance of continuous monitoring and adaptation. For instance, the widespread adoption of AI in TN tank reshapes societal norms and values in FH tank, redefining views on privacy and human-machine interactions. This is akin to the findings of McAfee and Brynjolfsson (2012), where imbalances in data-driven decision-making led to systemic impacts across industries. This interplay showcases the dynamic relationship between technology and society, reflected in the fluid interactions between the tanks.

The bidirectional influence between societal receptiveness (FH) and innovation diffusion (AG–S&D) is another example of this feedback loop. CE tank plays a crucial role in mediating societal dialogues around technology, influencing societal norms and values (FH), and fostering consensus. Forester (1999) provided empirical support for this in his discussion of TCA, demonstrating its ability to facilitate more democratic and inclusive dialogues. The norms established in FH tank guide individual and collective actions in IAEC tank, which can shape societal norms, further reinforcing the feedback loop. IH maintains system cohesion and mirrors its complexity, demonstrating intricate interdependencies and emergent behavior within framework.

The presence of feedback loops ensures the system's continuous adaptation to technological advancement and societal changes, impacting the overall functionality of the system significantly. Thus, the conceptual framework is a testament to the fluid, evolving relationship between society and technology, manifested in its dynamic interactions, feedback loops, and the phenomenon of AIsmosis.

As an example, regulation of product deployment is a multifaceted task in managing technological advancements. AG–S&D tank, informed by the diffusion of innovations theory, is critical in developing strategies for safe and effective technology adoption, akin to government agencies like FDA that regulate products for public safety. CE tank facilitates discourse around the technology, ensuring transparency and shared understanding. IAEC tank, reflecting human agency theory, address potential biases in AI systems, developing ethical guidelines that respect human values and autonomy. IH, in its overarching role, integrates inputs from all tanks for balanced AI development, deployment, and regulation.

As for creating completely unbiased AI, it's arguably unachievable if AI learns from human data. However, implementing measures to recognize, understand, mitigate, and monitor biases can prevent unjust outcomes or harmful stereotypes reinforcement. AG–S&D tank monitors AI dissemination and usage, ensuring ethical applications and bias mitigation. CE tank facilitates open dialogue about AI biases, promoting transparency and collective action. IAEC tank contributes to ethical AI guidelines, incorporating fairness, transparency, and accountability principles to guard against harmful biases. The importance of these measures can be seen in the work of Selbst et al. (2019), which emphasized the need to consider both the technical and social aspects of AI when addressing biases.

The tasks we face in managing technological advancements are intricate and multifaceted, necessitating collaborative efforts across different tanks in the controlled AIsmosis conceptual framework. This underscores the interconnected nature of these issues and the value of a comprehensive approach. Any imbalance within these tanks can trigger a domino effect, impacting the system's overall stability. Therefore, continuous monitoring and adjustments through feedback loops are essential for a balanced approach to technological evolution, ensuring stability in our ship.

CONCLUSIONS

In an increasingly intertwined global society, the pas de deux of AIsmosis echoes far beyond individual lifeworlds. This interaction can ripple across many realities, creating waves of turbulence in the vast ocean of technological innovation. As partners in this intricate dance, maintaining a balanced AIsmosis is not a prerogative but a necessity. However, disrupting the balance of this dance can lead to unforeseen consequences. It can widen the digital divide, spark socio-cultural turmoil, and undermine the fundamental principles of fairness, inclusivity, and ethical conduct that technology is intended to uphold. In contrast, achieving a harmonious controlled AIsmosis can guide us toward a future, where technology gracefully enhances our collective existence rather than disrupting it.

Sudden, careless changes can lead to a disruptive imbalance, like cells experiencing an osmotic shock. This unmeasured approach to AI integration might drive the adoption of technologies that do not align with societal values or foster policies ill-equipped to address new ethical, legal, or technical challenges. It is not a hurried race. AIsmosis is a continuous, progressive journey that requires careful choreography to preserve human interests and well-being with human agency. In any ballet, a sudden shift can cause the dancer to fall, but in the context of AIsmosis, it's the stability of society that hangs in the balance. Here, it is crucial to strike a balance between exploiting AI's potential and ensuring a fair transition for everyone. Drawing upon nautical wisdom, the saying 'a ship sinks at the port' serves as a potent warning. Just as an unbalanced load and misadjusted balance tanks can lead a ship to capsize, a lack of balance in AIsmosis can disrupt societal stability. We must ensure a careful and considered approach to avoid such a fate.

Controlled AIsmosis entails not only successfully navigating technological innovation but also harmonizing our physical and digital realities. Our dance–pas de deux–with AI is not just about survival but about progress, enlightenment, and the common good.

Funding: The author received no financial support for the research and/or authorship of this article.

Ethics declaration: The author declared that ethical approval was not applicable since this is a review article that does not involve human subjects or animal experiments.

Declaration of interest: The author declares no competing interest.

Data availability: Data generated or analyzed during this study are available from the author on request.

REFERENCES

Adorno, T. W. (1991). *The culture industry: Selected essays on mass culture*. Routledge.

Baker, R. S., Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education, 32*, 1052-1092. https://doi.org/10.1007/s40593-021-00285-9

BBC News. (2023). Drake and the weekend song 'pulled' by Universal Music. *BBC*. https://www.bbc.com/news/entertainment-arts-65309313

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code*. John Wiley & Sons. https://doi.org/10.1093/sf/soz162

Bickmore, T. W., & Picard, R. W. (2004). Towards caring machines. In *Proceedings of the 2004 Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/985921.986097

Bonhoeffer, K. (1904). Der Korsakowsche Symptomenkoplex in seinen Beziehungen zu den verschiedenen Krankheitsformen [Korsakoff's complex of symptoms in relation to the various forms of the disease]. *Allgemeine Zeitung Psychiatrie* [*General Newspaper Psychiatry*], *61*, 744e752.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023, March). Sparks of Artificial General Intelligence: Early experiments with GPT-4. https://www.microsoft.com/en-us/research/publication/sparks-of-artificial-general-intelligence-early-experiments-with-gpt-4/

Büchi, M., Festic, N., & Latzer, M. (2022). The chilling effects of digital dataveillance: A theoretical model and an empirical research agenda. *Big Data & Society, 9*(1). https://doi.org/10.1177/20539517211065368

Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Allas, T., Dahlström, P., Henke, N., & Trench, M. (2017). *The rise of AI: What do its advances mean for business strategy?* McKinsey Quarterly.

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

Capital.com. (2023). Top AI stocks: Best artificial intelligence companies to invest in. *Capital.com*. https://capital.com/top-ai-stocks-companies-to-invest

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165-1188. https://doi.org/10.2307/41703503

Council of Europe. (2022). Artificial intelligence and education: A critical view through the lens of human rights, democracy, and the rule of law. *Council of Europe*. https://rm.coe.int/artificial-intelligence-and-education-a-critical-view-through-the-lens/1680b7c9b8

Crawford, K. (2015). Can an algorithm be agonistic? Ten scenes from life in calculated publics. *Science, Technology, & Human Values, 41*(1), 77-92. https://doi.org/10.1177/0162243915589635

Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

Financial Times. (2023). We must slow down the race to God-like AI. *Financial Times*. https://www.ft.com/content/03895dc4-a3b7-481e-95cc-336a524f2ac2

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. *JMIR Mental Health, 4*(2), e19. https://doi.org/10.2196/mental.7785

Floridi, L. (2014). *The fourth revolution: How the infosphere is reshaping human reality*. Oxford University Press.

Foot, P. (1967). Abortion and the doctrine of double effect. *Oxford Review, 5*, 5-15.

Forbes. (2020). Why AI is transforming the banking industry. *Forbes*. https://www.forbes.com/sites/cognitiveworld/2020/04/05/why-ai-is-transforming-the-banking-industry/

Forester, J. (1999). *The deliberative practitioner: Encouraging participatory planning processes*. MIT Press.

Fortune. (2023a). Scammers are using voice-cloning A.I. tools to sound like victims' relatives in desperate need of financial help. It's working. *Fortune*. https://fortune.com/2023/03/05/scammers-ai-voice-cloning-tricking-victims-sound-like-relatives-needing-money/

Fortune. (2023b). *Snapchat influencer launches CarynAI, a virtual girlfriend bot using OpenAI's GPT-4*. https://fortune.com/2023/05/09/snapchat-influencer-launches-carynai-virtual-girlfriend-bot-openai-gpt4/

Foucault, M. (1977). *Discipline and punish: The birth of the prison*. Vintage.

Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Pitman.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerization? *Technological Forecasting and Social Change, 114*, 254-280. https://doi.org/10.1016/j.techfore.2016.08.019

Fuchs, C. (2013). *Social media: A critical introduction*. SAGE. https://doi.org/10.4135/9781446270066

Gerbner, G., Gross, L., Morgan, M., & Signorielli, N. (1980). The "mainstreaming" of America: Violence profile no. 11. *Journal of Communication, 30*(3), 10-29. https://doi.org/10.1111/j.1460-2466.1980.tb01987.x

Gerkin, R. C., & Wiltschko, A. B. (2022). Digitizing smell: Using molecular maps to understand odor. *Google AI Blog*. https://ai.googleblog.com/2022/09/digitizing-smell-using-molecular-maps.html

Goldhaber, M. H. (1997). *Attention shoppers! Wired*. https://www.wired.com/1997/12/es-attention/

Gujral, R. (2023). The implications of emotion AI in our lives [YouTube video]. https://www.youtube.com/watch?v=jHoCfJlMHdM

Habermas, J. (1984). *The theory of communicative action: Reason and the rationalization of society*. Beacon Press.

Habermas, J. (1987). *The theory of communicative action: Lifeworld and system, a critique of functionalist reason*. Beacon Press.

Haraway, D. J. (1991). *Simians, cyborgs, and women: The reinvention of nature*. Routledge.

Harris, T., & Raskin, A. (2023). The A.I. dilemma. *YouTube*. https://www.youtube.com/watch?v=xoVJKj8lcNQ

Hatfield, E., Cacioppo, J. L. & Rapson, R. L. (1993). Emotional contagion. *Current Directions in Psychological Sciences, 2*, 96-99. https://doi.org/10.1111/1467-8721.ep10770953

Hofstadter, D. R. (2008). *I am a strange loop*. Basic Books.

Huang, J., Gu, S. S., Hou, L., Wu, Y., Wang, X., Yu, H., & Han, J. (2022). Large language models can self-improve. *arXiv*, 11610. https://doi.org/10.48550/arXiv.2210.11610

Innis, H. (1950). *Empire and communications*. Oxford University Press.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389-399. https://doi.org/10.1038/s42256-019-0088-2

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255-260. https://doi.org/10.1126/science.aaa8415

Layton, E. (1971). Mirror-image twins: The communities of science and technology in 19th-century America. *Technology and Culture, 12*(4), 562-580. https://doi.org/10.2307/3102571

McAfee, A., & Brynjolfsson, E. (2012). Big data: The management revolution. *Harvard Business Review, 90*(10), 60–68.

McKinsey Global Institute. (2017). *Harnessing automation for a future that works*. https://www.mckinsey.com/featured-insights/digital-disruption/harnessing-automation-for-a-future-that-works

Metz, C. (2023, May 16). *Microsoft says new A.I. shows signs of human reasoning*. The New York Times. https://www.nytimes.com/2023/05/16/technology/microsoft-ai-human-reasoning.html

Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society, 3*(2). https://doi.org/10.1177/2053951716679679

Moore, S. (2023). *How bias impacts AI* [Paper presentation]. Rise of AI Conference 2023.

MSNBC. (2023). Hip-hop artists are leading the AI discussion. Are you listening? *MSNBC*. https://www.msnbc.com/the-reidout/reidout-blog/hip-hop-timbaland-ai-young-guru-rcna85875

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81-103. https://doi.org/10.1111/0022-4537.00153

New York Post. (2023). *AI-generated photo of fake Pentagon explosion sparks brief stock selloff*. https://nypost.com/2023/05/22/ai-generated-photo-of-fake-pentagon-explosion-sparks-brief-stock-selloff/

Noelle-Neumann, E. (1974). *The spiral of silence: Public opinion–Our social skin*. University of Chicago Press. https://doi.org/10.1111/j.1460-2466.1974.tb00367.x

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447-453. https://doi.org/10.1126/science.aax2342

OpenAI. (2023). GPT-4 technical report. *OpenAI*. https://cdn.openai.com/papers/gpt-4.pdf

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin.

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press. https://doi.org/10.4159/harvard.9780674736061

Piaget, J. (1929). *The child's conception of the world*. Brace Jovanovich.

Pinch, T. J., & Bijker, W. E. (1984). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science, 14*(3), 399-441. https://doi.org/10.1177/030631284014003004

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1*(4), 515-526. https://doi.org/10.1017/S0140525X00076512

Rogers, E. M. (1962). *Diffusion of innovations*. Free Press of Glencoe.

Russell, S. (2023). *How not to destroy the world with AI* [YouTube video]. https://www.youtube.com/watch?v=ISkAkiAkK7A&t=2865s

Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and abstraction in sociotechnical systems. In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 59-68). ACM. https://doi.org/10.1145/3287560.3287598

Selman, R. L. (1980). *The growth of interpersonal understanding: Developmental and clinical analyses*. Academic Press.

Simon, H. A. (1971). Designing organizations for an information-rich world. In M. Greenberger (Ed.), *Computers, communications, and the public interest* (pp. 37-72). The Johns Hopkins Press.

Smith, A., & Anderson, M. (2018). Automation in everyday life. *Pew Research Center*. https://www.pewresearch.org/internet/2017/10/04/automation-in-everyday-life/

Starke, C., Baleis, J., Keller, B., & Marcinkowski, F.(2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data & Society, 9*(2), 1-16. https://doi.org/10.1177/20539517221115189

Sunstein, C. R. (2001). *Republic.com*. Princeton University Press.

Taddeo, M., & Floridi, L. (2018). Regulate artificial intelligence to avert cyber arms race. *Nature*. https://www.nature.com/articles/d41586-018-04602-6

The Verge. (2023). Grimes says anyone can use her voice for AI-generated songs. *The Verge*. https://www.theverge.com/2023/4/24/23695746/grimes-ai-music-profit-sharing-copyright-ip

Tian, Z., Cui, L., Liang, J., & Yu, S. (2022). A comprehensive survey on poisoning attacks and countermeasures in machine learning. *ACM Computing Surveys, 55*(8), 166. https://doi.org/10.1145/3551636

Turkle, S. (2015). *Reclaiming conversation: The power of talk in a digital age*. Penguin Books.

van Dijk, J. A. G. M. (2020). The digital divide: Aspects of inequalities resulting from network participation. *International Journal of Communication, 14*, 3915-3931.

Veblen, T. (1899). *The theory of the leisure class: An economic study in the evolution of institution*. Macmillan.

Veblen, T. (1904). *The theory of business enterprise*. Charles Scribners.

Veblen, T. (1921). *The engineers & the price system*. Batoche Book.

Virvou, M. (2022). The emerging era of human-AI interaction: Keynote address. In *Proceedings of the 13th International Conference on Information, Intelligence, Systems & Applications* (pp. 1-10). https://doi.org/10.1109/IISA56318.2022.9904422

Warschauer, M., & Matuchniak, T. (2010). New technology and digital worlds: Analyzing evidence of equity in access, use, and outcomes. *Review of Research in Education, 34*(1), 179-225. https://doi.org/10.3102/0091732X09349791

Webb, A. (2023). *Amy Webb launches 2023 emerging tech trend report* [YouTube video]. https://www.youtube.com/watch?v=vMUpzxZB3-Y

Winner, L. (1986). *The whale and the reactor: A search for limits in an age of high technology*. University of Chicago Press.

World Economic Forum. (2018). *How artificial intelligence is shaking up the job market*. https://www.weforum.org/agenda/2018/09/artificial-intelligence-shaking-up-job-market/

World Economic Forum. (2022). *AI's trolley problem debate can lead us to surprising conclusions*. https://www.weforum.org/agenda/2022/05/ai-s-trolley-problem-debate-can-lead-us-to-surprising-conclusions/

Zhang, C., Patras, P., & Haddadi, H. (2018). Deep learning in mobile and wireless networking: A survey. *IEEE Communications Surveys & Tutorials, 21*, 2224-2287. https://doi.org/10.1109/COMST.2019.2904897

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. Public Affairs.