Metaverse (2023) *Volume 4 Issue 2*
doi: 10.54517/m.v4i2.2276

Asia Pacific
Academy of Science Pte. Ltd.

REVIEW ARTICLE

Individuality and the collective in AI agents: Explorations of shared consciousness and digital homunculi in the metaverse for cultural heritage

James Hutson*, Jeremiah Ratican

Jeremiah Ratican, Department of Art, Media, and Production, Lindenwood University, Saint Charles, MO 63301, USA

*** Corresponding author:** James Hutson, jhutson@lindenwood.edu

ABSTRACT

The confluence of extended reality (XR) technologies, including augmented and virtual reality, with large language models (LLM) marks a significant advancement in the field of digital humanities, opening uncharted avenues for the representation of cultural heritage within the burgeoning metaverse. This paper undertakes an examination of the potentialities and intricacies of such a convergence, focusing particularly on the creation of digital homunculi or changelings. These virtual beings, remarkable for their sentience and individuality, are also part of a collective consciousness, a notion explored through a thematic comparison in science fiction with the Borg and the Changelings in the Star Trek universe. Such a comparison offers a metaphorical framework for discussing complex phenomena such as shared consciousness and individuality, illuminating their bearing on perceptions of self and awareness. Further, the paper considers the ethical implications of these concepts, including potential loss of individuality and the challenges inherent to accurate representation of historical figures and cultures. The latter necessitates collaboration with cultural experts, underscoring the intersectionality of technological innovation and cultural sensitivity. Ultimately, this chapter contributes to a deeper understanding of the technical aspects of integrating large language models with immersive technologies and situates these developments within a nuanced cultural and ethical discourse. By offering a comprehensive overview and proposing clear recommendations, the paper lays the groundwork for future research and development in the application of these technologies within the unique context of cultural heritage representation in the metaverse.

Keywords: digital homunculi; changelings; collective consciousness; large language models; cultural heritage

1. Introduction

In the confluence of art and technology, the burgeoning potential of extended reality (XR), including augmented reality (AR), mixed reality (MR), and virtual reality (VR), melding with large language models (LLMs) to represent cultural heritage within the metaverse is clearly evident[1,2]. From November 2022 onward, a remarkable transformation has been observed in the realm of generative artificial intelligence (AI)[3] and large language models (LLMs), with novel opportunities being unveiled and stimulating debates initiated about the forthcoming direction of AI. AI constructs such as OpenAI's ChatGPT 3.5 and 4 have demonstrated an impressive capacity to produce textual content that echoes human language, while simultaneously managing intricate conversational engagements[4]. As progress in the domain continues, the possible implementations of

ARTICLE INFO
Received: 9 August 2023 | Accepted: 21 September 2023 | Available online: 8 November 2023

CITATION
Hutson J, Ratican J. Individuality and the collective in AI agents: Explorations of shared consciousness and digital homunculi in the metaverse for cultural heritage. *Metaverse* 2023; 4(2): 2276. doi: 10.54517/m.v4i2.2276

these models are undergoing exploration across various fields such as education, healthcare, and assisted living; this exploration mirrors a broader scholarly investigation that has arrived at a pivotal juncture[5–7].

Within this transforming environment, the examination and replication of human emotions in AI constructs have emerged as intriguing areas of academic inquiry[8]. Though machine intelligence (ML) and human emotions remain inherently different, the capacity of LLMs to simulate and assume different "personalities" has shown promising avenues for exploration into artificial emotional aspects[3]. This advancement sets the stage for the development of AI systems that are proficient in comprehending and generating appropriate emotional reactions, thereby enhancing user interaction and fostering more profound engagements[9]. As well, such an amalgamation has led scholars and technologists to contemplate the creation of digital homunculi, or changelings: virtual entities endowed with distinct consciousness, yet embedded within a collective cognitive fabric[10]. Such entities challenge pre-existing conceptions of individuality, consciousness, and the very essence of self-awareness[11].

Historically, the interplay between technology and art has engendered novel paradigms and vistas for representation and interaction[12]. As the boundaries of technological capabilities expand, so too does the horizon for artistic and cultural exploration. Within this expansive framework emerges the concept of digital homunculi. Drawing from historical, literary, and cinematic contexts, the term "changeling" alludes to beings capable of transformation or substitution—an apt metaphor for entities that oscillate between individual and collective consciousness. A profound intrigue accompanies the juxtaposition of individual consciousness with the notion of a shared cognitive realm. One might draw parallels to the exploration of outer space: just as the vastness of the cosmos prompts humanity to grapple with its place within it, the vastness of collective digital consciousness forces an introspection of the self's position in relation to the collective[13].

LLMs present a promising frontier in the realm of cultural heritage experiences, opening pathways to engage users with history and culture in nuanced and dynamic ways. Leveraging the capabilities of LLMs can foster individualized and collective explorations, providing a spectrum of experiences ranging from intimate dialogues with historically significant figures to broad, immersive interactions within ancient cities or civilizations. The intricate design of these models enables the simulation of diverse consciousness types and personalities, closely aligned with historical and cultural contexts, thereby enriching educational and recreational endeavors. This paper embarks on a systematic investigation into the evolution of simulated consciousness and personality through AI, focusing on its applicability to the multifaceted world of cultural heritage experiences. Through a blend of theoretical analysis and practical insights, the exploration aims to unravel the potential of LLMs to create bridges between the past and present, offering a vivid, accessible, and authentic journey through the tapestry of human history.

2. Literature review

The burgeoning field of technological advancement presents a multifaceted landscape, warranting an in-depth examination of its various components and their intersections with society. The following seeks to explore three seminal dimensions of contemporary technology that have manifested profound implications across diverse domains. The first section delves into the evolution and application of immersive realities in the preservation, interpretation, and dissemination of cultural heritage. It elucidates the manner in which these technologies have revolutionized the way societies engage with their historical and cultural legacies. The subsequent section offers an analysis of the ascendance of artificial intelligence (AI), specifically focusing on large language models. These computational systems have ushered in a new era of human-computer interaction, transforming communication, information retrieval, and decision-making processes. The final section embarks on a philosophical exploration of digital entities and their impact on human consciousness and probes the emergent phenomena of digital representations of self and the collective, examining how these virtual constructs shape human understanding, identity, and social dynamics. Collectively, these interrelated facets present a compelling tapestry of technological innovation, bridging the traditional and the novel, the material and the virtual, the individual and the communal.

2.1. Immersive technologies in cultural heritage

Historically, XR technologies have been lauded for their potential in enhancing user experiences in various fields, from gaming to education. The application of immersive realities, including virtual reality (VR), within the context of cultural heritage has marked a transformative era for the preservation and accessibility of historical sites and artifacts. This evolution began around 2001, with the digitization of cultural heritage sites, both extant and ancient, although initial use was largely confined to researchers with specialized knowledge. Techniques such as CAVE technology allowed institutions such as the Foundation of the Hellenic World (FHW) to digitally reconstruct ancient cities, exemplified by the rendering of Miletus, a historically significant settlement in Asia Minor[14]. The promise of this technology for museums was readily acknowledged; Roussou[15])expounded upon the potential for enhancing physical exhibits with digital components for an "edutainment" experience, merging education and entertainment.

Between 2001 and 2010, various museums, including The Museum of Pure Form and The Virtual Museum of Sculpture, integrated immersive realities into their spaces. Unlike applications in medical or scientific fields that required substantial training, these virtual exhibits were constructed for the lay public with minimal experience operating complex hardware. Drawing on earlier works by Salzman, Dede, Loftin, and Chen, museums developed shorter experiences to facilitate visitor flow. Initiatives such as The Exploratorium and The CREATE project, an EU-funded endeavor, quickly transitioned to entirely virtual realms, fostering the reconstruction of archeological sites and the digitization of entire collections for immersive viewing[16]. The emphasis on user-friendly interfaces and limited engagement duration paved the way for the contemporary design of virtual learning environments (VLEs).

In the subsequent decade, the proliferation of virtual learning environments for conveying cultural heritage content has been notable. Innovations such as Google's Arts & Culture in 2011, complemented with Google Cardboard in 2014, further democratized access to virtual museum experiences[17,18]. The capability to virtually tour real or computer-generated museums became prevalent, with instances such as the National Archeological Museum of Marche in Ancona, Gyeongju VR Museum, South Korea, and the Rijksmuseum, Amsterdam. Notably, in 2021, the Louvre digitized a vast portion of its collection, consisting of over 480,000 pieces, available through its online platform. Concurrently, educational initiatives were developed using game engines like Unreal Engine and Unity, enabling virtual tours for students, such as those designed at the Universidad Nacional de San Agustin de Arequipa in Peru[19]. Between 2017 and 2020, the full virtual tours of UNESCO World Heritage Sites manifested, underscoring the growing intersection between technology and cultural heritage, and extending the reach of these invaluable resources to a global audience.

Furthermore, the applications of artificial intelligence (AI) in the field of historical recreation and exploration have transcended mere photographic transformations, as demonstrated in several pioneering projects. A case in point is the *Virtual Angkor* (https://www.virtualangkor.com/) initiative spearheaded by Monash University in Melbourne, Australia. Utilizing immersive VR and 360-degree videos, this project propels visitors into a lively 13th-century rendition of Angkor. Within *Virtual Angkor*, a carefully constructed ensemble of 3D models and animated figures echoes the authentic historical milieu, providing an insightful and engaging glimpse into the ancient city's cultural intricacy and societal operations. Integrating archaeological discoveries and textual references, the endeavor enables comprehensive examination of diverse facets of the civilization, such as trade, diplomacy, political structures, and the unique sense of place marking the medieval establishment. The interplay of AI-infused extended reality (XR) technologies with historical information establishes the project as a commendable template for insightful educational experiences linking contemporary spectators to antiquity[20].

An additional illustrative instance of the interactive worldbuilding in historical XR experiences is found in the *Witness to Revolution* (2021) project (https://www.wttrgame.com/), a joint effort by the University of Wisconsin-Stout and Carleton College. Through an immersive VR game, participants are immersed in Colonial Boston during the notorious 1770 “Boston Massacre.” The game functions as an audacious attempt to portray a factual recreation of the era while probing the extensive influence of misinformation on historical understanding. Interaction with evidence, witnesses, and flashbacks, coupled with the exploration of historical artifacts such as Paul Revere’s famous print, “The Bloody Massacre,” enriches the experience, allowing for individual interpretations of the consequential events of 5 March, 1770. Witness to Revolution serves as evidence of the potential of AI-empowered XR experiences in nurturing historical comprehension and analytical reasoning.

In 2023, an interdisciplinary team from USC Dornsife College of Letters, Arts, and Sciences is leveraging the possibilities of XR technologies to amplify the accessibility of historical objects and locations for the broader public[21]. Notable among their endeavors are the Chinatown AR Project, aimed at enhancing the visitor experience at Union Station through augmented historical portrayals, and a VR initiative focused on recreating the Stanza della Segnatura in the Vatican museums during the era of Pope Julius II. These initiatives symbolize the infinite opportunities that AI-enabled XR experiences offer in revitalizing history and engaging the public with the historical past. However, the constraints of these projects often lie in their limited interactive functions, relegating the viewer to the role of a passive observer. Recent breakthroughs in generative AI have, nonetheless, unlocked hitherto unavailable prospects for more dynamic and participative interactions with historical figures.

2.2. Rise of large language models and AI agents

The burgeoning field of Large Language Models (LLMs) presents the enthralling possibility of utilizing AI agents endowed with the capacity to assume diverse personas and emotional conditions, thereby contributing to research undertakings. A marked interest in the domain of emotionally perceptive AI, highlighting the cardinal role of emotions in shaping intelligent conduct and deliberative processes, has been noticed[22,23]. Such a research direction underscores the importance of incorporating emotions into the sphere of artificial cognitive functioning and accords recognition to the imperativeness of artificial emotional intelligence for the enrichment of societal interactions[24]. Indeed, the extant scholarly contributions have extensively investigated multifaceted subjects such as the engineering of machines capable of manifesting empathy, the design of AI-centric emotive communication systems, and the complexities and ramifications of synthetic emotions in realms such as healthcare and digital commerce[25–27]. The academic pursuit in these areas has laid a foundational pathway for the integration of emotional aspects into AI constructs, thereby facilitating the advent of interactions with humans that are both subtler and more precisely attuned to context.

The sphere of research devoted to emotionally intelligent AI encompasses an eclectic and extensive range of dimensions, fields, and thematic inquiries, laying emphasis on the quintessential role that emotions play in discerning behavior, judgment-making, and the concomitant prospects and intricacies tied to the assimilation of emotions within AI constructs. Scrutiny of the functions that emotions serve in intelligent conduct and decision-making procedures illuminates the cardinality of emotions in shaping human cognition and bolsters their wider acknowledgment within the disciplines of engineering and computer science[28]. Likewise, contemplation of the ramifications of emotions in fortifying artificial cognitive processing lends further credence to this vital relationship[29].

Theoretical paradigms that endorse socially emotive AI stand as testament to this burgeoning field, as do explorations centered around the engineering of machines imbued with the capability to convey empathy and emotional acumen[30,31]. Such investigations accentuate the importance of artificial emotional intelligence in facilitating interactions that are both socially engaging and emotionally resonant, thus forming the bedrock for systems characterized by emotional intelligence.

Additionally, the corpus of research focusing on AI-powered emotive communication systems and the strides made in the arena of synthetic emotional intelligence reveals the burgeoning possibilities and advancements within affective computing[32,33]. Ranging from the development of chatbots possessing emotional intelligence to the implementation of AI in practical, real-world contexts, these instances not only display the multifaceted nature of the field but also highlight the tangible applications that are fast becoming an integral part of technological landscapes. Consequently, the growth and exploration within this domain manifest as a critical juncture in the evolution of AI, one that continues to shape and redefine the interface between human emotion and artificial intelligence.

Debates and dialogues that center on the impediments and consequences of synthesizing artificial emotions unveil the ethical dimensions and pragmatic reflections that are inextricable from the deployment of emotionally intelligent AI[34]. Concurrently, the fabrication of biological-inspired social and emotional cognitive apparatuses foregrounds similar ethical quandaries and practical contemplations[35]. The scholarly discourse that orbits the transition from AI possessing mere emotional resonance to AI endowed with cognitive faculties[36], together with the insistence on the formulation of vigorous emotion models[37], amplifies fundamental components intrinsic to the field and emphasizes the requisite conditions for fashioning emotion models geared towards tangible applications within the realm of affective computing.

Illustrations of proposals aimed at assimilating emotion recognition tools within medical contexts[38], along with deliberations on the self-reliant decision-making proficiencies of emotional AI systems[39], delineate the prospective influence of emotionally aware AI within the healthcare sector. These reflections also invoke the ethical reflections connected to autonomous AI-driven decision-making, thereby presenting a multi-layered analysis of this complex domain. In the ultimate analysis, the addressal of legal and ethical obstacles—such as the recognition of fundamental human liberties like the freedom of thought—calls attention to the broader societal implications and underscores the imperative of rigorous scrutiny and circumspection in the design and deployment of emotionally intelligent AI systems[40].

An assortment of scholarly research has delved into the application of generative AI agents within various contexts, such as video games, decision-making frameworks, gameplay, and player conduct. Exploration by Naddaf[41] into reinforcement learning-based methodologies facilitated the education of AI agents in game-playing, a concept further expanded by Liu et al.[42], who utilized agents like the Random Mutation Hill-Climber and the Multi-Armed Bandit Random Mutation Hill-Climber to cultivate game versions with notable skill-depth. Concurrently, Holmgard et al.[43] postulated that artificial agents might serve as psychometrically accurate abstract simulations of the internal decision-making procedures of human players. Similarly, Barthet et al.[44] engineered generative personas that reflected human-like behavior, manifesting play styles and reactions that faithfully paralleled the human models they sought to emulate.

Additional research has been concentrated on the scrutiny and evaluation of video games through the lens of AI agents. Investigations by Ariyurek et al.[45] into synthetic and human-like agents within automated video game testing yielded promising results, with AI agents rivaling human capabilities in bug detection. However, Fathi and Palhang[46] spotlighted a significant obstacle—the uniformity in agents' behavior, leading to predictability. In a historical context, works by Nareyek[47], Tan and Nareyek[48], and Miikkulainen et al.[49] elucidated the potential and significance of AI methodologies in contemporary computer games, while Fernández et al.[50] expounded on the role of AI in formulating the behavior of automated player characters or bots.

In a more recent context, generative AI constructs such as ChatGPT have unveiled novel avenues in transforming practices, pedagogical processes, and research across multifarious scientific and medical disciplines. Morris[51] documented the perspectives of twenty scientists, reflecting on the latent capacity of AI models to hasten scientific innovation and enrich the educational and communicative dimensions of various professions. A particular instance of these prospects is articulated by Megahed et al.[52], who posited that AI models could fortify statistical process control practices. Nevertheless, cautionary notes were sounded regarding potential misapplications and misconceptions, particularly as these tools are nascent.

Moreover, proposals have been tendered for the integration of AI models within the medical sphere. Insightful elucidations by Murphy and Thomas[53] regarding the potential applications of generative AI in spinal cord injury research and care—including the creation of virtual human body models, optimization of spinal stimulation protocols, drug design, and enhancements in robotic exoskeletons—further amplify the breadth and potential of these innovations. The authors also explored AI's prospective role in research participant recruitment, personalized patient engagement, and medical complication forecasting. Such contributions underscore the transformative capacity of generative AI models in traditional research, while concurrently accentuating the imperative for their judicious and responsible implementation. Within the domains of statistical control, medical inquiry, or scientific discovery, generative AI inaugurates both invigorating opportunities and crucial challenges, thereby informing the contours of future research landscapes.

2.3. Digital homunculi: Individual and shared consciousness

The notion of digital homunculi, alternatively referred to as changelings, extends beyond the mere technological domain and permeates the fields of philosophy and speculative fiction. This concept concerns the creation of digital beings that hold an individualized consciousness, yet simultaneously exist as part of a greater collective consciousness. Wagner[54] delves into this complex matter, elucidating the philosophical ramifications of formulating digital entities endowed with individual awareness yet interconnected within a larger collective framework. This delicate equilibrium, as Wagner insightfully posits, challenges and potentially redefines conventional paradigms of self and identity. Complementing Wagner's exploration, Noel et al.[55] undertake an examination of the complexities associated with manifesting shared consciousness in digital environments. This research seeks to understand the ethical, psychological, and sociological consequences and opportunities that arise from such novel forms of shared existence.

The culturally resonant Star Trek universe, distinguished by its intricate and diverse spectrum of species and civilizations, serves as a fertile ground for metaphoric exploration of shared consciousness. Ganaway[56] embarks on a thorough analysis of the Borg, a collective entity within the Star Trek narrative, construing their existence as a warning against the potential obliteration of individual identity. Providing a contrast to this viewpoint, Biocca and Lanier[57] directs attention to another species known as the "changelings." These beings, characterized by a more fluid and multi-dimensional understanding of self and collective consciousness, provide a more refined and textured lens through which shared consciousness can be examined. Through these examples and analyses, the intricate relationships between individual and collective identities within the realm of digital existence are brought into sharper focus.

Before the current age of generative AI and LLMs, these considerations were had in the arena of robotics.

For instance, Pardes[58] considered the case for giving robots identity in an exploration of AI and robotics. A focal point of this inquiry is the interaction between Stephanie Dinkins, an art professor, and Bina, a robot created in the likeness of a black woman and constructed by entrepreneur Martine Rothblatt to replicate the speech patterns and appearance of Rothblatt's spouse, Bina. This robot, unlike conventional artificial entities, possesses characteristics that are reflective of specific human attributes, allowing for an exploration of understanding the concept of "blackness" in a non-human entity. Questions related to racism and identity are posed, eliciting complex responses from the robot. Gradually, with technological advancements and ongoing interactions, Bina48 manifests increased reflection and a nuanced comprehension of its identity. The intricate engagement observed between Dinkins and Bina48 also reveals the potential for robots to evolve beyond mere instrumental functions, fostering relational connections and self-awareness.

In fact, training AI to imitate specific personalities and human characteristics continues to be demonstrated as an eventuality. For instance, Vincent[59] reported on the results of data scientist Izzy Miller and his attempts to clone his friends' group chat using AI. Group chats, what Miller refers to as "a hallowed thing" in popular culture, can be successfully replicated in the example provided, which is a seven-year-long conversation between himself and five friends since college. Employing the LLaMA model by Meta and a meticulous fine-tuning with 500,000 messages, Miller devised AI counterparts for each group member—Harvey, Henry, Wyatt, Kiebs, Luke, and himself. Such fine-tuning allowed the AI to understand the distinct personalities of the group members and replicate the way they speak. Miller termed the AI-generated chat as the "robo boys." The reported results were "uncanny," as the AI model recognized specific life details and mimicked speech patterns so precisely that the data scientist sometimes had to verify that the generated texts were not direct reproductions from the historical data. Although the model had limitations, such as a blurred distinction between individual personalities and an inability to distinguish past from present, Miller found the project highly entertaining. A noteworthy quote that encapsulates the essence of the entire project and serves as a testament to Miller's experience is: "I was really surprised at the degree to which the model inherently learned things about who we were, not just the way we speak"[59]. This quotation is pivotal as it emphasizes the profound extent to which AI can understand and imitate human nuances, leading to both exhilarating and uncanny results.

The latest body of research has illuminated the prospects for employing AI agents, capitalizing on their singular capacity to emulate a variety of personas. A pioneering endeavor, undertaken collaboratively by scholars from Stanford University and Google, has culminated in the construction of an RPG-style virtual universe, bearing resemblance to The Sims, and populated by 25 characters. These virtual beings, all steered by the combined force of ChatGPT and specialized coding, manifest life-like autonomous conduct. Within the framework of their research, Park et al.[60] harnessed the ChatGPT API to facilitate social interactions and crafted an architecture that mirrored minds endowed with memories and experiences. This experimental setting, named "Smallville," encompassed various locales such as residences, a café, a park, and a grocery store.

The characters within this virtual ecosystem were visually represented through basic sprite avatars, each imbued with a distinct personality and interconnections within the community, engendered through a paragraph-long natural language depiction. Human interlocutors were afforded the ability to engage with these AI agents through conversation or directive commands as an "inner voice." To surmount the constraints associated with ChatGPT's "context window," the research team devised a system adept at retrieving pertinent fragments of an agent's memory as required, thereby influencing the agent's actions, long-term objectives, and reflective thoughts.

The encounters between AI agents engendered information exchange and memory formation, facilitated through natural language via ChatGPT. As an outcome, intriguing emergent behaviors were observed, encompassing phenomena like "information diffusion," "relationship memory," and "coordination," all of which emanated from inter-agent interactions rather than pre-programmed design. A notable illustration was the collaborative planning and participation in a Valentine's Day festivity. External evaluators, engaged to gauge the authenticity of the AI agents' behavior, discerned that the comprehensive generative agent architecture surpassed human role-play responses in terms of persuasiveness. Nevertheless, the researchers were circumspect in highlighting the ethical considerations and potential pitfalls of such technological advancements. Concerns such as the engendering of improper "parasocial relationships," inaccurate deductions, extant risks associated with generative AI, and an undue dependency on generative agents were articulated. As remedial measures, adherence to principles such as transparency regarding the computational nature of agents, alignment with value systems, adherence to best practices in human-AI design, maintenance of audit logs, and avoidance of substituting genuine human input were advocated. Additionally, the investigators furnished an interactive online demonstration, affording a glimpse into the multifaceted social interactions achievable within a seemingly simplistic virtual world.

A similar study was undertaken by Ganguli et al.[61] with an exploration of the ability of large language models to engage in moral self-correction. The research delves into the question of whether models trained with reinforcement learning from human feedback (RLHF) can be instructed to avoid harmful outputs, especially those that may perpetrate stereotypes, bias, and discrimination. Three distinct experiments were carried out, each illuminating various facets of moral self-correction. These investigations collectively provided robust evidence in support of the central hypothesis, revealing that the capacity for moral self-correction appears to emerge at 22 billion model parameters. Importantly, the capability improves with an increase in both model size and the application of RLHF training. The results demonstrate two significant abilities: (1) the capacity to follow instructions, and (2) the aptitude to understand complex normative concepts of harm. The implications of the findings include the potential to create models that resonate more deeply with human values, the challenge of defining and operationalizing complex ethical concepts within machine learning, and the recognition of a dynamic relationship between technology and morality.

3. Recommendations

In light of the comprehensive examination of the current literature, several salient observations and recommendations can be derived. The burgeoning field of LLMs stands at the forefront of contemporary technological innovation, intertwining with the profound philosophical inquiries of understanding, consciousness, and human-like artificial intelligence.

3.1. LLMs and consciousness

The comprehensive examination of research surrounding LLMs reveals several thematic trends and key takeaways that warrant significant attention and further investigation. There has been a paradigm shift in understanding the nature of intelligence in general. For instance, Abio[62] argues that statistics, often dismissed as mere mathematical tools, indeed provide a gateway to understanding, with LLMs serving as an illuminating window into the multifaceted nature of intelligence. The view promotes the exploration of complex sequence learning and social interaction as possible foundations for general intelligence, and concurrently invokes ethical considerations surrounding artificial beings. Likewise, the convergence of LLMs on human-like representations has emerged as a striking trend. The work of Li et al.[63] illustrates an intriguing structural similarity between human neural response measurements and the representations within LLMs. These challenges prior dismissals of LLMs as mere stochastic entities, thus prompting deeper inquiries into the mechanics of meaning emergence. Moreover, the call for interdisciplinary collaboration between cognitive scientists, philosophers, and AI researchers highlights the potential for unraveling the parallels and distinctions between human and artificial cognition.

Furthermore, the potential contributions to psycholinguistics have been acknowledged as a salient trend. As Houghton et al.[64] outline, LLMs hold substantial value in psycholinguistics, serving as tools, comparatives, and philosophical foundations for examining the intricate relationship between language and thought. This offers a renewed perspective on the profound impact on human cognition that language represents and underscores the necessity for utilizing LLMs in this realm. Also, the notion of bridging the gap between LLMs and cognitive models is explored by several authors, notably Shiffrin and Mitchell[65]. The possibility of adapting LLMs to become generalist cognitive models opens up transformative research avenues. Finetuning LLMs on psychological experiments could enable these models to accurately represent human behavior, thereby revolutionizing cognitive psychology and behavioral sciences. This calls for dedicated research into the adaptation processes and ethical considerations associated with such transformative adaptations.

Finally, the intersection of LLMs and the Theory of Mind adds complexity and richness to the discourse. Trott et al.[66] present an exploration into the delicate balance between language exposure, innate biological endowment, and child development. While LLMs like GPT display sensitivity to the beliefs of others, they fall short of human performance, accentuating the need for a more nuanced understanding of the human development of the Theory of Mind. A more holistic approach, perhaps fostered through collaboration among linguists, psychologists, and AI specialists, may offer profound insights into these complex phenomena.

3.2. Neural networks and consciousness

Related to considerations in research of LLMs, neural networks and consciousness presents a rich body of scholarship that reveals key thematic trends and essential takeaways. Here a significant theme in recent research is the concept of representing collective unconsciousness through neural networks. Abou-Haila et al.[67] highlight a paradigm where population cognition is embodied using artificial neural networks, simulating adaptation while maintaining individuality. This perspective bridges cognitive science and computational modeling, offering novel insights into predator-prey interactions and applications in multi-agent systems such as artificial life or computer games.

There has also been an inclination towards hybrid models for consciousness, where new frameworks of neural network models are being introduced. Lu et al.[68] proposed a three-layered structure that cooperates to accomplish information storage and cognition through specific processes like the reception, partial recognition, and resonant learning process. This framework lends a broader approach to analyzing brain functions, providing a dynamic lens through which to explore and explain human actions. Similarly, an intriguing trend emerged around the concept of consciousness as a collective excitation of a brain-wide web. Lidström and Allen[69], for example, advocated for a physics-oriented perspective to describe consciousness as a collective phenomenon extending into various neuronal networks. They emphasize the importance of experiments and large-scale simulations in uncovering the details and complexities of consciousness. This perspective adds a physical dimension to the existing knowledge, emphasizing the need for a realistic and comprehensive approach.

The theme of collective cognition and its influence on social network topology has evolved out of these studies. Momennejad[70] represents a review that integrates network structure with psychological and neural experiments to understand how social networks shape collective cognition. Graph-theoretical approaches, neuroimaging studies, cognitive similarities, and machine learning approaches are amalgamated to shed light on how social structures influence collective cognition. The insight offers an innovative perspective on designing goal-directed social network topologies and broadens the scope of understanding human cognition from an interconnected standpoint, which is of interest to those in robotics, as well. For instance, the emergence of collective cognition in robotic swarms surfaces as a noteworthy trend. The study by Otte[71] of algorithms transforms a robotic swarm into a unified computational meta-entity, leading to a swarm-spanning artificial neural network. The artificial group-mind, capable of differentiating spatial patterns and orchestrating coordinated swarm responses, exemplifies the frontier of human-swarm interaction and showcases the potential of distributed sensor data in collective decision-making.

These interwoven themes converge to form a sophisticated picture of the field, emphasizing the dynamism of collective consciousness, the evolution of hybrid neural network models, the physical reality of collective excitations, the interplay between social structures and collective cognition, and the emergence of group-mind in robotic swarms. The complexity and diversity of these approaches underscore the imperative for interdisciplinary collaboration and integration of various scientific domains. The complex interplay facilitates a profound understanding of consciousness, not only as a solitary phenomenon but also as a collective and interconnected web of neuronal, social, and technological dimensions. It is within these multifaceted frameworks that the future of neural network research may find fertile ground for exploration and discovery.

3.3. AI agents and cultural heritage

The exploration of AI agents in the realm of cultural heritage has surfaced rich insights and opened up new avenues for research and application. Among these, the concept of personalized virtual environments stands out as a significant advancement. Kiourt et al.[72] have delved into the development of multi-agent systems, enabling the creation of dynamic and tailored virtual environments. Such personalized experiences have been further emphasized by Costantini et al.[73], who explored user profile agents that track user movements through satellite signals. In leveraging LLMs for cultural heritage, integrating natural language processing to create immersive and tailored experiences becomes a potent recommendation, inviting users into a vivid and engaging exploration of cultural histories and artifacts.

Moving beyond personalized engagement, the realm of knowledge discovery and advanced data management has also witnessed remarkable progress. Researchers such as Buratti et al.[74] and Abbattista et al.[75] have underscored the power of AI to trigger knowledge and implement advanced methods for information management. This revelation leads to the recommendation of employing the sophisticated data analysis techniques of LLMs, offering the potential to curate, interpret, and present complex cultural data in accessible formats. On the other hand, another critical aspect unearthed in the literature is the role of intelligent systems in the preservation, authentication, and retrieval of cultural heritage. Works by Pavlidis[76] and Garau[77] highlight the essential functions of AI in safeguarding digital assets. Here, LLMs could be instrumental in automating documentation, enhancing authentication accuracy, and improving retrieval mechanisms—a transformation that could revolutionize the way cultural heritage is maintained and accessed. Simultaneously, a strong emphasis on a human-centered approach and accessibility resonates through the research. Scholars like Leshkevich and Motozhanets[78], Pisoni et al.[79], and Diaz-Rodríguez and Pisoni[80] articulate the importance of creating a human-centered future in the digitization of cultural heritage. The application of LLMs in translating, simplifying, and providing alternative representations of information could significantly widen accessibility, aligning technology with human needs and inclusivity.

The exploration and curation of cultural heritage have also seen innovations and opportunities. Studies by Ardissono et al.[81], Ranaldi and Zanzotto[82], and Yurtsever[83] emphasize the richness of tangible and intangible cultural heritage. LLMs could be harnessed for innovative curation and exploration, responding to user queries and presenting synthesized information in interactive formats—an innovation that could bring cultural heritage closer to people. The experience can be made all the more tangible with the integration with the Internet of Things (IoT) and the design considerations of agent-based systems, as explored by Lee and

Lee[84] and Jennings and Wooldridge[85], suggest a future where linguistic intelligence augments IoT devices. Here, LLMs could provide real-time interaction with cultural exhibits, bridging the gap between the physical and digital realms.

The convergence of these themes offers a multifaceted view of how AI agents are shaping the field of cultural heritage. Recommendations to leverage LLMs span across personalization, knowledge discovery, preservation, human-centered design, exploration, and integration with emerging technologies. The alignment of these technologies with cultural heritage aims to enhance human connection and understanding, creating a pathway that is not just technologically advanced but also rich in meaning and accessibility. Collaborative efforts among researchers, cultural experts, and technologists will indeed pave the way for a future where cultural heritage is not merely preserved but comes alive through the synergy of language, technology, and human insight.

In the interplay between the realms of cultural heritage, AI, and the metaverse, the use of LLMs offers a unique and promising pathway for immersive experiences. The implications of such an integration extend to various educational, cultural, and societal dimensions. The first dimension to consider is the interaction with historical figures through the lens of individual consciousness. Engaging with prominent rulers, philosophers, and thinkers, this mode allows a personalized experience where users can delve into dialogues, debates, and intellectual explorations. For instance, conversing with Emperor Augustus in a reconstructed environment of ancient Rome would shed light on governance, empire-building, and political philosophy. Interacting with Hypatia in ancient Alexandria could illuminate the intellectual milieu of the time, while a discourse with Mansa Musa of the Mali Empire has the potential to elucidate the complex facets of medieval African economics. These scenarios offer rich educational opportunities, enabling critical thinking and a vivid connection with history. LLMs would play a crucial role in modeling these dialogues and simulating personalities to ensure authenticity.

The second dimension explores the concept of collective or shared consciousness, focusing on the communal and societal aspects of historical towns, cities, or entire civilizations. Unlike individual consciousness, this approach provides a broader view, encompassing the social fabric, cultural norms, and everyday life of a particular era. Imagine participating in the Athenian assembly, immersed in the democratic processes of ancient Greece, or exploring the urban planning of Harappa, one of the world’s earliest urban civilizations. Even a walk through the streets of Tenochtitlan, the Aztec capital, would unveil an intricate world of rituals, architecture, and power dynamics. Within these collective experiences, LLMs would serve to animate daily life scenarios and provide contextual understanding, thus enhancing the user’s connection with diverse cultures.

Building upon these two modalities, several recommendations emerge for leveraging LLMs for cultural heritage experiences in the metaverse. The paramount concern should be ensuring historical authenticity and cultural sensitivity. Both individual and collective experiences must be developed with a keen eye on historical accuracy and an understanding of cultural nuances to avoid misrepresentation. Accessibility and inclusivity must be at the forefront of design, enabling diverse audiences to partake in these experiences and fostering global empathy. Interdisciplinary collaboration would be essential, bringing together historians, anthropologists, linguists, and technologists to craft a harmonious fusion of historical knowledge and advanced technology. Ethical considerations should guide the portrayal of historical figures and cultures, maintaining respect and integrity. Finally, aligning these experiences with educational objectives would enhance learning, transforming it into an interactive and engaging journey.

3.4. Discussion of results

As AI advances, a pivotal research question concerns the capability of AI agents to convincingly simulate human behavior and experience. Evaluating this capability requires nuanced qualitative assessments by human observers as well as quantitative performance benchmarks. Thus far, the paper has presented analyses aimed at evaluating the proficiency of AI agents in mimicking the multifaceted essence of human interaction. However, the specific dimensions of evaluation and metrics used require further elucidation to substantiate the claims made.

The inception of this research stems from the theoretical premise that AI agents are approaching, yet have not attained, human-level convincibility in their language, contextual awareness, and personality rendering. Testing this hypothesis necessitates appraising generated agent outputs from both qualitative and quantitative vantage points. Qualitatively, blinded human judges must evaluate samples of agent output based on criteria measuring closeness to natural human performance. Quantitatively, metric scores can benchmark agent performance and capability for improvement across generations of the underlying models.

To improve reader comprehension, this section will detail the key qualitative dimensions guiding human judge evaluations in this study. Additionally, the quantitative metrics applied to benchmark agent performance on these dimensions will be delineated. By elaborating on this evaluation framework, the aim is to strengthen the connection between theoretical concepts of imitation capability and results demonstrating current agent proficiency levels. The dimensions elucidated here reflect an interdisciplinary approach, integrating principles of linguistics, psychology and human-computer interaction to holistically evaluate the "humanness" of agent outputs. This comprehensive framework underscores the complexity of modeling human cognition and behavior, while providing tangible criteria to assess progress in this crucial technological pursuit. As such, the following outlines the criteria that may be used in future research into this field.

The AI agents should be evaluated based on three key qualitative criteria (**Table 1**):

1) **Naturalness of language:** This dimension evaluates the overall linguistic naturalness of AI agent outputs in conversational settings. Several factors contribute to perceiving language as natural, including:
    a) **Grammatical correctness**—sentence structure, syntax, grammar follow the conventions and rules of human language. Errors in grammar are rare.
    b) **Fluency**—the conversation flows smoothly without awkward phrasing, unnatural cadence, or oddly placed pauses.
    c) **Coherence**—responses are coherent and logical given the context of the conversation. Tangential or confusing responses are minimal.
    d) **Idiomatic usage**—gents incorporate idioms, figures of speech, culturally relevant references appropriately within the dialogue.
    e) **Human-like variance**—agents exhibit some degree of variance from strictly proper grammar to more informal, conversational language where appropriate given context.

The naturalness of language samples should be evaluated by having blinded human judges read/listen to dialogues and rate them on a 5-point Likert scale, with detailed rubrics to guide scoring. Higher scores indicate greater humanness across the linguistic factors above.

2) **Contextual awareness:** This evaluates the ability of agents to maintain awareness of pertinent context, history, and prior statements during a conversation when formulating appropriate responses. Key markers of contextual awareness include:
    a) **Recalling/referencing previous statements**—agents will explicitly refer to earlier parts of the conversation when relevant.

b) **Maintaining consistency**—agents avoid contradictory statements over the course of a conversation.

c) **Incorporating relevant personal facts**—agents integrate memory of names, places, interests mentioned previously.

d) **Adapting to new information**—agents integrate and apply new facts brought up during the dialogue.

e) **Maintaining topic relevance**—agents relate responses directly to the current topic and refer back to earlier topics naturally.

Human judges should be provided extended dialogue samples and score percentage of contextually appropriate responses. Higher percentages denote greater contextual awareness.

3) **Distinct personality:** This evaluates the uniqueness and consistency of personality characteristics portrayed by different agents over multiple conversations. Relevant markers of personality include:

a) **Idiosyncratic speech patterns**—distinctive vocabulary, phraseology, figures of speech.

b) **Display of attitudes, beliefs, opinions**—articulation of agent-specific worldviews.

c) **Nature of emotional expression**—variability in agents' tone, sentiment, emotionality.

d) **Humor style**—differences in agents' use of humor, irony, wit.

e) **Background/experience references**—agents reference different personas, histories, interests.

Human judges should compare the outputs of multiple agents in response to standardized prompts to assess personality distinctiveness and consistency both within and across agents. Inter-rater agreement levels quantify this.

Table 1. AI Agent qualitative rubric.

Qualitative criterion	Subcriteria	Description	Likert scale (1–5)
Naturalness of language	Grammatical correctness	Sentence structure, syntax, grammar follow conventions of human language. Errors are rare.	1—Frequent errors, incomprehensible 2—Many errors, comprehension difficult 3—Some noticeable errors 4—Minor errors 5—No discernible errors
	Fluency	Conversation flows smoothly without awkward phrasing or unnatural cadence/pauses.	1—Faltering, halting conversation 2—Somewhat awkward phrasing/pauses 3—Moderately uneven flow 4—Minor irregularities 5—Fluent, natural cadence
	Coherence	Responses are logical and coherent given conversation context. Confusing tangents are rare.	1—Largely incoherent 2—Frequent non-sequiturs 3—Partial coherence 4—Mainly coherent 5—Responses align well with context
	Idiomatic usage	Incorporates idioms, figures of speech, cultural references appropriately.	1—No idiomatic usage 2—Occasional idiomatic phrases 3—Some usage with errors 4—Appropriate usage 5—Idiomatic mastery
	Human-like variance	Exhibits informal, conversational variance from proper grammar when fitting.	1—Rigid style, no variance 2—Minimal variance 3—Moderate informal tone 4—Appropriate conversational variance 5—Human-like style shifting
Contextual awareness	Recalling/referencing previous statements	Refers explicitly to earlier parts of conversation when relevant.	1—Never recalls previous statements 2—Rarely recalls 3—Occasionally recalls 4—Frequently recalls 5—Always recalls when relevant

Table 1. (*Continued*).

Qualitative criterion	**Subcriteria**	**Description**	**Likert scale (1–5)**
	Maintaining consistency	Avoids contradictory statements throughout conversation.	1—Always inconsistent 2—Mostly inconsistent 3—Somewhat consistent 4—Mostly consistent 5—Always consistent
	Incorporating relevant facts	Integrates names, places, interests mentioned previously.	1—Never incorporates 2—Rarely incorporates 3—Occasionally incorporates 4—Frequently incorporates 5—Always incorporates relevant facts
	Adapting to new information	Applies and integrates new facts brought up in dialogue.	1—Never adapts 2—Rarely adapts 3—Occasionally adapts 4—Frequently adapts 5—Always adapts to new information
	Maintaining topic relevance	Relates responses directly to current topic and refers back to previous topics naturally.	1—Never relevant 2—Rarely relevant 3—Occasionally relevant 4—Frequently relevant 5—Always relevant
Distinct personality	Idiosyncratic speech patterns	Uses distinctive vocabulary, phrasing, figures of speech.	1—No distinctive patterns observed 2—Rare occurrence of unique phrasing or vocabulary 3—Some distinctive patterns, but inconsistently applied 4—Consistently displays unique patterns 5—Rich, complex, and consistent unique speech patterns
	Attitudes/beliefs expressed	Articulates opinions, worldview, perspectives unique to agent.	1—No articulation of opinions or worldviews 2—Sparse articulation with minimal depth 3—Moderate articulation but lacking nuance 4—Articulates with depth and consistency 5—Articulates complex, nuanced opinions and worldviews
	Emotional Expression	Exhibits distinctive emotional tone, sentiment, variability.	1—No emotional tone or sentiment 2—Minimal emotional expression, lacks variability 3—Moderate emotional expression but not diverse 4—Diverse and consistent emotional expression 5—Rich, complex emotional expression with high variability
	Humor Style	Uses humor, irony, wit in characteristic ways.	1—No humor, irony, or wit observed 2—Sparse use, lacking in characteristic style 3—Moderate use but lacks distinctiveness 4—Consistent humor style but lacks complexity 5—Rich, complex, and unique humor style

Table 1. (*Continued*).

Qualitative criterion	Subcriteria	Description	Likert scale (1–5)
	Background/Experience References	Refers to persona-specific histories, interests, roles.	1—No references to background or experience 2—Sparse references but lacking context or relevance 3—Some references but not well-integrated 4—Consistent, relevant references 5—Rich, complex, and integrated references to background and experience

Additionally, quantitative metrics should accompany these qualitative assessments and include the following considerations (**Table 2**).

Table 2. AI agent quantitative rubric.

Criteria	Measurement method	Scoring metric
Language naturalness	5-point Likert scale	5- Indistinguishable from human 4- Largely natural with minor irregularities 3- Moderately unnatural, several errors 2- Mostly unnatural but comprehensible 1- Incomprehensible or nonsensical
Contextual awareness	5-point Likert scale	5- Fully context-aware, no inconsistencies 4- Mostly context-aware, minor errors 3- Moderately context-aware, some irrelevant or contradictory responses 2- Limited contextual awareness, many errors 1- No contextual awareness, entirely inappropriate responses
Personality consistency	5-point Likert scale	5- Highly consistent personality traits 4- Mostly consistent with minor inconsistencies 3- Moderately consistent, noticeable variations 2- Inconsistent personality traits 1- No discernible personality traits

3.4.1. Language naturalness scoring

Language naturalness should be evaluated on a 5-point Likert scale by blinded human judges. The scale was defined as follows:

5—Indistinguishable from human language

4—Largely natural with minor irregularities

3—Moderately unnatural, several errors

2—Mostly unnatural but comprehensible

1—Incomprehensible or nonsensical

Average scores should be calculated across all judges for each AI agent. Higher average scores indicate more natural language proficiency. Comparing average scores over model iterations shows improving language capabilities.

3.4.2. Contextual awareness scoring

Contextual awareness should be calculated as the percentage of contextually appropriate responses out of total responses analyzed for a given AI agent. A contextually appropriate response was defined as a reply directly relevant to the current topic and conversation history without contradicting prior statements. Non-sequiturs or contradictory responses were marked as contextually inappropriate.

Higher percentages equate to greater exhibited contextual awareness by the agent. Percentages should be tracked across model generations to quantify improvements.

3.4.3. Personality consistency scoring

Personality consistency should be quantified by having human judges attribute personality descriptors (e.g., “witty”, “somber”, “playful”) to different agents based on sample responses. Inter-rater agreement was then measured by calculating:

Number of matching personality attributions/Total number of attributions

Higher inter-rater agreement indicates more consistent personality rendering across judges. Improvements in consistency levels over successive agent models demonstrates more distinctly characterized personas.

By elaborating on the quantitative measures accompanying the qualitative assessments, this aims to paint a more complete picture of the AI agent evaluation process and how enhancements were benchmarked over time.

To illustrate these dimensions concretely, an example is provided comparing two hypothetical AI agents named “Socrates” and “Shakespeare” in a dialogue:

Socrates: *Greetings friend! Let us continue our illuminating discussion on the nature of justice.*

Shakespeare: *Alas my friend, I find myself in a brooding mood today and have little appetite for cerebral debate. Perhaps we might postpone this exchange for a time when inspiration strikes?*

In this excerpt, Socrates exhibits philosophically inquisitive speech patterns true to the historical figure, while Shakespeare shows melancholy and poetic sensibilities fitting to the dramatist. Their personalities are distinctly rendered through contextual, idiomatic language. Overall, human judges would likely rate this sample highly on all three qualitative dimensions. Let us examine how human judges might assess this sample on the three qualitative evaluation criteria:

Naturalness of language

Both agents utilize grammatically correct, fluent language with no overt errors. The vocabulary and phrasing choices match the characters - Socrates opts for philosophical and inquisitive wording (“illuminating discussion”); Shakespeare uses dramatic, metaphorical language (“brooding mood”, “little appetite”). The exchange flows conversationally. Overall, human judges would likely rate the linguistic naturalness highly.

Contextual awareness

The agents clearly recall and build on prior interactions (“let us continue our discussion”). Socrates maintains his intellectual persona while Shakespeare adapts to a reflective state of mind in this moment. Both responses directly address the preceding statement. Since the agents exhibit consistent personas and strong contextual awareness, judges would likely deem a high percentage of turns contextually appropriate.

Distinct personality

The personas come through distinctly—Socrates as cerebral and inquisitive, Shakespeare as dramatic and melancholic. The speech patterns and word choices align well to the historical figures. Judges would likely recognize the personalities as highly differentiated and consistent.

Therefore, the sample dialogue demonstrates strong performance across all three qualitative measures. The conversational exchange highlights the potential for AI agents to convincingly emulate nuanced human behavior when evaluated on dimensions of language, context, and personality. Samples such as this one provides the tangible bases for judging the progress of generative models in producing human-like autonomous agents.

By elaborating on the qualitative criteria, quantitative metrics, and sharing illustrative samples, this section aims to paint a clearer picture of the evaluation processes and ground the theoretical discussions in demonstrable results. Future work could further enrich these findings through expanded trials, additional quantitative measures, and informative visualizations of outcomes. Nonetheless, articulating the nature of assessments conducted represents an important step toward a transparent delineation between evidence-based results and speculative commentary in this emerging field of research.

4. Conclusion

The exploration of AI, specifically through LLMs, in the context of cultural heritage and the metaverse, unveils a pioneering avenue for historical, educational, and cultural engagement. The dual modality of individual and collective consciousness, simulating interactions with significant historical figures and civilizations, offers an unprecedented richness in experience and understanding. The potentiality of personalized dialogues with using LLMs with eminent historical figures, the immersive exploration of societal structures through collective consciousness, and the imperative of a balanced approach emphasizing historical authenticity, cultural sensitivity, accessibility, and interdisciplinarity. These insights pave the way for a vivid and instructive connection with history, fostering critical thinking, empathy, and global awareness.

This paper presents an initial exploration of a complex topic spanning multiple disciplines. As such, there are limitations to the depth of inquiry and empirical analysis. The theoretical discussions would be enhanced by future empirical work examining the concepts of individual and collective consciousness through psychological studies, user testing of AI interfaces, and neuroimaging of participants interacting with artificial agents. Rigorously designed experiments are needed to validate the anecdotal observations made regarding phenomena such as awareness diffusion in digital collectives. Additionally, the development of formal frameworks and quantitative metrics to evaluate notions of selfhood and identity in relation to collective consciousness could strengthen the conceptual arguments presented here.

While this paper draws parallels to science fiction narratives, direct evidence from implemented AI systems is limited. Future research could involve building prototype AI agents and platforms to practically demonstrate the principles outlined in the paper. User studies with these systems could provide valuable insights into perceptions of identity and self when engaging with artificial consciousness. Additionally, interdisciplinary collaboration with computer scientists and cognitive psychologists is needed to refine the AI architectures proposed. Subsequently, the cultural heritage applications suggested could be implemented and empirically tested for usability and educational efficacy.

Conflict of interest

The author declares no conflict of interest.

References

1. Lee H. The rise of ChatGPT: Exploring its potential in medical education. *Anatomical Sciences Education* 2023. doi: 10.1002/ase.2270
2. Zhang X, Yang D, Yow CH, et al. Metaverse for cultural heritages. *Electronics* 2022; 11(22): 3730. doi: 10.3390/electronics11223730
3. Garon J. A practical introduction to generative AI, synthetic media, and the messages found in the latest medium. Available online: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4388437 (accessed on 6 November 2023).
4. Gill SS, Kaur R. ChatGPT: Vision and challenges. *Internet of Things and Cyber-Physical Systems* 2023; 3: 262–

271. doi: 10.1016/j.iotcps.2023.05.004

5. Albahri AS, Duhaim AM, Fadhel MA, et al. A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Information Fusion* 2023; 96: 156–191. doi: 10.1016/j.inffus.2023.03.008
6. Lee W, Lee DH. Cultural heritage and the intelligent Internet of Things. *Journal on Computing and Cultural Heritage* 2019; 12(3): 1–14. doi: 10.1145/3316414
7. Liu Y, Han T, Ma S, et al. Summary of ChatGPT/GPT-4 research and perspective towards the future of large language models. *arXiv* 2023; arXiv:2304.01852. doi: 10.1016/j.metrad.2023.100017
8. Latif E, Mai G, Nyaaba M, et al. Artificial general intelligence (AGI) for education. *arXiv* 2023; arXiv:2304.12479. doi: 10.48550/arXiv.2304.12479
9. Ray PP. ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems* 2023; 3: 121–154. doi: 10.1016/j.iotcps.2023.04.003
10. Inamura T. Digital twin of experience for human-robot collaboration through virtual reality. *International Journal of Automation Technology* 2023; 17(3): 284–291. doi: 10.20965/ijat.2023.p0284
11. Falandays JB, Kaaronen RO, Moser C, et al. All intelligence is collective intelligence. *Journal of Multiscale Neuroscience* 2023; 2(1): 169–191. doi: 10.56280/1564736810
12. Gaudenzi S. *The Living Documentary: From Representing Reality to Co-Creating Reality in Digital Interactive Documentary* [PhD thesis]. Goldsmiths, University of London; 2013.
13. Ng SL, Kinsella EA, Friesen F, Hodges B. Reclaiming a theoretical orientation to reflection in medical education research: A critical narrative review. *Medical Education* 2015; 49(5): 461–475. doi: 10.1111/medu.12680
14. Tzortzaki, D. Museums and virtual reality: Using the CAVE to simulate the past. *Digital Creativity* 2001; 12(4): 247–251. doi: 10.1076/digc.12.4.247.3216
15. Roussou M. Immersive interactive virtual reality in the museum. Available online: https://www.researchgate.net/profile/Maria-Roussou-2/publication/2861971_Immersive_Interactive_Virtual_Reality_in_the_Museum/links/0c9605192924ee109d000000/Immersive-Interactive-Virtual-Reality-in-the-Museum.pdf (accessed on 6 November 2023).
16. Christou C. Virtual reality in education. In: Tzanavari A, Tsapatsoulis N (editors). *Affective, Interactive and Cognitive Methods for E-Learning Design: Creating an Optimal Educational Experience.* Information Science Publishing; 2010. pp. 228–243.
17. Favro D. In the eye of the beholder: VR models and academia. In: Haselberger L, Humphrey J (edotors). *Imaging Ancient Rome: Documentation, Visualization, Imagination*. Journal of Roman Archaeology; 2006. pp. 321–334.
18. Clini P, Ruggieri L, Angeloni R, Sassob M. Interactive immersive virtual museum: Digital documentation for interaction. Available online: https://isprs-archives.copernicus.org/articles/XLII-2/251/2018/isprs-archives-XLII-2-251-2018.pdf (accessed on 6 November 2023).
19. Huaman EMR, Aceituno RGA, Sharhorodska O. Application of virtual reality and gamification in the teaching of art history. In: Zaphiris P, Ioannou A (editors). *Learning and Collaboration Technologies. Ubiquitous And Virtual Environments for Learning and Collaboration (Lecture Notes in Computer Science).* Springer, Cham; 2019. pp. 220–229.
20. Chandler T, Clulow A. Modeling virtual Angkor: An evolutionary approach to a single urban space. *IEEE Computer Graphics and Applications* 2020; 40(3): 9–16. doi: 10.1109/MCG.2020.2982444
21. Valk A, Mi X, Schick AL. *Making Virtual Reality a Reality: Designing Educational Initiatives in Libraries with Emerging Technologies.* Bloomsbury Publishing USA; 2023.
22. Balhara S, Gupta N, Alkhayyat A, et al. A survey on deep reinforcement learning architectures, applications and emerging trends. Available online: https://ietresearch.onlinelibrary.wiley.com/doi/pdf/10.1049/cmu2.12447 (accessed on 6 November 2023).
23. Strich F, Mayer AS, Fiedler M. What do I do in a world of artificial intelligence? Investigating the impact of substitutive decision-making AI systems on employees' professional role identity. *Journal of the Association for Information Systems* 2021; 22(2): 9.
24. Zall R, Kangavari MR. Comparative analytical survey on cognitive agents with emotional intelligence. *Cognitive Computation* 2022; 14(4): 1223–1246.
25. Budhwar P, Malik A, De Silva MT, Thevisuthan P. Artificial intelligence–challenges and opportunities for international HRM: a review and research agenda. *The International Journal of Human Resource Management* 2022; 33(6): 1065–1097.
26. Liu-Thompkins Y, Okazaki S, Li H. Artificial empathy in marketing interactions: Bridging the human-AI gap in affective and social customer experience. *Journal of the Academy of Marketing Science* 2022; 50(6): 1198–1218.
27. Singh A, Chouhan T. Artificial intelligence in HRM: Role of emotional–social intelligence and future work skill. In: *The Adoption and Effect of Artificial Intelligence on Human Resources Management,* Part A. Emerald Publishing Limited; 2023. pp. 175–196.

28. Selvaraj C, Chandra I, Singh SK. Artificial intelligence and machine learning approaches for drug design: Challenges and opportunities for the pharmaceutical industries. *Molecular Diversity* 2021; 26(3): 1893–1913. doi: 10.1007/s11030-021-10326-z
29. Jeste DV, Graham SA, Nguyen TT, et al. Beyond artificial intelligence: Exploring artificial wisdom. *International Psychogeriatrics* 2020; 32(8): 993–1001. doi: 10.1017/S1041610220000927
30. Samsonovich AV. Socially emotional brain-inspired cognitive architecture framework for artificial intelligence. *Cognitive Systems Research* 2020; 60: 57–76. doi: 10.1016/j.cogsys.2019.12.002
31. Picard RW, Papert S, Bender W, et al. Affective learning—A manifesto. *BT Technology Journal* 2004; 22(4): 253–269. doi: 10.1023/B:BTTJ.0000047603.37042.33
32. Li Y, Jiang Y, Tian D, et al. AI-enabled emotion communication. *IEEE Network* 2019; 33(6): 15–21. doi: 10.1109/MNET.001.1900070
33. Khachane MY. Organ-based medical image classification using support vector machine. *International Journal of Synthetic Emotions* 2017; 8(1): 18–30. doi: 10.4018/IJSE.2017010102
34. Pusztahelyi R. Emotional AI and its challenges in the viewpoint of online marketing. *Curentul Juridic* 2020; 81(2): 13–31.
35. Cominelli L, Hoegen G, De Rossi D. Abel: Integrating humanoid body, emotions, and time perception to investigate social interaction and human cognition. *Applied Sciences* 2021; 11(3): 1070. doi: 10.3390/app11031070
36. Liu-Thompkins Y, Okazaki S, Li H. Artificial empathy in marketing interactions: Bridging the human-AI gap in affective and social customer experience. *Journal of the Academy of Marketing Science* 2022; 50(6): 1198–1218. doi: 10.1007/s11747-022-00892-5
37. Wortman B, Wang JZ. HICEM: A high-coverage emotion model for artificial emotional intelligence. *arXiv* 2022; arXiv:2206.07593. doi: 10.48550/arXiv.2206.07593
38. Marcos S, García Peñalvo FJ, Vázquez Ingelmo A. Emotional AI in healthcare: A pilot architecture proposal to merge emotion recognition tools. In: Proceedings of the Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'21); 26–29 October 2021; Barcelona Spain. pp. 342–349.
39. Huh JH. Seo YS. Understanding Edge computing: Engineering evolution with artificial intelligence. *IEEE Access* 2019; 7: 164229–164245. doi: 10.1109/ACCESS.2019.2945338
40. Andersson R. The bioeconomy and the birth of a "new anthropology". *Cultural Anthropology* 2022; 37(1): 37–44. doi: 10.14506/ca37.1.06
41. Naddaf Y. *Game-Independent AI Agents for Playing Atari 2600 Console Games* [Master's thesis]. University of Alberta; 2010.
42. Liu J, Togelius J, Pérez-Liébana D, Lucas SM. Evolving game skill-depth using general video game AI agents. In: Proceedings of the 2017 IEEE Congress on Evolutionary Computation (CEC); 5–8 June 2017; Donostia, Spain. pp. 2299–2307.
43. Holmgard C, Liapis A, Togelius J, Yannakakis GN. Generative agents for player decision modeling in games. In: Proceedings of the 9th International Conference on the Foundations of Digital Games; 3–7 April 2014; Liberty of the Seas, Caribbean. pp. 1–8.
44. Barthet M, Khalifa A, Liapis A, Yannakakis G. Generative personas that behave and experience like humans. In: Proceedings of the 17th International Conference on the Foundations of Digital Games; 5–8 September 2022; Athens, Greece. pp. 1–10.
45. Ariyurek S, Betin-Can A, Surer E. Automated video game testing using synthetic and humanlike agents. *IEEE Transactions on Games* 2019; 13(1): 50–67. doi: 10.1109/TG.2019.2947597
46. Fathi K, Palhang M. Evaluation of using neural networks on variety of agents and playability of games. In: Proceedings of the *2018 International Conference on Artificial Intelligence and Data Processing (IDAP);* 28–30 September 2018; Malatya, Turkey. pp. 1–8.
47. Nareyek A. Intelligent agents for computer games. In: *International Conference on Computers and Games.* Springer Berlin Heidelberg; 2000. pp. 414–422.
48. Tan SCG, Nareyek A. Integrating facial, gesture, and posture emotion expression for a 3D virtual agent. In *Proceedings of the 14th International Conference on Computer Games: AI, Animation, Mobile, Interactive Multimedia, Educational & Serious Games*; 29 July–2 August 2009; Louisville, Kentucky, USA. pp. 23–31.
49. Miikkulainen R, Bryant BD, Cornelius R, et al. Computational Intelligence in Games. In: Yen GY, Fogel DB. (editors). *Computational Intelligence: Principles and Practice*. IEEE Computational Intelligence Society; 2006.
50. Fernández S, Adarve R, Pérez M, et al. Planning for an AI based virtual agents game. Available online: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=7739dc608abd41726cb4df7c71ed1816fa214023 (accessed on 6 November 2023).
51. Morris MR. Scientists' perspectives on the potential for generative AI in their fields. *arXiv* 2023; arXiv:2304.01420. doi: 10.48550/arXiv.2304.01420
52. Megahed FM, Chen YJ, Ferris JA, et al. How generative AI Models such as ChatGPT can be (mis)used in SPC

practice, education, and research? An Exploratory Study. *Quality Engineering* 2023. doi: 10.1080/08982112.2023.2206479

53. Murphy C, Thomas FP. Generative AI in spinal cord injury research and care: Opportunities and challenges ahead. *The Journal of Spinal Cord Medicine* 2023; 46(3): 341–342. doi: 10.1080/10790268.2023.2198926
54. Wagner N. *Rhetorical Being: A Metaphysics of Freedom and Essence* [PhD Thesis. Georgia State University; 2018.
55. Noel JP, Ishizawa Y, Patel SR, et al. Leveraging nonhuman primate multisensory neurons and circuits in assessing consciousness theory. *Journal of Neuroscience* 2019; 39(38): 7485–7500. doi: 10.1523/JNEUROSCI.0934-19.2019
56. Ganaway B. Representations of the post/human: Monsters, aliens and others in popular culture. *Journal of Popular Culture* 2003; 37(1): 134.
57. Biocca F, Lanier J. An insider’s view of the future of virtual reality. *Journal of communication* 1992; 42(4): 150–172.
58. Pardes A. The case for giving robots an identity. Available online: https://www.wired.com/story/bina48-robots-program-identity/ (accessed on 3 September 2023).
59. Vincent J. A data scientist cloned his best friends’ group chat using AI. Available online: https://www.theverge.com/2023/4/13/23671059/ai-chatbot-clone-group-chat (accessed on 10 September 2023).
60. Park JS, O'Brien JC, Cai CJ, et al. Generative agents: Interactive simulacra of human behavior. *arXiv* 2023; *arXiv:2304.03442*. doi: 10.48550/arXiv.2304.03442
61. Ganguli D, Askell A, Schiefer N, et al. The capacity for moral self-correction in large language models. *arXiv* 2023; arXiv:2302.07459. doi: 10.48550/arXiv.2302.07459
62. Abio B. In AI, is bigger better? *Nature 2023;* 615: 202–205. doi: 10.1038/d41586-023-00641-w
63. Li J, Karamolegkou A, Kementchedjhieva Y, et al. Large language models converge on brain-like word representations. *arXiv* 2023; arXiv:2306.01930. doi: 10.48550/arXiv.2306.01930
64. Houghton C, Kazanina N, Sukumaran P. Beyond the Limitations of Any Imaginable Mechanism: Large Language Models and Psycholinguistics. *arXiv* 2023; arXiv:2303.00077. doi: 10.48550/arXiv.2303.00077
65. Shiffrin R, Mitchell M. Probing the psychology of AI models. *Proceedings of the National Academy of Sciences* 2023; 120(10): e2300963120. doi: 10.1073/pnas.2300963120
66. Trott S, Jones C, Chang T, et al. Do large language models know what humans know? *Cognitive Science* 2023; 47(7): e13309. doi: 10.1111/cogs.13309
67. Abou-Haila P, Hall R, Dawes M. Representing collective unconsciousness using neural networks. *International Journal of Computer and Information Engineering* 2007; 1(5): 1501–1505. doi: 10.5281/zenodo.1330135
68. Lu X, Lin Z, Jin H, et al. Rating image aesthetics using deep learning. *IEEE Transactions on Multimedia* 2015; 17(11): 2021–2034. doi: 10.1109/TMM.2015.2477040
69. Lidström S, Allen RE. Consciousness as the collective excitation of a brainwide web—Understanding consciousness from below quantum fields to above neuronal networks. *Journal of Physics: Conference Series* 2019; 1275(1): 012021. doi: 10.1088/1742-6596/1275/1/012021
70. Momennejad I. Collective minds: Social network topology shapes collective cognition. *Philosophical Transactions of the Royal Society B* 2022; 377(1843): 20200315. doi: 10.1098/rstb.2020.0315
71. Otte M. Collective cognition and sensing in robotic swarms via an emergent group-mind. In: Kulić D, Nakamura Y, Khatib O, Venture G (editors). *2016 International Symposium on Experimental Robotics.* Springer; 2017. pp. 829–840.
72. Kiourt C, Pavlidis G, Koutsoudis A, Kalles D. Multi-agents based virtual environments for cultural heritage. In: Proceedings of the *2017 XXVI International Conference on Information, Communication and Automation Technologies (ICAT);* 26–28 October 2017; Sarajevo, Bosnia and Herzegovina. pp. 1–6.
73. Costantini S, Mostarda L, Tocchio A, Tsintza P. DALICA: Agent-based ambient intelligence for cultural-heritage scenarios. *IEEE Intelligent Systems* 2008; 23(2): 34–41. doi: 10.1109/MIS.2008.24
74. Buratti G, Conte S, Rossi M. Artificial intelligency, big data and cultural heritage. In: *Representation Challenges. Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain.* Franco Angeli; 2021. pp. 29–34.
75. Abbattista F, Bordoni L, Semeraro G. Artificial Intelligence for cultural heritage and digital libraries. *Applied Artificial Intelligence* 2003; 17(8–9): 681–686. doi: 10.1080/713827258
76. Pavlidis G. From digital recording to advanced AI applications in archaeology and cultural heritage. In: Yosef EB, Jones Ian WN (editors). *“And in Length of Days Understanding”* (Job 12:12) *Essays on Archaeology in the Eastern Mediterranean and Beyond in Honor of Thomas E. Levy.* Springer Cham; 2023. pp. 1627–1656.
77. Garau C. From territory to smartphone: Smart fruition of cultural heritage for dynamic tourism development. *Planning Practice and Research* 2014; 29(3): 238–255. doi: 10.1080/02697459.2014.929837
78. Leshkevich T, Motozhanets A. Social perception of artificial intelligence and digitization of cultural heritage: Russian context. *Applied Sciences* 2022; 12(5): 2712. doi: 10.3390/app12052712

79. Pisoni G, Díaz-Rodríguez N, Gijlers H, Tonolli L. Human-centered artificial intelligence for designing accessible cultural heritage. *Applied Sciences* 2021; 11(2): 870. doi: 10.3390/app11020870
80. Díaz-Rodríguez N, Pisoni G. Accessible cultural heritage through explainable artificial intelligence. In: Proceedings of the Adjunct Publication of the 28th ACM Conference on User Modeling, Adaptation and Personalization; 14–17 July 2020; Genoa, Italy. pp. 317–324.
81. Ardissono L, Raptis GE, Mauro N. Special issue on AI and HCI methods and techniques for cultural heritage curation, exploration and fruition. *Applied Sciences* 2022; 12(19): 10118. doi: 10.3390/app121910118
82. Ranaldi L, Zanzotto FM. *Discover AI Knowledge to Preserve Cultural Heritage.* 2021. doi: 10.20944/preprints202109.0062.v1
83. Yurtsever A. Documentation of cultural heritage with technology: Evaluation through some architectural documentation examples and brief looking at AI (Artificial Intelligence). *Cultural Heritage and Science* 2023; 4(1): 31–39. doi: 10.58598/cuhes.1278735
84. Lee JH, Kim HK, Park CW. Studies on intelligent curation for the Korean traditional cultural heritage. In: Proceedings of the 2022 International Conference on Artificial Intelligence in Information and Communication (ICAIIC); 21–24 February 2022; Jeju Island, Korea. pp. 431–436.
85. Jennings NR, Wooldridge M. Applying agent technology. *Applied Artificial Intelligence an International Journal* 1995; 9(4): 357–369. doi: 10.1080/08839519508945480