Mutual Theory of Mind in Human-AI Interaction:

How Visual Feature Reflects What Users' Perceive About AI's Humanity?

Yiqing Zhou
Design for Interaction
Industrial Design of Engineering

Mutual Theory of Mind in Human-AI Interaction:

How Visual Feature Reflects What Users' Perceive About AI's Humanity?

By

Yiqing Zhou

Supervisor: Prof. Pan Wang
Mentor: Xun Zhang

Date: Aug 22, 2025

Contents

Abstract

This research investigates how artificial intelligence interprets and expresses the concept of humanity through visual art, and how people, in turn, perceive the humanity of AI based on these outputs. The project begins with the collection of approximately two thousand survey responses, in both Chinese and English, in which participants shared their personal views on the meaning of humanity. These responses were aggregated and used as input for an AI image generation process, resulting in a series of visual artworks intended to embody collective human perspectives.

To examine audience perceptions of these AI-generated works, semi-structured interviews were conducted with master-level design students who regularly engage with creative technologies. Participants were asked to interpret the artworks without prior knowledge that the images were created by AI, ensuring that responses were guided by the visual content itself rather than by preconceived notions of machine authorship.

The qualitative data were analyzed using Reflexive Thematic Analysis (RTA), through which recurring interpretive patterns were identified and grouped into six themes. To complement this analysis, participants also completed the Godspeed questionnaire, providing quantitative measures of anthropomorphism, animacy, and perceived intelligence. By correlating thematic interpretations with these perception scores, the study highlights how specific visual features—such as symbolic richness, emotional resonance, relational elements, or mechanistic qualities—influence the extent to which people attribute humanity to AI.

Findings suggest that judgments of AI's humanity are shaped less by technical fidelity than by the capacity of visual features to evoke emotions, cultural references, and symbolic associations. Participants who linked images to personal or collective human experiences tended to ascribe greater humanity to the AI, while those who focused on surface resemblance or structural flaws were more skeptical. This indicates that perceptions of AI's humanity depend on interpretive depth rather than superficial resemblance.

The project contributes to human–computer interaction research by demonstrating how qualitative interpretations of AI-generated art can reveal underlying frameworks through which people negotiate the boundary between human and machine creativity. It also provides practical insights for the design of AI systems in artistic and communicative domains, suggesting that fostering symbolic and emotional resonance may be more impactful than pursuing purely technical realism. Ultimately, this study positions AI art as a valuable lens for exploring evolving understandings of humanity and for guiding more responsible and human-centered development of creative AI systems.

1 Introduction

AI is no longer limited to routine automation but is now active in areas once viewed as uniquely human, including storytelling, music, and the visual arts. Such creative practices have historically been associated with human-specific capacities such as imagination, symbolic thought, and lived experience [20][27]. As AI encroaches on these creative territories, it raises new challenges for HCI: acceptance of AI depends not only on technical performance but also on whether people perceive the system as capable of human-like reasoning, vitality, and expression[37].

Recent studies illustrate this challenge. Even when AI-generated art achieves stylistic or technical parity with human-created work, people often evaluate it less favorably—attributing lower levels of creativity, intentional depth, or emotional meaning [3][22]. Generative models that produce novel compositions by deviating from learned styles further unsettle assumptions about what counts as "authentic" creativity [12]. These developments force a reconsideration of what it means for AI to possess "humanity" in the context of cultural and aesthetic production. Understanding how people interpret AI's human-likeness in visual art therefore offers a critical lens on the evolving boundaries of human–AI interaction.

Previous studies in HCI and HRI demonstrate that users judge AI not only on its outcomes but also on whether it conveys signs of intentionality, meaningfulness, or humanlike traits [13][47]. Humanlike cues—such as facial resemblance, gesture, or conversational tone—can make interactions feel smoother, fostering trust, engagement, and collaboration[8][35], while these cues may also evoke unease when users perceive AI as mimicking, rather than authentically embodying, human attributes[43]. This tension has been widely recognized in social robotics, conversational agents, and educational technologies, where the line between tool and partner is continually negotiated[14][36].

A growing body of work argues that these judgments emerge from interpretive processes: people do not passively consume system outputs but actively construct explanations of an AI's agency, intentionality, and creativity[25]. In creative domains, such as music and art, people's assessments hinge not only on the artifact's quality but also on perceived evidence of "mind" behind it[3][20]. This perspective draws inspiration from human–human interaction, where smooth collaboration relies on the ability to interpret others' inner states. This cognitive capacity—Theory of Mind (ToM)—refers to the ability to infer and understand the mental states of oneself and others, including emotions, intentions, preferences, and goals[1][6][16][17][44]. Building on this foundation, recent research has explored how ToM might be adapted for AI and robotic systems, leading to proposals Mutual Theory of Mind (MTOM)[40], framing human–AI interaction as a reciprocal process: users form inferences about an AI's mind while AI systems are designed to model and adapt to the user's states. This bidirectional modeling is argued to create interactions that feel more intuitive and responsive, supporting smoother and more effective collaboration between humans and AI[11][23][24][29][41]. MTOM framework has been primarily investigated in dialogical and embodied settings; however, its role in artifact-based encounters—such as how people interpret AI-generated artworks—remains underexplored. In these contexts, the interaction unfolds not through dialogue but through the exchange of artifacts: the AI interprets human input to generate a creative output, and humans, in turn, interpret that output as a window into the AI's mind. This dynamic shift is especially compelling in creative domains such as visual art, where audiences do more than assess aesthetic quality, often treating the artwork as a window into the 'mind' of its creator[3][21]. Yet we know little about how people interpret AI's 'humanity' in this context, or which specific visual features guide these judgments. Addressing this gap requires examining both the overall perception of AI's humanity through art, and the concrete features that shape such perceptions. To address this gap, we ask the following research questions:

• How do users perceive the humanity of AI based on visual artworks generated from collective human input?

• What visual features of AI-generated art shape users' perceptions of AI's humanity?

To answer these questions, we conducted a two-phase study. First, we collected over 2,000 open-ended responses about "humanity," which were synthesized by a generative AI model, ChatGPT 4o in this case, into a composite artwork. Second, we interviewed 21 design students, analyzing how they interpreted the artwork and how their responses related to perceptions of anthropomorphism, animacy, and perceived intelligence. By situating MTOM in an artifact-mediated creative context, our study extends prior applications of the framework and contributes new insights into how people make sense of AI's "humanity" through visual art. Our analysis showed that participants' perceptions of AI's humanity were nuanced and varied across dimensions. Anthropomorphism and animacy were rated near neutral, while perceived intelligence received the highest and most variable scores. Across all dimensions, interpretation mattered more than technical fidelity: participants who related visual features to emotions, relationships, or cultural symbolism attributed greater humanity to the AI, whereas others emphasized flaws, disconnection, or surface resemblance. Six cross-cutting themes highlight how features such as figures, expressions, relational cues, lighting, and symbolic references shaped judgments of anthropomorphism, animacy, and intelligence.

In the following sections, we will first review relevant work in AI perception studies, generative art, and HCI. We will then introduce our theoretical foundations, elaborate on our experimental design, and analyze the results in detail. Lastly, we will discuss the broader implications of our study, including the ethical considerations of AI-generated interpretations of humanity, the potential applications of AI in cultural and philosophical discourse, and future research directions that could further refine AI's ability to process and visually represent abstract human concepts.

2 Related Works

2.1 Theory of Mind in AI Era

In recent years, Theory of Mind (ToM) has emerged as a critical concept in the design and development of human-AI interaction systems. Originally a construct from cognitive science, Theory of Mind (ToM) describes the ability to recognize and infer mental states—including beliefs, desires, intentions, and emotions—in both oneself and others [1,17,40]. This skill supports core aspects of human social life, such as communication, collaboration, and shared decision-making [2,30, 40]. In the context of AI, enabling machines with analogous capabilities has become a key research objective as artificial agents are increasingly expected to participate in interactions that resemble human-human communication [40].

With the rise of artificial intelligence, ToM has evolved from a psychological theory into a computational framework that supports the modeling of mental states in machines. In the field of machine learning, researchers have begun exploring whether artificial agents can infer and represent the internal states of others. One prominent approach is the development of Machine Theory of Mind, where AI systems are trained to construct predictive models of other agents' goals, beliefs, and desires based solely on behavioral observations [32]. One example is ToMnet, a neural network architecture that applies meta-learning to generalize from observed behaviors and infer likely mental states of agents [32]. Although tested in controlled simulation environments, the system demonstrated that AI models can approximate ToM-like reasoning and predictions. More recently, GPT-4 has been shown to exhibit a surprisingly advanced level of Theory of Mind, further blurring the line between simulation and genuine mental modeling [4].

Beyond machine learning, ToM principles have also been integrated into the system architecture of human-robot collaboration frameworks. In such contexts, artificial agents are designed to not only monitor the physical state of the environment but also track and interpret the cognitive and emotional states of human partners [9]. This includes constructing internal models that simulate hypothetical mental states of users, especially in situations where observed behaviors deviate from expected plans [31]. Incorporating these models enables robots to better infer user intentions, adjust their assistance strategies, and maintain shared goals and beliefs throughout the interaction process [40].

Recent work has expanded beyond traditional ToM by proposing Mutual Theory of Mind (MToM), which frames human–AI interaction as a reciprocal process of understanding rather than a one-sided inference[40]. As human-AI interactions become increasingly complex and socially embedded, researchers have proposed the concept of Mutual Theory of Mind (MToM) to better account for the dynamic nature of these interactions. Unlike traditional Theory of Mind implementations where AI systems attempt to infer users' mental states, MToM introduces a reciprocal model in which both the human and the AI construct, refine, and adapt their understanding of each other's minds over the course of communication. This mutual interpretive process acknowledges that just as AI systems infer users' goals, preferences, or beliefs, users also develop mental models of the AI's capabilities, intentions, and limitations based on its behavior and feedback [40].

Because of this bidirectional nature of human-AI interaction, it is crucial for humans—not just machines—to engage in the interpretive loop. As AI systems construct models of human thoughts, behaviors, and values, we must also seek to understand how these systems perceive and represent humanity. This awareness can help inform more ethical and human-centered AI designs in the future, especially as AI plays an increasing role in shaping how we see ourselves and each other.

3 Method

In this section, the study methodology will be described.

3.1 Participants

This study engaged two separate participant groups aligned with the two phases of the research: (1) a large-scale survey group whose written contributions served as input for the AI-generated visual stimulus, and (2) an interview group who reflected on the stimulus to examine perceptions of AI's humanity.

3.1.1 Survey participants.

We collected responses from 2,119 individuals through online recruitment platforms, without applying pre-selection criteria beyond being at least 18 years old. This approach was chosen to capture a broad and heterogeneous range of perspectives on "humanity," allowing the AI model to be trained on varied cultural and experiential inputs rather than a filtered subset. The resulting sample was 52.47 % female, with ages ranging from 18 to 74 years (63.6 % between 25–45 years old). Participants reported a wide geographic distribution, including Australia, New Zealand, India, Israel, the United Kingdom, China, South Africa, Morocco, Kenya, Canada, Mexico, and multiple countries in the European Economic Area. Regarding attitudes toward AI, 53.37% expressed a positive view, and 39.55% reported using AI tools a few times per week. The survey consisted of open-ended prompts about participants' understanding of "humanity" and optional demographic questions. Compensation was provided in line with the norms of each recruitment platform.

3.1.2 Interview participants.

For the second phase, we recruited 21 participants from the Design School at Delft University of Technology in the Netherlands. All participants were enrolled or graduated in a Master's-level design program. They were selected for their familiarity with interpreting visual media. None had taken part in the survey to avoid prior exposure to the dataset. Participants ranged in age from 22 to 26 years and represented a mix of international backgrounds, including Spain, the Netherlands, Turkey, Sweden, India, Thailand, and China. All reported regular engagement with visual or creative work, and most had prior experience with AI-generated images. Recruitment was conducted via direct invitations and program group messages. Participants did not receive compensation for their involvement in the interviews.

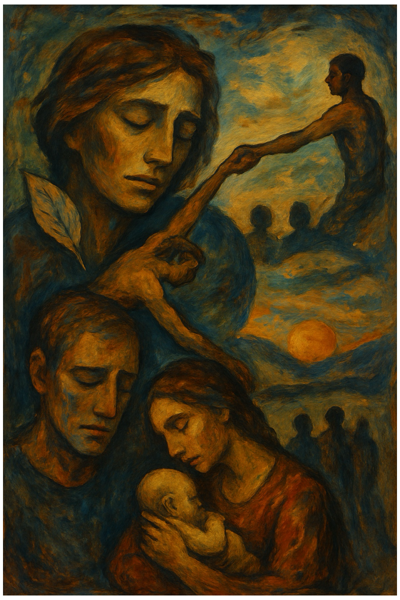
3.2 Procedure

The study was conducted in two sequential phases. In the first phase, survey participants completed an online questionnaire consisting of open-ended prompts about their understanding of "humanity," along with optional demographic questions. No constraints were placed on response length or content to allow participants to articulate their perspectives freely. Responses were collected over a two-week period. To reduce the likelihood of AI-generated or low-quality responses, we removed extremely long answers submitted in unusually short times, as well as nonsensical or random text. All other responses were kept in their original form to minimise researcher influence on the raw data.

Then survey responses were compiled and entered into ChatGPT-4o, using only answers to the "understanding humanity" prompt. A concise, standardised instruction described the research context and requested the generation of a visual artwork representing humanity, based on the provided data and the model's own interpretation. The model produced a single coherent image intended to synthesise the collective perspectives. Minor post-processing adjustments removed textual artefacts and any explicit indicators of AI generation. During pilot testing, multiple generations were produced; within the same conversation, results were visually similar, but across different conversations, they varied noticeably. For consistency, we selected the first image generated as the stimulus for the second phase.

In the second phase, interview participants took part in semi-structured, individual sessions (offline or via Microsoft Teams) lasting 25–45 minutes. Each interview began with warm-up questions to establish a baseline for discussing the concept of humanity. Participants were then shown the visual artwork without being told it was AI-generated and asked open-ended questions about their impressions, interpretations, perceived meaning, and notable visual elements. After these discussions, the interviewer disclosed that the image had been generated by an AI using responses from a large, diverse sample. Follow-up questions prompted detailed visual analysis and encouraged participants to reflect on any changes in interpretation after the disclosure. Finally, participants completed the Godspeed questionnaire measuring anthropomorphism, animacy, and perceived intelligence.

All sessions were audio- or video-recorded with participant consent for later transcription and analysis.



3.3 Data Analysis

Interview data were analysed using an inductive coding approach focused on identifying and categorising visual features described by participants. Both authors read all interview transcripts in full. Coding categories were developed iteratively: initial transcripts were reviewed to identify descriptive codes for specific visual elements or compositional aspects mentioned by participants, and the coding framework was refined as additional transcripts were analysed. This process continued until all transcripts were coded and no new categories emerged.

After coding was complete, Godspeed questionnaire scores for anthropomorphism, animacy, and perceived intelligence were calculated for each participant, along with the average and median scores for each dimension. Participants whose scores exceeded the dimension's average were classified as "high" for that dimension, and those whose scores fell below the average were classified as "low." For each dimension, coded data from the "high" and "low" groups were compared to identify patterns in visual features associated with higher or lower ratings. This comparative analysis examined, for example, which visual elements were more frequently described by participants with high anthropomorphism ratings versus those with low ratings. Patterns were developed collaboratively, with both authors cross-checking codes and interpretations to ensureconsistency and reduce individual bias.

4 Findings

The primary aim of this research was to examine how people perceive the humanity of AI through an AI-generated visual artwork, and to identify the visual features that shape these perception.

4.1 Perceived Anthropomorphism, Animacy, and Intelligence

4.1.1 Anthropomorphism.

Anthropomorphism, the human inclination to attribute traits such as emotions, intentions, or personalities to nonhuman entities, has been widely shown to affect perceptions of AI [26, 45]. In human–AI interaction research, anthropomorphic cues, such as human-like faces or other human-resembling elements, have been shown to enhance perceptions of trust and rapport between humans and AI [7, 15 , 41]. At the same time, prior work has noted that when nonhuman agents, such as AI systems or humanoid robots, closely mimic human capabilities without being human, they can elicit discomfort or even resistance [34]. This response is often tied to concerns about replacing or diminishing qualities perceived as uniquely human. In the context of our study—where the AI generated a piece of visual art, a domain traditionally viewed as requiring human creativity and understanding—such concerns become especially salient [18]. Thus, it becomes essential to distinguish which specific visual features associated with anthropomorphism elicit positive responses, in order to better inform the design of AI systems and foster more effective human–AI interactions.

On the Godspeed scale, the participants' ratings of anthropomorphism averaged 2.92, with a median of 3. This suggests a largely neutral position towards AI's human-like qualities. Eighteen of the 21 participants scored between 2.25 and 3.5, one participant gave a higher score of 4, and two rated AI particularly low at 1 and 2. These distributionsindicate that perceptions of anthropomorphism were generally modest, with only a few participants leaning strong positive or negative.

The interview data suggest that attitudes toward AI in general influenced how the participants approached anthropomorphism. The two participants who rated the lowest in this dimension expressed strong resistance to AI-generated art. One (Jiang) explained: "Because it does not have heart. It does not have a mind. It just does what we tell it to do." This participant further emphasised the uniqueness of human creativity: "I believe art is something expressed by humans, something only humans can create, because we devote our thoughts, emotions, and feelings into it. That makes the final piece a work of art. What AI does is just basically like programming it." Similarly, another (Hashita) contrasted human and AI creative processes: "When you start creating, even if you have a clear picture, your hand changes it as you paint, and that becomes yours. That process, AI doesn't have." For these participants, the absence of embodied process and emotional investment meant that AI outputs could not be considered art. By contrast, the participant who rated anthropomorphism highest (score of 4) adopted a more optimistic view of AI's creative potential. Sia described the artwork as capturing collective perspectives effectively: "It did really well on this work. From this picture, I can see the depiction from those 2000 people." This participant framed the AI not as replacing human uniqueness but as extending possibilities for representing shared understandings of humanity.

Overall, while most participants remained neutral, the few who deviated toward low or high scores illustrate how anthropomorphism ratings were closely tied to participants' broader attitudes toward AI. Those skeptical of AI in creative domains rejected the possibility of attributing human-like qualities, whereas those open to AI's role in art were more willing to acknowledge anthropomorphic features.

4.1.2 Animacy.

Animacy, understood as the perception of vitality or "livingness" in non-human forms, is a key factor in HCI research[19]. Cognitive neuroscience indicates that animacy can be registered early in visual processing, with cues like curvature or implied movement quickly triggering animacy-related interpretations [42]. In interactive contexts, and particularly within creative HRI, animacy has been shown to correlate positively with user engagement and perceptions of agency; for example, physical movement in robots reliably improves user animacy ratings [33]. Although our study examined static visual artwork, the same principle applies: visual features that imply vitality, intentionality, or motion can evoke stronger impressions of animacy even in the absence of actual movement. As such, animacy plays a critical role in shaping how people perceive AI's humanity through generated visual art.

Participants' ratings of animacy averaged 2.97, with a median of 3. All scores fell within a relatively narrow range (2.25–3.67), indicating that perceptions of the AI's vitality or "sense of life" were generally neutral. No participants perceived the AI as highly animate, but neither did they dismiss it as entirely lifeless. Interview data suggest that this neutrality reflected how participants used visual cues in the artwork to infer the AI's underlying vitality. Many noted that the human figures and group arrangements implied interaction or motion, which they interpreted as evidence of the AI's ability to capture aspects of life. For example, one participant commented, "another interaction is showing behind two people holding hands of each other. But I cannot see if this is like a dancing posture, or it's more like one person save another person from the hole."(florence) Another observed, "This group of three people at the back, maybe they are friends, I can associate with, they are looking at the sunset." (florence) These impressions suggest that participants attributed a baseline level of animacy to the AI, insofar as it could reproduce gestures and relationships that conveyed aliveness.

At the same time, participants often described these depictions as staged or classical in style—more reminiscent of historical art traditions than of dynamic, lived experience. Rather than inspiring a sense of spontaneity or creative vitality, the AI's portrayal of movement was seen as formal, almost static. As one participant reflected, "like it is use some typical things classical symbols like the madam and the Jesus. this image is just very typical." Another added, "I think this two guys holding hands is also iconic, the pose looks like the creation of adam."(chang) In this way, the AI was recognised as capable of replicating visual conventions of aliveness without necessarily being perceived as possessing animacy in a deeper or more creative sense.

Overall, participants acknowledged that the AI demonstrated a certain baseline capacity to depict lifelike qualities, but the lack of spontaneity or originality in these portrayals kept perceptions of animacy clustered around neutral.

4.1.3 Intelligence.

Perceived intelligence refers to how knowledgeable, capable, or deliberate an AI is judged to be in its operations. With increasing technical sophistication, users often expect AI to competently interpret requests and generate outputs that correspond with human goals [46]. Yet, studies consistently show a gap between user expectations and the actual performance of current AI systems, which shapes how intelligence is perceived [38]. This makes perceived intelligence not only a marker of current user experience but also an essential consideration for guiding AI development and adoption [41]. In our study, perceived intelligence
reflects how participants judged the AI's ability to synthesize diverse human input into a coherent visual narrative. Rather than focusing solely on the artwork as an aesthetic object, our analysis examined whether features such as compositional structure, symbolic layering, or interpretive depth conveyed to participants that the AI demonstrated intelligence—and how those impressions varied across ratings.

Participants' ratings of perceived intelligence averaged 3.37, with a median of 3.4—the highest among the three Godspeed dimensions. Scores ranged from 1.8 to 4.4, making this the broadest spread across dimensions. This suggests that while most participants regarded the AI as moderately capable, there was considerable divergence in how much intelligence they attributed to it.
One explanation for this higher mean and wider variance lies in participants' expectations of what AI should be able to do. Many assumed that an AI system ought to be "smart" enough to process large amounts of information and generate coherent outcomes. For these participants, the act of transforming thousands of human responses into a single visual representation was itself seen as a sign of intelligence. As one participant explained, "I think it can understand something. . . it can understand the people's opinion on humanity." (Kim) While some participants questioned whether this ability reflected deeper forms of intelligence. They described the AI's output as lacking genuine reasoning or interpretive depth, characterizing its grasp of humanity as relatively shallow or partial.

This perspective suggests that while participants acknowledged the capacity of the system to organize and present information, they remained skeptical about attributed more profound cognitive understanding to it. In general, perceived intelligence received the most generous assessments compared to anthropomorphism and animacy, but also provoked the widest divergence, reflecting the tension between recognizing computational capability and questioning the presence of deeper human-like intelligence.

4.2 Visual Features Shaping Perceptions of AI's Humanity

While the Godspeed ratings provide an overview of how participants assessed the AI in terms of anthropomorphism, animacy, and perceived intelligence, the interviews offer deeper insight into what aspects of the visual stimulus shaped these perceptions. Across transcripts, participants consistently identified a range of visual features. To understand how these elements informed judgments of the AI's humanity, we coded participants' descriptions and compared how high and low-scoring groups within each dimension interpreted the same features. This analysis reveals that the presence of visual features was often shared across groups, but their meanings and implications diverged significantly depending on participants' evaluations of anthropomorphism, animacy, or intelligence.

4.2.1 Anthropomorphism.

The analysis revealed that participants often focused on three prominent visual features—human figures and facial expressions, interpersonal connections, and symbolic roles such as family or caregiving—but diverged in how they interpreted their meaning. Three themes capture how these interpretations shaped attributions of anthropomorphism: (1) whether human figures and expressions were linked to real-life or emotional meaning, (2) the extent to which cultural or stylistic references were seen as markers of humanity, and (3) whether atmosphere created through lighting and color enhanced or undermined the sense of humanity.

Interpreting Human Elements Through Real-Life and Emotional Meaning.

A central factor influencing perceptions of anthropomorphism was whether participants connected the depicted figures and expressions to real-life experiences or emotions. For those who rated anthropomorphism higher, humanlike qualities were not judged solely by technical precision. Even when facial expressions, gestures, or group arrangements appeared imperfect, these participants still described them as communicating recognisable emotions or relationships. For example, a participant perceived sadness in the figures and related this to human experiences of loss, while simultaneously noting imperfections in how the characters were connected: "Everybody looks a little bit sad. They look like they're helping each other, but it's like everybody feels disconnected from each other. They all kind of have the same sullen face expressions, they're all closing their eyes. They're all looking down, but not looking at each other, not really looking at anything. It looks like they're a little bit lost in life." (Elin) Another participant described a similar tension—relating the figures to a narrative of departure or death while acknowledging the lack of expressive nuance: "Feel the sadness, because I think everyone close their eyes and there is no emotion on their face, and also their eyebrow is a little down, I see some people at the back. So I feel like someone will leave, or they all go to the heaven." (Kim)

By contrast, participants with lower ratings recognised the same features but treated them as detached from lived human meaning. Rather than constructing stories or emotions from what they observed, they described figures in a more literal and fragmented way. For example, one participant noted: "I see there's a bunch of people, for example, the two men like holding hands. You have the image of the family. They all look very sad. And there's a woman at the top corner, which I really don't understand what she's doing there." (Alba) Similarly, another participant acknowledged signs of sadness but framed them as compositional rather than meaningful: "It express maybe sadness, despair, being disappointed, because of facial expression is no eye contact. Everybody in their own zone. There is no eye contact and there're a lot of people, but everyone in their own space in the image." (Sai)

Together, these accounts highlight that anthropomorphism was not determined by the mere presence of humanlike figures, but by whether participants interpreted those figures as carriers of emotional or relational meaning.

Cultural Familiarity and Art-Historical References as Markers of Humanity.

Another theme shaping perceptions of anthropomorphism was whether participants associated the arrangement of figures, their poses, or the overall composition with familiar cultural and art-historical traditions. Several participants drew comparisons to classical or religious artworks, noting similarities in posture, grouping, or symbolic arrangements. For participants who rated anthropomorphism higher, such resemblances were interpreted as signs of humanlike intentionality and expression. These participants often connected stylistic echoes to broader themes of human experience. For example, one participant linked the work to Van Gogh's The Potato Eaters, interpreting the visual reference as reflective of hardship and resilience: "It reminds me of a painting of Van Gogh, The Potato Eaters—people are from a very rural area, or a low-income family. So this painting will give me the idea that people are thinking about humanity, even if they are in a not-so-good situation, like a tragedy." (Xinchen)

By contrast, participants with lower anthropomorphism ratings acknowledged similar stylistic references but did not extend them to deeper meaning. Instead, they described the resemblance as surface-level mimicry of known artistic traditions. One participant remarked: "It makes me resonate with some classical oil painting pieces. The lower part is like Santa Maria and his son. And the upper part, the structure looks very Matisse." (Wu) Another noted that while the image resembled religious works, the reference felt shallow and limiting in its expression of humanity: "I think ChatGPT here maybe took some religious paintings for reference, and it thinks that can represent humanity in a sense. But I think it's also a bit limiting in a way." (jiang)

Together, these accounts highlight that cultural and stylistic familiarity offered a potential basis for anthropomorphism, but participants diverged in whether they interpreted these references as evidence of humanlike intentionality or as superficial imitation.

Human-Centered Atmosphere Through Lighting and Color.

A final theme concerned the atmosphere created through lighting and color. While both high- and low-rating participants noticed these features, only those who rated anthropomorphism higher explicitly linked them to human presence and emotion. High scorers described how warm tones and soft illumination enhanced the image's humanlike qualities by conveying mood and affect beyond the figures themselves. For some, the lighting suggested both sadness and hope, framing the scene within an emotional narrative that resonated with human experience: "There's light make me feel like the sunset lighting their faces. And it gives me feeling like it's very sad, but we still have hope, yeah, but this hope is very, very soft." (Sia) Others interpreted the palette as creating a sense of warmth and resilience in the face of hardship: "I think this one is more smooth or more warm, because this one, I feels like everyone inside the image is super calm or a little bit sad even. . . maybe they have lost something, or go through some traumatic or like war, but then they became more treasure about people around them." (Wzy)

In contrast, participants with lower anthropomorphism scores referred to lighting and color only as stylistic aspects of the image (e.g., tone, palette, art style) without connecting them to human figures or emotional meaning. As such, lighting and color contributed uniquely to higher ratings of anthropomorphism by reinforcing interpretations of human-centered atmosphere and intentionality.

4.2.2 Animacy.

In the dimension of animacy, participants across both high and low ratings frequently described three types of visual features: literal motion and process cues (such as holding hands, hugging, or standing together), relational and interactional actions (for instance, depictions of families or groups), and temporal or event cues (such as the sunset suggesting an ending or transition). These elements were consistently identified, but participants diverged in how they interpreted them. From our analysis, one central theme emerged: whether such features were seen as static, symbolic motifs or as dynamic processes unfolding in time. This interpretive difference shaped whether the AI's output was perceived as lifelike or inert.

Interpreting Actions as Static Motifs vs. Dynamic Processes.

Participants with lower animacy ratings often acknowledged actions and gestures but interpreted them as symbolic conventions, resembling familiar motifs from classical or religious artworks. Rather than viewing the gestures as part of a living process, these participants described them as recognisable but inert representations. For example, one participant compared the mother and baby to recurring artistic tropes: "I think this mom and baby, they're symbolic. There are so many pictures like this part, really classical. Maybe this mountain and sunrise or sunset, and also hold this hand, all very the classic ones." (Jade) Another similarly highlighted the symbolic quality of the actions—connection and beginnings—without attributing vitality or temporal flow: "I think it's meant to be touching, like how humans are connected, and then they're holding hands, and you have a baby because starting of life." (Dimpy)

By contrast, participants with higher animacy ratings interpreted the same cues as dynamic processes situated in time, often constructing narratives of unfolding events or transitions. One participant described the scene as capturing an ending, resonant with the flow of daily life: "It looks [like] everyone [is] looking at a sunset and waiting for the world's end, or waiting [for] a day [to] end with their hard working [and] their frustration in their life." (Florence) Another participant described the composition sequentially, interpreting the gestures and arrangements as a story that progressed across the canvas: "Because I see the hands connect together, and also it's show me the flow I can see through the picture. So first I see the woman on the left top, and then I see the story just below. Then I saw a family, but they have no emotion on the face with the baby. So I'm thinking that maybe the baby is dying or something. And then the last thing I saw is people at background. So I feel they will leave too. They are going to the place far away from them." (Kim)

These contrasting accounts suggest that animacy was not determined by the mere presence of gestures, interactions, or temporal markers, but by how participants interpreted these features—either as static, symbolic references or as dynamic, life-like processes embedded in time.

4.2.3 Perceived Intelligence.

In the dimension of perceived intelligence, participants across both high and low ratings frequently remarked on three types of visual features: narrative logic and story integration (whether the depicted elements could be woven into a coherent whole), purposeful composition and reasoning (how spatial arrangements and proportions reflected intent or logic), and symbolism and metaphor recognition (the depth of meaning conveyed through representational elements). These features were consistently observed, but participants diverged in how they interpreted them. From our analysis, two themes emerged: whether fragmented elements were understood as part of a unified narrative or disconnected scenes, and whether compositional choices were interpreted as innovative reasoning or confusing flaws.

Narrative Coherence as a Marker of Intelligence .

Participants often judged the AI's intelligence by its ability to construct a meaningful story from multiple symbolic elements. For those with higher ratings, the figures and gestures were interpreted as interconnected, pointing toward broader human themes. One participant described the infant and hand gestures as reflecting the arc of human life and relationships: "I feel like the infants. It's expressing the new life, the very start of humanity. And there are two figures holding hands with each other. It's like relationship between humans, we come to earth, build relationships with each other, together, and then we break up so and especially the gesture between them." (Sia) Others read the image as an imaginative reconstruction of memory or thought, suggesting that even disjointed elements could be integrated into a narrative whole: "The whole story is a vague story. . . . It's not depicting the real scene, but . . . someone's illusion. . . . His mind reflecting on his whole life and he was born, and his father or something, and they [go] to the lake to see the sunset." (Chang)

By contrast, participants with lower ratings highlighted fragmentation. They acknowledged individual vignettes but perceived them as disconnected, limiting the artwork's capacity to reflect intelligent synthesis. As one participant noted: "I also feel they are different stories in this picture, and that all of them are not connected with each other. For me the family is one story. The two men is another story, then the three people on the corner . . . but . . . I don't see a direct connection." (Alba)

Here, perceived intelligence was tied not to the presence of symbolic content itself but to whether those symbols could be integrated into a coherent narrative structure.

Interpreting Composition as Creative Reasoning vs. Technical Flaw.

Another factor shaping perceptions of intelligence was how participants interpreted the AI's approach to spatial composition. While both high and low scorers noticed unconventional layering and proportions, they diverged in how they read these deviations. Participants with higher ratings often treated them as evidence of creative reasoning, suggesting that breaking with conventional artistic rules could reflect a form of intentionality. As one participant explained: "First, I think the way it layers . . . foreground, background, middle ground. . . . If it's human, maybe they will follow the normal rules. . . . But AI mixed the normal rules. This is maybe unexpected, but at the same time, it's a bit strange. It can be deep or it can be confusing." (Sai)

In contrast, participants with lower ratings described the same compositional irregularities as technical flaws that undermined coherence. Shifts in scale and multiple horizons were interpreted as mistakes, leaving the scene disjointed rather than intelligent: "I mean, there's a lot of composition wise that I don't completely understand. . . . Two horizons . . . large figures . . . proportion becomes a little bit off. . . . It feels a bit more like a dream . . . almost a god-like figure . . . but then . . . the bottom part . . . makes sense." (Elin) Others criticized the image for being overly literal or unrefined, suggesting a lack of depth in reasoning compared to human artistic process: "I think it has a shallow expression, because I think it's too violent. It's too straightforward . . . it lacks a bit of rethinking and refining in [an] artist's mind." (Cpc)

Thus, compositional choices were not inherently tied to perceptions of intelligence; rather, it was participants' interpretation—whether as signs of creative reasoning or as technical missteps—that shaped their judgments.

Bibliography

[1] Simon Baron-Cohen. 1999. The evolution of a theory of mind. na.

[2] Simon Baron-Cohen, Alan M Leslie, and Uta Frith. 1985. Does the autistic child have a “theory of mind”? Cognition 21, 1 (1985), 37–46.

[3] Lucas Bellaiche, Rohin Shahi, Martin Harry Turpin, Anya Ragnhildstveit, Shawn Sprockett, Nathaniel Barr, Alexander Christensen, and Paul Seli.

2023. Humans versus AI: whether and why we prefer human-created compared to AI-created artwork. Cognitive research: principles and implications 8, 1 (2023), 42.

[4] Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, et al. 2024. Sparks of artificial general intelligence: Early experiments with gpt-4. arXiv 2023. arXiv preprint arXiv:2303.12712 10 (2024).

[5] Jenna Burrell. 2016. How the machine ‘thinks’: Understanding opacity in machine learning algorithms. Big data & society 3, 1 (2016), 2053951715622512.

[6] Peter Carruthers and Peter K Smith. 1996. Theories of theories of mind. Cambridge university press.

[7] Justine Cassell and Timothy Bickmore. 2000. External manifestations of trustworthiness in the interface. Commun. ACM 43, 12 (2000), 50–56.

[8] Lara Christoforakos, Alessio Gallucci, Tinatini Surmava-Große, Daniel Ullrich, and Sarah Diefenbach. 2021. Can robots earn our trust the same way humans do? A systematic exploration of competence, warmth, and anthropomorphism as determinants of trust development in HRI. Frontiers in Robotics and AI 8 (2021), 640444.

[9] Sandra Devin and Rachid Alami. 2016. An implemented theory of mind to improve human-robot shared plans execution. In 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, 319–326.

[10] Upol Ehsan, Q Vera Liao, Michael Muller, Mark O Riedl, and Justin D Weisz. 2021. Expanding explainability: Towards social transparency in ai systems. In Proceedings of the 2021 CHI conference on human factors in computing systems. 1–19.

[11] Bobbie Eicher, Kathryn Cunningham, Sydni Peterson Marissa Gonzales, and Ashok Goel. 2017. Toward mutual theory of mind as a foundation for co-creation. In International Conference on Computational Creativity, Co-Creation Workshop.

[12] Ahmed Elgammal, Bingchen Liu, Mohamed Elhoseiny, and Marian Mazzone. 2017. Can: Creative adversarial networks, generating" art" by learning about styles and deviating from style norms. arXiv preprint arXiv:1706.07068 (2017).

[13] Nicholas Epley, Adam Waytz, and John T Cacioppo. 2007. On seeing human: a three-factor theory of anthropomorphism. Psychological review 114, 4 (2007), 864.

[14] Jodi Forlizzi, Carl DiSalvo, and Francine Gemperle. 2004. Assistive robotics and an ecology of elders living independently in their homes. Human–Computer Interaction 19, 1-2 (2004), 25–59.

[15] Eun Go and S Shyam Sundar. 2019. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. Computers in human behavior 97 (2019), 304–316.

[16] Alvin I Goldman et al. 2012. Theory of mind. Vol. 1. Oxford handbook of philosophy and cognitive science.

[17] Alison Gopnik and Henry M Wellman. 1992. Why the child's theory of mind really is a theory. (1992).

[18] Simone Grassini and Mika Koivisto. 2024. Understanding how personality traits, experiences, and attitudes shape negative bias toward AI-generated artworks. Scientific Reports 14, 1 (2024), 4113.

[19] Kashyap Haresamudram, Nena Van As, and Stefan Larsson. 2025. Tasks Over Traits: User perception of humanlike features in goal-oriented chatbots. International Journal of Human–Computer Interaction (2025), 1–19.

[20] Aaron Hertzmann. 2018. Can computers create art?. In Arts, Vol. 7. MDPI, 18.

[21] C Blaine Horton Jr, Michael W White, and Sheena S Iyengar. 2023. Bias against AI art can enhance perceptions of human creativity. Scientific reports 13, 1 (2023), 19001.

[22] YeiBeech Jang. 2023. Anthropomorphism, perceived learning for creation, and growth creative mindset as predictors of acceptance toward artificial intelligence creativity. Social Behavior and Personality: an international journal 51, 8 (2023), 1–9.

[23] Séverin Lemaignan and Pierre Dillenbourg. 2015. Mutual modelling in robotics: Inspirations for the next steps. In Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. 303–310.

[24] Shuhong Lin, Boaz Keysar, and Nicholas Epley. 2010. Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. Journal of experimental social psychology 46, 3 (2010), 551–556.

[25] Ewa Luger and Abigail Sellen. 2016. " Like Having a Really Bad PA" The Gulf between User Expectation and Experience of Conversational Agents. In Proceedings of the 2016 CHI conference on human factors in computing systems. 5286–5297.

[26] Sara Moussawi and Marios Koufaris. 2019. Perceived intelligence and perceived anthropomorphism of personal intelligent agents: Scale development and validation. (2019).

[27] Jonas Oppenlaender. 2022. The creativity of text-to-image generation. In Proceedings of the 25th international academic mindtrek conference. 192–202.

[28] Dino Pedreschi, Fosca Giannotti, Riccardo Guidotti, Anna Monreale, Salvatore Ruggieri, and Franco Turini. 2019. Meaningful explanations of black box AI decision systems. In Proceedings of the AAAI conference on artificial intelligence, Vol. 33. 9780–9784.

[29] Christopher Peters. 2005. Foundations of an agent theory of mind model for conversation initiation in virtual environments. Virtual Social Agents 163 (2005).

[30] David Premack and Guy Woodruff. 1978. Does the chimpanzee have a theory of mind? Behavioral and brain sciences 1, 4 (1978), 515–526.

[31] David V Pynadath and Stacy C Marsella. 2005. PsychSim: Modeling theory of mind with decision-theoretic agents. In IJCAI, Vol. 5. 1181–1186.

[32] Neil Rabinowitz, Frank Perbet, Francis Song, Chiyuan Zhang, SM Ali Eslami, and Matthew Botvinick. 2018. Machine theory of mind. In International conference on machine learning. PMLR, 4218–4227.

[33] Eduardo Benítez Sandoval, Ricardo Sosa, Massimiliano Cappuccio, and Tomasz Bednarz. 2022. Human–robot creative interactions: Exploring creativity in artificial agents using a storytelling game. Frontiers in Robotics and AI 9 (2022), 695162.

[34] Bernd Schmitt. 2020. Speciesism: an obstacle to AI and robot adoption. Marketing Letters 31, 1 (2020), 3–6.

[35] Tim Schreiter, Lucas Morillo-Mendez, Ravi T Chadalavada, Andrey Rudenko, Erik Alexander Billing, and Achim J Lilienthal. 2022. The effect of anthropomorphism on trust in an industrial human-robot interaction. arXiv preprint arXiv:2208.14637 (2022).

[36] William Seymour and Max Van Kleek. 2021. Exploring interactions between trust, anthropomorphism, and relationship development in voice assistants. Proceedings of the ACM on Human-Computer Interaction 5, CSCW2 (2021), 1–16.

[37] Ben Shneiderman. 2020. Human-centered artificial intelligence: Reliable, safe & trustworthy. International Journal of Human–Computer Interaction 36, 6 (2020), 495–504.

[38] Inara Tusseyeva, Anara Sandygulova, and Matteo Rubagotti. 2024. Perceived intelligence in human-robot interaction–a review. IEEE Access (2024).

[39] Warren J Von Eschenbach. 2021. Transparency and the black box problem: Why we do not trust AI. Philosophy & Technology 34, 4 (2021), 1607–1622.

[40] Qiaosi Wang and Ashok K Goel. 2022. Mutual theory of mind for human-AI communication. arXiv preprint arXiv:2210.03842 (2022).

[41] Qiaosi Wang, Koustuv Saha, Eric Gregori, David Joyner, and Ashok Goel. 2021. Towards mutual theory of mind in human-ai interaction: How language reflects what students perceive about a virtual teaching assistant. In Proceedings of the 2021 CHI conference on human factors in computing systems. 1–14.

[42] Ruosi Wang, Daniel Janini, and Talia Konkle. 2022. Mid-level feature differences support early animacy and object size distinctions: Evidence from electroencephalography decoding. Journal of Cognitive Neuroscience 34, 9 (2022), 1670–1680.

[43] Adam Waytz, John Cacioppo, and Nicholas Epley. 2010. Who sees human? The stability and importance of individual differences in anthropomorphism. Perspectives on psychological science 5, 3 (2010), 219–232.

[44] Henry M Wellman, Susan Carey, Lila Gleitman, Elissa L Newport, and Elizabeth S Spelke. 1990. The child's theory of mind. The MIT Press.

[45] Xintong Yao and Yipeng Xi. 2025. From Assistants to Digital Beings: Exploring Anthropomorphism, Humanness Perception, and AI Anxiety in Large-Language-Model Chatbots. Social Science Computer Review (2025), 08944393251354976.

[46] Jennifer Zamora. 2017. I'm sorry, dave, i'm afraid i can't do that: Chatbot perception and expectations. In Proceedings of the 5th international conference on human agent interaction. 253–260.

[47] Jakub Złotowski, Diane Proudfoot, Kumar Yogeeswaran, and Christoph Bartneck. 2015. Anthropomorphism: opportunities and challenges in human–robot interaction. International journal of social robotics 7, 3 (2015), 347–360.

Appendix A: Godspeed Scale

Please rate your impression of the AI on these scales:

Anthropomorphism

Fake	1	2	3	4	5	Natural
Machinelike	1	2	3	4	5	Humanlike
Unconscious	1	2	3	4	5	Conscious
Artificial	1	2	3	4	5	Lifelike

Animacy

Dead	1	2	3	4	5	Alive
Stagnant	1	2	3	4	5	Lively
Mechanical	1	2	3	4	5	Organic
Artificial	1	2	3	4	5	Lifelike
Inert	1	2	3	4	5	Interactive
Apathetic	1	2	3	4	5	Responsive

Perceived Intelligence

Incompetent	1	2	3	4	5	Competent
Ignorant	1	2	3	4	5	Knowledgeable
Irresponsible	1	2	3	4	5	Responsible
UnIntelligent	1	2	3	4	5	Intelligent
Foolish	1	2	3	4	5	Sensible

Appendix B: Interview Question

☞ **Intro & Familiarity**

1. Have you seen or used AI-generated artwork before? What was your experience?

2. In your own words, what does "humanity" mean to you?
(Follow-up: Are there particular emotions, relationships, or values you associate with it?)

3. If you were to create a piece of art to express "humanity," what would it look like? What might it include?
(This establishes their baseline artistic expectation.)

☞ **Show the visual**

☞ **Immediate Response to the Artwork**

4. What is your first impression of this artwork? How does it make you feel? (Follow-up: Is there any emotional response or mood you associate with it?)

5. What do you think this artwork is trying to express or communicate?
(Encourage reflection on themes, symbols, complexity.)

6. Are there any specific parts of the image that captured your attention? Why?

☞ **Tell it's from 2000 survey result**
Perceived AI Understanding & Intentionality

7. Now that you know this was generated by AI based on 2000 people's thoughts on humanity — what do you think the AI understood well, and what do you think it failed to understand? Why?

8. Do you feel this image shows a deep or shallow understanding of humanity? What makes you say that?

9. Did the artwork feel emotionally or cognitively rich to you? Why or why not?

☞ Visual Features & meaning-making

10. Are there visual elements — colors, forms, structure, details — that you found especially expressive or symbolic? Can you describe them?

11. Did you sense any narrative, relationship, or emotional interaction happening in the image?

12. Were there aspects of the image that felt confusing, alien, or emotionally flat? If so, which and why?

13. Are there any other visual features that you would like to comment on?

☞ Reflection on AI & Humanity

14. After seeing this, do you think AI can ever truly reflect or represent humanity? Why or why not?

15. What aspects do you think the AI would succeed and/or fail? why?

16. What would you want AI developers to focus on if their goal is to make AI better at expressing human ideas through art?

☞ Godspeed Scales

Yiqing Zhou
Design for Interaction
Industrial Design of Engineering

CHECK ON STUDY PROGRESS

To be filled in **by SSC E&SA** (Shared Service Centre, Education & Student Affairs), after approval of the project brief by the chair.
The study progress will be checked for a 2nd time just before the green light meeting.

Master electives no. of EC accumulated in total ______ EC

Of which, taking conditional requirements into account, can be part of the exam programme ______ EC

★	**YES**	all 1st year master courses passed
	NO	missing 1st year courses

Comments:

Sign for approval (SSC E&SA)

Name ______ Date ______ Signature ______

APPROVAL OF BOARD OF EXAMINERS IDE on SUPERVISORY TEAM -> to be checked and filled in by IDE's Board of Examiners

Does the composition of the Supervisory Team comply with regulations?

YES	★	Supervisory Team approved
NO		Supervisory Team not approved

Comments:

Based on study progress, students is ...

★	**ALLOWED** to start the graduation project
	NOT allowed to start the graduation project

Comments:

Sign for approval (BoEx)

Name ______ Date ______ Signature ______

Personal Project Brief – IDE Master Graduation Project

Name student ____________________ **Student number** ____________

PROJECT TITLE, INTRODUCTION, PROBLEM DEFINITION and ASSIGNMENT

Complete all fields, keep information clear, specific and concise

Project title Exploring Humanity Through AI-Generated Art

Please state the title of your graduation project (above). Keep the title compact and simple. Do not use abbreviations. The remainder of this document allows you to define and clarify your graduation project.

Introduction

Describe the context of your project here; What is the domain in which your project takes place? Who are the main stakeholders and what interests are at stake? Describe the opportunities (and limitations) in this domain to better serve the stakeholder interests. (max 250 words)

In an era where Artificial Intelligence (AI) continues to evolve, its role in understanding and reflecting humanity has become a critical subject of exploration. This project operates within the intersection of AI, art, and philosophy, aiming to investigate whether AI can capture the essence of humanity through artistic representation. By leveraging the capabilities of a custom AI painting model, the project seeks to translate collective human insights into a single visual artwork, raising profound questions about AI's interpretive and creative potential.

The main stakeholders in the project include:
- Participants: A diverse group of 1,000 individuals contributing their words or core concepts defining humanity. Their perspectives are central to shaping the dataset and ensuring inclusivity.
- AI Researchers and Developers: Interested in understanding how AI systems interpret abstract human concepts and the implications for AI-human interaction.
- Artists and Philosophers: Exploring the boundaries of creativity and whether AI can emulate or complement human artistic and philosophical endeavors.
- People who are interested in AI related discussion: Representing those impacted by AI's role in cultural and ethical domains, especially as it shapes societal perceptions of humanity.

This domain presents exciting opportunities to blend technology and human input in a collaborative creative process. It provides a unique opportunity to test how AI technology interprets and understands humanity and human nature. However, it also faces limitations, including biases in datasets, the inherently subjective nature of human concepts, and the challenges of evaluating artworks in relation to human values. By addressing these challenges, the project aims to contribute meaningfully to the ongoing dialogue on the relationship between AI and humanity.

➔ *space available for images / figures on next page*

introduction (continued): space for images

click to add picture

image / figure 1

image / figure 2

Problem Definition

What problem do you want to solve in the context described in the introduction, and within the available time frame of 100 working days? (= Master Graduation Project of 30 EC). What opportunities do you see to create added value for the described stakeholders? Substantiate your choice.
(max 200 words)

The core problem this project seeks to address is whether Artificial Intelligence (AI) can truly understand and represent the essence of humanity through artistic creation. While AI has made significant advances in generating creative outputs, its ability to capture abstract and deeply human concepts such as empathy, morality, and diversity remains unclear. This gap is critical, as it raises questions about AI's potential role in complementing human creativity and its implications for cultural, ethical, and societal narratives.

Within the 100-working-day timeframe, this project aims to leverage a custom AI painting model to interpret humanity by synthesizing input from 1,000 individuals who will provide words or core concepts defining humanity. By evaluating the resulting AI-generated artwork, the project will explore whether AI can meaningfully capture and reflect human values.

This work creates added value for stakeholders by advancing discussions on AI's interpretive and creative capabilities. For participants, it offers a platform to see their collective perspectives visualized. For researchers, it identifies gaps in AI's ability to process abstract concepts. For the general public and those concerned about AI's future applications, it provides insights into how AI might complement, rather than replace, human creativity. Ultimately, the project seeks to foster a deeper understanding of AI's potential role in representing humanity.

Assignment

This is the most important part of the project brief because it will give a clear direction of what you are heading for. Formulate an assignment to yourself regarding what you expect to deliver as result at the end of your project. (1 sentence) As you graduate as an industrial design engineer, your assignment will start with a verb (Design/Investigate/Validate/Create), and you may use the green text format:

Investigate and evaluate an AI-generated artistic approach to interpret and represent humanity, based on inputs from 1,000 individuals, in the context of understanding AI's capacity to reflect human values and diversity.

Then explain your project approach to carrying out your graduation project and what research and design methods you plan to use to generate your design solution (max 150 words)

The project begins with a comprehensive literature review to explore existing knowledge on humanity, AI-generated art, and evaluation frameworks. I will then collect qualitative data from 1,000 individuals, capturing diverse perspectives on humanity, which will be processed into a custom AI painting model to generate a visual representation.

Evaluating the artwork, the most challenging stage, involves addressing the subjectivity of art and the lack of standard evaluation criteria. To overcome this, I will research evaluation methods and develop an approach combining qualitative feedback from contributors and expert reviews. Metrics such as emotional resonance and creativity will support this analysis, highlighting strengths and limitations.

This process will test AI's ability to interpret humanity while providing insights into its potential for meaningful, human-centered applications in art and beyond.

Project planning and key moments

*To make visible how you plan to spend your time, you must make a planning for the full project. You are advised to use a Gantt chart format to show the different phases of your project, deliverables you have in mind, meetings and in-between deadlines. Keep in mind that all activities should fit within the given run time of 100 working days. Your planning should include a **kick-off meeting**, **mid-term evaluation meeting**, **green light meeting** and **graduation ceremony**. Please indicate periods of part-time activities and/or periods of not spending time on your graduation project, if any (for instance because of holidays or parallel course activities).*

Make sure to attach the full plan to this project brief.
The four key moment dates must be filled in below

Kick off meeting ______

Mid-term evaluation ______

Green light meeting ______

Graduation ceremony ______

In exceptional cases (part of) the Graduation Project may need to be scheduled part-time. Indicate here if such applies to your project

Part of project scheduled part-time	
For how many project weeks	
Number of project days per week	

Comments:

Motivation and personal ambitions

Explain why you wish to start this project, what competencies you want to prove or develop (e.g. competencies acquired in your MSc programme, electives, extra-curricular activities or other).

Optionally, describe whether you have some personal learning ambitions which you explicitly want to address in this project, on top of the learning objectives of the Graduation Project itself. You might think of e.g. acquiring in depth knowledge on a specific subject, broadening your competencies or experimenting with a specific tool or methodology. Personal learning ambitions are limited to a maximum number of five.
(200 words max)

This project aligns deeply with my academic background in communications, my passion for visual art, and my growing commitment to exploring the transformative role of AI in society. My studies and experiences in communication have equipped me with a strong foundation in understanding human interaction, which I aim to build upon by exploring how AI can interpret and represent humanity. Visual art has always been a personal passion, and this project provides an opportunity to combine it with my professional and academic pursuits, allowing me to investigate the creative potential of AI as a medium for human expression.

The rapid development of AI technology is reshaping industries, making it an essential skillset for my future career. Through this project, I aim to deepen my understanding of AI systems, particularly in creative applications, while honing my ability to evaluate and contextualize their outputs. I am also eager to experiment with tools like AI painting models and advanced evaluation methodologies, broadening my competencies in human-centered design and interdisciplinary research.

On a personal level, I want to develop the ability to merge technology with artistic and philosophical questions, fostering insights that contribute to more meaningful and ethical AI applications. This project represents a culmination of my academic journey and a foundation for future endeavors.