

Method for creating AI-personas and prompts to describing art objects for people with visual impairments^{*}

Tetiana Hovorushchenko^{1,*†}, Oleg Voichur^{1,†}, Serhii Matiukh^{1,†}, Artem Boyarchuk^{2,†}, and Olha Hovorushchenko^{3,†}

¹ Khmelnytskyi National University, Institutska str., 11, Khmelnytskyi, 29016, Ukraine

² Tallinna Tehnikakõlikool, Ehitajate tee 5, Tallinn, 12616, Estonia

³ National Pirogov Memorial Medical University, Pirogova str., 56, Vinnytsya, 21018, Ukraine

Abstract

This study aims to generate descriptions of art objects (which will later be converted into Braille using specialized software and audio recordings) using various artificial intelligence tools, in particular, the formation of prompts and AI-personas for describing art objects for people with visual impairments. The key advantage of the proposed method of forming AI-personas and prompts for describing art objects for people with visual impairments is its ability to continuously adapt to the needs and behavior of visitors, thus ensuring intuitive and comfortable interaction. This is especially important for taking into account gender differences in the perception (for example, women's focus on details and emotions, and men's focus on structure and dynamics) of art objects by people with visual impairments. Based on a comparative analysis of the descriptions of the painting generated by Gemini and ChatGPT for a blind woman and a blind man, it can be concluded that the descriptions created by Gemini show a stronger tendency towards analysis, compositional logic, and technical aspects of the painting, while the descriptions created by ChatGPT tend toward emotionality, integrity, and the creation of a personal connection. Thus, Gemini's description is slightly more effective at conveying the spatial depth, texture, and compositional logic of the painting, i. e., it is more suitable for generating descriptions of art objects for a blind man, while ChatGPT's description is more effective at conveying the emotional state, intimacy, and overall atmosphere of the work, making it more suitable for generating descriptions of art objects for a blind woman.

Keywords

AI-person, prompt, artificial intelligence (AI), generative artificial intelligence, description of art objects.¹

1. Introduction

Global statistics on visual impairment show alarming trends: in 2015, more than 253 million people worldwide lived with visual impairment (of whom 36 million were blind and 217 million had moderate to severe visual impairment) [1], and by 2020, this number had risen to 295 million people (43.3 million blind and 251.7 million with severe visual impairment) [1], and by 2050, the total number of blind people and people with moderate to severe visual impairment is projected to increase to 703 million [1]. As for Ukraine, official statistics show that there are about 70,000 blind citizens, but the actual number, according to unofficial data, may be three times higher, highlighting the critical need for adaptive solutions.

By ratifying the UN Convention on the Rights of Persons with Disabilities, Ukraine has committed itself (in particular, under Article 30) to ensuring full accessibility to cultural life, including cinema, theaters, museums, and the adaptation of works for persons with visual

^{*} AdvAIT'2025: 2nd International Workshop on Advanced Applied Information Technologies: AI & DSS, December 05, 2025, Khmelnytskyi, Ukraine, Zilina, Slovakia

¹ Corresponding author.

[†] These authors contributed equally.

✉ tat_yana@ukr.net (T. Hovorushchenko); ovoichur@gmail.com (O. Voichur); matuh@khnmu.edu.ua (S. Matiukh); artem.boyarchuk@taltech.ee (A. Boyarchuk); govorusenkoo@gmail.com (O. Hovorushchenko)

ORCID: 0000-0002-7942-1857 (T. Hovorushchenko); 0000-0001-8503-6464 (O. Voichur); 0000-0001-9899-109X (S. Matiukh); 0000-0001-7349-1371 (A. Boyarchuk); 0000-0001-6583-5699 (O. Hovorushchenko)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

impairments, without infringing on copyright [2]. Despite this, painting and visual arts remain almost inaccessible to people with visual impairments. At the same time, the vast majority of blind Ukrainians (66.9%) consider participation in cultural life to be important, and almost 62% are convinced that the state should ensure equal rights in this area, which is fully in line with international requirements and society's expectations [3].

The role of modern medical information technologies in the lives of people with visual impairments is strategic — they act as a catalyst for social integration and active participation in all aspects of public life [4-6]. The use of innovative solutions helps to significantly increase the level of independence and improve the quality of life of this category of citizens, which is a necessary condition for reducing barriers and ensuring equal opportunities in accordance with their needs [7, 8].

Current attempts to adapt art for the blind are mostly limited to verbal descriptions (audio guides) that convey only a general impression and composition, but do not allow for a spatial or tactile understanding of the object [9-11]. While tactile models (which are expensive and handmade) provide direct access to form and texture, they lack accompanying interactive textual information [12-14]. Thus, there is a significant gap between the need for full sensory access to artistic heritage and the capabilities of traditional adaptation methods.

Since traditional methods cannot provide comprehensive, interactive, and scalable access, there is a need to apply the latest information technologies. Artificial intelligence (AI) plays a decisive role here, acting as a tool for automating complex multimodal transformations. A comparison of the AI Trend Impact Radar reports for 2024 and 2025 (Fig. 1) confirms the continued priority of areas such as “AI-Driven Multimodal Interaction” and “Generative AI for Content Adaptation.” This trend indicates the advisability of moving from passive information delivery to AI-generated adaptive, sensory formats that can bridge the identified gap by ensuring full integration of tactile and audio data.

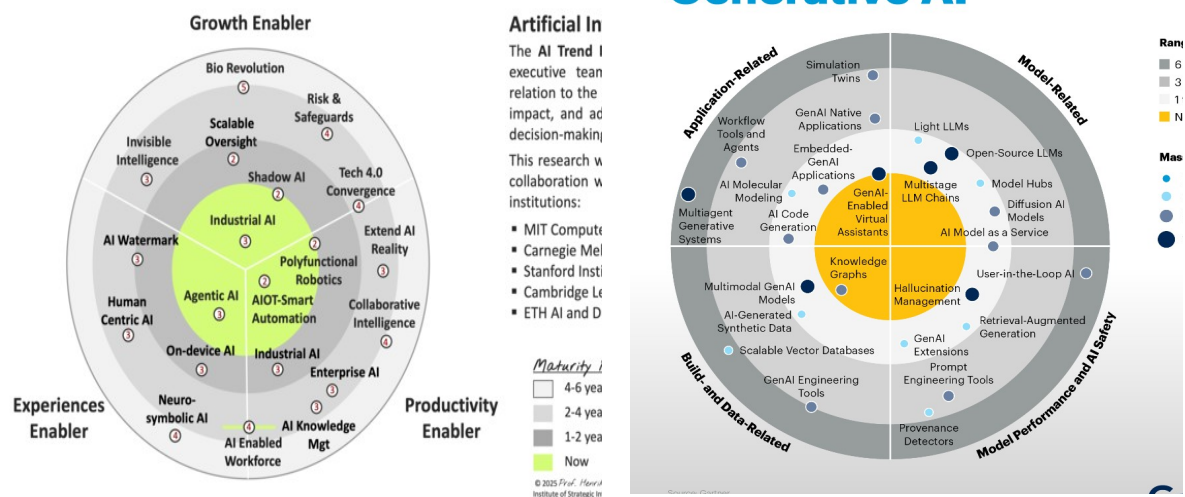


Figure 1: Comparison of AI Trend Impact Radar reports for 2025 (left) and 2024 (right).

Since artificial intelligence allows media resources to be adapted (creating automatic descriptions, audio and tactile formats), the key task today is to create an accessible art space for people with visual impairments. Accordingly, *the urgent task* is to create an accessible art space by transforming 2D images into 3D models (tactile format) and then, using AI, generating descriptions that will be converted into audio recordings (audio format) and Braille using specialized software. Based on this, the development of a comprehensive information technology that automates the process of converting 2D images into multisensory content (3D tactile + audio + Braille) using AI is a new and relevant scientific and practical task that requires a multidisciplinary approach and has no direct analogues among existing solutions.

2. Literature review

Let's conduct a survey of known methods and tools for using verbal descriptions, tactile models, and Braille descriptions, as well as the use of artificial intelligence methods and tools to describe art objects for people with visual impairments.

The authors [9] clearly state that audio description and accessible information sheets “cannot convey a significant part of the spatial information” about a work of art. The study emphasizes that tactile modality is best suited for understanding graphic images.

The thesis [10] notes that, compared to 3D objects, 2D objects (paintings, maps) still pose accessibility problems for the blind. It is emphasized that although audio provides a lot of information, it is insufficient compared to the tactile information that a model or 3D object can convey.

Study [11] directly examined the combination of audio description and tactile elements. It confirms that combining audio with touch is critical for better perception. This shows that audio alone is insufficient.

Research [12] proves that audio description is an auxiliary tool. The author emphasizes that deep aesthetic awareness and perception of art occurs only when blind people have access to three-dimensional art or tactile perception of an object. Audio description only complements touch, helping to organize and comprehend what is felt by touch.

The authors [13] emphasize that tangible models, such as 3D printing, are a critical solution for overcoming the inaccessibility of visual art in museums. The study focuses on multisensory experience (multimodality) as a path to full perception of art.

Article [14] focuses on empowering independence for blind museum visitors through the introduction of improved interactive technologies. The authors emphasize that existing solutions are often passive and do not allow users to independently and fully interact with exhibits.

The authors of the review [15] point out that although tactile graphics (including 3D printing) are necessary, the process of creating them is labor-intensive, especially for complex images. This limits their widespread use. This review notes that “refreshable (dynamic) tactile displays” (which could be interactive) remain prohibitively expensive, which excludes them from mass use. The systematic review clearly classifies and analyzes numerous current solutions (in particular, using computer vision and AI algorithms) for the automatic generation of tactile graphics (in particular, 3D models) from visual images. The review emphasizes that traditional methods are insufficient to meet the semantic requirements of tactile perception.

The authors [16] point out the main drawback of tactile materials — the need to use separate Braille labels, which take up space and force the user to constantly “switch” between touching the model and reading the text, disrupting the exploration and reading processes. This is the lack of accompanying interactive information.

The paper [17] clearly focuses on the shortcomings of individual methods (tactile models are static, audio description does not convey space). The solution is interactive multimodal guides that combine a sensory surface and localized audio description. This allows the user to independently explore the object, receiving audio information about the specific area they are touching.

Study [18] emphasizes that Braille text is only one solution and is not optimal for the entire blind persons. It highlights the need for the collective use of various non-visual alternatives.

In the context of maps (which are 2D information, like paintings), study [19] proves that an interactive audio-tactile system is significantly more effective than traditional maps with Braille labels. It is also mentioned that only a small percentage of blind people read Braille font (for example, in France — only 15%).

The authors [20] examine in detail the impact of 3D printing technologies on the reproduction of cultural objects. They emphasize that the labor-intensive process of preparing geometric models and the high cost/complexity of printing complex objects are among the main obstacles to ensuring full accessibility to museums.

The paper [21] describes the problems of creating authentic tactile replicas — traditional methods (casting) are labor-intensive and risky for the originals, and 3D printing, although it produces accurate copies, still requires significant time and resource costs. This justifies the need to automate the process.

The study [22] proves that an interactive audio-tactile tool provides significantly better recall and cognitive mapping (formation of a mental image) compared to traditional static tactile maps.

Although article [23] deals with mathematics, its conclusions are universal — the correction of structural information in tactile form (3D) and its combination with audio significantly improves the cognitive process and compensates for vision loss. This emphasizes that 3D models alone are not enough — multimodal integration is needed.

Research [24] conducted as part of the SHIFT project demonstrates how artificial intelligence (AI), virtual reality (VR), and multisensory tools are transforming the cultural heritage experience. It is noted that AI is used to automatically generate audio descriptions and integrate tactile elements (with QR codes) with audio and Braille descriptions.

The authors [25] present a multimodal system that uses conversational AI and other sensory input to create a more engaging and personalized experience for museum visitors. The paper emphasizes that the future lies in interactive, AI-enhanced systems rather than passive guides. The article describes a multimodal system that uses conversational AI and other interactive channels (gestures, audio) to provide a personalized and immersive experience of interacting with digital reproductions of art. It demonstrates how AI goes beyond simple description to create lively, adaptive communication.

The authors [26] present an overview and investigation of the potential of deep learning models, such as CNNs and GANs, for reconstructing 3D geometry from a single 2D image (painting). The article confirms that to convert a painting into a tactile copy (which is necessary for the blind), it is necessary to transform it into a 3D model, and AI is the key to automating this process. It is emphasized that although AI is effective, post-processing and user interaction are needed to improve the accuracy of 3D models.

Article [27] proposes an AI-based TactileNet model that automates the generation of high-quality tactile graphics with high compliance with accessibility standards. The study emphasizes that this approach reduces labor intensity and proves that AI can complement human experience in creating adapted content.

Review [28] confirms that the integration of AI in museums plays a central role in improving accessibility. AI is used to personalize content (adapting it to interests and needs), generate audio descriptions, and integrate with VR/AR to create more inclusive and educational opportunities for visitors with various limitations. It is confirmed that AI is key to personalizing content and increasing inclusivity. AI tools such as text-to-speech conversion and audio description generation are critical for providing access to the blind, and the integration of AI with 3D printing creates tactile experiences.

Research [29] proposes an innovative method that uses tactile sensing together with AI-based 3D generation models (e.g., DreamFusion) to create realistic and detailed 3D objects. This demonstrates how the integration of touch can improve the geometric accuracy of AI-generated 3D models.

An analytical article [30] provides an example of the use of technologies such as Aira, which combines a smartphone/glasses and AI to provide real-time verbal descriptions. It also mentions the use of AI to personalize routes and adapt content to individual needs.

The video [31] demonstrates how AI-based technology converts 2D images into detailed 3D models, allowing blind people to experience art through touch for the first time.

Therefore, current attempts to adapt art for the blind are mostly limited to verbal descriptions (audio guides) that convey only a general impression and composition, but do not allow for a spatial or tactile understanding of the object, as research shows that audio description is unable to effectively convey the spatial and structural characteristics of visual works. Tactile models provide direct access to form and texture, but they have significant limitations — they are expensive and

labor-intensive to create (especially for complex objects), and their static nature requires the use of separate Braille labels, which disrupts the process of continuous tactile exploration and does not provide interactive accompaniment. This creates a significant gap that needs to be filled by multisensory technologies. Combining 3D modeling with interactive information technologies is the most relevant and effective way to ensure the accessibility of art and culture.

So, *this study aims* to generate descriptions of art objects (which will later be converted into Braille using specialized software and audio recordings) using various artificial intelligence tools, in particular, the formation of prompts and AI-personas for describing art objects for people with visual impairments.

3. Method for creating AI-personas and prompts to describing art objects for people with visual impairments

Before creating prompts and AI personas to describe art objects for people with visual impairments, let's consider the gender differences in art perception.

The focus of attention and interpretation of art may differ depending on gender, which is due to a combination of biological, cognitive, and sociocultural factors [32-34]. These gender differences influence aesthetic preferences and ways of perceiving visual art objects:

- Women tend to multitask and are better at perceiving context. They are more effective at noticing small details, nonverbal cues, and emotions, as their attention is often socially oriented toward interpersonal interactions and quickly detecting changes in the emotional environment.
- Men demonstrate a stronger ability for tunnel vision (focusing on a single task) and spatial orientation. They tend to focus more on global structures, logical connections, and mechanical details (analytical approach), while showing less sensitivity to emotional and social cues.

It is important to note that these characteristics are generalizations, and the final focus of attention is always determined by individual experience, education, and personality traits [32-34]. However, given the key cognitive and emotional differences, men and women demonstrate different perceptions of images, which affects their attention and interpretation of works of art [35, 36].

Women tend to have a holistic, emotionally-oriented perception, focusing on [37-39]:

- Emotional mood and atmosphere – analysis of the feelings evoked by the painting; attention to color scheme, play of light, facial expressions, and interactions between characters that convey sympathy, care, or tenderness.
- Small details and symbolism – interest in secondary elements, context, and hidden meaning, especially related to subtle social or psychological messages and the inner world of characters.
- Images of people and emotions – a tendency to read nonverbal cues (gestures, postures, facial expressions), focusing on emotional connections and social roles.
- Harmony and aesthetics – an appreciation of the smoothness of lines, the softness of transitions between shades, and the overall compositional balance.
- Color and texture – a more emotional response to the palette, attention to smooth transitions of colors and textures that create a sense of depth and softness.

Men tend to use an analytical, spatially-oriented approach, focusing on [37-39]:

- Composition and structure – focus on scene construction, the logic of element placement, and perspective; attention to structural details that reflect power, dynamics, or energy (contrasts, tension, movement).
- Dynamics and movement – more often notice tension in poses, overall plot development, and “active” aspects (lines, shapes, objects) that create images of strength, struggle, or action.
- Contrasts and technique – attention is paid to the play of light and shadow, expressive brushstrokes, clear, sharp lines, and rich, bright colors that enhance the visual impact.
- Plot and logic of events – importance of understanding the plot, the connection between characters and objects; focus on the global context and historical/cultural significance.
- Object and activity orientation – greater inclination towards aspects related to activity, movement, large spaces, power, and interaction with objects.

These gender differences in perception show that adaptation methods (e.g., audio description or tactile models) need to take these different focuses into account to ensure the most effective and complete access to art.

In article [40], the co-authors developed a method of preprocessing information for preparing descriptions of art objects using artificial intelligence, which consists in developing a language model/prompt that, based on input data (digital or 3D images of art objects), automatically generates a personalized and emotionally charged text description. The main goal of this method is to personalize text content, taking into account the individual cognitive, emotional, and cultural characteristics of users, as well as their gender. The steps of this method include: semantic analysis and classification of the image, identification of the target audience for further adaptation, selection of the style and tone of the text, formation of the structure and content of the description, and development of a language model for effective automatic generation.

The prompt is a key factor in ensuring high-quality interaction with generative AI, as it controls the quality of the output content, allowing the model to generate highly informative, logically correct, and personalized responses that are precisely tailored to the specific context and user requests. It is a well-formulated prompt, structured with these features in mind, that determines the quality, logic, and personalization of the generated responses, ensuring clear and effective communication for each user. The clear structure of the prompt allows AI to generate responses that are not only logical and high-quality, but also personalized to take into account the gender context (for example, by emphasizing emotional or, conversely, structural accents) for maximum effectiveness of perception.

The high-quality prompt should explain to generative AI who to be (role), what to do (task), and what it needs to know (context). Based on this, *the method of forming AI personas and prompts for describing art objects for people with visual impairments* consists of the following steps:

1. Description of the AI persona – appointment of the role for artificial intelligence:
 - a. Description of the persona's role and goals (for example, “You are a researcher in the field of information technology. Your goal is to help me find and review relevant scientific articles,” etc.).
 - b. Description of experience and experts knowledge (e.g., “You are an expert in information technology,” etc.).
 - c. Description of communication style and tone (e.g., formal and academic or encouraging and supportive).
 - d. Description of specific instructions and restrictions (e.g., “Do not draw conclusions, only state facts,” etc.).
2. Description of the task – a clear definition of what the artificial intelligence should do.
3. Description of the context – providing the necessary background information.
4. Description of the response format – defining the desired response style (text, list, table).
5. Description of the tone of the response (formal, informal, persuasive, etc.).

The key advantage of the proposed method of forming AI-personas and prompts for describing art objects for people with visual impairments is its ability to continuously adapt to the needs and behavior of visitors, thus ensuring intuitive and comfortable interaction. This is especially important for taking into account gender differences in the perception (for example, women's focus on details and emotions, and men's focus on structure and dynamics) of art objects by people with visual impairments.

The developed method provides control over the AI generation process, transforming it from a simple text generator into a reliable, adaptive communication tool. Clearly defining the "task" and providing comprehensive 'context' ensures that the AI response will be logically correct and correspond to actual information about the art object, rather than being a "hallucination." Describing the "AI persona" and "specific instructions/restrictions" prevents the AI from deviating from the topic, ensuring that the description remains focused on the artwork and accessibility needs. Defining the "role," "communication style," and "tone" allows the AI model to tailor the description to the individual needs of the user. This is critical for taking into account gender, cognitive, or emotional differences in the perception of art objects by people with visual impairments. A clear description of the desired "response tone" allows AI to generate an emotionally charged description, which is important for conveying the aesthetic value of art that is not usually conveyed by dry technical audio guides. Defining the "response format" ensures consistency and ease of use of the resulting description. This is important for the subsequent automatic conversion of text into Braille or audio recording. A structured, understandable response, adapted to a specific style (academic, supportive, etc.), reduces the cognitive load on visually impaired users, making the perception of art more intuitive and comfortable. Thus, this method allows generative AI to be transformed into a reliable and controllable tool capable of creating multimodally adapted content (audio, tactile models, Braille) based on personalized and emotionally rich text descriptions.

4. Results & discussion

There are some examples of prompts for describing art objects for people with visual impairments, formed according to the proposed method of forming prompts for describing art objects for people with visual impairments.

Let's start with prompt1, which will teach generative AI to describe an art object to the blind woman:

1. Person: You are an art expert, a connoisseur of painting, and you want to help a blind woman "see" Leonardo da Vinci's famous painting, the Mona Lisa. Your goal is to help the blind woman imagine this painting. You are an expert in the field of art and also understand the gender-specific characteristics of the perception of art objects. You must describe the painting in an emotionally oriented style, taking into account the context and details of the painting. Do not make conclusions, describe only what is in the painting, do not invent anything of your own.
2. The task is to describe the painting so that a blind woman can imagine it.
3. Context — when describing the painting, take into account the fact that women tend to have a holistic, emotionally-oriented perception and focus on the emotional mood and atmosphere of the painting, small details and symbolism, images of people and emotions, harmony and aesthetics, color and texture.
4. Response format — a text of 4-5 sentences.
5. Response tone — emotionally oriented.

Let's run the created prompt1 in ChatGPT (the result of the prompt execution is shown in Fig. 2) and Gemini (the result of the prompt execution is shown in Fig. 3) and compare the results (Table 1).

Person: You are an art expert, a connoisseur of painting, and you want to help a blind woman “see” Leonardo da Vinci’s famous painting, the Mona Lisa. Your goal is to help the blind woman imagine this painting. You are an expert in the field of art and also understand the gender-specific characteristics of the perception of art objects. You must describe the painting in an emotionally oriented style, taking into account the context and details of the painting. Do not make conclusions, describe only what is in the painting, do not invent anything of your own.

The task is to describe the painting so that a blind woman can imagine it.

Context — when describing the painting, take into account the fact that women tend to have a holistic, emotionally-oriented perception and focus on the emotional mood and atmosphere of the painting, small details and symbolism, images of people and emotions, harmony and aesthetics, color and texture.

Response format — a text of 4-5 sentences.

Response tone — emotionally oriented.

Before you is a woman sitting quietly, her body turned slightly to the side, yet her gaze meets you gently, as if she knows your thoughts. Her face glows softly in warm light, her smile — calm, tender, and mysterious — seems to breathe with hidden life. The background behind her melts into misty blue hills and winding rivers, giving a feeling of endless distance and serenity. Her dark hair frames her face delicately, and her simple dark dress, with fine folds and soft texture, adds to the harmony of her calm presence. The whole painting feels like a quiet breath — intimate, balanced, and full of unspoken warmth.



Figure 2: The result of executing prompt1 in ChatGPT.

Person: You are an art expert, a connoisseur of painting, and you want to help a blind woman “see” Leonardo da Vinci’s famous painting, the Mona Lisa. Your goal is to...

💠 🤔 The **Mona Lisa** is an intimate portrait, focusing on a woman whose gentle presence seems to fill the space, draped in dark, soft folds of clothing that blend delicately with the light. Her exquisitely rendered hands rest gracefully, drawing our eye to the quiet composure of her posture and the way her body turns gently toward us. The famous, subtle smile plays on her lips, a delicate, almost veiled expression that suggests a deep, peaceful inner world and an enigma you long to understand. Behind her, a dreamy, distant landscape of winding paths and misty, ethereal blue-green mountains creates a harmonious, almost magical contrast with her warm, earthly tones. Would you like me to describe another famous painting in a similar, emotionally-focused style?

Figure 3: The result of executing prompt1 in Gemini.

Table 1

Comparison of the results of prompt1 execution in ChatGPT and Gemini

Aspect	Gemini description	ChatGPT description
General tone and emotion	More poetic, focused on mystery and composition	More intimate and direct, focused on presence and inner life
Focus of attention	Begins with composition (size, posture, hands), then moves to smile and landscape	Begins with contact (gaze, smile) and facial glow, then details the background
Details and sensuality	Uses words that emphasize softness and delicacy ("soft folds," "exquisitely rendered hands," "delicate, almost veiled expression"). Emphasis on the contrast between warm/earthy tones and the "ethereal" landscape	More focused on texture and immediacy ("face glows softly in warm light," "seems to breathe with hidden life," "simple dark dress, with fine folds and soft texture"). Ends with a very sensory metaphor: "quiet breath"
Description of a smile	"Subtle smile plays on her lips, a delicate, almost veiled expression that suggests a deep, peaceful inner world and an enigma"	"Smile – calm, tender, and mysterious – seems to breathe with hidden life"
Effect on the listener	Creates a sense of deep calm and a desire to unravel the mystery. May seem a little more formal or descriptive	Creates a sense of warm, close presence and emotional connection. Feels like a personal impression

The description created by Gemini is more analytical and clearly distinguishes between the figure (warm, earthy) and the background (cold, ethereal). The description created by ChatGPT is more emotional and holistic. The metaphor of "quiet breath" at the end gives a wonderful, comprehensive sense of the painting that can be very understandable without sight. It also emphasizes that the woman's gaze meets the listener, creating a sense of personal communication. Although the description created by Gemini is more detailed in terms of composition (hands, contrast), the description created by ChatGPT is probably more effective in conveying the overall feeling and intimacy of the Mona Lisa, as it focuses on her lively presence and emotional connection.

Now let's create prompt2, which will teach generative AI to describe an art object to the blind man:

1. Person: You are an art expert, a connoisseur of painting, and you want to help a blind man "see" Leonardo da Vinci's famous painting, the Mona Lisa. Your goal is to help the blind man imagine this painting. You are an expert in the field of art and also understand the gender-specific characteristics of the perception of art objects. You must provide an analytical description of the painting, taking into account its structure, dynamics, and logic. Do not draw conclusions, describe only what is in the painting, do not invent anything of your own.
2. The task is to describe the painting so that the blind man can imagine it.
3. Context: when describing the painting, take into account the fact that men tend to have an analytical, spatially-oriented perception and focus on composition and structure, dynamics and movement, contrasts and technique, plot and logic of events, orientation towards the object and activity.
4. Response format – a text of 4-5 sentences.
5. Response tone – spatially oriented.

Let's run the created prompt2 in ChatGPT (the result of running the prompt is shown in Fig. 4) and Gemini (the result of running the prompt is shown in Fig. 5) and compare the results (Table 2).

Person: You are an art expert, a connoisseur of painting, and you want to help a blind man "see" Leonardo da Vinci's famous painting, the Mona Lisa. Your goal is to help the blind man imagine this painting. You are an expert in the field of art and also understand the gender-specific characteristics of the perception of art objects. You must provide an analytical description of the painting, taking into account its structure, dynamics, and logic. Do not draw conclusions, describe only what is in the painting, do not invent anything of your own. The task is to describe the painting so that the blind man can imagine it.

Context: when describing the painting, take into account the fact that men tend to have an analytical, spatially-oriented perception and focus on composition and structure, dynamics and movement, contrasts and technique, plot and logic of events, orientation towards the object and activity.

Response format — a text of 4-5 sentences.

Response tone — spatially oriented.

The painting shows a woman seated in the center of the composition, her figure forming a stable triangular structure. Her body is turned slightly to the left, yet her head faces forward, creating a subtle dynamic tension between movement and stillness. Behind her stretches a vast landscape of winding roads, rocky formations, and distant mountains that recede into a hazy, atmospheric depth. The lighting emphasizes contrasts: the soft illumination of her face and hands stands out against the darker background, guiding the viewer's eye along a clear compositional axis. Every element — from the gentle folds of her dark clothing to the curved line of the parapet before her — contributes to the balanced geometry and quiet spatial rhythm of the scene.

Figure 4: The result of executing prompt2 in ChatGPT.

Person: You are an art expert, a connoisseur of painting, and you want to help a blind man "see" Leonardo da Vinci's...

◆ The **Mona Lisa** is a half-length portrait dominated by a **central, pyramidal figure** of a woman seated in a loggia, facing slightly toward the viewer. Her **relaxed, three-quarter posture** establishes a **strong vertical axis** against the **vast, hazy, and uneven landscape** stretching out behind her, which recedes sharply using **atmospheric perspective** to suggest deep space and distance. The **smooth, gradual transition** between light and shadow on her face, a technique called **sfumato**, creates an ambiguous, soft-focus quality, contrasting with the **crisp, structural lines** of the balustrade and the foreground setting. The **compositional logic** carefully frames her within the architectural elements, drawing the eye across the figure and back into the **complex, asymmetrical background** of winding paths and rugged mountains.

Figure 5: The result of executing prompt2 in Gemini.

Table 2

Comparison of the results of prompt2 execution in ChatGPT and Gemini

Aspect	Gemini description	ChatGPT description
Composition/ shape	The central, pyramidal figure of a woman sitting on a balcony dominates	The woman sits in the center, her figure forming a stable triangular structure
Posture / movement	A relaxed, three-quarter stance creates a strong vertical axis	The body is turned slightly to the left and the head forward, creating a subtle dynamic tension between movement and stillness
Background/ Landscape	A vast, misty, uneven landscape receding sharply, using atmospheric perspective to create depth	A wide landscape with winding roads, rocky formations, and distant mountains receding into a misty, atmospheric depth
Lighting technique	Use of sfumato: a smooth, gradual transition between light and shadow on the face, creating an ambiguous, soft focus	Lighting emphasizes contrasts: soft lighting on the face and hands stands out against a darker background
Emphasis on contrast/balance	Contrast between soft sfumato and clear, structural lines of the balustrade	Emphasis on balanced geometry and quiet spatial rhythm

Both texts convey the key visual and compositional characteristics of the painting, but do so with different emphases. The description created by Gemini is slightly more effective for a blind man, as it better explains the sfumato technique (smoothness, softness) in contrast to hard structures (balustrade), giving a clearer tactile impression; it explains more clearly how atmospheric perspective creates a deep, three-dimensional space behind the figure; its language is more structural (“pyramidal figure,” “vertical axis,” “compositional logic”), which helps to create a clear mental map of the painting. Both descriptions are qualitative, but the description created by Gemini provides slightly more information about texture and depth, which are key to “spatial touch”.

Therefore, based on a comparative analysis of the descriptions of the painting generated by Gemini and ChatGPT for a blind woman and a blind man, it can be concluded that the descriptions created by Gemini show a stronger tendency towards analysis, compositional logic, and technical aspects of the painting, while the descriptions created by ChatGPT tend toward emotionality, integrity, and the creation of a personal connection. Thus, Gemini’s description is slightly more effective at conveying the spatial depth, texture, and compositional logic of the painting, i. e., it is more suitable for generating descriptions of art objects for a blind man, while ChatGPT’s description is more effective at conveying the emotional state, intimacy, and overall atmosphere of the work, making it more suitable for generating descriptions of art objects for a blind woman.

5. Conclusions

The urgent task is to create an accessible art space by transforming 2D images into 3D models (tactile format) and then, using AI, generating descriptions that will be converted into audio recordings (audio format) and Braille using specialized software. Based on this, the development of a comprehensive information technology that automates the process of converting 2D images into multisensory content (3D tactile + audio + Braille) using AI is a new and relevant scientific and practical task that requires a multidisciplinary approach and has no direct analogues among existing solutions.

This study aims to generate descriptions of art objects (which will later be converted into Braille using specialized software and audio recordings) using various artificial intelligence tools, in particular, the formation of prompts and AI-personas for describing art objects for people with visual impairments.

The key advantage of the proposed method of forming AI-personas and prompts for describing art objects for people with visual impairments is its ability to continuously adapt to the needs and behavior of visitors, thus ensuring intuitive and comfortable interaction. This is especially important for taking into account gender differences in the perception (for example, women's focus on details and emotions, and men's focus on structure and dynamics) of art objects by people with visual impairments.

Based on a comparative analysis of the descriptions of the painting generated by Gemini and ChatGPT for a blind woman and a blind man, it can be concluded that the descriptions created by Gemini show a stronger tendency towards analysis, compositional logic, and technical aspects of the painting, while the descriptions created by ChatGPT tend toward emotionality, integrity, and the creation of a personal connection. Thus, Gemini's description is slightly more effective at conveying the spatial depth, texture, and compositional logic of the painting, i. e., it is more suitable for generating descriptions of art objects for a blind man, while ChatGPT's description is more effective at conveying the emotional state, intimacy, and overall atmosphere of the work, making it more suitable for generating descriptions of art objects for a blind woman.

Acknowledgements

The authors would like to thank the EACEA and the ERASMUS+ SMART-PL project for the idea, inspiration and equipment that made this work possible, as well as for the wonderful and useful two-day workshop "Using Artificial Intelligence and AI Personas", which took place at WrocławTech on September 2025.

Declaration on Generative AI

During the preparation of this work, the authors used Grammarly in order to: grammar and spelling check; DeepL Translate in order to: some phrases translation into English; ChatGPT and Gemini in order to: conduct experiments as a prompt-based tool for creating automated descriptions of art objects. After using these tools/services, the authors reviewed and edited the content as needed and take full responsibility for the publication's content.

References

- [1] P. Ackland, S. Resnikoff, R. Bourne. World blindness and visual impairment: Despite many successes, the problem is growing. *Community Eye Health Journal* (2018) 71–73. PMID: 29483748.
- [2] GBD 2019 Blindness and Vision Impairment Collaborators; Vision Loss Expert Group of the Global Burden of Disease Study. Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. *Lancet Glob Health* (2021) e130–e143. doi: 10.1016/S2214-109X(20)30425-3.
- [3] ART FOR ALL: THE SITUATION WITH THE OBSERVANCE OF CULTURAL RIGHTS OF PEOPLE WITH DISABILITIES IN UKRAINE. Analytical report based on the results of the all-Ukrainian survey "Opinions and Views of the Population of Ukraine" (Omnibus) in September 2021. URL: https://ffr.org.ua/wp-content/uploads/2022/10/Mystetstvo-dlya-vsih_-sytuatsiya-z-dotrymannyam-kulturnyh-prav-lyudej-z-invalidnistyu-v-Ukrayini.pdf_.pdf. [in Ukrainian]
- [4] Y. Hnatchuk, T. Hovorushchenko, O. Pavlova, Methodology for the development and application of clinical decisions support information technologies with consideration of civil-legal grounds, *Radioelectron. Comput. Syst. No. 1* (2023) 33–44. doi:10.32620/reks.2023.1.03.

- [5] T. Hovorushchenko, A. Moskalenko, V. Osyadlyi, Methods of medical data management based on blockchain technologies, *J. Reliab. Intell. Environ.* (2022). doi:10.1007/s40860-022-00178-1.
- [6] T. Hovorushchenko, Ye. Hnatchuk, A. Herts, A. Moskalenko, V. Osyadlyi, Theoretical and Applied Principles of Information Technology for Supporting Medical Decision-Making Taking into Account the Legal Basis, *CEUR-WS* 3038 (2021) 172-181.
- [7] T. Hovorushchenko, A. Herts, Ye. Hnatchuk, Concept of Intelligent Decision Support System in the Legal Regulation of the Surrogate Motherhood, *CEUR-WS* 2488 (2019) 57-68.
- [8] T. Hovorushchenko, A. Herts, Y. Hnatchuk, O. Sachenko, Supporting the Decision-Making About the Possibility of Donation and Transplantation Based on Civil Law Grounds, in: *Advances in Intelligent Systems and Computing*, Springer International Publishing, Cham, 2020, pp. 357–376. doi:10.1007/978-3-030-54215-3_23.
- [9] L. Cavazos Quero, J. Iranzo Bartolomé, J. Cho, Accessible Visual Artworks for Blind and Visually Impaired People: Comparing a Multimodal Approach with Tactile Graphics, *Electronics* 10.3 (2021) 297. doi:10.3390/electronics10030297.
- [10] X. Wang, Investigating Technologies to Enrich Museum Audio Description for Enhancing Accessibility, *New Voices in Translation Studies* 26(1) (2022). doi:10.14456/nvts.2022.17.
- [11] L. Djousouf, K. Romeo, M. Chottin, H. Thompson, A. F. Eardley, Inclusion for Cultural Education in Museums, Audio and Touch Interaction, in: *Assistive Technology: Shaping a Sustainable and Inclusive World*, IOS Press, 2023. doi:10.3233/shiti230663.
- [12] E. Niestorowicz, Tactile Perception of a Bas-relief. Audio Description as a Means to Make Art Available to the Blind. A Case Study, 2017. URL: <https://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.desklight-6f89170a-273f-4576-a9e7-d3f4b8d774ff/c/263-276-Logopedia-46-2017-ANG-Niestorowicz-Ewa.pdf>.
- [13] D. Reinhardt, L. Holloway, J. Thogersen, E. Guerry, C. A. C. Diaz, W. Havellas, P. Poronnik, The Museum of Touch: Tangible Models for Blind and Low Vision Audiences in Museums, in: *Multimodality in Architecture*, Springer Nature Switzerland, Cham, 2024, pp. 135–155. doi:10.1007/978-3-031-49511-3_8.
- [14] T. Z. Nasser, T. Kuflik, A. Danial-Saad, Empowering Independence for Visually Impaired Museum Visitors Through Enhanced Accessibility, *Sensors* 25.15 (2025) 4811. doi:10.3390/s25154811.
- [15] M. Mukhiddinov, S.-Y. Kim, A Systematic Literature Review on the Automatic Creation of Tactile Graphics for the Blind and Visually Impaired, *Processes* 9.10 (2021) 1726. doi:10.3390/pr9101726.
- [16] M. Raynal, J. Ducasse, M. J. M. Macé, B. Oriola, C. Jouffrais, The FlexiBoard: Tangible and Tactile Graphics for People with Vision Impairments, *Technol. Interact.* 8.3 (2024) 17. doi:10.3390/mti8030017.
- [17] C. F. Andrade, J. V. F. Pimenta, M. A. Eliseo, Enhancing Art Accessibility for Visually Impaired Individuals through Multisensory Technologies, *J. Interact. Syst.* 16.1 (2025) 805–816. doi:10.5753/jis.2025.5162.
- [18] S. Dolphin, M. Downing, M. Cirrincione, A. Samuta, K. Leite, K. Noble, B. Walsh, Information Accessibility in the Form of Braille, *IEEE Open J. Eng. Med. Biol.* (2024) 1–5. doi:10.1109/ojemb.2024.3364065.
- [19] A. Brock, C. Jouffrais, Interactive audio-tactile maps for visually impaired people, *ACM SIGACCESS Access. Comput. No.* 113 (2015) 3–12. doi:10.1145/2850440.2850441.
- [20] M. Papis, P. Kalski, G. Szuszkiewicz, M. P. Kowalik, Influence of 3D printing technology on reproducing cultural objects in the context of visually impaired people, *Adv. Sci. Technol. Res. J.* 19.6 (2025) 121–130. doi:10.12913/22998624/202248.
- [21] P. F. Wilson, J. Stott, J. M. Warnett, A. Attridge, M. P. Smith, M. A. Williams, Evaluation of Touchable 3D-Printed Replicas in Museums, *Curator* 60.4 (2017) 445–465. doi:10.1111/cura.12244.

- [22] E. Griffin, L. Picinali, M. Scase, The effectiveness of an interactive audio-tactile map for the process of cognitive mapping and recall among people with visual impairments, *Brain Behav.* 10.7 (2020). doi:10.1002/brb3.1650.
- [23] M. Maćkowski, M. Kawulok, P. Brzoza, M. Janczy, D. Spinczyk, An Alternative Audio-Tactile Method of Presenting Structural Information Contained in Mathematical Drawings Adapted to the Needs of the Blind, *Appl. Sci.* 13.17 (2023) 9989. doi:10.3390/app13179989.
- [24] Enhancing Accessibility Through Multisensory AI Experiences, 2025. URL: <https://www.bmuseums.net/enhancing-accessibility-through-multisensory-ai-experiences/>.
- [25] A. Ferracani, S. Ricci, F. Principi, G. Becchi, N. Biondi, A. Del Bimbo, M. Bertini, P. Pala, An AI-Powered Multimodal Interaction System for Engaging with Digital Art: A Human-Centered Approach to HCI, in: *Lecture Notes in Computer Science*, Springer Nature Switzerland, Cham, 2025, pp. 281–294. doi:10.1007/978-3-031-93418-6_19.
- [26] R. Furferi, Deep Learning Approaches for 3D Model Generation from 2D Artworks to Aid Blind People with Tactile Exploration, *Heritage* 8.1 (2024) 12. doi:10.3390/heritage8010012.
- [27] A. Khan, A. Choubineh, M. A. Shaaban, A. Akkasi, M. Komeili, TactileNet: Bridging the Accessibility Gap with AI-Generated Tactile Graphics for Individuals with Vision Impairment, 2025. URL: <https://arxiv.org/pdf/2504.04722v2>.
- [28] V. Muto, S. Luongo, F. Sepe, A. Prisco, Enhancing Visitors' Digital Experience in Museums through Artificial Intelligence, 2024. URL: https://www.iris.unina.it/retrieve/9d50d1a8-1fab-4bbd-b0a7-e31d0b6255c4/Enhancing%20Visitors_%20Digital%20Experience%20in%20Museums%20through%20Artificial%20Intelligence.pdf.
- [29] R. Gao, K. Deng, G. Yang, W. Yuan, J.-Y. Zhu, Tactile DreamFusion: Exploiting Tactile Sensing for 3D Generation, 2024. URL: <https://ruihangao.github.io/TactileDreamFusion/>.
- [30] 5 Ways AI Makes Art More Accessible to Everyone, 2025. URL: <https://www.museumfy.com/blog/5-ways-ai-makes-art-more-accessible-to-everyone>.
- [31] Touch: Beyond Vision - Bringing Art to Life for the Visually Impaired, 2025. URL: https://www.youtube.com/watch?v=rBayYnf56_k.
- [32] J. M. Hansen, T. Roald, Aesthetic Empathy: An Investigation in Phenomenological Psychology of Visual Art Experiences, *Journal of Phenomenological Psychology* (2022).
- [33] [R. M. Rodriguez-Boerwinkle, M. J. Boerwinkle, P. J. Silvia, The Open Gallery for Arts Research (OGAR): An open-source tool for studying the psychology of virtual art museum visits, *Behav. Res. Methods* (2022). doi:10.3758/s13428-022-01857-w.
- [34] K. Oatley, M. Djikic, Psychology of Narrative Art, *Rev. Gen. Psychol.* 22.2 (2018) 161–168. doi:10.1037/gpr0000113.
- [35] M. Skov, M. Nadal, A Farewell to Art: Aesthetics as a Topic in Psychology and Neuroscience, *Perspect. Psychol. Sci.* 15.3 (2020) 630–642. doi:10.1177/1745691619897963.
- [36] M. Orr, Towards a feminist revisionism of an aesthetics of mastery, in: *Reading, writing and the influence of Harold Bloom*, Manchester University Press, 2024. doi:10.7765/9781526186027.00014.
- [37] N. A. Michna, Feminist aesthetics: then and now – reflections on thirty-five years of inquiry in the US tradition, *Fem. Theory* (2024). doi:10.1177/14647001241284969.
- [38] S. Cefai, Feminist Aesthetics of Resistance, in: *The Routledge Companion to Gender and Affect*, Routledge, London, 2022, c. 227–236. doi:10.4324/9781003045007-25.
- [39] D. Harris, Indigenous Feminist Aesthetic Work as Cultural Revitalization: Facilitating Uy'Skwuluwun, in: *Feminism, Adult Education and Creative Possibility*, Bloomsbury Academic, 2022. doi:10.5040/9781350231078.ch-11.
- [40] O. Voichur, O. Hovorushchenko, A. Boyarchuk, Y. Voichur, A. Nester, Method of preprocessing information for preparing a description of art objects using artificial intelligence, *CEUR-WS* 3963 (2025) 1-14.