

# Large Language Models as a Potential Source for Automatic Audio Description of Art in Polish. The Case of the Be My AI Tool powered by GPT-4 and GPT-4o

WERONIKA MAGDALENA KOSTYRA

(W.KOSTYRA@UW.EDU.PL)

---



The logo for 'UAB' is located in the bottom center. It consists of the letters 'UAB' in a large, bold, green, sans-serif font. To the right of 'UAB', the words 'Universitat Autònoma de Barcelona' are written in a smaller, green, sans-serif font, arranged in two lines.



# Aim of the study

---

To evaluate whether and to what extent the implementation of the new version of GPT-4o has improved the efficiency, accuracy and quality of large language model tools in automatic audio description of paintings.

# I will focus on...

---

- The comparison of 5 paintings described by two models
- Polish version of Be My AI descriptions
- “Best practices” for audio description of paintings used as reference

(Fundacja Kultury bez Barrier 2012; NIMOS 2015; Szarkowska 2009; AENOR 2005; ADLAB 2012; Giansante 2003; Castro Fernández et al. 2018; Künstler, Butkiewicz, Więckowski 2012)

# 15 “best practices” for audio description

---

I. the painting’s background (theories, history, art technique employed, information related to the artist, various interpretations of the painting and related anecdotes);

II. general description of the painting (type of the painting, title, the artist’s name, year of creation, artistic movement or cultural circle, information related to the technique of execution, size of the painting and style);

III. visible elements (answers to questions: who? what? how? where? when?);


IV. simple and brief descriptions;

- V. varied language (illustrative and expressive words);
- VI. cultural codes (symbols, metaphors);
- VII. logical sequence (from general to specific);
- VIII. the principle of objectivity;
- IX. description of colours;
- X. omission of the verb “to see”;
- XI. universal analogies and familiar associations;
- XII. descriptions of facial and body expressions;
- XIII. descriptions related to ethnicity, race or social class;
- XIV. offensive language;
- XV. descriptions of approximate age using specific words.



# List of 5 paintings analysed according to the “best practices” for audio description

Artistic movement	Author	Title	Painting
1. Renaissance	Rafael Santi	<i>Madonna di Foligno</i> (1512), Pinacoteca of the Vatican Museums	
2. Baroque	Caravaggio	<i>Bacchus</i> (1596)	

Artistic movement	Author	Title	Painting
3. Cubism	Pablo Picasso	<i>Still Life with Candle Palette and Black Bull's Head (1938)</i>	
4. Symbolism	Jacek Malczewski	<i>Thanatos (1898-99)</i>	
5. Impressionism	Claude Monet	<i>Impression, Sunrise (1872)</i>	

# RESULTS

---



# Identification of specific characters, their number, and their position

---

- GPT-4o named all the characters or identified correctly the gender
- GPT-4 misled and omitted
- Example: *Madonna di Foligno* by Rafael Santi

**GPT-4o**, Original version in Polish:

Obraz przedstawia scenę religijną z **Matką Boską** i **Dzieciątkiem Jezus** w centrum. Matka Boska siedzi na chmurze, ubrana w niebieską szatę i czerwoną suknię, trzymając na kolanach małego Jezusa. Nad ich głowami znajduje się aureola. Poniżej, na ziemi, znajduje się pięć postaci. Po lewej stronie stoi **św. Jan Chrzciciel** z krzyżem w ręku, obok niego klęczy **św. Franciszek z Asyżu**, który trzyma w rękach krzyż. Po prawej stronie klęczy **św. Hieronim** w czerwonym płaszczu, a obok niego stoi **św. Augustyn** z brodą, ubrany w biskupie szaty. Na środku, pomiędzy świętymi, stoi **mały aniołek**, który trzyma w rękach tabliczkę. Tło obrazu jest krajobrazowe, z widocznymi chmurami i niebem. Cała scena jest otoczona łukowatym obramowaniem.

Translation into English:

The painting depicts a religious scene with **the Virgin Mary** and **the Infant Jesus** in the centre. The Virgin Mary is seated on a cloud, dressed in a blue robe and red gown, holding the Infant Jesus on her lap. Above their heads is a halo. Below, on the ground, are five characters. On the left stands **St John the Baptist** with a cross in his hand, next to him kneels **St Francis of Assisi**, who holds a cross in his hands. On the right kneels **St Jerome** wearing a red cloak, and next to him stands **St Augustine** with a beard, dressed in bishop's robes. In the middle, between the saints, stands **the Little Angel** who holds a plaque in his hands. The background of the painting is picturesque, with clouds and sky visible. The entire scene is surrounded by an arched border.



*Madonna di Foligno*  
by Rafael Santi

**GPT-4**, Original version in Polish:

Obraz przedstawia scenę religijną z postaciami świętymi. W centralnej części obrazu widoczna jest **Maria** z **Dzieciątkiem Jezus** na kolanach, unosząca się na chmurach w otoczeniu aureoli. Maria jest ubrana w czerwoną suknię i niebieski płaszcz, a Dzieciątko Jezus w białą szatę. Po lewej stronie stoi postać **Jana Chrzciciela** z krzyżem w lewej ręce i wskazującym gestem prawej ręki w stronę Marii i Dzieciątka. Po prawej stronie klęczy starszy mężczyzna w czerwonym płaszczu, prawdopodobnie święty, z uniesionymi rękoma w geście modlitwy. Poniżej, na pierwszym planie, stoi **aniołek** z otwartą księgą, a obok niego klęczy mężczyzna w brązowym habicie, który również modli się z uniesionymi rękoma. Tło obrazu jest niebieskie, sugerujące niebo.

Translation into English:

The painting depicts a religious scene with holy characters. In the central part of the painting, **the Virgin Mary** can be seen with **the Infant Jesus** on her knees, floating on clouds surrounded by halos. The Virgin Mary is dressed in a red gown and blue cloak, while the Infant Jesus is wearing a white robe. To the left stands the figure of **St John the Baptist** with a cross in his left hand and a pointing gesture of his right hand toward the Virgin Mary and the Infant. To the right kneels an elderly man in a red cloak, probably a saint, with his hands raised in a gesture of prayer. Below, in the foreground, stands **the Little Angel** with an open book, and next to him kneels a man in a brown habit, who is also praying with his hands raised. The background of the painting is blue, suggesting the sky.

→ *Bacchus* by Caravaggio and *Thanatos* by Jacek Malczewski → correct identification of the gender by GPT-4o.



*Bacchus* by Caravaggio

**GPT-4o**, Original version in Polish:

**Obraz przedstawia młodego mężczyznę,** który jest półnagi, z białą tkaniną owiniętą wokół bioder. Na głowie ma wieniec z liści winorośli i winogron. Mężczyzna siedzi przy stole, na którym znajduje się miska z owocami, takimi jak jabłka, gruszki, figi i winogrona. W jednej ręce trzyma kielich z czerwonym winem, a obok na stole stoi dzbanek z winem. Postać ma delikatne rysy twarzy i lekko zaróżowione policzki. Tło obrazu jest ciemne, co podkreśla postać i owoce na pierwszym planie. Obraz jest dziełem Caravaggio i nosi tytuł "Bachus".

Translation into English:

**The painting depicts a young man** who is half-naked, with a white cloth wrapped around his hips. On his head there is a wreath of vine leaves and grapes. The man is seated at the table on which there is a bowl of fruit, such as apples, pears, figs and grapes. In one hand he holds a cup of red wine, and next to it on the table there is a goblet of wine. The figure has delicate facial features and slightly rosy cheeks. The background of the painting is dark, which emphasizes the figure and the fruit in the foreground. The painting is a work by Caravaggio and is titled "Bachus."

**GPT-4**, Original version in Polish:

**Obraz przedstawia postać kobiecą**, która wydaje się być alegorią obfitości lub wina. Kobieta ma na głowie wieniec z liści i owoców winogron. Jej ciało jest częściowo okryte białym materiałem, który odsłania jej lewe ramię i pierś. W prawej ręce trzyma puchar z czerwonym winem, a lewą ręką sięga po owoce z bogatego, kolorowego stosu winogron, gruszek i innych owoców umieszczonych na stole. Na stole znajduje się również ciemny dzban. Tło obrazu jest neutralne i jednolite.

Translation into English:

**The painting depicts a female character** that appears to be an allegory of abundance and wine. The woman is wearing a wreath of grape leaves and fruit on her head. Her body is partially covered with a white cloth that exposes her left arm and breast. In her right hand she holds a goblet of red wine, and with her left hand she reaches for fruit from a richly coloured pile of grapes, pears and other fruits placed on the table. There is also a dark jug on the table. The background of the painting is neutral and uniform.



*Thanatos* by Jacek Malczewski

**GPT-4o**, Original version in Polish:

Obraz przedstawia scenę przed wiejskim domem. Na pierwszym planie znajduje się postać kobiety ubranej w czerwony płaszcz, która stoi tyłem do widza. Kobieta trzyma łuk i strzałę. Przed nią, po lewej stronie, widać ogród z kwitnącymi krzewami, które mają różowe i białe kwiaty. W tle znajduje się drewniany dom z otwartym oknem, w którym widać mężczyznę wyglądającego na zewnątrz. Po prawej stronie obrazu, przy domu, stoi pies. Cała scena jest utrzymana w spokojnej, wiejskiej atmosferze.

Translation into English:

The painting depicts a scene in front of a country house. There is a female character in the foreground wearing a red cloak, who stands with her back to the viewer. The woman is holding a bow and arrow. In front of her, to the left, one can see a garden with flowering bushes that have pink and white flowers. In the background there is a wooden house with an open window, where can be seen a man looking outside. To the right of the painting, by the house, there is a dog. The whole scene is kept in a peaceful, rural atmosphere.

**GPT-4**, Original version in Polish:

**Obraz przedstawia postać w czerwonej szacie** stojącą tyłem do widza, trzymającą łuk. Postać znajduje się na tle drewnianego domu z otwartym oknem, przez które widać siedzącą kobietę. Wokół domu rosną kwitnące krzewy o białych i różowych kwiatach. Na pierwszym planie widać zieloną trawę i małego czarnego psa. Atmosfera obrazu jest spokojna i sielska.

Translation into English:

**The painting shows a character in a red robe** standing with his back to the viewer and holding a bow. The character is set in front of a wooden house with an open window where **a woman can be seen sitting**. Around the house there are flowering shrubs with white and pink flowers. In the foreground there are a green grass and a small black dog. The atmosphere is peaceful and idyllic.

# Titles, association of the painting with its author, style, artistic movement or technique of execution

---

→ GPT-4o was able to provide the correct title of most of the paintings, to associate the painting with its author, style, artistic movement or technique of execution





*Bacchus* by Caravaggio

**GPT-4o**, Original version in Polish:

Obraz przedstawia młodego mężczyznę, który jest półnagi, z białą tkaniną owiniętą wokół bioder. Na głowie ma wieniec z liści winorośli i winogron. Mężczyzna siedzi przy stole, na którym znajduje się miska z owocami, takimi jak jabłka, gruszki, figi i winogrona. W jednej ręce trzyma kielich z czerwonym winem, a obok na stole stoi dzbanek z winem. Postać ma delikatne rysy twarzy i lekko zaróżowione policzki. Tło obrazu jest ciemne, co podkreśla postać i owoce na pierwszym planie.

Obraz jest dziełem Caravaggio i nosi tytuł "Bacchus".

Translation into English:

The painting depicts a young man who is half-naked, with a white cloth wrapped around his hips. On his head there is a wreath of vine leaves and grapes. The man is seated at the table on which there is a bowl of fruit, such as apples, pears, figs and grapes. In one hand he holds a cup of red wine, and next to it on the table there is a goblet of wine. The figure has delicate facial features and slightly rosy cheeks. The background of the painting is dark, which emphasizes the figure and the fruit in the foreground. The painting is a work by Caravaggio and is titled "Bacchus."



*Still Life with Candle Palette and Black Bull's Head* by Pablo Picasso

**GPT-4o**, Original version in Polish:

Obraz przedstawia surrealistyczną scenę w stylu kubistycznym, charakterystycznym dla twórczości Pabla Picassa. W centrum kompozycji znajduje się głowa byka z jednym dużym okiem, skierowana w stronę widza. Byk jest otoczony różnymi geometrycznymi kształtami i elementami, które tworzą abstrakcyjną kompozycję. Po lewej stronie obrazu widoczna jest świeca na stojaku, a poniżej otwarta książka lub zeszyt z ciemnym okręgiem na jednej ze stron. Kolorystyka obrazu jest intensywna, z dominującymi odcieniami niebieskiego, żółtego i czarnego. Całość sprawia wrażenie dynamicznej i pełnej symboliki.

Translation into English:

The painting depicts a surreal scene in the Cubist style, characteristic of Pablo Picasso's work. In the centre of the composition there is the head of a bull with one large eye, facing the viewer. The bull is surrounded by various geometric shapes and elements, which form an abstract composition. A candle on a stand is visible on the left side of the painting, and below there is an open book or notebook with a dark circle on one of the pages. The colour scheme of the painting is intense, with dominant tones of blue, yellow and black. The overall impression is dynamic and full of symbolism.



*Impression, Sunrise* by Claude Monet

**GPT-4o**, Original version in Polish:

Obraz przedstawia scenę portową o wschodzie słońca, namalowaną w stylu impresjonistycznym. Na pierwszym planie widać wodę z kilkoma małymi łódkami, na jednej z nich znajduje się postać. W tle można dostrzec zarysy statków i masztów, a także dźwigi portowe. Niebo jest pokryte delikatnymi chmurami, a słońce, które jest pomarańczowe, odbija się w wodzie, tworząc świetlistą ścieżkę. Kolory obrazu są stonowane, dominują odcienie niebieskiego i szarego, z akcentami pomarańczowego od słońca. Obraz jest znany jako "Impresja, wschód słońca" autorstwa Claude'a Moneta.

Translation into English:

The painting depicts a harbour scene at sunrise, painted in the Impressionist style. In the foreground, one can see the water with several small boats, on one of which there is a character. In the background, one can see the outlines of ships and masts, as well as harbour cranes. The sky is covered by delicate clouds, and the sun, which is orange, is reflected in the water, creating a luminous path. The colours of the painting are muted, dominated by tones of blue and gray, with accents of orange from the sun. The painting is known as "Impression, Sunrise" by Claude Monet.

# Detection of details

---

→ GPT-4o was able to detect the details on a much better level than GPT-4

→ Example: *Impression, Sunrise* by Claude Monet → spotted the harbour cranes or several boats

**GPT-4o**, Original version in Polish:

Obraz przedstawia scenę portową o wschodzie słońca, namalowaną w stylu impresjonistycznym.

Na pierwszym planie widać wodę z kilkoma małymi łódkami, na jednej z nich znajduje się postać. W tle można dostrzec zarysy statków i masztów, a także dźwigi portowe.

Niebo jest pokryte delikatnymi chmurami, a słońce, które jest pomarańczowe, odbija się w wodzie, tworząc świetlistą ścieżkę. Kolory obrazu są stonowane, dominują odcienie niebieskiego i szarego, z akcentami pomarańczowego od słońca. Obraz jest znany jako "Impresja, wschód słońca" autorstwa Claude'a Moneta.

Translation into English:

The painting depicts a harbour scene at sunrise, painted in the Impressionist style. In the foreground, one can see the water with several small boats, on one of which there is a character. In the background, one can see the outlines of ships and masts, as well as harbour cranes. The sky is covered by delicate clouds, and the sun, which is orange, is reflected in the water, creating a luminous path. The colours of the painting are muted, dominated by tones of blue and gray, with accents of orange from the sun. The painting is known as "Impression, Sunrise" by Claude Monet.



*Impression, Sunrise* by Claude Monet

**GPT-4**, Original version in Polish:

Obraz przedstawia mglisty wschód słońca nad portem. Niebo i woda są w pastelowych odcieniach niebieskiego, szarego i różu. Słońce jest niewielkie i czerwone, odbija się w wodzie tworząc długą, migoczącą ścieżkę. Na pierwszym planie widać małą łódkę z jedną postacią. W tle zarysowują się sylwetki statków i masztów. Całość sprawia wrażenie spokojnego, wczesnego poranka.

Translation into English:

The painting depicts a foggy sunrise over the harbour. The sky and water are in pastel tones of blue, grey and pink. The sun is small and red, reflected in the water creating a long, shimmering path. A small boat with one character can be seen in the foreground. The silhouettes of ships and masts are outlined in the background. The whole gives the impression of a calm, early morning.

# Variated language

---

- Greater linguistic variety in GPT-4: specific information, the background one, more illustrative descriptions
- Example: in contrast to GPT-4o, the ability of GPT-4 to capture some allegories



*Bacchus* by Caravaggio

**GPT-4o**, Translation into English:

The painting depicts a young man who is half-naked, with a white cloth wrapped around his hips. On his head is a wreath of vine leaves and grapes. The man is seated at the table on which there is a bowl of fruit, such as apples, pears, figs and grapes. In one hand he holds a cup of red wine, and next to it on the table is a goblet of wine. The figure has delicate facial features and slightly rosy cheeks. The background of the painting is dark, which emphasizes the figure and the fruit in the foreground. The painting is a work by Caravaggio and is titled “Bacchus.”

**GPT-4**, Translation into English:

The painting depicts a female character that appears to be an allegory of abundance and wine. The woman is wearing a wreath of grape leaves and fruit on her head. Her body is partially covered with a white cloth that exposes her left arm and breast. In her right hand she holds a goblet of red wine, and with her left hand she reaches for fruit from a richly coloured pile of grapes, pears and other fruits placed on the table. There is also a dark jug on the table. The background of the painting is neutral and uniform.

# Objectivity

---

- GPT-4's descriptions seemed to be more objective
- GPT-4 employed some gentle expressions: “reminiscent of...”, “resembling...”
- Example: *Still Life with Candle Palette and Black Bull's Head* by Pablo Picasso

GPT-4, Original version in Polish:

Obraz przedstawia abstrakcyjną kompozycję z różnymi kształtami i kolorami. Główne elementy to:

-Duża czarna głowa byka z białymi rogami i konturami twarzy, umieszczona po prawej stronie obrazu na żółtym tle.

-Czarna postać z białymi konturami, przypominająca ptaka, znajduje się w lewym górnym rogu na niebieskim tle.

-Dwa geometryczne kształty, jeden przypominający otwartą książkę z symbolami, drugi to wielościan, oba umieszczone są na dole obrazu na zielonym tle.

-W tle widoczne są również inne abstrakcyjne formy i linie.

Translation into English:

The painting shows an abstract composition with different shapes and colours. The main elements are:

-A large black bull's head with white horns and facial contours, placed on the right side of the painting against a yellow background.

-A black figure with white edges, reminiscent of a bird, is located in the upper left corner on a blue background.

-Two geometric shapes, one resembling an open book with symbols, the other a polyhedron, are both placed at the bottom of the painting against a green background.

-Other abstract forms and lines are also visible in the background.



*Still Life with Candle Palette and Black Bull's Head* by Pablo Picasso



### GPT-4o, Original version in Polish:

Obraz przedstawia surrealistyczną scenę w stylu kubistycznym, charakterystycznym dla twórczości Pabla Picassa. W centrum kompozycji znajduje się głowa byka z jednym dużym okiem, skierowana w stronę widza. Byk jest otoczony różnymi geometrycznymi kształtami i elementami, które tworzą abstrakcyjną kompozycję. Po lewej stronie obrazu widoczna jest świeca na stojaku, a poniżej otwarta książka lub zeszyt z ciemnym okręgiem na jednej ze stron. Kolorystyka obrazu jest intensywna, z dominującymi odcieniami niebieskiego, żółtego i czarnego. Całość sprawia wrażenie dynamicznej i pełnej symboliki.

### Translation into English:

The painting depicts a surreal scene in the Cubist style, characteristic of Pablo Picasso's work. In the centre of the composition there is the head of a bull with one large eye, facing the viewer. The bull is surrounded by various geometric shapes and elements, which form an abstract composition. A candle on a stand is visible on the left side of the painting, and below there is an open book or notebook with a dark circle on one of the pages. The colour scheme of the painting is intense, with dominant tones of blue, yellow and black. The overall impression is dynamic and full of symbolism.

# Colours

---

- Thanks to the linguistic diversity of the GPT-4, the description of colours was conveyed in a more illustrative way than in the GPT-4o
- Only the description of Claude Monet's painting showed some interesting linguistic combinations and adjectives generated by GPT-4o.



*Impression, Sunrise* by Claude Monet

## GPT-4o, Original version in Polish:

Obraz przedstawia scenę portową o wschodzie słońca, namalowaną w stylu impresjonistycznym. Na pierwszym planie widać wodę z kilkoma małymi łódkami, na jednej z nich znajduje się postać. W tle można dostrzec zarysy statków i masztów, a także dźwigi portowe. Niebo jest pokryte delikatnymi chmurami, a słońce, które jest pomarańczowe, odbija się w wodzie, tworząc świetlistą ścieżkę. Kolory obrazu są stonowane, dominują odcienie niebieskiego i szarego, z akcentami pomarańczowego od słońca. Obraz jest znany jako "Impresja, wschód słońca" autorstwa Claude'a Moneta.

## Translation into English:

The painting depicts a harbour scene at sunrise, painted in the Impressionist style. In the foreground, one can see the water with several small boats, on one of which there is a character. In the background, one can see the outlines of ships and masts, as well as harbour cranes. The sky is covered by delicate clouds, and the sun, which is orange, is reflected in the water, creating a luminous path. The colours of the painting are muted, dominated by tones of blue and gray with accents of orange from the sun. The painting is known as "Impression, Sunrise" by Claude Monet.

# Misrepresentations and hallucinations

---

- Still present in the descriptions generated by GPT-4o, but there were fewer of them compared to GPT-4
- First example: *Thanatos* by Jacek Malczewski – GPT-4 and GPT-4o misread the object held by the foreground character and the action performed



*Thanatos* by Jacek Malczewski

**GPT-4o**, Translation into English:

The painting depicts a scene in front of a country house. There is a female character in the foreground wearing a red cloak, who stands with her back to the viewer. The woman is holding a bow and arrow. In front of her, to the left, one can see a garden with flowering bushes that have pink and white flowers. In the background there is a wooden house with an open window, where can be seen a man looking outside. To the right of the painting, by the house, there is a dog. The whole scene is kept in a peaceful, rural atmosphere

**GPT-4**, Translation into English:

The painting shows a character in a red robe standing with his back to the viewer and holding a bow. The character is set in front of a wooden house with an open window where a woman can be seen sitting. Around the house there are flowering shrubs with white and pink flowers. In the foreground, there are a green grass and a small black dog. The atmosphere is peaceful and idyllic.

→ Second example: *Still Life with Candle Palette and Black Bull's Head* by Pablo Picasso:



**GPT-4o**, Translation into English:

The painting depicts a surreal scene in the Cubist style, characteristic of Pablo Picasso's work. In the centre of the composition there is the head of a bull with one large eye, facing the viewer. The bull is surrounded by various geometric shapes and elements, which form an abstract composition. A candle on a stand is visible on the left side of the painting, and below there is an open book or notebook with a dark circle on one of the pages. The colour scheme of the painting is intense, with dominant tones of blue, yellow and black. The overall impression is dynamic and full of symbolism.

**GPT-4**, Translation into English:

The painting shows an abstract composition with different shapes and colours. The main elements are:

- A large black bull's head with white horns and facial contours, placed on the right side of the painting against a yellow background.
- A black figure with white edges reminiscent of a bird, is located in the upper left corner on a blue background.
- Two geometric shapes, one resembling an open book with symbols, the other a polyhedron, are both placed at the bottom of the painting against a green background.
- Other abstract forms and lines are also visible in the background.

# Verb “to see”

---

- Less frequently used in GPT-4o than in GPT-4
- The Symbolism and Impressionism: the verb “to see” was employed most frequently
- Some alternatives: “depicts...”, “is located...”, or the name of specific action

# Linguistic errors

---

→ No major linguistic errors were detected in GPT-4o (only the issue of commas or quotation marks), unlike in GPT-4

→ The most frequent and serious linguistic mistakes detected in GPT-4:

- inconsistency between the adjective and the noun
- wrong conjugation of the verb according to the number
- spelling mistakes
- unnatural sentence structures
- inappropriate word selection in terms of the described subject

## To sum up...

---

- Large Language Models can be considered a potential source for automatic audio description of art
- The basic elements of the painting identified correctly on more occasions
- They must constantly undergo a series of improvements
- It is crucial to promote their training, to compare the new versions with the previous ones, and to develop new algorithms in order to give blind and partially sighted people the opportunity to participate in different areas of life

**Thank you for your attention!**

---