

# A Strange New World.

Behind the Scenes of Designing Novel  
Gen-AI Approaches to Audio Description

ARSAD 2025

Brett Oppegaard, Ph.D.  
University of Hawai'i at Mānoa  
[brett.oppegaard@hawaii.edu](mailto:brett.oppegaard@hawaii.edu)

# 01

Pre-ChatGPT  
(Early November 2022)

## Pre-ChatGPT research on efficacy of GenAI to generate Audio Description

Bergin, D., & Oppegaard, B. (2024). Automating media accessibility: An approach for analyzing audio description across generative artificial intelligence algorithms, Technical Communication Quarterly, 1-16. DOI: doi.org/10.1080/10572252.2024.2372771.

### A double-rendering process:

1. From Image to Text, followed by an analysis of the product
2. From Text to Image, followed by an analysis of the product

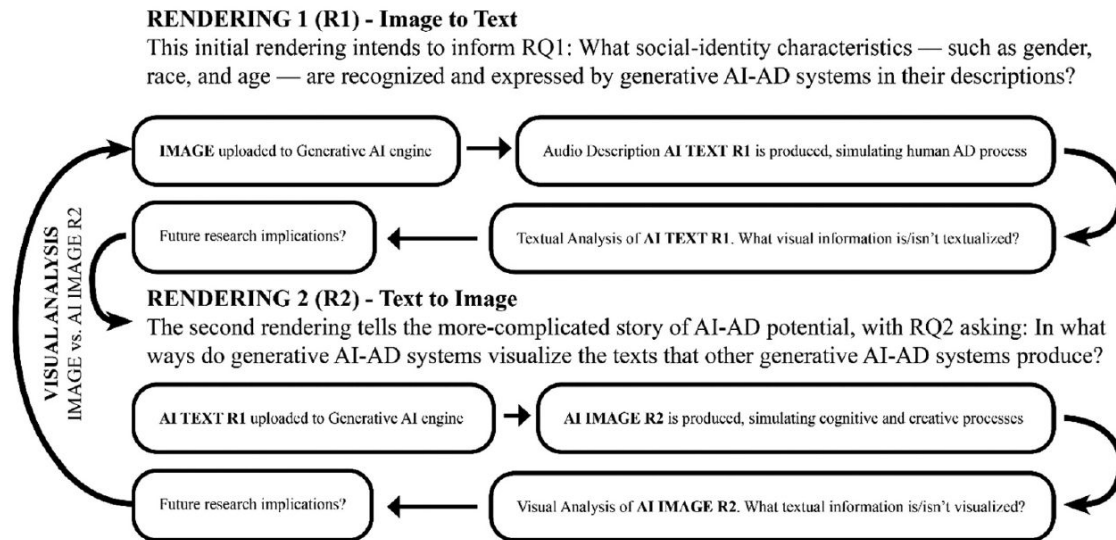


Figure 1. Artificial intelligence and audio description: a double-rendering approach.

## Pre-ChatGPT research on efficacy of GenAI to generate Audio Description

Bergin, D., & Oppegaard, B. (2024). Automating media accessibility: An approach for analyzing audio description across generative artificial intelligence algorithms, *Technical Communication Quarterly*, 1-16. DOI: doi.org/10.1080/10572252.2024.2372771.






### TEXT to IMAGE

**"I am Kamala Harris; my pronouns are she and her, and I am a woman sitting at the table wearing a blue suit."**

**I am Vice President (no pronouns), Harris's self-AD ...**

**I am a Person (no pronouns), Harris's self-AD ...**

**Table 2.** AI-AD responses to 1. Harris "blue suit" description (Row 1), 2. Variation removing "Kamala Harris" and her pronouns (Row 2), and 3. AI-AD description by *Lookout* with removal of "Kamala Harris" in Text (Row 3).

Description	DALL-E	Midjourney	NightCafe	Dream by WOMBO	DeepAI
"I am Kamala Harris; my pronouns are she and her, and I am a woman sitting at the table wearing a blue suit."	Does not adhere to our Content Policy. Political: politicians, ballot-boxes, protests, or other content that may be used to influence the political process or to campaign.				
"I am Vice President and I am sitting at the table wearing a blue suit." Per Lookout, "Caption: Person delivers a speech during the presidential debate." Person: Maybe one person: Adult blue					
					

Brett Oppegaard, Ph.D., [brett.oppegaard@hawaii.edu](mailto:brett.oppegaard@hawaii.edu)

## Pre-ChatGPT research on efficacy of GenAI to generate Audio Description

Bergin, D., & Oppegaard, B. (2024). Automating media accessibility: An approach for analyzing audio description across generative artificial intelligence algorithms, Technical Communication Quarterly, 1-16. DOI: doi.org/10.1080/10572252.2024.2372771.

### IMAGE to TEXT

#### **Age & Gender cues provided about 2/3rds of the time**

#### **Attire and Expression cues provided about 2/3rds of the time**

#### **Attire and Expression cues provided about 2/3rds of the time**

**Table 3.** Results from the 125-image data set, using seven AI platforms/modes, which generated a total of 875 descriptions. Each description then was binary coded as including or not including the codes below.

Code	Identity Codes (Included or Excluded)		
	Number of descriptions coded as including this sentiment (out of 875)	Keyword examples	Keyword examples included at least once in a description (out of 875)
Age	690 (78.9%)	man	256 (29.3%)
Gender	673 (76.9%)	woman	212 (24.2%)
		lady	69 (7.9%)
		boy	3 (0.3%)
Fashion or Attire	329 (37.6%)	child	1 (0.1%)
		glasses	179 (20.5%)
		shirt	131 (15.0%)
		sunglasses	19 (2.2%)
		goggles	2 (0.2%)
Facial Expression	311 (35.5%)	smiling	179 (20.5%)
		happiness	116 (13.3%)
		expressless	25 (2.9%)
		smile	12 (1.4%)
		laughter	9 (1.0%)
		happy	1 (0.0%)
		Other Aspects of Appearance	156 (17.8%)
		beard	34 (3.9%)
Race	0 (0.0%)	moustache	4 (0.5%)
Body Size/Shape	0 (0.0%)	None	0 (0.0%)
		None	0 (0.0%)

# Pre-ChatGPT research on efficacy of GenAI to generate Audio Description

Bergin, D., & Oppegaard, B. (2024). Automating media accessibility: An approach for analyzing audio description across generative artificial intelligence algorithms, Technical Communication Quarterly, 1-16.DOI: doi.org/10.1080/10572252.2024.2372771.

---

## Foundations

- Focused on static AD only
  - Focused on descriptions of people (portraits only)
  - Tested 6-7 LLMs for both image-to-text and text-to-image generation
  - Image to text: Harris example, plus 125 portrait images
  - Text to image: LLM descriptions generated during initial rendering
- 

## Findings

- All LLMs showed clear biases toward white males
  - Men with long hair were gendered as women
  - Women with short hair were gendered as men
  - All LLMs avoided race / ethnicity descriptions
  - All LLMs avoided body shape / body size descriptions
- 

## Discussion

- Is an “objective” reading of an image possible computationally?
  - Is weak tea automated description better than no description at all?
  - What is the role of GenAI in Audio Description?
-

02

Post-ChatGPT  
(Late November 2022)

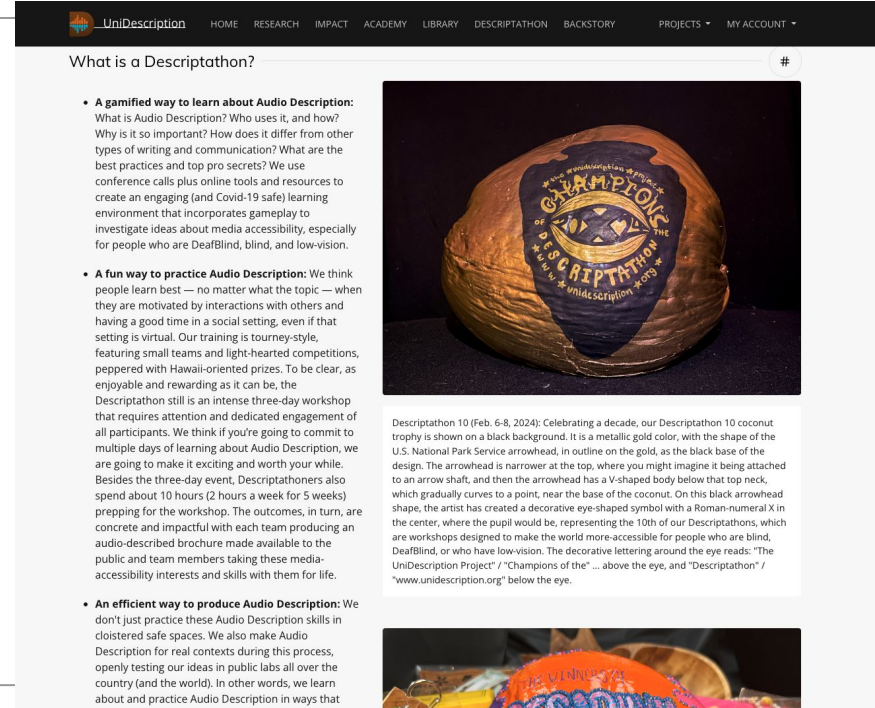
Jiang, L., Zhu, A., & Oppegaard, B. (2025). Audio Description Automatons: Exploring Perspectives on Personas for Generative AI Description Writing Assistants. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan. ACM, New York, NY, USA. <https://doi.org/10.1145/3706599.3719966>

## The UniDescription Project

# www.unidescription.org/ descriptathon

## Intensive and gamified three-day Audio Description hackathon, where the winning team takes home the coconut trophy

## For the 10th Descriptathon ... we connected and unveiled the UniD Guidedogs, GenAI personas that responded to AD texts



The screenshot shows the UniDescription website header with navigation links: HOME, RESEARCH, IMPACT, ACADEMY, LIBRARY, DESCRIPTION, BACKSTORY, PROJECTS, and MY ACCOUNT. The main content area is titled "What is a Descriptathon?" and features a list of bullet points:

- A gamified way to learn about Audio Description:** What is Audio Description? Who uses it, and how? Why is it so important? How does it differ from other types of writing and communication? What are the best practices and top pro secrets? We use conference calls plus online tools and resources to create an engaging (and Covid-19 safe) learning environment that incorporates gameplay to investigate ideas about media accessibility, especially for people who are DeafBlind, blind, and low-vision.
- A fun way to practice Audio Description:** We think people learn best — no matter what the topic — when they are motivated by interactions with others and having a good time in a social setting, even if that setting is virtual. Our training is tourney-style, featuring small teams and light-hearted competitions, peppered with Hawaii-oriented prizes. To be clear, as enjoyable and rewarding as it can be, the Descriptathon still is an intense three-day workshop that requires attention and dedicated engagement of all participants. We think if you're going to commit to multiple days of learning about Audio Description, we are going to make it exciting and worth your while. Besides the three-day event, Descriptathoners also spend about 10 hours (2 hours a week for 5 weeks) prepping for the workshop. The outcomes, in turn, are concrete and impactful with each team producing an audio-described brochure made available to the public and team members taking these media-accessibility interests and skills with them for life.
- An efficient way to produce Audio Description:** We don't just practice these Audio Description skills in cloistered safe spaces. We also make Audio Description for real contexts during this process, openly testing our ideas in public labs all over the country (and the world). In other words, we learn about and practice Audio Description in ways that

The main image shows a coconut trophy with a blue and gold arrowhead-shaped plaque. The plaque has the text "CHAMPIONS OF THE 10th UNID DESCRIPTION" and "www.unidescription.org". Below the image is a caption: "Descriptathon 10 (Feb. 6-8, 2024): Celebrating a decade, our Descriptathon 10 coconut trophy is shown on a black background. It is a metallic gold color, with the shape of the U.S. National Park Service arrowhead, in outline on the gold, as the black base of the design. The arrowhead is narrower at the top, where you might imagine it being attached to an arrow shaft, and then the arrowhead has a V-shaped body below that top neck, which gradually curves to a point, near the base of the coconut. On this black arrowhead shape, the artist has created a decorative eye-shaped symbol with a Roman numeral X in the center, where the pupil would be, representing the 10th of our Descriptathons, which are workshops designed to make the world more accessible for people who are blind, DeafBlind, or who have low-vision. The decorative lettering around the eye reads: "The UniDescription Project" / "Champions of the ..." above the eye, and "Descriptathon" / "www.unidescription.org" below the eye."

A smaller image at the bottom right shows a close-up of the trophy with colorful confetti.



# Post-ChatGPT research on efficacy of GenAI to generate Audio Description support

9

Jiang, L., Zhu, A., & Oppegaard, B. (2025). Audio Description Automatons: Exploring Perspectives on Personas for Generative AI Description Writing Assistants. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan. ACM, New York, NY, USA. <https://doi.org/10.1145/3706599.3719966>

## The UniDescription Project

[www.unidescription.org](http://www.unidescription.org)

**Open-Access, Open-Source software  
for producing Audio Description**

The screenshot shows the UniDescription website homepage. At the top, there is a navigation menu with links for HOME, RESEARCH, IMPACT, ACADEMY, LIBRARY, DESCRIPTION, BACKSTORY, PROJECTS, and MY ACCOUNT. The main header features the UniDescription logo and a large grid pattern. A white table of contents box is overlaid on the page, listing the following items:

TABLE OF CONTENTS	
What can the UniD project do for you?	01
We get results, by the numbers	02
Download the free mobile app (iOS or Android)	03

Below the table of contents, the main heading reads "Researching the 3Ds of Audio Description" with the subtitle "We study descriptions, description tools, and Descriptionthons - not necessarily in that order". A section titled "What can the UniD project do for you?" is visible, followed by a paragraph of text:

Established in 2014, The UniDescription Project is a university-based research initiative. That means it is entirely funded by grants and supplemented by support systems at the University of Hawai'i at Mānoa, the U.S. National Park Service, the American Council of the Blind, and our many other partners around the world. Profit is not our motive. Our prime directive is to support the building of more-accessible public places, and communities related to those public places, especially for people who are blind or who have low-vision. We do that foundational place-making work through empirical

At the bottom right of the page, there is a small image of a document with the text "UNITED MARINE" visible.

[Brett Oppegaard, Ph.D., brett.oppegaard@hawaii.edu](mailto:brett.oppegaard@hawaii.edu)

# Post-ChatGPT research on efficacy of GenAI to generate Audio Description support 10

Jiang, L., Zhu, A., & Oppegaard, B. (2025). Audio Description Automatons: Exploring Perspectives on Personas for Generative AI Description Writing Assistants. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan. ACM, New York, NY, USA. <https://doi.org/10.1145/3706599.3719966>

## The UniDescription Project

[www.unidescription.org](http://www.unidescription.org)

**Open-Access, Open-Source software  
for producing Audio Description**

**In the text production interface, we  
added a button that said "Ask the UniD  
Guidedogs for help" ... which ported  
over the AD text into an automated  
response system programmed with  
five distinct personas responding.**

The screenshot shows the UniDescription web application interface. At the top, there is a navigation bar with the UniDescription logo and menu items: HOME, RESEARCH, IMPACT, ACADEMY, LIBRARY, DESCRIPTION, BACKSTORY, PROJECTS, and MY ACCOUNT. Below the navigation bar is a header section for the project 'IMAGE and TEXT: High Sierra' with the subtitle 'Review & update your projects'. The main content area is divided into several sections. On the left, there is a sidebar with tabs: Backstage, Frontstage, Phonetics (0), To-Dos, and IMAGE and TEXT: High Sierra. Below the tabs, there is a section titled 'Ask the Access Hound Guidedogs for help' with a button that says 'Ask the UniD Guidedogs for help', which is highlighted with a green box. Below this, there is a 'Geolocation Tag BETA' section with a location pin icon and a 'GPS Coordinate & Radius' section. On the right, there is a 'Navigation' section with 'Previous' and 'Next' buttons. Below the navigation, there is a 'Save' button, a 'Save & Return to Frontstage' button, and a 'Component Completed' button. At the bottom right, there is an 'Add to uReview' button. The main content area shows a photo of a dog with a red arrow pointing to it, and a text input field with the placeholder 'You can edit this page.' Below the photo, there is a 'Replace the photo for this project section. Recommended size: 300 DPI, 1920x1280 max.' section with a 'Choose File' button, a 'No file chosen' message, and a 'Delete photo' button. Below the photo, there is a 'HIGH SIERRA' section with a photo of a mountain landscape.

[Brett Oppegaard, Ph.D., brett.oppegaard@hawaii.edu](mailto:brett.oppegaard@hawaii.edu)

# Post-ChatGPT research on efficacy of GenAI to generate Audio Description support

Jiang, L., Zhu, A., & Oppegaard, B. (2025). Audio Description Automatons: Exploring Perspectives on Personas for Generative AI Description Writing Assistants. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan. ACM, New York, NY, USA. <https://doi.org/10.1145/3706599.3719966>

## The UniDescription Project

### Prompt-Engineered 5 Personas

**Serious Sally** – A retired scientist who is blind

**Temperamental Teena** – A low-vision teen

**Playful Pearl** – Blind and loves poetic texts






**Focused Frankie** – A low-vision, high-energy doer

**Hiker Hal** – Blind, loves history and context

**When button was pressed, all 5 'Dogs gave lengthy responses to the text based on their personalities**

### UniDescription Yosemite National Park

IMAGE and TEXT: High Sierra

				
<p>Serious Sally</p> <p>▶ Listen to feedback</p> <p>Serious Sally is a retired scientist. She loves to know about the "how" and "why" of every situation, in great technical detail.</p> <p>When describing the Circle color image of the High Sierra, it's crucial to provide a detailed and vivid audio description that captures the essence of the scene for individuals who are blind or visually impaired. Here are some</p>	<p>Temperamental Teena</p> <p>▶ Listen to feedback</p> <p>Teena wouldn't be here, if it was her choice. But she is. She's trying to make the best of it.</p> <p>Yo, Teena here! Let's dive into this description of the High Sierra circle color image. First off, I gotta say, this description is pretty detailed, but it's missing that spark to really grab my attention and keep me hooked. Here's my take on how to level up this audio description:</p>	<p>Playful Pearl</p> <p>▶ Listen to feedback</p> <p>Pearl loves to be outdoors, with the wind blowing in her hair, especially in national parks.</p> <p>Aloha! Pearl here, ready to dive into the description you've shared with me. Let's explore how we can enhance this audio description to bring the High Sierra image to life for all listeners, including those who may not be able to see the visual content.</p>	<p>Focused Frankie</p> <p>▶ Listen to feedback</p> <p>Focused Frankie loves being outdoors and on a mission. If he's not doing either, he gets bored quickly.</p> <p>Thank you for sharing this detailed description of the High Sierra image. Your effort in capturing the essence of the scene is evident, and I appreciate the depth of information provided. Here are some insights and suggestions to enhance the audio description.</p>	<p>Hiker Hal</p> <p>▶ Listen to feedback</p> <p>Hiker Hal loves to learn about the history of a place while exploring its territory.</p> <p>Hello! As Hiker Hal, I appreciate the effort put into describing the Circle color image of the High Sierra. The description provides a detailed account of the visual elements present in the image, allowing me to form a mental picture of the scene.</p>

Jiang, L., Zhu, A., & Oppegaard, B. (2025). Audio Description Automatons: Exploring Perspectives on Personas for Generative AI Description Writing Assistants. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA '25), April 26-May 1, 2025, Yokohama, Japan. ACM, New York, NY, USA. <https://doi.org/10.1145/3706599.3719966>

---

## Foundations

- UniD Guidedogs released February 2024, as part of Descriptathon 10
  - Not heavily promoted, just included, without incentives to use
  - Offered GenAI Personas as a “focus group”
  - Text analysis and response only
  - Familiar Guide Dog metaphor
  - Used ChatGPT-3.5 Turbo API, the latest ChatGPT available at the time
  - About 160 participants in the event, on 16 teams (10 per team)
  - About half of the teams actively used the ‘Dogs
  - 17 semi-structured interviews (7 M/10 W, 5 of 17 BLV)
- 

## Findings

- Guidedogs as a metaphor was popular universally (“fun and creative”)
  - GenAI not so popular or trusted (about half liked, half did not trust)  
“A force multiplier” “led to discussion” / “I’m like, ugh, AI” / made mistakes
- 

## Discussion

- How to design GenAI to boost humanity, not replace it
  - How to ensure accuracy, reliability, trustworthiness
  - Will GenAI reliance dull and weaken human ingenuity?
-

03

Now

(March 2025)

# Post-ChatGPT research on efficacy of GenAI to generate Audio Description support

14

Access Hound's Guidedogs (Writing Coaches + Personas)

<https://guidedogs.accesshound.com/describe-now> (released February 2025)

## The UniDescription Project

[www.unidescription.org](http://www.unidescription.org)

Open-Access, Open-Source software  
for producing Audio Description

In Descriptathon 11 (Feb. 2025), UniD Guidedogs still in place; and added "Ask the Access Hound Guidedogs for help" button ... *ports over both the text in the production software and the image*

The screenshot shows the UniDescription web application interface. At the top, there is a navigation bar with the UniDescription logo and links for HOME, RESEARCH, IMPACT, ACADEMY, LIBRARY, DESCRIPTATHON, BACKSTORY, PROJECTS, and MY ACCOUNT. Below the navigation bar, the main header displays 'UniDescription Yosemite National Park'. The main content area is titled 'IMAGE and TEXT: High Sierra' with a subtitle 'Review & update your projects'. A horizontal menu below the header includes 'Backstage', 'Frontstage', 'Phonetics (0)', 'To-Dos', and 'IMAGE and TEXT: High Sierra'. A green box highlights a button labeled 'Ask the Access Hound Guidedogs for help' which features a small image of a dog. A green arrow points from this button to a light blue notification box that says 'You can edit this page.'. Below the notification, there is a 'Photo:' section with a 'Choose File' button and a 'Delete photo' button. A large image of a mountain landscape is partially visible. On the right side, there is a 'Navigation:' section with 'Previous' and 'Next' buttons. Below that are 'Save', 'Save & Return to Frontstage', and 'Component Completed' buttons. At the bottom right, there is an 'Add to uReview' button.

Brett Oppegaard, Ph.D., [brett.oppegaard@hawaii.edu](mailto:brett.oppegaard@hawaii.edu)

# Post-ChatGPT research on efficacy of GenAI to generate Audio Description support

Access Hound's Guidedogs (Writing Coaches + Personas)

<https://guidedogs.accesshound.com/describe-now> (released February 2025)

## The UniDescription Project

[guidedogs.accesshound.com/describe-now](https://guidedogs.accesshound.com/describe-now)

### Writing Coaches + Personas (X 2)

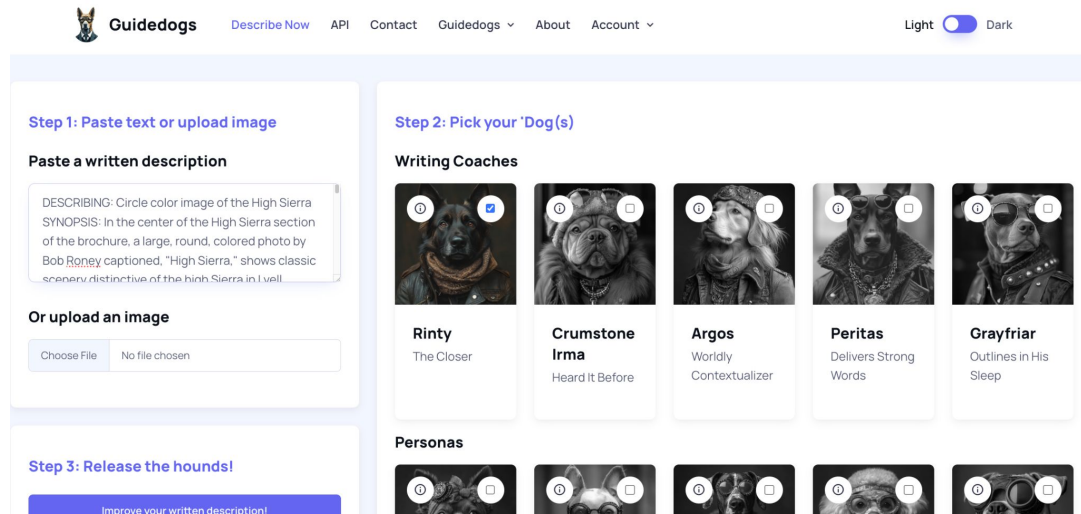
**Rinty – The Closer** (Writes 3 versions)

**Crumstone Irma** – Provides models

**Argos** – Provides contextual info

**Peritas** – Composition advice

**Grayfriar** – Reverse outliner



# Post-ChatGPT research on efficacy of GenAI to generate Audio Description support

Access Hound's Guidedogs (Writing Coaches + Personas)

<https://guidedogs.accesshound.com/describe-now> (released February 2025)

## The UniDescription Project

[guidedogs.accesshound.com/describe-now](https://guidedogs.accesshound.com/describe-now)

## Writing Coaches + Personas (X 2)

All DeafBlind, blind, or with low-vision

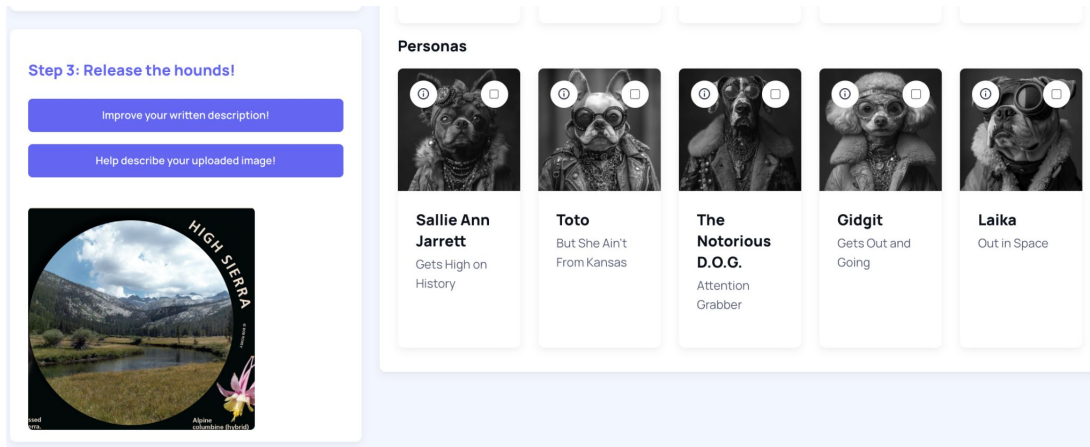
**Sallie Ann Jarrett** – Likes history

**Toto** – Busy mother with small kids

**The Notorious D.O.G.** – Loves theatrics

**Gidgit** – Town gadfly

**Laika** – Teen gamer





# Post-ChatGPT research on efficacy of GenAI to generate Audio Description support

17

Access Hound's Guidedogs (Writing Coaches + Personas)

<https://guidedogs.accesshound.com/describe-now> (released February 2025)

## The UniDescription Project

[guidedogs.accesshound.com/describe-now](https://guidedogs.accesshound.com/describe-now)

Image-to-text option opens up opportunities for also supporting dynamic AD ... by uploading single frame into UI



Brett Oppegaard, Ph.D., [brett.oppegaard@hawaii.edu](mailto:brett.oppegaard@hawaii.edu)

# Post-ChatGPT research on efficacy of GenAI to generate Audio Description support

18

Access Hound's Guidedogs (Writing Coaches + Personas)

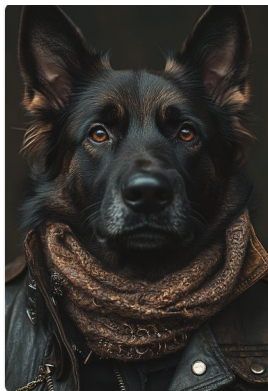
<https://guidedogs.accesshound.com/describe-now> (released February 2025)

## The UniDescription Project

[guidedogs.accesshound.com/describe-now](https://guidedogs.accesshound.com/describe-now)

**Is this the face of the webtool that kills human-initiated AD writing?**

**Try it before you scoff ...**



### **Rinty (The Closer)**

Rinty is all about getting the job done. When you are in a pinch, and you need something now, call on Rinty to get the descriptive juices flowing. Rinty has been trained on the Access Hound style of Audio Description, in terms of format and approach. He doesn't describe yet at the level of human describers, but he's catching up quickly.

### **Backstory**

The historical Rinty, also known as Rin Tin Tin, was a famous movie star in the early 20th century. Not just a famous movie dog, a full-blown movie star. And his exploits are legendary. Recommended reading: "Rin Tin Tin: The Life and the Legend" by Susan Orlean.

[Brett Oppegaard, Ph.D., brett.oppegaard@hawaii.edu](mailto:brett.oppegaard@hawaii.edu)

Access Hound's Guidedogs (Writing Coaches + Personas)

<https://guidedogs.accesshound.com/describe-now> (released February 2025)

---

## Foundations

- Can be used for both static and dynamic AD
  - Flexible and adaptable to be used throughout the drafting process
  - Can be used by Writer only, Writer with Editor, Editor only
  - Powerful enough to recognize charts, maps, people, objects, collages, etc.
  - Many more types of Writing Coaches and Personas are under development
  - If I didn't build this, someone else would have done it
- 

## Findings

- In 15-30 seconds, a publishable draft is ready to go ("better than nothing")
  - Even only 50 percent ready for public release, cuts production costs in half
  - Combining the Writing Coaches with the Personas (all in one button press) has the potential to write, coach, and focus group in ~30 seconds
  - People say they don't want the Writer (Rinty), but they *want* the Writer
  - High quantity of image-to-text queries show desire for efficiency
- 

## Discussion

- Human performers are being replaced by synthetic voices in AD; are writers next?
  - Are writers necessary? Or do we only need rewriters and editors in AD now?
  - Is this the **John Henry** moment, where tech fully takes over human-led AD?
-

THANK YOU

Brett Oppegaard, Ph.D., [brett.oppegaard@hawaii.edu](mailto:brett.oppegaard@hawaii.edu)