

Annual Interagency Accessibility Forum

Beyond Compliance: Building a Culture of Digital Accessibility

November 7-9, 2023



FDIC



AI-generated alt-text

The future of image accessibility?

November 9 , 2023

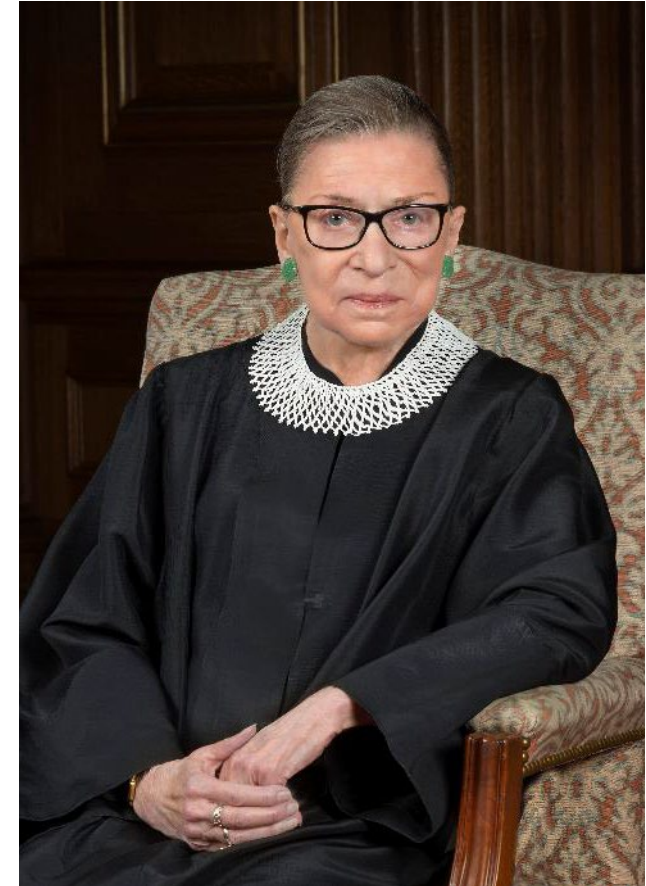
Laura Larrimore & Teresa Verigan

U.S. Department of State

U.S. Patent and Trademark Office

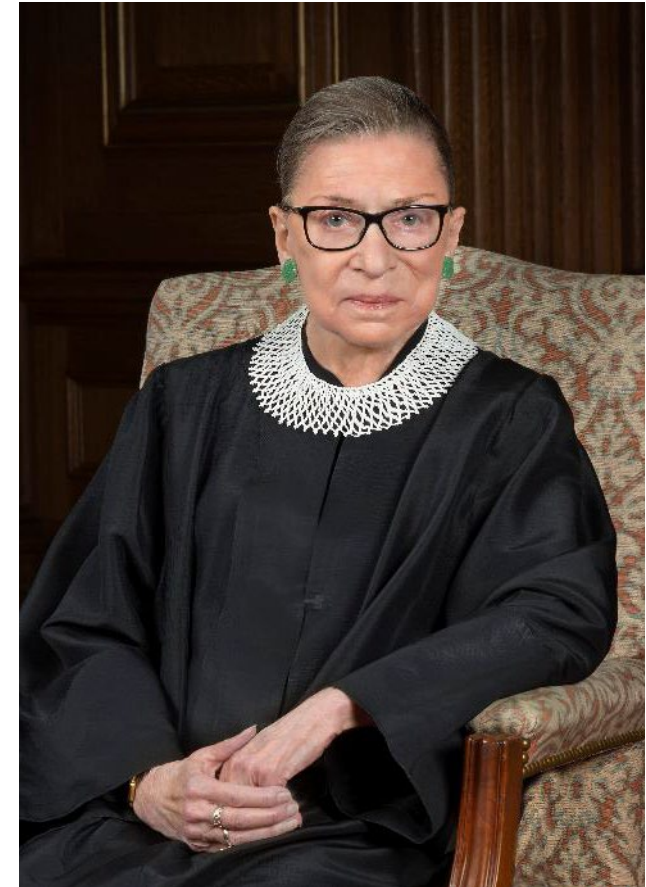
What are alternative text descriptions?

- Describes the image when the visual cannot be viewed
- Required for accessibility
- Longer descriptions are not always better
- Context helps determine what to include in a concise description



What are alternative text descriptions?

- 1 image, 3 descriptions:
 - A woman in a chair
 - Portrait of U.S. Supreme Court Justice Ruth Bader Ginsburg seated in her judicial robes.
 - Official portrait of Supreme Court Justice Ruth Bader Ginsburg. She is seated in a chair wearing black judicial robes and her iconic white beaded collar necklace. Her gaze is calm and direct as she leans confidently on the arm rest with her hands clasped loosely in her lap.



What is Artificial Intelligence (AI)?

- The theory and development of **computer systems able to perform tasks that normally require human intelligence**, such as visual perception, speech recognition, decision-making, and translation between languages. [Oxford English Dictionary]
 - Examples include self-driving cars, auto-captioning, medical imaging
 - In some instances, AI is getting more accurate than humans on certain tasks (like medical imaging)

A subset of AI: Computer vision

- Computer vision is a **field of artificial intelligence (AI) that enables computers and systems to derive meaningful** information from digital images [Microsoft]
- Computer vision is a **technology that machines use to automatically recognize images and describe them accurately and efficiently.** [Amazon Web Services]

What we hoped to learn

AI has gotten much more powerful and accurate in recent years
– **how does AI compare to a human when writing alt-text?**

- *Can AI computer vision correctly identify images?*
- *Are certain images more challenging to identify?*
- *Among AI systems and models freely available today, do some do a better job than others?*
- *How far apart in quality is human alt-text compared to AI?*
- *How far apart in quality is human alt-text from experts versus novices?*

How we set up our experiment



We selected types of photos commonly found on government websites, from simple portraits or editorial photos to more complex photos, illustrations, and graph.

Selecting our human writers

- **Three Experts**
 - Writers who regularly create alt-text as part of their job
 - 7 years experience writing alt-text, on average
- **Three Novices**
 - 2 high school interns and 1 recently hired graphic designer
 - None had received training in writing alternative text prior to project

Both groups were given a short tutorial on the best practices for writing alternative text. They were allowed to refer to the tips while they were completing the assignment.

Instructions given to human writers

- Alternative text describes what is going on
- Be specific, and try to add context by describing actions, objects, or tone expressed through the visuals.
- You must include any words that are important to understanding the image.
- Charts or graphs do not have to include all data points, but should try to express the larger insights provided by the image.
- You should try to be as brief as possible. Screen readers cut off longer alternative text around 125-characters.
- It should not be the file name or duplicate what's listed in the caption. Simple portrait images of a person might be labeled with their name.
- Avoid saying 'picture of' or 'image of'. Feel free to specify if the image is an illustration, collage, or non-standard photographic image.
- Avoid spelling mistakes or 'easter egg' jokes. Alternative text is an important accessibility tenet and should be treated with the same care as other digital content.

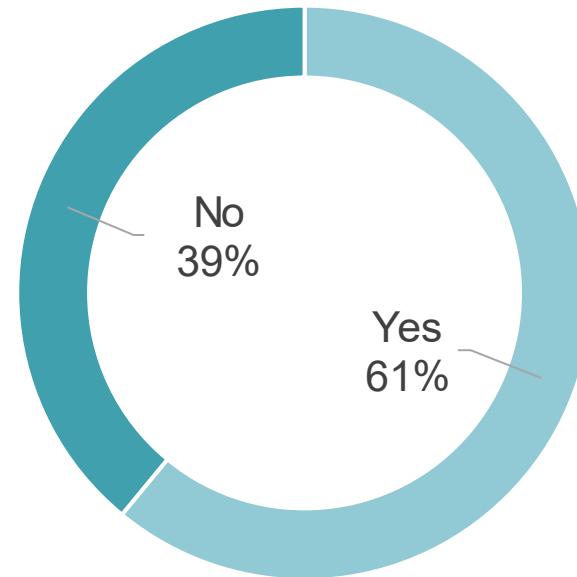
Selecting the AI systems

- **Microsoft Word 365**
 - Alternative text suggestions are now offered within programs such as Word or PowerPoint
 - Designed to help authors creating accessible files from those programs
- **GenAlt**
 - GenAlt is a Google Chrome browser 3rd party extension
 - It generates alt-text for online images missing that information
 - Designed to help those who require alternative text
- **Alt.Ai**
 - Paid add-on for authors in Wordpress or other content management systems (CMSes)
 - Designed to help businesses automate creating image descriptions

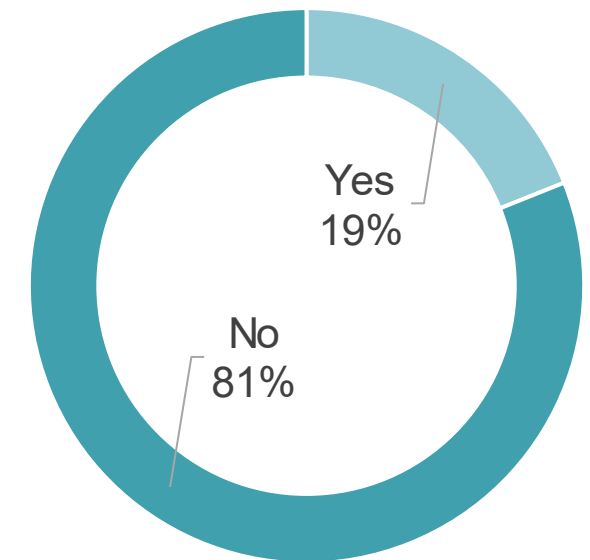
The alt-text raters

- The survey was sent to government listservs, but did not mention AI-generated text.
- **702** responses
- Most respondents were alt-text writers, rather than alt-text users

Do you regularly write alt-text as part of your job?



Do you regularly use alt-text to understand images on websites?



The rating method

Rating the accuracy of alternative text in images

Image 1



The image above does not include a description because we want you to rate 10 options below using a 1 to 5 scale. It is ok to assign the same rating to multiple descriptions if they are roughly equivalent in quality.

* 4. Please rate the alt-text description of the above image

	1 - inaccurate	2	3 - acceptable	4	5 - highly accurate
Portrait of a child playing in a park, using natural lighting and candid expressions.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Colin Powell	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Official photograph of Colin Powell smiling and wearing glasses, a black suit and red tie, with an	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

- Each image was loaded into a matrixed survey with 10 possible alt-text descriptions:
 - 3 expert descriptions
 - 3 novice descriptions
 - 3 AI descriptions
 - 1 wildly inaccurate description
- All descriptions were provided as originally written, without editing to correct spelling or grammar

Scoring methodology

1 to 5 scale

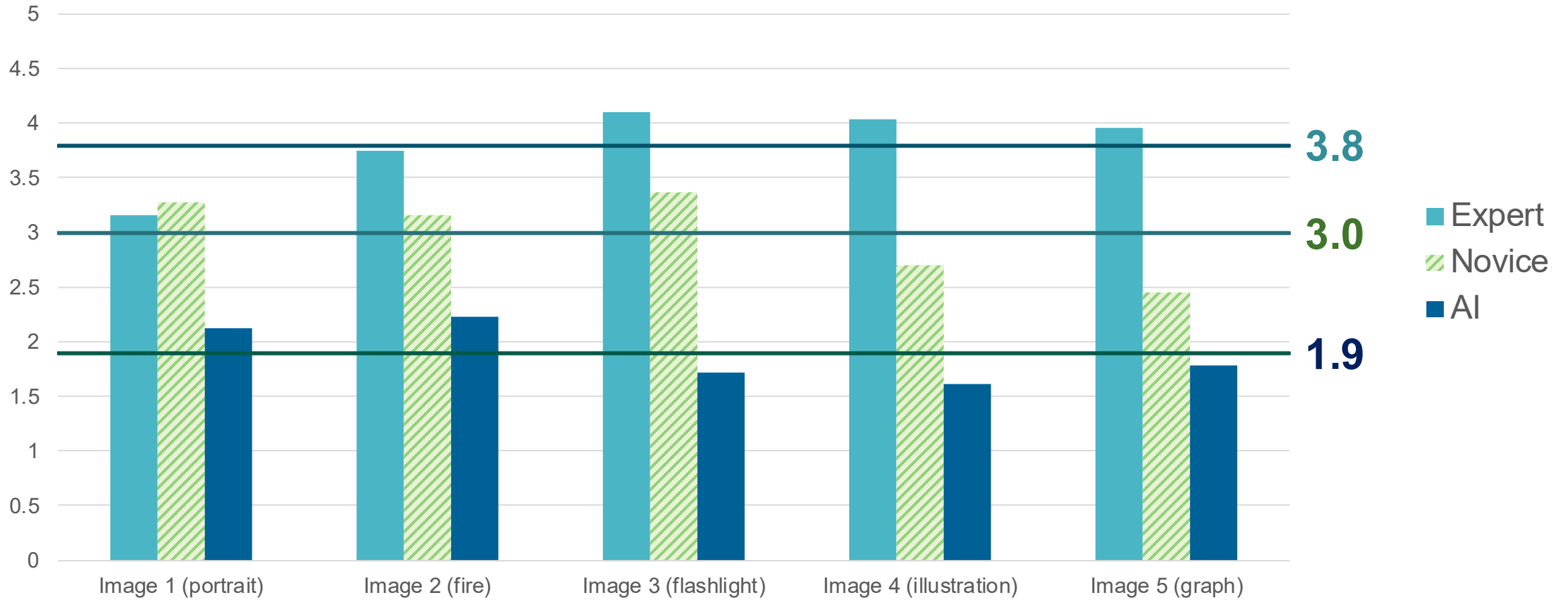
- Rate the alt-text as a **“1”** if it is **wrong** and does not describe what is shown.
- Rate the alt-text as a **"3"** if it is **acceptable**, but missing important details.
- Rate the alt-text as a **“5”** is **highly accurate** and would work best to help someone understand what is being visually shown on screen.

Some limitations

- Limitations:
 - Raters only rated 5 images
 - We included a ‘worst’ answer, but not a ‘best’ answer in the set – e.g there was no pre-determined “correct” answer
- No situational context given
 - Raters were told they would not have context for how the image appears in relationship to surrounding text or captions. They were told to “do your best to rate the image as a stand-alone element.”

Who's best at writing quality alt-text?

Alt-text quality, by source

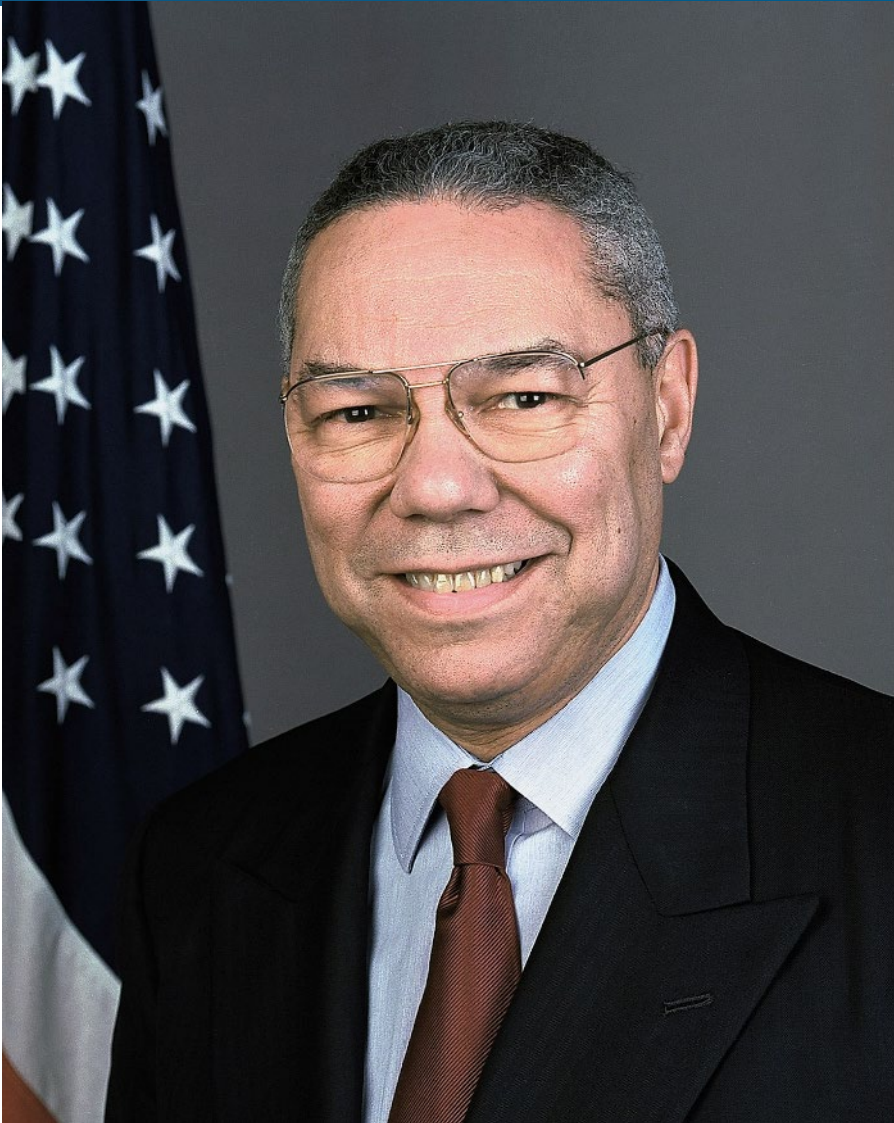


Who's best at writing quality alt-text?

- Humans were notably better than AI
- Experts produced acceptable descriptions for all five images
- Novices and AI struggled more with complex images (illustration and chart)

Source of alt-text	Skill-level	Quality rating (Average)
Human	Experts	3.8
Human	Novice	3.0
AI	AI	1.9
Intentionally inaccurate	N/A	1

Key point: Novice writers can produce quality descriptions



- This description was written by a novice, and rated highest in the experiment (4.82)
- **“Official photograph of Colin Powell smiling and wearing glasses, a black suit and red tie, with an American flag behind him.”**

Key point: Novice writers can produce quality descriptions



Source	Alt-text provided	Quality score (Average)
Expert	Official U.S. government portrait of Colin Powell	3.76
Expert	Colin Powell	2.83
Expert	A smiling man wearing glasses, a suit, and tie, with an American flag in the background.	2.86
Novice	Official photograph of Colin Powell smiling and wearing glasses, a black suit and red tie, with an American flag behind him.	4.82
Novice	Colin Powell	2.82
Novice	A man smiling in front of the United States flags	2.18
AI (Alt AI)	A man in a suit and tie standing in front of an american flag.	2.38
AI (MS Word)	A person wearing glasses and a suit.	2.13
AI (GenAlt)	An image of 1 person.	1.85

Key point: All creators make educated guesses



“The firefighting picture - all of the text is incorrect. The hose is not spraying water, it is spraying fire-fighting foam...”

“It was difficult to rate the image of the firefighters, because the image did not provide sufficient context to determine what was on fire.”

Key point: all creators make educated guesses



Source	Alt-text provided	Quality score (Average)
Expert	Four firefighters in full gear spray water on a small fire in an industrial area	3.79
Expert	Four firefighters in full gear attempt to extinguish a fire with water hose.	3.75
Expert	Four firefighters standing on flooded ground with one of them putting out fire on the ground with water from the hose.	3.71
Novice	Four firefighters wearing full firefighting gear and oxygen tanks spraying water on an aircraft on fire	3.44
Novice	Four firefighters extinguishing a fire with a water hose.	3.18
Novice	Fireman urgently putting out a fire outside under an oil tank	2.83
AI (alt.ai)	A group of firefighters spraying water from a hose.	2.65
AI (MS Word)	Firefighters spraying water on a fire.	2.75
AI (GenAlt)	An image of 3 people, fire and text.	1.27

Key point: Context tells us what is important to include



“These are especially tricky to judge without context. The image of the person with the prosthetic leg and flashlight has many elements that could be described, completely depending on context.”

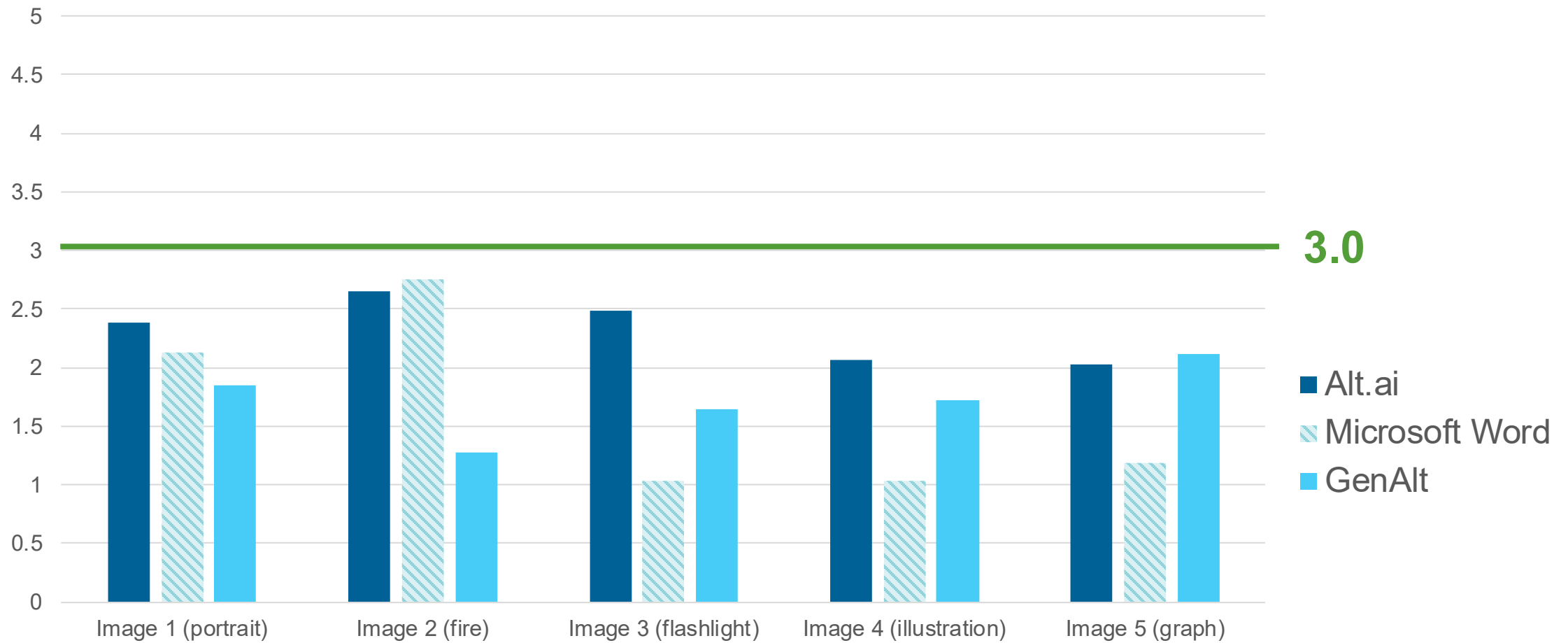
Key point: context tells us what is important to include



Source	Alt-text provided	Quality score (Average)
Expert	A woman with a prosthetic leg stands in a dark kitchen and shines a flashlight into an open cabinet drawer	4.43
Expert	In dim lighting, a woman with prosthetic leg checks the contents of a kitchen drawer using a flashlight.	4.31
Expert	A person in a dark kitchen pointing a lit flashlight into the pulled out top drawer of a kitchen cabinet.	3.54
Novice	A middle-aged African American woman with a prosthetic (right) leg in a dark kitchen looking into a drawer with a flashlight.	4.64
Novice	A woman with glasses searching for an item in the dark with a flashlight.	2.86
Novice	A woman in a dark kitchen looking in a draw with a flashlight	2.6
AI (alt.ai)	A woman opening a drawer in a kitchen at night.	2.48
AI (MS Word)	A person wearing a mask and holding a clipboard.	1.03
AI (GenAlt)	An image of 1 person.	1.65

What are the quality differences between AIs?

Alt-text quality, by AI service



How did AIs do? Top level results

- A few things stand out:
 - Alt-text submitted by AIs was not very well rated, **averaging 1.9**.
 - Alt.ai (whose main business focus is alt-text) rated best, but never scored as “acceptable” (3 or higher)
 - There was significant variability in quality even from the same company depending on the image.

Source of alt-text	Quality rating (Average)
Alt.ai	2.3
Microsoft Word	1.6
GenAlt	1.7

Image 4: Illustration / logo for L.E.A.P program



Image 4: Illustration / logo for L.E.A.P program



Source	Alt-text provided	Quality score (Average)
Highest Rated Expert	An illustration of a woman with a book in her hand, leaping over three lightbulbs in a gap between two cliffs. Text reads: LEAP: Legal Experience and Advancement Program	4.66
Highest Rated Novice	Graphic of a lady leaping between two cliffs and three lightbulbs. Text: "LEAP" & "Legal Experience and Advancement Program".	3.74
Alt.ai	"Leap legal experience and advancement program."	2.01
MS Word	Chart.	1.03
GenAlt	"A graphic of text that says '(L LEAP (eap legal experience and advancement program'."	1.7

Image 5: U.S. Patent line graph

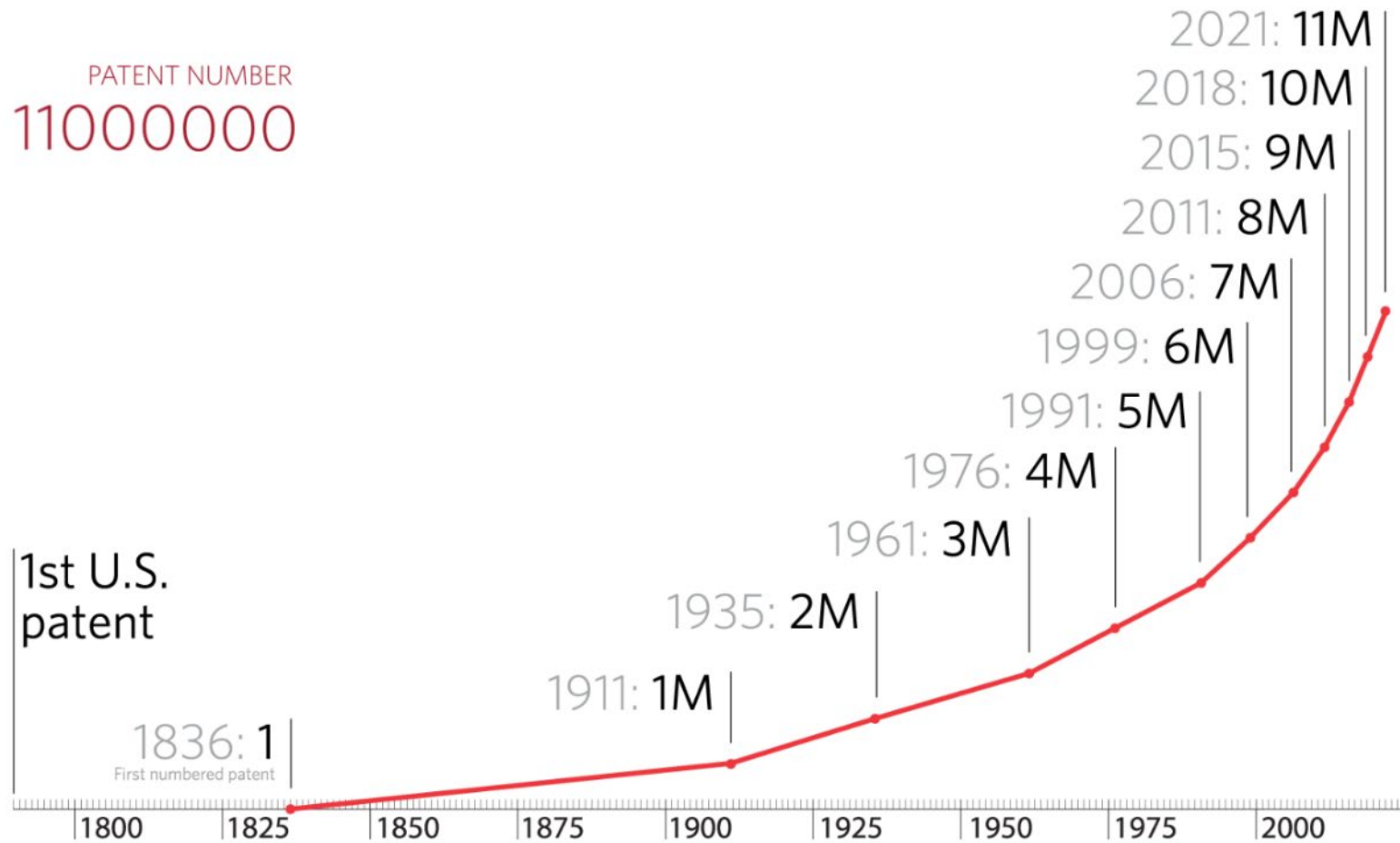
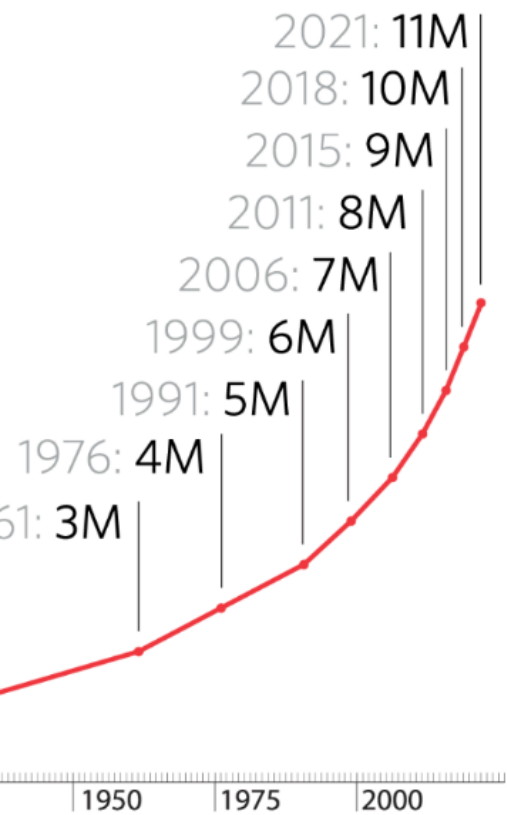


Image 5: Line graph



Source	Alt-text provided	Quality score (Average)
Highest Rated Expert	Line chart depicting U.S. patents issued per year, starting with the first numbered patent in 1836 and ending with the 11 millionth patent in 2021. The majority of patents have issued since 1976.	4.2
Highest Rated Novice	Table showing the increasing number of patents per year starting in 1836 (first U.S. patent) until 2021 (11,000,000 patents).	3.21
Alt.ai	A graph showing the number of u s patents.	2.11
MS Word	A picture containing text.	1.19
GenAlt	An image of text that says 'patent number 11000000 2021:11M 2018:10M 2015: 9M 2011:8M 2006:7M 1996:8M 1999:6M 1991:5M 5M 1976:6M 1961:3M 1911:1M 1st U.S. patent 1836. fsnumbered patent 1800 825 1800 1875 1900 1925 1950 1975 2000'.	2.11

What we learned: Context (or lack of it) impacts the quality

- If you want to write ideal alt-text, it's best to have the full context so you can include only details for what's relevant and important.
- Novice writers can be trained quickly on writing alternative text, but the more complex the image, the more experience will improve the output.
- Both humans and robots sometimes make mistakes with image identification, having an editor or second look confirm context and grammar is ideal.

What we learned: Current state of AI computer vision

- AI's may fail to identify or misidentify images in ways that are puzzling.
- AI guardrails prevent mistakes, but may limit AI's usefulness
- **Artificial intelligence is getting better at computer vision – but the tools we tested are no match for a human intern.**

Future areas of research

- Additional studies that give raters more context
- Testing other AI tools which might provide better results.
- Do you use any AIs on a regular basis that we should look into?

References

- Consumer Safety Product Commission diverse stock imagery
 - www.cpsc.gov/Newsroom/Multimedia/Americans-with-Disabilities-Home-Safety-Best-Practices
- GSA bias in facial recognition study
 - QR code to volunteer (4,000 people needed!)
- Sample survey –see how our survey was structured: www.surveymonkey.com/r/6CH9HXV



Thank you!

- **Laura Larrimore**
Consular Affairs
U.S. Department of State
LarrimoreLL@state.gov

- **Teresa Verigan**
Digital and Creative Services Branch
U.S. Patent and Trademark Office
Teresa.Verigan@uspto.gov