# Transcript for the ILOTA Things - Episode 2 – A Picture is Worth A Thousand Words, or is it? Podcast

**Announcer:** Welcome to ILOTA Things, the AD CET podcast where we explore Inclusive Learning Opportunities through AI. In this series, we'll explore the exciting convergence of universal design for learning, UDL, artificial intelligence, AI, and accessibility, and examine ways in which we can utilise emerging technologies to enhance learning opportunities for educational designers, educators, and students. Now, here are your hosts, Darren, Elizabeth, and Joe.

**Darren:** Hello and welcome from whenever, wherever, and however you are joining us and thank you for your time as we investigate ILOTA Things. That is Inclusive Learning Opportunities through AI. My name is Darren Britten and joining me once again on the artificial intelligence, universal design and accessibility merry-go-round are my co-hosts, Joe Hou ghton,

**Joe:** Hi from Dublin in Ireland

**Darren:** and Elizabeth Hitches,

**Elizabeth:** Hi there for Australia.

**Darren:** In this episode, titled ‘a picture is worth a thousand words, or is it?’ we are going to dive into the world of image descriptions and alternate text, commonly known as alt text, and look at how AI can assist in creating descriptions and support educators in providing inclusive learning environments through universal design for learning.

First up, I feel maybe you need to provide some context in terms of what we mean when we talk about alt text and image descriptions. This is fundamentally the text that accompanies visual imagery that's intended to convey the same information or function as the image or the non-text content, so that users who cannot perceive the visual element can still understand its meaning or purpose.

Now alt text may be hidden in the background, and only available to those with the right tools to read it, or it could be alongside the image so that everybody has access to the text equivalent. The idea of alt text has been around for almost as long as we've been putting images online, and the need for describing them has been a fundamental criteria of the web content accessibility guidelines, otherwise known as WCAG, which is the international standard for web accessibility, since back in 1999.

Creating alt text is one of the easiest and at the same time most confusing issues for educational staff to deal with when they try and make their content more accessible. If the adage that a picture is worth a thousand words is true, then is it not just a simple process then to take a picture and go back to a thousand words? And does every image need an actual text equivalent? This is what we'll explore in this episode.

But before we dive in, I think we might need your input Elizabeth on how and where does alt text intersect with the notion of universal design for learning?

**Elizabeth:** That's a really good question. And you know, when we think about where it actually intersects with UDL, we're really thinking about that area of, you know, how can we proactively design for multiple means of perceiving information? So we're making sure that where we have information presented in a visual way, we also have that non-visual way of accessing this. So if we have an image, we can provide a text-based description that alt text and that can actually be read aloud by screen reader or it can be displayed in a tactile way on a braille display. So if we think about multiple means of perception, we don't just have that visual option, we also have that audio and tactile option as well because we have alt text attached to that image. So we've got multiple means of accessing that information.

Now you might think this may be for specific types of students, but it's actually something that benefits everybody and much like the rest of that UDL framework. So you know, you might have experienced this when you have a web page that may not be loading or let's imagine you have challenges with your download speed or limited data downloads. When that page isn't loading, it's actually alt text that gets displayed in place of that image. So you may not have that image displayed on the screen, but you're going to have that text description in its place even if it's not displaying visually.

Now what I'd also like to mention is that if we're taking that UDL approach, we're thinking in that UDL way, we're going to be considering this proactively. So in education or any field, really, we already know that we're going to have a diverse range of audience members. So in education, that's students, parents and carers, it's our colleagues or even the broader community that engages with that education sector. And we want to be sure that we're not simply making these resources or communications accessible in a reactive way. We don't want someone to have to say to us, I'm sorry, but this resource is not accessible or I can't access that image. We want to be doing this proactively. We want to make sure that we are designing or multiple means of access from the very start. So we're removing barriers to access before they're even encountered.

 So that really takes us to how do we actually do this? You know, how do we generate alt text? And I think, you know, Darren, you might be best to jump in here because you've got a lot of experience with accessibility. So, you know, how do we usually generate alt text?

**Darren:** Well, traditionally creating alt text has been done by a human manually describing the image, meaningfully, and I suppose in as few words as possible. Alt text was not designed to contain multiple paragraphs of information. If you've ever, I suppose, put an image into a Microsoft Word document, then you would have seen the alt text panel that opens where you can type in whatever text that you want to be associated with that image, or there's also the option to mark an image as decorative. This process of adding alt text is similar in many programs, whether you're adding to a web page or even into an online learning management system. And for those that are unfamiliar with writing alternative text or wondering where to start or need a bit of help along the way, we will be providing links to some websites that can help you in stepping through that process because it isn't necessarily intuitive to a lot of people. And these resources are really useful in taking you through the process of understanding when to describe information, how to describe that information, and more importantly, what not to describe, just because a picture can be worth a thousand words doesn't mean it needs to be.

 I mentioned earlier the option of marking images as decorative. This is useful for images that are just there for aesthetics or to create a particular look and feel and help break up blocks of text. These images are referred to as decorative images, and these are not read out by screen readers when marked as such. However, if they're not marked as decorative and they have no alt text with them, they can behave in different ways with different technologies, such as just reading out the file name, and trust me, a lot of file names are not very descriptive, such as A154389976432\_1012.jpg. So adding alt text to an image is just as important as marking those that do not need alt text as decorative images. Now, this brings me to AI and why we're here. We now have a whole array of tools that can infer information from an image and describe them in a way that hasn't been possible before. And it's improving with each iteration of AI. So, Joe, I'll let you jump in here and help us understand where we are with AI describing images and what do we need to know?

**Joe:** Well, let's rewind to where you were talking about this long file name that means nothing. One thing that you can do and get into the habit of doing is, you know, when you're going to use an image, save it into a folder on your computer and give it a meaningful name. Call the image something that actually refers to what's in the image, because as we move on to the next step that I'm going to take you through of kind of then then putting that image up into some AI program. The program will see the file name and will use that as part of its kind of like processing. So that's an immediate input that you can use. And also, it has the benefit that even if you forgot to do the alt text, for instance, you've now got a meaningful file name. So if the web page doesn't load properly, it reads out this meaningful file name dot jpeg rather than 10157, whatever it is.

So in terms of AI, I mean, there's lots of different ways of doing this and the show note links will will have links to a number of different AI programs that will allow you and assist you in generating alt text. But I mean, let's just go back to chatGPT or Copilot or Claude, any of the normal bots that you might already be using . I think most of them now have a button where there's there's normally a little paperclip where you where you can upload a file. So if you take your image file, your jpeg image or your tiff image or your png and it doesn't really matter which one of those it is normally, some of the AI tools will only accept JPEGs. So you might need to convert a tiff file or a png file into a jpeg file to get it into these AI tools.

So all these AI's are now multi-modal. So they can read text, they can analyze an image, you know, in first stuff about it. So if you take your image, you upload that into chat GPT say, you can now say the image that I've just uploaded generate me some alt text and that's the most basic ask to do this. Generate me some alt text and it will do that. It'll give you a paragraph and it will describe what's what it sees in the image. So you might think, right, well that's the end of the episode. Yeah, we know how to do it now. Yeah, it's not actually the end of the episode. There's a little bit more that we need to look at. So I'm going to leave it there for now and I'm going to throw this back over to Darren. What else? What else do we need to think about Darren?

**Darren:** I was thinking of going on holidays when you mentioned how easy it was Joe, but of course, you're right again, It's not that simple. Context is everything at the prompt and what you ask AI for is extremely important. Who is the audience for the image and what do you want to describe? Maybe very different if your audience has a class of 10- year-old students versus some postgraduate students. What assumptions are we making about any prior knowledge that you may need to understand the image that's there? Using a prompt that asks ‘describe this image for a neurodiverse audience’ will give you a hugely different response to ‘provide alt text for this image’. Without context, AI can give you the bare minimum or completely inaccurate information.

So being thoughtful about what you want described and the context in relation to the surrounding text will make all the difference. And Elizabeth, you've got a great example of this. I believe from Harvard University.

**Elizabeth:** That's right. And the best way to really see how alt text can work and how it operates is to see examples and get a real feel for how it can be done. And one of the really excellent examples is from Harvard University. And the reason that this is a really excellent example is that they start with a single image, but then they demonstrate how that alt text for that single image can actually differ based on the context or purpose of the image. So they give an example, it's an image of a stadium, and they show that one description may be that it's describing it as an empty stadium. But if the context or purpose of that image is actually to show a particular aspect of that stadium, that's why the image is there, then the alt text needs to focus in on describing that particular aspect.

So one image can be described in various ways. And this is where we need to really consider as whoever's creating the content or whoever's putting resources together, whether that's educators or administrative staff, anyone who is involved in putting information together, we need to think about what is it we're intending our audience, our students, whoever it is. What are we intending them to see or understand from that image and making sure that alt text provides that equivalent in text form?

**Darren:** Absolutely. And I'm sure listeners will hear the same message in every episode that we do, and that's the importance of a human, you know, and that context in the AI educational environment. There may well be a thousand words, but what if only 20 of them are relevant or even necessary? So just to put another example forward, you know, imagine a series of images, the first showing a couple holding a young baby, the next showing an awkward teenager posing for a photo, the next, an image of an excited young adult holding up keys next to a car, the next a wedding photo of a couple and their guests on the steps of a church, and lastly a photo of a couple holding a newborn baby. Now, without me providing any more context, AI may make the same assumptions that you may have made. Is the baby in the first photo the same teenager in the next? Are they the same men or women that's getting married? Are they the same parents of the child in the final photo? Maybe? Maybe it is a series of images from an adopted baby who goes on to get married and adopts a child of their own. Did you assume that the baby is the same person throughout all the photos? I didn't tell you if they were the same person or if they were even the same nationality. Maybe it was a series of images just showing the circle of life or maybe a series of significant milestones in a single person's life.

Now, I've seen images of people described in ways that can only be listed as, you know, offensive. I've seen people misgendered, people with disabilities described with complete ableist language, and in one case a short statured person being automatically assumed to be a child due to their height. There's also an assumption that automating image descriptions using AI will just be good enough, and that for students with a vision impairment, they'll be able to use various technologies that will describe these images for them.

Now, don't get me wrong, AI in this space will be an invaluable addition to assistive technology sphere and a game-changer in many ways. However, it can only go so far with understanding the context of an image. Another important consideration is that each response, even with the same prompt, will be different. And, you know, as educators, do we really want every student having a different description?

So the assumption that this will do it for us is still far from reality, you know, as an image is really in the eye of the beholder, as they say. We can, however, work with the AI, and it's important to carefully think about what we're asking the AI to do and what we're willing to have the AI do for us. So, back to you, Joe, touching on some of these things, how important is that instruction or the prompt that we give it, given the context we've been talking about, and how can we avoid some of those biases?

**Joe:** Maybe this whole podcast series is about communication, isn 't it? And words are important. And the words we use and the words we don't use are important. And I think, you know, perhaps even more so as educators, we're probably sensitized to that already, to some extent. You know, how we get our messages across two students, how we hear what our students are telling us. And, you know, going back to what Elizabeth was saying earlier on about multiple means of representation, multiple means of perception, just because I say something to you, doesn't mean that you will process those words in the same way as Elizabeth does, or as Darren does, or as Susan does, you know, or whoever else it is. So, we have to really be, you know, mindful of that and really think about how we both prompt the AI, if we're going to use AI, to generate words.

But then we've also, you know, you've got to then bring the domain knowledge and bring the context awareness, and bring the awareness of your students, and the message that you are trying to convey, to read what the AI comes up with. Because I mean, we've all seen, well, I'm saying we've all seen it now. I know that some people still haven't used chatGPT or AI, so we can't assume that everybody's seen it. But it's, it's, I mean, every time I use it, you know, and it just goes, and there's this huge page full of information. It's like Harry Potter, it's magic, and I still get a buzz out of seeing it work. And there's that assumption almost, oh, that must be fine.

But I was at a conference yesterday, and a senior academic was talking at the conference and she said she'd, she'd asked chatGPT for, for some stuff and it had hallucinated all the references. And she was just shocked that, that this was it. So you can't actually just assume that even though these AI's are amazing, that what they give you is going to be right. So you still got to read the alt text.

All of us have got inherent biases. If you went now and added an image of a person to the article that you're writing or the lesson plan that you're creating today, you would probably default to bringing up a picture of somebody that looks something like you. So I'm a white middle aged male, and I probably, even unconsciously, default to putting in white people in my image. Now I'm fortunate enough to, you know, be teaching highly multicultural, globally diverse, pluses from India, China, you know, all over the world and stuff. So I suppose maybe my sensitivity to that stuff has been raised a little bit and certainly talking to Elizabeth and Darren and other people in the UDL, you know, space as sensitized me to this. But we have to be aware of our own biases.

And then we also have to be aware that AI's have bias because the AI's have been trained on basically the data set of what's out there. Well, if we're talking about, you know, the data set that's out there, it's probably quite Western skewed at the moment. I think most of the big AI's have still been trained mostly on Western data. Now I know China's doing its own thing with AI's at the moment and India's doing its own thing with AI's and stuff. So, you know, perhaps that's almost going to be a regional thing, which AI do we use? And if you use the Chinese AI and you use, you know, chat GPT or Copilot, are you going to get very different responses? If you put the same image up, that would be an interesting thing to check. There's a paper in there somewhere, isn't that, for somebody?

So we've got to think about bias, both our own biases, but then also being conscious that there may be a bias coming through from the chat GPT's as well. So, yeah, I mean, there's so much stuff there to think about, isn't there? So, I don't know. I've said enough. Elizabeth, what are your thoughts?

**Elizabeth:** Well, I think, you know, it's really interesting. We're touching on the really simplistic images. We're talking about just a simple image that, you know, it 's not a figure or a diagram or a graph, you know, that something will come to another episode. This is just thinking purely about describing a simple image. And there's a lot of complexity there. We can create a really accessible digital resource, you know, our images can have that alt text attached, but what happens then if we print it? You know, all of that alt text will disappear on that paper -based print, we won't see any of that data that's attached to it, but what we may not always realise is that this can also happen if we do that function of printing to a PDF. So, this is not that saving to a PDF, printing to a PDF is like, essentially like taking a photo of your resource and having a digital photo. So, whether we print in paper or we print a PDF, all of that effort that you've put into creating that accessible document, that extra information in the back end usually gets removed .

So, we need to think about then, if you started with an accessible document, you've printed it, that hidden information's disappeared, how then are your students going to have access to that alt text? So, if you're using a screen read or a text to speech technology, how's it going to happen? So, if you're in a classroom that's really working on that, you know, that printed resource, you know, things get printed out for each student. We can think about maybe taking a UDL approach and making options available. So, you can have your printed version, you know, many students work well with printed versions of resources, but we can also provide access through having an accessible digital copy available that has your alt text to those images. So, you can have options, make sure there's access through one of those modes and ideally have options for how students can engage with that.

And if you aren't doing a printed copy, you can even add something like a QR code so that, you know, students can even scan that QR code on that paper-based resource and have, you know, instant access to that digital material, that really accessible digital material. And, you know, I've learned from Darren, there are lots of free ways that we can create those. You've got the CANVA QR code app. You've also got Bitly, that's bit.ly. and, you know, images don't just appear on our flat resources on and our paper documents. We also might be presenting a PowerPoint slide in a classroom and have that on the projector. And so, we think, well, you know, we're going to have it visually projecting onto that screen, but how have we got that access that alt text would usually provide? And what we can do again, UDL, we've got options, we can make those slides available to our students that digitally accessible copy. And while you're presenting those slides in class, you can naturally build in some of that image description.

Now, when we think about how you'll actually make that alt text available to all users, we can also think about the length of that alt text. So, you know, maybe you only need a really short single sentence that describes that image itself. And then to have a more in depth, more details and relevant description, we can add that below the image so everybody can understand the purpose and context of that image. So, we can have a really short description for that alt text and then have a larger description that all users of our resources can actually access and understand.

You know, if it doesn't need a description, let 's imagine that you have a resource and you've created some really, you know, beautiful ways to chunk the text up and you've put an orange shape in to add visual interest somewhere or you've, you know, you've added a big arrow from one section to another, but it's not actually really adding meaning, it's just to add that visual interest and that visual decoration. Then we don't need to tell people, you know, somewhere on this page is a, you know, an orange rectangle or somewhere on this page is, you know, an arrow. If it's not actually adding meaning, if it's just for that visual interest but not communicating information, then we can mark it as decorative.

Now, getting us back to thinking about, you know, what do we do then using AI in this process? I'd really love to bring up a really funny learning experience that happened and, you know, this example for me just really reminds us that as much as AI can do in this space, we still need that human lens on whatever it is that's generated. So, Joe, I would absolutely love if you could please recreate the image in our minds of what happened with the bracelet.

**Joe:** I will, yes. But before I do, just let me jump back to what you were just saying about kind of, you know, the visual or the verbal description of alt text, you know, potentially when you're presenting a slide deck. And, you know, as a learner, I always look to find somebody who's really good at doing something to kind of learn how to do it well. And one of the best people that I've ever seen at doing this kind of description is Tom Tobin. So, Thomas J. Tobin, you know, who kind of wrote the book on, you know, plus one with Kirsten Behling and he's a great friend of ADCET. And, you know, so dig out one of his videos, you know, just go on YouTube, there's plenty of stuff from Tom Tobin. And just look at how he presents a slide deck. And he just, he's so good. Because, you know, he talks about the slide by describing it first and then talked about, you know, the meaning and all the rest of it. So, there's a good exponent of what Elizabeth was talking about.

My daughter April, who is 12, uses Canva to produce all the materials for her little business. And she makes jewelry. So, she orders all the stones and things and the wire and everything. And she's got all the tools. And she makes little bracelets and necklaces and stuff like that. And Elizabeth and I, month two ago, were, you know, looking at some of the stuff that she did on a Zoom call because we never met each other. We're Zoom friends and Elizabeth and I. And she'd asked me to make help her make a little photo book of all her jewelry. So, I set up a background just a white background and I put one of her multicolored bracelets. So, you imagine, you know, bracelet will go around your wrist, multicolored stones. And we shot this just on a white background. So, it's a very simple image that you can't get wrong, really. You know, I mean, how difficult is it to describe a multicolored stone bracelet, you know, on a white background? I mean, that would be the old text from it.

So, I uploaded the bracelet into chatGPT and said, generate me the alt text, please, for this bracelet. And the alt text that came back said, yes, this is a beautiful sunset against a pale sky. And Elizabeth, I looked at this and I'm, what? So, this was a great example of an AI hallucination, you know, it completely misread this image. Now, would it do that today? Maybe, maybe not because the algorithms are improving every day, every week, another estimate, but it just shows that you can't actually believe it, you have to process it and check that what you're being told is actually what's there. So, I mean, you've got to go off and you've got to try these tools. You know, upload images into chatGPT, upload images into any of the tools. So in, in the show notes, ahrefs.com is one tool that can do this. Another one is Protoolio.com. And that does an alt text generator. So again, these, these links will be in the show notes. But also all the main chatGPT, Copilot, Claude, they will, will let you do this stuff. But, you know, as with everything we talk about in this series, don't just listen. Okay, go play. Go, go do. Because hopefully we're putting this time in to help you improve your practice one thing at a time, you know, plus one. So just go try this and see whether it makes things better for both you and your learners.

**Darren:** Thank you, Joe. And just to re-emphasize, go and have a play with these tools. And you can find all the relevant information from today's show and previous episodes there on the website at [www.adcet.edu.au/ilotathings](http://www.adcet.edu.au/ilotathings). And of course, we'd love to hear from you because we want this series to be an ongoing discussion. If you have a question or a comment about artificial intelligence, UDL, accessibility, anything we've discussed or would like to share how you're using AI in your practice, we'd love to hear from you. And you can contact us via email at feedback@ilotathings.com.

**Elizabeth:** Now we really hope that we 've given an insight into how AI can help you to take that UDL approach to provide multiple means of representation and different ways that students can perceive information and how. And you know, with just a little bit of context and thought, you can ensure that all your visual imagery is available to all your students or to anybody else who's going to be interacting with that resource or that digital image. So thank you so much for listening and we really hope you join us next episode so were going to continue to explore a lot of things. So till then, take care and keep on learning.

**Joe:** See you next time.

**Darren:** Bye, thanks for joining us.

**Announcer:** Thank you for listening to this podcast brought to you by the Australian Disability Clearinghouse on Education and Training. For further information on universal design for learning and supporting students through inclusive practices, please visit the ADCET website. ADCET is committed to the self-determination of First Nations people and acknowledge the Palawa and Pakana peoples of Lutruwita upon whose lands ADCET is hosted. We also acknowledge the traditional custodians of all the lands across Australia and globally from wherever you may be listening to this podcast and pay our deep respect to Elders past, present and emerging and recognize that education and the sharing of knowledge has taken place on traditional lands for thousands of years.