

# A Note of Caution: Uses of AI for Accessible Publishing

**Simon Mellins**

11 September 2024



**Keywords:** Accessibility, Content Creators, Large Language Models.

## Licence agreement

© 2024 Simon Mellins. This article is distributed under the terms of the Creative Commons Attribution Licence (CC-BY) 4.0 <https://creativecommons.org/licences/by/4.0/>, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## Peer review

This article has been peer reviewed through the journal's standard double-blind peer review.

## Article type

Research Article

## DOI

10.14324/111.444.2398-4732.2003

## First published

28 March 2025

## Abstract

As accessibility legislation becomes increasingly stringent, publishers are turning to AI to help meet their obligations. However, the integration of AI into accessible publishing presents significant challenges. While AI tools show promise, their application in accessibility remediation is far from straightforward and requires careful consideration. Tasks such as providing image descriptions, creating accessibility metadata, and structuring content for assistive technologies require complex, subjective editorial decisions that AI currently struggles to handle effectively. AI's limitations in contextual analysis and its inability to fully replicate human judgment mean that its use for these tasks is not yet reliable without substantial human oversight. Additionally, issues like copyright risk and the need for tailored content checks further complicate AI's role in accessibility. This article explores the current state of AI in accessible publishing, emphasizing the need for cautious, informed implementation and further research to assess its true potential in this critical area.

## Biography

Simon Mellins is a Digital Publishing Consultant ([simonmellins.com](https://simonmellins.com)) specialising in accessibility transformation, for organisations needing to be compliant with the European Accessibility Act and other similar legislation. He has worked in digital publishing for over 15 years, including spearheading the accessibility transformation at Penguin Random House UK and now as a consultant working with many of the UK's top publishers, large and small.

As publishers and other content creators see accessibility legislation looming, it makes absolute sense that time and resource-strapped organisations are looking to AI to help meet their accessibility obligations. Whilst this is worth exploring, significant caution must be brought to bear, as the application of AI in accessible publishing is not straightforward and creates its own unique set of issues to solve.

As a consultant working with a range of publishing organisations to help with accessibility strategy, within the first few minutes I can almost guarantee ‘AI’ will be raised by the person I’m speaking with. And who can blame them - publishing companies are not tech firms, and as such haven’t necessarily thought much about accessibility and don’t fully understand its application or relevance to their work - yet they find themselves suddenly tasked with fully implementing it in a matter of months if they want to retain access to European markets, as well as meet the increasingly high bar for accessibility set by other legislation and organisational procurement rules around the world. With this historical moment handily coinciding with the rise of the next generation of Artificial Intelligence tools, of course many see it as a neat dovetail as well as a perfect opportunity to show themselves at the forefront of technological development.

But this might be a misnomer. AI has only very limited application for accessibility remediation at the moment, and understanding *why* this is the case is key to understanding Large Language Models and related technologies themselves. It also helps illuminate the nature of accessibility work and semantics.

All this begins with a foundational principle: accessibility work is fundamentally editorial/content work, *not* a technical task - the tasks involved here are relatively simple from a technical perspective. Let’s review the core tasks of accessibility remediation with that in mind:

- 1) **Text alternatives to non-text media (e.g. image description, video description):** This is normally the first use of AI that people suggest. But, as we’ll discuss later, describing visuals is not just simple observation, but something that happens in context. The meaning of an image, and whether it even warrants description, is inextricably related to the context of the image’s use.
- 2) **Accessibility tags in content (epub:type and ARIA):** Labelling the parts of a publication is not always simple and there can be nuance, especially in works with unusual structure.
- 3) **Logically ordered and structured code, rather than using visual tricks to infer structure:** It’s hard to explain this to anyone not passingly familiar with frontend development or basic HTML, but there are ways you can create an eBook or web page that ‘look’ correct but are structurally all over the place and will be impossible for assistive technologies like screen readers to navigate correctly. Getting this structure right can involve a bit of thinking, especially when it comes to hierarchy and order of objects/sections that may have not had a strict linear hierarchy or order in the original text, which was designed for visual consumption only rather than for assistive technology.

- 4) **Accessibility Metadata:** Creating metadata that describes the accessibility properties of a publication, including features and benefits as well as warnings and even noting where testing hasn't yet been conducted. This requires knowledge of works as a whole to accurately assess.
- 5) **Visual Clarity:** Ensuring sufficient levels of colour contrast, clear and legible fonts, and many more design and functionality considerations

All the above require complex and subjective thinking, where two humans may not even agree on a 'correct' answer. Many academics and practitioners writing and talking about AI at the moment are proposing that the best use of AI is for tasks where there is an indisputable 'right' answer, but not for anything subjective or nuanced. That is the nature of Large Language Models (the type of AI we are discussing here, which is trained on large amounts of text), which have only relational ways of 'knowing' and cannot easily perform contextual analysis in the same way as a human brain - at least not yet. And that 'yet' is a huge computer science problem that is out of scope of us mere publishing mortals - we'll just have to be patient.

When we say that AI is 'nearly there', remember that the effort is not linear. This 'last bit' needed to make AI capable of the very-human skill of *global* contextualisation is huge. I like to compare this to how the effort required to increase an object's speed rises exponentially, rather than linearly. This means that doubling the speed of an object requires significantly more than double the energy. In a car this is due to factors like mass, rolling resistance, aerodynamic drag and engine efficiency. For AI, this is due to the sheer size of contextual frames, and the different ways that 'knowing' and 'thinking' exist beyond the merely relational.

Coming back to the most popular current topic of image descriptions, it's worth dwelling on where AI struggles, as it better illustrates the broader point of its limitations.

How an image description should be written needs to take into account many overlapping contexts: The intended audience/s, assumptions about their prior knowledge, reading level, familiarity with different terms etc; context of the image within the book and what a reader's progress so far will mean for their understanding of the image; context of the subject area and known norms; the publisher/imprint/editor's preferences and norms; and anything else affecting that image's informational payload in this specific context. Think about all the different things that a photo of da Vinci's *Mona Lisa* could mean in different books through different lenses, for instance an art book, or something from a sociological lens, or a tourist manual.

Getting an AI to take all of that into account may be possible with careful prompt engineering (the process of refining the questions asked of AIs), but it would certainly require a lot of that each time, as well as careful checking by a human afterwards. Is this actually saving time or money? If we wanted to fully automate it, the AI would need to fully ingest and 'understand' the book and its cultural/social/pedagogical context. Is that realistic? And that's to say nothing of copyright risk, depending on how the AI is hosted.

There is another level to context which is crucial here: Did the image even need describing in the first place? Is it already described by surrounding text or a suitably descriptive caption? If so, then we don't want to serve print disabled people needless repetition. Or is the image purely decorative? Whilst some debate surrounds this, there is broad consensus that assistive technology users would rather not have a poetic description of an ornate 'swirly line separator' ornament you've chosen to use between two paragraphs, or an abstract shape or some clip art stock image you've used to help visually balance whitespace. Again, these assertions are relatively straightforward for a human, but a significant challenge for AI.

Finally, remember that to users of assistive technology, image descriptions are as much part of the 'editorial voice' as any other part of the book. The Platonic ideal of a totally neutral and functional image description is a persistent falsehood, and most publishers will want to have some influence on the AI's tone and linguistic choices. One can imagine some of the awkward and potentially offensive image descriptions that are likely coming soon to a social media feed near you.

The sum of all this is not that AI *cannot* be used, but rather that if it is, a human *must* check its work. The marketing message from companies offering this service is generally that having an AI write a draft speeds up human creation of image descriptions. Great if true, but that is a pretty significant cognitive assertion being made without any objective data to back it up, and the stakes are high for budget-strapped publishers with tight schedules. **The cognitive load of checking an AI's image description is not necessarily any lower than that of writing the image description in the first place.** It may even be more, depending on context. This will all depend on a publisher's specific content, along with workflow and scheduling needs.

One other thing to note in closing: Copyright risk. Terms and conditions are of paramount importance when assessing an AI, given few publishers will self-host their own instance. The last thing a publisher wants is its copyrighted works becoming part of a corpus of 'general knowledge'. Copyright *erosion* is a subtle risk that can be harder to spot and fight than outright theft. This is a huge factor that must be weighed up and thoroughly checked before bringing in an AI to help ease the accessibility workload. Any service provider should be ready to discuss this, and robust contract provisions are a must.

None of this is intended as a blanket call to 'avoid AI for accessibility work'. Anything that potentially eases transformation and transition to accessible publishing is worth exploring, and there are few nobler causes. But the reality of AI's usefulness for this kind of problem is not at all settled or fully understood at this moment, and more research is needed to help guide publishers in making these decisions, and to counter often overly optimistic claims made by some service providers.