

Using artificial-intelligence tools to make \LaTeX content accessible to blind readers

Gerd Kortemeyer

Abstract

Screen-reader software enables blind users to access large segments of electronic content, particularly if accessibility standards are followed. Unfortunately, this is not true for much of the content written in physics, mathematics, and other STEM disciplines, due to the strong reliance on mathematical symbols and expressions, which screen-reader software generally fails to process correctly.

A large portion of such content is based on source documents written in \LaTeX , which are rendered to PDF or HTML for online distribution. Unfortunately, the resulting PDF documents are essentially inaccessible, and the HTML documents vary greatly in accessibility, since their accessibility-standards compliant rendering is cumbersome at best.

This paper explores the possibility of generating standards-compliant, accessible HTML from \LaTeX sources using Large Language Models. It is found that the resulting documents are highly accessible, with possible complications occurring when the artificial intelligence tool starts to interpret the content.

1 Introduction

Compared to paper printouts, electronic documents generally have the advantage of being accessible to blind readers using screen-reader technology [2]: the software reads the text on the screen using speech synthesis (often at amazing speeds!), and users can navigate the document using the keyboard by following hyperlinks or jumping from section to section. How a screen reader works, unfortunately, is hard to convey without observing users [22, 24].

Particularly in physics and mathematics, but also in other STEM disciplines, \TeX (and for several journals particularly \LaTeX) continues to be the dominant typesetting environment for scientific articles and reports, as well as for lecture scripts and other teaching materials. The \LaTeX source is most frequently processed into a PDF-presentation document for distribution. While screen-reader technology generally works well for web pages, particularly if accessibility standards like the Web Content Accessibility Guidelines (WCAG) [26] are followed, it frequently fails to adequately make PDF documents accessible, which leaves large segments of scientific content out of immediate reach for blind users [17]. Particularly in Europe, where lectures are often accompanied by instructor-authored scripts rather than standard

textbooks, these shortcomings of PDF also strongly disadvantage blind students while pursuing higher education [10, 21].

Currently, either particular \LaTeX packages with some restrictions on syntax and later rendering need to be applied when writing the \LaTeX source [1, 16], or specialized tools need to be used to later mark up the rendered PDF [6, 13]. In an alternative approach, the \LaTeX source is rendered to HTML with embedded MathJax [3, 4] using tools like `htlatex`; unfortunately, this approach is also not compatible with all of \LaTeX 's syntactical variants. Given the current restrictions, some blind users prefer having the original \LaTeX source available to their screen-reader software [23]; however, students in introductory courses are likely not proficient in \LaTeX .

Narrating mathematical formulas requires either explicit semantic markup, which is usually not provided, or other sense-making technologies. When automated mechanisms fail, this work currently needs to be done by human experts. The Large Language Model GPT has been shown to have remarkable mastery of physics and mathematics [15], and it might be up to the task. In this exploratory study, GPT-4 [19] is used to translate the \LaTeX source to HTML while narrating the formulas and graphics in plain language.

2 Presentation versus semantics

Any kind of document we see on paper or on screen is a presentation of content, and we are using certain visual clues and symbols to convey meaning (“semantics”). For example, something that is larger and boldface is interpreted as a heading for a new section or subsection of a document, and when we glance over the document, we might jump from heading to heading to see what we might like to read. Screen-reader software cannot glance over a document; it needs to rely on semantic clues in the document. To the eye, `\section*{Findings}` and `\textbf{\Large FINDINGS}` will look the same, but the former has embedded meaning, namely, this is where an unnumbered section starts, while the latter just specifies presentation that the eye (actually, our whole image-processing pipeline) needs to interpret and make sense of.

One of the many reasons for using \LaTeX is its ability to beautifully typeset mathematical equations. A large challenge in narrating these formulas is that mathematical typesetting is not necessarily semantic, nor even designed to be, and in fact there are ambiguities [11, 14]. \LaTeX , and for that matter also handwriting, provides a presentation view of an equation [7]. An expert usually sees what is meant,

Testing \LaTeX Accessibility

Zaphod Beeblebrox
Sirius Cybernetics Corporation
 (Dated: June 3, 2023)

This is a document to explore the ability of GPT-4 to translate scientific content written in \LaTeX into WCAG-compliant HTML. The document intentionally uses several cumbersome \LaTeX constructs while at the same time not claiming to make much physics sense.

I. INTRODUCTION

When GPT-4 first came out, it immediately gained the attention of scientists [1]. Here we test if it can be used to create WCAG-compliant HTML [2] based on \LaTeX -source code.

II. SOME USEFUL FORMULAS

When formulating what is called Special Relativity today, Einstein started with the Maxwell equations, like this one [3]:

$$\oiint \vec{E} \cdot d\vec{S} = \frac{1}{\epsilon_0} \iiint \rho dV \quad (1)$$

Table I lists some of the most commonly used relativistic equations [3].

And then there is one of the most famous formulas of physics [4]:

$$E = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \stackrel{v=0}{=} mc^2 \quad (2)$$

Of course, all of these formulas are merely consequences of the Lorentz transformation between two moving frames of reference. Consider a rotation matrix about the z -axis in three-dimensional space,

$$\vec{x}' = D\vec{x} \quad (3)$$

$$D = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4)$$

This is illustrated in Fig. 1.

TABLE I. Some relativistic formulas

Common Name	Meaning	Formula
Time Dilation	Transformation of time-difference between moving frames	$\Delta t' = \frac{\Delta t}{\sqrt{1 - \frac{v^2}{c^2}}}$
Length Contraction	Transformation of a length of a moving object	$L' = L\sqrt{1 - \frac{v^2}{c^2}}$

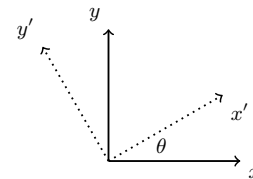


FIG. 1. Rotating a coordinate system.

The Lorentz Transformation is a four-dimensional rotation with a matrix Λ^ν_μ :

$$x'^\nu = \Lambda^\nu_\mu x^\mu \quad (5)$$

where we imply the summation convention,

$$x'^\nu = \sum_{\mu=0}^3 \Lambda^\nu_\mu x^\mu \quad (6)$$

III. OTHER REMARKS

One could say a lot more about Special Relativity, but *don't panic* [5]! At the very least, remember equation 2!

- [1] K. Sanderson, GPT-4 is here: what scientists think, *Nature* **615**, 773 (2023).
 [2] B. Caldwell, M. Cooper, L. G. Reid, G. Vanderheiden, W. Chisholm, J. Slatin, and J. White, Web content accessibility guidelines (WCAG) 2.0, WWW Consortium (W3C) **290**, 1 (2008).

- [3] A. Einstein, Zur Elektrodynamik bewegter Körper, *Annalen der Physik* **17**, 891 (1905).
 [4] A. Einstein, Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?, *Annalen der Physik* **18**, 639 (1905).
 [5] D. Adams, *The Ultimate Hitchhiker's Guide to the Galaxy: Five Novels in One Outrageous Volume* (Del Rey, 2010).

Figure 1: \LaTeX rendering of the test document.

(Prompt text:)

```

Model: GPT-4

Translate the following LaTeX and BibTeX into screenreader-accessible HTML, translating the formulas into English (including the indices), and translating labels and citations into anchors and links. For figures, just give a description in English, no code for embedding an image:

\documentclass[aps,reprint,amsmath,amssymb,prper,floatfix]{revtex4-2}
\usepackage{wasysym}
\usepackage{tikz}

\newcommand{\thesquareroot}{\sqrt{1-\frac{v^2}{c^2}}}

\begin{document}

...

\section{Other Remarks}
One could say a lot more about Special Relativity, but \textit{don't panic}-\cite{adams2010}! At the very least, remember equation-\ref{eq:famous}!
\bibliography{testdoc}

\end{document}

@article{einstein1905a,
  title={Zur Elektrodynamik bewegter K{\o}rper},
  author={Einstein, Albert},
  journal={Annalen der Physik},
  volume={17},
  pages={891--921},
  year={1905}
}

```

Response:

```

Here's a translation of your LaTeX and BibTeX document into screenreader-accessible HTML:

html
Copy code

<!DOCTYPE html>
<html>
<head>
  <title>Testing LaTeX Accessibility</title>
</head>
<body>
  <header>
    <h1>Testing LaTeX Accessibility</h1>
    <p>Author: Zaphod Beeblebrox, Sirius Cybernet</p>
    <p>Date: Today's Date</p>
  </header>

```

Figure 2: ChatGPT dialogue.

as marked up by `<h1>...</h1>`, `<h2>...</h2>`, etc. GPT correctly interpreted

```

\begin{document}

\title{Testing \LaTeX{} Accessibility}
\author{Zaphod Beeblebrox}
\affiliation{Sirius Cybernetics Corporation}
\date{\today}

\begin{abstract}
This is a document to explore ...
and inserted structures such as
<body>
  <header>
    <h1>Testing LaTeX Accessibility</h1>
    <p>Author: Zaphod Beeblebrox, Sirius ...</p>
    <p>Date: Today's Date</p>
  </header>
  <section>
    <h2>Abstract</h2>
    <p>This is a document to explore ...</p>
  </section>

```

which not only use the correct section headers, but also, in one of the responses, provide additional semantic markup like the HTML5 tags `<header>` and `<section>`; those potentially helpful tags were missing in the second rendering. However, both versions were perfectly accessible to and navigable by screenreader software.

4.2 New commands

Several of the regular HTML rendering tools fail to interpret `\newcommand` correctly. The test document included

```

\newcommand{\thesquareroot}{\sqrt{1-\frac{v^2}{c^2}}}

```

which was used in nearly all subsequent formulas and interpreted correctly in every case. For example,

$$L' = L \sqrt{1 - \frac{v^2}{c^2}}$$

was narrated by GPT as

L prime equals L times the square root of 1 minus v squared over c squared.

4.3 Equations

GPT-4 overall did very well in narrating equations. As an example,

```

\begin{equation}
\oiint \vec{E} \cdot d\vec{S} =
\frac{1}{\epsilon_0} \iiint \varrho \, dV
\end{equation}

```

was translated in the first version (Fig. 3) into

The closed integral of E dot dS equals one over epsilon zero times the triple integral of rho dV.

and, in the second version (Fig. 5):

Testing LaTeX Accessibility

Author: Zaphod Beeblebrox, Sirius Cybernetics Corporation

Date: Today's Date

Abstract

This is a document to explore the ability of GPT-4 to translate scientific content written in LaTeX into WCAG-compliant HTML. The document intentionally uses several cumbersome LaTeX constructs while at the same time not claiming to make much physics sense.

Introduction

When GPT-4 first came out, it immediately gained the attention of scientists^[1]. Here we test if it can be used to create WCAG-compliant HTML^[2] based on LaTeX source code.

Some Useful Formulas

When formulating what is called Special Relativity today, Einstein started with the Maxwell equations, like this one^[3]:

The closed integral of $\mathbf{E} \cdot d\mathbf{S}$ equals one over epsilon zero times the triple integral of $\rho \, dV$.

Table 1 lists some of the most commonly used relativistic equations^[3].

Table 1: Some relativistic formulas

Common Name	Meaning	Formula
Time Dilation	Transformation of time-difference between moving frames	$\Delta t' = \Delta t / \sqrt{1 - v^2/c^2}$
Length Contraction	Transformation of a length of a moving object	$L' = L \sqrt{1 - v^2/c^2}$

And then there is one of the most famous formulas of physics^[4]:

$E = mc^2 / \sqrt{1 - v^2/c^2}$, which simplifies to $E = mc^2$ when $v = 0$.

Figure 3: First HTML rendering of the test document, top part.

Of course, all of these formulas are merely consequences of the Lorentz transformation between two moving frames of reference. Consider a rotation matrix about the z-axis in three-dimensional space, $x' = D x$ and $D = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$. This is illustrated in Figure 1.

Figure 1: Rotating a coordinate system.

A coordinate system is represented with x and y axes. Two additional axes, x' and y' , are represented, rotated from the original x and y axes by an angle θ .

The Lorentz Transformation is a four-dimensional rotation with a matrix $\lambda_{\nu\mu}$:

$x'_{\nu} = \lambda_{\nu\mu} x_{\mu}$

where we imply the summation convention,

$x'_{\nu} = \sum_{\mu=0}^3 \lambda_{\nu\mu} x_{\mu}$.

Other Remarks

One could say a lot more about Special Relativity, but *don't panic*^[5]! At the very least, remember equation ²!

References

- Einstein, Albert. "Zur Elektrodynamik bewegter Körper." Annalen der Physik 17.891-921 (1905).
- Einstein, Albert. "Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?." Annalen der Physik 18.639-641 (1905).
- Adams, Douglas. The Ultimate Hitchhiker's Guide to the Galaxy: Five Novels in One Outrageous Volume. Del Rey, 2010.
- Caldwell, Ben, et al. "Web content accessibility guidelines (WCAG) 2.0." WWW Consortium (W3C) 290.1-34 (2008).
- Sanderson, Katharine. "GPT-4 is here: what scientists think." Nature 615.7954: 773 (2023).

Figure 4: First HTML rendering of the test document, bottom part.

Testing LaTeX Accessibility

Author: Zaphod Beeblebrox

Affiliation: Sirius Cybernetics Corporation

Date: 3rd June, 2023

Abstract

This is a document to explore the ability of GPT-4 to translate scientific content written in LaTeX into WCAG-compliant HTML. The document intentionally uses several cumbersome LaTeX constructs while at the same time not claiming to make much physics sense.

Introduction

When GPT-4 first came out, it immediately gained the attention of scientists. [Sanderson 2023](#). Here we test if it can be used to create WCAG-compliant HTML based on LaTeX-source code. [Caldwell 2008](#).

Some Useful Formulas

When formulating what is called Special Relativity today, Einstein started with the Maxwell equations, like this one [Einstein 1905](#):

"The surface integral over the vector E dotted with the infinitesimal area vector dS equals 1 over the permittivity of free space times the volume integral over the volume charge density times the infinitesimal volume dV ."

Table 1 lists some of the most commonly used relativistic equations [Einstein 1905](#).

Some relativistic formulas		
Common Name	Meaning	Formula
Time Dilation	Transformation of time-difference between moving frames	$\Delta t' = \Delta t / \sqrt{1 - (v^2/c^2)}$
Length Contraction	Transformation of a length of a moving object	$L' = L \sqrt{1 - (v^2/c^2)}$

And then there is one of the most famous formulas of physics [Einstein 1905](#):

" $E = mc^2$ when v is not equal to 0 . When v equals 0 , E equals mc^2 ."

Figure 5: Second HTML rendering of the test document, top part.

The surface integral over the vector E dotted with the infinitesimal area vector dS equals 1 over the permittivity of free space times the volume integral over the volume charge density times the infinitesimal volume dV .

Both narratives are adequate and could be spoken by screen-reader software. The first rendering is rather literal, though, as it fails to mention that the first integral is a surface integral, and that \vec{E} and $d\vec{S}$ are vectors. A physicist will understand this concise narrative and probably be thankful that it is not too verbose, while a learner (actually, any learner, blind or seeing) would be thankful for the second, interpretive rendering. It is remarkable that GPT provided the physical meaning of the symbols alongside the narrative of their mathematical connections.

Another example is

```
\begin{equation}\label{eq:famous}
E=\frac{mc^2}{\sqrt{1-v^2/c^2}}
\end{equation}
```

which is rendered

E equals mc^2 squared divided by the square root of 1 minus v squared over c squared, which simplifies to mc^2 squared when v equals 0 .

in one version, and

E equals mc^2 squared divided by the square root of $(1 - (v^2/c^2))$ when v is not equal to 0 . When v equals 0 , E equals mc^2 squared.

in the other. Both are essentially equivalent, and both do a remarkable job of interpreting `\stackrel{rel}{v=0}{=}mc^2` (although the full expression is correct regardless of v being zero or not).

Of course, all of these formulas are merely consequences of the Lorentz transformation between two moving frames of reference. Consider a rotation matrix about the z-axis in three-dimensional space,

"The vector x prime equals the matrix D times the vector x ."

The rotation matrix D is a 3×3 matrix with elements $\cos \theta$, $\sin \theta$, 0 on the first row, $-\sin \theta$, $\cos \theta$, 0 on the second row, and 0 , 0 , 1 on the third row.

This is illustrated in Figure 1 which shows a rotation of a coordinate system with x , y axes and the rotated x prime, y prime axes.

The Lorentz Transformation is a four-dimensional rotation with a matrix λ^{ν}_{μ} :

" x^{ν} prime equals λ^{ν}_{μ} times x^{μ} "

where we imply the summation convention,

" x^{ν} prime equals the sum from μ equals 0 to 3 of λ^{ν}_{μ} times x^{μ} ."

Other Remarks

One could say a lot more about Special Relativity, but *don't panic* [Adams 2010](#)! At the very least, remember equation E equals mc squared divided by the square root of $(1 - (v^2/c^2))$!

References

Einstein, Albert. (1905). Zur Elektrodynamik bewegter Körper. Annalen der Physik, 17, 891–921.

Einstein, Albert. (1905). Ist die Trägheit eines Körpers von seinem Energieinhalt abhängig?. Annalen der Physik, 18, 639–641.

Adams, Douglas. (2010). The Ultimate Hitchhiker's Guide to the Galaxy: Five Novels in One Outrageous Volume. Del Rey.

Caldwell, Ben, Cooper, Michael, Reid, Loretta Guarino, Vanderheiden, Gregg, Chisholm, Wendy, Slatin, John, & White, Jason. (2008). Web content accessibility guidelines (WCAG) 2.0. WWW Consortium (W3C), 290, 1–34.

Sanderson, Katharine. (2023). GPT-4 is here: what scientists think. Nature, 615(7954), 773. Nature.

Figure 6: Second HTML rendering of the test document, bottom part.

A true challenge could have been posed by the rotation matrix:

```
\begin{eqnarray}
\vec{x}' = D \vec{x} \\
D = \begin{pmatrix}
\cos \theta & \sin \theta & 0 \\
-\sin \theta & \cos \theta & 0 \\
0 & 0 & 1
\end{pmatrix}
\end{eqnarray}
```

Intriguingly, this was incorporated into the running narrative in one version (Fig. 4),

Consider a rotation matrix about the z-axis in three-dimensional space, x prime equals $D x$ and D equals $\cos \theta$, $\sin \theta$, 0 in the first row, minus $\sin \theta$, $\cos \theta$, 0 in the second row, and 0 , 0 , 1 in the third row.

and spelled out rather literally in the other (Fig. 6),

"The vector x prime equals the matrix D times the vector x ."

The rotation matrix D is a 3×3 matrix with elements $\cos \theta$, $\sin \theta$, 0 on the first row, $-\sin \theta$, $\cos \theta$, 0 on the second row, and 0 , 0 , 1 on the third row.

Again, both versions correctly describe the situation, and a blind person would be able to envision the matrix in his or her head based on the description.

On the other hand, the equation

```
\begin{equation}
x'^{\nu} = \sum_{\mu=0}^3 \Lambda^{\nu}_{\mu} x^{\mu}
\end{equation}
```

is the first example of a situation where one of the narratives falls short:

x^{ν} prime equals the sum from μ equals 0 to 3 of $\Lambda^{\nu}_{\mu} x^{\mu}$.

is less helpful than

x^{ν} prime equals the sum from μ equals 0 to 3 of λ^{ν}_{μ} times x^{μ} .

since it fails to distinguish between covariant and contravariant indices. The nonchalant treatment of indices is also the reason for parenthetically introducing "including the indices" into the prompt; without that phrase, the GPT algorithm recognized the matrix multiplication for what it is and simply stated

The transformed x vector equals the sum from μ equals 0 to 3 of Λ times the original x vector.

In fact, even before that, it already assumed the summation convention and rendered $x^\nu = \Lambda^\nu_\mu x^\mu$ as

The transformed x vector equals Lambda times the original x vector.

This highly concise narration would likely be appreciated by an expert, but might generally be too far from the original.

In one of the renderings, uppercase Λ was replaced by lowercase λ . Again, one needs to remind oneself that it does not matter what the document looks like, but what it “sounds” like: screen-reader software would read both renderings simply as “lambda” without distinction between uppercase and lowercase symbols. In any case, independent of using GPT or more traditional HTML conversion methods, this could be a problem in a document that uses both uppercase and lowercase versions of the same symbol. As it turns out, GPT could recognize this; for example, `\[\Lambda=\sum_i\lambda_i\]` results in “The capital lambda (Λ) is equal to the sum of lambda (λ) over all i ” (here, the Greek characters have been rendered as Unicode in the HTML).

4.4 Links and references

Links and references were handled differently in the two renderings. For example, toward the bottom of the L^AT_EX source, it says

```
One could say a lot more about Special Relativity,
but \textit{don't panic}~\cite{adams2010}!
At the very least, remember equation~\ref{eq:famous}!
```

In the first version, the GPT algorithm had included `<p id="eq:famous">`

```
E equals mc squared divided ... </p>
```

at the place where the formula appeared, as well as `<li id="adams2010">`

```
Adams, Douglas. The Ultimate ... </li>
```

into the reference list, and it rendered the L^AT_EX as but `<i>don't panic</i>[5]!`
 At the very least, remember equation
`2!`

This in fact only works if following the hyperlinks, since neither the equation nor the reference have numbers appearing in the output, as the reference list used an unordered list `` instead of an ordered list ``. A blind user, however, could follow those hyperlinks to get to the correct equation and references.

Within the references themselves, some encoding problems are visible with the German umlaut in Fig. 4. However, “visible” is irrelevant, since the real question, common to any non-ASCII characters, is whether or not the screen-reader software handles their encoding (rather than the Safari web browser used to generate Fig. 4 for this visually-oriented paper), which will depend on the product used. As an

aside, for the German-language paper titles, WCAG would expect `...`, which was unfortunately not generated; missing that tag, the screen reader will attempt to read the German with English pronunciation, which makes the considerations about the encoding moot.

The second version uses a different approach:

```
... but <em>don't panic</em>
<a href="#adams2010">Adams 2010</a>!
At the very least, remember equation
E equals mc squared divided by the square root
of (1 minus (v squared divided by c squared))!
```

This approach works for both blind and seeing readers, but actually quoting the referenced equation could become awkward if the equation is a long one; in this case, GPT only quoted the first part, which may not have been what the author wanted to reference. This literature reference also works for seeing readers, since the author name is used. In this version, though, the table reference does not work for any readers, since it simply says

Table 1 lists some of the most commonly used ... but the table has no number.

4.5 Tables

The table in the source document was coded as

```
\begin{table}
\caption{\label{tab:formulas}Some relativistic
formulas}
\begin{tabular}{p{1.8cm}p{4cm}l}
Common Name&Meaning&Formula\\ \hline
Time\newline Dilation&Transformation ...
\end{tabular}
\end{table}
```

which, as an example, in one of the HTML versions, GPT interpreted as

```
<table id="tab:formulas">
<caption>Some relativistic formulas</caption>
<tr>
<th>Common Name</th>
<th>Meaning</th>
<th>Formula</th>
</tr>
<tr>
<td>Time Dilation</td>
<td>Transformation ...
</table>
```

Both versions generated by GPT were WCAG-compliant and can be navigated by screen readers. Particularly helpful is that the first row was interpreted as column headers (`<th>` instead of `<td>`), which is not even specified in the L^AT_EX markup.

4.6 Figures

While GPT-4 itself is multimodal, that is, it can support audio, image, and video input and output, the ChatGPT interface only allows for text-based conversations. The figure in the test document is generated using the popular TikZ package [25]:

```
\begin{figure}
\begin{tikzpicture}
\draw[thick,->] (2,0) -- (4,0)
    node[anchor=north west] {$x$};
\draw[thick,->] (2,0) -- (2,2)
    node[anchor=south east] {$y$};
\draw[dotted,thick,->] (2,0) -- (3.732,1)
    node[anchor=north west] {$x'$};
\draw[dotted,thick,->] (2,0) -- (1,1.732)
    node[anchor=south east] {$y'$};
\draw(3,0) node[anchor=south east] {$\theta$};
\end{tikzpicture}
\caption{Rotating a coordinate
system.\label{fig:coord}}.
\end{figure}
```

GPT was prompted to describe this figure, but not draw it. The first version states

```
<figure><figcaption>
    Figure 1: Rotating a coordinate system.
</figcaption>
<p>A coordinate system is represented
with x and y axes.
Two additional axes, x prime and y prime,
are represented,
rotated from the original x and y axes
by an angle theta.</p>
</figure>
```

which includes semantic tags such as `<figure>` and `<figcaption>`, while the second version has a simpler description that is woven into the narrative:

```
<p>This is illustrated in Figure 1 which shows
a rotation of a coordinate system with x, y
axes and the rotated x prime, y prime axes.</p>
```

This second version correctly recognizes the contents of the figure, but fails to mention the angle θ that is also in the figure. Overall, both descriptions are useful, which is fairly amazing: interpreting TikZ and subsequently narrating the contents of the resulting image could be considered an emergent ability of GPT (if such things exist).

5 Discussion

Ironically, if this paper gets published as a PDF, it will not be accessible to blind users. How can this be remedied, based on the findings of this exploratory study? In recent years, there has been a strong movement toward Open Access publication; based on this exploratory study, maybe publications should also become Open Source.

Over the coming years, it is to be expected that artificial-intelligence tools will increasingly become personalized agents; this is already suggested by the concepts that assistants like Siri and Alexa put forward, and by new products like Copilot. In this study, the document rendered differently every time, and the prompt was fine-tuned in an attempt to avoid particular interpretations that experienced physicists might in fact prefer. One could argue that AI-assisted sense-making should be under the control of the user, and this means any user, having a disability or not. Thus, instead of mangling the document into a pure presentation format, the source code should be made available, so AI tools can render it in ways preferred by the user.

This would mean free publication, not just in the sense of “free beer”, that is, free-of-charge, but in the sense of “free speech”, to reiterate an old free software adage [9]. Publishing the source code of manuscripts would likely have to go along with adequate licensing terms such as those defined by Creative Commons [5]. Publishers and preprint servers could make a corresponding link available to manuscripts published under those terms.

6 Limitations

The exact results of this study are not reproducible, since the GPT algorithm is not deterministic. Also, constructing a prompt for GPT is essentially trial-and-error, and better or more specific prompts than the one used for this study almost certainly exist. Finally, the study is purely exploratory: while the document was designed to be demanding in terms of L^AT_EX structures, there was no way to even remotely represent the creativity of L^AT_EX users. For large documents, GPT might encounter limitations due to its token limit.

7 Conclusion

Overall, GPT was found to be a useful tool to generate screen-reader compatible renderings of L^AT_EX documents, including document structure, equations, tables, references, and even figures that are based on textual source code. The output is compliant with accessibility standards.

As a probabilistic algorithm, the output is not reproducible, and it was found that each time a document is processed, some elements are translated better or worse than in other versions. Notably, in situations where GPT is able to interpret equations based on its training data in physics, it is capable of generating narratives that are concise and desirable for experts, as well as more elaborate narratives that are suitable for learners.

References

- [1] D. Ahmetovic, T. Armano, et al. Axessibility: A \LaTeX package for mathematical formulae accessibility in PDF documents. In *Proceedings of the 20th international ACM SIGACCESS conference on computers and accessibility*, pp. 352–354, 2018. See also tug.org/TUGboat/tb39-3/tb123ahmetovic-axess.pdf.
- [2] Y. Borodin, J.P. Bigham, et al. More than meets the eye: a survey of screen-reader browsing strategies. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*, pp. 1–10, 2010.
- [3] D. Cervone. MathJax: a platform for mathematics on the web. *Notices of the AMS*, 59(2):312–316, 2012. doi.org/10.1090/noti794
- [4] D. Cervone, P. Krautzberger, V. Sorge. New accessibility features in MathJax. *The Journal on Technology and Persons with Disabilities*, 4:167–175, 2016.
- [5] Creative Commons. Share your work. creativecommons.org/share-your-work/.
- [6] A. Darvishy. PDF accessibility: tools and challenges. In *Computers Helping People with Special Needs: 16th International Conference, ICCHP 2018, Linz, Austria, July 11–13, 2018, Proceedings, Part I*, pp. 113–116. Springer, 2018.
- [7] R. Fateman. How can we speak math. *Journal of Symbolic Computation*, 25(2), 1998.
- [8] J. Fenn. Managing citations and your bibliography with Bib \TeX . *The Prac \TeX Journal*, 4:2006–4, 2006. tug.org/pracjourn/2006-4/fenn
- [9] Free Software Foundation. What is free software? www.gnu.org/philosophy/free-sw.en.html.
- [10] J. Gilligan. Competencies for educators in delivering digital accessibility in higher education. In *Universal Access in Human-Computer Interaction. Applications and Practice: 14th International Conference, UAHCI 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22*, pp. 184–199. Springer, 2020.
- [11] T. Groza, S. Handschuh, et al. SALT-semantically annotated \LaTeX for scientific publications. In *The Semantic Web: Research and Applications: 4th European Semantic Web Conference, ESWC 2007, Innsbruck, Austria, June 3–7, 2007. Proceedings 4*, pp. 518–532. Springer, 2007.
- [12] S. Iwarsson, A. Ståhl. Accessibility, usability and universal design — positioning and definition of concepts describing person-environment relationships. *Disability and rehabilitation*, 25(2):57–66, 2003.
- [13] A. Jembu Rajkumar, J. Lazar, et al. PDF accessibility of research papers: What tools are needed for assessment and remediation? In *Proceedings of the 53rd Hawaii International Conference on System Sciences (HICSS)*, 2020.
- [14] M. Kohlhas. Using \LaTeX as a semantic markup format. *Mathematics in Computer Science*, 2(2):279–304, 2008.
- [15] G. Kortemeyer. Could an artificial-intelligence agent pass an introductory physics course? *Phys. Rev. Phys. Educ. Res.*, 19:010132, May 2023. doi.org/10.1103/PhysRevPhysEducRes.19.010132
- [16] A. Manzoor, S. Arooj, et al. ALAP: Accessible \LaTeX -based mathematical document authoring and presentation. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pp. 1–12, 2019.
- [17] J.T. Nganji. The Portable Document Format (PDF) accessibility practice of four journal publishers. *Library & Information Science Research*, 37(3):254–262, 2015.
- [18] A. Ogawa. Rev \TeX version 4.0, an authoring package by the American Physical Society. *TUGboat* 22(3):131–133, 2001. tug.org/TUGboat/tb22-3/tb71ogawa.pdf
- [19] OpenAI. GPT Release Notes. help.openai.com/en/articles/6825453-chatgpt-release-notes.
- [20] H. Persson, H. Åhman, et al. Universal design, inclusive design, accessible design, design for all: different concepts — one goal? On the concept of accessibility — historical, methodological and philosophical aspects. *Universal Access in the Information Society*, 14:505–526, 2015.
- [21] A. Petz, K. Miesenberger. Supporting blind students in STEM education in Austria. In *AAATE Conf.*, pp. 27–31, 2015.
- [22] Salt Lake Community College. Screen reader demo. www.youtube.com/watch?v=q_ATY9gim0M.
- [23] V. Sorge, C. Chen, et al. Towards making mathematics a first class citizen in general screen readers. In *Proceedings of the 11th web for all conference*, pp. 1–10, 2014.
- [24] University of California San Francisco. Screen reader demo for digital accessibility. www.youtube.com/watch?v=dEb15jvLKGQ.
- [25] Z. Walczak. Graphics in \LaTeX using TikZ. *TUGboat* 20:176–179, 2008. tug.org/TUGboat/tb29-1/tb91walczak.pdf
- [26] WWW Consortium (W3C). Web content accessibility guidelines (WCAG) 2.0. www.w3.org/TR/WCAG20/.
- ◇ Gerd Kortemeyer
ETH Zurich
Rämistrasse 101
8092 Zurich
Switzerland
kgerd (at) ethz (dot) ch
<https://gerdkortemeyer.com/>
ORCID 0000-0001-6643-9428