

What are tags?

Microsoft Word, Indesign, PDFs -- they all use tags, but what *are* tags?

Take a look at the below blurb from Microsoft Word:



Behind the scenes, Microsoft Word builds the page based on structured *tags*. Think of these tags as containers with names that explain what's *inside* the container. Every single element on the page is contained in a tag. A heading at level 1 would be h1, level 2 = h2, and so forth. Paragraphs would be “p”. Images as “image”. Each tag “container” has a start <> and an end </>. Every tag is in a specific order. The blurb above would look something like this:

```
<h1>Annual Report 2024</h1>
<p> July 20, 2032</p>
<h2>Executive Summary</h2>
<p>The company performed well this year...</p>
<image alt="bar chart showing revenue growth. See table
2.1 for accessible data from this chart." />
<p>Revenue increased by 12% over the prior year...</p>
<list>
  <item>Revenue up 12%</item>
  <item>Costs down 4%</item>
</list>
```

A screen reader reads the tags from top to bottom. If you were to have a floating text box, where would it appear in that tag tree? Microsoft Word doesn't know your intention, so it might place it pretty much anywhere in the tree, and you have very limited control over it. The screen reader might read your text box in an inappropriate position on the page, which can confuse the reader. Now imagine having multiple text boxes on the page. Things can get messy very quickly. This also rings true for images on the page. When should the alt text for the image be read to the user? This is why it's important to try and keep your images inline with the text, and limit the use of text boxes altogether.

Below is an example of a tag tree as seen in Adobe Acrobat:

