



# HTML is easy(?)

Lessons from the Web



**Andy Carter**  
andycarter.dev

**Front End Sheffield**  
January 2024

# 0. Introduction

## HTML is easy(?) - Lessons from the Web

What can **200 major websites**, from across the Web, tell us about the state of HTML?



## Why does well written HTML matter?

- Accessibility
- User-experience
- SEO
- Robust




# 1. Semantics

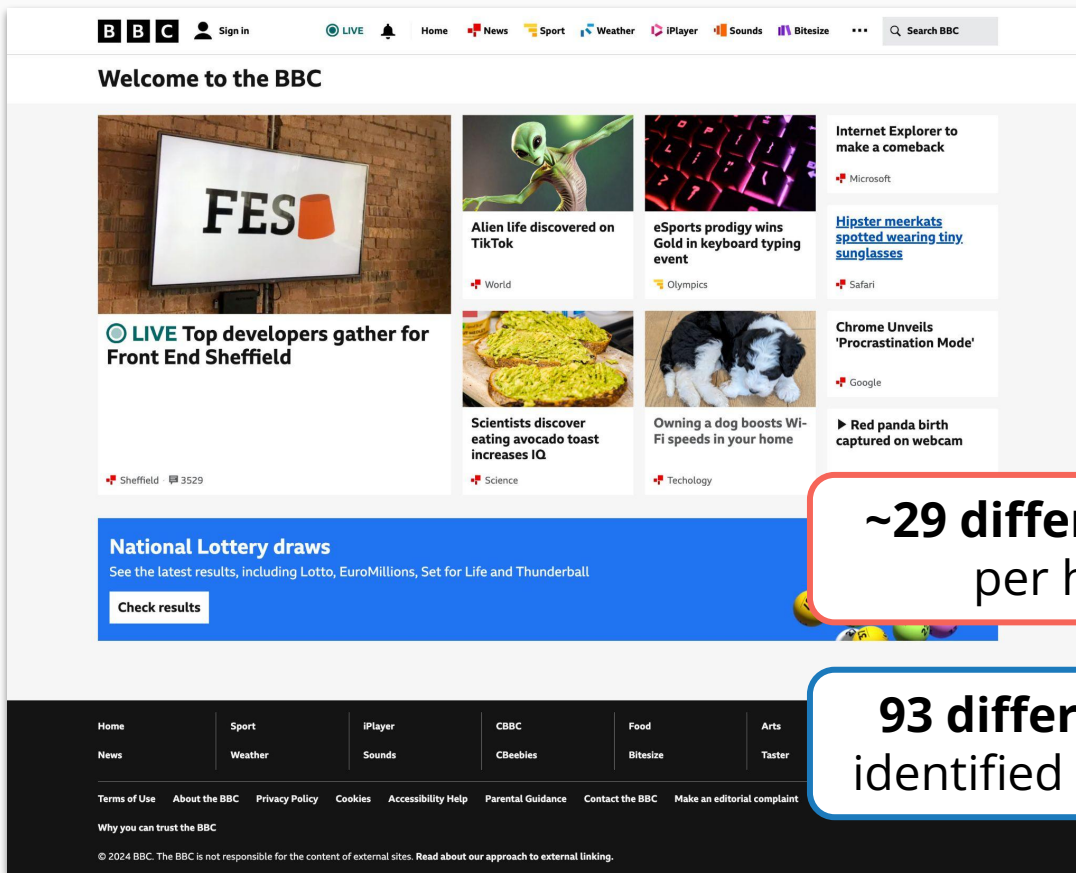
## Confusion

`<article>` **VS**  
`<section>`?



 Smashing Magazine - [bit.ly/article-vs-section](https://bit.ly/article-vs-section)

# Semantic elements

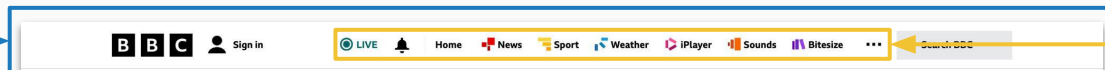


~29 different elements  
per homepage

93 different elements  
identified across the sites

# Semantic elements

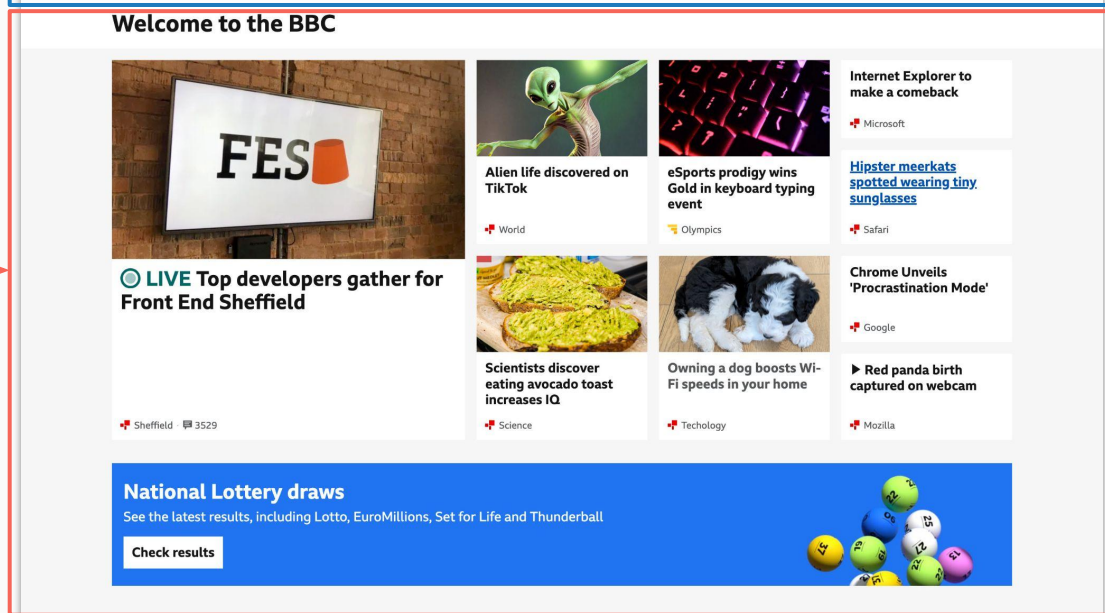
header



The header of the BBC website, containing the BBC logo, a 'Sign in' link, a navigation menu with links for LIVE, Home, News, Sport, Weather, iPlayer, Sounds, and Bitesize, and a search bar.

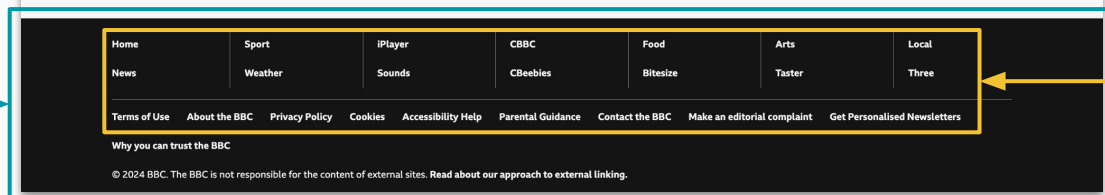
nav

main



The main content area of the BBC website, featuring a 'Welcome to the BBC' heading, a grid of news stories, and a National Lottery draws banner. The news stories include: 'FES Top developers gather for Front End Sheffield', 'Alien life discovered on TikTok', 'eSports prodigy wins Gold in keyboard typing event', 'Hipster meerkats spotted wearing tiny sunglasses', 'Owning a dog boosts Wi-Fi speeds in your home', and 'Red panda birth captured on webcam'. The National Lottery banner includes a 'Check results' button.

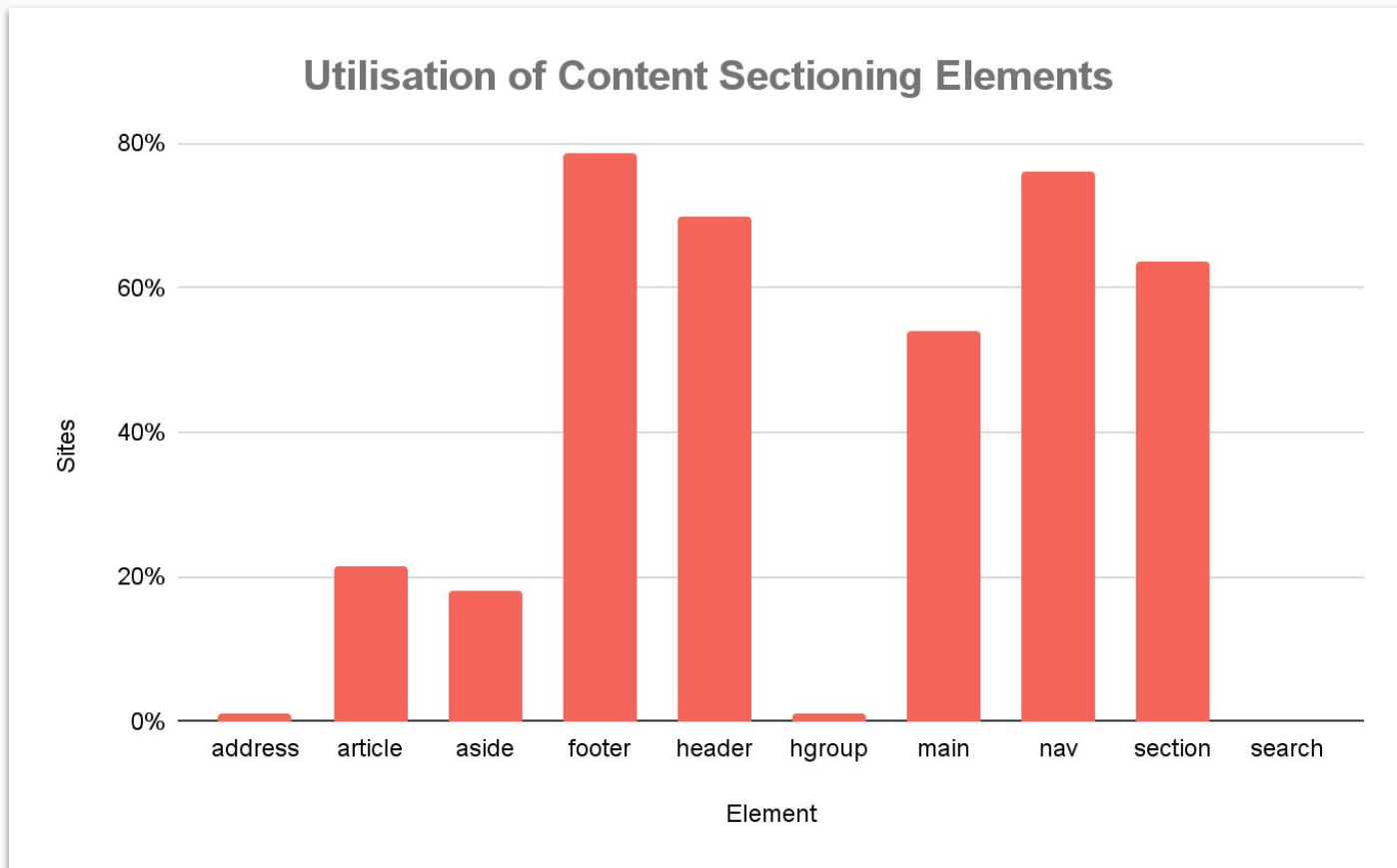
footer



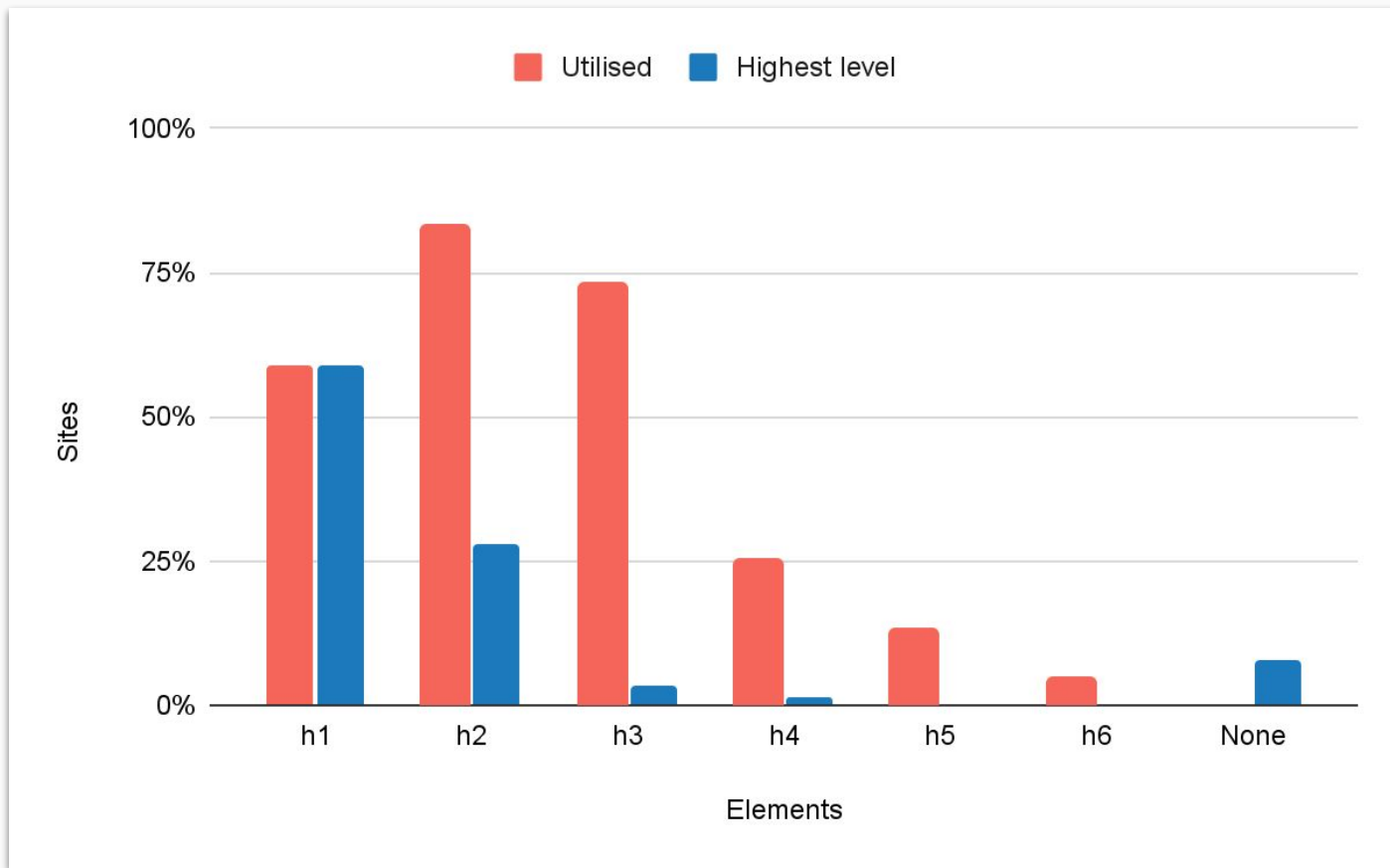
The footer of the BBC website, containing a grid of links for Home, Sport, iPlayer, CBBC, Food, Arts, Local, News, Weather, Sounds, CBeebies, Bitesize, Taster, and Three. It also includes links for Terms of Use, About the BBC, Privacy Policy, Cookies, Accessibility Help, Parental Guidance, Contact the BBC, Make an editorial complaint, and Get Personalised Newsletters. The footer ends with the text 'Why you can trust the BBC' and a copyright notice: '© 2024 BBC. The BBC is not responsible for the content of external sites. Read about our approach to external linking.'



## Content Sectioning elements



# Headings



**Semantics can be hard to understand / implement**



## 2. Common mistakes

## Validation errors

1. Invalid child elements (~67%)
2. Bad attribute values (~66%)
3. Disallowed attributes (~65%)
4. CSS errors (~47%)
5. Missing required attributes (~42%)

**94%** of sites had errors

**~141** errors per homepage

## Validation errors

1. Invalid child elements (~67%)
  2. Bad attribute values (~66%)
  3. Disallowed attributes (~65%)
  4. CSS errors (~47%)
  5. Missing required attributes (~42%)
- Elements
- Attributes
- 
- The diagram shows a list of five validation error categories. A red box highlights the first category, 'Invalid child elements (~67%)'. A red line with an arrow points from this box to the label 'Elements' on the right. A blue box highlights the remaining four categories: 'Bad attribute values (~66%)', 'Disallowed attributes (~65%)', 'CSS errors (~47%)', and 'Missing required attributes (~42%)'. A blue line with an arrow points from this box to the label 'Attributes' on the right.

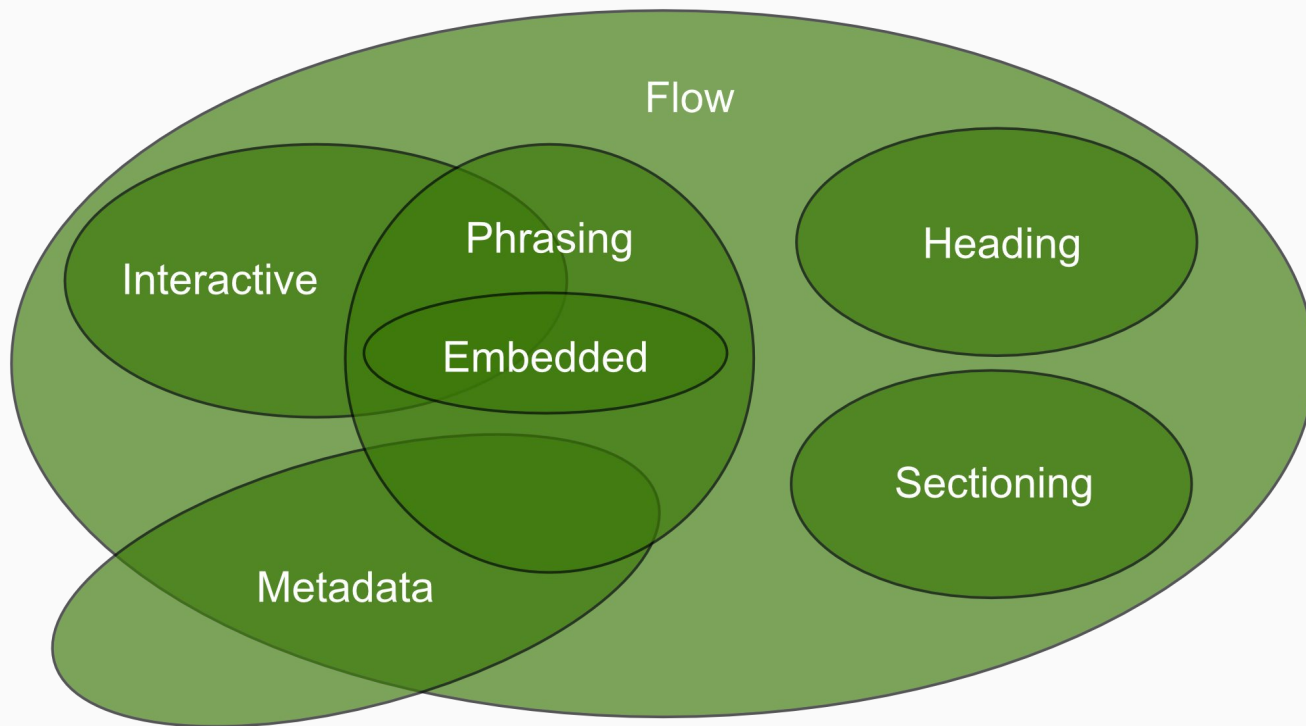
**94%** of sites had errors

**~141** errors per homepage

Common mistakes

## 2.a Elements

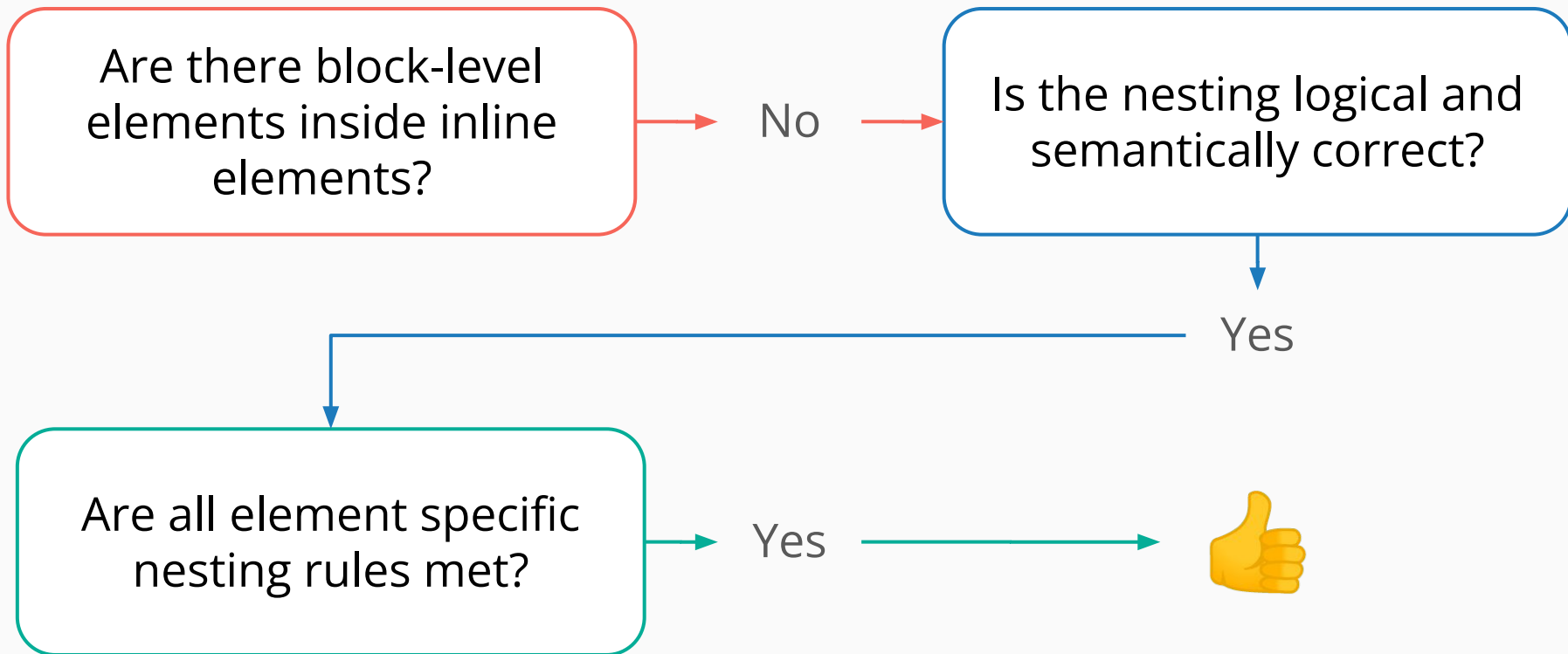
## HTML content categories



Source: <http://html.spec.whatwg.org/>



## Nesting flow content



e.g. ``<h1>`` can only contain phrasing content

## Divs do not belong in spans

✓ Inline-level element in block-level element



```
<div>
  <span>
    ✓ Good HTML
  </span>
</div>
```

✗ Block-level element in inline-level element



```
<span style="display: block;">
  <div style="display: inline;">
    ✗ Bad HTML
  </div>
</span>
```

1 in 4 sites do this

## Is it a button? Is it a link? ...No it's a ` `!

- **51% of sites** have clickable `

`/`span` elements
- To make a `

` a button we need to replicate all of these:
  - Make it focusable (`tabindex=0`)
  - Style the interactive states
  - Set `role="button"`
  - Handle *synthetic click activation*
  - Handle `disabled` state



▶ Just use button (A11ycasts) - [youtu.be/CZGqnp06DnI](https://youtu.be/CZGqnp06DnI)

**KISS** *(keep it simple, stupid)*

**Work with HTML, not against it!**



Common mistakes

## 2.b **Attributes**

## Validation errors

---

1. Invalid child elements (~67%)
2. Bad attribute values (~66%)
3. Disallowed attributes (~65%)
4. CSS errors (~47% of sites)
5. Missing required attributes (~42%)

## Bad attribute values

### ✗ Example of bad HTML

```

```

## Bad attribute values

### ✗ Example of bad HTML

Encode special characters

```

```

Invalid integer values

Missing pixel density descriptor



## Bad attribute values

✓ Example of good HTML

```

```



## Disallowed / Missing required attributes

### ✗ Example of bad HTML

```

```

## Disallowed / Missing required attributes

✗ Example of bad HTML

“Home pages with ARIA present averaged 68.6% more detected errors than those without ARIA.”

- [webaim.org/projects/million/](http://webaim.org/projects/million/)

```

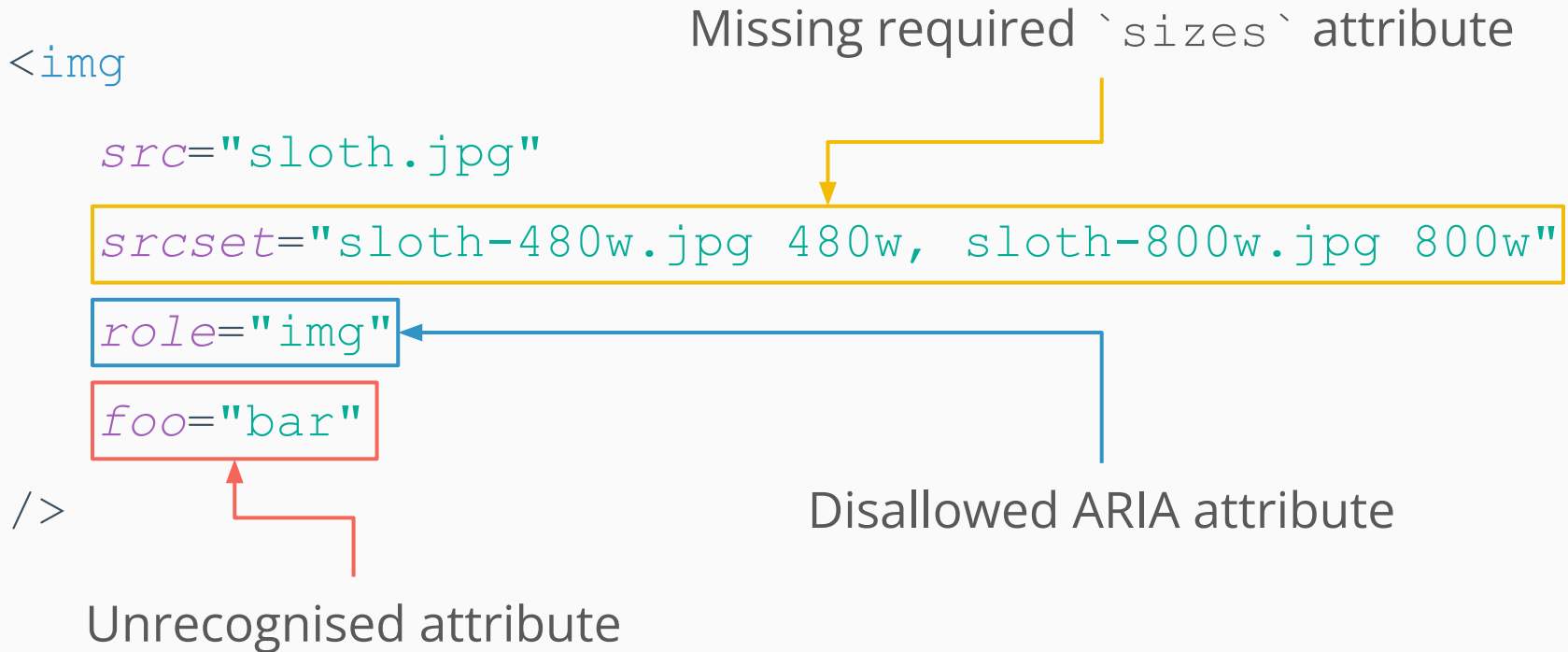
```

Disallowed ARIA attribute

Unrecognised attribute

## Disallowed / Missing required attributes

### ✗ Example of bad HTML



## Disallowed / Missing required attributes

### ✓ Example of good HTML

```

```



1 in 3 sites omit the `alt` attr

# Is HTML easy?

Maybe, but mastering HTML is difficult



## Links

- HTML specification - [html.spec.whatwg.org/](https://html.spec.whatwg.org/)
- MDN - [developer.mozilla.org/en-US/docs/Learn/HTML](https://developer.mozilla.org/en-US/docs/Learn/HTML)
- Learn HTML - [web.dev/learn/html](https://web.dev/learn/html)
- Validator - [validator.w3.org/nu](https://validator.w3.org/nu)
- The WebAIM Million - [webaim.org/projects/million](https://webaim.org/projects/million)
- HTMHell - [htmhell.dev](https://htmhell.dev)

