

pup

Command line HTML parsing tool.

- Transform a raw HTML file into a cleaned, indented, and colored format:

```
cat {{index.html}} | pup --color
```

- Filter HTML by element tag name:

```
cat {{index.html}} | pup '{{tag}}'
```

- Filter HTML by id:

```
cat {{index.html}} | pup '{{div#id}}'
```

- Filter HTML by attribute value:

```
cat {{index.html}} | pup '{{input[type="text"]}}'
```

- Print all text from the filtered HTML elements and their children:

```
cat {{index.html}} | pup '{{div}} text{'
```

- Print HTML as JSON:

```
cat {{index.html}} | pup '{{div}} json{'
```