

Data Migration - Round 2

Team TheThree - Task 2

a) HTML to RST parser

- We implemented a class `MyHTMLParser` by inheriting from the `html.parser.HTMLParser` and overridden its handler methods to carry out the desired behavior when start tags, end tags, and data are encountered.
- List of tags that our parser can handle:
 - `<body></body>`
 - `<div></div>`
 - ``
 - `<p></p>`
 - ``, ``, `<code></code>`
 - `
`, `<a>`
 - ``, ``, ``
- Example:

```
text = '<ol><li>Coffee<ul><li>Arabica</li><li>Robusta</li></ul></li><li>Tea</li><li>Milk</li></ol>'
parser = MyHTMLParser()
parser.feed(text)
parser.get_rst()

#### Output:
#
#
# #. Coffee
#
#     * Arabica
#     * Robusta
#
# #. Tea
# #. Milk
#
#
```

- Reference: [python-html2rest](#)

b) JSON to RST parser