Page 1 - With Content

This page has substantial text content.

It should be marked as non-empty by the extractor.

The word count and character count should be positive.

Page 4 - With Content

This page also has text content.

It follows two empty or near-empty pages.