

Sample Markdown Document

Test Author

2025-01-15

Introduction

This is a sample Markdown document used for testing library document ingestion. It contains multiple heading levels, code blocks, and various formatting elements to validate the stream-based extraction pipeline.

Section 1: Text Content

 Lorem ipsum dolor sit amet, consectetur adipiscing elit. Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.

 Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Subsection 1.1: Lists

- First item in unordered list
 - Second item with **bold text**
 - Third item with *italic text*
 - Fourth item with `inline code`
1. First ordered item
 2. Second ordered item
 3. Third ordered item

Subsection 1.2: Code Blocks

```
def fibonacci(n: int) -> int:  
    """Calculate the nth Fibonacci number."""  
    if n <= 1:  
        return n  
    return fibonacci(n - 1) + fibonacci(n - 2)
```

```

# Example usage
for i in range(10):
    print(f"F({i}) = {fibonacci(i)}")

fn main() {
    let numbers: Vec<i32> = (1..=10).collect();
    let sum: i32 = numbers.iter().sum();
    println!("Sum of 1..10 = {}", sum);
}

```

Section 2: Tables and Links

Format	Family	Extension
PDF	Page-based	.pdf
Markdown	Stream-based	.md
EPUB	Stream-based	.epub
DOCX	Page-based	.docx

For more information, see the project documentation.

Section 3: Block Quotes and Emphasis

“The best way to predict the future is to invent it.” – Alan Kay

This section tests **bold**, *italic*, **bold italic**, and ~~strikethrough~~ formatting.

Conclusion

This document serves as a comprehensive test fixture for the stream-based document extraction and chunking pipeline. It validates that heading hierarchy, code blocks, tables, and inline formatting are correctly preserved through the ingestion process.