

Text-Based PDF Test Document

This is a standard text-based PDF document with an embedded text layer.
It contains regular text that can be extracted without OCR.
The content is rendered with proper font encoding.

Multiple paragraphs help test text processing capabilities.
Each paragraph is positioned with proper spacing and formatting.

Testing different content: numbers 12345, symbols @\$%^&*().
This document simulates a typical born-digital text page.

The text layer is embedded in the PDF structure, not as an image overlay.
This is the standard format for documents created digitally.

Additional content to reach adequate text length for testing.
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.
This approach ensures reliable text extraction.