Canvas/DOMMatrix Build Error Fix

Issue

Build failing during page data collection with canvas-related errors:

```
Warning: Cannot polyfill `DOMMatrix`, rendering may be broken.
ReferenceError: DOMMatrix is not defined

Error: Failed to collect page data for /api/documents/[id]/extract
```

Root Cause

The `canvas` package (used by `pdf-parse`) was being imported during the Next.js build phase. Canvas requires browser APIs (DOMMatrix, ImageData, Path2D) that don't exist in Node.js.

Solution

1. Dynamic Imports

Changed all pdf-parse and tesseract.js imports to dynamic imports:

```
// Before
import pdf from 'pdf-parse';

// After
const pdf = await import('pdf-parse');
const pdfParse = pdf.default || pdf;
```

2. Webpack Externals

Updated next.config.js to externalize canvas:

```
webpack: (config, { isServer }) => {
  if (isServer) {
    config.externals.push({
      canvas: 'commonjs canvas',
      'pdf-parse': 'commonjs pdf-parse',
    });
  }
  return config;
}
```

Files Changed

1. src/lib/documents/pdf-extractor.ts - Dynamic imports
2. src/lib/documents/ocr.ts - Dynamic imports
3. src/lib/documents/extraction.ts - Dynamic imports
4. src/app/api/documents/[id]/extract/route.ts - Dynamic imports

5. next.config.js - Webpack externals

Testing

✅ Local build completed successfully
✅ No canvas/DOMMatrix errors
✅ Page data collection passed

Deployment

Ready for production deployment to Render.

Status: ✅ Fixed and Tested
Date: December 20, 2025