

# PDF Processing Fix Summary

---

## Problem

---

PDF bank statements were showing “FAILED” status after upload. The errors logged were:

- “Processing API call failed”
- “PdfParse is not a function”

## Root Cause

---

The upload endpoint was trying to trigger processing via an internal fetch call to `/api/bank-statements/process`, which was failing due to:

1. Potential network/internal routing issues
2. Missing authentication context in internal API calls
3. Circular dependency concerns

## Solution Implemented

---

### 1. Created Dedicated Statement Processor Module

**File:** `/home/ubuntu/cfo_budgeting_app/app/lib/statement-processor.ts`

- Extracted all processing logic into a dedicated module
- Implements complete PDF and CSV processing pipeline:
  - File download from S3
  - PDF data extraction using AI (Abacus AI LLM)
  - Transaction categorization
  - Financial insights generation
  - Transaction creation in database
  - Financial metrics updates
  - User notifications

### 2. Updated Upload Route

**File:** `/home/ubuntu/cfo_budgeting_app/app/app/api/bank-statements/upload/route.ts`

Changed from:

```
// OLD: Using fetch to internal API (unreliable)
const response = await fetch(`${process.env.NEXTAUTH_URL}/api/bank-statements/process`, {
  method: 'POST',
  headers: { 'Content-Type': 'application/json' },
  body: JSON.stringify({ statementId })
});
```







To:

```
// NEW: Direct module import (reliable)
setImmediate(async () => {
  const { processStatement } = await import('@lib/statement-processor');
  await processStatement(statementId);
});
```

### 3. Enhanced Error Handling

- Added detailed console logging at each processing stage
- Improved error messages
- Database updates track exact failure points
- Non-critical operations (like metrics update) don't block processing

### 4. Processing Pipeline

Upload  S3 Storage  Background Processing  AI Extraction   
Transaction Categorization  Database Creation  Completion

## Key Improvements

1. **Reliability:** No dependency on internal API calls
2. **Performance:** Async background processing doesn't block uploads
3. **Debugging:** Comprehensive logging throughout the pipeline
4. **Error Recovery:** Graceful handling of failures at each stage

## Testing Instructions

### Option 1: Through the UI (Recommended)

1. Log in with credentials: khoustons721@gmail.com / n2vGWELvGMSd2m
2. Navigate to Data Import → Bank Statements
3. Upload a PDF bank statement
4. Wait 30-60 seconds for processing
5. Refresh the page to see results

### Option 2: Check Existing Statements

The two previously failed statements are now in PENDING status:

- Personal Statement\_Sep\_11\_2025.pdf
- Business Statement\_Jan\_8\_2024.pdf

These will be automatically processed when:

- A new file is uploaded
- The bank statements page is loaded
- The system triggers background processing

## Technical Details

### AI Processing

The system uses Abacus AI's LLM API to:

1. Extract transaction data from PDFs

2. Identify bank information, account details, balances
3. Categorize transactions into meaningful categories
4. Generate financial insights and recommendations

## Database Schema




BankStatement:

- status: PENDING → PROCESSING → COMPLETED (or FAILED)
- processingStage: UPLOADED → EXTRACTING\_DATA → CATEGORIZING\_TRANSACTIONS → ANALYZING\_PATTERNS → DISTRIBUTING\_DATA → COMPLETED

## Files Changed

1. `/home/ubuntu/cfo_budgeting_app/app/lib/statement-processor.ts` - NEW
2. `/home/ubuntu/cfo_budgeting_app/app/app/api/bank-statements/upload/route.ts` - UPDATED

## Environment Variables Required

- `ABACUSAI_API_KEY` -  Already configured
- `DATABASE_URL` -  Already configured
- `AWS_BUCKET_NAME` -  Already configured (via S3 setup)

## Checkpoint Saved

 Checkpoint: "Fixed PDF processing functionality"

## Next Steps for User

1. Test by uploading a new PDF through the UI
2. Verify transactions are extracted and categorized
3. Check that status shows COMPLETED (not FAILED)
4. Review extracted transactions in the Transactions page

## Support

If issues persist:

- Check server logs for detailed error messages
- Verify the PDF is a valid bank statement format
- Ensure `ABACUSAI_API_KEY` has sufficient credits
- Check network connectivity to Abacus AI API