

Tóm tắt: Khắc phục lỗi “Không đọc được PDF”

Vấn đề

Người dùng báo cáo vấn đề “không đọc được PDF” khi click vào nút “Xem PDF Flipbook” cho số 05/2025:

- Màn hình hiển thị **đen/trống**
- Thông báo lỗi: “**Error - Failed to load PDF document.**”
- PDF viewer không hiển thị nội dung

Chẩn đoán

Bước 1: Kiểm tra trình duyệt

- Mở trang `/issues/[id]/viewer` cho số 05/2025
- Quan sát lỗi: màn hình đen với thông báo lỗi
- Mở DevTools Console: không có lỗi JavaScript cụ thể
- Mở DevTools Network:
- Request tới `issue-05-2025.pdf` trả về **304 Not Modified** (thành công)
- **KHÔNG** có request nào tới `pdf.worker.min.js`

Bước 2: Kiểm tra file PDF


```
# Kiểm tra loại file
file public/issues/issue-05-2025.pdf
# Output: data (không phải PDF document)

# Kiểm tra magic bytes
head -c 100 public/issues/issue-05-2025.pdf | od -c
# Output: \0 \0 \0 \0 ... (toàn null bytes)
```

Bước 3: So sánh với các file khác

```
# Kiểm tra tất cả PDF files
for f in public/issues/*.pdf; do
  head -c 20 "$f" | od -An -tx1
done
```

Kết quả:

-  `issue-01-2025.pdf` : 25 50 44 46 (%PDF-) - **HỢP LỆ**
-  `issue-02-2025.pdf` : 25 50 44 46 (%PDF-) - **HỢP LỆ**
-  `issue-03-2025.pdf` : 25 50 44 46 (%PDF-) - **HỢP LỆ**
-  `issue-04-2025.pdf` : 25 50 44 46 (%PDF-) - **HỢP LỆ**
-  `issue-05-2025.pdf` : 00 00 00 00 (null bytes) - **BỊ CORRUPT**



Nguyên nhân

File `public/issues/issue-05-2025.pdf` **bị corrupt** - chứa toàn null bytes (`\0`) thay vì dữ liệu PDF hợp lệ.

Lý do có thể:

1. File bị ghi đè trong quá trình copy/move
2. Lỗi khi tạo file ban đầu
3. Lỗi trong quá trình upload/seed



Giải pháp

Thay thế file corrupt bằng file hợp lệ:

```
# Tìm file PDF hợp lệ trong Uploads
ls -lh Uploads/*.pdf
# Tìm thấy: So05.2025.pdf (7.5M)

# Kiểm tra file hợp lệ
head -c 20 Uploads/So05.2025.pdf | od -An -tx1
# Output: 25 50 44 46 (bắt đầu bằng %PDF-) ✓

# Thay thế file corrupt
cp Uploads/So05.2025.pdf tapchi-hcqs/nextjs_space/public/issues/issue-05-2025.pdf
```

Kết quả sau khi thay thế:



PDF hiển thị hoàn hảo:

- Trang bìa và mục lục hiển thị rõ ràng
- Toolbar PDF hoạt động (download, print, zoom)
- Scroll qua các trang mượt mà
- Không còn màn hình đen



Kiểm tra toàn bộ hệ thống

Issues PDFs:

- ✓ issue-01-2025.pdf: 3.0M - Hợp lệ
- ✓ issue-02-2025.pdf: 14M - Hợp lệ
- ✓ issue-03-2025.pdf: 3.2M - Hợp lệ
- ✓ issue-04-2025.pdf: 5.1M - Hợp lệ
- ✓ issue-05-2025.pdf: 7.5M - Hợp lệ (đã thay thế)

Article PDFs:

- ✓ article-1.pdf: 617 bytes - Placeholder hợp lệ
- ✓ article-2.pdf: 617 bytes - Placeholder hợp lệ
- ✓ article-3.pdf: 617 bytes - Placeholder hợp lệ



Tài liệu tham khảo

Người dùng đã cung cấp tư vấn chi tiết về các nguyên nhân phổ biến:

Các lỗi thường gặp với PDF viewer:

1. **✗ CORS** - Trình duyệt chặn cross-origin requests
2. **✗ Đường dẫn sai** - File không tồn tại hoặc path không đúng
3. **✗ Mixed Content** - HTTPS/HTTP không khớp
4. **✗ WebGL lỗi** - Canvas rendering không hoạt động
5. **✓ File corrupt** - Đây là nguyên nhân trong trường hợp này

Giải pháp áp dụng:

- **Đơn giản và hiệu quả:** Sử dụng `<iframe>` với browser's built-in PDF viewer
- **Không cần PDF.js:** Tránh phức tạp với worker scripts
- **Fallback options:** Cung cấp nút "Mở tab mới" và "Tải về" cho người dùng

Cấu trúc code hiện tại

Trang viewer (`app/(public)/issues/[id]/viewer/page.tsx`):

```
// Generate PDF URL
const pdfUrl = `/issues/issue-${String(issue.number).padStart(2, '0')}-${issue.year}.pdf`;

// Simple iframe viewer
<iframe
  src={` ${pdfUrl}#toolbar=1`}
  className="w-full h-full border-0"
  title={`Số ${issue.number} (${issue.year})`}
/>
```

Ưu điểm:

- **✓** Đơn giản, không dependency phức tạp
- **✓** Sử dụng trình đọc PDF tích hợp của trình duyệt
- **✓** Hỗ trợ tất cả browser hiện đại
- **✓** Toolbar mặc định (download, print, zoom)
- **✓** Không có vấn đề CORS hay worker scripts

Bài học kinh nghiệm

1. **Luôn kiểm tra file integrity** trước khi upload/deploy
2. **Sử dụng magic bytes** để validate PDF files:


```
bash
# PDF hợp lệ phải bắt đầu bằng: 25 50 44 46 (%PDF-)
head -c 4 file.pdf | od -An -tx1
```
3. **Giữ code đơn giản:** Iframe với built-in viewer tốt hơn PDF.js cho use case cơ bản
4. **Cung cấp fallback:** Luôn có nút "Tải về" và "Mở tab mới" cho người dùng

✓ Trạng thái hiện tại

- **✓** Vấn đề đã được khắc phục hoàn toàn
- **✓** Tất cả PDF files đều hợp lệ
- **✓** PDF viewer hoạt động mượt mà

- ☒ Không cần thay đổi code
- ☒ Sẵn sàng deploy

Các bước tiếp theo

1. ☒ Thay thế file corrupt (đã hoàn thành)
 2. ⌚ Run tests để đảm bảo không có regression
 3. ⌚ Save checkpoint
 4. ⌚ Deploy to production
-

Tác giả: DeepAgent Assistant

Ngày: 2025-11-13

Trạng thái: ☒ Resolved