PipeWrench MVP - Issues Analysis Report

Generated: 2025-10-31
Repository: https://github.com/rmkenv/pipewrench_mvp

Executive Summary

After analyzing the codebase, I've identified **15 critical issues** that need to be addressed for production readiness.

Critical Issues Found

1. CRITICAL BUG: Missing HTML Template

- **File:** `api/main.py:58`
- **Issue:** HTML_TEMPLATE is a placeholder string instead of actual HTML content
- **Impact:** The home route (`/`) returns invalid HTML, breaking the entire frontend
- **Fix:** Load the actual `index.html` file content

2. Missing Environment Variable Validation

- **File:** `api/main.py:49`
- **Issue:** No validation if ANTHROPIC_API_KEY is present or valid
- **Impact:** Application crashes with unclear error messages
- **Fix:** Add startup validation and clear error messages

3. Missing CORS Configuration

- **File:** `api/main.py`
- **Issue:** No CORS middleware configured
- **Impact:** Frontend requests from different origins may fail
- **Fix:** Add FastAPI CORS middleware

4. No Logging Implementation

- **File:** All files
- **Issue:** No structured logging for debugging production issues
- **Impact:** Difficult to debug production issues
- **Fix:** Add proper logging with log levels

5. No Rate Limiting

- **File:** `api/main.py`
- **Issue:** API endpoints have no rate limiting
- **Impact:** Vulnerable to abuse and high API costs
- **Fix:** Add rate limiting middleware

6. Session Storage is In-Memory

- **File:** `api/main.py:52`
- **Issue:** Sessions stored in memory will be lost on serverless restarts

- **Impact:** Data loss on every deployment
- **Fix:** Document the limitation clearly

7. No Session Cleanup

- **File:** `api/main.py:52`
- **Issue:** No mechanism to clean up expired sessions
- **Impact:** Memory leak, eventual server crash
- **Fix:** Add session expiration and cleanup

8. Missing Input Validation

- **File:** Multiple endpoints
- **Issue:** No validation on session IDs format
- **Impact:** Potential security issues and crashes
- **Fix:** Add Pydantic models for request validation

9. Hardcoded Model Version

- **File:** `api/main.py:187, 232`
- **Issue:** Claude model version hardcoded
- **Impact:** No flexibility to switch models
- **Fix:** Move to environment variable or config

10. No Error Recovery

- **File:** Multiple endpoints
- **Issue:** Generic exception handling with exposed error details
- **Impact:** Information leakage and poor user experience
- **Fix:** Add specific error handling with user-friendly messages

11. Missing API Health Check

- **File:** `api/main.py`
- **Issue:** No health check endpoint for Vercel monitoring
- **Impact:** Cannot monitor application health
- **Fix:** Add `/health` endpoint

12. No Request Size Limits

- **File:** `api/main.py`
- **Issue:** No global request size limits beyond file upload
- **Impact:** Potential DOS attacks
- **Fix:** Add FastAPI middleware for request size limits

13. Missing Tests

- **File:** Project root
- **Issue:** No test suite
- **Impact:** No confidence in code changes
- **Fix:** Add pytest with unit and integration tests

14. No .env.example File

- **File:** Project root
- **Issue:** No example environment file

- **Impact:** Developers don't know what env vars are needed
- **Fix:** Add .env.example

15. Missing .gitignore

- **File:** Project root
- **Issue:** No .gitignore file
- **Impact:** Risk of committing sensitive data
- **Fix:** Add comprehensive .gitignore

Security Issues

High Priority

- ✓ No critical security vulnerabilities (Bandit scan passed)
- ❌ No input sanitization for HTML injection in reports

Medium Priority

- ❌ No request authentication (relies on client-side API keys)
- ❌ Session IDs are UUIDs without additional security

Dependency Issues

Current Dependencies

```
fastapi - ✓ (no version pinning)
uvicorn[standard] - ✓ (no version pinning)
anthropic - ✓ (no version pinning)
python-multipart - ✓ (no version pinning)
PyPDF2 - ✓ (no version pinning)
python-docx - ✓ (no version pinning)
jinja2 - ✓ (not used in code!)
pycryptodome - ✓ (not used in code!)
```

Issues

- No version pinning (can cause deployment failures)
- Unused dependencies (jinja2, pycryptodome)
- Missing dependencies for production:
- python-dotenv (env variable management)
- pytest (testing)
- pytest-asyncio (async testing)
- httpx (API testing)

Production Readiness Checklist

- [ ] Environment variable validation
- [ ] Proper error handling and logging
- [ ] CORS configuration
- [ ] Health check endpoint
- [ ] Input validation with Pydantic

- [ ] Session cleanup mechanism
- [ ] Comprehensive tests
- [ ] Documentation updates
- [ ] .env.example file
- [ ] .gitignore file
- [ ] Dependency version pinning
- [ ] Remove unused dependencies
- [ ] Security headers
- [ ] Request size limits