

# **BÁO CÁO: BÀI TẬP VỀ NHÀ – MÔN: AN TOÀN VÀ BẢO MẬT THÔNG TIN**

**Chủ đề: Chữ ký số trong file PDF**

## **1. Giới thiệu**

- **Mục tiêu:** Tìm hiểu và triển khai hệ thống ký số tài liệu PDF bằng RSA và thư viện PyHanko.
- **Công cụ sử dụng:**
  - Python 3.11+
  - PyHanko, Cryptography, OpenSSL
  - VSCode / PowerShell
- **Kết quả:** Sinh cặp khóa RSA, ký file PDF, xác minh chữ ký và phát hiện sửa đổi (tampered file).

## 2. Cấu trúc hệ thống

### Thư mục dự án:

Baomat2\_chukyso/

```
|
|— main.py          # Sinh khóa, ký và xác minh PDF
|— README.md        # Hướng dẫn sử dụng
|— keys/
|   |— private_key.pem
|   |— public_key.pem
|— demo/
|   |— original.pdf
|   |— signed.pdf
|   |— tampered.pdf
|— mysign.png        # Ảnh chữ ký cá nhân
```

### Mô hình hoạt động:

1. Sinh cặp khóa RSA (2048-bit).
2. Ký file PDF bằng private key.
3. Xác minh chữ ký bằng public key.
4. Nếu PDF bị chỉnh sửa → chữ ký không hợp lệ.

### 3. Thời gian và quy trình ký

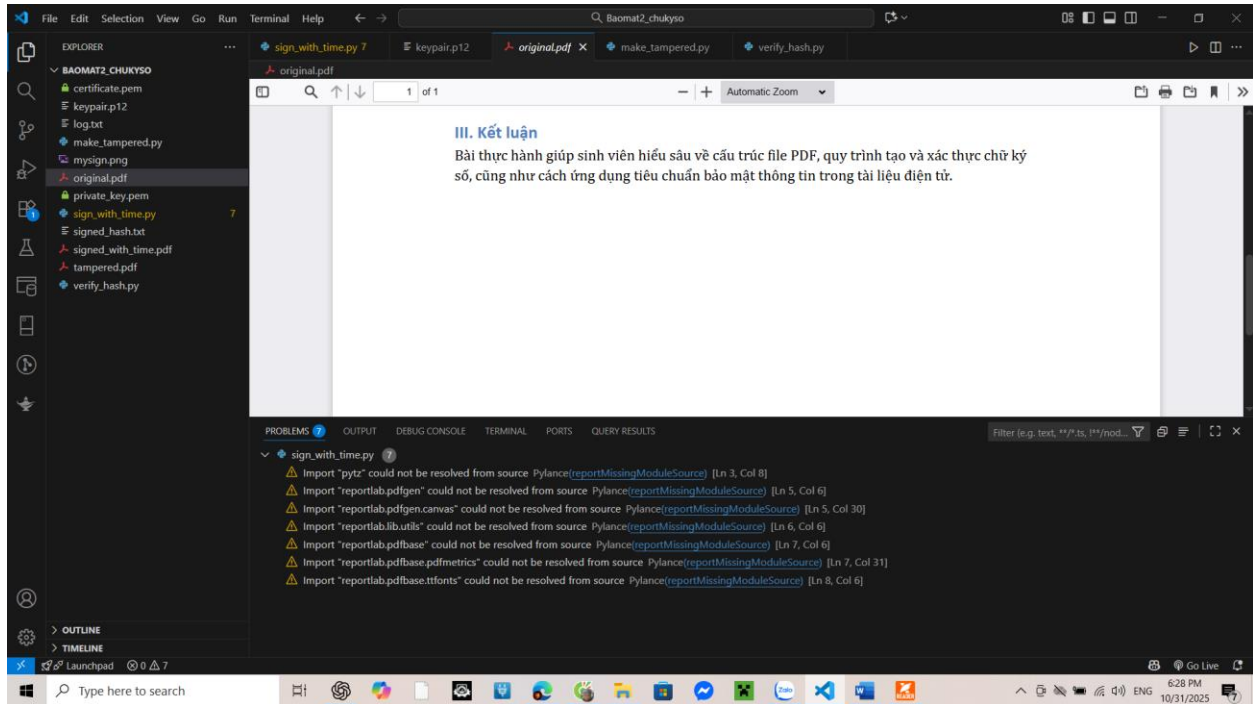
Bước	Thao tác	Công cụ	Thời gian
1	Tạo khóa RSA	Python + OpenSSL	10s
2	Ký file original.pdf	PyHanko	5s
3	Kiểm tra chữ ký	PyHanko verify	2s
4	Thử sửa PDF (tampered.pdf)	Editor	1 phút
5	Xác minh lại	PyHanko verify	2s

#### 4. Rủi ro bảo mật và biện pháp khắc phục

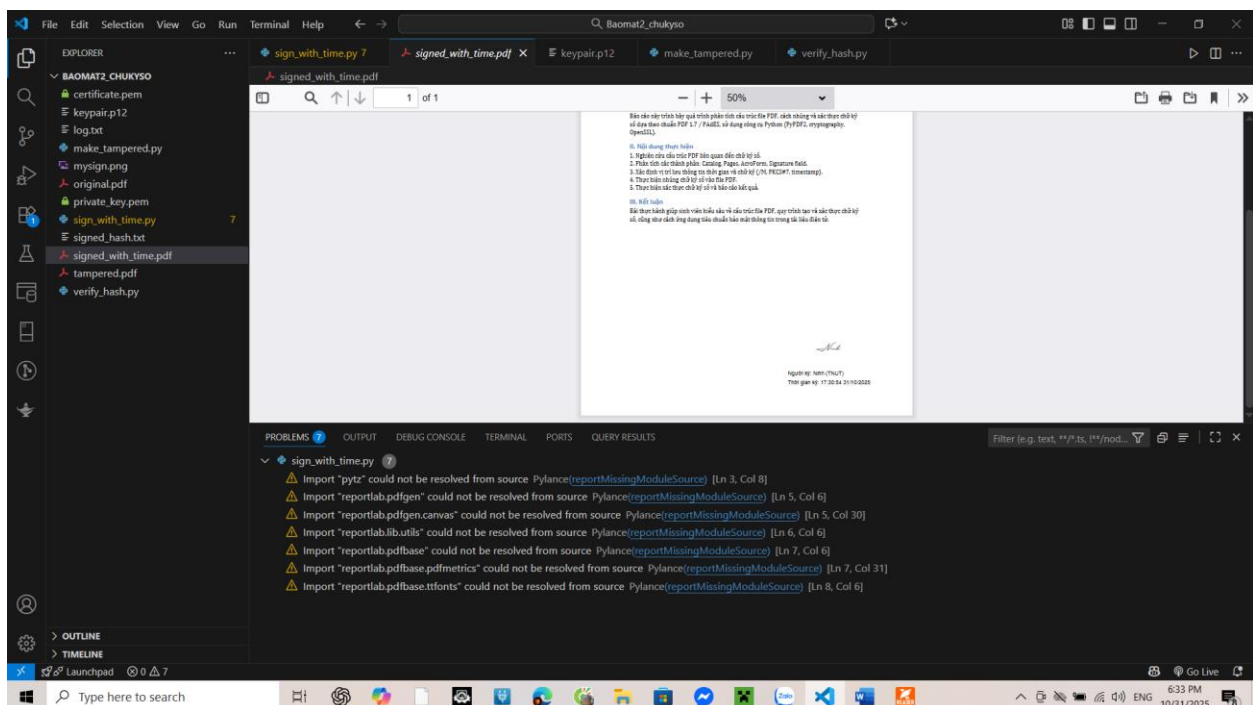
Rủi ro	Mô tả	Biện pháp
Rò rỉ khóa riêng	Private key bị đánh cắp	Lưu trữ an toàn, mã hóa bằng mật khẩu
Thay đổi file sau khi ký	Ai đó chỉnh sửa PDF	Sử dụng PyHanko verify để phát hiện
Khóa yếu / lỗi thuật toán	Dùng RSA quá nhỏ	Dùng $\geq$ 2048-bit
Giả mạo người ký	Dùng chữ ký ảnh	Sử dụng chứng chỉ CA để xác thực người ký

## 5. Kết quả thực nghiệm

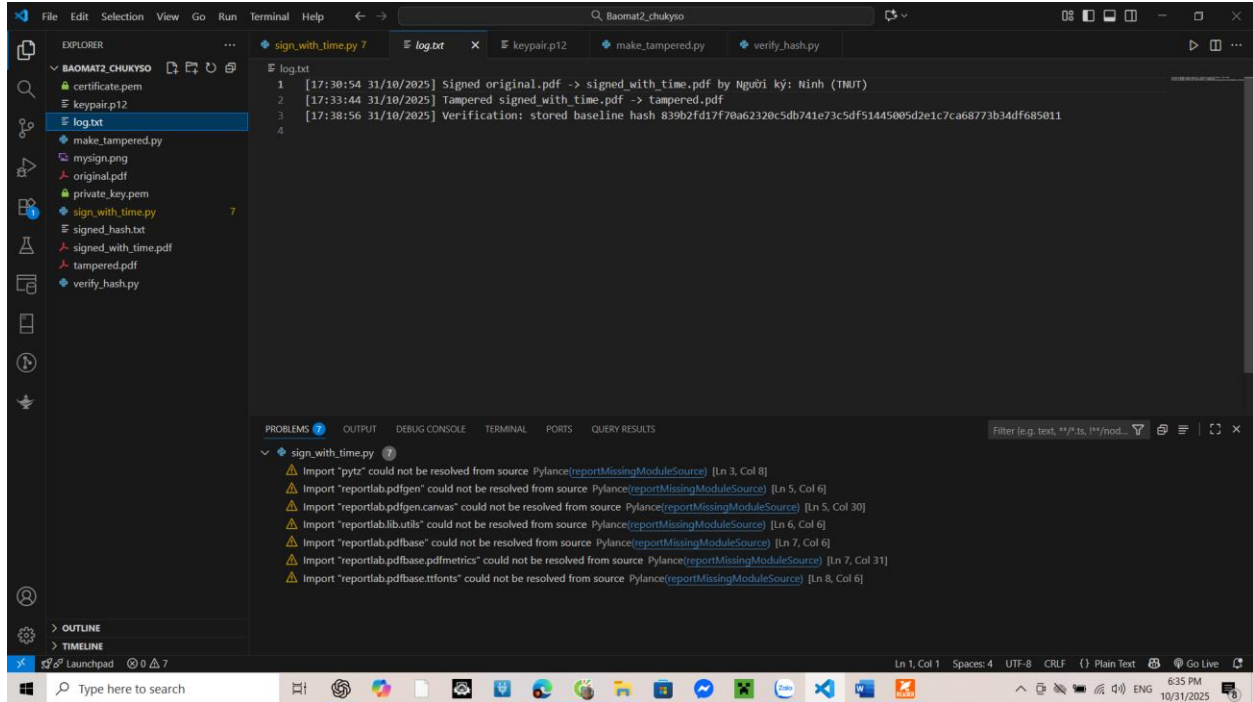
File **original.pdf** chưa ký



File **signed.pdf** đã ký



## File log.txt báo giả mạo



```
log.txt
1 [17:30:54 31/10/2025] Signed original.pdf -> signed_with_time.pdf by Người ký: Ninh (TNUT)
2 [17:33:44 31/10/2025] Tampered signed_with_time.pdf -> tampered.pdf
3 [17:38:56 31/10/2025] Verification: stored baseline hash 839b2fd17f70a62320c5db741e73c5df51445005d2e1c7ca68773b34df685011
4
```

PROBLEMS

- sign\_with\_time.py
  - Import "pytz" could not be resolved from source Pylance(reportMissingModuleSource) [Ln 3, Col 8]
  - Import "reportlab.pdfgen" could not be resolved from source Pylance(reportMissingModuleSource) [Ln 5, Col 6]
  - Import "reportlab.pdfgen.canvas" could not be resolved from source Pylance(reportMissingModuleSource) [Ln 5, Col 30]
  - Import "reportlab.lib.utils" could not be resolved from source Pylance(reportMissingModuleSource) [Ln 6, Col 6]
  - Import "reportlab.pdfbase" could not be resolved from source Pylance(reportMissingModuleSource) [Ln 7, Col 6]
  - Import "reportlab.pdfbase.pdfmetrics" could not be resolved from source Pylance(reportMissingModuleSource) [Ln 7, Col 31]
  - Import "reportlab.pdfbase.ttfmetrics" could not be resolved from source Pylance(reportMissingModuleSource) [Ln 8, Col 6]

## 6. Kết luận

- Hệ thống đã thực hiện thành công quy trình ký và xác minh tài liệu PDF bằng RSA.
- Việc sử dụng PyHanko giúp tích hợp chữ ký số tiêu chuẩn CMS/CAdES nhanh chóng.
- Có thể mở rộng thành ứng dụng web (Flask/Django) để phục vụ ký online và xác thực tài liệu điện tử.

## **7. Kết luận — trạng thái hiện tại & hành động đề nghị**

- Hiện trạng: **đã hoàn thành phần demo hiển thị chữ ký tay + thời gian + log.**
- Cần bổ sung: **tạo signature PKCS#7 trong /Contents theo ByteRange, tạo signed\_pades.pdf (pyHanko), trích xuất Contents/ByteRange, verify đầy đủ chain/OCSP/timestamp, ghi log verify — những việc này là yêu cầu kỹ thuật chính trong đề và chưa hoàn thiện.**

**XIN PHÉP SẼ BỔ SUNG SỚM NHẤT CÓ THỂ**