# QUY TRÌNH TẠO VÀ LƯU CHỮ KÝ SỐ TRONG FILE PDF

## 1. Chuẩn bị file PDF gốc

Chọn file PDF cần ký, ví dụ: original.pdf. File này có thể chứa văn bản, hình ảnh hoặc form. Mục tiêu là thêm trường chữ ký (Signature Field) mà không làm thay đổi nội dung gốc.

## 2. Tạo trường chữ ký (AcroForm / Signature Field)

Tạo AcroForm Dictionary trong file PDF nếu chưa có, thêm Signature Field (/SigField) để chứa chữ ký. Reserve vùng /Contents (8192 bytes) để chèn chữ ký nhị phân.

## 3. Xác định vùng /ByteRange

ByteRange là danh sách 4 số [start1, length1, start2, length2], chỉ định vùng dữ liệu được hash. Vùng /Contents được loại trừ để có thể chèn chữ ký sau.

## 4. Tính hash trên vùng ByteRange

Đọc hai vùng dữ liệu theo ByteRange, nối lại và tính giá trị băm SHA-256 hoặc SHA-512.

## 5. Tạo cấu trúc PKCS#7 (CMS Detached Signature)

Tạo chữ ký theo chuẩn PKCS#7 gồm các thành phần: messageDigest, signingTime, contentType, certificate chain, signerInfo và tùy chọn timestamp RFC3161.

## 6. Chèn chữ ký vào vùng /Contents

Sau khi có chữ ký PKCS#7 dạng DER, chèn vào /Contents dưới dạng hex. Chiều dài phải khớp với vùng dự trữ (8192 bytes).

## 7. Ghi Incremental Update

Thực hiện lưu file PDF theo dạng incremental (không ghi đè), giúp bảo toàn nội dung và hỗ trợ nhiều chữ ký.

## 8. Cập nhật DSS (Document Security Store - LTV)

Thêm trường /DSS vào Catalog chứa Certs, OCSP, CRL, VRI để đảm bảo xác thực lâu dài (Long-Term Validation).

## 9. Thông số kỹ thuật

Hash algorithm: SHA-256  
RSA key size: 2048 bits  
Padding: PKCS#1 v1.5  
Signature type: CMS detached (CAdES-BES)  
Encoding: DER (hex trong /Contents).

## 10. Kết quả đầu ra

Mã nguồn Python ký PDF (pyHanko, pikepdf, cryptography).  
File PDF gốc: original.pdf  
File PDF đã ký: signed\_output.pdf  
Có thể xác thực bằng Adobe Acrobat Reader hoặc script verify\_pdf\_signature.py.

## Sơ đồ quy trình

original.pdf → Thêm AcroForm → Tính hash → Tạo PKCS#7 → Ghi vào /Contents → Lưu incremental update → Thêm DSS → signed\_output.pdf

## Kết luận

Quy trình tuân thủ chuẩn ETSI EN 319 142-1 (CAdES) và ISO 32000-2 (PDF 2.0), hỗ trợ mở rộng timestamp, OCSP, CRL để tăng tính pháp lý.