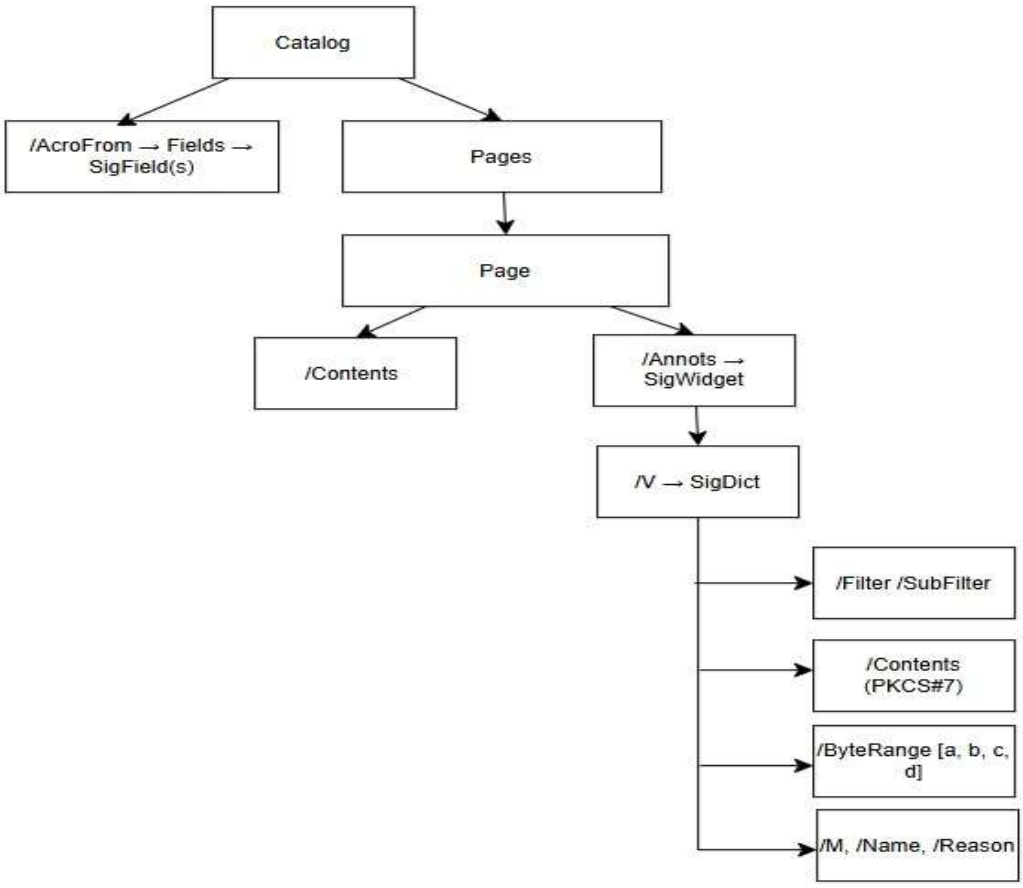
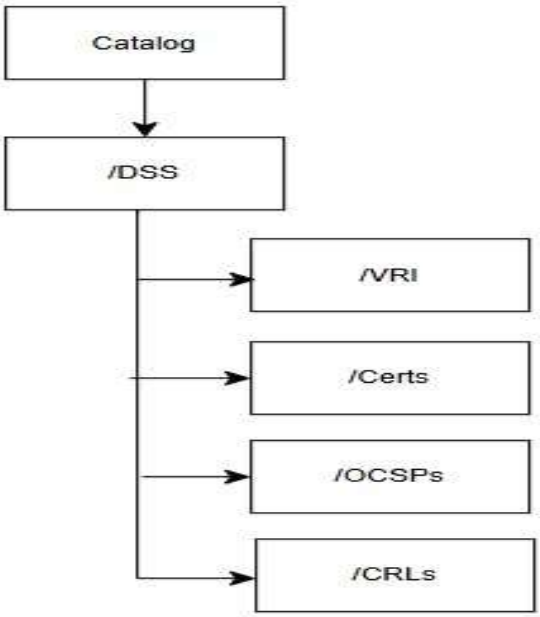
BÁO CÁO MÔN: AN TOÀN VÀ BẢO MẬT THÔNG TIN

1. Cấu trúc PDF liên quan chữ ký

Tóm tắt lý thuyết:

Thành phần	Vai trò
Catalog (/Root)	Gốc của toàn bộ cấu trúc PDF, trỏ tới /Pages và /AcroForm
Pages tree	Danh sách các trang PDF
Page object	Mỗi trang có /Contents (dòng lệnh vẽ, text)
Resources	Phông chữ, ảnh, XObject
AcroForm	Biểu mẫu chứa các trường tương tác (form field)
Signature field (Widget)	Field trong AcroForm có loại /Sig
Signature dictionary (/Sig)	Chứa dữ liệu chữ ký số, thuộc tính /Contents, /ByteRange, /M…
/ByteRange	Vùng byte của file được hash và ký
/Contents	Dữ liệu chữ ký PKCS#7 (dạng hex hoặc binary)
Incremental update	Lần cập nhật sau cùng để thêm chữ ký mà không ghi đè file gốc
DSS (Document Security Store)	Theo PAdES, chứa Certs, OCSP, CRLs, Timestamp


Catalog
/AcroFrom → Fields → SigField(s)
Pages
Page
/Contents
/Annots → SigWidget
/V → SigDict
/Filter /SubFilter
/Contents (PKCS#7)
/ByteRange [a, b, c, d]
/M, /Name, /Reason


Catalog
/DSS
/VRI
/Certs
/OCSPs
/CRLs

2. Thời gian ký

2.1 /M trong Signature Dictionary

- Là thuộc tính bắt buộc của từ điển chữ ký PDF (/Type /Sig).
- Chứa **thời gian do phần mềm ký ghi lại** (thường lấy theo đồng hồ máy người ký)
- Dạng chuỗi text, **không được bảo vệ bởi chữ ký số** → có thể bị sửa
- hiển thị thời điểm ký trong trình xem PDF (Adobe, Foxit, v.v.)
- Không có giá trị pháp lý chỉ dùng để hiển thị thời điểm ký trên giao diện PDF viewer (Adobe, Foxit...)

2.2 Timestamp Token (RFC 3161) trong PKCS#7 (timeStampToken)

- Là **tem thời gian chuẩn RFC 3161** được chèn vào **trong cấu trúc PKCS#7** của chữ ký PDF (thuộc tính timeStampToken)
- Token này được **TSA – Time Stamping Authority** phát hành và ký riêng, đảm bảo rằng **chữ ký đã tồn tại tại thời điểm được ghi trong token**
- Cấu trúc token gồm:
    - Hash của dữ liệu cần đóng dấu (messageImprint)
    - Thời gian chính xác (genTime)
    - Chữ ký của TSA xác nhận
- Có giá trị pháp lý dùng để chứng minh thời điểm chữ ký tồn tại

2.3 Document Timestamp Object (PadES)

- Là một **chữ ký đặc biệt** trong PDF/PAdES dùng để **đóng dấu toàn bộ tài liệu**, không gắn với người ký cụ thể
- Được lưu với /SubFilter /ETSI.RFC3161
- Dùng để chứng minh rằng **toàn bộ tài liệu PDF** đã tồn tại tại thời điểm genTime do TSA cung cấp

2.4 DSS – Document Security Store

- Là **vùng lưu trữ bảo mật** trong PDF (theo PAdES-LTV)
- Có thể chứa:
    - Timestamp tokens
    - OCSP responses
    - CRL (Certificate Revocation List)
    - Chứng chỉ TSA, CA
- Mục đích: duy trì khả năng xác thực **lâu dài (Long-Term Validation)** cho chữ ký

2.5Khác biệt giữa thông tin thời gian /M và timestamp RFC3161

Tiêu chí	/M trong Signature Dictionary	RFC 3161 Timestamp Token
Nguồn	Ứng dụng ký tự ghi	Cơ quan TSA ký và phát hành
Kiểu dữ liệu	Chuỗi văn bản (string)	Dữ liệu mã hóa (signed token)
Độ tin cậy	Thấp, có thể giả mạo	Cao, được ký bởi TSA
Giá trị pháp lý	Không có	Có giá trị chứng minh thời gian tồn tại tài liệu
Chuẩn liên quan	PDF 1.7 / PAdES Basic	RFC 3161, PAdES-B-T trở lên

3. Các bước tạo và lưu chữ ký trong PDF (đã có private RSA)- Viết script/code thực hiện tuần tự:

- Chuẩn bị file PDF gốc.
- Tạo Signature field (AcroForm), reserve vùng /Contents (8192 bytes).
- Xác định /ByteRange (loại trừ vùng /Contents khỏi hash).
- Tính hash (SHA-256/512) trên vùng ByteRange.
- Tạo PKCS#7/CMS detached hoặc CAdES:- Include messageDigest, signingTime, contentType.- Include certificate chain.- (Tùy chọn) thêm RFC3161 timestamp token.
- Chèn blob DER PKCS#7 vào /Contents (hex/binary) đúng offset.
- Ghi incremental update.
- (LTV) Cập nhật DSS với Certs, OCSPs, CRLs, VRI.- Phải nêu rõ: hash alg, RSA padding, key size, vị trí lưu trong PKCS#7.- Đầu ra: mã