PHẦN 1: CẤU TRÚC PDF LIÊN QUAN CHỮ KÝ SỐ

1. Khái niệm chung

Trong định dạng PDF, toàn bộ tài liệu được tổ chức dưới dạng các đối tượng (object) được liên kết với nhau thành một cấu trúc cây (tree structure).
Khi người dùng thêm chữ ký số vào file PDF, hệ thống không thay đổi nội dung gốc mà chỉ thêm mới một tập các đối tượng mô tả chữ ký ở cuối file.
Cách này được gọi là incremental update – giúp phát hiện nếu file bị chỉnh sửa sau khi ký.

2. Giải thích từng đối tượng

2.1. Catalog

Catalog là đối tượng gốc (root object) của file PDF, tương tự như "mục lục tổng".
Nó trỏ đến hai thành phần chính: /Pages (cây trang chứa nội dung hiển thị) và /AcroForm (vùng chứa biểu mẫu, bao gồm chữ ký số).

2.2. Pages Tree

Là cấu trúc quản lý toàn bộ các trang trong tài liệu PDF.
Mỗi trang (Page) là một đối tượng con của /Pages.
Cây trang giúp PDF xử lý nhiều trang nhanh hơn và tái sử dụng nội dung chung.

2.3. Page Object

Là đối tượng mô tả nội dung cụ thể của một trang PDF, gồm:
/Contents (nội dung hiển thị), /Resources (font, hình ảnh), và /Annots (chú thích, vùng chữ ký).

2.4. Resources

Lưu trữ tài nguyên dùng chung như font chữ, hình ảnh, màu nền.
Các trang có thể dùng lại cùng một /Resources để giảm kích thước file.

2.5. Content Stream (/Contents)

Là luồng dữ liệu mô tả nội dung hiển thị của trang PDF, gồm các lệnh PDF (BT, ET, Tj...).
Trình đọc PDF sử dụng phần này để hiển thị nội dung lên màn hình.

2.6. XObject
Là đối tượng con có thể tái sử dụng, ví dụ hình ảnh hoặc mẫu vẽ được chèn nhiều lần trong tài liệu.

2.7. AcroForm
AcroForm là đối tượng chứa toàn bộ các form tương tác trong PDF, như ô nhập tên, checkbox, nút bấm hoặc trường ký (signature field).
Bên trong có thuộc tính /Fields: danh sách các field bao gồm chữ ký.

2.8. Signature Field (Widget Annotation)
Signature Field là một form field đặc biệt dùng để chứa chữ ký số.
Nó có kiểu /FT /Sig và được đặt trong /Annots của trang (hiển thị vị trí vùng ký) và /AcroForm /Fields (biểu mẫu chính).
Trường này trỏ tới Signature Dictionary (/Sig).

2.9. Signature Dictionary (/Sig)
Đây là nơi lưu trữ thông tin chữ ký số thực tế, gồm:
/Type /Sig, /Filter /Adobe.PPKLite, /SubFilter /adbe.pkcs7.detached, /Contents (chữ ký PKCS#7), /ByteRange, /M (thời gian ký).

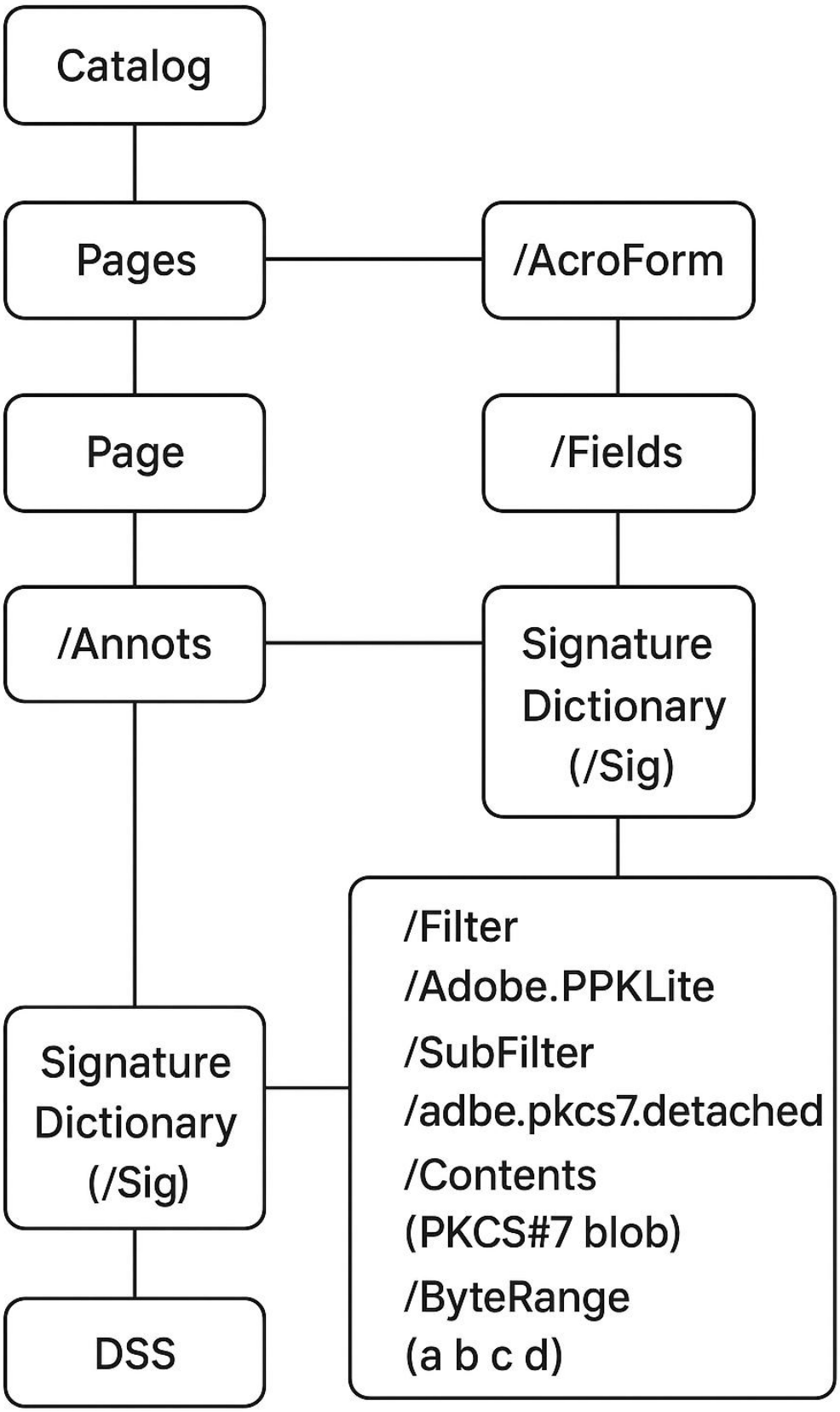
2.10. /ByteRange
Là mảng gồm 4 số quy định phạm vi dữ liệu được ký, ví dụ /ByteRange [0 12345 56789 456].
Hash được tính trên hai đoạn tách biệt, loại trừ vùng /Contents.

2.11. /Contents
Là vùng chứa chữ ký số thực tế, thường được mã hóa theo chuẩn PKCS#7 hoặc CMS, lưu dạng hex.

2.12. Incremental Update
Khi ký, PDF không bị ghi đè mà phần chữ ký được thêm mới vào cuối file như một lớp cập nhật.
Nhờ đó, dữ liệu gốc vẫn nguyên vẹn và mọi chỉnh sửa đều bị phát hiện.

2.13. DSS (Document Security Store)
Là kho lưu trữ bảo mật (theo chuẩn PAdES), chứa chứng chỉ, OCSP, CRL và timestamp token để xác thực chữ ký lâu dài (LTV).

3. Sơ đồ cấu trúc PDF chứa chữ ký số
(Chèn hình sơ đồ cấu trúc PDF tại đây, file hình: A_flowchart_diagram_in_the_digital_illustration_sh.png)


Catalog
Pages
/AcroForm
Page
/Fields
/Annots
Signature Dictionary (/Sig)
/Filter
/Adobe.PPKLite
/SubFilter
/adbe.pkcs7.detached
/Contents
(PKCS#7 blob)
/ByteRange
(a b c d)
Signature Dictionary (/Sig)
DSS

4. Kết luận phần 1

Cấu trúc chữ ký số trong PDF dựa trên mối quan hệ giữa các đối tượng Catalog → AcroForm → Signature Field → Signature Dictionary (/Sig).
/Contents chứa chữ ký số thực tế, /ByteRange xác định vùng được ký, /M ghi thời gian ký, và DSS giúp xác minh chữ ký lâu dài.
Cơ chế incremental update đảm bảo mọi chỉnh sửa sau khi ký đều có thể phát hiện, duy trì tính toàn vẹn và xác thực của tài liệu.

PHẦN 2: THỜI GIAN KÝ ĐƯỢC LƯU Ở ĐÂU

Trong chữ ký số PDF, thông tin về **thời gian ký (timestamp)** đóng vai trò quan trọng giúp xác định **thời điểm tài liệu được ký và xác minh chữ ký có còn hợp lệ hay không**.
PDF có thể lưu trữ thời gian ký tại nhiều vị trí khác nhau, tùy theo định dạng chữ ký và chuẩn bảo mật áp dụng (PDF 1.7, PDF 2.0, PAdES).

Dưới đây là các vị trí phổ biến lưu thông tin thời gian trong file PDF:

1. **Trường /M trong Signature Dictionary (/Sig)**
Đây là thuộc tính mặc định trong mỗi chữ ký số.
- Giá trị của /M là chuỗi văn bản dạng "D:YYYYMMDDHHmmSS+TZ" (ví dụ: D:20251024T114053+07'00').
- Nó chỉ mang tính **thông tin mô tả**, không có giá trị pháp lý, vì người ký có thể tự thay đổi mà không cần xác nhận từ bên thứ ba.

2. **Timestamp token (RFC 3161)**
Là dấu thời gian do **TSA (Time Stamping Authority)** phát hành.
- Token này được chèn vào trong **PKCS#7/CMS** dưới dạng thuộc tính `timeStampToken`.
- TSA ký vào token bằng khóa riêng của họ → đảm bảo thời gian là xác thực.
- Có giá trị pháp lý vì nó chứng minh được thời điểm tài liệu được ký.

3. **Document Timestamp Object (PAdES)**
Chuẩn PAdES mở rộng PDF để cho phép **gắn dấu thời gian cho toàn bộ tài liệu**, không chỉ riêng người ký.
- Object này thường được chèn vào phần cuối file, hoạt động tương tự chữ ký số nhưng không thay đổi nội dung cũ.
- Hữu ích khi muốn xác thực "tài liệu tồn tại tại thời điểm X".

4. **DSS (Document Security Store)**
Khi tài liệu PDF hỗ trợ LTV (Long-Term Validation), DSS sẽ chứa toàn bộ dữ liệu cần để xác thực lại chữ ký trong tương lai:
- Bao gồm **chứng chỉ (Certs)**, **OCSP (trạng thái chứng chỉ)**, **CRL (danh sách thu hồi)** và **timestamp token**.
- Giúp xác minh chữ ký kể cả khi chứng chỉ đã hết hạn hoặc máy chủ TSA không còn hoạt động.

So sánh giữa /M và timestamp RFC 3161

Đặc điểm	/M trong /Sig	Timestamp RFC3161
Kiểu dữ liệu	Chuỗi text (metadata)	Token PKCS#7 được TSA ký
Ai tạo ra	Phần mềm ký PDF	Cơ quan cấp dấu thời gian (TSA)
Giá trị pháp lý	Không	Có (vì có chữ ký TSA)
Có thể sửa được không?	Có	Không thể (vì token được ký)
Vị trí lưu	Trong Signature Dictionary	Trong PKCS#7/CMS (attribute timeStampToken)
Mục đích	Hiển thị thời gian ký	Chứng minh thời điểm ký là thật

Kết luận phần 2
Thông tin thời gian ký trong PDF có thể xuất hiện ở nhiều nơi, nhưng chỉ **timestamp theo chuẩn RFC 3161** mới có giá trị chứng minh hợp pháp về mặt pháp lý.
Trong khi đó, trường **/M** trong `/Sig` chỉ là thông tin hiển thị đơn giản.
Khi áp dụng chuẩn **PAdES**, việc lưu thêm **Document Timestamp Object** và **DSS** sẽ giúp đảm bảo khả năng **xác thực lâu dài (LTV)** cho chữ ký điện tử.