1) Cấu trúc PDF liên quan chữ ký (Nghiên cứu)- Mô tả ngắn gọn: Catalog, Pages tree, Page object, Resources, Content streams, XObject, AcroForm, Signature field (widget), Signature dictionary (/Sig), /ByteRange, /Contents, incremental updates, và DSS (theo PAdES).- Liệt kê object refs quan trọng và giải thích vai trò của từng object trong lưu/truy xuất chữ ký.- Đầu ra: 1 trang tóm tắt + sơ đồ object (ví dụ: Catalog → Pages → Page → /Contents ; Catalog → /AcroForm → SigField → SigDict)

- Các objects chính trong PDF:

- **1. Catalog Object** Là "gốc" (Root) của tài liệu PDF. Từ đây có thể truy cập tới toàn bộ cây trang (Pages tree) và đến AcroForm – nơi chứa các trường form (Form fields). Trong trường hợp có chữ ký, Catalog có một key `/AcroForm` trỏ đến AcroForm dictionary.

2. Pages Tree Là cấu trúc dạng cây mô tả toàn bộ các trang trong tài liệu. Mỗi Page Object nằm trong cây này.

3. Page Object Đại diện cho từng trang cụ thể trong file PDF. Nó chứa các tham chiếu tới nội dung hiển thị (Content Stream), tài nguyên (Resources), và các annotation (bao gồm chữ ký).

4. Resources Lưu thông tin tài nguyên dùng cho trang: phông chữ, hình ảnh, màu sắc, XObject (đối tượng nhúng),…

5. Content Streams Là nơi chứa các lệnh vẽ (văn bản, hình khối, hình ảnh) – nói cách khác là "nội dung nhìn thấy được" của trang.

6. XObject Là các đối tượng nhúng, ví dụ hình ảnh hoặc mẫu template. Không liên quan trực tiếp đến chữ ký, nhưng có thể tồn tại trong cùng trang.

7. AcroForm Dictionary Là thành phần quản lý toàn bộ các **form fields** trong PDF (ví dụ: textbox, checkbox, signature field...). Khi có chữ ký, phần này sẽ chứa một danh sách các Signature Fields.

8. Signature Field (Widget Annotation) Là trường chữ ký hiển thị trên trang. Nó có vị trí (tọa độ x, y), kích thước (width, height), và tham chiếu đến **Signature Dictionary** (object chứa dữ liệu chữ ký thật).

9. Signature Dictionary (/Sig) Đây là thành phần cốt lõi của chữ ký PDF. Nó lưu dữ liệu PKCS#7 (chứa chữ ký số RSA hoặc ECDSA) cùng với thông tin thời gian, tên người ký, thuật toán băm, và phạm vi dữ liệu được ký. Các key quan trọng trong dictionary này gồm:

- `/Filter` : Tên bộ lọc xử lý chữ ký (thường là `/Adobe.PPKLite`).

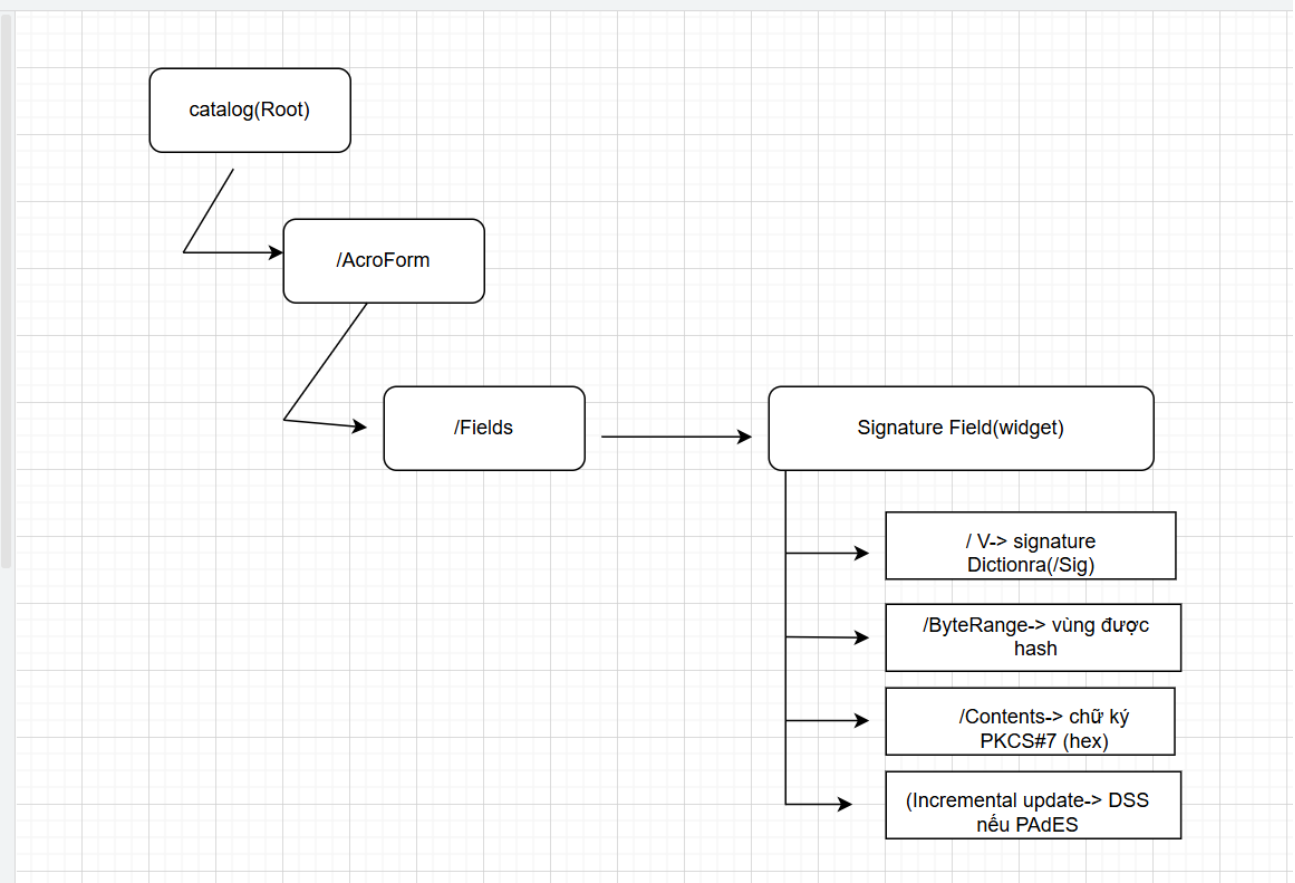
- `/SubFilter` : Định dạng chữ ký (ví dụ `/adbe.pkcs7.detached` hoặc `/ETSI.CAdES.detached`).

- `/ByteRange` : Xác định vùng byte trong file được hash và ký (loại trừ vùng `/Contents`).

- `/Contents` : Chứa chữ ký PKCS#7/CMS dưới dạng chuỗi hex hoặc binary.

10. /ByteRange	Là mảng gồm 4 giá trị (offset, length, offset, length). Nó cho biết phần nào của file được dùng để tính hash. Phần `/Contents` (chứa chữ ký) bị loại ra khỏi vùng này để tránh thay đổi dữ liệu sau khi ký.
11. /Contents	Là vùng chứa dữ liệu chữ ký số thực tế – thường là chuỗi PKCS#7 (chuẩn CMS) đã được mã hóa bằng khóa riêng của người ký.
12. Incremental Update	Khi ký, PDF không bị ghi đè mà chỉ thêm phần mới vào cuối file. Phần mới này chứa chữ ký và cập nhật lại bảng `xref` và `trailer`. Điều này cho phép xác minh toàn vẹn vì các phần cũ vẫn giữ nguyên.
13. DSS (Document Security Store)	Thành phần mở rộng theo chuẩn **PAdES (ETSI EN 319 142)**. Nó lưu trữ các dữ liệu cần thiết cho việc xác minh chữ ký lâu dài (Long-Term Validation – LTV): danh sách chứng chỉ, CRL, OCSP và timestamp.

❖ Sơ đồ quan hệ object



2) Thời gian ký được lưu ở đâu?- Nêu tất cả vị trí có thể lưu thông tin thời gian: + /M trong Signature dictionary (dạng text, không có giá trị pháp lý). + Timestamp token (RFC 3161) trong PKCS#7 (attribute timeStampToken). + Document timestamp object (PAdES). + DSS (Document Security Store) nếu có lưu timestamp và dữ liệu xác minh.- Giải thích khác biệt giữa thông tin thời gian /M và timestamp RFC

Trong tài liệu PDF có chữ ký số, thông tin thời gian ký có thể được lưu tại nhiều vị trí khác nhau, tùy thuộc vào chuẩn ký và mức độ xác thực. Các vị trí lưu thời gian gồm:

1. **Giá trị /M trong Signature Dictionary**
   - Đây là thuộc tính nằm trong từ điển chữ ký (/Sig) của PDF.
   - Dữ liệu thời gian được lưu dưới dạng chuỗi văn bản theo định dạng:

- Thời gian này do **phần mềm ký** tự ghi vào, không có ràng buộc xác thực.

- Vì vậy **không có giá trị pháp lý**, do người ký hoặc thiết bị có thể thay đổi thời gian hệ thống máy tính trước khi ký gian hệ thống máy tính trước khi ký.

2. Timestamp token (RFC 3161) bên trong chữ ký PKCS#7/CMS

- Token thời gian được nhúng bên trong dữ liệu chữ ký /Contents.
- Token này được tạo bởi **Timestamp Authority (TSA)** — một tổ chức tin cậy có chứng thư số riêng.
- Timestamp chứa:
  - Hash của dữ liệu được ký
  - Thời gian ký
  - Chữ ký số của TSA xác nhận thời gian
- Đây là bằng chứng hợp pháp về thời điểm chữ ký được tạo ra.

3. Document Timestamp (theo chuẩn PAdES)

- PDF có thể chứa một chữ ký đặc biệt dựa trên timestamp, không gắn với người ký mà gắn với tài liệu.
- Được dùng để đóng dấu thời gian cho cả tài liệu sau khi chữ ký được tạo.
- Không yêu cầu private key của người ký, chỉ cần timestamp của TSA.

4. DSS – Document Security Store

- Lưu ở phần incremental update cuối cùng của PDF.
- DSS chứa các dữ liệu hỗ trợ xác minh chữ ký lâu dài:
  - Certificate chain
  - CRL và OCSP
  - Timestamp token (nếu có)
- Hỗ trợ chế độ **LTV (Long Term Validation)**, đảm bảo tài liệu vẫn xác thực được dù chứng thư số hết hạn.

3) Các bước tạo và lưu chữ ký trong PDF (đã có private RSA) - Viết script/code thực hiện tuần tự: 1. Chuẩn bị file PDF gốc. 2. Tạo Signature field (AcroForm), reserve vùng /Contents (8192 bytes). 3. Xác định /ByteRange (loại trừ vùng /Contents khỏi hash).

SĐT: 0345678210
Ngày ký: 31/10/2025 18:57:49
Location: VIỆT NAM

TRIỆU TRÀ MY