1. **Giới thiệu khái niệm**

❖ **Chữ ký số (Digital Signature)** là một dạng chữ ký điện tử được tạo ra bằng cách sử dụng **hệ thống mật mã không đối xứng**, gồm một cặp khóa: **khóa bí mật (private key)** và **khóa công khai (public key)**

❖ Cách hoạt động của chữ ký số:

+ **Tạo chữ ký**:

- Người gửi sử dụng **khóa bí mật** để mã hóa một bản tóm tắt (hash) của tài liệu.

+ **Xác minh chữ ký**:

- Người nhận dùng **khóa công khai** của người gửi để giải mã chữ ký số.

❖ Quy trình ký số một file Quy

+ **Tạo bản tóm tắt (Hash) của nội dung PDF:**

- Phần mềm ký sẽ chọn một thuật toán băm (thường là SHA-256) để tạo ra một chuỗi hash duy nhất từ nội dung PDF.

+ **Ký chuỗi hash bằng khóa bí mật:**

- Chuỗi hash được mã hóa bằng **khóa bí mật** của người ký để tạo ra chữ ký số.

+ **Nhúng chữ ký số vào file PDF:**

- Chữ ký số được nhúng vào file PDF dưới dạng một trường đặc biệt (Signature Field).
- Ngoài chữ ký, phần mềm cũng có thể nhúng **chứng thư số** (certificate) để người nhận xác minh danh tính người ký.

❖ **Lưu file PDF đã ký:**

- File PDF sau khi ký vẫn giữ nguyên nội dung gốc, nhưng có thêm phần chữ ký số.

2. **Chuẩn tham chiếu**

- PDF version: **PDF 1.7 hoặc PDF 2.0**

 PDF 1.7: Là phiên bản được chuẩn hóa thành ISO 32000-1:2008. Đây là nền tảng phổ biến cho các hệ thống ký số hiện nay.

 PDF 2.0: Chuẩn ISO 32000-2:2017, cải tiến về bảo mật, cấu trúc và hỗ trợ tốt hơn cho các tính năng như chữ ký số, xác thực và mã hóa.

- Chuẩn ký số cho PDF: **PAdES (ETSI TS 102 778, ETSI EN 319 142-1)**.

- Chuẩn timestamp: **RFC 3161**.

3. Công cụ sử dụng để thực hiện:

- Python (thư viện pikepdf, cryptography, PyHanko)
- Hoặc OpenSSL, hoặc iText7 (Java)
- Mục tiêu: tạo chữ ký PKCS#7 (CMS) và chèn vào PDF.

1) Cấu trúc PDF:

Các thành phần chính trong cấu trúc PDF ký số

Object	Vai trò
/Catalog	**Điểm bắt đầu của cấu trúc PDF, trỏ đến /Pages và /AcroForm.**
/Pages	**Cây các trang (Pages tree).**
/Page	**Trang cụ thể, chứa /Contents (stream hiển thị).**
/Contents	**Dữ liệu nội dung trang (text, ảnh, v.v.).**
/Resources	**Định nghĩa font, ảnh dùng trong trang.**
/AcroForm	**Vùng chứa form fields, bao gồm chữ ký.**
/SigField (Widget)	**Trường chữ ký hiển thị trên trang.**
/Sig (Signature dictionary)	**Nơi chứa thông tin chữ ký.**
/ByteRange	**Phạm vi byte được ký (ngoại trừ phần /Contents).**
/Contents	**Chứa dữ liệu PKCS#7 (CMS).**
/DSS (Document Security Store)	**Dữ liệu chứng thực mở rộng (cert, CRL, OCSP, timestamp).**

🖇️ Sơ đồ cấu trúc object PDF liên quan chữ ký

```
Catalog
├── /Pages → Pages tree
│   └── /Kids → Page object(s)
│       └── /Contents → Content streams
│       └── /Resources → Fonts, XObject
├── /AcroForm
│   └── /Fields → Signature field (Widget)
│       └── /FT /Sig → Signature dictionary
│           ├── /Contents → Chữ ký số
│           ├── /ByteRange → Vùng được ký
│           ├── /Filter /SubFilter → Loại chữ ký (e.g. Adobe.PPKLite, ETSI.CAdES.det
└── DSS (nếu có)
    ├── /Certs → Chứng thư số
    ├── /OCSPs → Trạng thái chứng thư
    └── /CRLs → Danh sách thu hồi
```

2) Thời gian ký được lưu ở đâu?

a. Các vị trí chứa thông tin thời gian trong PDF:

Vị trí	Mô tả	Giá trị pháp lý
/M (trong Signature dictionary)	Chuỗi thời gian định dạng D:YYYYMMDDHHmmSS+TZ	❌ Không có giá trị pháp lý (chỉ metadata)
RFC 3161 Timestamp Token	Trong thuộc tính PKCS#7 (timeStampToken)	✅ Có giá trị pháp lý (ký bởi TSA)
Document Timestamp Object	Chữ ký thời gian dạng riêng trong PAdES	✅ Có giá trị pháp lý
/DSS	Có thể lưu timestamp + OCSP + CRL	✅ Có giá trị pháp lý (LTV)

b. Phân biệt rõ:

- /M: chỉ là metadata, có thể bị sửa sau khi ký.
- RFC3161 timestamp: chứng thực thời điểm file tồn tại, do TSA ký → có giá trị chứng minh pháp lý.

3️⃣ Các bước tạo và lưu chữ ký trong PDF

a. Mô tả lý thuyết

Bước	Mô tả	Thực hiện
1	Chuẩn bị file PDF gốc	Có thể dùng file bất kỳ
2	Tạo Signature field (/FT /Sig)	Dùng pikepdf hoặc iText
3	Dự trữ vùng /Contents (8192 bytes)	Chuỗi <0000...>
4	Xác định /ByteRange (4 giá trị)	Vùng được hash
5	Tính hash SHA-256 trên ByteRange	openssl dgst -sha256
6	Tạo chữ ký PKCS#7 (CMS detached)	openssl cms -sign
7	Chèn blob PKCS#7 vào /Contents	Bằng script
8	Ghi incremental update	Giữ nguyên nội dung cũ
9	(Tùy chọn) Thêm timestamp & DSS	Với TSA hoặc PyHanko