# Firmas solidarias: Tecnología criptográfica para los derechos humanos

Solidarity signatures: Cryptographic technology for human rights

**Equipo:**

Ricardo Marín Pérez | A01174384

Daniel Ríos Zúñiga | A01174445

José Antonio Torres Villegas | A00835737

Mario Alberto Landa Flores | A00836172

Victor Adid Salgado Santana | A01710023

**Instituto Tecnológico y de Estudios Superiores de Monterrey**

Uso de álgebras modernas para seguridad y criptografía (Gpo 602)

**Socio formador: Casa Monarca**

**Profesor:**

Luis Miguel Méndez Díaz

Daniel Otero Fadul

16 de marzo del 2025

# Índice



# Introducción

Identificar documentos falsos es una tarea que ha representado un reto desde hace varios años. Empresas y organizaciones en muchos países han dedicado sus esfuerzos para encontrar soluciones a este problema. Los gobiernos tienen un interés prioritario en la detección de documentos falsificados debido a su impacto en la seguridad nacional, la prevención del crimen y el control migratorio. La falsificación de documentos puede facilitar actividades ilícitas como el tráfico de personas, el fraude financiero y el terrorismo, al permitir que individuos oculten su identidad o accedan ilegalmente a beneficios y derechos. En el contexto migratorio, la verificación rigurosa de documentos es esencial para garantizar procesos justos de regularización, evitar la explotación de personas en situación vulnerable y prevenir el uso indebido de recursos públicos.

# Marco de referencia

El estudio desarrollado en la Ecole des Sciences Criminelles de la Universidad de Lausana y aplicado en Suiza propone un método eficiente basado en el análisis de datos forenses digitalizados para detectar documentos fraudulentos. Este método se implementa con la ayuda de la Base Intercantonale des Documents d'Identité Frauduleux (BIDIF), una base de datos utilizada por la policía suiza para la identificación de documentos falsos mediante la comparación de características visuales y técnicas.

Uno de los principales enfoques del estudio es el análisis visual de imágenes digitalizadas, para lo cual los peritos forenses examinan aspectos como las técnicas de impresión, reacciones bajo luz ultravioleta y alineación del texto, con el fin de identificar inconsistencias que indiquen falsificación.

Además, el método se basa en la comparación de características de perfilado, distinguiendo entre características generales y específicas. Las generales incluyen el tipo de documento, tipo de fraude (como falsificación total, alteración de un documento genuino o uso de documentos en blanco robados), país de emisión y material de fabricación. Por otro lado, las características específicas incluyen errores de alineación, fuentes incorrectas, errores ortográficos y sintácticos, defectos en logotipos o marcas de seguridad y fallos en el proceso de impresión, como variaciones de color o marcas accidentales.

Para mejorar la identificación de documentos fraudulentos, la base de datos BIDIF permite la detección automática de patrones mediante algoritmos de visión por computadora. Los documentos ingresados en el sistema se comparan con registros previos para identificar similitudes que sugieran un origen común, facilitando la detección de redes de falsificación y la vinculación de documentos falsos a grupos delictivos.

Los estándares de seguridad para detectar documentos digitales falsos incluyen una combinación de prácticas tecnológicas y procedimientos de verificación que se emplean para identificar y prevenir fraudes. Una de las medidas clave es la verificación de la autenticidad de la firma digital. Esto se logra mediante el uso de certificados digitales emitidos por autoridades certificadoras confiables, que permiten validar que la firma no ha sido alterada. Además, algunas soluciones de seguridad implementan blockchain para registrar la autenticidad de los documentos, proporcionando un medio inviolable para verificar la integridad del contenido.

Otro enfoque importante es el análisis de metadatos. Los metadatos de un documento pueden revelar información crucial sobre su autenticidad, como las fechas de creación y modificación, el software utilizado y los usuarios que han interactuado con el archivo. Herramientas especializadas permiten examinar estos metadatos para identificar cualquier inconsistencia o alteración que podría ser indicativa de falsificación.

Finalmente, el análisis forense de documentos juega un papel esencial en la detección de documentos falsos. El uso de herramientas avanzadas de OCR (Reconocimiento Óptico de

Caracteres) permite extraer el texto de los documentos y compararlo con bases de datos de documentos legítimos, buscando patrones que puedan sugerir modificaciones. Además, el análisis de los aspectos visuales del documento, como las marcas de agua, la resolución y la calidad de las imágenes, también puede ser útil para identificar documentos manipulados o falsificados.

Para determinar si existen trabajos previos relacionados con el reto planteado, se realizó una investigación documental basada en el análisis del informe: "Automating Immigration and Asylum: The Uses of New Technologies in Migration and Asylum Governance in Europe" (Ozkul, 2023), publicado por la Universidad de Oxford. Este documento analiza el uso de tecnologías emergentes en la gestión migratoria y de asilo en Europa, incluyendo procesamiento automatizado de documentos, reconocimiento de texto y análisis de datos. Misma que, tras indagar a profundidad, detalló múltiples semejanzas, con respecto a sus soluciones tecnológicas, que se asemejan a la propuesta para Casa Monarca.

La investigación se estructuró en tres fases:

1. **Identificación de tecnologías utilizadas:** Se analizaron las tecnologías documentadas en el informe, comparándolas con las soluciones propuestas en este reto.
2. **Clasificación de los trabajos identificados:** Se consideró toda aquella información cuyo contenido estuviera relacionado con las tecnologías identificadas tales como automatización de documentos, uso de inteligencia artificial y digitalización en migración y asilo.
3. **Selección del trabajo más relevante:** Se determinó cuál de las implementaciones documentada**s** en el informe tiene mayor similitud con la solución buscada para Casa Monarca.

Es de esta manera que en cada ámbito de las tecnologías de interés se encontró lo siguiente con respecto a las implementaciones que se les dieron para el caso del informe:

## Automatización de Procesamiento de Documentos

Se implementaron tecnologías para la verificación de documentos en la Unión Europea, utilizadas para validar la identidad de solicitantes de asilo y migrantes. Estas implementaciones se han llevado a cabo en países como Países Bajos, Bélgica y Francia, de los cuales podría haber semejanza con el uso de OCR con Pytesseract y OpenCV.

## Inteligencia Artificial en la Clasificación de Solicitudes

Se menciona el uso de algoritmos de categorización y análisis de documentos, como en el EU Settlement Scheme del Reino Unido, donde se automatizó la evaluación de solicitudes migratorias. Esto se relaciona con el uso de Azure AI Document Intelligence para estructurar datos de documentos en Casa Monarca.

## Evaluación de Riesgo y Priorización de Casos

Se han desarrollado modelos automatizados para clasificar solicitudes según su nivel de prioridad, como el sistema usado en el Reino Unido para evaluar riesgos en aplicaciones de matrimonio y ciudadanía. Aunque el reto de Casa Monarca no busca clasificación de riesgo, la metodología puede aplicarse para ordenar y estructurar información de beneficiarios.

## Reconocimiento de Texto y Traducción Automática

Algunos países han probado el reconocimiento de dialectos y nombres en solicitudes de asilo, con IA aplicada en Alemania y Turquía. Esto se asemeja al uso de OCR en español e inglés para digitalizar documentos de migrantes en Casa Monarca.

Casa Monarca utiliza PowerApps, por lo que los servicios proporcionados por Microsoft son opciones a considerar. Inteligencia de Documentos es un servicio de Azure que permite extraer la información y estructura de documentos y formularios. Se pueden entrenar modelos con este servicio junto con documentos originales para detectar inconsistencias en documentos alterados.

Por otro lado, Open CV es una librería de visión computacional que se puede utilizar en Python, con esta herramienta se pueden procesar y analizar las imágenes, esto sirve para comparar documentos auténticos y documentos falsos y encontrar manipulaciones visuales. Pytesseract (wrapper de Google's Tesseract-OCR Engine), es una herramienta de OCR (Reconocimiento Óptico de Caracteres) para Python, la cuál permite leer y extraer texto de imágenes, esto resulta útil para validar campos como fechas de los documentos.

Con estas 2 librerías de Python se puede proponer una solución que analice tanto el contenido (texto) de los documentos como su apariencia (imágen). Complementando estas herramientas, se pueden construir modelos de Machine Learning con Sci-kit Learn o Tensor Flow que permitan hacer una clasificación de documentos y detección de anomalías.

El principal uso de recursos informáticos los cuales utilizan Casa Monarca son las PowerApps, debido al cierto grado de facilidad, su entorno de trabajo cuenta con su mayoría operado por este recurso, además, no disponen de una aplicación o algoritmo con la capacidad de detección de documentos falsos o no permitidos.

Con las tecnologías investigadas y el contexto de Casa Monarca, se generan dos propuestas de solución. Aplicación local utilizando las librerías en Open CV y Pytesseract hecha en Python. Aplicación con en el ecosistema Microsoft, en la cuál se utilizarían los servicios de Azure: Blob Storage, Document Intelligence y App service, resultaría en un costo estimado de 71.05 USD por mes (si la organización tiene algunos servicios ya incluídos o está en

Microsoft for Nonprofits los costos disminuirían), a continuación se desglosa una estimación del costo de esta solución.

Document Intelligence (Form Recognizer) 41.50 USD por mes:

Azure Form Recognizer, Pago por uso, S0: 1 x 1000 páginas personalizadas, 0 x 1000 páginas pregeneradas, 1 x 1000 páginas de lectura, 0 x 1000 páginas de complemento, 1 x 1000 páginas de consulta

Azure Blob Storage 11.55 USD por mes:

Almacenamiento de blobs en bloque, Uso general V2, Espacio de nombres plano, LRS Redundancia, Acceso frecuente Nivel de acceso, Capacidad: 500 GB - Pago por uso, 10 x 10 000 operaciones de escritura, 10 x 10 000 operaciones de lista y operación de creación de contenedores, 10 x 10 000 operaciones de lectura, 1 x 10 000 otras operaciones. 1000 GB de recuperación de datos, 1000 GB de escritura de datos, SFTP deshabilitado

App Service 18 USD por mes:

Nivel Basic; 1 B1 (1 núcleos, 1.75 GB de RAM, 10 GB de almacenamiento) x 240 Horas; Sistema operativo Windows; 0 SSL SNI Conexiones; 0 IP SSL Conexiones; 0 Dominios personalizados; 0 Certificados SLL estándar; 0 Certificados SSL con caracteres comodín

# Referencias


1.- Calculadora de precios | Microsoft Azure. (s. f.). Microsoft Azure. https://azure.microsoft.com/es-es/pricing/calculator/

2.- Laujan. (2025, February 6). *What is Azure AI Document Intelligence? - Azure AI services*. Microsoft Learn. https://learn.microsoft.com/en-us/azure/ai-services/document-intelligence/overview?view=doc-intel-4.0.0

3.- Moulin, S., Weyermann, C., Baechler, S. 2022 An efficient method to detect series of fraudulent identity documents based on digitised forensic data. Science & Justice. Pages 610-620. https://doi.org/10.1016/j.scijus.2022.09.003.

4.- OpenCV. (2025, February 21). OpenCV - Open Computer Vision Library. https://opencv.org/

pytesseract. (n.d.). PyPI. https://pypi.org/project/pytesseract/

5.- Ozkul, D. (2023). Automating immigration and asylum: The uses of new technologies in migration and asylum governance in Europe. Refugee Studies Centre, University of Oxford. https://www.rsc.ox.ac.uk/files/files-1/automating-immigration-and-asylum_afar_9-1-23.pdf