

Reconocimiento de escenas

Carlos Bienvenido Ogando Montas

Proyecto “Caso Policía Nacional”

Sustentantes

Teresa Meridith Díaz
Vinicio Arturo Castillo Castillo

Santo Domingo, R.D.
18 de abril del 2022

Tabla de contenido

<i>Proyecto “Caso Policía Nacional”</i>	3
Introducción	3
Recursos locales que se emplearon.	4
Recursos en la nube que se utilizaron.....	4
Infraestructura del sistema que proponen como solución.	5
Diagrama de flujo de los datos y proceso.	6
Procedimiento realizado para resolver el problema.....	7
Método de recolección de datos y recopilación de la información de entrenamiento.....	7
Imágenes	7
Creación de los datos	7
Formato de salida propuesto de la información. Inteligencia Artificial Aprendizaje automático.	7
Rol de los miembros del equipo y trabajo realizado.....	8
Vinicio Castillo	8
Teresa Meridith Díaz	8
Costos de la propuesta.....	9
Tabla de uso de MySQL	10
Métricas de efectividad del sistema propuesto.	10
Resultados por persona en el set de prueba:	10
Dificultades al realizar el proyecto.....	11
Posibles mejoras de la propuesta.	11
Riesgos y limitaciones.	11
Bibliografía:	12
Anexos.....	13
Fase inicial del proyecto (primera reunión)	13
Diseño la salida visual del PDF desde la aplicación	13
Imágenes de la interfaz grafica	14

Proyecto “Caso Policía Nacional”

Introducción

La foto de prontuario, también llamada ficha policial, es un retrato fotográfico de un individuo desde la cintura hacia arriba, tomada luego de un arresto.

Sus orígenes datan a principios de los años 1800 en Bélgica y Dinamarca. Donde se introdujo el sistema antropométrico de la fotografía en el registro y clasificación de los individuos considerados peligrosos. Los malhechores eran fotografiados con la finalidad de guardar un registro documentado, luego se contrataron fotógrafos profesionales para hacer tomas de los criminales.

De acuerdo con Schvarzstein (2020), el retrato fotográfico estandarizado de frente y perfil, asociado a una ficha, contenía una serie de detalles (información sensible) donde se encuentra descripción física, antecedentes judiciales, entre otras, fue la primera forma “científica” de identificación policial.

En 1888, después de la derrota de la comuna de París, el francés *Alphonse Bertillon* desarrolló un método que años después fue implementado por la policía de París, este se propagó por el mundo, impulsando la circulación transnacional de información delictiva.

Posteriormente en 1890, el sistema del “Bertillonage” quedó descrito en un precioso libro de instrucción fotográfica titulado “La photographie judiciaire” el cual contemplaba imágenes impresas por reproducción fotomecánica con resultados para lograr representaciones uniformes que servirían al reconocimiento y la clasificación de los imputados.

En el año 1854, en Lausanne, Suiza, por primera vez estas fotografías comenzaron a circular en las comisarías. Las primeras 'galería de ladrones' se implementó en 1858 en el Departamento de Policía de Nueva York y en la década del setenta su uso se había extendido en Estados Unidos y Europa”.

En la actualidad, diversos países ponen a circular publicación de fotografías con los criminales más buscado especificando datos relevantes.

El proyecto Policía Nacional, aplicación Identificador de rostros, se basa en estos formatos, pero orientado en una forma visual simple y rápida, creando una base de datos del perfil criminológico con patrones particulares de las personas capturadas.

Recursos locales que se emplearon.

- a. Se utilizó Python para hacer el desarrollo e implementación de la aplicación y para establecer las conexiones con los recursos necesarios. Las librerías que utilizamos para poder conseguir esto son:
- **PyQt5 y tkinter** para implementar la GUI.
 - **Pyodbc** para establecer conexiones con la base de datos.
 - **azure.cognitiveservices.vision y msrest.authentication** para establecer el cliente de comunicación con el Face API de Azure.
 - **Request** para poder descargar las imágenes desde GitHub para entrenar nuestro modelo.
 - **Fpdf** para implementar la generación de reportes en formato PDF.
- b. Utilizamos Microsoft SQL Server Management Studio para poder administrar localmente una base de datos llamada “**PNRec**” que creamos por medio de los servicios de Azure para así disponer de un grupo de recursos y un servidor.

Recursos en la nube que se utilizaron.

Azure: utilizamos este servicio de Microsoft para poder alojar y administrar nuestra base de datos y también para tener acceso a su servicio de reconocimiento facial (Face API) y poder entrenar el modelo de IA.

Git y GitHub: utilizamos estos recursos para poder compartir un mismo repositorio de datos del proyecto para los integrantes del equipo y así poder compartir y colaborar conjuntamente. También utilizamos GitHub para almacenar las imágenes de entrenamiento para nuestro modelo de reconocimiento facial.

Infraestructura del sistema que proponen como solución.

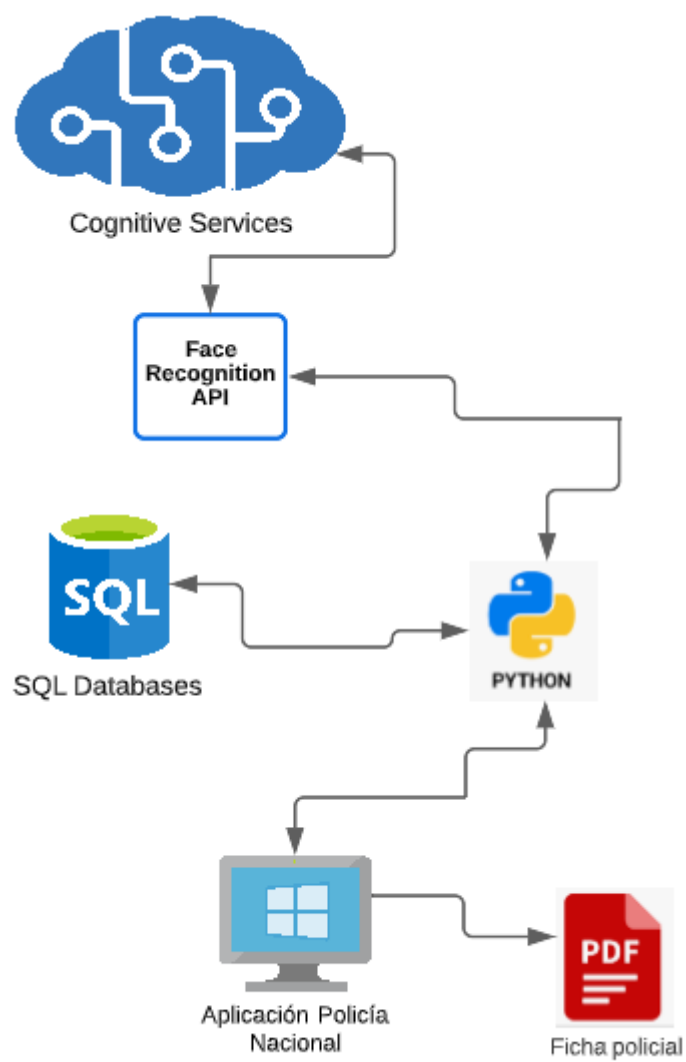
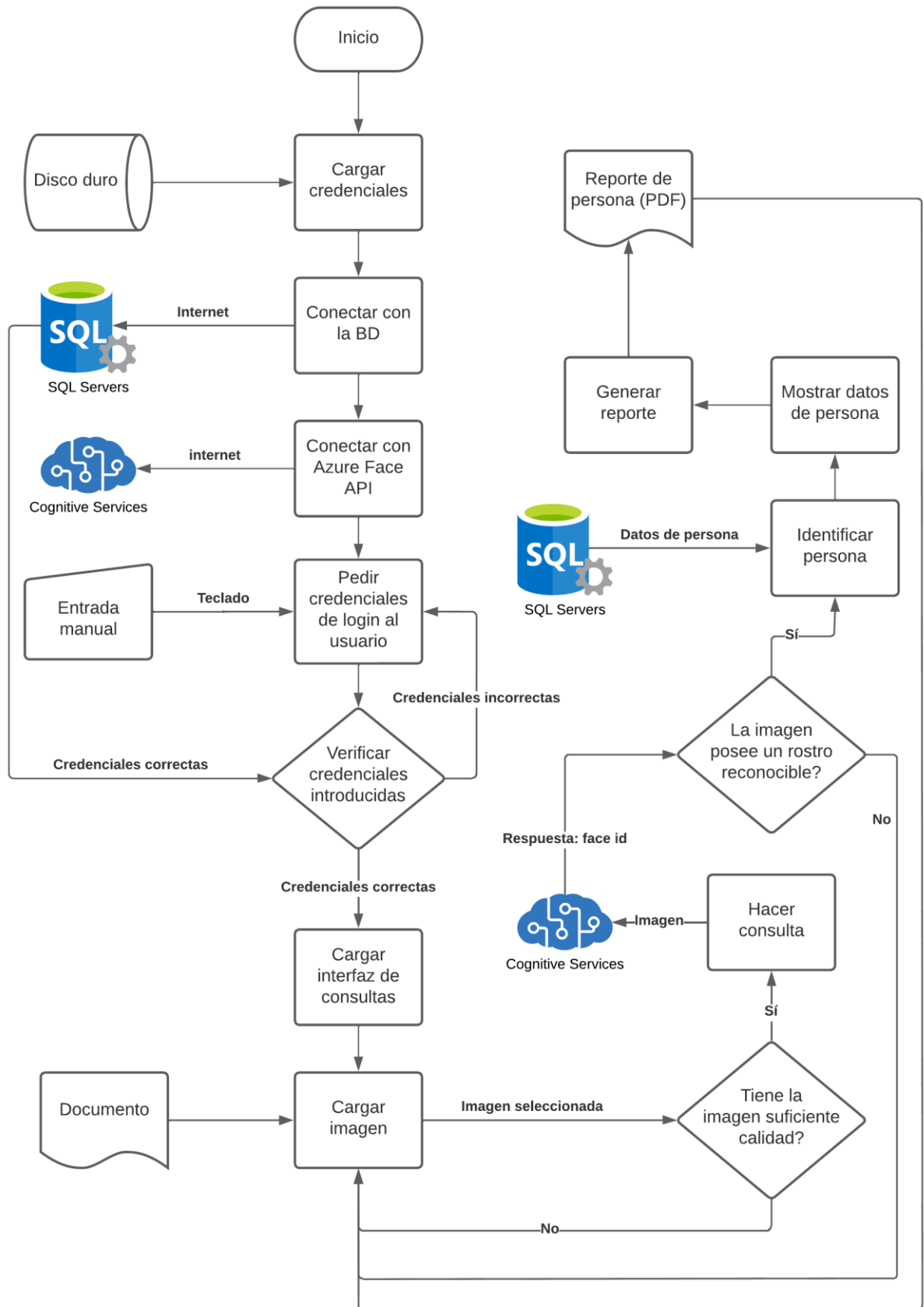


Diagrama de flujo de los datos y proceso.



Procedimiento realizado para resolver el problema.

Luego de haber identificado el problema, se ideó una solución basada en el uso de un sistema que por medio de un modelo de reconocimiento facial y acceso a una base de datos, le da la capacidad a la Policía Nacional de identificar sospechosos y recuperar todos sus datos con tan solo poseer una foto del rostro de tal persona. Para así ayudar a eficientizar y mejorar el proceso de búsqueda y captura en nuestra sociedad.

Pasos que ejecutamos para resolver el problema:

- ✓ Se partió de las necesidades existentes en la Policía Nacional, haciendo una proyección a futuro de las posibles ventajas de implementación de esta solución basada en IA.
- ✓ Se creó una gráfica de como sería el desarrollo de la aplicación y la arquitectura de esta.
- ✓ Se empezó a investigar cuales recursos serian necesarios para implementar este proyecto y como se utilizan tales recursos.
- ✓ Se dividió el proyecto en partes:
 - a. Base de datos
 - b. Modelo de reconocimiento facial
 - c. Interfaz grafica
 - d. Filtrado y recolección de los datos para entrenamiento y prueba del modelo.
 - e. Programa principal que se encargará de comunicar todas las partes del proyecto.
 - f. Programa para encargarse de la creación de reportes de datos en PDF. Y se diseñó una plantilla principal para estos documentos.
- ✓ Se empezó a hacer pruebas de comunicación entre todas las partes.
- ✓ Se creó un repositorio compartido para el manejo del proyecto.
- ✓ Luego se empezó el desarrollo de la aplicación.

Método de recolección de datos y recopilación de la información de entrenamiento.

Imágenes

- Se utilizó la plataforma *Getty Images* por la calidad y alta resolución de las imágenes.
- Fotos originales (fotografías digitales a través de móvil).

Creación de los datos

- Visita a diferentes portales con documentación de perfiles criminológicos tanto nacionales como en el exterior.
- Visita al portal de la ONE para revisar los crímenes habituales en República Dominicana.

Formato de salida propuesto de la información. Inteligencia Artificial Aprendizaje automático.

Nuestro modelo devuelve como resultado un tipo de dato llamado "Face ID" que representa las características espaciales y contextuales de el rostro que se encuentra en la imagen proporcionada a la aplicación. Luego este Face ID es comparado con los Face ID de las personas que están almacenadas en el entrenamiento del modelo, así devolviéndonos el nombre de la persona a la que cuyo rostro le pertenece el Face ID, entonces se utiliza el nombre encontrado como llave para recuperar todos los datos disponibles sobre esta en la base de datos y generar un reporte en formato PDF para que el usuario pueda hacer uso de la información efectivamente.

Rol de los miembros del equipo y trabajo realizado.

Se dividieron los recursos en las cuentas académicas de Azure con la finalidad de evitar que se agotaran antes de la presentación del proyecto.

Vinicio Castillo

- ✓ Desarrolló los programas de interconexión de la aplicación en Python.
- ✓ Desarrollo de aplicación en de consulta Policía Nacional.
- ✓ Desarrollo de salida de PDF de aplicación.
- ✓ Recopiló imágenes para el entrenamiento del modelo.
- ✓ Entrenó el modelo de Azure.

Teresa Meridith Díaz

- ✓ Creó los perfiles policiales (información) para alimentar la base de datos.
- ✓ Desarrolló la base de datos MySQL y conexiones de Azure.
- ✓ Diseño la salida visual del PDF de la aplicación.
- ✓ Recopiló imágenes para el entrenamiento del modelo en distintas posiciones.
- ✓ Realizó el documento del proyecto en base a la rúbrica solicitada.

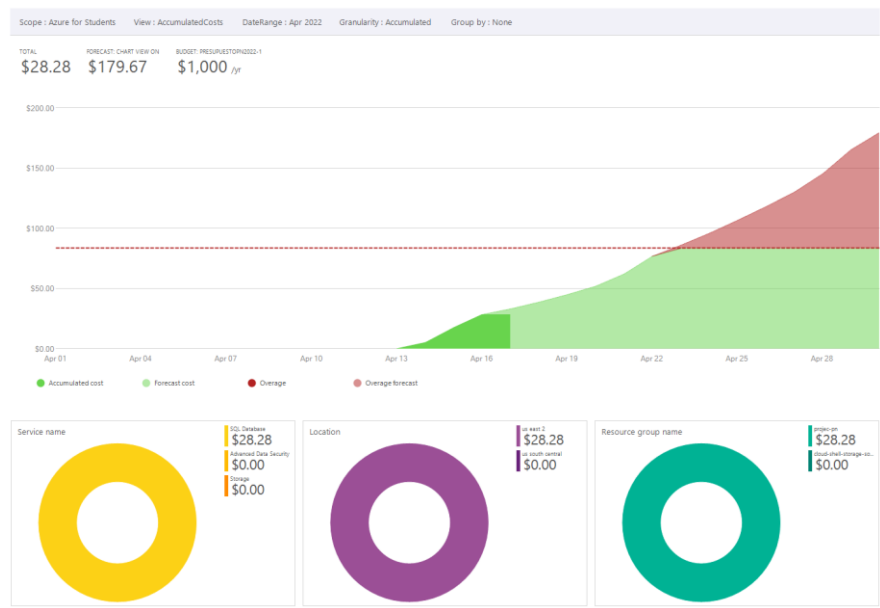
Costos de la propuesta.

Azure cobra la por cada mil (1000) transacciones un costo de aproximadamente \$0.5USD y por almacenar las imágenes de las caras de las personas cobra \$0.01 USD por cada mil (1000) imágenes almacenadas, por lo que se si hiciera un estimado de que existieran 2000 usuarios por provincia en la Rep. Dom. Y cada uno de estos usuarios hicieran un promedio de 6 o 7 transacciones al día por todo un mes, entonces el costo total seria de \$4,821.20 USD al mes, para darle servicio a todo el país y si se guardaran 10 millones de personas y se entrenaran 4 imágenes por cada persona, entonces el costo mensual de almacenar esos datos seria de \$400 USD.

Azure cobra por almacenamiento en base de datos (única), con un almacenamiento con redundancia geográfica, almacena varias copias de seguridad donde asegura fallas transitorias, desastres naturales, entre otras. Una instancia (redundancia) por cada 730 horas tendría un costo de \$888,95 USD por computadora y la licencia de SQL sería de \$583.80 USD, con un almacenamiento de 32G y una restauración de 5G, con un valor de \$3.68 USD.

Microsoft Azure Estimate				
Your Estimate				
Service type	Region	Description	Estimated monthly cost	Estimated upfront cost
Azure Cognitive Services	East US 2	Face API, Standard tier: 12,000,000 transactions with 40,000,000 faces stored per month.	\$5,221.20	\$0.00
Azure SQL Database	East US	Single database, service tier: Backup storage RA-GRD, hardware Standard-series (Gen 5), purchame model vCore and instance 8 vCore. Compute Local 1 x 730 hours	\$1,476.43	\$0.00
Support		Support	\$0.00	\$0.00
		Licensing Program	Microsoft Customer Agreement (MCA)	
		Billing Account		
		Billing Profile		
		Total	6,697.63	\$0.00
Disclaimer				
All prices shown are in United States – Dollar (\$) USD. This is a summary estimate, not a quote. For up-to-date pricing information please visit https://azure.microsoft.com/pricing/calculator/				
This estimate was created at 4/17/2022 9:35:26 PM UTC.				

Tabla de uso de MySQL



Métricas de efectividad del sistema propuesto.

Para el entrenamiento y evaluación de este proyecto, creamos un dataset de 15 personas de diferentes: edades, géneros, etnias, cabello y accesorios faciales. Donde se contiene 4 imágenes por cada persona, las cuales se dividen 75% para entrenamiento del modelo y 25% para evaluar el modelo.

Como métricas de evaluación para modelos de clasificación utilizamos:

- Accuracy:** $ACC = (TP + TN) / (P + N)$. El modelo logro una calificación de 1.0.
- Precision:** $P = TP / (TP + FP)$. El modelo logro una calificación de 1.0.

Resultados por persona en el set de prueba:

Persona	Confianza del modelo
Adrián	92%
Alexandra	80%
Misael	93%
Arturo	97%
Nelson	95%
Pedro Alfredo	95%
Dlan Emmanuel	94%
Alfred	95%
Hugo	93%
Juan Pablo	94%
Marola	96%
Martin	95%
Paola	98%
Nicol	92%
Roberto José	95%

Al revisar estos resultados pudimos darnos cuenta de que el modelo realmente es muy capaz, pero para poder evaluarlo correctamente necesitaremos mas datos para poder conseguir información de rendimiento mas significativa, lo cual, por falta de recursos, esta fuera del alcance de este proyecto en este momento.

Dificultades al realizar el proyecto.

- Resolver errores de conexión API Azure con base de datos y aplicación final de Policía Nacional.
- Documentación estandarizada en Azure para implementación del código, conexiones de bases de datos y aplicación.
- Aprender desde cero MySQL para crear las tablas, columnas y las interconexiones.

Posibles mejoras de la propuesta.

- ✓ Crear una base de datos que las tablas estén estandarizadas por etnia, genero (sexo), permitirá una búsqueda más rápida.
- ✓ La presentación del PDF incluir datos biométricos, acceso con código QR y código de barras para numeración de la ficha.
- ✓ Mejorar la seguridad e integridad de los datos manejados en las comunicaciones de la aplicación y proteger su código fuente para ocultar credenciales.
- ✓ Agregar un sistema de registro de sesiones de los usuarios que acceden a la aplicación, para así poder saber quien accede y cuando lo hace. Para poder agregar otra capa de seguridad a los datos sensibles que esta aplicación maneja.

Riesgos y limitaciones.

- El sistema de Azure solo puede soportar hasta 1 millón de personas diferentes y cada persona puede tener hasta un máximo de 256 imágenes asociadas a esta. Por lo que si se espera hacer una implementación de la aplicación que exceda estas capacidades, pues habrá que reevaluar la arquitectura y el plan de suscripción a Azure.
- Azure solo permite un máximo de transacciones por segundo, por lo que hay que determinar cuidadosamente la cantidad de usuarios por segundo que estarán haciendo consultas a el modelo a cada momento.

Instance	Transactions Per Second (TPS) *	Features	Price
Free - Web/Container	20 transactions per minute	Face Detection Face Verification Face Identification Face Grouping Similar Face Search	30,000 transactions free per month
Standard - Web/Container	10 TPS	Face Detection Face Verification Face Identification Face Grouping Similar Face Search	0-1M transactions - \$1 per 1,000 transactions 1-5M transactions - \$0.80 per 1,000 transactions 5-100M transactions - \$0.60 per 1,000 transactions 100M+ transactions - \$0.40 per 1,000 transactions
		Face Storage	\$0.01 per 1,000 faces per month

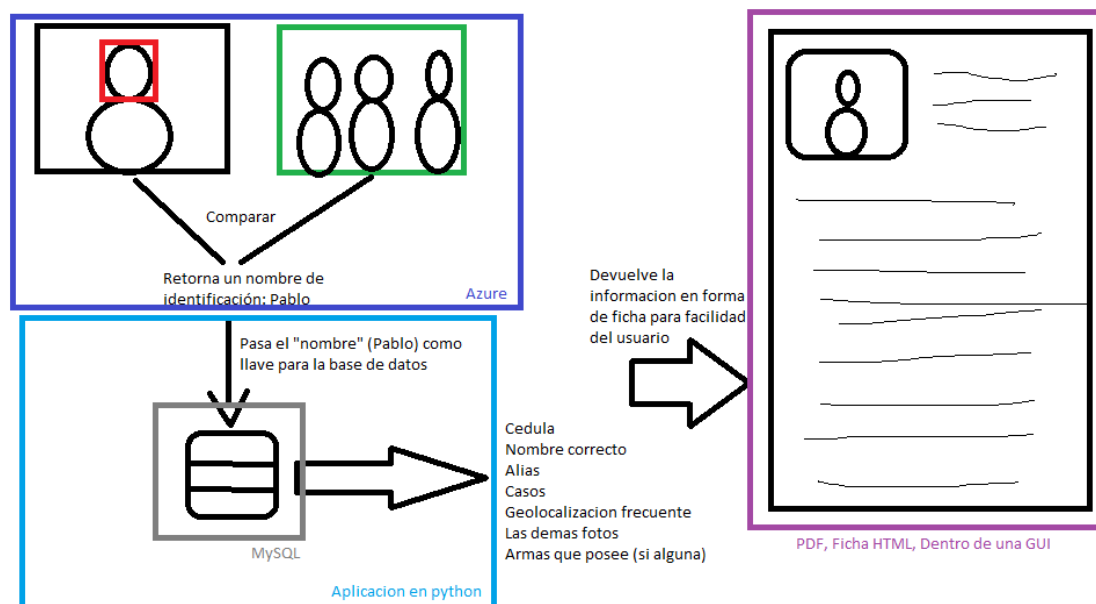
* TPS only applies to web endpoint

Bibliografía:

1. Schvarzstein, D. (2020). Cómo nació la fotografía de identificación policial: creado en el siglo XIX para reconocer delincuentes, el sistema antropométrico es un capítulo apasionante en la historia judicial. Recuperado de: <https://noticias.perfil.com/noticias/cultura/como-nacio-la-fotografia-de-identificacion-policial.phtml#:~:text=En%20las%20%C3%BAltimas%20d%C3%A9cadas%20del,%E2%80%9Ccient%C3%ADfica%E2%80%9D%20de%20identificaci%C3%B3n%20policial.>
2. Flores, J. (2020?). [Como] CREAR una BASE DE DATOS en la NUBE con SQL Server. Recuperado de: <https://www.youtube.com/watch?v=HiOO46X4LxA>.
3. García, F. (2020). Crear TABLAS en SQL Server con y sin CODIGO. Recuperado de: <https://www.youtube.com/watch?v=rHcB0KGoysw>.
4. García, F. (2020). Insertar Datos en TABLAS en SQL SERVER (INSERT). Recuperado de: <https://www.youtube.com/watch?v=PfY6Ypzp0Io>.
5. García, F. (2020). Crear Base de Datos en Sql Server y Mysql. Recuperado de: <https://codigosql.top/sql-server/como-crear-tablas-en-sql-server/>
6. Inicio rápido: usar la biblioteca del cliente Face (2022). Recuperado de: <https://docs.microsoft.com/en-us/azure/cognitive-services/face/quickstarts/client-libraries?tabs=visual-studio&pivots=programming-language-python>
7. Person Group Person (2022). Recuperado de: <https://docs.microsoft.com/en-us/rest/api/faceapi/person-group-person>
8. arXiv.org e-Print archive. (s. f.). Metric Classification Network in Actual Face Recognition Scene. <https://arxiv.org/ftp/arxiv/papers/1910/1910.11563.pdf>

Anexos

Fase inicial del proyecto (primera reunión)



Diseño la salida visual del PDF desde la aplicación

FICHA TÉCNICA PN20220415		Alejandra Ortega Núñez	
DESCRIPCIÓN FÍSICA			
	ID: 00120678786 Fecha de nacimiento: 1/1/1976 Edad: 46 Género: F Estatura: 5' 5" Peso: 140 LBS Tono de piel: Negra Ojos: Negros Cabello: Negros Cicatrices: Ninguna Alias e apodo: Torpedo		
OTROS DATOS			
Padre: Franklin Ortega Santana Madre: Claudia del Carmen Núñez Fernández Dirección: C/Ana Josefa Puella, No.20, Mirador Sur			
RÉCORD DELICTIVO			
Delitos: Soborno, corrupción, caso Odebrecht. Otros casos: Fraude en las transacciones y cambio de dólares Caso Baninter. Instrumento usado: Pistola Magnum Beretta 9mm.			

POLICÍA NACIONAL DOMINICANA
Palacio de la Policía Nacional (Sede Central)
Avenida Leopoldo Navarro No. 402, Gacaca, Santo Domingo,
D.N., República Dominicana
Teléfonos: (809) 682-2151 / (809) 685-2020
info@policia.gob.do

Imágenes de la interfaz gráfica



Costo de recursos Azure anual Plan Académico, pueden variar según el contrato

