Tarea Flujo de Datos

Juan Manuel García Moyano IABD Informática y comunicaciones

Índice

Enunciado	
Ejercicio 1: Diseño del flujo de datos	
Ejercicio 2: Implementación del flujo de datos	
Ejercicio 3: Pruebas del sistema	
Eiercicio 4: Análisis del impacto en el negocio.	

Enunciado

Esta tarea consta de cuatro ejercicios a través de los cuales trabajara el uso y aplicación de los flujos de datos aplicado a un caso práctico real:

Ejercicio 1: Diseño del flujo de datos

Enuncie un caso práctico en el que sea necesario automatizar un proceso mediante un flujo de datos. Identifique las herramientas y servicios que podrían intervenir en la solución. Describa los pasos del proceso y los resultados esperados. Si es necesario diseñe un esquema del sistema.

Caso práctico

Una startup necesita realizar una búsqueda sobre las noticias de Inteligencia Artificial y comprobar si en el título existe las siglas IA. Además, le interesaría guardar algunos datos extra como la URL y su descripción. Por tanto, esto conlleva abrir el navegador, realizar una búsqueda, guardar cada noticia en un Excel/csv con las columnas necesarias y en una de ellas poner si el título contiene la palabra IA.

Para esta tediosa labor deciden contratar a una persona a tiempo parcial en teletrabajo durante 7 meses. Antonio, que está estudiando un curso de especialización de Inteligencia Artificial y Big Data y necesita sacarse un dinerillo, decide ofrecerse para el trabajo con la suerte de que la semana pasada estuvieron viendo en clase de PIA los flujos de trabajo automatizados con Power Automate.

Así que Antonio, en vez de abordar la tarea con la tediosa labor de copiar y pegar campo a campo y comprobar si contiene la palabra indicada, decide diseñar un flujo que haga el trabajo automáticamente.

Tras varios borradores decide implementar el siguiente flujo que para cada informe hará lo siguiente:

Paso 1. Lanzará un script de PowerShell que obtendrá las noticias en las últimas 24h. Este sacará los siguientes datos:

- Título
- Descripción.

- URL.
- Fecha de publicación.
- ContieneIA (en un principio tendrá una cadena vacia).

Los guardará en un DataFrame, el cuál, luego será guardado en el csv.

- **Paso 2.** Leer el archivo csv recién creado.
- **Paso 3.** Mediante un bucle irá comprobando si en el título existe las siglas IA.
- **Paso 4.** Escribe el nuevo archivo csv con las noticias ya procesadas.
- Paso 5. Mostrará un mensaje donde indique cuando ha finalizado.

Herramientas y servicios:

- o Python: recolección de datos.
- Power Automate Desktop: Automatización del flujo.
- Almacenamiento: local en un CSV.

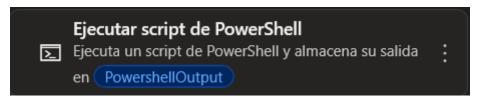
El resultado debe ser de un csv que contenga todas las columnas anteriormente indicadas y en la columna ContieneIA debe tener los valores de Si/No, es decir, las noticias ya clasificadas.

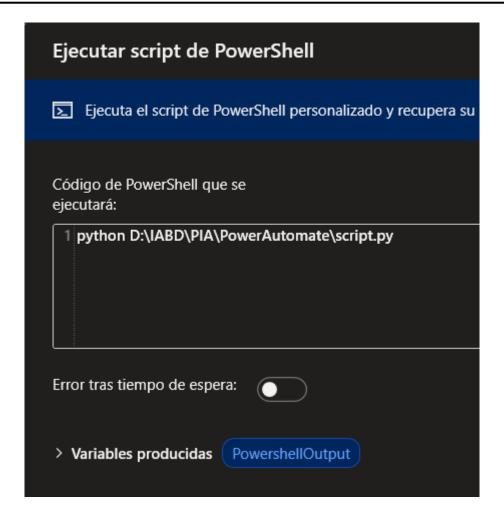
Ejercicio 2: Implementación del flujo de datos

Utilice Power Automate para crear el flujo de datos descrito en el ejercicio anterior. Asegúrese de integrar las herramientas y servicios identificados, aplicando lógicas de automatización que optimicen el proceso.

Puede implementarlo en Power Automate de escritorio, en línea o a través de la integración de ambas.

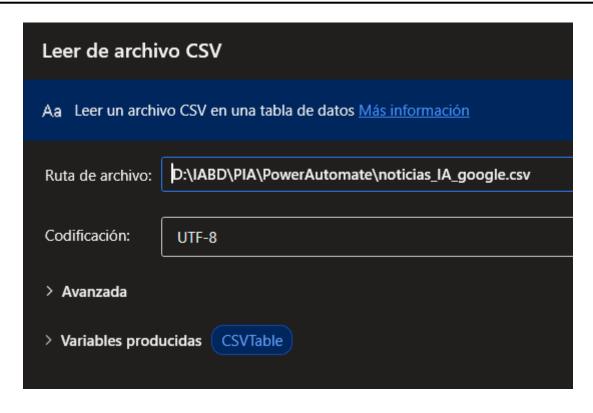
1. Ejecutar script de PowerShell.





2. Leer de archivo CSV.



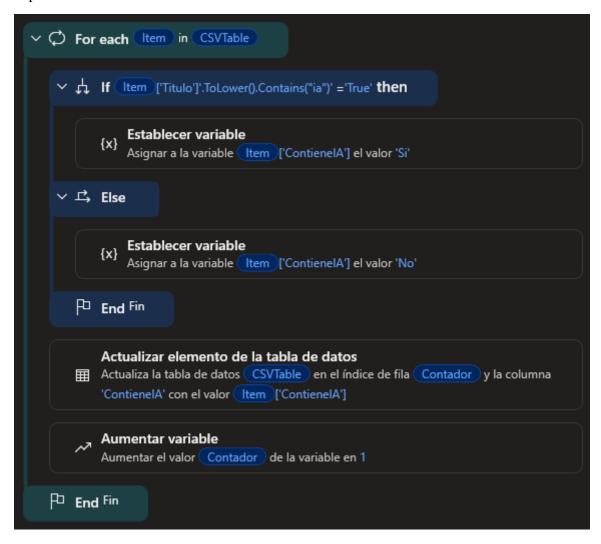


3. Establecer variable.

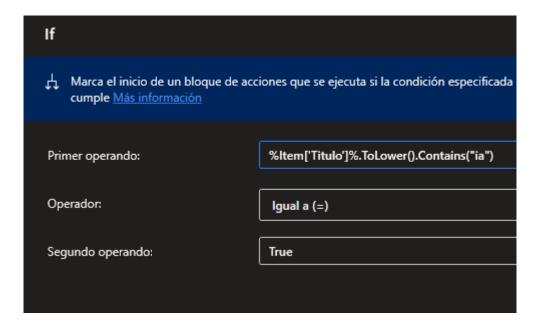




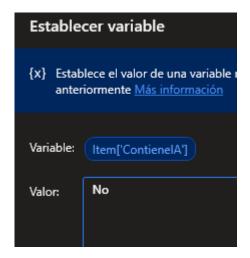
4. Bucle para recorrer la variable CSVTable. Condición para comprobar si en el título existe la palabra IA. Actualizo el elemento de la tabla de datos.

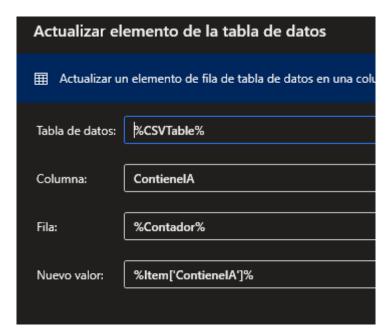


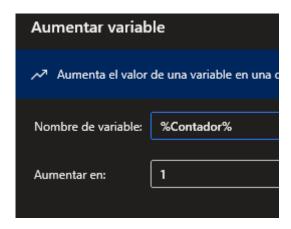
IES Zaidín Vergeles Tarea Flujo de Datos





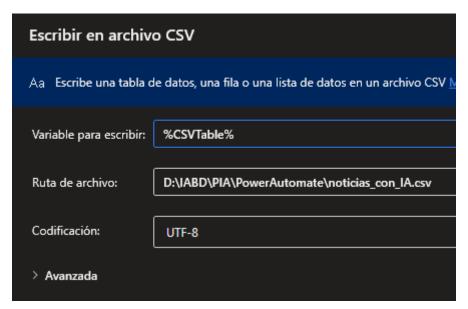




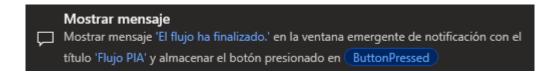


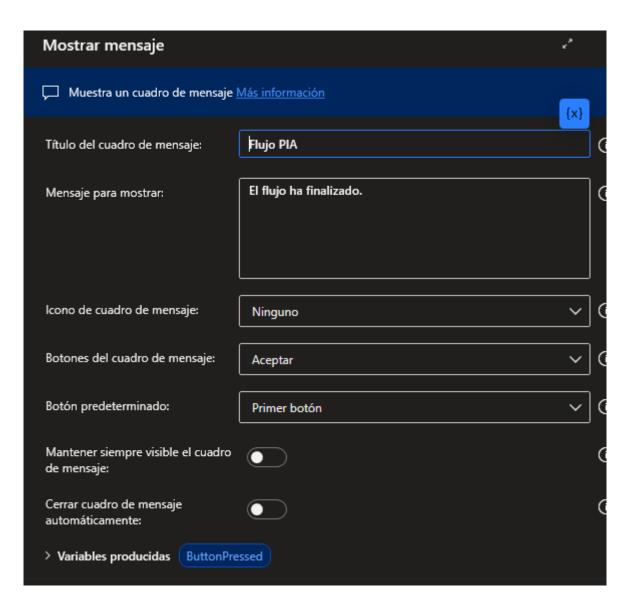
5. Escribir en archivo CSV.





6. Mostrar mensaje.





Ejercicio 3: Pruebas del sistema

Realice pruebas del flujo implementado, identificando posibles errores y proponiendo soluciones. Documente el proceso de prueba y los resultados obtenidos, si procede, proponer posibles mejoras que se podrían aplicar de acuerdo con los resultados analizados.

De primeras se han realizado diferentes pruebas, las cuales han ido surgiendo diferentes errores. Entre ellos y aunque parezca básico, Excel debe estar instalado en la máquina local si se utiliza la aplicación de escritorio, por eso ahora se utiliza con un CSV. Otro problema que surgió, fue que tuve que poner CSV diferentes, es decir, uno para leer los datos y otro para escribirlos. Si nos fijamos, el script es de PowerShell, porque si ponía un script de Python no me creaba el fichero correctamente.

Ejecuto el flujo y vemos los resultados. De primeras se ve que se han creado correctamente los ficheros.



Vamos a comprobar que se obtenga el resultado esperado.

• noticias_IA_google: este fichero guarda los resultados obtenidos de la búsqueda.

Titulo, Descripción, URL, Fecha de Publicación, ContieneIA

"8ill Gates, cofundador de Microsoft: "En una década, la inteligencia artificial hará innecesarios a los humanos para la mayoría de las cosas" - National Geographic", "ca href=mcXkz.CFDMMdPMDEAZOKDIXKVZxBMVgITnJSDWgIQZMSVTBADFFJWJMCVXMVMDZXGDKAZYVAMKSIDNFYSGMaEPVQIQIdFF3VJMCYZRIXIk7oc=5, "Mon, 31 Mar 2025 08:56:00 GNT", Inteligencia artificial: la 'revolución' que la Torre de Babel no vio ventr - El Mundo; "ca href="Thttps://news.google.com/rss/articles/CBMJAFBFW95CMXPXJFNRAMMMZTBEDFY3ZmFpZW5C49YZ3hWDUTNSFhidWxZXDh1ZMGPremPeUVBSEFTSIVQZ055dFFOWZZ3hFRDVZ5VEQZNKAUYAHFXXdWQZbH7ZamPseFTZ3SH100ZZFPRDFBVY9VUNIBBNJOTEZOMSKARXVHYZOXY8ImRjhuUUpTOEhFdzFTZzH H&M usará clones de modelos generados con IA en sus anuncios y redes sociales - Reason (hy, "ca href=""https://news.google.com/rss/articles/CBMIXAFBV99StLWAUNDFbPRPhPadpFyzbybbdmACTSCJCMXYAFBVP9StLWAUNDFbPRPhPadpStcJCMXMDPSbLXDMAPSHZJCMXPSLWAUNDFbPRPhPadpStcJCMXMDPSbLXDMAPSHZJCMXDPSLWAUNDFbPRPhPadpStcJCMXMPSBLXDMAPSHZJCMXDWpFURTASSCJCMXPSJCMXDWFXSBZKNKWAUNDFBPRPADPSCJCMXDPSLWAUNDFBPRPADPSCJCMXDPSLWAUNDFBPRPADPSCJCMXDPSLWAUNDFBPRADPSCJCMXDPSLWAUNDFBPRADPSCJCMXDPSLWAUNDFBPRADPSCJCMXDPSLWAUNDFBPRADPSCJCMXDPSLWAUNDFBPRADPSCJCMXD

• noticias_con_IA: este fichero guarda los resultados anteriores más una columna que pone si contiene o no la palabra IA.

|Titulo,Descripción,URL,Fecha de Publicación,ContieneIA
"Bill Gates, cofundador de Microsoft: "En una década, la inteligencia artificial hará innecesarios a los humanos para la mayoría de las cosas" - National Geographic", "<a href="mcKzcytDMwdPMJEd2V6bIJKVZx0NVgITnJSbWgIQZNSMTB4bFFJWJDFONVxMDZ8ZGGKa2YwWkSIbnFyaGdmaEpYQlQIdFF3VJNCYzRIX1k?oc=5,"Mon, 31 Mar 2025 08:56:00 GMT",No
Inteligencia artificial: la 'revolución' que la Torre de Babel no vio venir - El Mundo, "<a href=""https://news.google.com/rss/articles/CBMipaFBW95cUkPXipTRVRAMMMIZTBEbEY3ZmFp
ZM5CcV9YZ3hWWJUm5FhidWwz20hzNGJrem9reUVBSEpTSJUQ2055dFfoWWZ2aiGtRUYzbyV6QZhXaUYAHFhXdWQZbHhZam9seE94MFRGbEZISSILOTQ2ZFpROFBVVY9VUNIBDMJOTEZoWGARXVHYZOXYDIMRjhuUUpTOEhFdzeTIZh
NEMU usará clones de modelos generados con IA en sus anuncios y redes sociales - Reson Why, "<a href=""https://news.google.com/rss/articles/CBMizAFBWV95cUkNZoYPbbmdncXds
Visa vs estafadores: cómo la inteligencia artificial está redefiniendo la lucha contra el fraude - Forbes España, "<a href="https://news.google.com/rss/articles/CBMizAFBWV95cUkNZeiBWy5cUkN

Una mejora que haría que fuese más polivalente sin tener que tocas el flujo, es poner dos acciones que pidan al usuario el texto de la búsqueda y las palabras a buscar.

Ejercicio 4: Análisis del impacto en el negocio

Evalúe cómo el flujo de datos definido puede influir en la toma de decisiones del modelo de negocio. Identifique las ventajas que aporta en términos de eficiencia, seguridad, confiabilidad y capacidad estratégica.

Este flujo de datos no necesariamente se tiene que realizar para la búsqueda de noticias de Inteligencia Artificial. Lo importante del flujo de datos propuesto (extracción, clasificación y análisis automatizado de noticias sobre IA) tiene un impacto significativo en el modelo de negocio, especialmente en empresas donde la vigilancia tecnológica o el análisis de tendencias son clave. En cuanto a ventajas que aporta me gustaría destacar las siguientes:

- 1. Eficiencia: se puede observar una reducción del tiempo. Automatizar la recolección y clasificación de noticias elimina horas de trabajo manual. Además, de esta forma se pueden identificar tendencias rápidamente.
- 2. Seguridad y confiabilidad: al realizar un flujo minimizamos los errores como los sesgos o los errores de hacerlo manualmente.
- 3. Capacidad estratégica: podemos detectar oportunidades tempranas, tomar decisiones en función de los Si obtenidos y el flujo podría adaptarse a rastrear palabras clave.