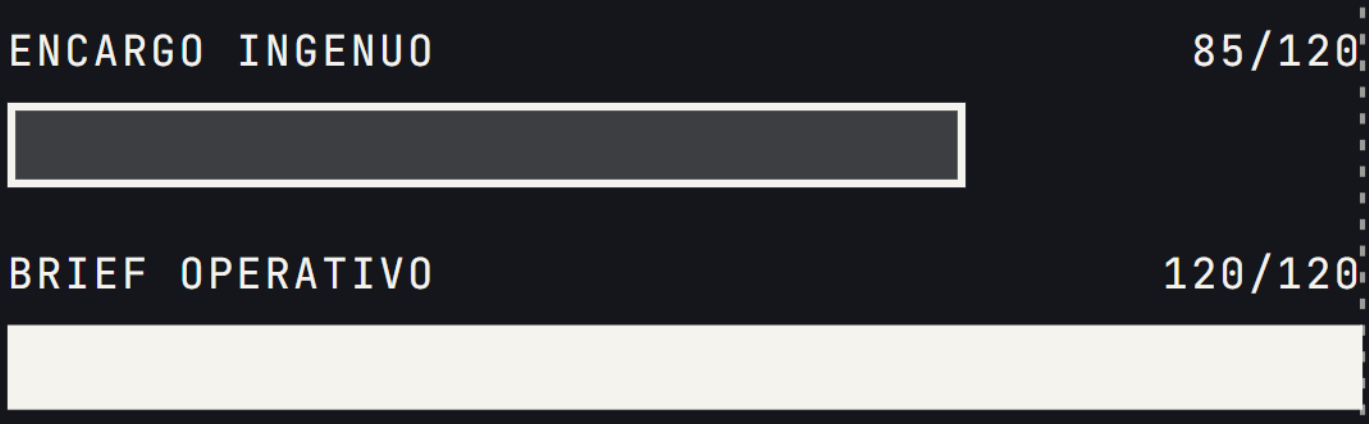
Facturas con IA, sin errores



Facturas y escaneos con IA, sin errores silenciosos

De una carpeta de facturas, tickets y escaneos a una hoja de Excel que marca sola lo dudoso

Edición: 05/11/2026. **Para quién:** eres autónomo, trabajas en una gestoría pequeña o llevas la administración de una pyme, y quieres pasar facturas, tickets y escaneos a una hoja de cálculo con IA sin que se te cuele un dato mal leído. Manejas Excel con soltura (filtros, arrastrar fórmulas) y sabes adjuntar un archivo en un chat. No hace falta programar: los capítulos que usan Python lo explican paso a paso y son opcionales.

Alcance

Las pruebas de este libro se hicieron el 04/11/2026 con facturas ficticias, en un PC con Windows 11, 31 GB de RAM y una NVIDIA RTX 5060 Ti de 16 GB. Los modelos de nube fueron Claude Haiku 5.5 y Claude Sonnet 5.5 (vía Claude Code 2.1.317); los locales, gemma4:12b y qwen3.5:9b con Ollama 0.40.14. Otros modelos u otras versiones pueden acertar o fallar en sitios distintos; las comprobaciones de la hoja sirven igual con cualquiera. La normativa citada se consultó el 05/11/2026 en el BOE y en la sede de la Agencia Tributaria. Lo legal se explica para que sepas qué preguntar, no sustituye el consejo de tu asesor. Es el único aviso de este tipo que vas a leer.

Cómo usar este libro

Haz el capítulo 01 entero antes que nada: en menos de una hora tienes tus primeras facturas en una hoja que marca sola lo que no cuadra, y ves en datos reales el fallo que da título al libro. Después ve a lo que te haga falta:

Lo que quieres	Capítulo
Que la IA lea mejor tus escaneos, fotos y PDF	02
Elegir entre un chat de nube y un modelo en tu PC, y no regalar datos de clientes	03
Que lo extraído entre en Excel sin números convertidos en texto ni columnas movidas	04
Comprobar NIF, fechas, duplicados y numeración, en Excel o con un script	05
Organizar la revisión de lo marcado sin rehacerlo todo	06
Saber qué tienes que guardar, cuánto tiempo y si puedes tirar el papel	07

Qué necesitas

- Excel 2010 o posterior en español (las fórmulas usan punto y coma). Con Excel en inglés, el paquete trae cada fórmula traducida.
- Un chat con IA que acepte imágenes, o un modelo local con visión en Ollama (capítulo 03).
- Unas cuantas facturas de prueba. Para la primera vuelta, mejor de un proveedor que conozcas bien.
- Opcional, para los capítulos 03 y 05: Python 3.9 o posterior. No hace falta instalar ningún paquete.

El paquete digital incluye la hoja `hoja_comprobaciones.xlsx` con todas las fórmulas ya puestas y unas filas trampa para probarla, el prompt de extracción, las fórmulas en un CSV con su explicación, y dos scripts: `extraer_ollama.py` (extrae con un modelo local) y `validar_facturas.py` (aplica las comprobaciones a lo extraído), con datos ficticios para probarlos.

Contenido

Facturas y escaneos con IA, sin errores silenciosos . . . 2
De una carpeta de facturas, tickets y escaneos a una hoja de Excel que marca sola lo dudoso . . . 2
Alcance . . . 2
Cómo usar este libro . . . 2
Qué necesitas . . . 2
01 · La primera carpeta en una hoja que avisa sola 5
Lo que pasó con una factura y cuatro modelos . . . 5
Paso 1 (10 minutos): monta la hoja . . . 5
Paso 2 (15 minutos): pide los datos a un chat . . . 6
Paso 3 (5 minutos): pásalo a la hoja . . . 6
Paso 4 (15 minutos): las comprobaciones . . . 6
Paso 5 (10 minutos): prueba la hoja con trampas . . . 7
Cómo sabes que ha salido bien . . . 8
02 · Escaneos y fotos que la IA no tenga que adivinar 9
Qué separó la imagen buena de la mala . . . 9
Si escaneas . . . 9
Si fotografías con el móvil . . . 9
Si te llegan en PDF . . . 10
Los tickets . . . 10
Cómo compruebas que tus imágenes valen . . . 10
03 · Extraer con un chat de nube o con un modelo en tu PC 11
Lo que medimos . . . 11
Cómo compararlos . . . 11
Antes de subir facturas a la nube: los datos personales . . . 11
El camino local, paso a paso . . . 12
Cómo compruebas que el camino elegido te vale . . . 13
04 · Un formato de salida que no se rompa al llegar a Excel 14
CSV o JSON: cuándo usar cada uno . . . 14
Los cuatro detalles que rompen un CSV en Excel en español . . . 14
Cómo arreglar números que han entrado como texto . . . 14
El JSON que lee el validador . . . 15
Los CSV que lee el validador . . . 15
Cuando la factura no cabe en las columnas . . . 15
Cómo compruebas que el formato aguanta . . . 16
05 · Las comprobaciones: cuentas, NIF, duplicados, fechas y numeración 17
Por qué las cuentas las hace una fórmula y no la IA . . . 17

Las cuentas: total, IVA, retención y líneas . . . 17
El NIF: la letra que delata una cifra mal leída . . . 17
La fecha . . . 18
Duplicados . . . 18
La numeración: la comprobación que caza el número mal leído . . . 19
El mismo trabajo en Python . . . 19
Cómo compruebas que tus comprobaciones funcionan . . . 20
06 · La revisión humana: qué mirar en lo marcado y cómo cerrarlo 21
Cómo organizar la sesión . . . 21
Qué mirar según el motivo . . . 21
Lo que sí puede aportar la IA en la revisión . . . 22
Revisar también un poco de lo aprobable . . . 22
Cuando la hoja se equivoca . . . 22
Cerrar la tanda . . . 22
Cómo compruebas que la revisión funciona . . . 23
07 · Conservar las facturas: lo que pide Hacienda y lo que no arregla un escaneo 24
Qué tienes que guardar . . . 24
Durante cuánto tiempo . . . 24
Cómo: papel o electrónico . . . 25
¿Puedo tirar el papel después de escanearlo? . . . 25
Los tickets . . . 25
Los datos personales que guardas . . . 26
Cómo compruebas que tu archivo cumple . . . 26
Fuentes . . . 26

01 · La primera carpeta en una hoja que avisa sola

Quieres meter en Excel las facturas de un mes sin teclearlas y sin repasar cada importe a ojo. La IA las lee bastante bien. El problema es cuando se equivoca: lo hace sin avisar, con un dato que parece correcto. En una hora vas a montar una hoja que hace ese repaso por ti y deja en «REVISAR A MANO» solo las facturas que lo merecen.

Antes de empezar, mira por qué no basta con fiarse de que «cuadre».

Lo que pasó con una factura y cuatro modelos

En el laboratorio creamos una factura ficticia de Taller Norte S.L.: número F-2026-0187, base 1.240,50 €, IVA del 21 % de 260,51 €, total 1.501,01 € y tres líneas de 540,00 €, 420,50 € y 280,00 €. La preparamos en tres versiones: una imagen limpia, un escaneo torcido y con ruido, y una foto mala, pequeña, desenfocada y con sombra. Cuatro modelos tenían que devolver nueve datos de cada imagen y un programa los comparaba con la factura real, dato a dato.

Imagen	Claude Haiku 5.5	Claude Sonnet 5.5	gemma4:12b (local)	qwen3.5:9b (local)
Limpia	9/9	9/9	9/9	9/9
Escaneada	8/9	9/9	9/9	9/9
Foto mala	7/9	8/9	1/9	8/9

Los números esconden lo importante: cómo falló cada uno.

- **Sonnet** no leyó el número de factura en la foto mala y escribió «ILEGIBLE», que es lo que le pedimos cuando dudara. Ese fallo no hace daño: salta a la vista.
- **Haiku y qwen3.5:9b** leyeron «F-2026-0167» donde ponía «F-2026-0187». Sin avisar. Todo lo demás estaba bien y cuadraba, así que ninguna suma lo detecta.
- **gemma4:12b** se inventó una factura entera: número «P-2020001», otro cliente, «Mantenimiento de maquinaria», base 1.000,00 €, IVA 210,00 € y total 1.210,00 €. Base más IVA da el total: cuadra. No marcó nada como ilegible. Añadió fuera de los datos una nota sobre «inconsistencias visuales», que se pierde si solo copias la tabla.

Lo que delata a la factura inventada son sus líneas: sumaban 1.000 + 1.200 + 600 = 2.800 €, no la base de 1.000 €. Lo que delata al 0167 es algo que la IA no ha leído: el último número que te mandó ese proveedor. La hoja de este capítulo comprueba las dos cosas.

Paso 1 (10 minutos): monta la hoja

Abre un libro nuevo de Excel con tres hojas y llámalas exactamente `Facturas`, `Lineas` (sin tilde, porque las fórmulas la nombran) y `Proveedores`. Escribe estas cabeceras en la fila 1:

Hoja	Columnas, de la A en adelante
Facturas	archivo, numero, fecha, nif_emisor, emisor, base, iva_pct, iva, retencion, total
Lineas	archivo, concepto, unidades, precio, importe
Proveedores	nif, nombre, ultimo_numero

`archivo` es el nombre del archivo de la imagen y une cada factura con sus líneas. `iva_pct` va como 21, no como 0,21 ni 21 %. `retencion` es el importe retenido de IRPF; si no hay, 0.

En `Proveedores` escribe tus proveedores habituales: el NIF sin espacios ni guiones, el nombre y el número de la última factura que registraste de cada uno. Esta hoja es lo único que la IA no toca, y por eso vale tanto. Con tres o cuatro proveedores basta para empezar.

Paso 2 (15 minutos): pide los datos a un chat

Elige 5 o 10 facturas de proveedores que estén en tu hoja `Proveedores`. Abre un chat con IA que acepte imágenes, adjunta las facturas y pega este encargo:

```
Te adjunto facturas en imagen. Extrae sus datos en dos tablas CSV, sin comentarios antes ni después.

Reglas:
1. Copia lo que ves. No calcules, no completes y no corrijas nada, aunque te parezca que no cuadra.
2. Si un dato no se lee con seguridad, escribe ILEGIBLE en esa celda. Es mejor un ILEGIBLE que un dato
adivinado.
3. Separador de columnas: punto y coma (;). Importes con coma decimal y sin punto de miles: 1240,50. Sin
el símbolo €.
4. Fechas como dd/mm/aaaa. Tipo de IVA como número sin %: 21.
5. "archivo" es el nombre del archivo de la imagen. Si no hay retención, pon 0.
6. Una fila por factura en la tabla 1 y una fila por línea de concepto en la tabla 2.
7. Si algo no cabe en estas columnas (dos tipos de IVA, recargo, una nota a mano), no lo resumas:
escribe ILEGIBLE en el importe afectado y sigue.

Tabla 1, con esta cabecera:
archivo;numero;fecha;nif_emisor;emisor;base;iva_pct;iva;retencion;total

Tabla 2, con esta cabecera:
archivo;concepto;unidades;precio;importe

Devuelve cada tabla en su propio bloque de código.
```

Si las facturas llevan datos de clientes o de personas, lee antes el apartado de privacidad del capítulo 03. Para esta primera vuelta, usa facturas que te hayan enviado empresas.

Lee la respuesta entera antes de copiarla. Si el chat ha escrito algo fuera de las tablas («algunos importes no se ven bien», «he supuesto que…»), apúntalo: es un aviso como el que gemma4 dejó fuera de los datos.

Paso 3 (5 minutos): pásalo a la hoja

Copia el contenido del primer bloque, sin la cabecera, y pégalo en la celda A2 de `Facturas`. Si todo cae en la columna A, usa **Datos > Texto en columnas**, elige **Delimitados** y marca **Punto y coma**. Haz lo mismo con el segundo bloque en la celda A2 de `Lineas`.

Comprueba dos cosas antes de seguir:

- Los importes salen alineados a la derecha. Así muestra Excel los números; si salen a la izquierda son texto y las fórmulas fallarán. El capítulo 04 explica cómo arreglarlo.
- Las fechas también salen a la derecha. Una fecha a la izquierda es texto, y la hoja la mandará a revisión.

Paso 4 (15 minutos): las comprobaciones

Escribe estas fórmulas en la fila 2 de `Facturas`, cada una en su columna, y arrástralas hasta la última fila con datos. Están para Excel en español; la tolerancia de 0,01 absorbe los redondeos de céntimo.

```
K  Total = base + IVA - retención
=SI(ABS(F2+H2-I2-J2)<=0,01;"OK";"REVISAR: TOTAL")

L  IVA = base × tipo
=SI(ABS(REDONDEAR(F2*G2/100;2)-H2)<=0,01;"OK";"REVISAR: IVA")

M  Las líneas cuadran y suman la base
=SI(Y(ABS(SUMAR.SI(Lineas!$A$2:$A$1000;A2;Lineas!$E$2:$E$1000)-F2)<=0,01;CONTAR.SI.CONJUNTO(Lineas!$A$2:
$A$1000;A2;Lineas!$F$2:$F$1000;"REVISAR*")=0);"OK";"REVISAR: LINEAS")

N  Ningún dato ILEGIBLE
=SI(CONTAR.SI(B2:J2;"ILEGIBLE")=0;"OK";"REVISAR: ILEGIBLE")

O  No está dos veces
=SI(CONTAR.SI.CONJUNTO($D$2:$D$1000;D2;$B$2:$B$1000;B2)=1;"OK";"REVISAR: DUPLICADA")

P  El proveedor está en tu registro
=SI(CONTAR.SI(Proveedores!$A$2:$A$500;D2)>0;"OK";"REVISAR: NIF DESCONOCIDO")

T  Veredicto
=SI(CONTAR.SI(K2:S2;"OK*")=CONTARA(K2:S2);"APROBABLE";"REVISAR A MANO")
```

Y esta, en la columna F de `Lineas`, arrastrada hasta la última línea:

```
F  Unidades × precio = importe
=SI(ABS(C2*D2-E2)<=0,01;"OK";"REVISAR")
```

Las columnas Q, R y S quedan libres a propósito: ahí van la letra del NIF, la fecha y la numeración del capítulo 05. El veredicto ya las tiene en cuenta. Tampoco depende de que cada comprobación funcione: cuenta cuántas celdas dicen «OK» y las compara con cuántas tienen algo, así que una celda con error (por ejemplo, porque un importe es texto) también manda la fila a revisión.

Por qué el IVA se compara con REDONDEAR: 1.240,50 × 21 % da 260,505 y la factura pone 260,51. Sin redondear a dos decimales, la fórmula marcaría una factura correcta.

Para cazar el 0167 te falta una comprobación más, la de la columna S: que el número sea posterior al último registrado de ese proveedor. Pégala ya si quieres:

```
S  Número posterior al último registrado
=SI.ERROR(SI(--DERECHA(B2;4)>--DERECHA(BUSCARV(D2;Proveedores!$A$2:$C$500;3;FALSO);4);"OK";"REVISAR:
NUMERO");"REVISAR: NUMERO")
```

Compara las cuatro últimas cifras del número, así que vale para formatos como F-2026-0187. El capítulo 05 la explica y da la versión para otros formatos.

Paso 5 (10 minutos): prueba la hoja con trampas

Una hoja de comprobaciones también tiene errores, y el peor es el que aprueba lo que no debe. Antes de fiarte de ella, añade a mano cuatro filas falsas al final de `Facturas`, con sus líneas en `Lineas`:

1. Una que no cuadre: base 100, IVA 21, total 120.
2. Una copia exacta de otra fila (mismo NIF y mismo número).
3. Una con un NIF que no esté en `Proveedores`.
4. La factura inventada de gemma4: base 1000, IVA 210, total 1210, con tres líneas de 1000, 1200 y 600.

Las cuatro tienen que salir en «REVISAR A MANO». La primera, por la columna K; la segunda, por la O (en ella y en su original); la tercera, por la P; la cuarta, por la M. Si alguna sale «APROBABLE», revisa la fórmula de esa columna antes de seguir. Cuando estén bien, borra las filas trampa.

Cómo sabes que ha salido bien

- Cada factura real tiene un veredicto y puedes filtrar la columna T por «REVISAR A MANO».
- Las filas trampa salieron marcadas por la columna que tocaba.
- Para cada factura marcada sabes qué mirar: la columna que dice «REVISAR» te lleva al dato concreto.

Ahora abre la imagen de cada factura marcada y corrige la fila a mano. No vuelvas a pedírsela a la IA hasta haber leído el capítulo 02: si la imagen es mala, lo más probable es que vuelva a fallar, y quizá de otra forma.

Lo que sale «APROBABLE» no es una garantía, es una factura que ha pasado todas las comprobaciones que tienes. Con la hoja de este capítulo, un número mal leído que siga siendo posterior al último registrado pasaría. El capítulo 06 explica cuánto revisar también de lo aprobable y cómo.