



Hopalys PDF

Manual del usuario

Versión 1.3.2

© 2026 Tag & Code, Wilhelm Polmann. Todos los derechos reservados.

Índice

1	Introducción	4
2	Símbolos, notas y convenciones	4
3	Requisitos del sistema y requisitos previos	4
4	Contenido del paquete / alcance del suministro	5
5	Instalación	5
6	Versión de prueba / activación	5
7	Vista general de la ventana principal	7
8	Flujo de trabajo general	8
9	Formatos de entrada y salida compatibles	8
10	Instrucciones por tarea	9
10.1	Analizar archivos PDF	9
10.2	Exportar texto desde archivos PDF	10
10.3	Convertir archivos al formato PDF	10
	Dividir archivos PDF en lotes	12
10.4	Ejecutar acciones en archivos PDF	12
10.5	Convertir archivos PDF en imágenes	13
10.6	Crear archivos PDF con capacidad de búsqueda	13
11	Selección de archivos y comportamiento de guardado	14
11.1	Selección de documentos	14
11.2	Comportamiento de guardado	15
12	Configuración	16
12.1	Configuración general y configuración de informes	16
12.2	Configuración de exportación	16
12.3	Configuración de OCR	17
12.3.1	Añadir y descargar archivos de idioma OCR	18
13	Archivo de configuración	19
13.1	Configuración disponible a través de la interfaz	19
13.2	Configuración avanzada almacenada solo en el archivo de configuración	19
13.2.1	Información detallada sobre la configuración de TextRank	20
13.2.2	Información detallada sobre la configuración de rendimiento de OCR	21
13.2.3	Información detallada sobre la configuración fina del análisis	21
13.2.4	Información detallada sobre las opciones de resumen de texto	22
14	Menús y comandos adicionales	23
15	Resultados, informes y archivos de registro	23

Manual del usuario – Hopalys PDF

16	Función de línea de comandos de Hopalys PDF	24
17	Indicaciones para la traducción y la edición	25
18	Solución de problemas	25
19	Best practice / consejos para obtener resultados óptimos	27
20	Soporte y contacto	28
21	Nuevas funciones / notas de versión	28
22	Aviso legal	28

Nota sobre las versiones idiomáticas

Este manual fue redactado originalmente en inglés. Las demás versiones lingüísticas se generaron mediante traducción automática. Por ello, pueden existir pequeñas desviaciones.

Si tiene preguntas o alguna duda, puede ponerse en contacto con nosotros en: info@tag-code.com

1 Introducción

Hopalys PDF es una aplicación para Windows destinada a tareas con PDF con un flujo de trabajo claramente estructurado paso a paso. La interfaz de usuario guía al usuario a través del procesamiento en cinco pasos: seleccionar la tarea, seleccionar los archivos de origen, revisar la configuración, definir el comportamiento de guardado e iniciar el procesamiento.

Según la tarea seleccionada, los usuarios pueden analizar archivos PDF, exportar texto editable, convertir archivos de origen a PDF, dividir archivos PDF grandes, aplicar acciones sobre archivos PDF, generar archivos de imagen a partir de PDF o añadir texto con capacidad de búsqueda a PDF escaneados.

Este manual se centra en el uso práctico. Explica el flujo de trabajo dentro de la interfaz de usuario, los formatos de entrada y salida compatibles, la gestión de informes y la configuración relevante para el trabajo diario.

2 Símbolos, notas y convenciones

Nota: información útil que facilita una tarea o aclara el comportamiento esperado.

Importante: información que ayuda a evitar retrabajos, salidas faltantes o sobrescrituras accidentales.

Configuración avanzada: información de configuración destinada principalmente a usuarios avanzados e integradores de sistemas.



Los nombres de menús, los botones y las denominaciones visibles de las tareas se escriben tal como aparecen en la aplicación.

Las rutas de archivo se indican con la notación habitual de Windows, por ejemplo ProgramData>tagandcode\HopalysPDF\config.

3 Requisitos del sistema y requisitos previos

- Entorno de escritorio de Windows.
- Acceso de lectura a la carpeta de origen y acceso de escritura a la ubicación de trabajo utilizada para salidas, informes y registros.
- Un navegador web predeterminado para abrir informes HTML y resultados de análisis.
- Un visor de PDF y un procesador de textos para revisar archivos PDF y DOCX exportados.
- Al menos un paquete de idioma OCR activo para tareas basadas en OCR, como la creación de PDF con capacidad de búsqueda o la exportación de texto mediante OCR.
- Espacio libre suficiente para los archivos PDF, DOCX, XML, de imagen, de informe y de registro generados.

Para integradores de sistemas: en despliegues gestionados, asegúrese además de que la aplicación pueda crear o leer sus carpetas compartidas en ProgramData>tagandcode\HopalysPDF.

4 Contenido del paquete / alcance del suministro

El software se suministra mediante un enlace de descarga. La estructura exacta del paquete puede variar según el método de entrega, pero el paquete incluye claramente:

- Paquete de aplicación proporcionado por su proveedor o mediante su proceso interno de distribución de software.
- Soporte para paquetes de idioma OCR para flujos de trabajo basados en OCR.
- Información relacionada con la licencia o un archivo de licencia, según su proceso de entrega.
- Archivo de configuración XML opcional para despliegues preconfigurados o configuraciones de proyecto recurrentes.

Mantenga juntos todos los archivos suministrados hasta que se complete el primer inicio. Si su organización distribuye el software de forma centralizada, siga las directrices internas de instalación y seguridad aplicables a su entorno.

5 Instalación

1. Descargue el paquete de la aplicación desde el enlace proporcionado.
2. Si la descarga se proporciona como archivo comprimido, descomprímala en una carpeta local antes del primer uso.
3. Inicie el programa de instalación suministrado o el archivo de la aplicación de acuerdo con sus instrucciones de entrega.
4. Deje que la instalación o el despliegue finalicen por completo. En entornos controlados, este paso puede ser realizado por su departamento de TI o por su equipo de integración.
5. Inicie Hopalys PDF y continúe con la activación o con el uso de prueba si la aplicación se lo solicita.

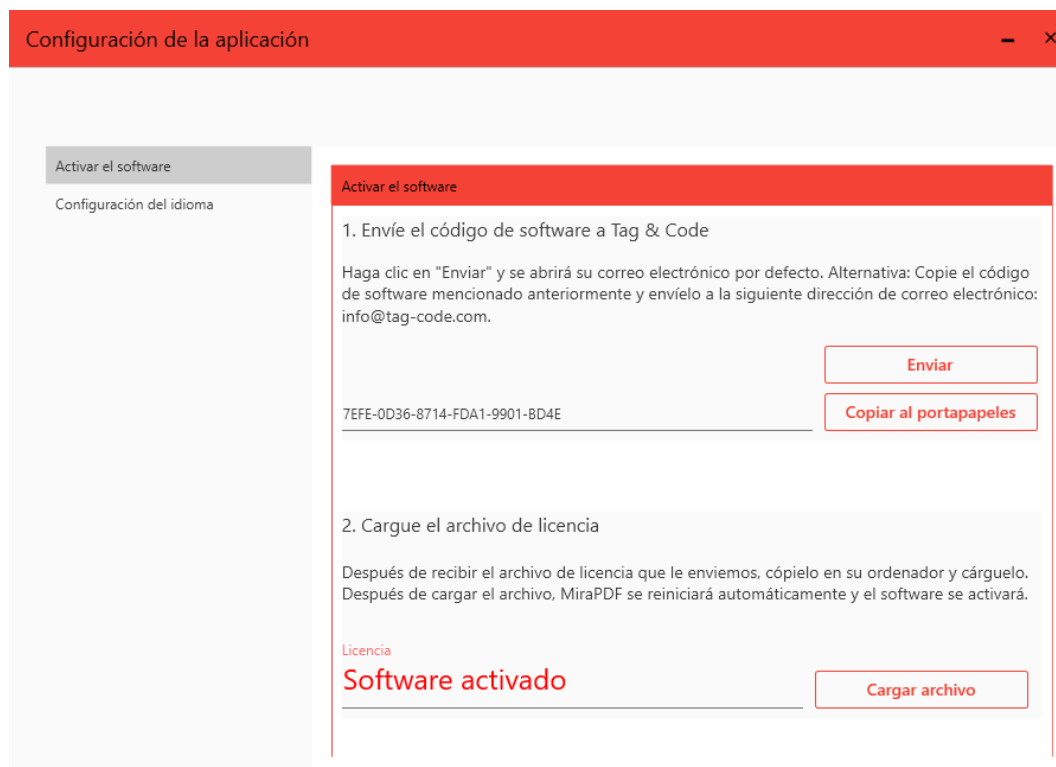
Para integradores de sistemas: la aplicación carga su configuración XML persistente desde ProgramData>tagandcode\HopalysPDF\config. Por ello, los archivos preconfigurados pueden prepararse como parte de un despliegue gestionado.

6 Versión de prueba / activación

Al iniciarse por primera vez, Hopalys PDF puede mostrar una pantalla de bienvenida, de prueba o de activación. Esta pantalla es una importante fuente de información sobre el producto y la licencia. Identifica la aplicación, muestra el estado actual de activación y proporciona el código de software utilizado para la asistencia de activación.

- El modo de prueba puede utilizarse gratuitamente para realizar pruebas iniciales.
- El flujo de trabajo de prueba permite un máximo de 10 archivos por ejecución.
- La pantalla de activación muestra el código de software necesario para la gestión de la licencia.

- Si su proveedor utiliza este método, se puede cargar un archivo de licencia desde la configuración de la aplicación.



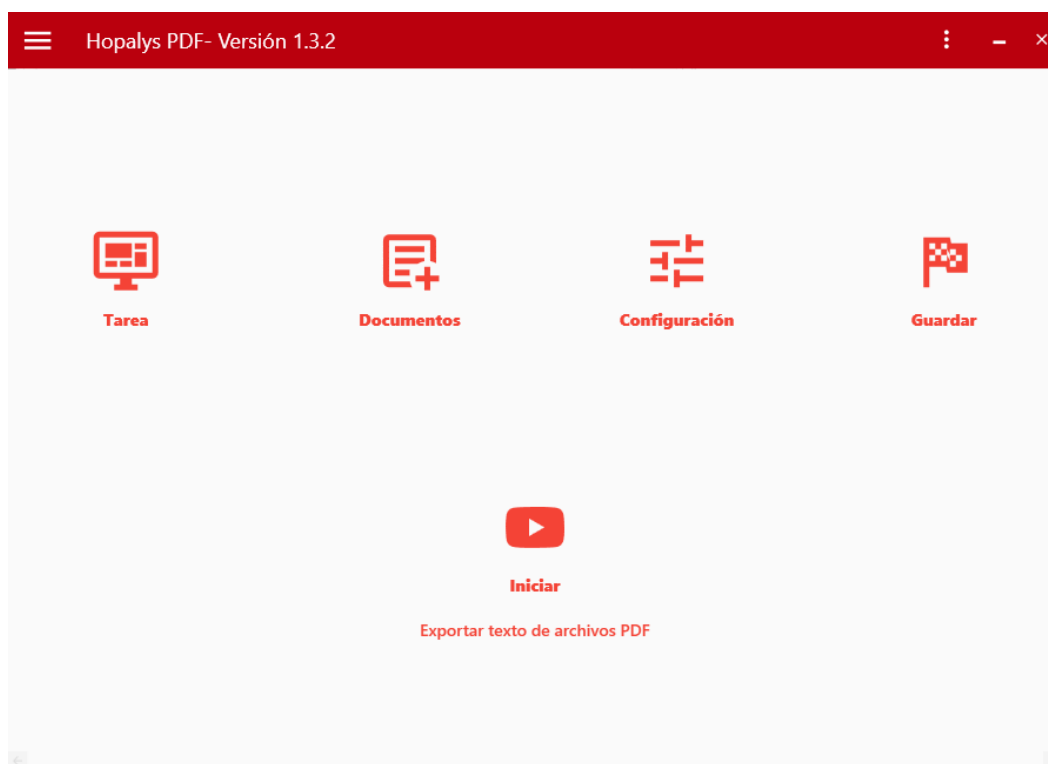
Configuración de activación y licencia. Tenga a mano el código de software mostrado cuando solicite ayuda para la activación.

Abra el menú de la aplicación y seleccione en el «menú de tres puntos» la entrada «Configuración de la aplicación» para acceder a las funciones relacionadas con la activación. Si la aplicación ya está activada, puede seguir utilizando esta misma área para comprobar o actualizar el estado actual de la licencia.

Importante: si la aplicación no está activada, pueden aplicarse restricciones de uso. En el modo de prueba, el límite es de 10 archivos por ejecución.

7 Vista general de la ventana principal

La ventana principal mantiene visible en todo momento el flujo de trabajo completo. Esta es una de las características de usabilidad más destacadas de Hopalys PDF, ya que permite a los usuarios seguir el proceso en un orden claro y repetible.



Ventana principal con el flujo de trabajo de cinco pasos y la vista general del proyecto actual.

Área	Uso
Tarea	Seleccione la tarea de procesamiento e introduzca la información básica del proyecto.
Documentos	Seleccione la carpeta de origen, revise la lista de archivos y elija los filtros de archivo relacionados con la tarea.
Configuración	Defina las opciones de procesamiento específicas de la tarea, las opciones de OCR y el comportamiento relacionado con PDF.
Guardar	Defina el comportamiento de guardado, la estructura de subcarpetas y el tratamiento asociado de la salida.
Iniciar	Ejecute la tarea seleccionada para los archivos marcados.

Área	Uso
Panel del proyecto actual	Antes de iniciar, revise la tarea seleccionada, las carpetas de trabajo y el estado actual del procesamiento.

8 Flujo de trabajo general

1. Seleccione la tarea deseada.
2. Seleccione la carpeta de origen y revise la lista de archivos.
3. Ajuste la configuración de la tarea, la configuración de OCR y las opciones de PDF que influyen en el resultado.
4. Revise el comportamiento de guardado, la ubicación del informe y las opciones de nomenclatura.
5. Inicie el procesamiento.
6. Abra la vista de resultados y revise el informe.
7. Revise los archivos de salida generados y entréguelos.



La misma estructura básica se aplica a todas las tareas principales. El formato exacto de entrada, el formato de salida y la salida secundaria opcional dependen de la tarea seleccionada.

9 Formatos de entrada y salida compatibles

Hopalys PDF utiliza un procesamiento de archivos específico según la tarea. La mayoría de las tareas trabajan directamente con archivos PDF de entrada. La tarea de conversión a PDF acepta tanto imágenes como archivos de documento. Los archivos de informe se tratan por separado de la salida principal y están disponibles después del procesamiento en la vista de resultados.

Tarea	Entrada compatible	Salida y salida secundaria opcional	Formato(s) del archivo de informe
Analizar archivos PDF	PDF	Resultados de clasificación y recomendaciones; opcionalmente, copias ordenadas de los PDF clasificados	Informe HTML más un informe de análisis HTML independiente
Exportar texto desde archivos PDF	PDF	DOCX con maquetación, DOCX mediante OCR y/o XML; archivo XFDF de comentarios opcional	Informe HTML

Tarea	Entrada compatible	Salida y salida secundaria opcional	Formato(s) del archivo de informe
Convertir archivos al formato PDF	JPG, JPEG, GIF, TIFF, PNG, DOCX, RTF	PDF; archivo PDF unido opcional al utilizar opciones de unión	Informe HTML
Dividir archivos PDF en lotes	PDF	Varios archivos PDF por lotes	Informe HTML
Ejecutar acciones en archivos PDF	PDF	Archivos PDF actualizados o renombrados, según las opciones de guardado	Informe HTML
Convertir archivos PDF en imágenes	PDF	PNG, JPG/JPEG, GIF	Informe HTML
Crear archivos PDF con capacidad de búsqueda	PDF	Archivos PDF con capacidad de búsqueda, normalmente guardados con el sufijo _searchable	Informe HTML

10 Instrucciones por tarea

10.1 Analizar archivos PDF

Objetivo: revisar archivos PDF y clasificarlos antes de la exportación o del OCR según su calidad, legibilidad e idoneidad para la planificación del flujo de trabajo.

Formatos de entrada compatibles: PDF

Formato(s) de salida: resultados de clasificación, recomendaciones y, opcionalmente, copias ordenadas de los PDF clasificados.

Formato(s) del archivo de informe: informe HTML y un informe de análisis HTML independiente.

Caso de uso típico: utilice esta tarea antes de un proyecto de exportación de mayor alcance cuando los PDF de origen difieran mucho en calidad o maquetación.

1. Seleccione «Analizar archivos PDF» en la lista de tareas.
2. Seleccione la carpeta de origen que contiene los PDF que se van a revisar.
3. Revise la lista de archivos y deje marcados únicamente los archivos que pertenecen a esta ejecución de análisis.
4. Revise todas las opciones relacionadas con el análisis. Los umbrales de clasificación avanzados también pueden estar guardados en el archivo de configuración.
5. Inicie la ejecución y espere hasta que el informe de análisis esté disponible.
6. Abra el informe de análisis y utilice las recomendaciones para decidir si los PDF de origen deben exportarse primero, procesarse mediante OCR o limpiarse.

Configuración avanzada: valores de configuración como los umbrales de superficie de imagen y de probabilidad de texto pueden influir en cómo se clasifican los resultados del análisis.

10.2 Exportar texto desde archivos PDF

Objetivo: crear contenido editable a partir de archivos PDF para revisión, traducción o edición.

Formatos de entrada compatibles: PDF

Formato(s) de salida: DOCX con maquetación, DOCX mediante OCR y/o XML, según la opción de exportación seleccionada. Opcionalmente, está disponible un archivo XFDF para la exportación de comentarios.

Formato(s) del archivo de informe: informe HTML.

Caso de uso típico: utilice esta tarea cuando necesite texto editable de documentos PDF y quiera mantener la maquetación o el soporte OCR.

1. Seleccione «Exportar texto desde archivos PDF».
2. Seleccione la carpeta que contiene los archivos PDF.
3. En la configuración de exportación, seleccione un modo de salida: «DOCX (conversión directa, manteniendo el diseño)», «DOCX con OCR» o «XML».
4. Si los PDF contienen páginas solo de imagen o contenido escaneado mixto, active el soporte OCR cuando sea necesario.
5. Opcional: active la exportación de comentarios, la copia de imágenes, la eliminación de imágenes del DOCX o el ajuste de tamaños de fuente muy pequeños.
6. Inicie la ejecución y revise los archivos DOCX, XML y XFDF opcionales generados.
7. Abra el informe para confirmar qué formato se generó para cada archivo procesado.

Importante: seleccione el modo de exportación antes de la ejecución. La aplicación utiliza el formato seleccionado para determinar qué archivos de resultado se escribirán.

10.3 Convertir archivos al formato PDF

Objetivo: crear archivos PDF a partir de imágenes o archivos de documento compatibles.

Formatos de entrada compatibles: JPG, JPEG, GIF, TIFF, PNG, DOCX, RTF

Formato(s) de salida: PDF. Según las opciones seleccionadas, los resultados individuales también pueden combinarse en un único archivo PDF unido.

Formato(s) del archivo de informe: informe HTML.

Caso de uso típico: utilice esta tarea para estandarizar material de origen mixto en formato PDF antes de la revisión, la entrega o el archivado.

1. Seleccione «Convertir archivos al formato PDF».
2. Seleccione la carpeta de origen y confirme el tipo de entrada esperado.

4. Si se necesita un resultado unido, active la opción de lote o de combinación adecuada para su flujo de trabajo.
5. Inicie la ejecución y revise la salida PDF generada.
6. Utilice el informe para confirmar qué archivos de origen se han convertido correctamente.

Dividir archivos PDF en lotes

Objetivo: crear paquetes PDF más pequeños a partir de PDF de origen de mayor tamaño.

Formatos de entrada compatibles: PDF

Formato(s) de salida: varios archivos PDF por lotes.

Formato(s) del archivo de informe: informe HTML.

Caso de uso típico: utilice esta tarea cuando sea necesario dividir archivos PDF grandes en partes manejables para su entrega, revisión o procesamiento posterior.

1. Seleccione «Dividir archivos PDF en lotes».
2. Seleccione la carpeta de origen que contiene los archivos PDF que deben dividirse.
3. Defina el número máximo de páginas por lote.
4. Defina el separador que debe utilizarse en los nombres de los archivos de salida si necesita un patrón de nomenclatura específico.
5. Inicie la ejecución y revise los PDF por lotes en la carpeta de resultados.

10.4 Ejecutar acciones en archivos PDF

Objetivo: aplicar acciones relacionadas con PDF, como gestión de contraseñas, actualización de metadatos, configuración de compresión, opciones de versión PDF y reglas de nomenclatura.

Formatos de entrada compatibles: PDF

Formato(s) de salida: archivos PDF actualizados. Según las opciones de nomenclatura y guardado, el resultado puede guardarse con un prefijo o un sufijo.

Formato(s) del archivo de informe: informe HTML.

Caso de uso típico: utilice esta tarea cuando necesite preparar archivos PDF para su distribución, cumplimiento normativo, limpieza de metadatos o cambio de nombre controlado.

1. Seleccione «Ejecutar acciones en archivos PDF».
2. Seleccione la carpeta de origen que contiene los PDF que deben modificarse.
3. Active únicamente las opciones de acción PDF que necesite, por ejemplo creación de contraseñas, campos de metadatos o configuración de compresión.
4. Revise antes de iniciar las opciones de nomenclatura, como prefijos o sufijos.
5. Inicie la ejecución y abra la carpeta de salida para confirmar los archivos PDF modificados.

Importante: la configuración de nomenclatura y guardado puede determinar si un PDF modificado se guarda como archivo nuevo o como parte de la estructura de resultados existente.

10.5 Convertir archivos PDF en imágenes

Objetivo: renderizar páginas PDF como archivos de imagen.

Formatos de entrada compatibles: PDF

Formato(s) de salida: archivos PNG, JPG/JPEG o GIF, según el formato seleccionado.

Formato(s) del archivo de informe: informe HTML.

Caso de uso típico: utilice esta tarea para la entrega de archivos PDF basada en imágenes.

1. Seleccione «Convertir archivos PDF en imágenes».
2. Seleccione la carpeta de origen que contiene los archivos PDF.
3. Seleccione el formato de imagen de salida requerido.
4. Inicie la ejecución y revise los archivos de imagen generados.

10.6 Crear archivos PDF con capacidad de búsqueda

Objetivo: añadir texto con capacidad de búsqueda mediante OCR a archivos PDF escaneados o basados en imágenes.

Formatos de entrada compatibles: PDF

Formato(s) de salida: archivos PDF con capacidad de búsqueda, normalmente guardados con el sufijo «_searchable» en el nombre.

Formato(s) del archivo de informe: informe HTML.

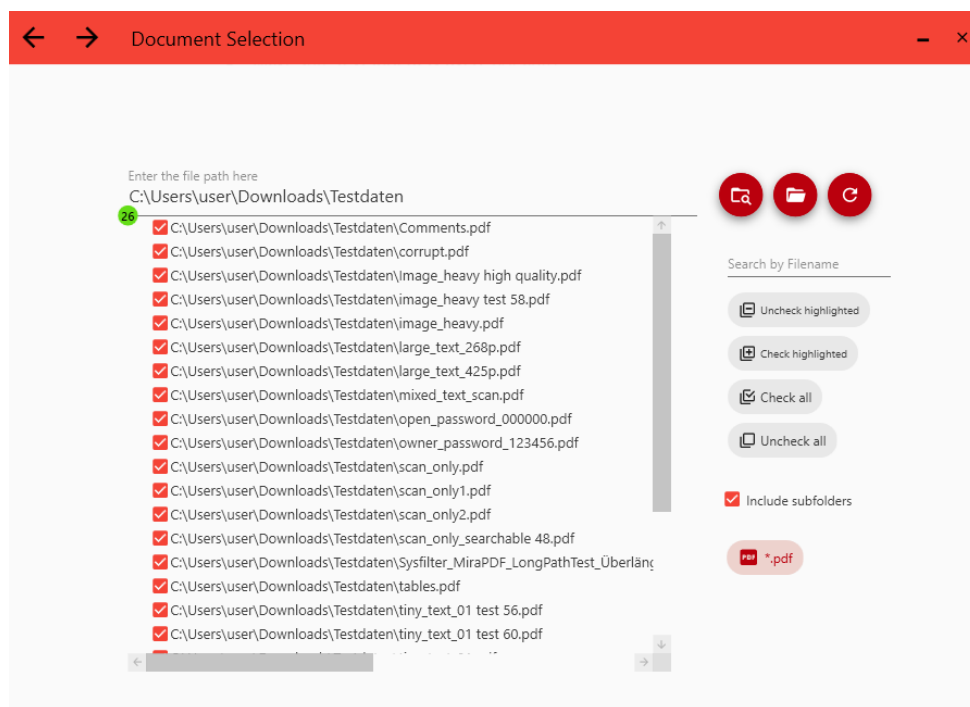
Caso de uso típico: utilice esta tarea cuando un PDF escaneado necesite texto con capacidad de búsqueda para buscar o copiar.

1. Seleccione «Crear archivos PDF con función de búsqueda».
2. Seleccione la carpeta de origen que contiene los PDF escaneados o basados en imágenes.
3. Inicie la ejecución y revise la salida PDF con capacidad de búsqueda.

11 Selección de archivos y comportamiento de guardado

11.1 Selección de documentos

La ventana de selección de documentos controla qué archivos se incluyen en la ejecución actual. La aplicación adapta automáticamente el filtro de archivos a la tarea seleccionada. Por ejemplo, las tareas basadas en PDF muestran archivos PDF, mientras que la tarea de conversión a PDF puede mostrar imágenes o archivos de documento compatibles.



Selección de documentos con carpeta de origen, lista de archivos, campo de búsqueda y opción para incluir subcarpetas.



Utilice el selector de carpetas para indicar la carpeta de trabajo.



Se utiliza para abrir el Explorador de Windows.

- Active o desactive archivos individuales o utilice, si están disponibles, las funciones para seleccionar todo o limpiar la selección.
- Active el procesamiento de subcarpetas solo cuando todos los archivos contenidos pertenezcan al mismo proceso.
- Revise la selección del tipo de archivo relacionada con la tarea cuando la pantalla ofrezca más de un tipo de entrada.

11.2 Comportamiento de guardado

El paso «Guardar» controla cómo organiza Hopalys PDF el conjunto de resultados. Además de la carpeta de trabajo/salida mostrada en la ventana principal, la aplicación también puede crear subcarpetas y copiar archivos relacionados en la estructura generada.

- Cree subcarpetas cuando necesite una estructura de resultados más clara para la entrega o la revisión.
- Copie los archivos originales en la estructura de subcarpetas generada solo cuando los archivos originales y los de resultado deban entregarse juntos.
- Utilice prefijos, sufijos o subcarpetas propias cuando necesite conservar varias ejecuciones de los mismos archivos de origen.
- Antes de iniciar, revise la carpeta de salida visible para saber dónde se escribirá el conjunto de resultados.



Configuración de guardado para carpetas de resultados, creación de subcarpetas y copia de archivos originales.

Importante: si vuelve a utilizar la misma ubicación de destino y el mismo patrón de nomenclatura, los archivos de resultado existentes pueden sustituirse. Utilice subcarpetas independientes o reglas de nomenclatura para conservar sin cambios las ejecuciones anteriores.

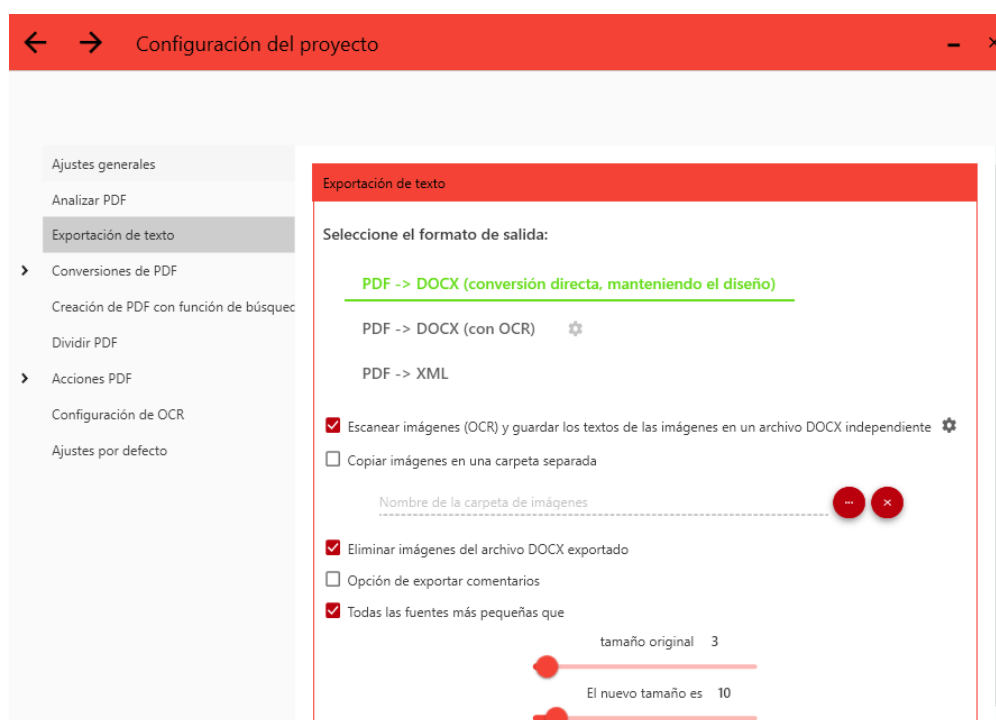
12 Configuración

Hopalys PDF agrupa la configuración por funciones para que la exportación, el OCR, las acciones PDF y el comportamiento de guardado sigan siendo fáciles de comprender. En el uso diario, esto reduce el riesgo de mezclar opciones no relacionadas.

12.1 Configuración general y configuración de informes

- El nombre del proyecto, el ID del proyecto y la fecha del proyecto ayudan a identificar claramente la ejecución en el informe.
- Puede definirse una ruta de informe cuando los informes deban guardarse en una ubicación conocida.
- Puede especificarse una contraseña de documento si los documentos de origen protegidos así lo requieren.
- «Restablecer la configuración por defecto» está disponible si desea volver a la configuración estándar.

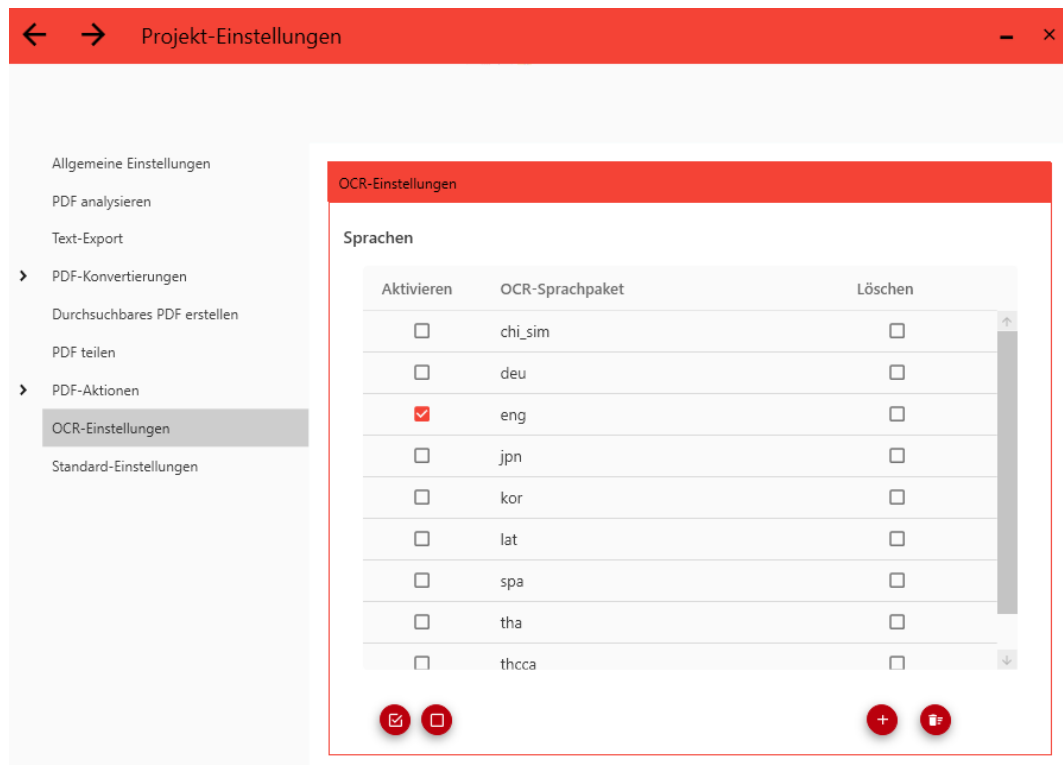
12.2 Configuración de exportación



Configuración de exportación para la exportación de texto.

- Seleccione un destino de exportación: DOCX with layout, DOCX via OCR o XML.
- Active OCR para las imágenes dentro de los PDF si también debe extraerse texto de las zonas escaneadas.
- Exporte comentarios cuando se requiera un archivo XFDF de comentarios independiente.
- Copie las imágenes incrustadas a una carpeta o elimínelas de la salida DOCX cuando su flujo de trabajo lo requiera.
- Ajuste tamaños de fuente muy pequeños en la salida DOCX cuando la legibilidad sea más importante que una reproducción exacta del tamaño.

12.3 Configuración de OCR



Configuración de OCR con gestión activa de paquetes de idioma.

- Active únicamente los paquetes de idioma necesarios para los documentos actuales.
- Utilice «Añadir entrada» para registrar archivos adicionales de paquetes de idioma OCR, cuando estén disponibles.
- Los archivos de paquetes de idioma OCR se gestionan como recursos traineddata en la carpeta compartida de OCR.

Importante: desactive los idiomas no utilizados para que las ejecuciones con mucho OCR sigan siendo más rápidas.

12.3.1 Añadir y descargar archivos de idioma OCR

Las tareas basadas en OCR requieren archivos válidos de paquetes de idioma OCR. Estos archivos se gestionan como recursos .traineddata y aparecen en la lista de configuración de OCR. Utilice esta sección cuando falte un idioma necesario en la lista o cuando deba prepararse un idioma OCR adicional para un proyecto.

Dónde descargar archivos de idioma OCR. Descargue archivos de idioma OCR exclusivamente de las fuentes oficiales de Tesseract. El resumen oficial de archivos de datos enumera los repositorios y paquetes de idioma disponibles.

- Resumen oficial de Tesseract sobre archivos de datos: <https://tesseract-ocr.github.io/tessdoc/Data-Files.html><https://tesseract-ocr.github.io/tessdoc/Data-Files.html>
- tessdata_fast: recomendado cuando la velocidad de procesamiento es más importante que la máxima precisión de OCR. https://github.com/tesseract-ocr/tessdata_fasthttps://github.com/tesseract-ocr/tessdata_fast
- tessdata_best: recomendado cuando la mayor precisión posible de OCR es más importante que la velocidad de procesamiento. https://github.com/tesseract-ocr/tessdata_besthttps://github.com/tesseract-ocr/tessdata_best
- tessdata: utilice este repositorio cuando, además del OCR basado en LSTM, también se requiera compatibilidad con flujos de trabajo de motores OCR antiguos. <https://github.com/tesseract-ocr/tessdata><https://github.com/tesseract-ocr/tessdata>

Cómo añadir un nuevo archivo de idioma OCR. En el área «Configuración de OCR», utilice la función «Añadir entrada» para registrar un archivo de idioma descargado.

1. Abra «Configuración» y vaya a «Configuración de OCR».
2. Descargue el archivo .traineddata necesario desde uno de los repositorios oficiales de Tesseract.
3. Mantenga sin cambios el nombre original del archivo, por ejemplo eng.traineddata, deu.traineddata o fra.traineddata.
4. Haga clic en «Añadir entrada» o en el botón más del área de Configuración de OCR.
5. Vaya al archivo .traineddata descargado y confirme la selección.
6. Compruebe si el nuevo idioma aparece en la lista de idiomas OCR.
7. Active el nuevo paquete de idioma para el proyecto actual y desactive los paquetes de idioma no utilizados.

Ubicación de almacenamiento. Los archivos de paquetes de idioma OCR se gestionan en la carpeta compartida ProgramData\tagandcode\HopalysPDF\ocrLanguages.

Importante. Utilice únicamente archivos .traineddata fiables y compatibles. Si los resultados de OCR son incompletos o deficientes, compruebe primero si se ha añadido y activado el paquete de idioma correcto para los documentos actuales.

13 Archivo de configuración

Hopalys PDF guarda la configuración persistente en un archivo de configuración XML. La aplicación carga este archivo al iniciarse y también permite a los usuarios guardar o cargar archivos de configuración desde el menú. Por ello, el archivo XML resulta especialmente útil para proyectos recurrentes, despliegues controlados e integración de sistemas.

Ubicación o archivo	Finalidad
ProgramData\tagandcode\HopalysPDF\config	Carpeta compartida predeterminada para el archivo de configuración XML.
Hopalys PDF-Config<version>.xml	Archivo de configuración relacionado con la versión para guardar ajustes persistentes.
ProgramData\tagandcode\HopalysPDF\ocrLanguages	Carpeta compartida para archivos de paquetes de idioma OCR.
ProgramData\tagandcode\HopalysPDF\logs	Carpeta compartida para archivos de registro técnicos.

13.1 Configuración disponible a través de la interfaz

- Tarea seleccionada e información del proyecto.
- Carpeta de origen, archivos seleccionados y comportamiento al incluir subcarpetas.
- Configuración visible de salida y guardado, incluida la creación de subcarpetas.
- Ruta del informe y campo general de contraseña.
- Selección de formatos de exportación como DOCX, OCR-DOCX o XML.
- Activación de paquetes de idioma OCR.
- Configuración de acciones PDF, como contraseñas, metadatos, versión, compresión, PDF/A y reglas de nomenclatura.

13.2 Configuración avanzada almacenada solo en el archivo de configuración

El modelo de configuración contiene valores adicionales que no se muestran por completo en la interfaz visible. Estos ajustes son relevantes principalmente para usuarios avanzados e integradores de sistemas.

Grupo	Ajuste(s)	Efecto práctico
Rendimiento de OCR	MaxParallelOcrFiles	Limita cuántos trabajos de OCR pueden ejecutarse en paralelo. Resulta útil cuando es necesario equilibrar el rendimiento y la carga del sistema.

Grupo	Ajuste(s)	Efecto práctico
Ajuste fino del análisis	ImageTextProbabilityThreshold, ImageWideAspectRatio, ImageDocumentAspectRatio, LargeImageArea, SmallImageArea, LargeCoverageThreshold, FullPageCoverageThreshold	Ajuste fino de cómo la tarea de análisis evalúa páginas con mucha imagen y clasifica contenidos PDF.
Opciones de resumen de texto	EnableTextSummary, SummaryMaxPages, SummarySentenceCount, SummaryMinWordCount	Controla si el proceso de análisis genera resúmenes breves y cuál puede ser su extensión.
Detalles del ranking de texto	TextRankDampingFactor, TextRankMaxIterations, TextRankMinDiff, TextRankMinSentenceLengthLatin, TextRankMinSentenceLengthCjk, TextRankMinSimilarityThreshold, TextRankMaxSentences	Ajuste fino avanzado de la generación de resúmenes para entornos de revisión controlados.

Configuración avanzada: cree una copia de seguridad del archivo de configuración antes de editarlo manualmente. Si una configuración modificada provoca un comportamiento inesperado, restaure la copia de seguridad o restablezca la configuración a los valores estándar desde la aplicación.

13.2.1 Información detallada sobre la configuración de TextRank

Los siguientes ajustes permiten un control más preciso de la generación de resúmenes basada en TextRank y están pensados para escenarios avanzados de revisión y ajuste.

Ajuste	Efecto práctico
TextRankDampingFactor	Controla hasta qué punto las puntuaciones de las frases se influyen entre sí durante la clasificación. Los valores más altos suelen hacer que el ranking dependa más de la interconexión entre las frases.
TextRankMaxIterations	Define el número máximo de iteraciones de clasificación. Aumente este valor si la convergencia se produce demasiado pronto; redúzcalo para disminuir el tiempo de procesamiento.
TextRankMinDiff	Define el cambio mínimo de puntuación necesario para continuar con nuevas iteraciones. Los valores más pequeños permiten una convergencia más precisa, pero pueden tardar más.
TextRankMinSentenceLengthLatin	Longitud mínima de frase para sistemas de escritura latinos antes de que una frase se tenga en cuenta para la clasificación.

Ajuste	Efecto práctico
TextRankMinSentenceLengthCjk	Longitud mínima de frase para idiomas CJK antes de que una frase se tenga en cuenta para la clasificación.
TextRankMinSimilarityThreshold	Similitud mínima necesaria para que dos frases se conecten en el grafo de clasificación. Los valores más altos crean vínculos más estrictos.
TextRankMaxSentences	Número máximo de frases que se devuelven en el resumen extraído.

Nota: CJK significa chino, japonés y coreano.

13.2.2 Información detallada sobre la configuración de rendimiento de OCR

El siguiente ajuste controla la paralelización de OCR y está dirigido a usuarios avanzados que necesitan equilibrar rendimiento, capacidad de respuesta y carga del sistema.

Ajuste	Efecto práctico
MaxParallelOcrFiles	Define cuántas operaciones de OCR pueden ejecutarse al mismo tiempo. Los valores más altos pueden mejorar el rendimiento en sistemas potentes, pero también aumentan el uso de CPU, memoria y disco. Los valores más bajos reducen la carga del sistema y pueden mejorar la estabilidad en equipos más pequeños o muy cargados.

13.2.3 Información detallada sobre la configuración fina del análisis

Los siguientes ajustes influyen en cómo el proceso de análisis evalúa el contenido visual, la estructura de la página y la cobertura de imagen al clasificar páginas PDF.

Ajuste	Efecto práctico
ImageTextProbabilityThreshold	Define el umbral utilizado para estimar si una zona de la página representa más bien texto o imagen. Los valores más altos hacen que el análisis sea más conservador antes de clasificar el contenido como relacionado con texto.
ImageWideAspectRatio	Define el umbral de relación de aspecto para identificar imágenes muy anchas. Esto puede ayudar a distinguir elementos tipo banner o panorámicos de objetos documentales más típicos.
ImageDocumentAspectRatio	Define el rango esperado de relación de aspecto para contenidos de imagen similares a un documento. Esto ayuda al análisis a decidir si una imagen se parece a una página escaneada, a un documento incrustado o a una ilustración no relacionada con documentos.
LargeImageArea	Define el umbral de superficie a partir del cual una imagen se considera grande. Esto influye en la medida en que los elementos visuales grandes afectan a la clasificación de la página.

Ajuste	Efecto práctico
SmallImageArea	Define el umbral de superficie por debajo del cual una imagen se considera pequeña. Los elementos pequeños, como iconos o gráficos menores, pueden tratarse entonces de forma distinta a los contenidos dominantes de la página.
LargeCoverageThreshold	Define cuándo la superficie de imagen combinada en una página es lo bastante grande como para influir significativamente en la clasificación. Los valores más altos requieren una mayor cobertura visual antes de que una página se considere cargada de imágenes.
FullPageCoverageThreshold	Define cuándo la cobertura de imagen es lo bastante alta como para tratar la página como contenido de imagen casi a página completa. Esto resulta útil para detectar páginas escaneadas o páginas dominadas por una única imagen grande.

13.2.4 Información detallada sobre las opciones de resumen de texto

Los siguientes ajustes controlan si se generan resúmenes durante el análisis y cuánto contenido puede considerarse para ello.

Ajuste	Efecto práctico
EnableTextSummary	Activa o desactiva la generación automática de resúmenes durante el análisis. Si esta función está desactivada, no se genera ningún texto de resumen.
SummaryMaxPages	Limita cuántas páginas se tienen en cuenta para la generación de resúmenes. Esto ayuda a mantener predecible la creación del resumen y evita un procesamiento excesivo en documentos muy grandes.
SummarySentenceCount	Define el número máximo de frases que pueden incluirse en el resumen generado. Los valores más altos producen resúmenes más detallados, mientras que los más bajos mantienen el resultado más breve y fácil de revisar.
SummaryMinWordCount	Define la cantidad mínima de texto que debe existir antes de generar un resumen. Esto evita crear resúmenes para documentos o rangos de páginas con muy poco texto útil.

14 Menús y comandos adicionales

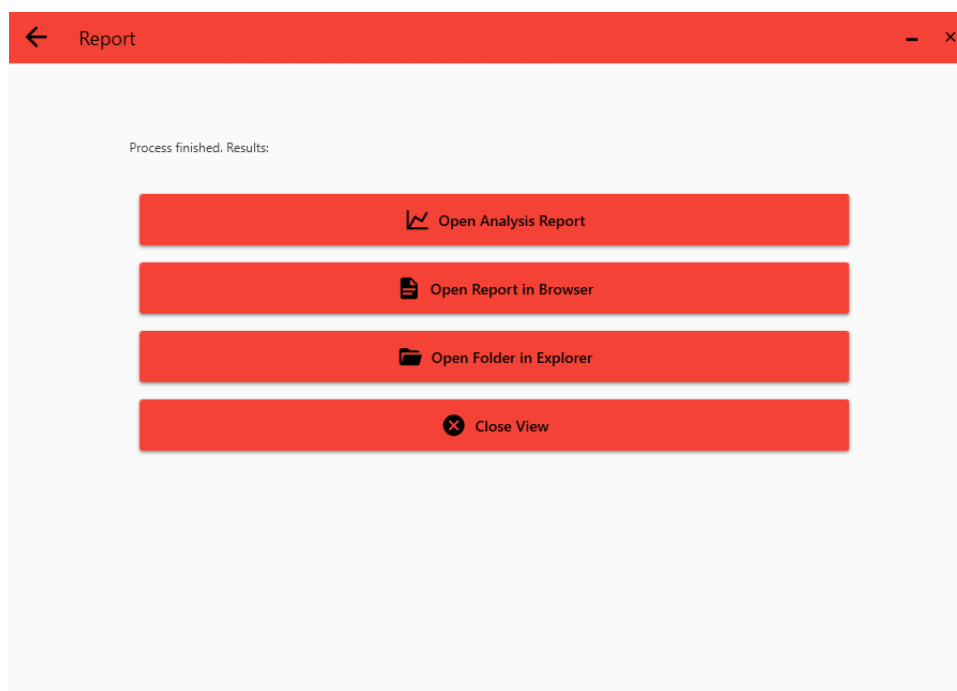
Además del flujo de trabajo principal de cinco pasos, Hopalys PDF ofrece comandos de menú para configuración y funciones relacionadas con la aplicación.

Área del menú	Funciones disponibles
Menú hamburguesa	Guardar configuración, cargar configuración, salir.
Más / menú de la aplicación	Configuración de la aplicación, Soporte, Acerca de Hopalys PDF.
Configuración de la aplicación	Activar el software, Configuración del idioma.

Utilice «Guardar» y «Cargar» cuando quiera reutilizar la misma configuración para proyectos recurrentes. Esto resulta especialmente útil en entornos de producción y para integradores de sistemas que necesitan una base de configuración controlada.

15 Resultados, informes y archivos de registro

Después del procesamiento, la ventana de resultados ofrece acceso directo a la salida de la ejecución. Este es el lugar preferido para comprobar el proceso antes de la entrega o de una edición posterior.



Ventana de resultados con acceso directo al informe, al informe de análisis y a la carpeta de salida.

- Informe de ejecución: informe HTML legible en el navegador, disponible desde la vista de resultados.
- Informe de análisis: informe HTML independiente para la tarea «Analizar archivos PDF».
- Archivos de salida: se crean de acuerdo con la tarea seleccionada y la configuración.

- Registros técnicos: se guardan en ProgramData\tagandcode\HopalysPDF\logs para soporte y diagnóstico.

El informe es el primer lugar para confirmar si se han procesado los archivos previstos y si la configuración seleccionada se ajusta al resultado entregado. Según la tarea, el informe también puede incluir métricas del documento y detalles de los archivos procesados relevantes para la revisión.

Para integradores de sistemas: la aplicación escribe archivos de registro continuos en la carpeta compartida de registros situada en ProgramData. Estos registros son útiles para diagnosticar problemas relacionados con despliegues, permisos o el entorno.

16 Función de línea de comandos de Hopalys PDF

Hopalys PDF puede iniciarse desde la línea de comandos pasando un archivo XML de configuración con /SettingsPath="...". Al iniciarse, la aplicación carga el archivo XML indicado y aplica la configuración almacenada en él para esa ejecución.

Procesamiento por línea de comandos con perfiles XML

Cada archivo XML es un perfil de procesamiento completo para una ejecución por línea de comandos. Define:

- la tarea que debe ejecutarse
- la carpeta de entrada
- el tipo de entrada esperado
- la configuración de OCR
- todas las demás opciones de procesamiento específicas de la tarea

Las distintas ejecuciones por línea de comandos se crean utilizando diferentes perfiles XML.

Configuración importante del perfil XML para el uso en la línea de comandos

- «SelectedTaskValue» define la tarea de Hopalys PDF utilizada para la ejecución.
- «InputFolder» define la carpeta que debe procesarse.
- Las opciones específicas de la tarea definen el resultado del procesamiento, por ejemplo «ConvertToJpg», «CreateSearchablePdfFile» o «OutputFormatPDF2DOCXviaOCR».
- «OCRLanguageFolder» y «OCRLanguagePackages» definen la ruta de idioma OCR y los idiomas OCR disponibles cuando el procesamiento OCR está activo.

Ejemplos para la línea de comandos:

```
"C:\Program Files\tagandcode\HopalysPDF\HopalysPDF.exe"  
/SettingsPath="D:\Config\Config_CL_Analyse.xml"  
"C:\Program Files\tagandcode\HopalysPDF\HopalysPDF.exe"  
/SettingsPath="D:\Config\Config_CL_ConvertPDF2Images_JPG.xml"  
"C:\Program Files\tagandcode\HopalysPDF\HopalysPDF.exe"  
/SettingsPath="D:\Config\Config_CL_Convert2PDF.xml"
```

Notas prácticas

- La ruta situada después de /SettingsPath debe ir entre comillas si contiene espacios.

- Actualice InputFolder en el archivo XML antes de iniciar Hopalys PDF si debe procesarse una carpeta de origen distinta.
- Los perfiles basados en OCR requieren un OCRLanguageFolder válido y entradas de idioma válidas en OCRLanguagePackages.
- Utilice un perfil XML independiente para cada flujo de trabajo recurrente, de modo que la configuración siga siendo coherente y fácil de reutilizar.

17 Indicaciones para la traducción y la edición

Cuando Hopalys PDF crea salidas DOCX o XML editables, estos archivos suelen convertirse en archivos de trabajo para traducción, edición o control de calidad. Trátelos como copias de trabajo controladas y no como sustitutos del PDF de origen.

- Trabaje con copias de los archivos DOCX o XML generados cuando se requiera un proceso formal de revisión.
- Siempre que sea posible, conserve los archivos de trabajo exportados en una carpeta separada de los PDF originales.
- No cambie el nombre de archivos secundarios relacionados, como un archivo XFDF de comentarios, ni los separe entre sí antes de que el proceso haya concluido.
- Revise la maquetación final después de la edición o la traducción, especialmente si el resultado va a devolverse a clientes o a utilizarse para impresión o salida PDF.
- Utilice el PDF de origen y, cuando esté disponible, una salida de referencia basada en PDF para la corrección y la comparación.

Importante: revise siempre el resultado editado antes de la entrega, ya que la maquetación, los saltos de línea, las fuentes y la colocación de las imágenes pueden cambiar en ediciones posteriores.

18 Solución de problemas

Problema	Causa probable	Qué hacer
No se muestran archivos en la lista de archivos	La tarea seleccionada o el tipo de archivo no coinciden con el contenido de la carpeta.	Revise la tarea seleccionada, el filtro de archivos y la carpeta de origen. Active el procesamiento de subcarpetas solo cuando sea necesario.
El proceso no se inicia	No hay archivos marcados o falta una configuración necesaria.	Revise la lista de archivos, marque los archivos deseados y confirme la configuración específica de la tarea antes de volver a iniciar.

Problema	Causa probable	Qué hacer
Se ha alcanzado el límite de prueba	La aplicación sigue en modo de prueba y se han seleccionado demasiados archivos.	Active la aplicación o reduzca la ejecución a un conjunto más pequeño de archivos.
El resultado de OCR es incompleto o deficiente	Los paquetes de idioma OCR necesarios no están activos o la calidad del escaneo es baja.	Active los paquetes de idioma correctos y compruebe la calidad del escaneo de origen.
No se abre ningún informe después de la ejecución	La ruta del informe o el paso de apertura en el navegador han fallado.	Revise la ubicación configurada del informe, los permisos de archivo y el navegador predeterminado.
Los archivos de salida se guardan en una ubicación inesperada	La configuración de guardado, las subcarpetas o las reglas de nomenclatura han cambiado la ruta del resultado.	Revise el paso «Guardar», la carpeta de salida visible y cualquier ajuste activo de prefijo, sufijo o subcarpeta.
Se han sobrescrito archivos existentes	Se ha vuelto a utilizar la misma ubicación de salida y el mismo patrón de nomenclatura.	Utilice subcarpetas independientes o reglas de nomenclatura para ejecuciones repetidas.
No se puede completar la activación	El código de software, el archivo de licencia o la información de soporte están incompletos.	Abra Configuración de la aplicación > Activar el software y contacte con el soporte utilizando el código de software mostrado por la aplicación.
La aplicación se comporta de forma inesperada después de un cambio en la configuración	El archivo de configuración contiene valores inadecuados o desactualizados.	Restaure una copia de seguridad o restablezca la configuración a los valores estándar y vuelva a intentarlo.

19 Best practice / consejos para obtener resultados óptimos

- Trabaje con copias cuando trate archivos importantes de clientes u originales escaneados.
- Utilice primero la tarea de análisis cuando los PDF de origen varíen mucho en calidad.
- Elija cuidadosamente el formato de exportación: DOCX basado en maquetación para PDF estructurados, DOCX basado en OCR para PDF con mucho contenido escaneado y XML solo cuando su flujo de trabajo posterior lo requiera.
- Active únicamente los paquetes de idioma OCR que correspondan al idioma de origen.
- Conserve los archivos de origen, los archivos de trabajo editables, los informes y los resultados entregados en carpetas claramente separadas.
- Utilice la creación de subcarpetas, prefijos o sufijos cuando deban conservarse varias ejecuciones del mismo conjunto de datos de origen.
- Revise el informe después de cada ejecución, incluso cuando la salida parezca correcta a primera vista.

20 Soporte y contacto




Cuando se ponga en contacto con el soporte, facilite suficiente información para que el problema pueda reproducirse rápidamente. Los archivos de informe y registro suelen ser más útiles que una breve descripción verbal por sí sola.

- Nombre del producto y versión instalada, tal como aparecen en la barra de título o en la ventana «Acerca de Hopalys PDF».
- La tarea seleccionada y la configuración más importante utilizada para la ejecución.
- El informe HTML y, cuando esté disponible, el informe de análisis HTML.
- Un archivo de ejemplo con el que pueda reproducirse el problema.
- Una breve descripción del resultado esperado y del resultado real.
- Archivos de registro relevantes de ProgramData>tagandcode\HopalysPDF\logs si el problema está relacionado con el entorno o aparece de forma esporádica.

La información de soporte relacionada con la activación es visible en el proceso de activación de la aplicación. Los textos de activación remiten además a la dirección de contacto info@tag-code.com para la comunicación relacionada con licencias.

21 Nuevas funciones / notas de versión

Revise la barra de título de la aplicación y la ventana «Acerca de Hopalys PDF» para determinar la versión exacta instalada. Las funciones, las entradas de menú y las opciones de licencia pueden cambiar entre versiones. Por ello, este manual describe el paquete de la aplicación y las funciones visibles disponibles en él.

-  Revise la información específica de la versión proporcionada por su proveedor, si está disponible.
-  Revise el archivo de configuración después de las actualizaciones, ya que los nombres de los archivos de configuración y los valores guardados pueden depender de la versión.
-  Vuelva a comprobar los paquetes de idioma OCR y las rutas de almacenamiento después de cambios importantes en el entorno o de despliegues de nuevas versiones.

22 Aviso legal

Tag & Code, Wilhelm Polmann
Dorfstr. 8, 86459 Gessertshausen
Alemania

Correo electrónico: info@tag-code.com

Este manual fue redactado originalmente en inglés. Las demás versiones lingüísticas se generaron mediante traducción automática. Pueden producirse pequeñas diferencias de redacción o de interpretación. En caso de duda, prevalece la versión en inglés.

Los nombres de productos, nombres de empresas y signos distintivos relacionados utilizados en la aplicación y en su entorno pueden estar protegidos por sus respectivos titulares.