# Pseudalis

*Use y comparta sus documentos sin exponer los datos personales que contienen*

Pseudalis detecta los datos personales de sus documentos y sustituye cada valor por una etiqueta que lo contiene en forma cifrada. Solo la clave de su equipo puede restaurar el documento, o cualquier texto derivado de él.

Versión **1.0.11** Plataforma **Windows 11, x64, sin controlador**

Licencia **Boost Software License 1.0**

Ejemplo *Con etiquetas*

```
Asunto: renovación del contrato

Estimado [[PERSON:Rnifypn Eouiitlrwky#EKR]]:

Tras nuestra llamada del [[DATE:28/08/2018#VLC]], le
remito la renovación para [[ORG:Xkekwmufwmcj Eouiitlrwky
et Wldwdst SARL#T4N]], [[LOCATION:Saint-Ntaahgp#QWB]].
El reembolso se abonará en la cuenta [[IBAN:GB45 NWBK 6152
7186 9288 16#K9P]].

Puede localizarme en el [[PHONE:06 07 09 00 95#M2X]] o en
[[EMAIL:rnifypn.eouiitlrwky@pspbshklkspj.com#7HD]].
```

**0,946**
F1 del detector en un conjunto de prueba en cinco idiomas (versión 1.0.11)

**41 000**
caracteres leídos por segundo con dos núcleos

**40**
países cuyos números nacionales se verifican

**17 MB**
de instalador que lo incluye todo, sin controlador

## En resumen

### Una etiqueta lleva su valor
Cada valor se convierte en una etiqueta que lo contiene, cifrado con FF1 sobre AES-256, con una comprobación. No se conserva ninguna tabla de correspondencias.

### Con o sin etiquetas
Los valores pueden escribirse solo con sus sustitutos, de modo que el archivo se lee como un documento corriente para un prompt, datos de prueba o un envío al exterior.

### Términos y columnas impuestos
Los administradores imponen hasta 100 términos y 100 columnas de archivos CSV, sustituidos antes de cualquier otra lectura, de modo que las uniones entre tablas exportadas se mantienen.

### Imágenes descritas por Albireo
Con Albireo en el mismo equipo, cada imagen se sustituye por una descripción que no nombra a ninguna persona.

### Los documentos conservan su formato
Los archivos Word, Excel, PowerPoint, PDF, de texto, CSV, Markdown, HTML, XML, JSON y YAML conservan su formato, y un PDF conserva sus páginas.

### Nombres y propiedades
Los nombres de los archivos, las entradas de los archivos comprimidos y las propiedades de los documentos reciben también sustitutos, y vuelven en la restauración.

### Originales guardados cifrados
Los originales van a un almacén cifrado, acotado en el tiempo y en el espacio, y vuelven a petición desde cualquier copia del archivo anonimizado.

### Sustitutos coherentes
El mismo valor recibe en todas partes el mismo sustituto, un IBAN protegido sigue superando su comprobación, y las fechas se desplazan juntas con un desfase secreto.

## Integrado en sus herramientas de trabajo

### Una unidad
Un archivo o una carpeta copiados en la unidad de Pseudalis se anonimizan una vez escritos, nombre incluido. No se instala ningún controlador.

### Comandos del Explorador
Anonimizar y Restaurar en archivos, carpetas y selecciones de hasta 100 elementos, con o sin etiquetas, en su lugar o como copia.

### Una aplicación web
Suelte archivos, archivos comprimidos y carpetas, o pegue un texto, y recupérelos anonimizados o restaurados.

### Una interfaz local
Una interfaz HTTPS documentada en la dirección de bucle local para otros programas: archivos, textos, nombres, registros y ajustes.

# Cuatro dispositivos de detección

Las reglas detectan los valores con estructura; dos pequeños modelos entrenados detectan, por el contexto, los que escapan a las reglas.

**Escáneres validados**
Direcciones de correo e IP, números de teléfono, IBAN, tarjetas y los números nacionales de unos cuarenta países, aceptados cuando sus dígitos de control se verifican.

**Etiquetador SVM estructural**
Un etiquetador de secuencias sobre las palabras, que lee nombres, lugares, direcciones y organizaciones por su contexto.

**Tablas de palabras**
44 tablas en cinco idiomas: tratamientos, etiquetas de campo, formas jurídicas, nombres de pila, organismos públicos y unas dos mil organizaciones.

**Transformer a nivel de byte**
Un modelo de tramos de 116 791 parámetros que lee cualquier grafía e indica dónde empieza y termina un valor.

## Las mediciones del detector

Detector completo, límites exactos, todas las categorías (Pseudalis 1.0.5)

| Conjunto de prueba | Precisión | Exhaustividad | F1 |
|---|---|---|---|
| Prueba, cinco idiomas | 0,955 | 0,935 | **0,945** |
| Valores nunca vistos en el entrenamiento | 0,951 | 0,935 | **0,943** |
| Seis veces más ruido | 0,941 | 0,897 | **0,918** |
| Documentos enteros | 0,961 | 0,943 | **0,952** |
| Solo lugares del extranjero | 0,960 | 0,939 | **0,949** |

## Los límites de Pseudalis

- La detección es probabilística en las categorías aprendidas: conviene releer un documento protegido antes de compartirlo.
- El modelo de confianza es el equipo: todo usuario del equipo comparte su clave, sus originales y su administración.
- Una página escaneada sin capa de texto se conserva como un recuadro gris: no se realiza OCR.
- Un documento en un idioma distinto de los cinco que leen los reconocedores conserva más valores legibles.

## Categorías de datos personales

| Categoría | Estado |
|---|---|
| Nombres, apodos, nombres de usuario | ● Cubierto |
| Direcciones, ciudades, códigos postales con su ciudad | ● Cubierto |
| Teléfono, correo electrónico, fechas | ● Cubierto |
| Tratamiento, sexo, nacionalidad | ● Cubierto |
| Identificadores con dígito de control: seguridad social, IBAN, tarjetas, unos 40 países | ● Cubierto |
| IP, MAC, UUID, IMEI, matrículas, VIN, GPS | ● Cubierto |
| Organizaciones y organismos públicos nombrados | ● Cubierto |
| Documento de identidad, permiso de conducir, permiso de residencia | ● En parte |
| Números de cliente, expediente, contrato y empleado | ● En parte |
| Salud, creencias o hechos penales escritos en frases enteras | ○ No cubierto |
| Texto dentro de las imágenes | ○ No cubierto |

## Características

| | |
|---|---|
| **Documentos** | Word, Excel y PowerPoint (.docx, .xlsx, .pptx y sus variantes con macros y plantillas), PDF |
| **Textos** | Más de treinta extensiones, entre ellas .txt, .csv, .md, .json, .xml, .html, .yaml y .sql, y cualquier archivo cuyo contenido sea texto |
| **Idiomas leídos** | Francés, inglés, italiano, español y alemán |
| **Cifrado de los valores** | FF1 sobre AES-256, cada etiqueta autenticada con HMAC-SHA-256 |
| **Almacén de los originales** | AES-256-CTR con HMAC-SHA-256; clave maestra protegida por DPAPI; claves derivadas con HKDF-SHA-256 |
| **Interfaz local** | https://127.0.0.1:8444, documentada con Swagger |
| **Dependencias** | Ninguna: el instalador lo incluye todo, y los programas importan solo las bibliotecas de Windows |

## Estándares y referencias

| | |
|---|---|
| **NIST SP 800-38G** | Cifrado FF1 de los valores con conservación del formato |
| **AES-256, HMAC-SHA-256** | Cifrado y autenticación de las etiquetas y del almacén |
| **HKDF-SHA-256, PBKDF2-SHA-256** | Derivación de las claves y protección de la clave exportada |
| **ECDSA P-256** | Firma del canal de actualizaciones |
| **RGPD, artículos 25 y 32** | Seudonimización y cifrado como medidas de protección desde el diseño y de seguridad |
| **Directrices 01/2025 del CEPD** | Seudonimización, leídas en su estado de noviembre de 2026 |

*Más información*
cydral-technology.eu/es/products/pseudalis/

*Descargas*
cydral-technology.eu/es/downloads/

*Contacto*
cydral-technology.eu/es/contact/

Libre y gratuito bajo la Boost Software License 1.0. © 2026 Cydral Technology, Francia.