Pseudalis

個人データをさらすことなく、文書を使い、共有する

Pseudalis はあなたの文書の個人データを検出し、それぞれの値を、暗号化した値そのものを収めたタグに置き換えます。文書も、そこから派生したテキストも、元に戻せるのはあなたのコンピューターの鍵だけです。

バージョン **1.0.11**

プラットフォーム **Windows 11, x64、ドライバー不要**

ライセンス **Boost Software License 1.0**

例 — タグあり

件名：契約更新のご案内

`[[PERSON:Rnifypn Eouiitlrwky#EKR]]` 様

`[[DATE:28/08/2018#VLC]]` のお電話に続き、`[[ORG:Xkekwmufwmcj Eouiitlrwky et Wldwdst SARL#T4N]]`（`[[LOCATION:Saint-Ntaahgp#QWB]]`）の更新書類をお送りします。返金は口座 `[[IBAN:GB45 NWBK 6152 7186 9288 16#K9P]]` に振り込まれます。

ご連絡は `[[PHONE:06 07 09 00 95#M2X]]` または `[[EMAIL:rnifypn.eouiitlrwky@pspbshklkspj.com#7HD]]` までお願いいたします。

0.946
5 言語のテストセットでの検出器の F1（バージョン 1.0.11）

41,000
2 コアで毎秒読み取る文字数

40
国民番号を検証できる国の数

17 MB
すべてを含み、ドライバー不要のインストーラー

概要

タグが値を運ぶ
それぞれの値は、AES-256 上の FF1 で暗号化した値を検査符号付きで保持するタグになります。対応表は一切保存しません。

タグあり、タグなし
値を置換値だけで書き出すこともでき、プロンプト、テストデータ、社外に送るファイル用に、普通の文書として読めるファイルになります。

指定語と指定列
管理者は最大 100 個の語と CSV ファイルの 100 列を指定でき、ほかの読み取りより先に置き換えるため、エクスポートした表どうしの結合が保たれます。

画像は Albireo が説明
同じコンピューターに Albireo があれば、各画像は、人物を特定しない説明文に置き換えられます。

文書はレイアウトを保つ
Word、Excel、PowerPoint、PDF、テキスト、CSV、Markdown、HTML、XML、JSON、YAML のファイルはレイアウトを保ち、PDF はページ構成を保ちます。

名前とプロパティ
ファイル名、アーカイブ内の項目名、文書のプロパティも置換され、復元時に元に戻ります。

原本は暗号化して保管
原本は、期間と容量に上限のある暗号化保管庫に入り、匿名化したファイルのどのコピーからでも必要に応じて取り出せます。

一貫した置換値
同じ値にはどこでも同じ置換値が与えられ、保護された IBAN は検査に通り、日付は秘密のずらし幅でそろって移動します。

いつもの業務ツールに組み込まれています

ドライブ
Pseudalis のドライブにコピーしたファイルやフォルダーは、書き込まれた時点で名前も含めて匿名化されます。ドライバーはインストールしません。

エクスプローラーのコマンド
ファイル、フォルダー、最大 100 項目の選択に対する「匿名化」と「復元」。タグあり・なし、その場での処理・コピーの作成を選べます。

ウェブアプリケーション
ファイル、アーカイブ、フォルダーをドロップするか、テキストを貼り付けると、匿名化または復元された結果を受け取れます。

ローカルインターフェース
他のプログラム向けの、ループバックアドレス上の文書化された HTTPS インターフェース。ファイル、テキスト、名前、レコード、設定を扱います。

4 つの検出機構

ルールが構造を持つ値を検出し、2 つの小さな学習済みモデルが、ルールでは拾えない値を文脈から検出します。

検証付きスキャナー
メールアドレスと IP アドレス、電話番号、IBAN、カード番号、約 40 か国の国民番号。チェックディジットが合うものだけを採用します。

構造化 SVM タガー
単語に対する系列ラベラーで、氏名、地名、住所、組織を文脈から読み取ります。

単語表
5 言語で 44 の表：敬称、項目ラベル、法人形態、名、公的機関、約 2,000 の組織。

バイト単位 transformer
116,791 パラメーターのスパンモデルで、どんな綴りも読み、値の開始位置と終了位置を示します。

検出器の実測値

検出器全体、スパンの完全一致、全カテゴリ（Pseudalis 1.0.5）

テストセット	適合率	再現率	F1
テスト、5 言語	0.955	0.935	**0.945**
学習で一度も見ていない値	0.951	0.935	**0.943**
6 倍のノイズ	0.941	0.897	**0.918**
文書全体	0.961	0.943	**0.952**
国外の地名のみ	0.960	0.939	**0.949**

個人データのカテゴリ

カテゴリ	状態
氏名、ニックネーム、ログイン名	● 対応
住所、都市、都市名付きの郵便番号	● 対応
電話、メール、日付	● 対応
敬称、性別、国籍	● 対応
検査符号付きの識別子：社会保障番号、IBAN、カード、約 40 か国	● 対応
IP、MAC、UUID、IMEI、ナンバープレート、VIN、GPS	● 対応
組織と名称のある公的機関	● 対応
身分証明書、運転免許証、滞在許可証	● 一部対応
顧客番号、案件番号、契約番号、従業員番号	● 一部対応
文章として書かれた健康、信条、犯罪に関する事実	○ 非対応
画像内の文字	○ 非対応

Pseudalis の限界

- 学習によるカテゴリでは、検出は確率的です。保護した文書は共有前に読み直してください。
- 信頼の単位はコンピューターです。そのコンピューターのすべてのユーザーが、鍵、原本、管理権限を共有します。
- テキストレイヤーのないスキャンページは灰色の枠として残ります。OCR は行いません。
- 認識器が読める 5 言語以外で書かれた文書では、読めるまま残る値が多くなります。

仕様

文書	Word、Excel、PowerPoint（.docx、.xlsx、.pptx と、そのマクロ付き・テンプレート形式）、PDF
テキスト	.txt、.csv、.md、.json、.xml、.html、.yaml、.sql など 30 種類以上の拡張子と、内容がテキストであるあらゆるファイル
読み取る言語	フランス語、英語、イタリア語、スペイン語、ドイツ語
値の暗号化	AES-256 上の FF1。各タグは HMAC-SHA-256 で認証
原本の保管庫	AES-256-CTR と HMAC-SHA-256。マスター鍵は DPAPI で保護。鍵は HKDF-SHA-256 で導出
ローカルインターフェース	https://127.0.0.1:8444、Swagger で文書化
依存関係	なし。インストーラーがすべてを含み、プログラムは Windows のライブラリだけを読み込みます

規格と参考文献

NIST SP 800-38G	値の FF1 フォーマット保持暗号
AES-256, HMAC-SHA-256	タグと保管庫の暗号化と認証
HKDF-SHA-256, PBKDF2-SHA-256	鍵の導出と、エクスポートした鍵の保護
ECDSA P-256	アップデートフィードの署名
GDPR 第 25 条・第 32 条	設計段階からのデータ保護措置およびセキュリティ措置としての仮名化と暗号化
EDPB ガイドライン 01/2025	仮名化（2026 年 11 月時点の内容に基づく）

詳しくは
cydral-technology.eu/ja/products/pseudalis/

ダウンロード
cydral-technology.eu/ja/downloads/

お問い合わせ
cydral-technology.eu/ja/contact/

Boost Software License 1.0 による無料のオープンソース。© 2026 Cydral Technology、フランス。