CamoText User Guide

Table of Contents

1. Introduction
2. System Requirements & Installation
3. Interface Overview
4. Step-by-Step Usage Instructions
    1. Input Text
    2. Anonymize
    3. Review the Output
    4. Copy or Save Output for Further Use
    5. De-Anonymize
5. Data Types & Meanings
6. FAQ
7. Best Practices
8. Contact Us

1. Introduction

CamoText is a file and text anonymization tool designed to detect and obscure sensitive data before sharing text or files, including in LLM prompts and other AI use cases, to preserve privacy. It operates as an offline desktop application with zero data retention by default.

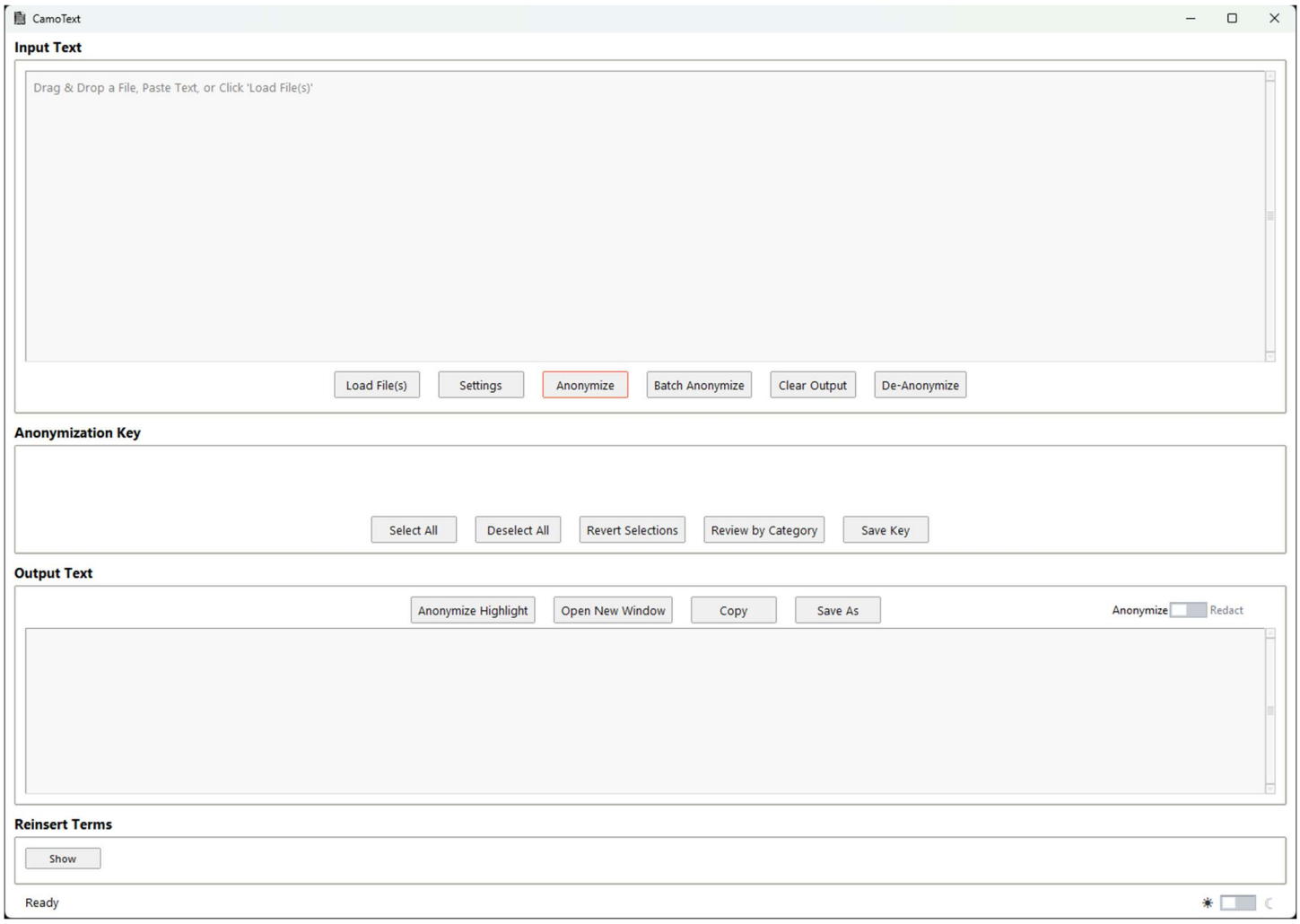
CamoText scans your input(s) for various personally identifiable information (PII) and other types of confidential details including any configured user settings, then securely anonymizes them with encrypted labels in newly-written output. The human-in-the-loop reviews the output to highlight and anonymize any other sensitive terms, and revert any false positives, before use.

2. System Requirements & Installation

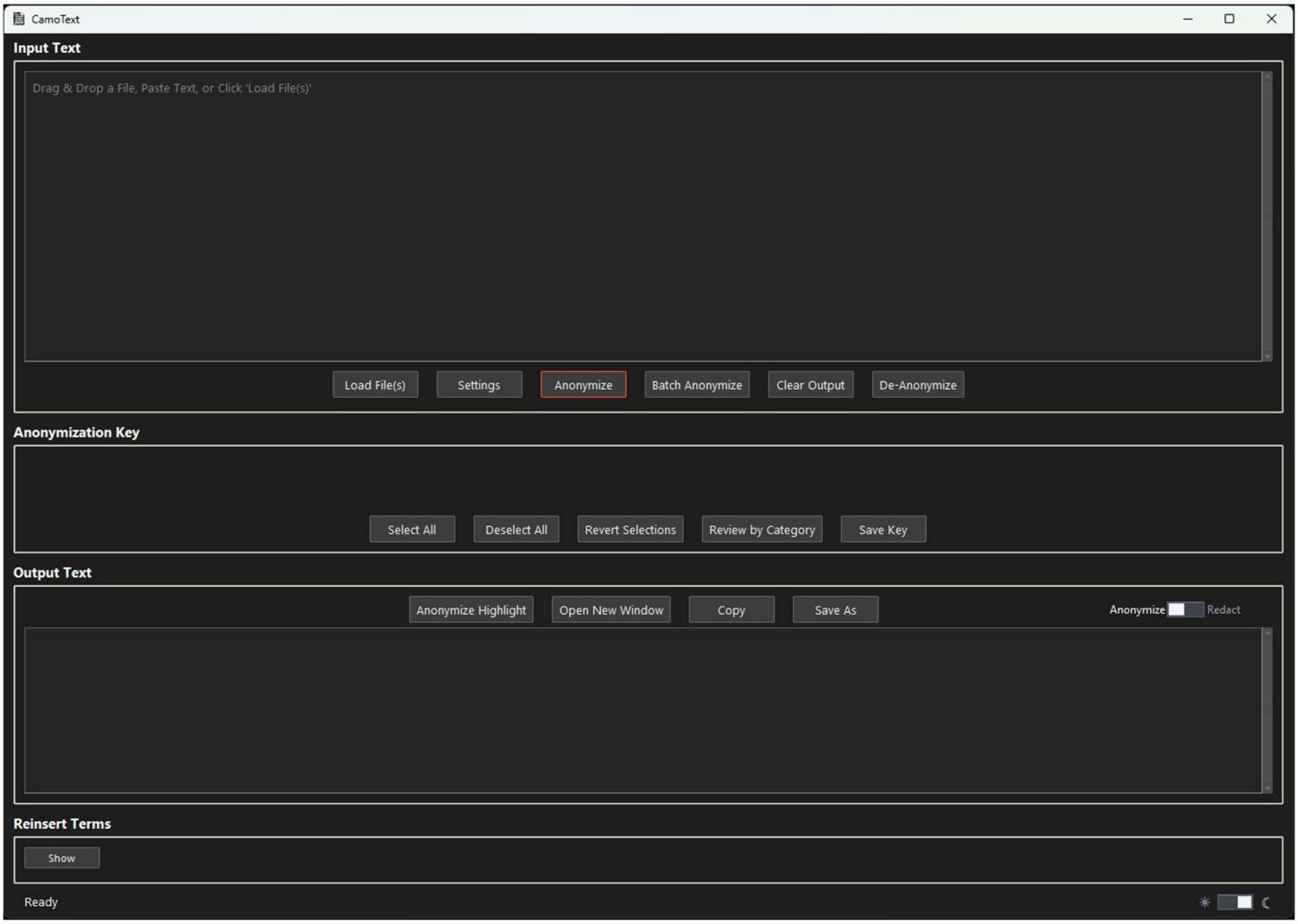
- **Operating System**: Windows 10 or later 64-bit, or macOS 15+
- **Memory and Space**: At least 4 GB RAM; 8 GB or higher recommended for large files. Approximately 340 MB disk space required for the full bundle installation.
- **Installation**: Double-click the downloaded installer. Follow on-screen prompts to complete. After installation, CamoText will be in the Start menu (Windows) or Applications folder (macOS).

3. Interface Overview

Light Theme

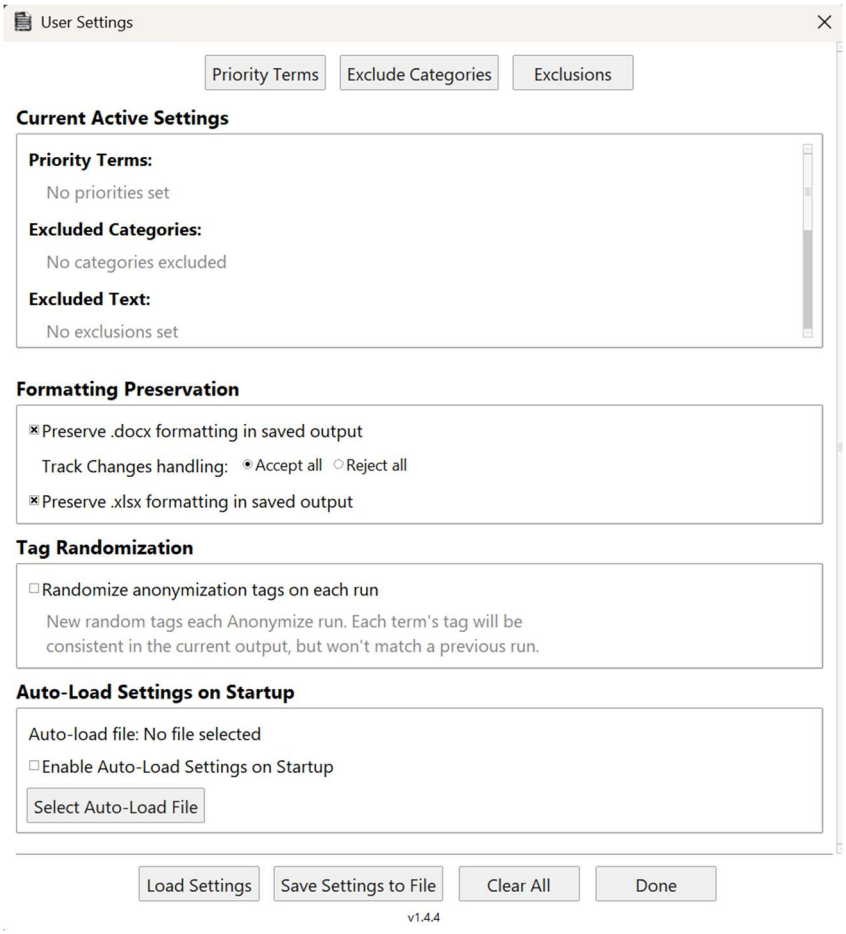


Dark Theme

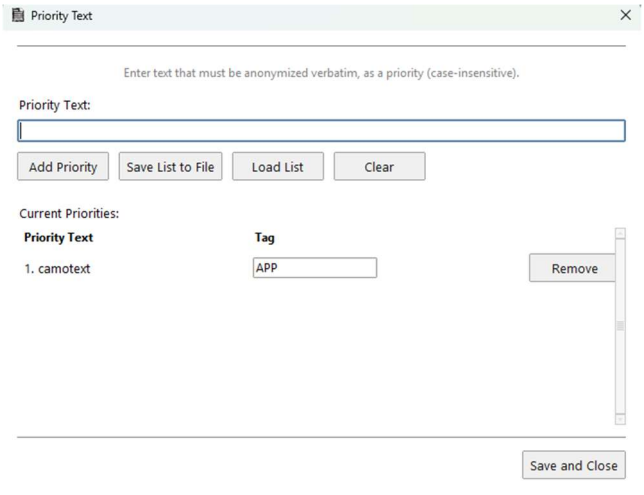


Once launched, the main window of CamoText is divided into sections. The default size is according to your display settings, but all windows may be manually resized. Hover your mouse over buttons to display helpful tips on their purposes.

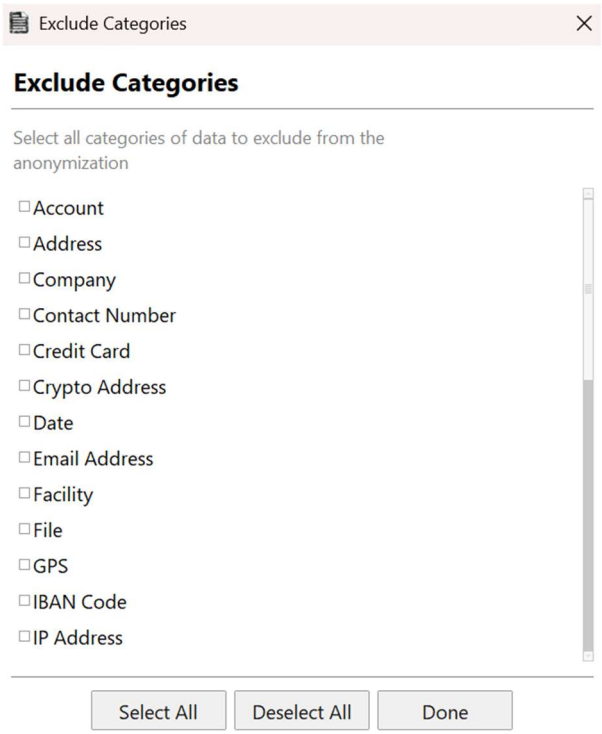
1. **Input Text**
    - A text box where you *paste or load* the document you wish to anonymize.
    - **Load File(s)** button for importing supported file types (*.txt, .md, .pdf, .docx, .xlsx, .csv, .rtf*).
    - **Settings** button for specifying User Settings for anonymizations, which may be saved for repeat use, loaded automatically at startup, and combined. Settings options:



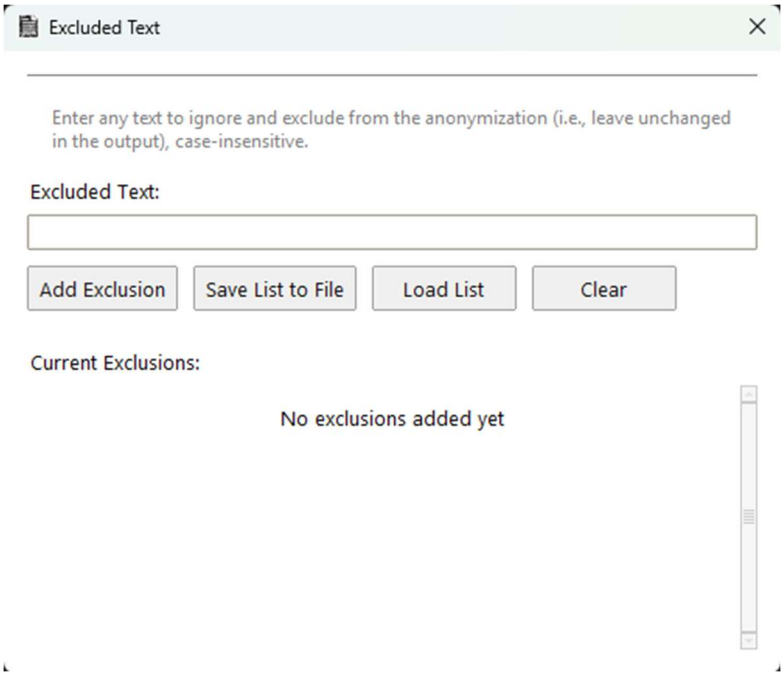
- **Priority Terms:** verbatim, case-insensitive text that MUST be anonymized as a priority before running the anonymization. Priorities are matched in order and ignore leading or trailing whitespace but **can** include punctuation. You may also specify a custom category tag to override the “PRIORITY” default, such as “LAWFIRM” or “SUSPECT”; handy for more specifically labeling anonymized terms for AI use.



- **Exclude Categories:** displays all data categories in anonymizations; select any that should be excluded (ignored) when anonymizing, if not contextually sensitive.



- **Exclusions**: verbatim, case-insensitive text that MUST be excluded and ignored in the anonymization (leaving these terms unchanged in the output), for example contextual false positives.

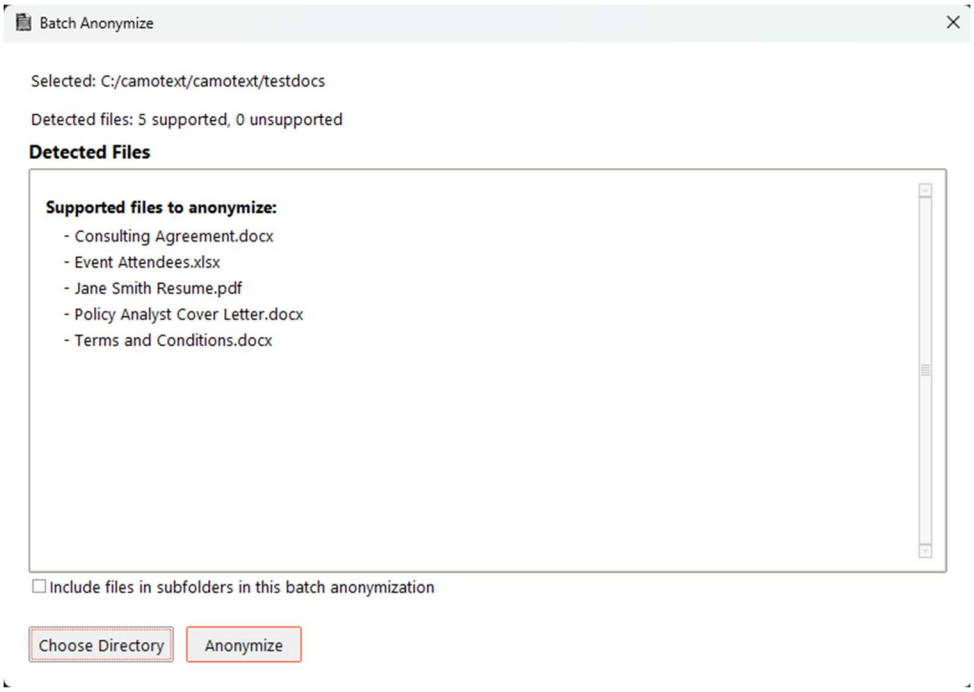


- **DOCX Formatting**: preserve input .docx formatting when saving Output Text to .docx, with the ability to accept or reject tracked changes if present.
- **XLSX Formatting**: preserve input .xlsx formatting when saving Output Text to .xlsx

- **Tag Randomization**: randomize the hash tags for each finding on each anonymization run, preventing correlation across sessions. For example, with this checked, “CamoText LLC” may be detected as <**ORGANIZATION_8a3f2794**> in one anonymization, then <**ORGANIZATION_c9c8b7e8**> in a subsequent anonymization.
- **Auto-Load Settings on Startup**: automatically load saved settings each time CamoText is opened

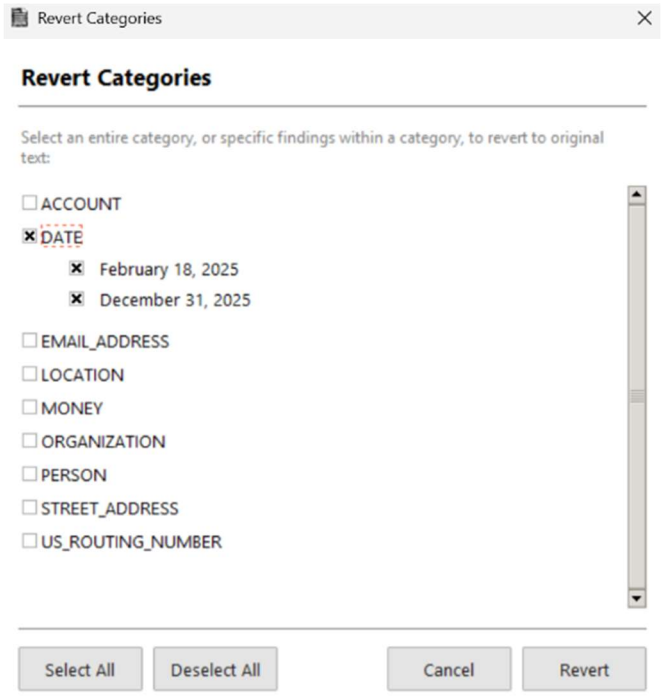
- **Anonymize** button to run CamoText’s anonymization of your input.

- **Batch Anonymize** button to run a batch anonymization of all supported files in a specified directory, with a checkbox to also anonymize files in any subfolders. Active User Settings will be applied.
  - When clicked, a window opens with a **Choose Directory** button to select the directory or folder containing the files to be anonymized, and an **Anonymize** button that is enabled (displayed with an orange outline) when a directory is selected. Clicking this button runs the anonymization on all applicable files in the directory, producing the anonymization key and output text in the main window.



- **Clear Output** button to clear the Anonymization Key and Output Text.

2. **Anonymization Key**
    - This pane displays the list of all anonymized terms along with their encrypted placeholders.
    - Tools include **Select All**, **Deselect All**, **Revert Selections**, **Review by Category**, and **Save Key**.
        - When **Review by Category** is clicked, a window opens displaying the list of data categories recognized and anonymized. From there, the user may select an entire category or specific items within a category to revert to their respective original text all at once by selecting and clicking **Revert**.



3. **Output Text**
    - Shows the text after anonymization.
    - Buttons to **Anonymize Highlight**, **Open New Window, Copy**, and **Save As**:
        - **Anonymize Highlight** — Highlight and anonymize any additional sensitive terms or text of any length you spot, in every location in the output, with an optional custom label.
        - **Open New Window** — see *Using "Open New Window"* below.
        - **Copy** — Copies the output text to your clipboard.
        - **Save As** — Saves the current state of the output as a file of your choice. If Batch Anonymization mode was selected, a window will pop up with checkboxes for whether the output files should be saved in the same directory and same filetypes as the input (for example, saving all .docx and

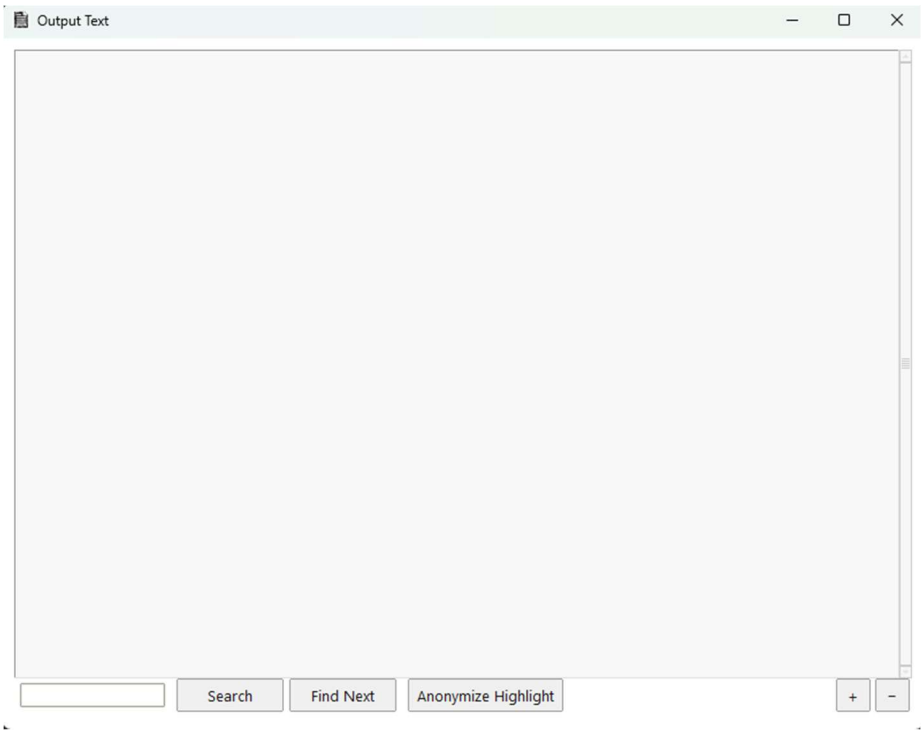
.pdf inputs in the same folder as anonymized .docx and .pdf files respectively), and whether the anonymization keys should be saved in the same directory. If unchecked, the user specifies the output directory.

Using "Open New Window"

In the **Output Text** panel, you will see a button labeled **Open New Window**.

When clicked:

1. **Separate Window Launch**
   - CamoText opens the entire output text in a new, larger window (screenshot below).
   - This new window includes the following features:
     - **Search** — A text field and a **Search** button to locate specific words or phrases in your anonymized text. If a searched term is found, easily click **Anonymize Highlight** to anonymize it in each location.
     - **Find Next** — skips to the next highlighted location of a searched term.
     - **Anonymize Highlight** — same functionality as in the main window.
     - **Zoom** — Small **-** and **+** buttons appear in the bottom-right corner for adjusting the font size.



2. **Making Changes in the Separate Window**
   - Any text you highlight and anonymize in the separate window will be automatically updated accordingly in the main window's Output Text upon closing the separate window.
   - Each newly anonymized item is also tracked in the **Anonymization Key**, just like in the main interface.

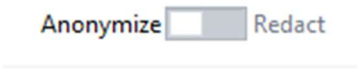
3. **Automatic Sync**
    - While the separate window is open, a message is displayed in the Output Text field and the main window will be frozen to avoid mismatches while editing.
    - When you close the separate window, **all changes** — including newly anonymized and tagged terms — automatically update the main window's **Output Text** and the **Anonymization Key**.

Use this feature whenever you have a large amount of text or prefer a dedicated space (or larger font) to review, search, and refine the final anonymized text before saving or distributing.

At the bottom, you'll see a status indicator that typically shows messages such as "Ready" or "Analyzing text...", as well as your application version.

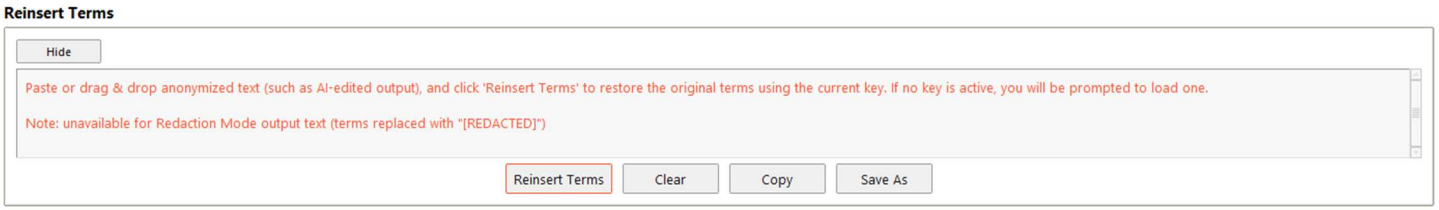
Redaction Mode

At the right of the frame is a toggle to switch between **Anonymize and Redact** modes, for toggling Redaction and Normal mode. After an anonymization has run, enabling Redaction Mode will replace all anonymized items (i.e. **<PERSON_cf624301>**) with **[REDACTED]** in all locations. Clicking the button again will reinstate the tagged anonymizations. Redaction Mode files cannot be de-anonymized in the app.



Reinsert Terms

Below the Output Text panel is a collapsible/expandable field to Reinsert Terms, to easily drag-and drop or paste an anonymized file or text (for example, if edited with an AI tool using the anonymized tags) and reinsert the original terms in the active Anonymization Key by clicking **Reinsert Terms.** If a .docx or .xlsx file is dragged and dropped into this field, formatting will be preserved after clicking **Reinsert Terms** and **Save As** to save the proper filetype. There are also buttons to **Clear** the field, and **Copy** plaintext to clipboard.



Also note the **Theme Toggle** on the bottom right of the app, to choose light or dark mode.



4. Step-by-Step Usage Instructions

Step 1: Input Text

In the **Input Text** box, you can either:

- **Drag & Drop** a file or files directly from your file explorer into the text box.
- **Paste** text from your clipboard.
- **Click** the **Load File(s)** button to open a file or files (*.txt, .md, .docx, .pdf, .xlsx, .csv, .rtf*).

If multiple files are inputted, they will be concatenated in the input text with a ***NEXT FILE*** indicator between them. If you prefer to preserve file and filetype separation and save outputs accordingly, use the **Batch Anonymize** feature.

If you batch anonymize files, the directory of the files being anonymized will be displayed in the Input Text field.

If you want to ensure certain terms are detected and hashed, categories are ignored, or certain terms are ignored, configure them in the **Settings** window. You can save settings for subsequent use, load automatically upon app open, and combine lists.

For .docx and .xlsx inputs, formatting is preserved by default. Headers and footers will also be anonymized, and hyperlinks, images, and logos are removed. You may also choose to **accept or reject redlines** if the loaded .docx includes tracked changes – your setting choice is reflected upon the next .docx input load, and will also be reflected if tracked-change documents are used in Deanonymize or Reinsert Terms.

If CamoText detects an unusually large file, it may prompt you regarding file size and performance considerations. CamoText automatically extracts the text in your file, so be aware that some file formats such as .pdf may introduce formatting issues (such as line breaks, spaces, or unknown characters) when automatically extracted that can affect the anonymizer's accuracy. Using plaintext, .txt, .md, or .docx files whenever possible is recommended.

Step 2: Anonymize

1. Once your text is visible in the **Input Text** area, click the **Anonymize** button.
2. If you're using the batch anonymization feature and have properly selected a directory in the Batch Anonymize window, click the **Anonymize Batch** button in this window.
3. **CamoText** automatically wipes metadata, scans the input and replaces detected sensitive information with hashed (encrypted) and categorized placeholders, for example: <PERSON_c94e6fa3>.
4. The anonymized version appears in the **Output Text** panel, and the **Anonymization Key** panel now lists each hashed term and its original value.

Step 3: Review the Output

1. In the **Anonymization Key** panel, review each hashed item.
2. If a term was hashed *unnecessarily*, check its box and click **Revert Selections**.

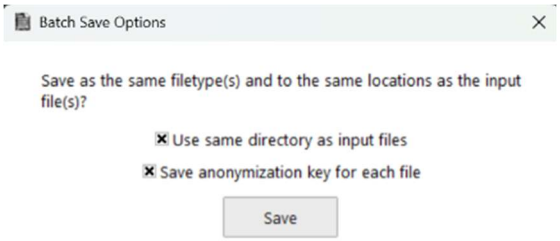
- Example: If **“Maryland”** was hashed but you prefer it visible as you don’t consider it to be confidential.

3. Browse items for any that were hashed unnecessarily withing each category by clicking **Review by Category**, then check the full category (unchecking individual items you want to remain masked) and click **Revert**.
    - Example: If **“DATE”** text findings were hashed but you prefer them all visible as you don’t consider them confidential, revert them in one easy step here.
4. The output updates so that the original text is reinserted for the reverted terms.
5. *Double-Check Sensitive Terms*: Look over the **Output Text** carefully to ensure that **all** critical confidential or identifying terms or passages are hashed. *Human-in-the-loop review* is essential as automated detection may skip oddly formatted words or partial matches.
6. **Anonymize Highlight**: If you spot new or missed terms or even full passages, highlight them in the Output Text box and click **Anonymize Highlight**. This hashes them in every location in the text and updates the Key panel accordingly**.**
7. **Open New Window**: For an easier review, you can open the **Output Text** in a separate window— especially helpful for larger documents and searching for specific terms in the output. Any anonymizations you make in that window autosync back to the main interface.
8. If you’d prefer the anonymizations in the output be redacted (displayed as [REDACTED]) rather than tagged, click **Redaction Mode**

Step 4: Copy or Save Output, or Reinsert Terms into adjusted Output

1. **Copy Output**: Copies the fully anonymized text to your clipboard.
2. **Save As**: Save output text in a supported format (.md, *.pdf, .txt, .docx, .xlsx, .csv, .rtf*), with formatting preserved for .docx and .xlsx.

   For Batch Anonymizations, this saves the anonymized files in the same filetype as their input, and shows checkboxes for whether you’d like to save in the same directory and save the anonymization keys.



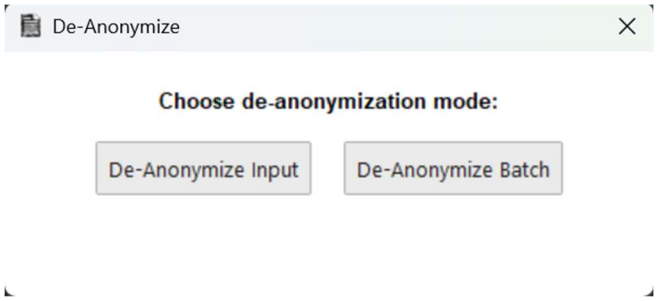
3. **Reinsert Terms**: If you have adjusted output text for which you’d like to reinsert the original terms in place of the anonymized placeholders, drop the document or paste the text into the Reinsert Terms field and click **Reinsert Terms** to use the active Anonymization Key to reinsert original text for each anonymized tag (if no key is active, you will be prompted to load one). Not available for [REDACTED] text.
4. **Keep or Discard the Key**:

- If you need to reference the original anonymized details in a future session, or want the ability to use CamoText to natively de-anonymize multiple files with the same key later, hold onto the Anonymization Key by clicking **Save Key** to save locally.
- If you want the data to remain inaccessible, do *not* share or store the Key.

5. **Close**: Once CamoText is closed, no text nor data is retained by it unless you have copied it to clipboard or saved it to your device (like Settings lists, keys, or output).

De-Anonymization

If the applicable text to be deanonymized corresponds to the Anonymization Key that's still loaded and active in the current session, use the **Reinsert Terms** feature described above for quick restoration of original terms. Alternatively, to deanonymize a file or folder of files:



The "De-Anonymize" button in the Input Text frame enables either single-file or full directory restoration of original terms from anonymization keys of anonymized files (redaction mode files, i.e. with global "[REDACTED]" tags, cannot be de-anonymized). For Batch de-anonymization, a directory containing at least one valid file and anonymization key (.JSON) must be selected. For single input de-anonymization, after loading the document into the Input field, a valid anonymization key (.JSON) must be selected.

In both modes, the relevant anonymization key(s) populate the Anonymization Key window, and the Output displays the restored text(s). Similar to the Batch Anonymization save process, when saving a Batch De-Anonymization you have the option to save the output as the same file extensions in the same directory as the input, with a "clean-" prefix to the filename. For .docx and .xlsx files, de-anonymized outputs will retain original formatting.

5. Data Types & Meanings

CamoText recognizes and replaces many different categories of text that are commonly considered PII (personally identifiable information) or otherwise commercially sensitive:

- **PERSON**: Full or partial names that might identify an individual.
- **EMAIL_ADDRESS**: Any recognized email format (e.g., *user@domain.com*).
- **CONTACT_NUMBER**: Telephone and other numbers in various formats, including international.
- **DATE**: Dates and times in various formats (e.g. 01/01/2001, D.O.B. Jan. 1, 2001)
- **COMPANY**: Formatted legal entity names, such as *CamoText Inc.*
- **ORGANIZATION:** Groups, organizations and associations
- **NRP**: Nationality, religious, and other demographic identifiers
- **MONEY**: Monetary amounts (with currency symbols or references, numerical or words).
- **LOCATION**: General location identifiers (e.g. *Baltimore, Maryland*)
- **ADDRESS / STREET_ADDRESS**: Address lines, road names, etc.
- **CREDIT_CARD**: Numerical string matching common credit card number patterns
- **ACCOUNT**: Account handles, PINs, and names (e.g. *@CamoText1*)
- **US_SSN**: Numerical string matching US social security number pattern
- **US_DRIVER_LICENSE**: Driver's license according to https://ntsi.com/drivers-license-format/
- **US_PASSPORT**: string matching US passport number pattern
- **US_BANK_NUMBER**: US Bank Account number
- **US_ROUTING_NUMBER**: US Bank routing number
- **US_ITIN**: US Individual Taxpayer Identification Number
- **UK_NHS**: United Kingdom National Health Service ID number
- **IBAN_CODE**: IBAN banking code
- **TRADEMARK**: trademarked symbol-adjacent terms
- **VIN:** vehicle identification numbers
- **IP_ADDRESS**: Both IPv4 and IPv6 patterns (e.g., *192.168.0.1*).
- **MEDICAL_LICENSE**: String matching common medical license patterns
- **MEDICARE_ID**: Medicare Beneficiary Identifier (MBI)
- **UUID**: Universally Unique Identifier strings (e.g., e-signed document identifiers, GUIDs).
- **CRYPTO_ADDRESS**: cryptocurrency addresses (Bitcoin, Ethereum/EVM, Solana)
- **GPS**: GPS Coordinates in common formats (e.g., 38°53'48.1"N 77°00'23.1"W)
- **FAC**: Facilities, such as buildings or airports
- **URL**: Website urls or hyperlink patterns.
- **FILE**: File path strings (e.g., *C:\Users\Username\Documents\secret.pdf*, *SurveillanceFootage.mp4*).
- **PRIORITY**: any text entered in the "Priorities" window by the user not otherwise custom-tagged.
- **HIGHLIGHT**: default tag for any text highlighted and anonymized by the user not otherwise custom-tagged.

Note: Some text (such as misspelled personal names, custom IDs, unusual terms, or text with line breaks or multiple whitespaces) might *not* be auto-detected. Due to the overlap among some categories' coverage, some text may be anonymized under an erroneous category. Many categories capture other types of data with similar formats, such as CIK or serial numbers.

Detection uses a natural language processing engine combined with probability-based pattern matching, which is not 100% accurate and may miss or improperly label detections.

Sensitivity of text is often subjective and recognizers are not perfect – ALWAYS do a manual review of the anonymized terms and output.

6. Frequently Asked Questions

Q: Can I run multiple anonymizations at once? What about batch processing?

A: Yes, just drag-and-drop or load multiple files into the Input Text window or paste the text from multiple documents into the Input Text window. For batch processing of an entire folder, use the Batch Anonymize button. For CLI batch processing, CamoTextCLI is bundled with Pro+ or International+. Contact us about Enterprise customization options.

Q: Why does my computer block installation, saying the app is from an "unknown" or "unidentified" publisher?
A: Windows' firewall and macOS's Gatekeeper have especially sensitive security settings for bundled executables like CamoText, so they may warn you when trying to install, despite the application and installers being code-signed and notarized.

On Windows, double-click the ".exe" file to open the installer, which should display "CamoText LLC" as the publisher. If the file is blocked, the window may give an option for "More Info" which, when clicked, should provide an option to "Run Anyway" – after selecting this, the installer will proceed as normal.

On MacOS, double-click the ".dmg" file to open it, drag the .app file to your Applications folder, Right-click the .app in Applications and select "Open." In the pop-up, click "Open" to bypass the warning if it appears.

On either operating system, after CamoText has been installed and opened, it should be accessible and usable in the future without any further warnings.

Q: Can CamoText handle large PDFs or Word documents?
A: Yes, up to 50MB and 800 pages, though performance will depend on your system's memory. PDFs will be slower and more prone to errors in formatting than text files. For extremely large files on less-performant computers, consider first copying the text itself to clipboard, converting to a txt file, or breaking them into parts.

Q: Do I need to keep the Anonymization Key?
A: Only if you foresee needing to restore or reference those portions of the original text in a future session without using the original text itself, or need it for auditing purposes. If not, you can ignore it to reduce the risk of unintentional exposure, as it is automatically deleted when CamoText is closed.

Q: Does CamoText detect sensitive images or scans of text?
A: In its current version, CamoText only output textual content. Content other than text (including scanned PDFs with no extractable text) will not appear in the output.

Q: Why does the tool sometimes over-match or skip certain words?
A: It is designed to often err on the side of caution by over-matching certain patterns, especially

considering the ease of reversing false positives. Text that's misspelled or oddly formatted may be skipped, so always double-check.

Q: Wait, can AI reverse-engineer the hashes CamoText creates when anonymizing?
A: No, the SHA256 hashing is effectively irreversible by current technology. While only the first 8 characters are displayed for readability, 32 characters are used in the code and a unique "salt" is added in the anonymization process. For additional protection, enable Tag Randomization in Settings to create unique per-run keys that are not correlated with other anonymizations of the same terms.

Q: Are there other reasons to use CamoText other than with AI?
A: CamoText can be useful anytime masking or redaction is needed or advisable in documents to quickly remove PII and as an easy-to-use interface for manual redactions.

Q: Can I have a special type of recognizer for my company or agency, or remove some recognizers I don't need?
A: Absolutely! Email contact@camotext.ai for custom versions and bulk pricing. The User Settings and auto-load feature can also be used to specify company-wide settings.

---

7. Disclaimers & Best Practices

1. **REVIEW THE END USER LICENSE AGREEMENT**

   Use of CamoText is expressly subject to the accompanying End User License Agreement — ensure you read it in full and understand it. It's also accessible at https://camotext.ai/eula.

2. **DO NOT RELY SOLELY ON AUTOMATED ANONYMIZATION**

   This software is intended to include *human-in-the-loop review*. Always inspect the output for missed or improperly hashed terms and leverage the User Settings to improve outcomes.

3. **AIM FOR THE STRICTEST PRIVACY STANDARDS POSSIBLE**

   If you identify entire paragraphs or other broad sections that pose a risk of re-identification due to their substance, highlight and anonymize them fully.

4. **SHARING THE ANONYMIZATION KEY DEFEATS ITS PURPOSE**

   Do not share the key with anyone unless absolutely necessary – it contains the removed PII and other sensitive terms mapped to their placeholders.

5. **CONSULT LEGAL OR COMPLIANCE EXPERTS**

   Data sensitivity and privacy requirements vary by jurisdiction and context. This tool does not replace professional counsel.

6. **FILE FORMATTING LIMITATIONS**

   Some file types may introduce stray characters or spacing that hamper detection. Check for line-breaks, concatenation, or mid-word splits and correct them if they reveal sensitive data.

---

8. Contact & Version

- **Website**: https://camotext.ai
- **Email Support**: contact@camotext.ai
- **Latest Version: 1.4.x**

Thank you for using CamoText!

If you have any questions or ideas on how to improve the software, please reach out.