

УДАЛЕНИЕ ПЕРСОНАЛЬНЫХ ДАННЫХ ИЗ ДОКУМЕНТОВ

В эпоху цифровой революции наша жизнь стала неразрывно связана с технологиями, которые, хоть и приносят нам множество преимуществ, также представляют собой потенциальную угрозу для нашей личной безопасности и конфиденциальности, от алгоритмов искусственного интеллекта, отслеживающих нашу онлайн-активность, до камер видеонаблюдения, следящих за нашими движениями в реальном мире. Наши личные данные оказываются под угрозой постоянного вмешательства и контроля. Если не подходить к защите своих данных со всей серьёзностью, может произойти их утечка или неправомерное использование. Нужно принимать меры для защиты нашей цифровой личности и сохранения приватности в мире, где границы между частной и публичной сферами становятся все более размытыми.

Сегодня для написания докладов, статей, рефератов, дипломных работ, и оформления других документов подавляющее большинство людей используют Microsoft Word, но немногие задумываются о том, что docx файл кроме написанного в нём текста, также хранит скрытые свойства. При сохранении документа автоматически заполняются сведения об авторе, дате создания и изменения, однако если документ создан на компьютере какой-либо организации утечка этих данных не желательна. Это не безопасно, так как любой человек, в руки которого попадает этот файл, может их прочитать [1]. Авторы статей на данную тему, предлагают использовать встроенные в редакторы документов функции очистки скрытых свойств, но в таком случае каждый файл требуется обрабатывать отдельно [2, 3]. Кроме того, при редактировании файла с удалёнными свойствами, могут быть добавлены новые сведения о компьютере, на котором выполнялось редактирование.

Для решения этой проблемы было разработано десктопное приложение на Python. При разработке приложения были задействованы следующие библиотеки: docx, tkinter, os. Docx позволяет работать с файлами формата Microsoft Word, предоставляет возможности для создания, чтения и модификации документов с расширением docx. Tkinter является стандартным модулем Python для создания графического интерфейса пользователя. Os предоставляет функции для взаимодействия с операционной системой, такие как создание, удаление и перемещение файлов и каталогов, выполнение команд в командной строке. Разработанное приложение позволяет изменять скрытые свойства docx файла. Для этого нужно выбрать файл (рис. 1), и ввести требуемые значения свойств в предоставленные поля (рис. 2).

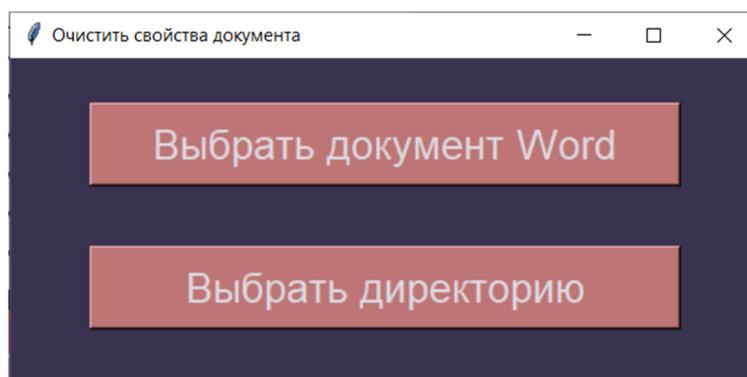


Рисунок 1 — Выбор файлов

Также реализована возможность удаления скрытых свойств сразу у нескольких файлов. Для этого необходимо поместить все файлы, подлежащие очистке, в одну директорию, после чего выбрать эту директорию в приложении (рис. 1). После завершения на экран будет выведено сообщение, информирующее об успешности выполнения операции, и количестве файлов, для которых удалось провести очистку свойств (рис. 3).

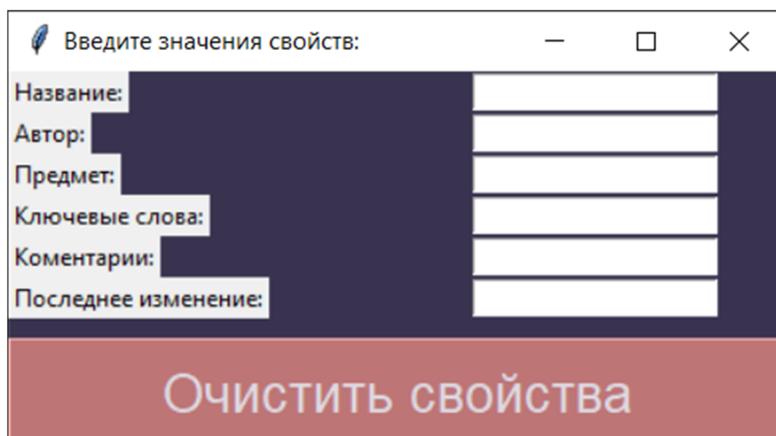


Рисунок 2 — Изменение свойств

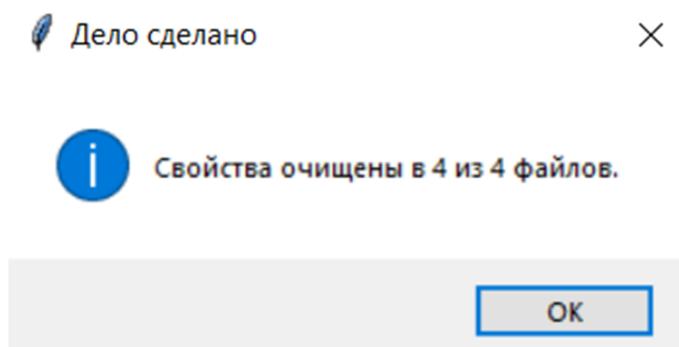


Рисунок 3 — Очистка нескольких файлов

Распространение word документов на ненадёжных платформах и отсутствие заботы о конфиденциальности скрытых свойств может привести к утечке личных данных. Благодаря тому, что приложение очистки скрытых свойств предоставляет возможность одновременного редактирования сразу нескольких документов, его использование не отнимает много времени. Это позволяет индивидуальным предпринимателям, и крупным компаниям организовать безопасный обмен документами, без потери времени.

Список источников

1. Каримова А. Ю. Метаданные как угроза конфиденциальности личной информации // E-Scio. 2023. № 4 (79). С. 41–49. EDN DMYYZF
2. Барамба С. Удаление персональных данных из метаданных файлов // Системный администратор. 2013. № 12 (133). С. 17. EDN RNHFED
3. Садирмекова Ж. Б., Самбетбаева М. А. Разработка технологии извлечения метаданных из документов в процессе интеграций информационных систем // Наука и инновационные технологии. 2020. № 1 (14). С. 209–214. EDN ZOESMV. DOI: 10.33942/sit.nes004