Руководство пользователя для приложения "Document Processor"

# Введение

Добро пожаловать в приложение 'Document Processor'! Это многофункциональное средство для работы с различными типами файлов. С помощью приложения вы можете извлекать данные из архивов, обрабатывать текст и изображения, создавать опись документов, наносить номера, переименовывать файлы и сравнивать метаданные документов со справочником. Данное руководство подробно описывает каждую функцию и шаги для её использования.

# Описание функций

## Поля для путей

В верхней части интерфейса расположены поля для указания путей: к архивам, директории для результатов и директории с файлами. Эти поля позволяют вам указать расположение файлов для обработки в каждой из функций.

## Извлечение архивов

Кнопка 'Извлечь архив/архивы' позволяет распаковать выбранные архивы (форматы .zip, .rar, .7z) и сохранить их содержимое в указанную директорию. Для использования этой функции выберите архивы и директорию для извлечения, затем нажмите 'Извлечь архив/архивы'.

## Извлечение текста и изображений

С помощью кнопки 'Извлечь текст и изображения' можно извлекать текст и изображения из файлов (PDF, DOCX, TXT, XLSX), находящихся в выбранной директории. Извлечённый текст сохраняется в текстовый файл, а изображения сохраняются в папку 'extracted\_images' внутри выбранной директории.

## Создание описи

Кнопка 'Сформировать опись' создаёт опись документов, находящихся в указанной директории. В опись включены такие данные, как наименование, обозначение, количество страниц и формат каждого документа. Для использования укажите директорию с файлами и нажмите 'Сформировать опись'.

## Нанесение номеров

Кнопка 'Нанести номера' автоматически пронумеровывает файлы в выбранной директории. Каждому файлу присваивается уникальный номер, который добавляется к содержимому файла (для документов) или к имени файла.

## Переименование файлов

Функция 'Переименовать файлы' проверяет файлы в выбранной директории на наличие недопустимых символов и переименовывает их в допустимый формат. Имена файлов обновляются, чтобы содержать только буквы, цифры, дефисы и символы подчеркивания.