

Wstępne założenia projektowe dla aplikacji „Duplikwidator”

Celem projektu jest opracowanie aplikacji pozwalającej na wyszukiwanie i eliminowanie duplikatów plików graficznych, niezależnie od formatu, wielkości czy też jakości obrazu.

Projekt zostanie zrealizowany z wykorzystaniem platformy GitHub.

Repozytorium projektu dostępne jest pod adresem:

<https://github.com/Sc00rpY/Duplikwidator>

Koordinacja prac nad projektem będzie przebiegać za pomocą diagramu Gantta. Aplikacja będzie posiadać podstawowe schematy UML, podręcznik użytkownika oraz dokumentację javadoc.

Aplikacja zostanie stworzona jako program desktopowy przy użyciu języka Java w środowisku programistycznym NetBeans IDE.

Język Java został wybrany ze względu na niezależność od platformy systemowej, wysoką wydajność oraz dostępność dużej ilości bibliotek usprawniających pracę nad aplikacją.

Interfejs graficzny aplikacji zostanie zaprojektowany z wykorzystaniem biblioteki Swing. Możliwa będzie również obsługa aplikacji za pomocą linii poleceń.

Do implementacji wyszukiwania duplikatów plików graficznych wykorzystane zostaną trzy biblioteki, stosujące różnego rodzaju algorytmy do porównywania zawartości obrazu:

- jpHash (<https://github.com/pragone/jphash>)
- ImagePHash (<https://gist.github.com/lihonosov/10988301>)
- imgDiff (<https://codepaste.net/tviae0>)

Aplikacja będzie funkcjonować w następujący sposób:

W pierwszej kolejności użytkownik musi utworzyć listę plików graficznych, dla których będą wyszukiwane duplikaty. Użytkownik może dodawać do listy pojedyncze pliki graficzne lub zawartość całych katalogów, z możliwością uwzględnienia ich podkatalogów.

Możliwe będzie usunięcie wybranych plików z listy, jak i całkowite wyczyszczenie jej zawartości.

W kolejnym kroku należy określić minimalny stopień podobieństwa plików, od którego będą one uznawane za duplikaty. Użytkownik będzie mógł tego dokonać za pomocą suwaka lub poprzez wpisanie wartości z przedziału 0-100 w polu tekstowym.

Następnie użytkownik musi wskazać katalogi, które mają zostać przeszukane w celu odnalezienia duplikatów, oraz wybrać metodę porównywania plików za pomocą przycisków wyboru. Możliwy będzie wybór kilku metod jednocześnie i połączenie ich wyników.

Wyszukiwanie duplikatów będzie odbywać się w osobnym wątku w taki sposób, aby umożliwić użytkownikowi korzystanie w tym samym czasie z interfejsu graficznego aplikacji.

W czasie wyszukiwania duplikatów aplikacja wyświetli okienko z informacją o znalezionych do tej pory duplikatach. Użytkownik będzie miał możliwość przerwania wyszukiwania w dowolnej chwili.

Wyniki wyszukiwania duplikatów zostaną zaprezentowane w tablicy, zawierającej takie informacje jak:

- lokalizacja porównanych plików na dysku twardym
- rozszerzenie plików
- szerokość oraz wysokość obrazów
- podobieństwo plików wyrażone w procentach

Użytkownik będzie mógł zobaczyć podgląd każdej pary obrazów, aby samemu ocenić ich podobieństwo. Jeżeli uzna, że obrazy nie są duplikatami, będzie mógł odrzucić wynik zaproponowany przez aplikację.

Następnie użytkownik będzie mógł usunąć znalezione duplikaty z dysku twardego, pozostawiając obraz o najlepszej jakości spośród wszystkich kopii.