

Janusz S. Bień

djview4poliqarp

8 marca 2018

Program jest przede wszystkim graficznym zdalnym klientem serwera oprogramowania korpusowego Poliqarp for DjVu; jako taki stanowi modyfikację przeglądarki djview4 (<http://djvu.sourceforge.net/djview4.html>). Dodatkowo w 2017 r. program został rozbudowany o obsługę indeksów do dokumentów w formacie DjVu.

Postać kwerend służących do wyszukiwania w korpusach jest udokumentowana w innych tekstach, między innymi w <http://nkjp.pl/poliqarp/help/pl.html> i <http://poliqarp.sourceforge.net/man/poliqarp-query-syntax.html>. Przykłady użycia można znaleźć m.in. w artykułach i prezentacjach dostępnych do 30 czerwca 2018 r. w Bibliotece Cyfrowej Katedry Lingwistyki Formalnej Uniwersytetu Warszawskiego (<http://bc.klf.uw.edu.pl/>); ich lokalizacja po tej dacie nie jest jeszcze znana.

1. Konfiguracja

1.1. Serwery

Program jest dystrybuowany w postaci częściowo skonfigurowanej, w szczególności dostępne są od razu adresy serwerów Poliqarp for DjVu, które oczywiście w razie potrzeby możemy zmienić w panelu Ustawienia — Skonfiguruj — Serwery; adres serwera można (trzeba?) podać w postaci samej nazwy domenowej bez oznaczenia protokołu.

Warto dodać, że oprogramowanie serwera i narzędzia do budowy korpusów są dostępne na otwartych licencjach, zainteresowani mogą więc tworzyć własne korpusy i serwery.

1.2. Dodatkowe możliwości przeglądania plików DjVu

Ponieważ funkcje programu są nieco inne niż przeglądarki djview4, istnieje możliwość łatwego uruchomienia djview4 dla aktualnie oglądanego tekstu. Dla poprawnego działania tej funkcji w wersji dla MS Windows jest niezbędne wskazanie lokalizacji programu djview4 w panelu Ustawienia — Skonfiguruj — Ogólne.

1.3. Wygląd

Można wskazać język interfejsu na 3 sposoby:

- język systemu
- polski
- angielski

Efekt zmiany jest widoczny dopiero po ponownym uruchomieniu programu.

Można też wybrać kolor wyróżnienia wyników wyszukiwania.

1.4. Font

Można wybrać font używany do wyświetlania wyników tekstowych i niektórych innych informacji w celu zmiany jego rozmiaru lub repertuaru znaków.

Rekomendowane są fonty zgodne z rekomendacjami Medieval Unicode Font Initiative, np. *Junicode*.

Niektóre informacje są wyświetlane jednak zawsze z wykorzystaniem domyślnego fontu systemu.

1.5. Podgląd graficzny

Można zmienić domyślną wysokość okienka podglądu wyników w panelu bocznym.

Można również zmienić domyślną wartość powiększenia lupy.

1.6. Dokument powitalny

Można wskazać dokument, który zostanie wyświetlony w panelu głównym po uruchomieniu programu. Możliwość w praktyce przydatna w razie dystrybucji programu z odpowiednim dokumentem — dotąd z tej możliwości nie korzystano.

1.7. Wprowadzanie znaków specjalnych

Można zdefiniować podstawienia ułatwiające wprowadzenia do pola kwerend i do indeksu znaków niedostępnych na klawiaturze. Na przykład poniższe podstawienie

```
/#=#
```

umożliwia wprowadzanie znaku DOUBLE OBLIQUE HYPHEN jako `/#`. Analogicznie znak SOFT HYPHEN można wprowadzać jako np. `/-`.

1.8. Korzystanie z plików lokalnych

W celu przyspieszenia działania programu i uniezależnienia się od serwera korpusu (jest to w pełni możliwe tylko przy korzystaniu z samego indeksu) można pobrać odpowiednie dokumenty na dysk lokalny i podać ich lokalizację za pomocą odpowiedniego podstawienia. Na przykład jeśli na komputerze z systemem Linux pierwszy tom słownika Lindego znajduje się w katalogu `/var/cache/Linde/1`, to odpowiednie podstawienie powinno mieć postać

```
http://korpusy.klf.uw.edu.pl/djvus/linde-t/01/=file:///var/cache/Linde/1/
```

Należy pamiętać, że w przypadku słowników mamy z reguły do czynienia z tzw. dokumentami rozłożonymi. Najprostszy sposób ich pobrania to otwarcie w programie `djview` za pomocą `Open Location`, a następnie użycie `Save As`.

1.9. Indeksy

Obsługa korpusów pojawiła się w 2017 r. w wersji 1.3. Początkowo dokładnie jeden indeks był przypisany do konkretnego korpusu, obecnie wybór indeksu lub indeksów jest niezależny od korpusu.

Indeksy to pliki tekstowe, hasła są reprezentowane przez wiersze zawierające pola oddzielone średnikami. Początkowo indeksy miały trzy pola, obecnie jest ich cztery (indeksy trzypolowe są nadal akceptowane, czwarte pole traktuje się jako puste).

Aktualnie pola indeksu są następujące:

1. wyrażenie hasłowe — służy do wyszukiwania i szeregowania haseł, wyświetlane w panelu listy haseł;
2. adres zaznaczenia w dokumencie DjVu, wyświetla się ono w panelu głównym po wybraniu hasła — pole może być puste, wtedy wyrażenie hasłowe jest wyświetlone kursywą;
3. komentarz — pole może być puste, po wybraniu hasła wyświetla się pod panelem listy haseł (we wcześniejszej wersji hasło nie wyświetlało się, jeśli pierwszym znakiem był wykrzyknik);

4. opis — wyświetla się za wyrażeniem hasłowym, ale nie wpływa na ich porządkowanie czy wyszukiwanie, stosowane np. do różnicowania homonimów. Aktualnie dostępny jest indeks do słownika Lindego, patrz <https://bitbucket.org/jsbien/ilindecsv>. Nowe indeksy można łatwo tworzyć samodzielnie dzięki odpowiednim funkcjom programu.

2. Użytkowanie

Sposób korzystania z programu jest dość oczywisty, najważniejsze funkcje są dostępne w menu głównym oraz w podręcznych menu.

Skróty klawiaturowe są w dużym stopniu odziedziczone po programach djview4 i djview3. Były one opisane w instrukcji do tzw. wydania elektronicznego DjVu-Libre słownika Knapskiego z 2005 r.; ponieważ jest ona obecnie niedostępna, prawie bez zmian jest ona powtórzona dalej (właściwy słownik jest nadal dostępny w e-BUW). Zaproponowana wtedy terminologia nie jest — jak się wydaje — powszechnie akceptowana.

2.1. Panel główny

Panel główny służy do przeglądania dokumentu DjVu, w szczególności do oglądania wyniku wyszukiwania za pomocą kwerendy lub indeksu.

- Rozkład tekstu na ekranie:
 - Ciągłe wyświetlanie stron: **F4**.
 - Wyświetlanie stron obok siebie: **F5**.
Wyświetlana para stron z reguły nie odpowiada tzw. rozkładówką oryginału: wyświetlają się strony 1-2, 3-4, a nie 2-3, 4-5 itd. W razie potrzeby wyświetlania rozkładówek należy skorzystać z programu djview, w którym taka funkcja jest dostępna.
- Skalowanie (ang. *zoom*), większość poniższych operacji dostępna również z menu:
 - Skalowanie względne
 - na szerokość strony: **w** (od angielskiego *Fit width*)
 - na rozmiar strony: **p** (od angielskiego *Fit page*)
 - Skalowanie bezwzględne
 - 100 %: **1**
 - 200 %: **2**
 - 300 %: **3**
 - kolejno 150 %, 300 %, 450 % itp.: **+**
 - kolejno 75 %, 50 %, 25 % itp.: **-**
 - Skalowanie płynne: rolka myszy przy wciśniętym **Ctrl**.
 - Lupa (ang. *lens*): **Ctrl** razem z **Shift**. Lupa wyświetla zawsze obraz nawet jeśli włączone jest wyświetlanie tekstu ukrytego.
- Pozycjonowanie (ang. *panning* i inne)
 - Początek strony (górnny lewy róg): **Home**
 - Koniec strony (dolny lewy róg): **End**
 - Przesunięcie strony w dół (kursora w górę): **Up**
 - Przesunięcie strony w górę (kursora w dół): **Down**
 - Przesunięcie strony w lewo (kursora w prawo): **Right**
 - Przesunięcie strony w prawo (kursora w lewo): **Left**
 - Przesunięcie strony myszą w dowolnym kierunku po naciśnięciu lewego klawisza
 - Skalowanie płynne w pionie: rolka myszy.
- Przewijanie (ang. *scrolling*)
 - Przewinięcie do przodu (tekst w górę, kursor w dół, w razie potrzeby przejście do następnej strony): **Space**
 - Przewinięcie do tyłu (tekst w dół, kursor w górę, w razie potrzeby przejście do poprzedniej strony): **Backspace**

- Nawigacja, poniższe operacje są dostępne również z menu:
 - Następna strona (zachowuje aktualną pozycję strony w oknie): **Page Down**
 - Poprzednia strona (zachowuje aktualną pozycję strony w oknie): **Page Up**
 - Początek dokumentu: **Ctrl + Home**
 - Koniec dokumentu: **Ctrl + End**
- Obrót
 - Obrót 90° w lewo (w kierunku przeciwnym do ruchu wskazówek zegara): **Ctrl+Left**.
 - Obrót 90° w prawo (w kierunku zgodnym z ruchem wskazówek zegara): **Ctrl+Left**.
- Zaznaczanie:
 - Wciśnięcie klawisza **Ctrl** i lewego klawisza myszy pozwala zaznaczyć prostokątny fragment skanu, a następnie wybrać z menu podręcznego odpowiednią operację:
 - Powiększ zaznaczenie
 - „Kopiuj odnośnik” czyli umieść w schowku adres tego zaznaczenia
 - Kopiuj tekst ukryty pod zaznaczeniem
 - Kopiuj zaznaczony fragment skanu
 - Zapisz zaznaczony fragment skanu
 - Dodaj hasło do indeksu (indeks musi być aktualnie otwarty, pierwsze pole jest wstępnie wypełnione przez tekst ukryty, drugie zawiera URL zaznaczenia)
 - Uaktualnij hasło (URL zaznaczenia zastępuje odpowiednie pole w bieżącym hasle)
 - Wyświetlanie tekstu ukrytego: **Ctrl+t**.
 - Dostęp do fragmentu tekstu ukrytego:
 - Wciśnięcie klawisza **Shift** wyświetla fragment tekstu ukrytego znajdującego się pod kursorem (fragment jest ograniczony przez strukturę tekstu zapisaną w dokumencie).

Następujące operacje dostępne są tylko z menu **Widok**:

- Wyświetlanie wierne (jednemu pikselowi dokumentu odpowiada jeden piksel na ekranie) — **Wiernie**.

3. Panel kwerend

Panel ten wyświetla się z lewej strony ekranu; jego wyświetlanie można włączyć i wyłączyć z menu **Widok**. Jego szerokość może być zmieniana przez użytkownika za pomocą myszy.

Zawiera on następujące elementy:

- Pole serwera
- Pole korpusu
- Pole kwerend
- Trzy podpanele:
 - Podpanel wyników graficznych
 - Podpanel wyników tekstowych
 - Podpanel metadanych
- Przycisk **Więcej**
- Informacja o liczbie trafień znalezionych i wyświetlonych.

Informacja o aktualnie znalezionych trafieniach może być wyeksportowana w formacie cvs. Plik zawiera następujące pola oddzielone przecinkami, w razie potrzeby ujęte w cudzysłowy:

 1. numer trafienia
 2. lewy kontekst
 3. lewe uzgodnienie
 4. prawe uzgodnienie

5. prawy kontekst
6. URL z zaznaczeniem trafienia
7. tekst kwerendy

3.1. Pole serwera

Pole serwera pozwala wybrać konkretny serwer z listy ustalonej w trakcie konfiguracji. Przycisk obok nazwy serwera wyświetla informację o serwerze.

3.2. Pole korpusu

Pole korpusu pozwala wybrać jeden z korpusów dostępnych na wybranym wcześniej serwerze. Dodatkowo dostępne są dwa przyciski. Jeden z nich wyświetla informację o korpusie, drugi pozwala zmienić domyślny sposób wyszukiwania. Dostępne są następujące opcje (są one przekazywane do programu **poliqarp**, w związku z tym są identyczne jak w innych korpusach obsługiwanych przez ten program, np. Narodowy Korpus Języka Polskiego):

- Ograniczenie przeszukiwania do losowej próbki i ustalenie jej wielkości.
- Sortowanie wyników (domyślnie są wyświetlane w kolejności występowania w korpusie):
 - według lewego lub prawego kontekstu, *a fronte* lub *a tergo*;
 - według „uzgodnienia” (jeśli kwerenda jest podzielona na części, to według lewego uzgodnienia), *a fronte* lub *a tergo*;
 - według prawego „uzgodnienia”, jeśli kwerenda jest dwuczęściowa,
 - według prawego kontekstu,
- Sposób wyświetlania elementów tekstu, określany niezależnie dla „uzgodnienia” i kontekstu. Można włączyć lub wyłączyć wyświetlanie:
 - formy pierwotnej (w typowych korpusach to forma fleksyjna),
 - formy podstawowej (w typowych korpusach to forma hasłowa),
 - znaczników (w typowych korpusach opisujących własności gramatyczne)
- Długość kontekstu w jednostkach wyświetlanego w konkordancjach i w oknie kontekstu.

3.3. Pole kwerendy

Program przekazuje kwerendy bez żadnych zmian do serwera, w związku z czym ich opisu należy szukać w innych źródłach, które były wspomniane we wstępie. Program przechowuje historię kwerend, zamiast wpisywać kwerendę od początku można wybrać już użytą kwerendę i w razie potrzeby ją zmodyfikować..

Najprostsza kwerenda to po prostu napis, np. broda. Należy pamiętać, że wielkość liter jest istotna, więc wyniki kwerend broda i Broda są różne.

Dowolny element oznaczamy w kwerendzie jako []. Oto jego użycie w kwerendzie dwuczęściowej (wskazane jest wybranie sortowania według prawego uzgodnienia): ku ^ [].

3.4. Podpanel wyników tekstowych

Ten podpanel można uznać za główny, wyświetla wyniki w formie konkordancji w sposób identyczny jak inne narzędzia wykorzystujące serwer **poliqarp**.

Podwójne kliknięcie lewym klawiszem myszy na wiersz trafienia wyświetla jego obraz graficzny w głównym panelu.

Pojedyncze kliknięcie lewym klawiszem myszy wybiera ten wiersz do następujących operacji:

- przejścia do okna metadanych, gdzie wyświetla się również dłuższy kontekst,
- jednoczesne naciśnięcie klawisza **Ctrl** i **Backspace** usuwa z listy trafień ten wiersz jako nieinteresujący (nie ma możliwości cofnięcia tej operacji inny sposób niż ponowne wykonanie kwerendy).
- przejścia do okna wyników graficznych z zachowaniem dokonanego wyboru (który może być niewidoczny).

Kliknięcie lewym klawiszem na numer trafienia wyświetla jego metadane bez zmiany zawartości panelu głównego.

3.5. Podpanel wyników graficznych

Część funkcji jest analogiczna do panelu wyników tekstowych:

Podwójne kliknięcie lewym klawiszem myszy na wiersz trafienia wyświetla jego obraz graficzny w głównym panelu.

Pojedyncze kliknięcie lewym klawiszem myszy wybiera ten wiersz do następnych operacji:

- przejścia do okna metadanych, gdzie wyświetla się również dłuższy kontekst,
- jednoczesne naciśnięcie klawisza **Ctrl** i **Backspace** usuwa z listy trafień ten wiersz jako nieinteresujący (nie ma możliwości cofnięcia tej operacjiw inny sposób niż ponowne wykonanie kwerendy).
- przejścia do okna wyników tekstowych z zachowaniem dokonanego wyboru (który może być niewidoczny).

Kliknięcie lewym klawiszem na numer trafienia wyświetla jego metadane bez zmiany zawartości panelu głównego.

Kliknięcie środkowym klawiszem myszy na tekst w bocznym panelu uruchamia nowy egzemplarz `djview4` wyświetlający odpowiednią stronę; pozwala to wykonać na tekście operacje aktualnie nie dostępne w `djview4poliarp` — np. wyświetlić tzw. konspekt (*outline*).

Każde z okienek podglądu graficznego działa podobnie, jak panel główny — możliwe jest w szczególności pozycjonowanie tekstu.

3.6. Podpanel metadanych i kontekstu

Funkcja tego podpanelu jest oczywista, nie ma z nim związanych jakichś specjalnych funkcji poza widocznymi przyciskami Poprzedni, Następny, Usuń. Jak w każdym panelu działa w nim zaznaczanie i kopiowanie tekstu. Kliknięcie hiperlink uruchamia `djview` z odpowiednim zaznaczeniem.

4. Panel indeksu

Panel ten wyświetla się z prawej strony ekranu; jego wyświetlanie można włączyć i wyłączyć z menu Widok. Jego szerokość może być zmieniana przez użytkownika za pomocą myszy.

Zawiera on następujące elementy:

- Informacja o nazwie indeksu, jego ewentualne modyfikacji i liczbie haseł.
- Pole wyszukiwania przyrostowego haseł; domyślnie wyszukuje się od początku wyrażenie hasłowego, ale jeśli wprowadzony zapis zaczyna się od znaku *, to wyszukuje się również podnapisy.
- Lista haseł; normalnie obok hasła wyświetla się również jego opis (z wyjątkiem sortowania *a tergo*).
- Pole komentarza, które można edytować; po edycji można przejść do pola wyszukiwania za pomocą **Ctrl+f**.
- Pole stanu: ścieżka do aktualnie otwartego indeksu, numer strony dokumentu wyświetlanej w panelu głównym,

Za pomocą menu głównego można wykonać następujące informacje na indeksie:

- Otwórz wskazany plik jako indeks,
- Otwórz ponownie jeden z ostatnio używanych indeksów
- Odśwież aktualnie otwarty indeks (w razie dokonania zmian w pliku za pomocą innego programu)
- Otwórz plik i dodaj jego zawartość do indeksu
- Zapisz aktualnie otwarty indeks
- Zamknij aktualnie otwarty indeks

Po wpisaniu w polu wyszukiwania pewnego napisu naciśnięcie klawisza Enter wyświetla cyklicznie kolejne hasła zaczynające się od tego napisu. Po liście haseł można poruszać się również za pomocą strzałek w górę i w dół oraz klawiszy Home i End.

Dostępne w menu i za pomocą skrótów klawiaturowych **Alt** + strzałka w lewo i **Alt** + strzałka w prawo operacje przejścia do poprzedniego i następnego hasła nie odnoszą się do kolejności na liście haseł, ale do historii oglądania haseł.

W trakcie zapisywania zmodyfikowanego indeksu początkowa wersja indeksu jest zachowana w formie kopii zapasowej. Jeśli otwartych było kilka plików, zapisywane są one jako jeden plik.

Dodatkowe operacje są dostępne w menu podręcznym otwieranym przez kliknięcie prawym klawiszem na wybrane hasło. Są to dwie operacje odnoszące się do wybranego hasła

- Edytuj hasło
- Oznacz jako usunięte (hasło zostanie usunięte dopiero przy zapisywaniu indeksu)

oraz — ze względów technicznych — opcję odnoszącą się do całego indeksu, czyli „Kolejność haseł”. Umożliwia ona wybór następujących sposobów ustalania kolejności:

- „Z pliku”
- Alfabetyczne *a fronte* „słowo po słowie”
- Alfabetyczne *a fronte* „litera po literze”
- *A tergo*

5. Uwagi końcowe

Tekst ten jest udostępniony na licencji GNU FDL i nie ma charakteru ostatecznego. Źródła tekstu są dostępne w repozytorium <https://bitbucket.org/jsbien/ndt>, uwagi i propozycje zmian można zgłaszać za pomocą narzędzi dostępnych w tym repozytorium.