

Ogonki 97

wersja 1.3

Konwerter polskich znaków w dokumentach HTML

Copyright (C) 1997 Michał Jaskólski.
Wszelkie prawa zastrzeżone.

Autor:

Michał Jaskólski

e-mail: jj@gdansk.sprint.pl

www: <http://www.gdansk.sprint.pl/~jj/>

programy: <http://www.gdansk.sprint.pl/~jj/apps/>

Wersja 1.3

15 listopada 1997

Spis treści:

- [1] Wstęp.
- [2] Instalacja i deinstalacja.
- [3] Dokładny opis możliwości programu.
- [4] Możliwe problemy.
- [5] Historia.
- [6] O autorze.
- [7] Umowa licencyjna i gwarancja.

[1] Wstęp.

Problem z polskimi literkami istnieje od zarania ery informatycznej :). Od czasu powstania "światowej pajęczyny" nabrał on nowego wymiaru. Jakiś czas temu na polskich stronach WWW wcale nie używano polskich znaków, ale dziś sytuacja jest zupełnie inna - większość przeglądarek potrafi już polskie znaki wyświetlać.

Kłopot polega na tym, że polskie znaki powinny być zakodowane w standardzie ISO-8859-2, a większość programów do tworzenia i edycji stron HTML działających pod Windows używa "standardu" Windows-1250. Część z tych edytorów (np. Internet Assistant) zapisuje polskie literki w jeszcze bardziej "zagnatwany" sposób. Jeśli zależy tobie, aby polskie litery na twoich stronach były wyświetlane prawidłowo w większości przeglądarek, musisz strony te skonwertować. Testowałem praktycznie wszystkie konwertery i doszedłem do wniosku, że żaden nie spełnia moich oczekiwań. Postanowiłem więc napisać własny konwerter...

[2] Instalacja i deinstalacja.

Instalacja programu jest niezwykle prosta: wystarczy umieścić plik ogonki97.exe w dowolnym folderze :). Przy każdym uruchomieniu programu program sprawdzi gdzie się znajduje i dostosuje odpowiednie ustawienia.

Deinstalacja również jest bardzo łatwa. Wystarczy uruchomić "Ogonki", nacisnąć przycisk "Odinstaluj program" w opcjach - usunie to wszystkie ustawienia z rejestru, a potem wyjść z programu. Teraz możesz już skasować folder z "Ogonkami".

[3] Dokładny opis możliwości programu.

"Ogonki 97" to bardzo szybki konwerter o największych możliwościach wśród podobnych aplikacji dla Windows 95/98/NT.

[a] obsługiwane standardy

"Ogonki 97" konwertują z i na ISO-8859-2, Windows-1250 (CP) oraz Unikod UTF-8 (ten ostatni obsługiwany jest np. przez Netscape Navigатора 4 oraz Microsoft Internet Explorera 4), potrafią również konwertować znaki na ich bezogonkowe odpowiedniki. Konwertowane są nie tylko polskie litery, ale wszystkie znaki, zgodnie z tabelami konwersji opracowanymi przez Qrczaka (mkowalcz@knm.org.pl). Jeśli odpowiednia opcja jest włączona, dodawany / zmieniany (!) jest tag <meta ... charset> mówiący przeglądarce w jaki sposób ma wyświetlić daną stronę. Program potrafi też zamieniać znaki zapisane w postaci € oraz Ó na ich 8-bitowe odpowiedniki. Rozpoznawanie standardu wejściowego jest automatyczne (Unikod - na podstawie <meta ... charset>, ISO i CP na podstawie zawartych w dokumencie polskich liter).

[b] wybieranie dokumentów do konwersji

Przed samą konwersją musisz stworzyć listę dokumentów HTML, które zostaną poddane konwersji. Dokumenty mogą być dodawane do listy w następujący sposób:

- przez zaznaczenie danego dokumentu / dokumentów i / lub folderów zawierających pliki do skonwertowania w eksploratorze i "przesunięcie" ich nad okno "Ogonków", czyli przez drag'n'drop :)
- przez naciśnięcie pierwszego przycisku z boku listy: pojawi się wtedy okienko "Otwórz..." i będziesz mógł wybrać odpowiedni dokument (lub dokumenty, używając Ctrl / Shift).
- przez naciśnięcie drugiego przycisku obok listy: będziesz mógł wtedy wybrać folder - wszystkie dokumenty HTML znajdujące się w nim zostaną dodane do listy. Jeśli wybierzesz dodanie "Wszystkich plików" to wszystkie pliki, niezależnie od rozszerzenia, zostaną sprawdzone pod kątem charakterystycznej dla dokumentów HTML zawartości - w taki sposób program doda do listy wszystkie dokumenty HTML, nawet o nietypowym rozszerzeniu.
- jeśli używasz polskiego edytora HTML Pajęczek w wersji 2.0 lub 3.0 Pro lub edytora HomeSite w wersji 2.5 lub 3, możesz dodać do listy wszystkie lub wybrane dokumenty należące do wybranego projektu (lub projektów) tych edytorów. Po wybraniu przycisku "Dodaj projekt" z ikonką danego programu, pojawi się nowe okno. Po lewej stronie znajdzie się lista dostępnych projektów, po prawej - lista dokumentów. Gdy zaznaczysz projekt (lub projekty, używając Ctrl), wszystkie dokumenty należące do tych projektów zostaną wyświetlone po prawej stronie. Jeśli nie zaznaczysz żadnego dokumentu, to po wybraniu "OK" do listy zostaną dodane wszystkie dokumenty. Jeśli jakieś dokumenty są zaznaczone, to tylko one zostaną dodane do listy.
- przez wybór "Konwertuj" z menu kontekstowego, pojawiającego w Eksploratorze się po naciśnięciu prawego klawisza myszy przy zaznaczonym dokumencie HTML lub folderze. "Ogonki" zostaną wtedy uruchomione automatycznie.

Jeśli zaznaczona jest opcja "Dodawanie tylko dokumentów HTML", na liście zostaną umieszczone tylko dokumenty rozpoznane jako HTML. Oznacza to, że na liście nie pojawi się np. plik tekstowy, nawet o rozszerzeniu .html.

Lista plików zawiera własne menu, wywoływane przez kliknięcie prawym klawiszem myszy. Wybierając "Otwórz" możesz obejrzeć stronę w używanej przez siebie przeglądarce.

Pliki z listy możesz usuwać, zaznaczając je i naciskając odpowiedni przycisk obok listy. Możesz też usunąć wszystkie pliki, naciskając kolejny przycisk. Usunięcie plików z listy nie usuwa ich z dysku.

Istnieje możliwość korzystania ze skrótów klawiaturowych: Ins to dodanie jednego (lub kilku) plików, Ctrl+Ins to dodanie plików z wybranego folderu, Del usuwa zaznaczony plik (lub pliki), Ctrl + Del usuwa całą zawartość listy.

Możesz też zapisać nazwy plików znajdujące się na liście pod wybraną nazwą, klikając na przycisku "+" i użyć tej listy później.

[c] konwersja

Gdy na liście znajdują się już pliki, które zamierzasz skonwertować, musisz zdecydować się _na co_ chcesz te pliki skonwertować.

ISO-8859-2 to standard, internetowy i nie tylko, i właśnie do ISO powinieneś skonwertować swoje strony przed umieszczeniem ich w Internecie. Windows-1250, zwany potocznie CePem, to strona kodowa Windows, używana w większości programów działających w Windows, z Notatnikiem włącznie. Gdy wybierzesz ASCII, polskie (i nie tylko) literki zostaną zamienione na bezogonkowe odpowiedniki. Możesz w ten sposób stworzyć bezogonkową wersję tworzonych przez siebie stron. Unicode UTF-8 to standard przyszłości, umożliwiający wyświetlenie na _jednej_ stronie WWW tekstu w kilku językach. Na razie obsługiwany jest tylko przez najnowsze przeglądarki. Możesz skonwertować swoje strony na Unikod po zaznaczeniu opcji "Obsługa Unicode".

Teraz możesz przystąpić do konwersji. Naciśnij przycisk "Konwertuj" i w chwilę później pliki zostaną skonwertowane. Po naciśnięciu przycisku "Szczegóły" możesz obejrzeć krótki raport, zawierający również ewentualne problemy, które wystąpiły podczas konwersji. Jeśli żadnych problemów nie znaleziono, przycisk ten jest wyłączony.

Jeśli masz jeszcze jakieś pytania dotyczące "Ogonków 97", napisz do mnie (e-mail: jj@gdansk.sprint.pl).

[4] Możliwe problemy.

Oto kilka kwestii o których warto wspomnieć:

[a] auto-rozpoznawanie HTML.

"Ogonki" automatycznie rozpoznają, czy dodawany do listy dokument to dokument HTML, na podstawie rozszerzenia i zawartości. "Ogonki" służą wyłącznie do konwersji dokumentów HTML i nie skonwertują "zwykłego" pliku tekstowego. Jeśli drag'n'drop wydaje się nie działać, sprawdź, czy przeciągany plik to rzeczywiście dokument HTML.

[b] dodawanie meta-charset.

"Ogonki" potrafią dodać, albo nawet zmienić tag HTML <meta ... charset>, który umożliwia przeglądarkom wyświetlenie stron HTML we właściwy sposób. Jeśli jednak tag ten już istnieje ale jest nieprawidłowy, zmiana tagu może być niemożliwa.

[c] rozpoznawanie standardu

W większości przypadków "Ogonki" potrafią rozróżnić, czy strona jest zapisana w CP-1250, ISO-8859-2 czy Unikodzie i w zależności od wykrytego kodowania podejmują odpowiednie działania. Gdy "Ogonki" wykryją, że plik zawiera zbyt dużo "nieznanych" znaków z przedziału #127-#255, nie dokonają konwersji.

[d] win-1250 a iso-8859-2

Program używa pełnych tablic konwersji zawierających wszystkie, nie tylko polskie znaki. Gdy dany dokument zawiera zarówno znaki charakterystyczne dla win-1250, jak i iso-8859-2, to przy konwersji na jedno z tych kodowań mogą pojawić się krzaczkę. W takiej sytuacji można dokonać konwersji w drugą stronę, a potem konwersję powtórzyć. Jeśli używasz edytora tekstowego, możesz w opcjach, na zakładce "Zaawansowane", wyłączyć opcję "Przy konwersji ISO?CP używaj pełnych tablic konwersji".

[e] inne

Nie zawsze działa drag'n'drop, mogą pojawić się kłopoty przy korzystaniu z list plików. Jeśli program jest wywoływany z menu kontekstowego jeśli w eksploratorze zaznaczonych jest kilka plików, do listy dodawany jest tylko pierwszy zaznaczony plik. Związane jest to z tym, że system Windows nie obsługuje takich sytuacji w poprawny sposób.

[5] Historia.

15/11/1997 - "Ogonki 97" 1.3

- zmieniony interfejs użytkownika
- obsługa projektów edytora HTML "Pajaczek" w wersji 2.0 i 3.0 Pro oraz HomeSite'a w wersji 2.5 i 3.
- większa kontrola nad przebiegiem konwersji
- poprawiony błąd polegający na tym, że czasem informacja o konwerterze wstawiana była kilkakrotnie. Teraz "Ogonki" wstawiają już ją prawidłowo, a "nadmiarowe" - usuwają
- zmienione okno "Dodaj zawartość foldera"
- możliwość używania wielu nazw plików i folderów jako parametrów przy wywoływaniu programu
- kilka mniej istotnych udogodnień i poprawek

5/10/1997 - "Ogonki 97" 1.2

- strona kodowa pliku zawierającego tekst "UTF-8" jest rozpoznawana poprawnie
- dodana nowa opcja w menu kontekstowym listy plików: "Edytuj w Notatniku"
- uaktualniony adres strony Marcina "Qrczaka" Kowalczyka ;-)

4/10/1997 - "Ogonki 97" 1.1

- "Ogonki" zachowują się poprawnie, jeśli zostały wywołane z menu kontekstowego plików lub folderów, których nazwy zawierają spację
 - nowa funkcja: można wyłączyć uaktualnianie czasu oraz daty ostatniej modyfikacji przy konwertowaniu plików.
- Gdy opcja ta jest włączona, a dany plik HTML był utworzony lub poprawiony np. 1 czerwca oraz konwertowany 5 czerwca, data modyfikacji tego pliku jest zmieniana na 5 czerwca. Jeśli opcja ta jest wyłączona, data i czas modyfikacji nie są zmieniane w czasie konwersji.

30/09/1997 - "Ogonki 97" 1.0

Nowa wersja "Ogonków 97", wiele zmian i poprawionych błędów:

- dokumenty > 64 KB konwertowane są już poprawnie
- automatyczne rozpoznawanie kodowania zostało udoskonalone
- konwersja entities -> 8-bit została poprawiona i zoptymalizowana
- "Ogonki" nie mają już problemów z plikami, których nazwa składa się z wielkich liter (np. stworzonymi w DOSie).
- pojawiły się skróty klawiaturowe (=> 2B, przedostatni paragraf). Można uruchomić i zakończyć konwersję oraz dodać zawartość katalogu klawiszem "Enter".
- wygląd programu trochę zmieniony: nowa ikona, logo, ikonki. "Ogonki" wykorzystują font "Tahoma" (jeśli jest dostępny).
- dokumentacja w formie Pomocy
- udoskonalona zmiana tagu <meta ... charset>
- przyspieszenie pracy programu
- ogonki w tematach plików zapisanych w ISO-8859-2 wyświetlane są prawidłowo

Dziękuję beta-testerom "Ogonków" za informacje o zauważonych błędach i kilka interesujących pomysłów.

22/08/1997 - "Ogonki 97" 1.0 beta

Pierwsza wersja "Ogonków" udostępniona w Internecie.

[6] O autorze.

Nazywam się Michał Jaskólski, mam 18 lat. Mieszkam w Rumi a uczę się w LO nr 3 w Gdyni. Zamierzam studiować zarządzanie. W wolnych chwilach programuję w Delphi :-) i zajmuję się "webmasterowaniem".

[7] Umowa licencyjna i gwarancja.

Freeware

Niniejszym autor udziela licencji na: używanie programu; wykonywanie dowolnej ilości kopii

programu; rozdawanie dokładnych kopii oryginału komukolwiek; dystrybucję programu w niezmienionej formie przez media elektroniczne. Nie ma konieczności uiszczania jakichkolwiek opłat za powyższe czynności.

Autor nie udziela zezwolenia na pobieranie opłat (za wyjątkiem rozsądnych kwot pokrywających koszty kopiowania itd.) za kopię tego programu, jakkolwiek dokonywane, ani też na rozprowadzanie tego programu i / lub dokumentacji razem z innymi produktami bez jego zgody.

Gwarancja

Oprogramowanie to i wszystkie towarzyszące pliki jest rozprowadzane "jak jest", bez żadnych gwarancji. Ponieważ program ten może być użyty w różnych środowiskach sprzętowych i programowych, autor nie oferuje żadnych gwarancji co do tego, czy program nadaje się do specyficznych zastosowań.

Praktyka komputerowa nakazuje, aby jakiegokolwiek oprogramowanie było przetestowane z niekrytycznymi danymi, zanim zacznie się na nim polegać. Użytkownik przyjmuje na siebie całe ryzyko używania tego programu. W żadnym wypadku autor programu, ani też żaden dystrybutor oprogramowania nie może być pociągnięty do odpowiedzialności za jakiegokolwiek szkody czy straty spowodowane przez użycie lub nadużycie tego programu.

Wszystkie użyte w programie oraz dokumentacji znaki firmowe lub towarowe należą do ich właścicieli i zostały użyte wyłącznie w celach informacyjnych.

