



**AKADEMIA GÓRNICZO-HUTNICZA  
IM. STANISŁAWA STASZICA W KRAKOWIE**

# **HTML 5.0**

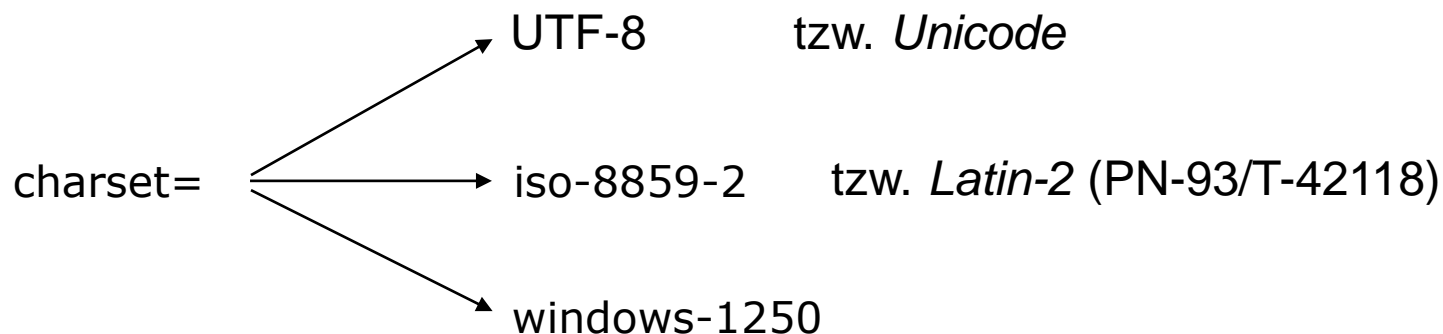
**metainformacje, polskie znaki diakrytyczne**

**Tomasz Bartuś**

**Wydział Geologii, Geofizyki i Ochrony Środowiska  
Katedra Geologii Ogólnej i Geoturystyki  
Kraków, 2023**

```
<head>
  <meta charset="UTF-8">
  <meta name="description" content="Osady jury górnej na Wyżynie
Krakowsko-Wieluńskiej">
  <meta name="keywords" content="Kraków, jura, amonit">
  <meta name="author" content="Tomasz Bartuś">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
</head>
```

## Wybór strony kodowej



```
<meta charset="UTF-8">
```

```
<meta charset="iso-8859-2">
```

```
<meta charset="windows-1250">
```

**Kod ASCII** - American Standard Code – brak „krzaczkowatych znaków”, **7 bitów**,

Nawet ósmy bit nie wystarczał do zakodowania wszystkich znaków. Utworzono zatem *strony kodowe*.

Np.: ISO 8859

JAK TO DZIAŁA? – to po prostu tabele

strona kodowa:	bajt kodu znaku:
	<b>0xA3</b>
ISO-8859-1	£ (znak funta)
ISO-8859-2	Ł (wielka polska litera Ł)

Jednak wszystkie kodowania mają jedną wspólną część – pierwsze **128 znaków** (te, które da się zapisać za 7 bitach) to **podstawowe kodowanie ASCII**.

## Strony kodowe

**iso-8859** (Latin 1) – 1987 r. ISO opracowało standard kodowania znaków w różnych zestawach charakterystycznych dla różnych języków (128 znaków ASCII + dodatkowe charakterystyczne dla języków zachodnioeuropejskich) – **8 bitów**.

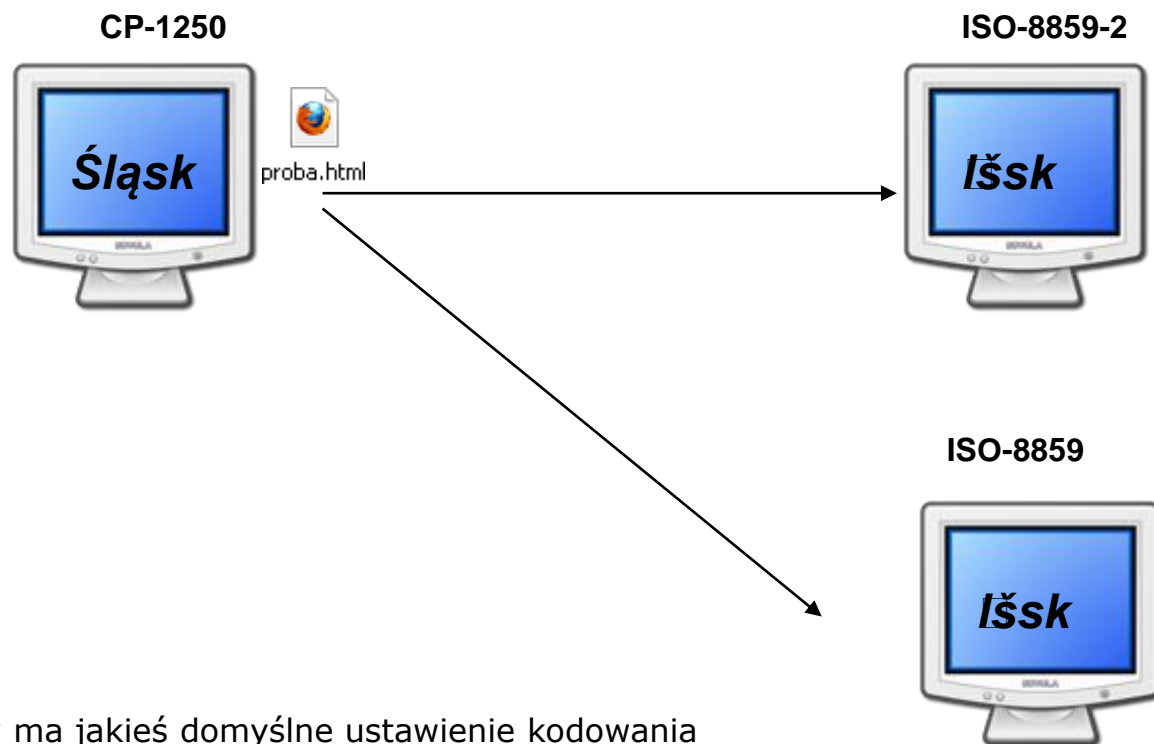
**iso-8859-2** (Latin 2) – 1991 r. Polski Komitet Normalizacji odpowiedzialny za opracowanie standardu kodowania polskich znaków wydał normę PN-91/T-42115, w której utworzył standard kodowania polskich znaków – **8 bitów**.

## Kodowanie znaków w Windows

**NIESTETY** opracowany standard nie zostały przyjęte przez Microsoft i w Windows ISO nie jest standardem.

Istnieje za to **Windows CP-1250** – wraz z Windows 3.11 Microsoft wprowadził własne **8 bitowe** kodowanie znaków bazujące na standardzie ASCII,

strona kodowa:	bajt kodu znaku:
	<b><i>0xB1</i></b>
CP-1250	± (znak +/-)
ISO-8859-2	ą (mała polska litera ą)



Każdy z komputerów ma jakieś domyślne ustawienie kodowania

Dane przesyłane przez sieć to po prostu strumień bajtów. Kodowanie odpowiada jedynie za ich graficzną prezentację i nie zmienia ich zawartości.

Dlatego trzeba w html dodawać informację o wybranej stronie kodowej.

**Problem z formularzami na stronach html !!!**

- Aby rozwiązać problem utworzono jedną, wspólną tablicę znaków dla wszystkich języków - **zarówno łacińskich**, jak i tych opartych o **cyrylicę**, pismo **arabskie**, **hebrajskie**, czy nawet języków takich jak **chińskie**, **japońskie**, czy **koreańskie**.
- Standard ten nazywa się **Unicode**. W odróżnieniu od poprzednio omawianych tablic znaków nie ogranicza on zapisu znaku do jednego bajta dzięki czemu do dyspozycji jest znacznie szersza przestrzeń wartości pozwalająca na pomieszczenie wszystkich znaków obecnie używanych alfabetów



- Mało tego unikod obejmuje także wiele znaków graficznych, symbolik (układy scalone, elektryczne, wzory matematyczne) oraz znaków interpunkcyjnych, które były pomijane w normalnych kodowaniach.

## Unicode vs. UTF

UTF - *Unicode Transformation Format* jest sposobem na binarny zapis znaków Unicode.

Istnieje kilka wariantów UTF (zależą od długości słowa):

**UTF-8** 8–16 bitów (1–2 bajty),

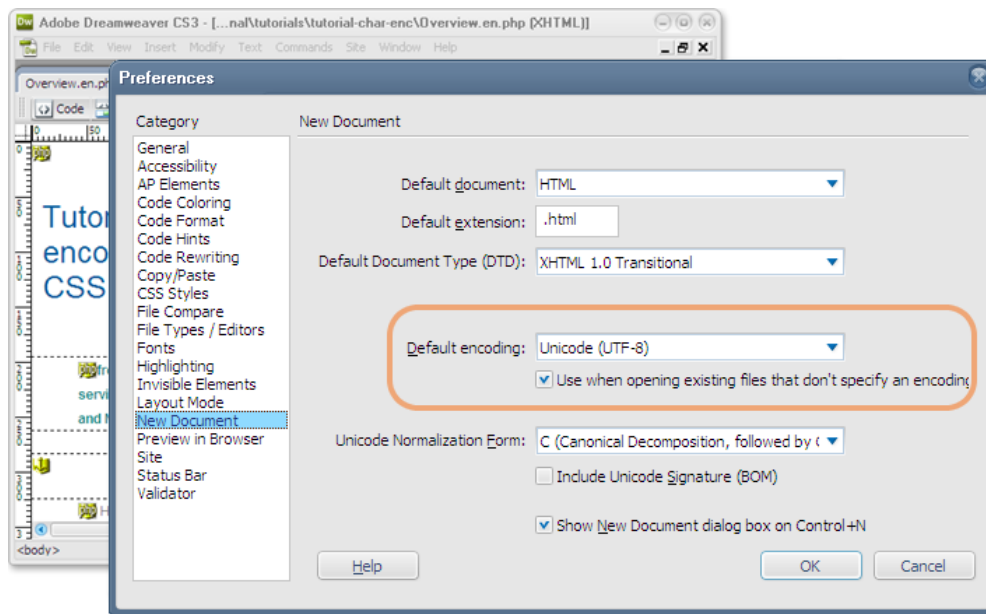
**UTF-16** 16 bitów (2 bajty),

**UTF-32** 32 bity (4 bajty),

UTF-8 znaki ASCII koduje w 8 bitach, a znaki diakrytyczne w 16 bitach.

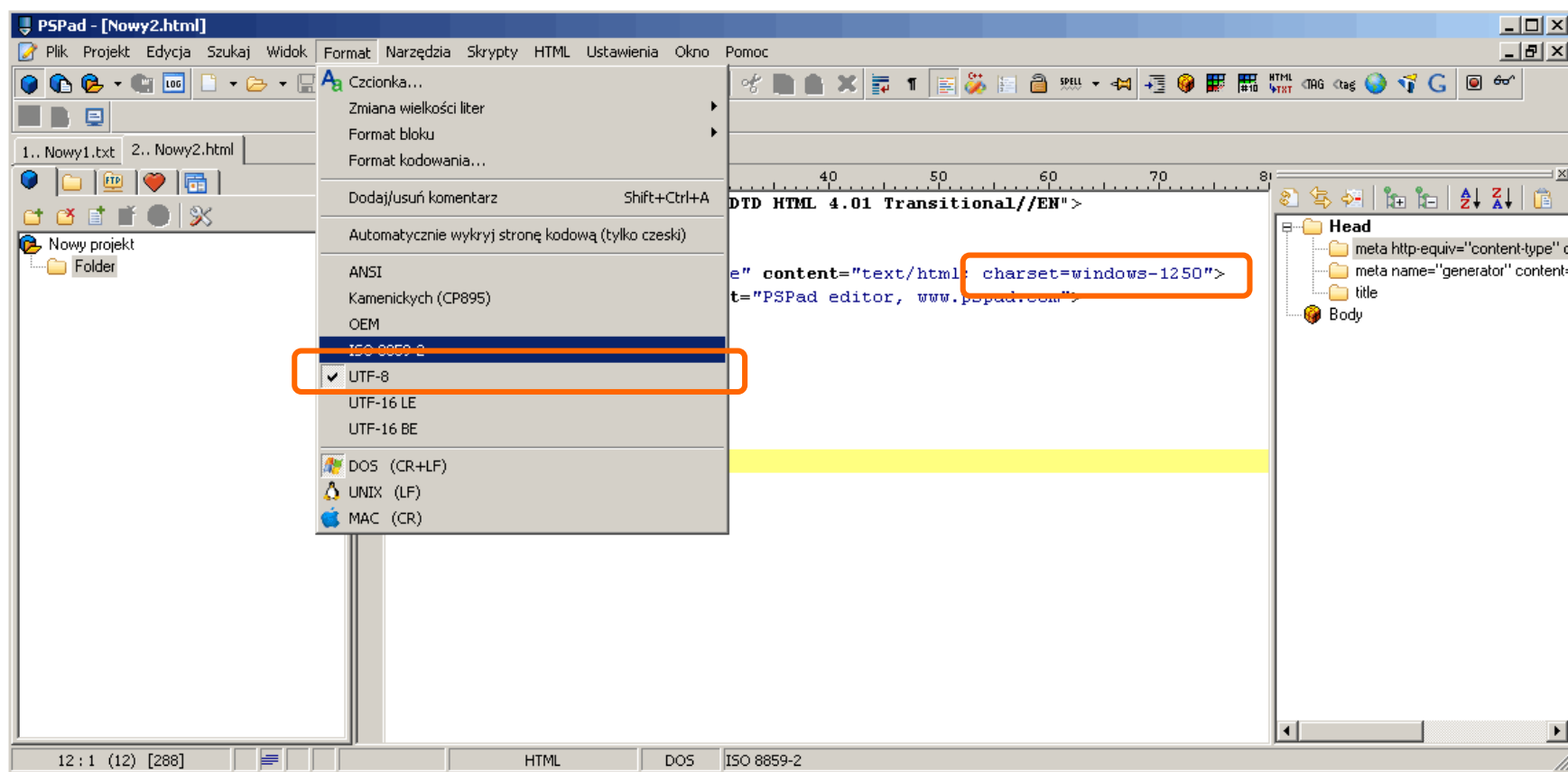
# Wybór strony kodowej

Komercyjny, profesjonalny edytor html **Adobe Dreamweaver**



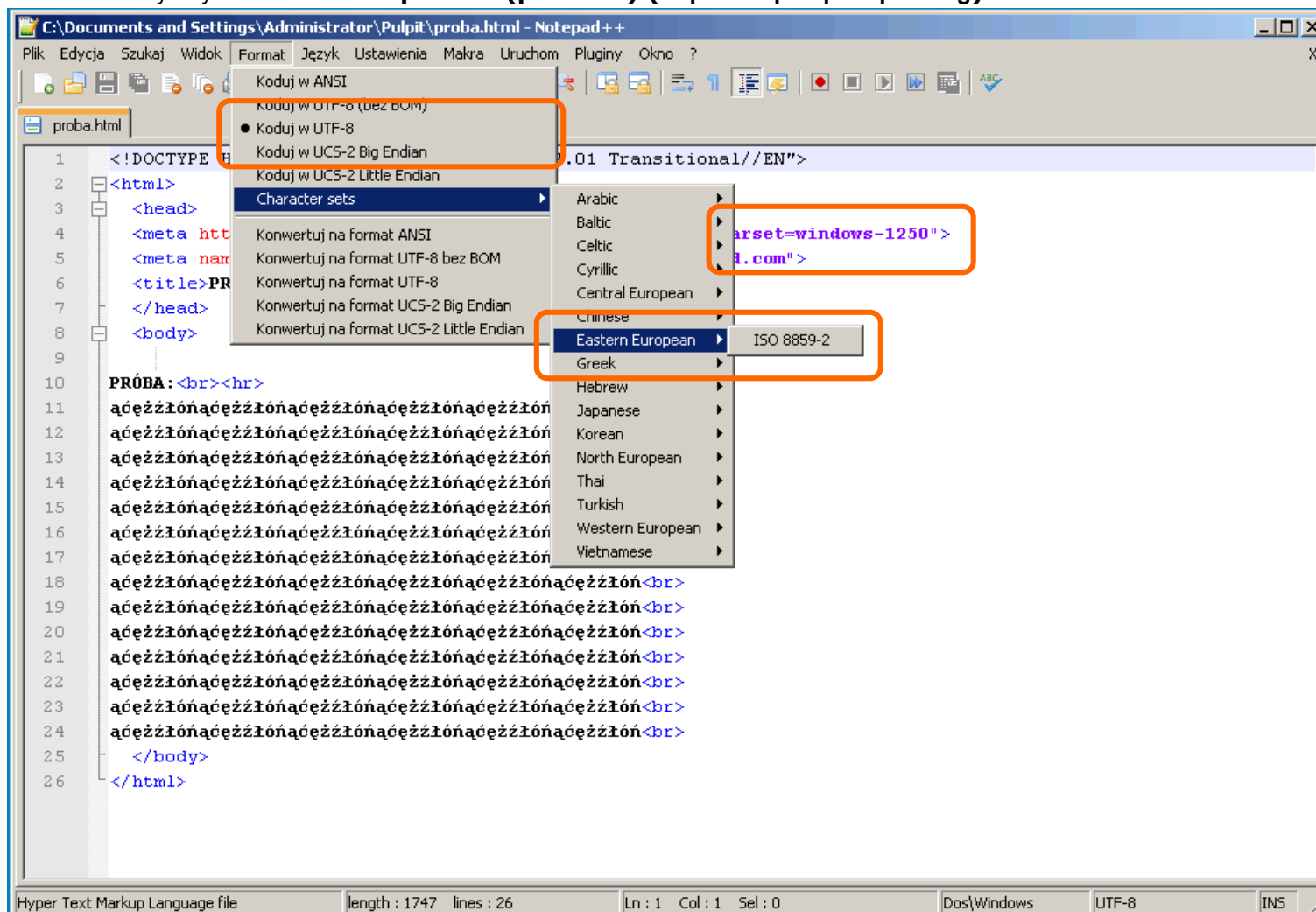
# Wybór strony kodowej

Freeware'owy edytor kodów **PSPad** [<http://www.pspad.com>]



## Wybór strony kodowej

Freeware'owy edytor kodów **Notepad ++ (portable)** (<http://notepad-plus-plus.org>)

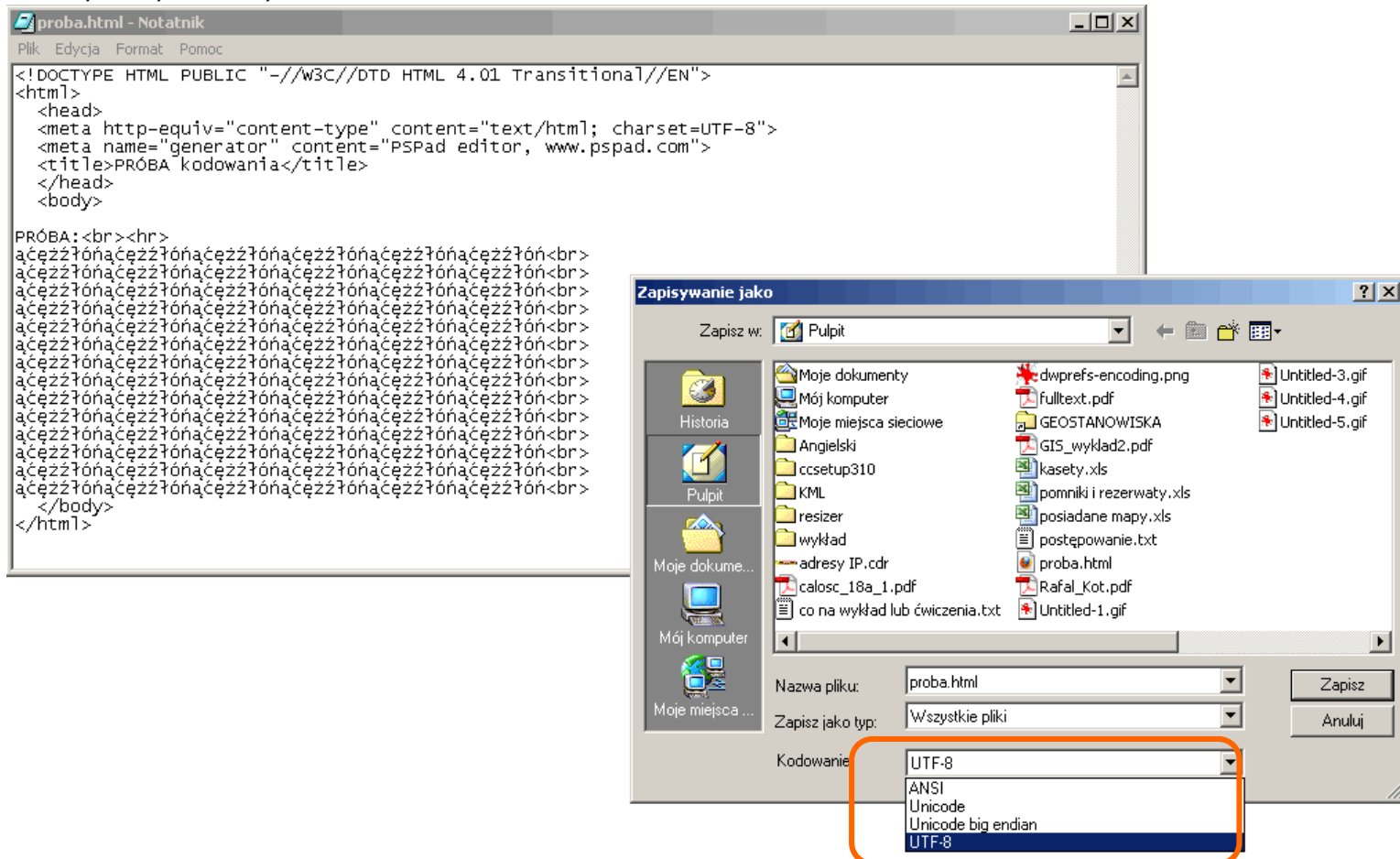


The screenshot shows a web browser window with a title bar that reads "proba.html - Notatnik". The address bar shows a file path: "C:\Program Files\Internet Explorer\proba.html". The main content area displays a page with a broken image icon and the text "PRÓBA: <br><br>". Below the text, there is a list of font names: "Lucida Console", "Lucida Sans Unicode", "M42\_FLIGHT 721", "Marlett", "Microsoft Sans Serif", "Modern", and "Monotype Corsiva". The "Lucida Console" font is selected. To the right of the font list, there are two columns: "Styl czcionki:" (Font style) and "Rozmiar:" (Size). The "Styl czcionki:" column shows "Regularne" (Regular) and "Kursywa" (Italic). The "Rozmiar:" column shows "10", "11", "12", "14", "16", "18", and "20". The "10" size is selected.

The screenshot shows the 'Czcionka' (Font) dialog box. The 'Czcionka' (Font) list has 'Lucida Console' selected. The 'Styl czcionki' (Font style) list has 'Regularne' (Regular) selected. The 'Rozmiar' (Size) list has '10' selected. The 'Przykład' (Preview) box shows the text 'AaBbÁáÓó'. The 'Skrypt' (Script) dropdown is highlighted with an orange box and shows 'Europa Środkowa' (Central European) selected.

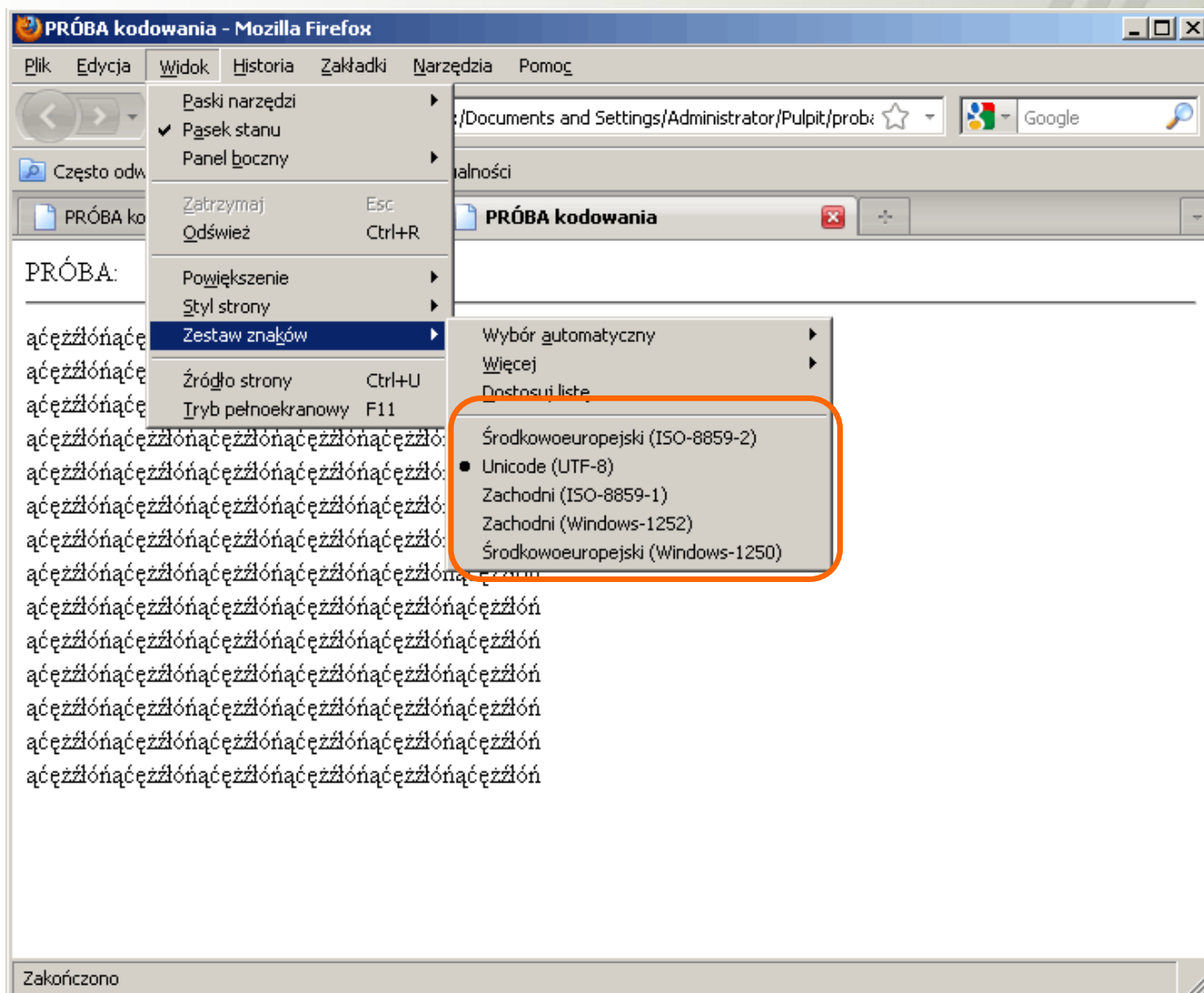
## Wybór strony kodowej

## Edytor systemowy Windows - Notatnik



## Notatnik obsługuje polskie znaki w standardzie **windows-1250** lub **UTF-8**

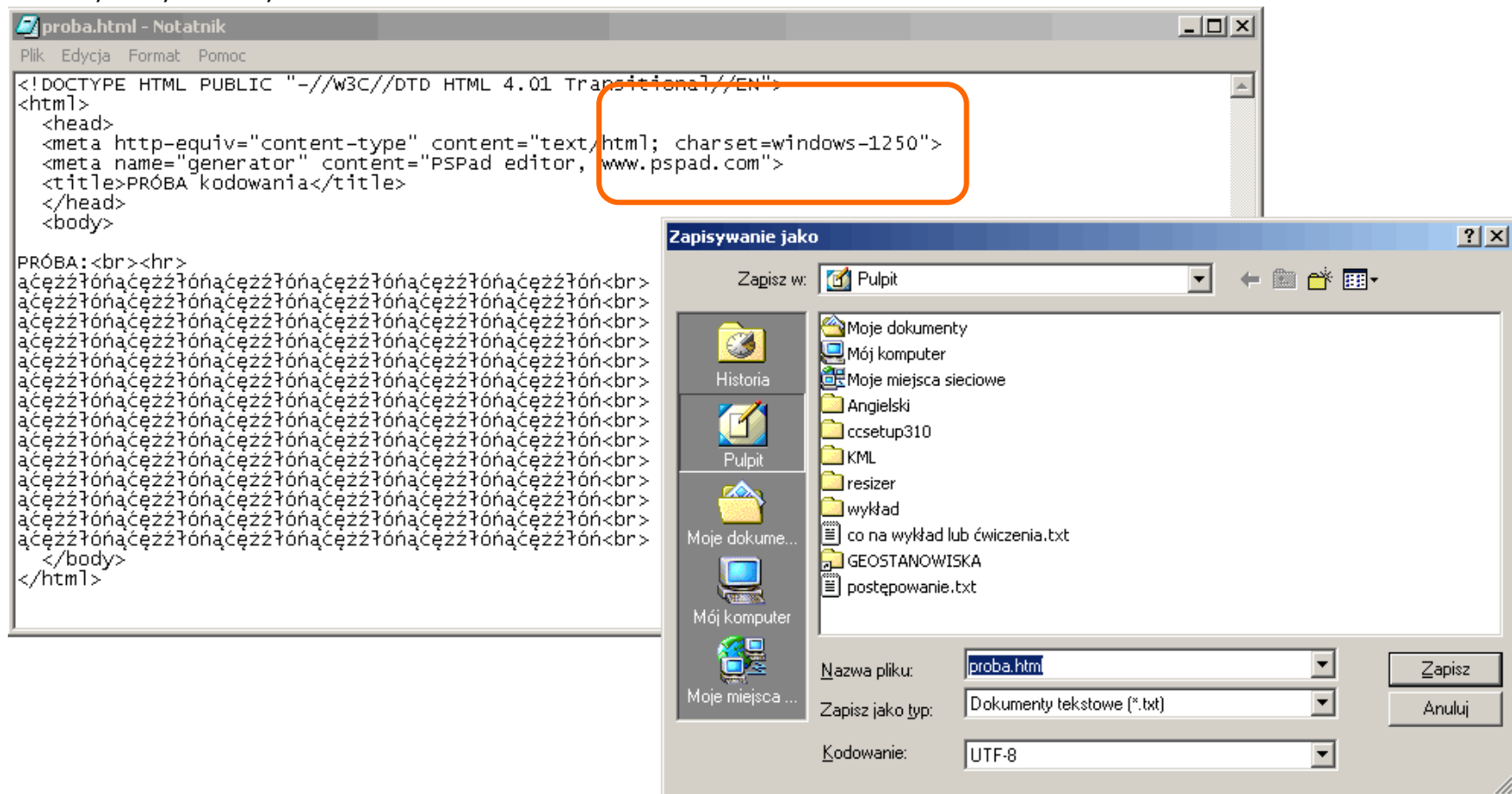
# Efekt





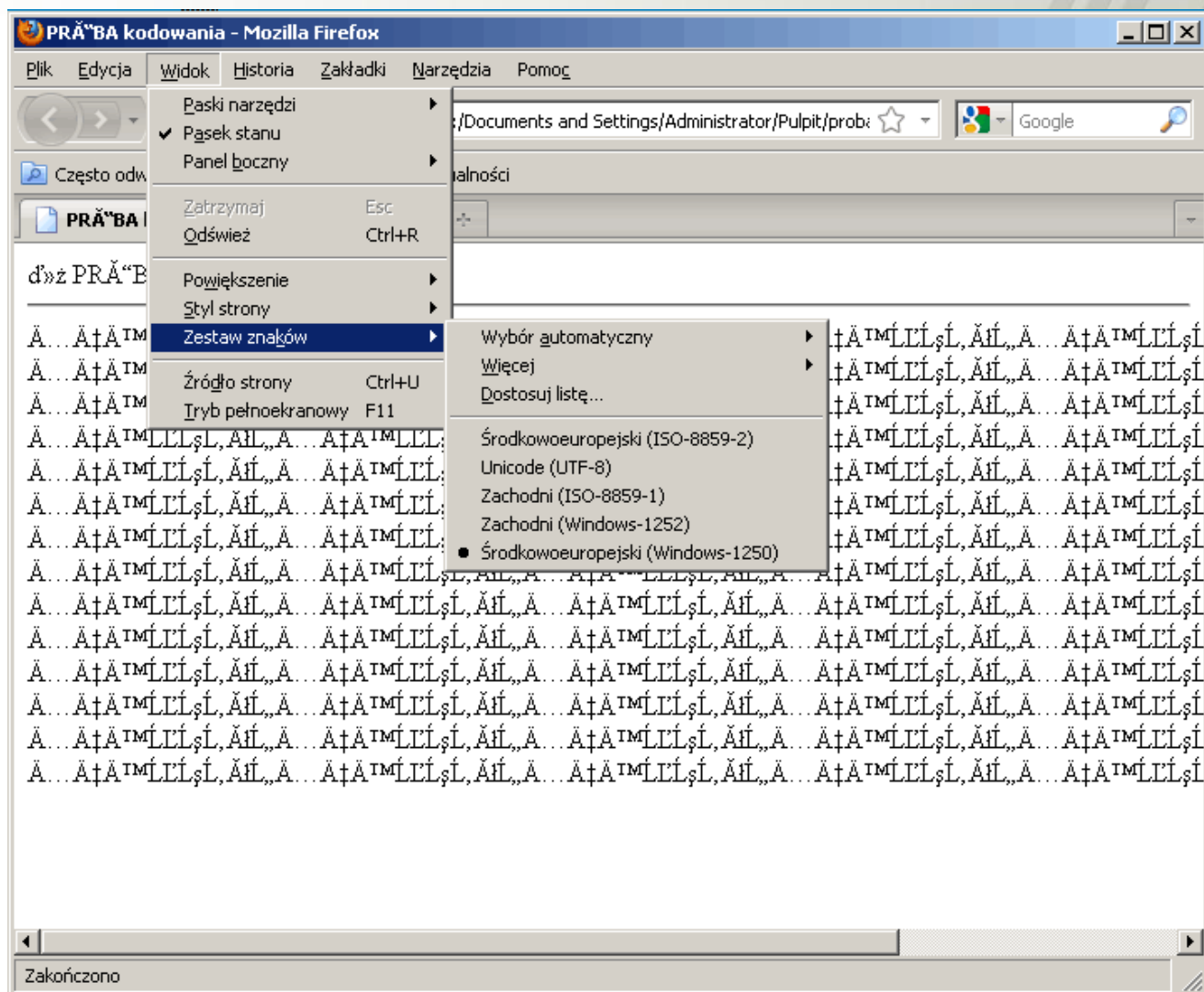
## Wybór strony kodowej (kodowanie znaków dokumentu)

## Edytor systemowy Windows - **Notatnik**



Pozostawiam kodowanie UTF-8, ale zmieniam deklarację charset na **windows-1250**

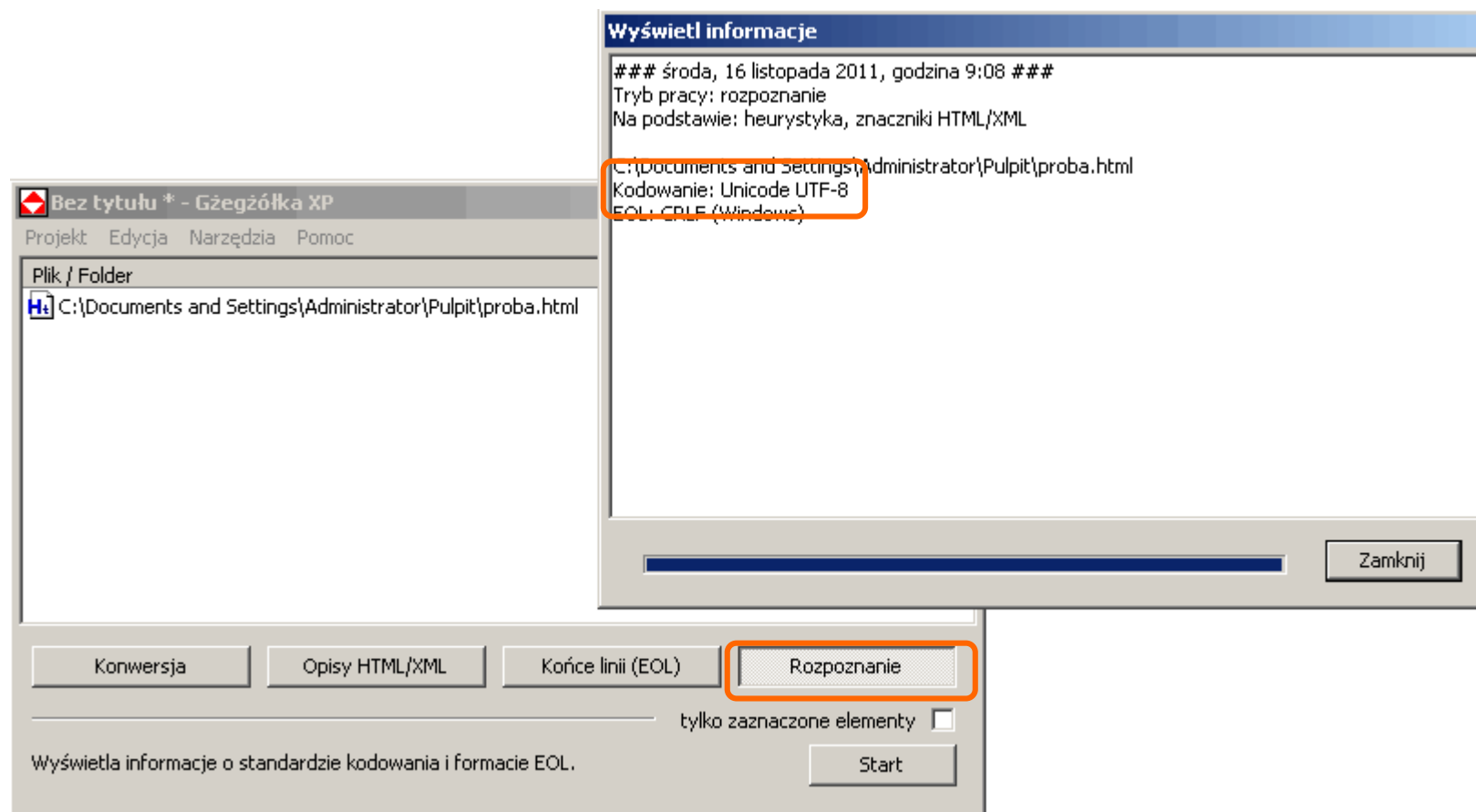
# Efekt zmiany deklaracji charset

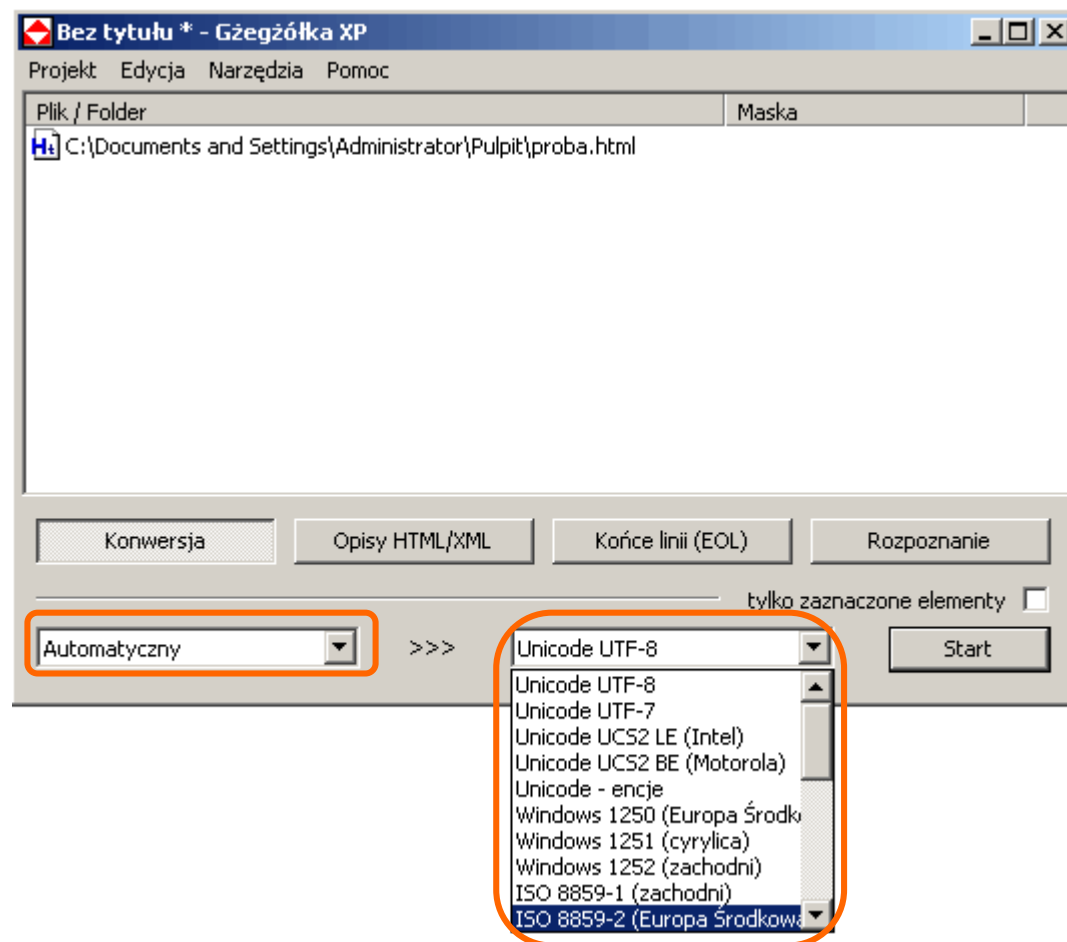


## Podsumowanie

Gdy edytujesz dokument HTML w Notatniku to:

- musisz zmienić deklarację sposobu kodowania w `<meta>` na zgodną z faktycznie wybraną stroną kodową (**UTF-8** lub **windows-1250**),
  - albo pozostawić deklarację kodowania bez zmian ale za to po zakończeniu edycji **przekonwertować** polskie znaki edytowane za pomocą Notatnika do wybranego standardu,
- zalecana zmiana edytora na „**PSPad**” lub „**Notepad ++**”





## Wybór strony kodowej

Jeśli zechcesz stosować stronę kodową **iso-8859-2** (nie zalecane), wykonaj następujące działania:

1. Zaopatrz się w dowolną czcionkę ekranową w standardzie iso-8859-2, za pomocą której będziesz wyświetlać tekst dokumentu w edytorze HTML,
2. Używaj polskiej klawiatury w standardzie iso-8859-2,
3. Wstaw deklarację strony kodowej iso-8859-2 w nagłówku dokumentu (**charset=iso-8859-2**),

Jeśli zamierzasz stosować stronę kodową **Windows-1250** (nie zalecane):

1. Użyj dowolnej czcionki ekranowej w tym standardzie (jest ich sporo w Windows),
2. Używaj standardowej klawiatury Windows,
3. Wstaw deklarację strony kodowej Windows-1250 w nagłówku dokumentu (**charset=windows-1250**),

## **Wybór strony kodowej**

Obecnie zaleca się używanie UTF-8

```
<meta name="description" content="Osady jury górnej na  
Wyżynie Krakowsko-Wieluńskiej">
```

< **150-200** znaków

max **20-25** wyrazów



## Słowa kluczowe

```
<meta name="keywords" content="HTML, CSS, JavaScript">
```

- Słowa kluczowe są oddzielane **przecinkami**,
- **kilka**, rzadziej **kilkanaście** wyrazów. Kilkadziesiąt to już absolutny wyjątek. Generalnie **maksymalnie 1000 znaków**,
- nie powtarzaj wyrazów kluczowych, chcąc "podbić" swój ranking (np. *sex, sex, sex*), gdyż wiele wyszukiwarek może to potraktować jako spam.

```
<meta http-equiv="X-UA-Compatible" content="IE=Edge,chrome=1">
```

- Zapis wymusza na przeglądarce IE renderowanie strony w najwyższej dostępnej w systemie wersji IE.
- Chrome=1 to wymóg wyświetlenia strony w emulatorze Chrome Frame dla IE.

```
<meta name="author" content="Tomasz Bartuś">
```

```
<meta name="generator" content="Notepad++">
```

## Odświeżanie strony www

```
<meta http-equiv="refresh" content="30">
```

polecenie spowoduje regularne odświeżanie strony co 30 sek. Może to mieć praktyczne zastosowanie w przypadku, gdy strona zawiera bardzo często aktualizowane informacje (wiele razy dziennie).

**Uwaga:** nie stosuj tego polecenia jako narzędzia do "podbijania" wartości swojego licznika wizyt. Goście Twojej strony prędko zrezygnują z wizyt, gdy w trakcie czytania strona będzie się nagle odświeżała.

```
<meta name="viewport" content="width=device-width,  
initial-scale=1.0">
```

Deklaracja definiuje obszar strony internetowej. Różni się w zależności od urządzenia - na telefonie komórkowym będzie mniejszy niż na ekranie komputera.

Daje to przeglądarce instrukcje dotyczące kontrolowania wymiarów i skalowania strony.

```
<meta name="viewport" content="width=device-width,  
initial-scale=1.0">
```

`width=device-width` ustawia szerokość strony zgodnie z szerokością ekranu urządzenia (która będzie się różnić w zależności od urządzenia).

Część `Initial-scale=1.0` ustawia początkowy poziom powiększenia, gdy strona jest ładowana po raz pierwszy przez przeglądarkę.

Oto przykład strony internetowej bez metatagu `viewport` i tej samej strony internetowej z metatagiem `viewport`:







# Bibliografia

HTML <meta> Tag [https://www.w3schools.com/tags/tag\\_meta.asp](https://www.w3schools.com/tags/tag_meta.asp)