

Archiwum  
Prac  
Dyplomowych

Michał Malinowski  
Dariusz Działak

# Spis treści

- pierwszy temat pracy
- mała zmiana... tematu
- standardy W3C
- template
- jaki system istnieje
- prezentacja naszej pracy
- co będzie zrobione

# Pierwszy temat pracy

- znajdowanie podobieństw w kodach źródłowych studentów
- rozpoznawanie kilku języków: c, java, itd.
- porównywanie tekstów i analiza semantyczna źródeł
- wybór języka programowania - D

# Co to jest język D

- następca C
- tworzony od grudnia 1999 przez jednego człowieka
- wiele cech z C++, Javy i C#
- binarna kompatybilność z C
- bardziej przenaszalny od C/C++



# Zalety języka D

- zarządzanie pamięcią
- szybkość C, a nie Javy
- testowanie programów: assertion, invariants, unittesting
- wbudowane ułatwienia w język - np. wersjonowanie, tablice asocjacyjne
- wiele innych usprawnień w stosunku do C i C++: moduły, interfejsy zamiast .h

# Kawałek kodu...

```
int[ char[] ] keywords;
```

```
char[] abc = "bar";
```

```
...
```

```
keywords["foo"] = 3; // initialize table
```

```
...
```

```
if (keywords[abc] == 3) // look up keyword
```





# Kawałek kodu...

```
class Day
{
    int d;
    invariant()
    {
        assert(d >= 1 && d <= 31);
    }
}
```



# Dlaczego nie warto używać D

- wciąż w rozwoju
- brak bibliotek
- brak dobrego kompilatora
- zbyt małe wsparcie
  - brak debbugera
  - edytora
  - tutoriali
- głównie rozwija go jedna osoba



# Nowy temat ;-)

- połowa stycznia,
- propozycja JMD – potrzeba nowego modułu do USOSa,
- małe odejście od plagiatów
- koncentracja na zarządzaniu pracami
- czyli Archiwum Prac Dyplomowych (APD)

# Obecne archiwum na UW

- jedna osoba odpowiedzialna za archiwum,
- ręczne przepisywanie danych z USOS do APD,
- archiwizacja prac z płyt CD
- czyli
  - nadmiar pracy
  - błędy i braki

# Co można z tym zrobić

- wyeliminować udział pracowników UW przy organizacji archiwum
  - zrzuć odpowiedzialność za wprowadzanie danych na (biednych) studentów
  - automatycznie zarządzać archiwum
  - wykorzystywać dane USOSa i kopiować zamiast przepisywać

# Dodatkowe oczekiwania

- elektroniczne archiwum prac dyplomowych (może i internetowe)
- udostępnianie prac i informacji o nich
- wyszukiwanie plagiatów

# Przykład Archiwum Prac Dyplomowych

- Wyższa Szkoła Informatyki i Zarządzania (WSIZ)
- udostępniają publicznie informacje nt. prac dyplomowych ich absolwentów
- w przypadku większości można zobaczyć dołączaną prezentację
- i/lub streszczenie w języku polskim lub angielskim
- [szybkie spojrzenie na stronę]

# Zarys systemu APD

- organizacja archiwum
- system webowy (PHP + MySQL + ???)
- konwersja prac dyplomowych do czystego tekstu (Java) i
- możliwość porównywania prac
- komunikacja z USOS'em



# Dodatkowe cechy APD

- zgodność ze standardami przy tworzeniu serwisu webowego:
  - XHTML 1.1
  - CSS 2.1
- używanie CSS zgodnie z przeznaczeniem
- bezpieczeństwo

# (X)Tag

- gwarancja zgodności ze standardem W3C (xhtml 1.1)
- posiada zaszyte wymagania xhtml 1.1
- bezlitośnie testuje poprawność stron wraz z każdym otwarciem
- druga wersja – komercyjna (bez testowania)
- [krótka prezentacja]

# Strony zgodne ze standardem

- [validator.w3.org](http://validator.w3.org)
- ???

# Strony nie zgodne

- USOSweb – 48 błędów
- MIMUW – 23
- UW – 29
- [rainbow.mimuw.edu.pl/SR](http://rainbow.mimuw.edu.pl/SR) – 163
- [www.icm.edu.pl](http://www.icm.edu.pl) – 85
- onet.pl – 150
- gazeta.pl – 592
- google.pl – 43
- era.pl – 123
- apache.org – 5
- cała masa innych stron (również te ze znaczkami 'valid page')

# Dlaczego omijamy standardy W3C

- bardzo trudno jest napisać większą stronę zgodną ze standardem
- jeszcze trudniej utrzymać stronę/portal, szczególnie przy częstych zmianach i wielu programistach
- dynamiczne strony sklejjane są z kawałków bez żadnego sprawdzania
- brak (super) środowisk ułatwiających zachowanie standardów

# Serwis webowy

- zrezygnowaliśmy z wykorzystania template'ów
- w zamian sami “próbowaliśmy”:
  - separacji warstwy prezentacji danych (HTML) od warstwy logiki
  - wykorzystywać własny mini-framework
- stworzyliśmy obiektowy dostęp do danych
- a podstawą był (X)Tag



# Standard i po standardzie

- co może dać zgodność ze standardem
  - dobre samopoczucie (bez gwarancji poprawnego wyświetlania strony)
- co daje stosowanie template'ów
  - cache'owanie
  - separacja warstw logiki i wyświetlania
  - przejrzystość kodu
  - łatwiejsze modyfikacje

# Serwis webowy - błędy

- nie skorzystaliśmy z żadnego frameworku
- nie udało się nam oddzielić warstwy logiki i wyświetlania
- stworzymy kod trudny do późniejszej modyfikacji i utrzymania

Co przed nami?



# Co przed nami?



- trochę dodatkowej pracy – wdrażanie nowej technologii
- i większe opóźnienia :(
- ale lepiej późno niż wcale

# USOS

- Uniwersytecki System Obsługi Studiów
- dla nas interesujące - baza danych
  - prac dyplomowych
  - osób
- wydziały dostarczają informacji
  - kto może modyfikować informacje o pracy
- “anomalie” w tabelach
  - klucze zewnętrzne mogą być NULL

# migrator

- łatwe przeniesienie danych między bazami danych ORACLE i MySQL
- mapowanie tabel i kolumn
- plik konfiguracyjny (migrator-mapping.xml)
- polskie znaczki – problem z MySQL 4.1



# Autoryzacja

- do ustalenia
- instalacje WWW
- baza użytkowników na UW –  
single sign-on przygotowywane przez  
ICM
- chwilowo zaślepka (brak autoryzacji)

# Aktualny system USOSweb

- wbudowany moduł Dyplomy
- głównie na MIM UW – wymagane jest logowanie
- dlaczego nie rozszerzyć do APD(?) - centralne archiwum

# moduł Dyplomy

- umożliwia
  - przeglądanie opisów prac
  - wyszukiwanie prac
- brakuje
  - przechowywania samych prac
  - możliwości uzupełniania danych dotyczący pracy przez autora
  - miejsca na podłączenie systemu szukającego plagiaty
- krótki przegląd modułu

# Aktualne archiwum

Krótką prezentacją pracy

# Wykrywanie plagiatów

- druga część pracy magisterskiej
- mniej znacząca
- planujemy na wakacje – szczególnie własny moduł szukający plagiatów
- Java + PDFBox

# Możliwości wykrywania plagiatów

- własne przeszukiwanie internetu
- moduł napisany przez ICM
- [plagiat.pl](http://plagiat.pl)
- udostępnianie czystego tekstu do szukania plagiatów w inny sposób



Pytania?

# Linki

- <http://validator.w3.org> – walidator W3C
- <http://stud.wsi.edu.pl/~czytelnia/> - APD na uczelni WSIZ
- <http://www.xisc.com/> - PRADO

Dziękujemy

za uwagę