



AKADEMIA GÓRNICZO-HUTNICZA  
IM. STANISŁAWA STASZICA W KRAKOWIE

# **Specjalistyczna internetowa wyszukiwarka ogłoszeń przetargowych**

**Andrzej Jasiński**

Wydział Elektrotechniki, Automatyki,  
Informatyki i Elektroniki  
Katedra Automatyki

Promotor: dr Adrian Horzyk

Kraków, 2009

# Plan prezentacji

- 1. Wstęp**
- 2. Cel pracy**
- 3. Rozwiązanie**
- 4. Wyniki**
- 5. Zakończenie**

## Wstęp

- Źródła informacji o przetargach
  - Dedykowane portale.
  - Wyszukiwarki ogólnego zastosowania.
  - Strony BIP (Biuletyn Informacji Publicznej).
  - Ogłoszenia na stronach instytucji.
  - Gazety.

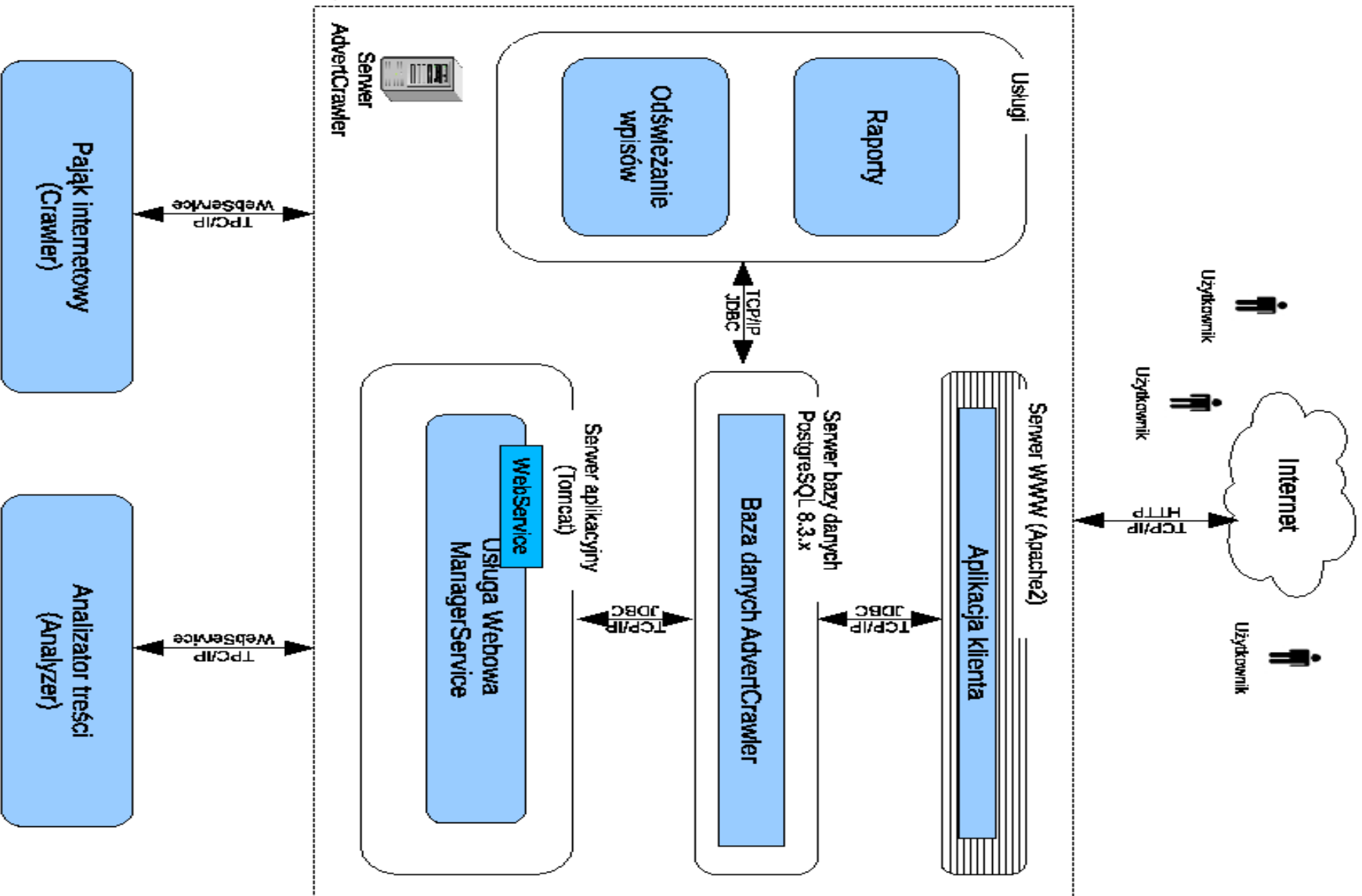
## Cel pracy

- Robot przeszukujący Internet
- Klasyfikacja dokumentów (przetargi, inne strony)
- Wydobywanie informacji z treści ogłoszenia
  - Branża
  - Daty
    - Publikacji
    - Otwarcia ofert
    - Zamknięcia
    - Realizacji

- Problemy
  - Duża ilość informacji w internecie
  - Wielość formatów plików z ogłoszeniami
  - Język HTML
    - Błędy programistów
    - Brak jasnego standardu
  - Niestandardowe postacie ogłoszeń

- Główne elementy systemu:
  - Pająk: pobiera strony z sieci, dokonuje szybkiej analizy i odszukuje odnośniki do innych stron,
  - Analizator: próbuje ustalić podstawową branżę przetargu, wyszukuje istotne daty,
  - Odświeżacz: dodaje strony, na których jest wiele ogłoszeń, do ponownej analizy,
  - Generator raportów: tworzy raporty dla użytkowników.

# Rozwiązanie



# Rozwiązanie

- Pająk
  - Pobranie dokumentu
  - Wstępna kategoryzacja
  - Wydobycie odnośników
  - Nadanie odnośnikom wag
  - Zapis do odnośników do kolejki pająka
  - Jeśli znaleziono potencjalny przetarg – zapis do kolejki Analizatora



- Analizator
  - Ocena dokumentu
  - Określenie branży
  - Wydobycie informacji o datach
  - Wysłanie dokumentu do zapisania w bazie

- Kategoryzacja dokumentów
  - **Metody wektorowe.**
  - **Metody probabilistyczne.**
  - Drzewa decyzyjne.
  - **Reguły decyzyjne.**
  - Sieci neuronowe.
  - Algorytmy genetyczne.

- Metody wektorowe
  - TF
- Metody probabilistyczne
  - Klasyfikator Bayesa (naiwny)
- Reguły decyzyjne
  - Tworzenie reguł
  - Możliwość rozróżniania homonimów

- Generator raportów
  - Okresowe przeglądanie bazy w poszukiwaniu nowych ogłoszeń, spełniających kryteria podane w definicji
- Odświeżacz
  - Dodanie do kolejki Pająka stron, na których znaleziono ogłoszenia przetargowe
  - Usuwanie z bazy nieistotnych odnośników

## • Etapy działania : ekstrakcja tekstu



**WOJEWÓDZKI SZPITAL PODKARPACKI im. Jana Pawła II w Krośnie**  
tel. (13) 4378000  
fax (13) 4378204

Strona główna  
Przetargi aktualne  
Przetargi archiwalne

**Informacja skrócona**

NZ/215/23/2009 *Termomodernizacja budynków Wojewódzkiego Szpitala Podkarpackiego im. Jana Pawła II w Krośnie.*

Numer : NZ/215/23/2009

Krosno: Nz/215/23/2009 Data składania: 2009-03-10  
Termomodernizacja budynków Data otwarcia: 2009-03-10  
Wojewódzkiego Szpitala Podkarpackiego im. Jana Pawła II w Krośnie

Numer ogłoszenia: 37769 - 2009; data zamieszczenia: 13.02.2009  
OGŁOSZENIE O ZAMÓWIENIU - roboty budowlane

Zamieszczanie ogłoszenia: obowiązkowe.  
Ogłoszenie dotyczy: zamówienia publicznego.  
SEKCJA I: ZAMAWIAJĄCY  
I. 1) NAZWA I ADRES: Wojewódzki Szpital Podkarpacki im. Jana Pawła II, ul. Korczyńska 57, 38-400 Krosno, woj. podkarpackie, tel. 013 4378497, 4378215, faks 013 4378497, 4378215.

- Adres strony internetowej zamawiającego: [www.krosno.med.pl](http://www.krosno.med.pl)

I. 2) RODZAJ ZAMAWIAJĄCEGO: Podmiot prawa publicznego.  
SEKCJA II: PRZEDMIOT ZAMÓWIENIA  
II.1) OKREŚLENIE PRZEDMIOTU ZAMÓWIENIA  
II.1.1) Nazwa nadana zamówieniu przez zamawiającego: Nz/215/23/2009 Termomodernizacja budynków Wojewódzkiego Szpitala Podkarpackiego im. Jana Pawła II w Krośnie.  
II.1.2) Rodzaj zamówienia: roboty budowlane.  
II.1.3) Określenie przedmiotu oraz wielkości lub zakresu zamówienia: Przedmiotem zamówienia są roboty budowlane w zakresie docieplenia ścian zewnętrznych i stropodachów oraz wymiany stolarki okiennej i drzwi zewnętrznych oraz bram garażowych określone przez: specyfikację techniczną wykonania i odbioru robót (załącznik 4) projekt budowlany i

Wojewódzki Szpital Podkarpacki im. Jana Pawła II w Krośnie  
Szpital  
I509001  
Mikrobiologia  
Akredytacja  
ZPORR

Informacja skrócona  
NZ/215/23/2009 Termomodernizacja budynków Wojewódzkiego Szpitala Podkarpackiego im. Jana Pawła II w Krośnie. <!>  
Numer : NZ/215/23/2009  
Data składania: 2009-03-10  
Data otwarcia: 2009-03-10  
Krosno: Nz/215/23/2009 Termomodernizacja budynków Wojewódzkiego Szpitala Podkarpackiego im. Jana Pawła II w Krośnie  
Numer ogłoszenia: 37769 - 2009; data zamieszczenia: 13.02.2009  
OGŁOSZENIE O ZAMÓWIENIU - roboty budowlane  
Zamieszczanie ogłoszenia: obowiązkowe.  
Ogłoszenie dotyczy: zamówienia publicznego.

SEKCJA I: ZAMAWIAJĄCY  
I. 1) NAZWA I ADRES: Wojewódzki Szpital Podkarpacki im. Jana Pawła II, ul. Korczyńska 57, 38-400 Krosno, woj. podkarpackie, tel. 013 4378497, 4378215, faks 013 4378497, 4378215.  
Adres strony internetowej zamawiającego: [www.krosno.med.pl](http://www.krosno.med.pl)  
I. 2) RODZAJ ZAMAWIAJĄCEGO: Podmiot prawa publicznego.  
SEKCJA II: PRZEDMIOT ZAMÓWIENIA  
II.1) OKREŚLENIE PRZEDMIOTU ZAMÓWIENIA  
II.1.1) Nazwa nadana zamówieniu przez zamawiającego: Nz/215/23/2009 Termomodernizacja budynków Wojewódzkiego Szpitala Podkarpackiego im. Jana Pawła II w Krośnie.  
II.1.2) Rodzaj zamówienia: roboty budowlane.  
II.1.3) Określenie przedmiotu oraz wielkości lub zakresu zamówienia: Przedmiotem zamówienia są roboty budowlane w zakresie docieplenia ścian zewnętrznych i stropodachów oraz wymiany stolarki okiennej i drzwi zewnętrznych oraz bram garażowych określone przez: specyfikację techniczną wykonania i odbioru robót (załącznik 4) projekt budowlany i projekty wykonawcze (załącznik 8) przedmiary robót (załączniki 5a, 6a, 7a), dokumentację projektową do docieplenia ścian zewnętrznych i stropodachów, wymiany stolarki okiennej i drzwi zewnętrznych oraz bram garażowych ( załącznik nr 9) oraz kosztorysy budowlane ślepe (załączniki 5b, 6b, 7b).  
II.1.4) Wspólny Słownik Zamówień (CPV): 45.00.00.00-7, 44.22.11.00-6.  
II.1.5) Czy dopuszcza się złożenie oferty częściowej: nie.  
II.1.6) Czy dopuszcza się złożenie oferty wariantowej: nie.  
II.1.7) Czy przewiduje się udzielenie zamówień uzupełniających: nie.  
II.2) CZAS TRWANIA ZAMÓWIENIA LUB TERMIN WYKONANIA: Zakończenie: 30.06.2010.  
(...)

- Etapy działania : analiza i zapis do bazy

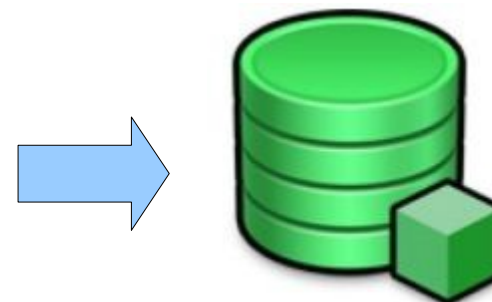
Otoczenie frazy kluczowej

Fraza kluczowa

ZP-340 -10/09 Wielka Wieś, 03.03.2009  
 Wójt Gminy Wielka Wieś ogłasza przetarg nieograniczony poniżej 5.150.000 euro na  
 Wykonanie ciągu pieszego jednostronnego wraz z remontem nawierzchni na szacunkowej długości 240 mb - droga dz. nr 530 w Czajowicach  
 Numer ogłoszenia: 48170 - 2009; data zamieszczenia: 03.03.2009  
 OGŁOSZENIE O ZAMÓWIENIU - roboty budowlane  
 Zamieszczanie ogłoszenia: obowiązkowe.  
 Ogłoszenie dotyczy: zamówienia publicznego.  
 SEKCJA I: ZAMAWIAJĄCY  
 I. 1) NAZWA I ADRES: Gmina Wielka Wieś, ul. Wesoła 48, 32-089 Wielka Wieś, woj. małopolskie, tel. 012 4191704, faks: 012 4191705.  
 • Adres strony internetowej zamawiającego: [www.wielka-wies.pl](http://www.wielka-wies.pl)  
 I. 2) RODZAJ ZAMAWIAJĄCEGO: Administracja samorządowa.  
 SEKCJA II: PRZEDMIOT ZAMÓWIENIA  
 II.1) OKREŚLENIE PRZEDMIOTU ZAMÓWIENIA  
 II.1.1) Nazwa nadana zamówieniu przez zamawiającego: Wykonanie ciągu pieszego jednostronnego wraz z remontem nawierzchni na szacunkowej długości 240 mb - droga dz. nr 530 w Czajowicach.

Data

Otoczenie daty



## Wyniki

- Ilość ogłoszeń według kategorii:

Kategoria	Ilość ogłoszeń
Art.Spoż	2530
Automaty	11 150
Budownictwo	42590
Energetyka	36780
Medycyna	23770
Nieruchomości	42490
Vending	110
Pozostałe	520 200

- Ilość błędnych klasyfikacji

Kategoria	Błędnych
Art.Spoż	5
Automaty	5
Budownictwo	0
Energetyka	3
Medycyna	1
Nieruchomości	5
Vending	10*



- Ilość odnalezionych ogłoszeń dla podanego zapytania

Hasło	Rozw. własne	egospodarka	Google	money.pl**
automat napój	50	0	22 100	0
dostawa leków	7290	8 179	77 600	ponad 1 500
dostawa energii elektrycznej	3960	2 709	57 000	ponad 1 500
marchewka ziemniaki	100	67	6	45

- Ilość prawidłowych ogłoszeń

Hasło	Rozw. własne	egospodarka	Google	money.pl
automat napój	7	nd.	5	nd.
dostawa leków	9*	10	10	8
dostawa energii elektrycznej	10	10	6	7
marchewka ziemniaki	10	10	7	10

## Podsumowanie

- System może dotrzeć do mniej pouplarnych ogłoszeń
- Nieograniczony zasięg
- Duże możliwości rozbudowy
- Elastyczność



# Zakończenie

Dziękuję za uwagę