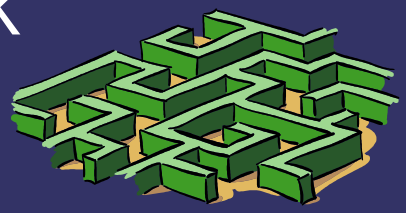
Mikser

mechanizm ułatwiający korzystanie z wyszukiwarki internetu

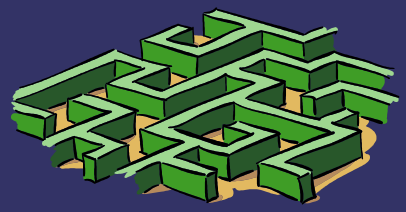
Michał Kosmulski
http://hektor.umcs.lublin.pl/~mikosmul/
mkosmul at users dot sourceforge dot net

Seminarium „Bazy Danych, Inżynieria Oprogramowania, Systemy Rozproszone”
Katedra Inżynierii Oprogramowania PJWSTK
7 stycznia 2008



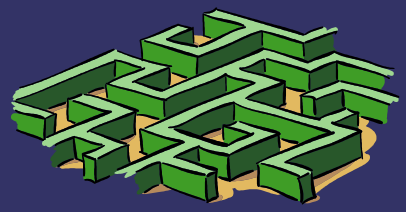
Plan prezentacji

- Definicja i krótki opis Miksera z punktu widzenia użytkownika
- Porównanie z konkurencyjnymi rozwiązaniami
- Struktura projektu
- Sposoby pozyskiwania i przetwarzania danych
- Dyskusja



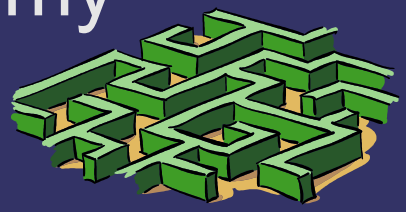
Czym jest Mikser

- Mechanizm współpracujący z wyszukiwarką internetu, który stawia sobie za zadanie usprawnienie korzystania z wyszukiwarki poprzez prezentowanie już na stronie z wynikami wyszukiwania gotowych informacji pochodzących z wiarygodnych źródeł
- Rozwiązanie dla wyszukiwarki NetSprint.pl
- Inspirowany przez ask.com



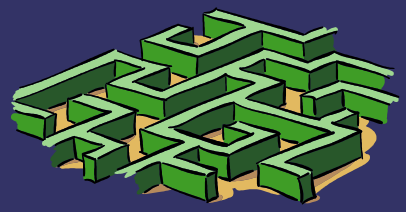
Powody powstania Miksera

- Ogromna ilość informacji w internecie
- Użytkownicy wyszukiwarki:
  - Często szukają bardzo ogólnych haseł, dla których zwykłe algorytmy rankingu kiepsko działają
  - Mogą nie wiedzieć o zaawansowanych funkcjach wyszukiwarki
  - Chcieliby mieć wszystkie informacje w jednym miejscu – również te, z którymi wyszukiwarka WWW kiepsko sobie radzi (np. szybko się zmieniające)
- Szukając w internecie, zwykle chcemy znaleźć informacje, nie strony

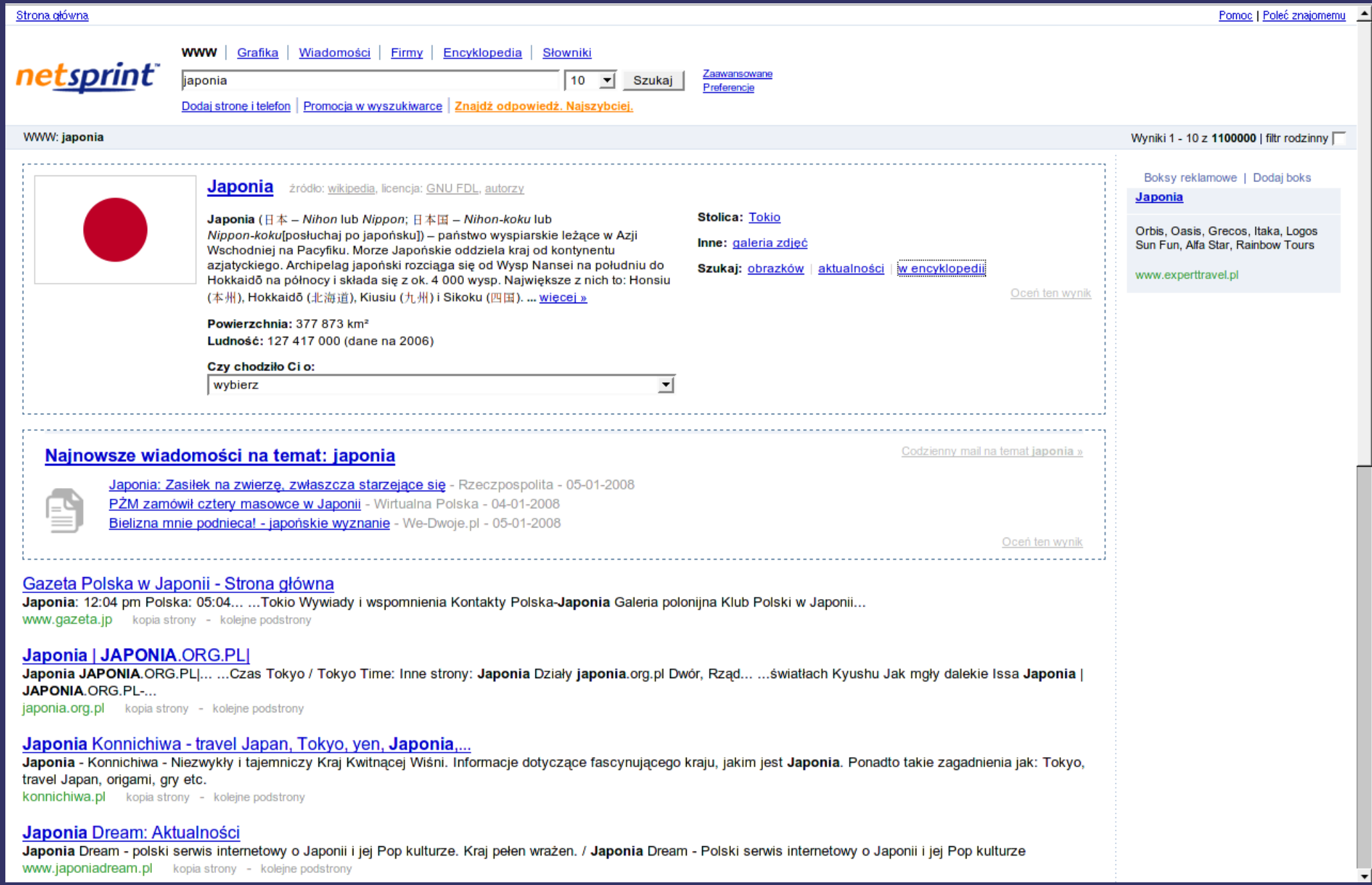


Podstawowe cechy

- Próbujemy zgadnąć „co użytkownik miał na myśli” i dostarczyć najlepszej odpowiedzi
- Mikser stanowi uzupełnienie zwykłych wyników wyszukiwania
- Dzięki zebraniu danych z wielu źródeł w jednym miejscu oraz prezentowaniu ich już na stronie z wynikami wyszukiwania, oszczędzamy czas i ułatwiamy dotarcie do informacji
- Pierwszy tego rodzaju system dla polskich użytkowników

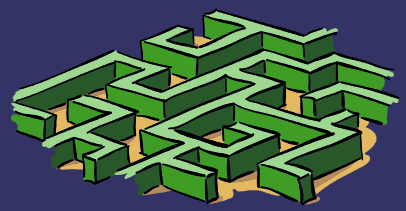


Przykład


Strona główna
Pomoc | Poleć znajomemu
netsprint
WWW | Grafika | Wiadomości | Firmy | Encyklopedia | Słowniki
japonia
10
Szukaj
Zaawansowane
Preferencje
Dodaj stronę i telefon | Promocja w wyszukiwarce | Znajdź odpowiedź. Najszybciej.
WWW: japonia
Wyniki 1 - 10 z 1100000 | filtr rodzinny
Japonia
źródło: wikipedia, licencja: GNU FDL, autorzy
Japonia (日本 – Nihon lub Nippon; 日本国 – Nihon-koku lub Nippon-koku[posłuchaj po japońsku]) – państwo wyspiarskie leżące w Azji Wschodniej na Pacyfiku. Morze Japońskie oddziela kraj od kontynentu azjatyckiego. Archipelag japoński rozciąga się od Wysp Nansei na południu do Hokkaidō na północy i składa się z ok. 4 000 wysp. Największe z nich to: Honsiu (本州), Hokkaidō (北海道), Kiusiu (九州) i Sikoku (四国). ... więcej »
Powierzchnia: 377 873 km²
Ludność: 127 417 000 (dane na 2006)
Stolica: Tokio
Inne: galeria zdjęć
Szukaj: obrazków | aktualności | w encyklopedii
Oceń ten wynik
Czy chodziło Ci o:
wybierz
Boksy reklamowe | Dodaj boks
Japonia
Orbis, Oasis, Grecos, Itaka, Logos Sun Fun, Alfa Star, Rainbow Tours
www.experttravel.pl
Najnowsze wiadomości na temat: japonia
Codzienny mail na temat japonia »
Japonia: Zasiłek na zwierzę, zwłaszcza starzejące się - Rzeczpospolita - 05-01-2008
PŻM zamówił cztery masowce w Japonii - Wirtualna Polska - 04-01-2008
Bielizna mnie podnieca! - japońskie wyznanie - We-Dwoje.pl - 05-01-2008
Oceń ten wynik
Gazeta Polska w Japonii - Strona główna
Japonia: 12:04 pm Polska: 05:04... ...Tokio Wywiady i wspomnienia Kontakty Polska-Japonia Galeria polonijna Klub Polski w Japonii...
www.gazeta.jp kopia strony - kolejne podstrony
Japonia | JAPONIA.ORG.PL|
Japonia JAPONIA.ORG.PL|... ...Czas Tokyo / Tokyo Time: Inne strony: Japonia Działy japonia.org.pl Dwór, Rząd... ...światłach Kyushu Jak mgły dalekie Issa Japonia | JAPONIA.ORG.PL-...
japonia.org.pl kopia strony - kolejne podstrony
Japonia Konnichiwa - travel Japan, Tokyo, yen, Japonia,...
Japonia - Konnichiwa - Niezwykły i tajemniczy Kraj Kwitnącej Wiśni. Informacje dotyczące fascynującego kraju, jakim jest Japonia. Ponadto takie zagadnienia jak: Tokyo, travel Japan, origami, gry etc.
konnichiwa.pl kopia strony - kolejne podstrony
Japonia Dream: Aktualności
Japonia Dream - polski serwis internetowy o Japonii i jej Pop kulturze. Kraj pełen wrażeń. / Japonia Dream - Polski serwis internetowy o Japonii i jej Pop kulturze
www.japoniadream.pl kopia strony - kolejne podstrony

Porównanie z konkurencją

- ask.com
  - Plusy
    - Są dostępne opcje *narrow your search* i *expand your search*
    - Prezentuje wyniki wyszukiwania wideo oraz grafiki wraz z miniaturami
  - Minusy
    - Brak polskiej wersji językowej (interfejs, dane, priorytety haseł)
    - Trudna nawigacja: nie da się łatwo trafić z wyszukiwania hasła „Warszawa” do opisu samochodu Warszawa ani po wpisania zapytania „Tusk” znaleźć informacji o Donaldzie Tusku (w starszej wersji było to łatwiejsze)



ask.com

Web · Images · City · News · Video · More▾ AskEraser MyStuff▾ Options▾

Ask.com

japan

Advanced

Narrow Your Search

Map of **Japan**
Facts about **Japan**
Japanese Culture
Japan Flag
Japan History
Japan Population
Tokyo **Japan**
Japan Information
Foods in **Japan**
Japan Government

More »

Expand Your Search

China
Japanese Writing
Mt Fuji
Mount Fuji

japan

Showing **1-10** of 139,400,000

Japan www.cia.gov | Save
Capital: Tokyo; **Population:** 127,417,244
Location: Eastern Asia, island chain between the North Pacific Ocean and the Sea of Japan, east of the Korean Peninsula
Chief of State: Emperor Akihito, **Head of Government:** Prime Minister Yasuo Fukuda
Languages: Japanese More »
Encyclopedia | BBC Profile | History | Anthem | Flag | Maps
US Government Travel Info | Tourist Attractions | Current Weather | Local Time

Japan National Tourist Organization Web Site
Official Japanese travel guide providing information on transportation, hotels, attractions, festivals & events, history, culture, tradition, ...
www.jnto.go.jp/ · Cached

Japan
Official government ministry homepage with visa information, events, announcements and foreign policy news, and economic afairs.
www.mofa.go.jp/ · Cached

Japan Guide
Extensive guide to tourism and living in **Japan**. ... The **Japan** Rail Pass is just one of many rail passes for train travel in **Japan**.
www.**japan**-guide.com/ · Cached

Japanese History
History of **Japan**. Overview, and ancient to modern **Japan** ... Keeping you up to date on **Japan** travel and living related issues and site ...
www.**japan**-guide.com/e/e641.html · Cached
More Results from www.japan-guide.com

Japan Times Online
Online extension of The **Japan** Times. ... **Japan** Info Guide Links for living in **Japan** ... Romaina outlasts **Japan** on final day to advance in ...
www.**japan**times.co.jp/ · Cached

Japan Information Network
A fun approach to a guide. Lots of current information. ... Supreme Court of **Japan** Government of **Japan** Parliament of **Japan** Web - **Japan** News ...
www.jin**japan**.org/ · Cached

Japan - A Country Study
January 1994 country profile provides information about its historical setting, society and environment, economy, government and politics, and ...
lcweb2.loc.gov/frd/cs/jptoc.html · Cached

Japan Information Network

Images

More »

News Images

Firefighters donning traditional outfits perform at the top of t...
View Related
Source

Akira Sasaki of Japan passes a gate on his way to setting the se...
View Related
Source

Current Time

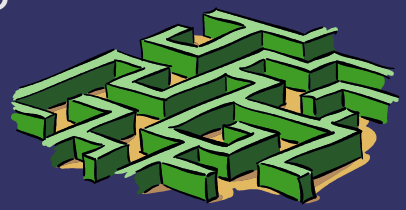
Tokyo, Japan
09:38:22 PM JST
Sunday, 6 January 2008

Encyclopedia

Japan
'(日本 Nihon or Nippon?, officially 日本国 Nihon-koku or Nippon-koku) is an island nation in East Asia. Located in the Pacific
Source

Porównanie z konkurencją – c.d.

- google.pl
  - Plusy
    - Bardzo dobre zwykłe wyniki wyszukiwania
    - Wyszukiwanie grafiki z prezentacją miniatur
    - Polska wersja interfejsu
  - Minusy
    - Niewiele informacji dostępnych już na stronie wyników (brak encyklopedii, inne dane pojawiają się rzadko)
    - Polska wersja googlowego odpowiednika Miksera jest znacznie uboższa od angielskiej
      - Brak wyszukiwania aktualności i firm
      - Przeliczanie walut domyślnie na dolary (i opis po ang.)
      - Tylko trzymanie się sztywnej składni prowadzi do wyników („gbp” zadziała, ale „kurs gbp” już nie)



google.pl

Sieć Grafika Wideo Grupy dyskusyjne Katalog Gmail więcej ▼ Zaloguj

Google japonia Szukaj Szukanie zaawansowane Ustawienia

Szukaj w Internecie Szukaj na stronach kategorii: język polski

Sieć Wyniki **1** - **10** spośród około **1,250,000** dla zapytania **japonia**. (Znaleziono w **0,07** sek.)

Wyniki wyszukiwania obrazów o **japonia**



Japonia - Wikipedia, wolna encyklopedia
Obszerny artykuł poświęcony **Japonii**. Zawiera wiele odnośników do innych artykułów z Wikipedii poszerzających wiedzę nt. Kraju Wschodzącego Słońca.
pl.wikipedia.org/wiki/**Japonia** - 361k - Kopia - Podobne strony

Japonia.xmc.pl **Japonia** Japan Tokyo :: Pogoda Mapy Hymn Tapety ...
Mieszkańcy współczesnej **Japonii** w większości są, jak przed wiekami, zwolennikami zielonej herbaty, którą zwyczajowo piją po każdym posiłku. ...
www.**japonia**.xmc.pl/ - 28k - Kopia - Podobne strony

Japonia Konnichiwa - travel Japan, Tokyo, yen, **Japonia**, pachinko ...
Japonia - Konnichiwa - Niezwykły i tajemniczy Kraj Kwitnącej Wiśni. Informacje dotyczące fascynującego kraju, jakim jest **Japonia**.
www.konnichiwa.pl/ - 29k - Kopia - Podobne strony

Japonia - Kompendium wiedzy o **Japonii** (Tokyo, Samuraje, Język ...
Japonia, **Japonii**, nihon, nippon, historia, samuraje, samuraj, nihon-go, manga, podróż, Tokio, Tokyo.
www.**japonia**.no-ip.net/ - 32k - Kopia - Podobne strony

Japonia | **JAPONIA**.ORG.PL | Historia, kultura, sztuka i ...
Informacje dotyczące kultury **japońskiej**, sztuki oraz najnowszej techniki.
www.**japonia**.org.pl/ - 26k - Kopia - Podobne strony

Japonia - WIEM, darmowa encyklopedia
Flaga i hymn Mapa Honsiu i Sikoku, **Japonia** Kiusiu, **Japonia Japonia**, Nihon, Nippon, państwo wyspiarskie we wschodniej części kontynentu...
portalwiedzy.onet.pl/902,,,,**Japonia**,haslo.html - 46k - Kopia - Podobne strony

Skarby Świata - **Japonia**
Japonia jest krajem wyspiarskim, obejmuje 3922 wyspy różnej wielkości o łącznej

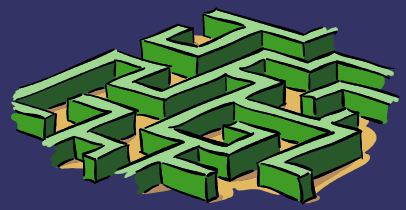
Linki sponsorowane

Japoński - kursy językowe
Od razu zaczniesz mówić ! Autorski program. Bezpłatne materiały.
www.kotonoha.pl
Mazowieckie

Japonia
Tanie przeloty z BA
Styczniowa promocja
www.ba.com

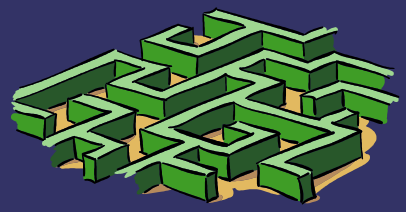
Mikser – wybrane założenia

- System przeznaczony dla osób szukających informacji w języku polskim
- Przeznaczony dla niekoniecznie doświadczonych użytkowników wyszukiwarki
  - Rozpoznawanie zapytań o „luźnej” składni
  - Mikser ułatwia odkrywanie zaawansowanych opcji wyszukiwania (np. wyszukiwarki aktualności)
- Powinien wspomagać standardowy mechanizm wyszukiwania, nie zastępować go



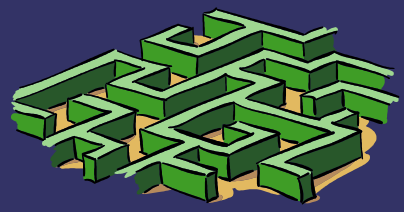
Budowa Miksera

- Biblioteka w Javie, możliwość uruchomienia zarówno wewnątrz serwisu WWW jak i w formie osobnej aplikacji (tryb tekstowy)
- Osobne aplikacje służą do importu danych statycznych oraz ich konwersji na we-wnętrzny format
- Elastyczny mechanizm wtyczek ułatwia do-dawanie nowych funkcjonalności



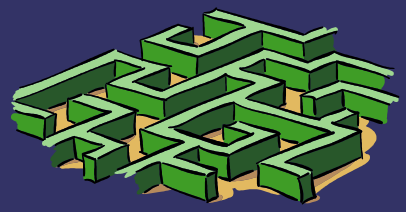
Rodzaje prezentowanych danych

- Dane statyczne
  - Wikipedia (wszystko)
  - WP (filmy, aktorzy, płyty, wykonawcy, książki)
  - FilmPolski.pl (filmy)
  - AutoCentrum.pl (samochody)
  - IDG (programy)
  - gry-online.pl (gry)
  - ...
- Dane dynamiczne
  - NetSprint
    - aktualności
    - kalkulator
    - imieniny
  - Panorama Firm
    - adresy firm
  - WP
    - prognoza pogody
    - kursy walut
    - repertuar kin
    - notowania giełdowe
    - ...



Dane dynamiczne

- Pobierane na bieżąco z zewnętrznych źródeł (możliwe keszowanie)
- Dane z różnych źródeł są zwykle obsługiwane przez oddzielne wtyczki
- Dla danych takich jak prognoza pogody czy baza firm, ważne jest określenie lokalizacji użytkownika
  - Na podstawie treści zapytania
  - Na podstawie zapisanych preferencji wyszukiwania
  - Na podstawie adresu IP



Przykłady danych dynamicznych

netsprint

WWW | Grafika | Wiadomości | Firmy | Encyklopedia | Słowniki

pogoda jutro | 10 | Szukaj | Zaawansowane Preferencje

Dodaj stronę i telefon | Promocja w wyszukiwarce | Znajdź odpowiedź. Najszybciej.

WWW: **pogoda jutro**

Warszawa mazowieckie **6 stycznia 2008** źródło: Wirtualna Polska

zmień miasto | Zapisz swoją lokalizację

	so 5.01	n 6.01	pn 7.01	wt 8.01	śr 9.01
temp. maks.:	-4°C	1°C	2°C	0°C	0°C
temp. min.:	-10°C	-3°C	-4°C	-5°C	-6°C
zachmurzenie:	zachmurzenie częściowe	pochmurno	pochmurno	zachmurzenie zmienne	zachmurzenie zmienne
opady:	-	śnieg	śnieg	śnieg	śnieg
wiatr:	płd. - wsch. 9m/s	płd. - wsch. 6m/s	płd. - zach. 5m/s	zach. 6m/s	płd. - wsch. 4m/s

Oceń ten wynik

netsprint

WWW | Grafika | Wiadomości | Firmy | Encyklopedia | Słowniki

10 dolarów | 10 | Szukaj | Zaawansowane Preferencje

Dodaj stronę i telefon | Promocja w wyszukiwarce | Znajdź odpowiedź. Najszybciej.

WWW: **10 dolarów**

10 USD = 24,53 PLN źródło: Wirtualna Polska

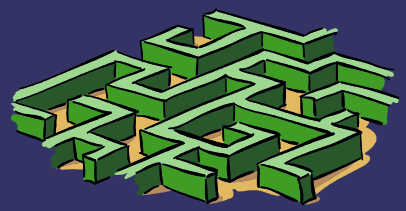
Kurs średni*: **2,4529 PLN** Zmiana: **-0,09%**

*Kurs waluty wg kursu średniego NBP z dnia 04.01.2008

Oceń ten wynik

Dane statyczne

- Dane pochodzące z różnych źródeł i udostępnione w różnych formatach są przez specjalne aplikacje konwertowane do ujednoliconego formatu
- Ujednolicone dane są indeksowane za pomocą USE (firmowej aplikacji indeksującej), ale można by też użyć zwykłej bazy danych
- Wszystkie dane statyczne są obsługiwane przez pojedynczą wtyczkę Miksera, BTS (Baza Treści Statycznych)



Przykłady danych statycznych

netsprint

WWW | Grafika | Wiadomości | Firmy | Encyklopedia | Słowniki

REJS [film] 10 Szukaj Zaawansowane Preferencje

Dodaj stronę i telefon | Promocja w wyszukiwarce | Znajdź odpowiedź. Najszybciej.

WWW: **REJS**

REJS źródło: FilmPolski.pl

Ten film otacza legenda. Istnieje liczne grono jego wielbicieli, a nawet fan club, którego członkowie znają go niemal na pamięć, scena po scenie. Skąd bierze się siła tej improwizowanej w dużej mierze opowiastki o rejsie wycieczkowym parowcem po Wiśle? Przyczyn jest chyba kilka. Po pierwsze, krytyka wszelkiego zniewolenia, na tyle zawoalowana, że mająca walor satyrycznego uogólnienia. ... więcej »

Rok: 1970
Produkcja: Polska

Czy chodziło Ci o: wybierz

Reżyseria: Marek Piwowski
Obsada: Stanisław Tym | Jolanta Lothe | Wanda Stanisławska-Lothe | Jerzy Dobrowolski | Andrzej Dobosz

Inne: strona domowa | galeria zdjęć

Szukaj: obrazków | aktualności

Oceń ten wynik

netsprint

WWW | Grafika | Wiadomości | Firmy | Encyklopedia | Słowniki

digikam 10 Szukaj Zaawansowane Preferencje

Dodaj stronę i telefon | Promocja w wyszukiwarce | Znajdź odpowiedź. Najszybciej.

WWW: **digikam**

DigiKam źródło: IDG

Narzędzie graficzne, które wspiera program dcraw, dzięki czemu jest w stanie dekodować pliki w formacie RAW. Aby rozszerzyć możliwości menedżera stworzono do niego wiele wtyczek (DigikamImagePlugins), jak i dostosowano go do współpracy z KipiPlugins. ... więcej »

Czy chodziło Ci o: Digikam [program komputerowy|nr2]

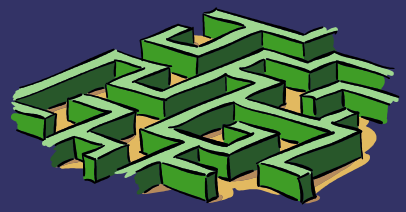
Inne: pobierz | Licencja - GPL | strona domowa | galeria zdjęć

Szukaj: obrazków | aktualności

Oceń ten wynik

Działanie BTS

- Rekordy w BTS zawierają m.in.:
  - tytuł (nazwę) rekordu
  - krótki opis i link do oryginalnego źródła
  - rysunek
  - linki do strony domowej, galerii zdjęć itp.
  - pola tabelaryczne (ustrukturyzowane dane statystyczne i inne): ludność miast, powierzchnia państw, nazwa łacińska dla roślin i zwierząt, ...
- Wyszukiwanie (w uproszczeniu)
  - po nazwie rekordu
  - po „synonimach”, np. „Mikołaj Kopernik”→„Kopernik”, „Wrocław”→„Breslau”

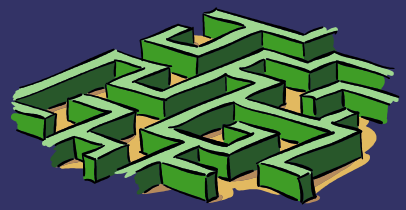


Typy rekordów

- Każdy rekord w BTS jest przypisany do jednego z kilkudziesięciu „typów”
  - Typy pozwalają odróżnić kilka rekordów o tej samej nazwie (Warszawa: miasto vs samochód)
  - „Triggery” wymuszają szukanie rekordu określonego typu oraz umożliwiają znajdowanie informacji w rodzaju „powierzchnia Gruzji”, „wilk nazwa łacińska”, „bitwa pod Grunwaldem data”

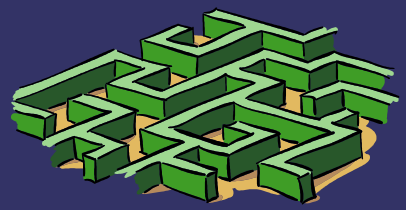
Rola Wikipedii w BTS

- Zawiera artykuły z wielu różnych dziedzin
- Dane są dostępne w zasadzie tylko jako tekst przeznaczony do prezentacji, pozbawiony struktury – trzeba z nich wydobyć:
  - Krótki opis hasła
  - Rysunek (najlepiej „właściwy” - np. flagę jeśli artykuł dotyczy państwa)
  - Typ rekordu
  - Synonimy nazwy rekordu
  - Dane tabelaryczne



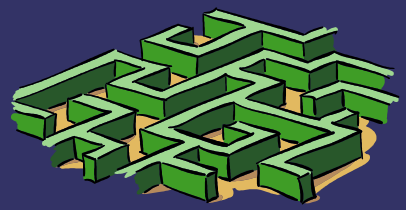
Wikitekst

```
{{Uczelnia infobox
|nazwa            = Polsko-Japońska Wyższa Szkoła Technik Komputerowych
|łacińska         =
|angielska        =
|ojczysta         =
|godło            =
|motto            =
|mapa             = Grafika:POL Warszawa map.svg
|miasto           = Warszawa
|kraj             = Polska
|dzień_założenia  =
|rok_założenia    = 1994
|tytuły_rektora   =
|rektor           = Jerzy Paweł Nowacki
|studenci         =
|adres            = ul. Koszykowa 86<br>02-008 Warszawa
|telefon          = 0-22 584-45-00
|e-mail           = pjwstk@pjwstk.edu.pl
|www              = http://www.pjwstk.edu.pl
|członkostwo      = [[Socrates-Erasmus]]
|}}
'''Polsko-Japońska Wyższa Szkoła Technik Komputerowych''' została założona w [[1994]] roku przez Fundację
   Rozwoju Technik Komputerowych powstałą na podstawie porozumienia rządów [[Polska|Polski]] i [[Japonia|
   Japonii]] z [[1993]] roku. Została wpisana do rejestru niepaństwowych szkół wyższych decyzją Ministra
   Edukacji Narodowej z dnia [[30 listopada]] [[1994]] r.

== Władze ==
* '''Rektor''' PJWSTK dr Jerzy Paweł Nowacki
* '''Prorektor ds. Ogólnych''' dr Maciej Dubejko
* '''Prorektor ds. Studenckich''' dr Aldona Drabik
* '''Kierownik Centrum Badawczego''' prof. Kazimierz Subieta
* '''Dyrektor Administracyjny''' Jan Jedliński
* '''Dziekan Wydziału Informatyki''' dr Aldona Drabik
* '''Prodziekan Wydziału Informatyki''' dr Adam Wierzbicki
* '''Dziekan Wydziału Sztuki Nowych Mediów''' prof. Marian Nowiński
* '''Prodziekan Wydziału Sztuki Nowych Mediów''' dr Włodzimierz Pastuszak
* '''Dziekan Wydziału Zarządzania Informacją''' p.o. dr Marek Kukulski

==Historia==
.....................................
```



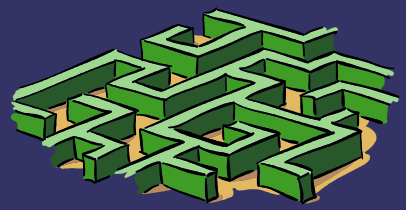
Opis i rysunek

- Wikitekst, język zapisu artykułów w Wikipedii, jest nastawiony na prezentację a nie na strukturę dokumentu
  - Nawet wydobycie pierwszego paragrafu tekstu nie jest proste jeśli ktoś użył tabel i nietypowego formatowania → heurystyki
- Rysunki: rysunek wskazany przez wybrane pole szablonu lub pierwszy rysunek w artykule (+czarna lista rysunków pomocniczych)



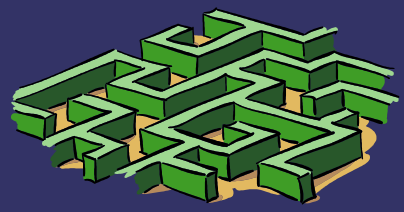
Typy rekordów z Wikipedii

- Są określane m.in. na podstawie:
  - Nazwy artykułu
  - Użytych w artykule szablonów
  - Kategorii Wikipedii, do których należy artykuł
    - Kategorii są tysiące, tworzą graf skierowany
    - Jeśli rekord Miksera o nazwie A jest typu B, oznacza to relację „A jest przedstawicielem B”, natomiast kategorie Wikipedii oznaczają w takim wypadku relację „A kojarzy się z B” - np. w kategorii „Warszawa” i podkategoriach znajdziemy zarówno dzielnice jak i osoby, organizacje, utwory literackie i wiele innych
    - Kategorie Wikipedii są w ciągłym ruchu, czasem trudno jest nadążyć za zmianami



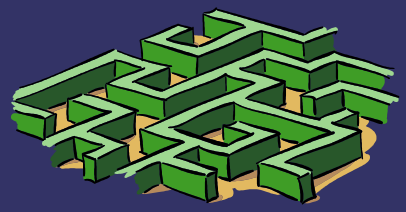
Synonimy

- Generowane na podstawie reguł, włączanych osobno dla poszczególnych typów rekordów
  - Algorytmiczne (np. „Jan III Sobieski”→(„Jan Sobieski”, „Sobieski”), „pies domowy”→„pies”)
  - Oparte o wyrażenia regularne
  - W połączeniu z pewnymi triggerami, można w zapytaniu używać haseł w dopełniaczu (z Wikipedii lub z osobnego pliku)
- Tworzone na podstawie przekierowań
- Tworzone na podstawie stron ujednoznaczniających
  - Niezbędne jest odfiltrowanie szumu

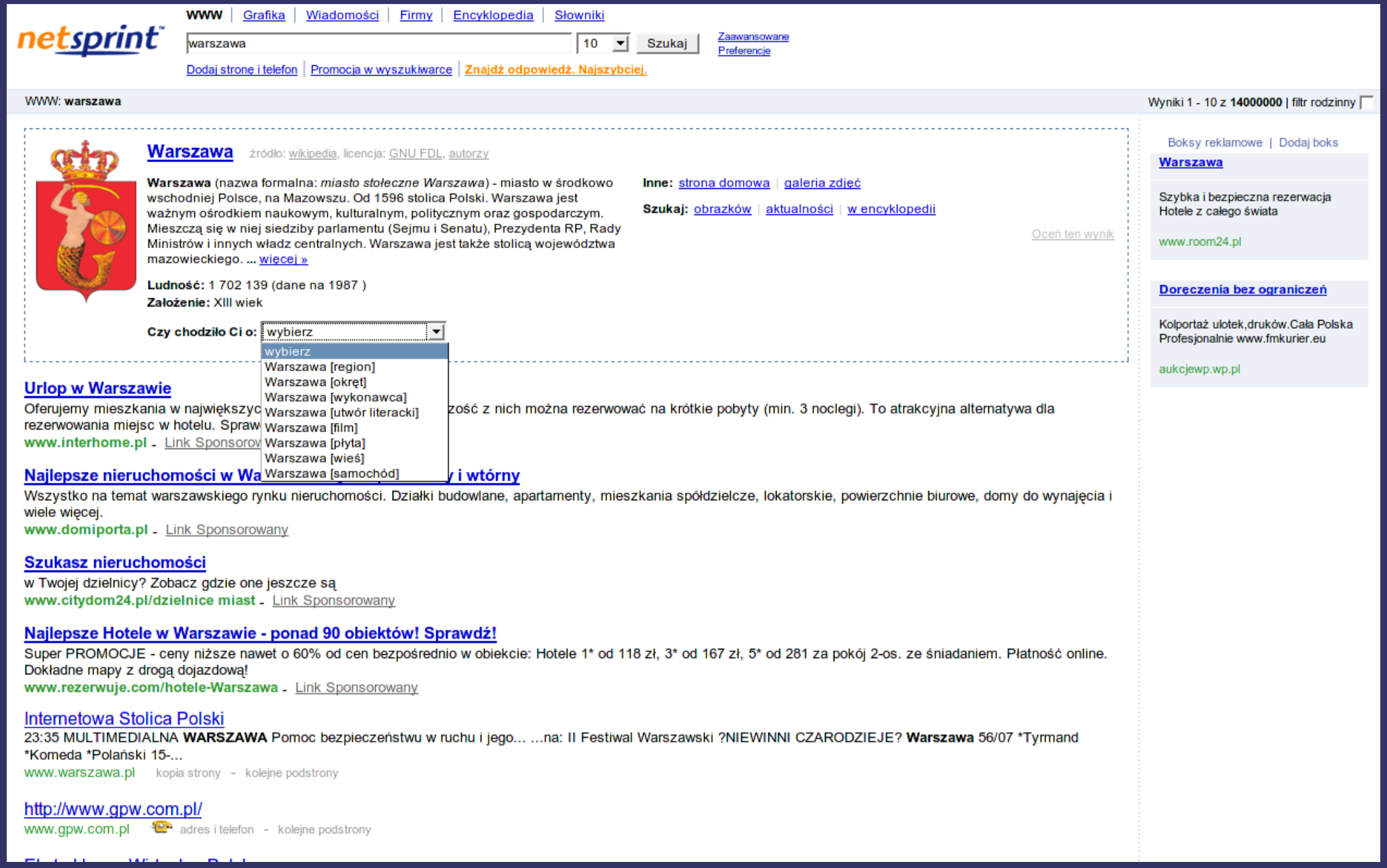


Dane tabelaryczne i linki

- Wydobywane z szablonów oraz na podstawie heurystyk z samego tekstu artykułu
- Wydobyte dane wymagają dostosowania do wspólnego formatu (np. adresy stron raz są w postaci linku, a czasem jako zwykły tekst, czasem z prefiksem http:// a czasem bez)

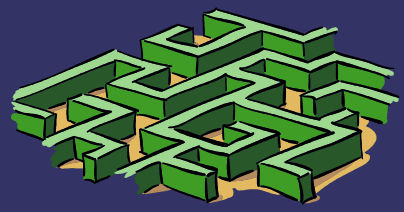


Problem:
wiele wyników dla jednego zapytania



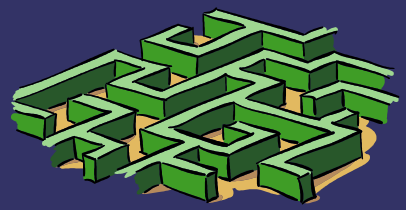
Sortowanie wyników

- Dla pojedynczego zapytania może zostać znalezionych wiele rekordów w BTS. Jak wybrać ten „najważniejszy”?
  - Nie ma jednoznacznej odpowiedzi
  - Rekordów jest bardzo wiele, żadna prosta reguła nie rozwiązuje problemu
  - Zasada najmniejszego zdziwienia
  - Kim jest „przeciętny użytkownik”?
    - Konieczność konfrontacji pomysłów teoretycznych z rzeczywistymi oczekiwaniami (analiza logów, ankiety)
      - Wczesna wersja Miksera: najpopularniejsze rekordy typu „pierwiastek” to „Tlen” oraz „Bar”
      - Wymagania zmieniają się w czasie



Sortowanie wyników – c.d.

- Obecnie
  - Częstości występowania hasła w indeksach wyszukiwarek internetu i aktualności
  - Analogiczne częstości dla zapytań z dodatkowymi członami powiązanymi z typami rekordów
  - Uwzględnianie wszystkich synonimów nazwy rekordu (założenie, że ich lista jest „kompletna”)
  - Wiele parametrów wpływających na sposób sortowania
- Pomysły
  - Klikalność
  - Ocena przez użytkowników



Dziękuję