#### Spejd Web Service

Szymon Acedański

# Specyfikacja projektu

Implement a <u>Spejd</u> web service.

# Czy już zaliczyłem?;)

```
from SimpleXMLRPCServer import SimpleXMLRPCServer
# Create server
server = SimpleXMLRPCServer(("localhost", 8000))
def whoami():
    return "Spejd, a shallow parser Web Service!"
# Register whoami() function.
server.register function(whoami)
# Run the server's main loop
server.serve forever()
```

 Parser powierzchniowy stworzony w IPI PAN

- Parser powierzchniowy stworzony w IPI PAN
- Rozpowszechniany jako Open Source

- Parser powierzchniowy stworzony w IPI PAN
- Rozpowszechniany jako Open Source
- Reguły kaskada gramatyk regularnych

- Parser powierzchniowy stworzony w IPI PAN
- Rozpowszechniany jako Open Source
- Reguły kaskada gramatyk regularnych
- Dezambiguacja

- Parser powierzchniowy stworzony w IPI PAN
- Rozpowszechniany jako Open Source
- Reguły kaskada gramatyk regularnych
- Dezambiguacja
- Parsowanie powierzchniowe (grupy składniowe)

Web Service

Strona internetowa

Web Service

Strona internetowa

Usługa, której nie widać. • HTML, CSS, te sprawy...

Web Service

- Usługa, której nie widać.
- Zestaw funkcji lub komponentów dostępnych przez sieć przy użyciu ustalonego protokołu, np. XML.

Strona internetowa

- HTML, CSS, te sprawy...
- Czasem bywa interfejsem użytkownika dla web servicu. W końcu ludzie nie mówią w XMLu.

#### Web Service

- Usługa, której nie widać.
- Zestaw funkcji lub komponentów dostępnych przez sieć przy użyciu ustalonego protokołu, np. XML.
- Jednolity sposób
   komunikacji, możliwość
   podpięcia różnych
   klientów.

#### Strona internetowa

- HTML, CSS, te sprawy...
- Czasem bywa interfejsem użytkownika dla web servicu. W końcu ludzie nie mówią w XMLu.

Bywa jednym z klientów.

• Edukacja: zajęcia dla studentów, szkolenia

- Edukacja: zajęcia dla studentów, szkolenia
- Doraźne eksperymenty, testowanie reguł tworzonych ręcznie bądź automatycznie

- Edukacja: zajęcia dla studentów, szkolenia
- Doraźne eksperymenty, testowanie reguł tworzonych ręcznie bądź automatycznie
- Wsparcie dla badań

- Edukacja: zajęcia dla studentów, szkolenia
- Doraźne eksperymenty, testowanie reguł tworzonych ręcznie bądź automatycznie
- Wsparcie dla badań
- Możliwość wykorzystania usługi na potrzeby innych serwisów
  - w końcu po to robimy Web Service, żeby nie tylko klikać przez przeglądarkę

# Dotychczasowy "rynek"

Results 1 - 10 of about 64,100 for shallow parser online. (0.20 seconds) Web Show options... Cognitive Computation Group - EOH Demonstration | T | X | Shallow parsing is the process of identifying syntactical phrases, such as noun phrases, ... [ See an online demonstration Back to the Main Page ] 12r.cs.uiuc.edu/~cogcomp/eoh/chunker.html - 4k - Cached - Similar pages -IPDFI A language-independent shallow-parser Compiler 🖟 🔀 File Format: PDF/Adobe Acrobat - View as HTML demo is available online on the author's homepage. 3.2 A Shallow-Parser for English. We wanted to evaluate our compiler on more ... www.ldc.upenn.edu/acl/P/P01/P01-1043.pdf - Similar pages by A Kinyon - Cited by 15 - Related articles - All 15 versions CiteULike: Skeletons in the parser: Using a shallow parser to ... \(\bar{\text{\te}\text{\texi\text{\texi}\text{\tex{\texit{\text{\texi}\text{\text{\text{\texi}\tiex{\tiinte\tan}\ Register and you can start organising your references online. ... We describe a simple approach for integrating shallow and deep parsing. ... www.citeulike.org/user/kvorg/article/1476088 - 38k - Cached - Similar pages -CiteULike: A language--independent shallow--parser compiler \(\bar{\pi}\) 15 Mar 2007 ... CiteULike is a free online bibliography manager. ... NJ, USA EP - 337 TI - A language--independent shallow--parser compiler PB - Association ... www.citeulike.org/user/yoavg/article/1165413 - 32k - Cached - Similar pages - P More results from www.citeulike.org » imlr-announce: JMLR special issue on shallow parsing is now available \text{ } 18 Mar 2002 ... is pleased to announce the Special Issue on Machine Learning Approaches to Shallow Parsing, available online at http://www.jmlr.org. ... www.ph.tn.tudelft.nl/PRInfo/journals/msg00253.html - 5k - Cached - Similar pages -

# Dotychczasowy "rynek"

#### **Software error:**

Can't connect to 3497: Reason: Connection refused

For help, please send mail to the webmaster (<a href="mailto:you@example.com">you@example.com</a>), giving this error message and the time and date of the error.

- Wybór reguł Spejda
  - z listy predefiniowanych zestawów
  - lub przez podanie własnego zestawu

- Wybór reguł Spejda
  - z listy predefiniowanych zestawów
  - lub przez podanie własnego zestawu
- Możliwość podania tekstu do analizy
  - w formacie IPI PAN
  - lub jako czystego tekstu

- Wybór reguł Spejda
  - z listy predefiniowanych zestawów
  - lub przez podanie własnego zestawu
- Możliwość podania tekstu do analizy
  - w formacie IPI PAN
  - lub jako czystego tekstu
- Dostęp do wyników
  - wizualizacja na stronie
  - oraz możliwość pobrania jako XML

# ... w języku API

"Simplicity is the ultimate sophistication"

Leonardo da Vinci

# ... w języku API

"Simplicity is the ultimate sophistication"

Leonardo da Vinci

String parse (Enum rulesMode, Enum textMode, String rules, String text)

rulesMode: PREDEFINED, CUSTOM

textMode: PREDEFINED, PLAIN, XML

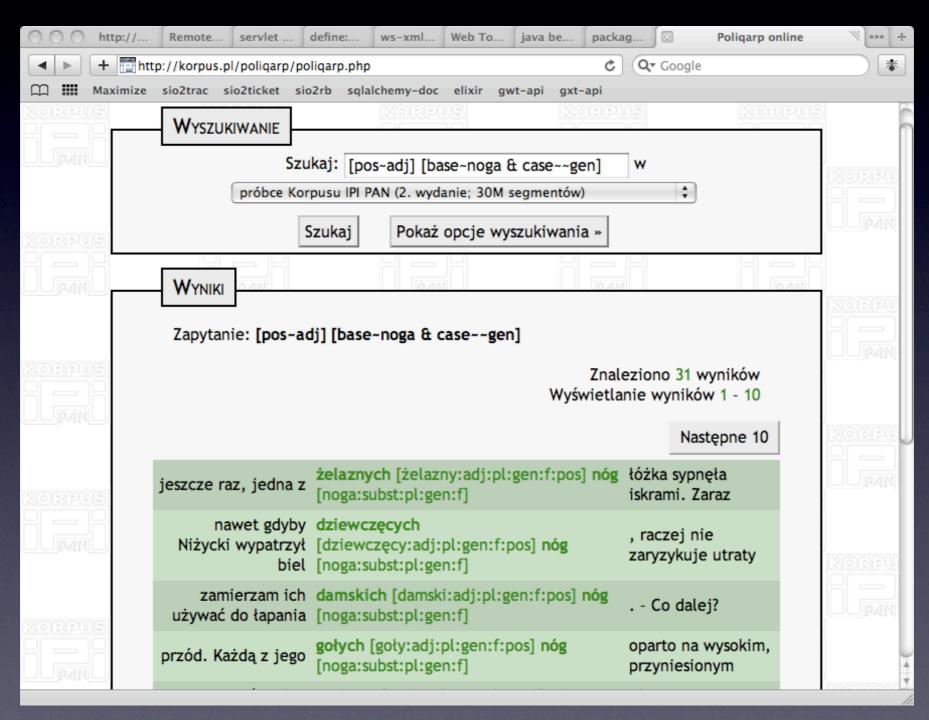
# ... w języku API

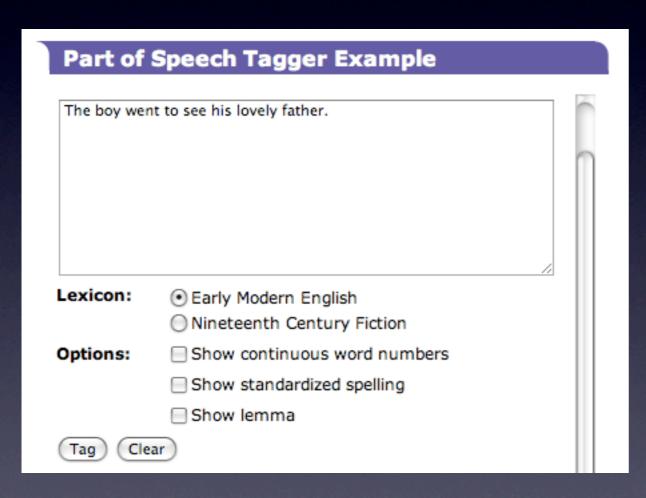
"Simplicity is the ultimate sophistication"

Leonardo da Vinci

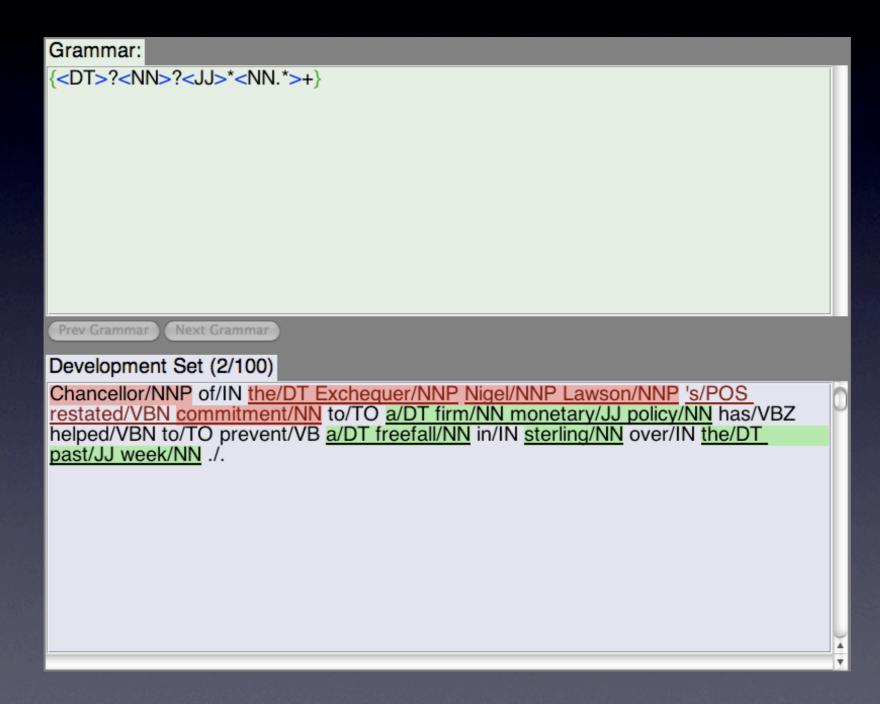
No dobrze, dobrze. Zrobię tych kilka dodatkowych funkcji, dzięki którym na przykład będzie można pobrać predefiniowane gramatyki oraz teksty. :)

W jaki sposób dobrze pokazać wynik działania parsera powierzchniowego?





S#	W#	Word	Tag	
1	1	The	dt	
1	2	boy	n1	
1	3	went	vvd	
1	4	to	pc-acp	
1	5	see	vvi	
1	6	his	po31	
1	7	lovely	j	
1	8	father	n1	
1	9			



Kotek	subst:sg:nom:m2
może	fin:sg:ter:imperf
skończyć	inf:perf
na	prep:loc
dnie	subst:sg:loc:f
rzeki	subst:sg:gen:f
!	interp

Kotek	subst:sg:nom:m2	
może	fin:sg:ter:ir.perf	
skonczyć	ınf:perf	
na	prep:loc	
inie	substloc:f	
rzeki	subst:sg:gen:f	
!	interp	

Kotek może skończyć dnie rzeki na subst:sg: subst:sg: fin:sg: subst:sg: inf:perf prep:loc interp :ter:imperf :nom:m2 :loc:f :gen:f

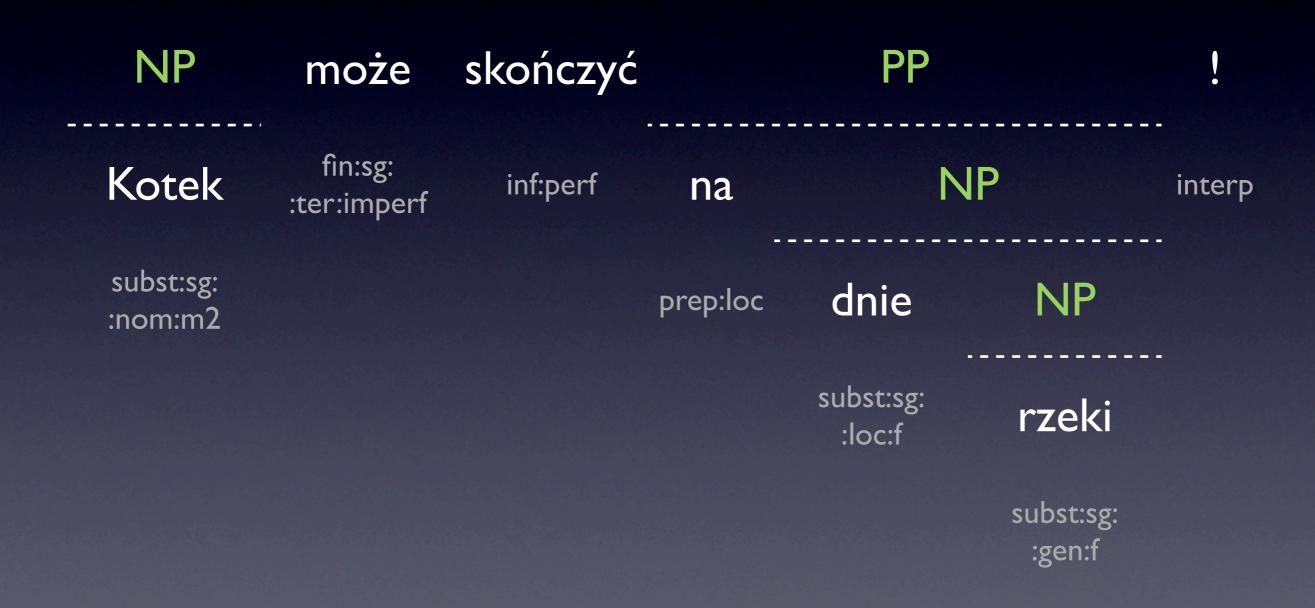
Kotek skończyć dnie rzeki może na subst:sg: subst:sg: fin:sg: subst:sg: inf:perf prep:loc interp :nom:m2 :ter:imperf :loc:f :gen:f

Kotek subst:sg:nom:m2 może fin:sg:ter:imperf skończyć inf:perf na prep:loc dnie subst:sg:loc:f rzeki subst:sg:gen:f! interp

		PP				
				NP		
NP					NP	
Kotek	może	skończyć	na	dnie	rzeki	!
subst:sg: :nom:m2	fin:sg: :ter:imperf	inf:perf	prep:loc	subst:sg: :loc:f	subst:sg: :gen:f	interp



```
[ NP Kotek subst:sg:nom:m2 ] może fin:sg:ter:imperf skończyć inf:perf [ PP na prep:loc [ NP dnie subst:sg:loc:f [ NP rzeki subst:sg:gen:f ] ] ] ! interp
```



# Pozostałe wymagania

# Pozostałe wymagania

 Do uruchomienia na Linuxie przy użyciu narzędzi Open Source

## Pozostałe wymagania

- Do uruchomienia na Linuxie przy użyciu narzędzi Open Source
- Dostępność w dwóch językach: polskim i angielskim

- Java
  - ponieważ Spejd jest napisany w Javie

- Java
  - ponieważ Spejd jest napisany w Javie
- Java RMI
  - "Java Remote Method Invocation (Java RMI) provides a simple and direct model for distributed computation with Java objects"

- Java
  - ponieważ Spejd jest napisany w Javie
- Java RMI
  - "Java Remote Method Invocation (Java RMI) provides a simple and direct model for distributed computation with Java objects"
- Google Web Toolkit
  - ponieważ jest mi znany :)

 Użycie Morfeusza jako analizatora morfosyntaktycznego dla wpisywanych tekstów.

- Użycie Morfeusza jako analizatora morfosyntaktycznego dla wpisywanych tekstów.
- Automatyczna ewaluacja gramatyki przy danym korpusie wzorcowym.

- Użycie Morfeusza jako analizatora morfosyntaktycznego dla wpisywanych tekstów.
- Automatyczna ewaluacja gramatyki przy danym korpusie wzorcowym.
- Parsowanie fragmentów korpusu IPI PAN, które można wybrać podając zapytanie w języku Morfeusza.

- Użycie Morfeusza jako analizatora morfosyntaktycznego dla wpisywanych tekstów.
- Automatyczna ewaluacja gramatyki przy danym korpusie wzorcowym.
- Parsowanie fragmentów korpusu IPI PAN, które można wybrać podając zapytanie w języku Morfeusza.

# Dziękuję za uwagę!