

java-web-bot-library



<https://github.com/Dyzio18/java-web-bot-library>

Spis treści:

- 1. Cele projektu**
- 2. Aspekty techniczne**
- 3. Aplikacja Demo**
- 4. Dokumentacja**
- 5. Organizacja pracy**
- 6. Podsumowanie**

Cele projektu

Projekt - *wstęp*

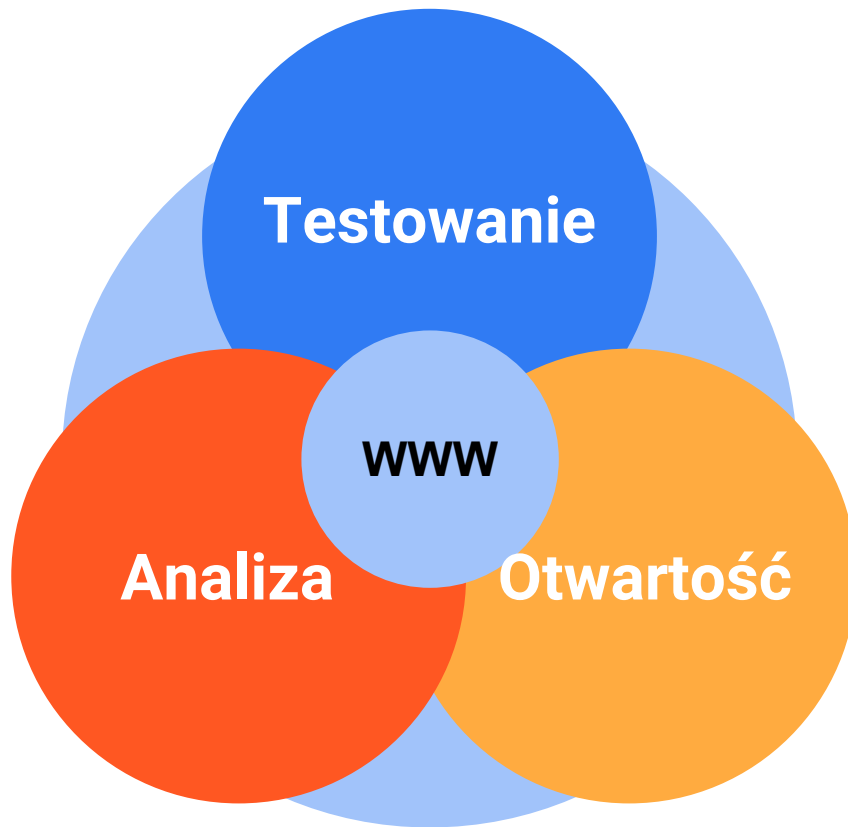
Biblioteka będzie rozwiązywać problemy z obszaru analizy i testowania rozwiązań webowych oraz komunikacji przez protokoły internetowe.

Biblioteka będzie dostarczała funkcjonalności potrzebne do wspomagania pracy developerów oraz osób zajmujących się testowaniem projektów www.

Zawierała będzie m.in. możliwość do korzystania z różnych protokołów komunikacyjnych, parsowanie stron WWW i tworzenie mapy serwisu internetowego, proste metody łamania haseł, otrzymywanie wyniku z analizy strony internetowej.

Biblioteka docelowo ma wspomagać pracę bota do testowania i analizy.
Będzie otwarta na rozwój i rozbudowę funkcjonalności.

Projekt - założenia



Problematyka i realna wartość

Wiele stron internetowych i dużych serwisów zmaga się z problemem skali.

Wraz z rozwojem aplikacji część hiperłączy lub źródeł staje się nieaktywna - nasza biblioteka pozwala na walidację “stron widmo” oraz źródeł które już nie istnieją.

Biblioteka może być również pomocna w **SEO** jako narzędzie do analizy powiązań na stronie i tworzenia **mapy strony**. Dzięki temu osoba zajmująca się pozycjonowaniem mogłaby określić hierarchię i określić poziomy zagnieżdżeń na stronie.

Osoba która chciała by wykryć pliki do których nie ma odnośnika na stronie mogłaby wykorzystać bibliotekę do wygenerowania listy plików w domenii i porównać ją z listą plików na serwerze (za pomocą np. prostego skryptu)

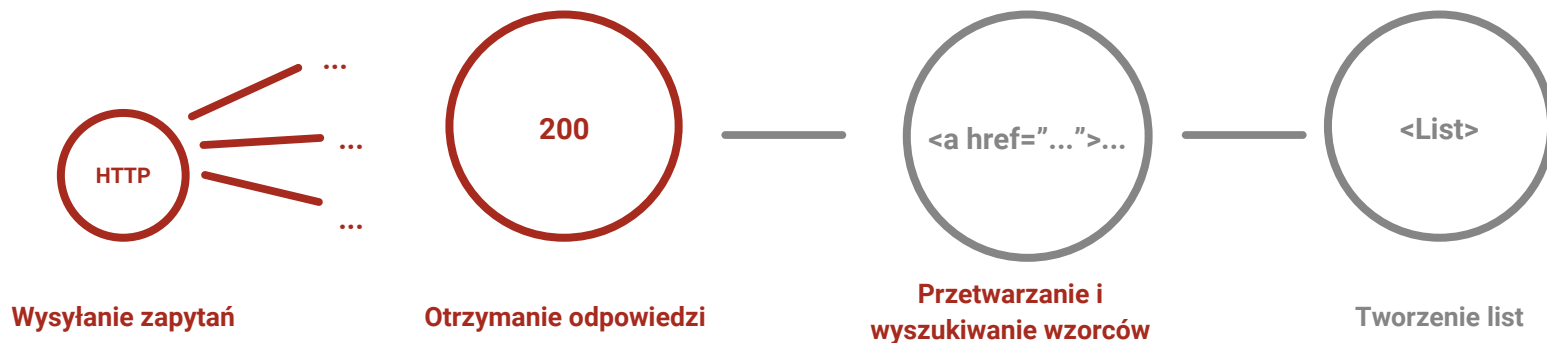


Aspekty techniczne

Zapytania HTTP



Bardzo dużo zapytań...

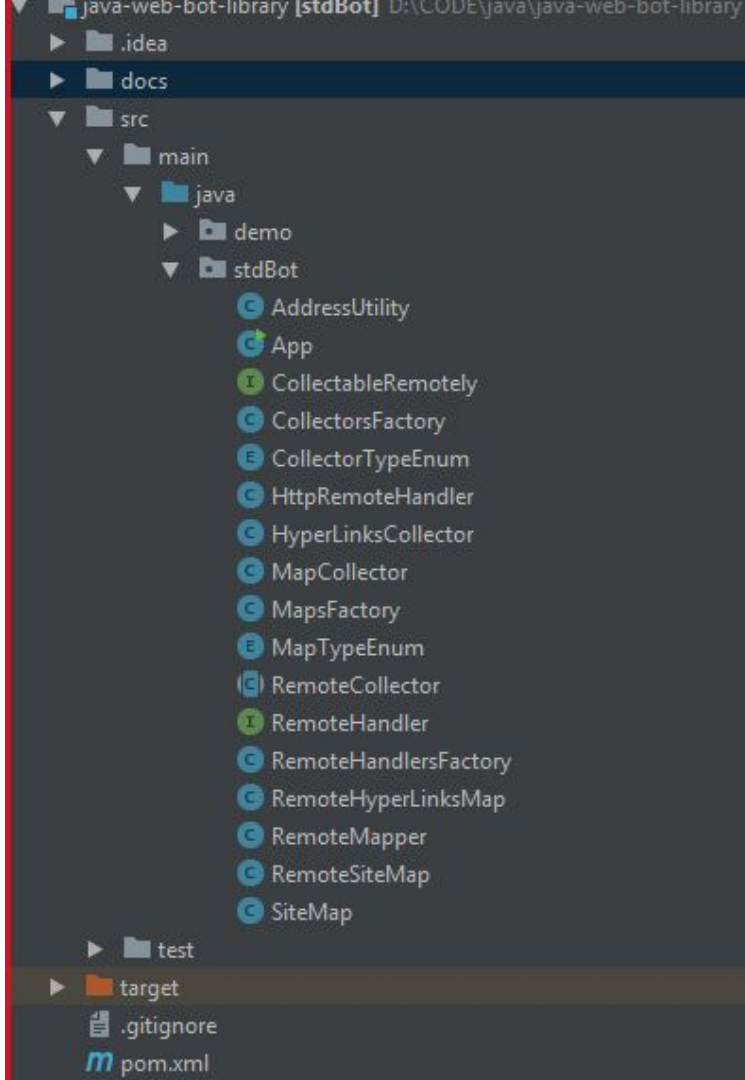


Kod aplikacji

Naszą bibliotekę pisaliśmy przy użyciu między innymi **wzorca projektowego fabryki statycznej**.

Na podstawie zadanego parametru (typu) tworzony jest dany obiekt, pozwala to na łatwe dodawanie kolejnych typów nie modyfikując kodu reszty aplikacji.

Całość jest otwarta na dodawanie funkcjonalności i rozwój biblioteki.



Wysyłanie zapytań

Żeby przyspieszyć działanie aplikacji używamy **wielowątkowości**. Poszczególne kolektory do zbierania danych dziedziczą po abstrakcyjnej klasie kolektora która obsługuje dany wątek.

```
/**
 * This class call fabric for create maps
 */
public class MapCollector extends RemoteCollector {

    public MapCollector(RemoteHandler handler, String address, int maxRecursionLevel) {
        super(handler, address, maxRecursionLevel);
        collectableRemotely = factory.create(MapTypeEnum.SiteMap);
    }
}
```

```
public class MapsFactory {

    public CollectableRemotely create(MapTypeEnum mapTypeEnum) {
        switch (mapTypeEnum) {
            case SiteMap:
                return new RemoteSiteMap();
            case HyperLinksMap:
                return new RemoteHyperLinksMap();
        }
    }
}
```

```
public abstract class RemoteCollector extends Thread {

    static protected MapsFactory factory = new MapsFactory();
    protected CollectableRemotely collectableRemotely = null;
    private String address;
    private RemoteHandler handler;
    protected int maxRecursionLevel;

    public RemoteCollector(RemoteHandler handler, String address, int maxRecursionLevel) {
        this.handler = handler;
        this.address = address;
        this.maxRecursionLevel = maxRecursionLevel;
    }

    @Override
    public void run() {
        collectableRemotely.setEntryPoint(address);
        collectableRemotely.setMaxRecursionLevel(maxRecursionLevel);
    }
}
```

```
        (!collectableRemotely.hasFinished())
        collectableRemotely.collectOne(handler);
    }
}
```

```
teMap collectAll() {
    while (true) {
        try {
            collectOne();
        } catch (InterruptedException e) {
            e.printStackTrace();
        }
    }
}
```

```
        collectableRemotely.getMap();
    }
}
```

Lista zagnieżdżeń, zliczanie łączy, graf powiązań

```
-> http://wggios.agh.edu.pl/studenci/pomoc-ma  
-> http://wggios.agh.edu.pl/studenci/erasmus/  
-> http://wggios.agh.edu.pl/kandydaci/kierunk  
-> http://wggios.agh.edu.pl/wydzial/rada-wydz  
http://wggios.agh.edu.pl/wydzial/wladze/  
-> http://wggios.agh.edu.pl/wydzial/wladze/#mm  
-> http://wggios.agh.edu.pl/wydzial/wladze/#map  
-> http://wggios.agh.edu.pl/wydzial/wladze/#sz  
-> http://wggios.agh.edu.pl/wydzial/wladze/#txt
```

Total number of internal links is: 416

```
-> http://wggios.agh.edu.pl/studenci/uprawnienia-i-kwalifikacje-zawodowe-absolwentow-wggios/  
-> http://wggios.agh.edu.pl/studenci/pomoc-materialna/  
-> http://wggios.agh.edu.pl/studenci/erasmus/  
-> http://wggios.agh.edu.pl/kandydaci/kierunki-studiow/ochrona-srodowiska/szczegolowy-opis/  
-> http://wggios.agh.edu.pl/wydzial/rada-wydzialu/sklad/  
http://wggios.agh.edu.pl/aktualnosci/info/article/egzamin-kierunkowy-i-stopnia-turystyka-i-rekreacja/  
-> -> http://wggios.agh.edu.pl/aktualnosci/info/article/egzamin-kierunkowy-i-stopnia-turysty  
-> -> http://wggios.agh.edu.pl/aktualnosci/info/article/egzamin-kierunkowy-i-stopnia-turysty  
-> -> http://wggios.agh.edu.pl/aktualnosci/info/article/egzamin-kierunkowy-i-stopnia-turysty  
-> -> http://wggios.agh.edu.pl/aktualnosci/info/article/egzamin-kierunkowy-i-stopnia-turysty  
http://www.wggios.agh.edu.pl/  
-> http://wggios.agh.edu.pl/studia/studia-podyplomowe/  
-> http://wggios.agh.edu.pl/pracownicy/  
-> http://wggios.agh.edu.pl/doktoranci/  
-> http://wggios.agh.edu.pl/studia/jakosc-kształcenia/  
-> http://wggios.agh.edu.pl/mapa-strony/  
-> http://wggios.agh.edu.pl/aktualnosci/info/article/oferta-stazu-1/  
-> http://wggios.agh.edu.pl/aktualnosci/info/article/godziny-rektorskie-w-zwiazku-z-juwenaliami-a
```

Aplikacja DEMO

Prosta **aplikacja webowa** pozwala na wyświetlenie listy powiązań na stronie podanej przez użytkownika.



localhost:2222/imback



Demo - strona www

Java-web-bot-library

Wpisz url:

<http://home.agh.edu.pl/~geoinf/pl/katedra/>

Wyślij

Mapa dla strony

<http://home.agh.edu.pl/~geoinf/pl/katedra/>

Linki:

- <http://home.agh.edu.pl/~geoinf/pl/studenci/aktualnosci/>
- <http://home.agh.edu.pl/~geoinf/pl/pracownicy/sklad-osobowy/>
- <http://home.agh.edu.pl/~geoinf/wp-json/oembed/1.0/embed?url=http%3A%2F%2Fhome.agh.edu.pl%2F%7Egeoinf%2Fpl%2Fbadania%2Fprojekty%2F>
- <http://home.agh.edu.pl/~geoinf/wp/pl/pracownicy/sklad-osobowy/dr-pawel-oleksik/>

Rezultat jako lista linków i powiązań

- <http://home.agh.edu.pl/~geoinf/pl/studenci/aktualnosci/>
- <http://home.agh.edu.pl/~geoinf/pl/pracownicy/sklad-osobowy/>
- <http://home.agh.edu.pl/~geoinf/wp-json/oembed/1.0/embed?url=http%3A%2F%2Fhome.agh.edu.pl%2F%2Fgeoinf%2Fpl%2Fbadania%2Fprojekty%2F>
- <http://home.agh.edu.pl/~geoinf/wp/pl/pracownicy/sklad-osobowy/dr-pawel-oleksik/>

Dokumentacja

JavaDoc

Kod pisaliśmy zgodnie ze standardem potrzebnym do wygenerowania dokumentacji.

Za pomocą **mavena** i wtyczki **javadocs** wygenerowaliśmy dokumentację w formie **HTML**.

W dokumentacji biblioteki znajdziemy potrzebne opisy do skorzystania z **API** oraz opisane główne komponenty projektu.

komenda:
mvn javadoc:javadoc

file:///D:/CODE/java/java-web-bot-library/target/site/apidocs/index.html

All Classes

Packages

demo
stdBot

All Classes

AddressUtility
App
CollectableRemotely
CollectorsFactory
CollectorTypeEnum
Common
HttpRemoteHandler
HyperLinksCollector
MapCollector
MapsFactory
MapTypeEnum
RemoteCollector
RemoteHandler
RemoteHandlersFactory
RemoteHyperLinksMap
RemoteMapper
RemoteSiteMap
SiteMap

Method Summary

| All Methods | Static Methods | Concrete Methods |
|-------------------|----------------------------------|----------------------------------------------------------------------------------------------|
| Modifier and Type | Method and Description | |
| static String | clearHttpAddress(String address) | This method can replace addresses startet to www. on domain address URL |
| static String | getScheme(String address) | This method return correct Uniform Resource Identifier (URI) address or null as exception |
| static boolean | isHttp(String address) | This method check and return true if address is correct HTTP URL in other cases return false |
| static boolean | isInternalRef(String address) | This method check internal references for address |

Methods inherited from class java.lang.Object

clone, equals, finalize, getClass, hashCode, notify, notifyAll, toString, wait, wait, wait

Constructor Detail

AddressUtility

```
public AddressUtility()
```

Method Detail

isHttp



```
public static boolean isHttp(String address)
```


This method check and return true if address is correct HTTP URL in other cases return false




Organizacja pracy

Repozytorium - GitHub

GitHub, Inc. [US] | <https://github.com/Dyzio18/java-web-bot-library>

 This repository Search Pull requests Issues Marketplace Explore + 

Dyzio18 / **java-web-bot-library** 

 Watch 2  Star 0  Fork 0






[Code](#) [Issues 7](#) [Pull requests 0](#) [ZenHub](#) [Projects 0](#) [Wiki](#) [Insights](#) [Settings](#)


Java library for analyze and testing websites [Edit](#)

[Add topics](#)

8 commits 1 branch 0 releases 2 contributors

Branch: master New pull request Create new file Upload files Find file [Clone or download](#)

| |
|------------------------------------------------------------------------------------------------------------------------------------------------------|
|  Dyzio18 [master] Create demo app Latest commit #015134 an hour ago |
|  src [master] Create demo app 2 hours ago |
|  .gitignore [master] Create demo app 2 hours ago |
|  README.md Initial commit 9 days ago |
|  pom.xml [master] Create demo app 2 hours ago |

 README.md

java-web-bot-library

Java library for analyze and testing websites

Podział zadań - ZenHub

The screenshot displays the ZenHub interface for the repository 'Dyzio18 / java-web-bot-library'. The top navigation bar includes the GitHub logo, a search bar, and links for 'Pull requests', 'Issues', 'Marketplace', and 'Explore'. The repository name and icons for 'Watch' (2), 'Star' (0), and 'Fork' (0) are shown. Below the navigation bar, tabs for 'Code', 'Issues' (1), 'Pull requests' (0), 'ZenHub' (active), 'Projects' (0), 'Wiki', 'Insights', and 'Settings' are visible. The left sidebar contains links for 'Boards', 'Reports', 'Create...', 'Invite your team', 'View tutorials', 'Shortcuts', and 'Open in web app'. The main area features a Kanban board with four columns: 'Backlog' (4 Issues - 8 Story Points), 'In Progress' (2 Issues - 0 Story Points), 'Review/QA' (1 Issue - 0 Story Points), and 'Closed' (3+ Issues - 5 Story Points). Each column contains task cards with titles, descriptions, and labels. The 'Backlog' column lists tasks like 'Create visualization of graph as png' and 'Write documentation'. The 'In Progress' column shows 'Create graph from lists' and 'Simple password validation'. The 'Review/QA' column has 'Create demo app'. The 'Closed' column includes 'Prepare UML diagram', 'Request for subpages', and 'Create maps of connections'. A 'New Issue' button is located in the top right of the board area.

Repository: Dyzio18 / java-web-bot-library

Navigation: Watch (2), Star (0), Fork (0), Code, Issues (1), Pull requests (0), ZenHub, Projects (0), Wiki, Insights, Settings

Left Sidebar: Boards, Reports, Create..., Invite your team, View tutorials, Shortcuts, Open in web app

Board Columns:

- Backlog (4 Issues - 8 Story Points)**
 - java-web-bot-library #4: Create visualization of graph as png (feature)
 - java-web-bot-library #6: Write documentation (feature)
 - java-web-bot-library #8: Create SSH module with API (feature)
 - java-web-bot-library #9: Simple password validation (feature)
- In Progress (2 Issues - 0 Story Points)**
 - java-web-bot-library #5: Create graph from lists (feature)
 - java-web-bot-library #10: Simple password validation (feature)
- Review/QA (1 Issue - 0 Story Points)**
 - java-web-bot-library #3: Create demo app (feature)
- Closed (3+ Issues - 5 Story Points)**
 - java-web-bot-library #7: Prepare UML diagram (5)
 - java-web-bot-library #2: Request for subpages (feature)
 - java-web-bot-library #1: Create maps of connections (feature)

Right Sidebar: Add a Pipeline ...

Buttons: Find Issues (f+i), New Issue

Dysk - Google Drive / GDocs

Mój ... > Studia - IS ... > Projekty Gr... > serwe... > PP...

Pliki Nazwa ↓

java-web-bot-library
biblioteka

[projekt][java][serw...

[Projekt] Java - PP...

PPAK-P1.pdf

java-web-bot-librar
https://github.com/Dyzio18/java-web-bot-library

[prezentacja] java-...

[Projekt] Java - PPAK [Serwer_13]

Plik Edycja Wyświetlanie Wstaw Formatuj Dane Narzędzia Dodatki Pomoc U

100% zł % .0 .00 123 Arial 10 B I

| | E | F |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|
| 1 | oty, serwisy www, zabezpieczenia) | |
| 2 | | |
| 3 | Opis | Status |
| 4 | Biblioteka umożliwia stworzenie mapy serwisu internetowego | open |
| 5 | Tworzenie listy powiązań do każdej ze stron(możliwe jest ograniczenie stopnia zagnieżdżenia) | done |
| 6 | Z każdej podstrony zliczana jest liczba odnośników (hiperłączy) wychodzących poza domene. (opcjonalnie) Wylistowane serwisy do które te odnośniki prowadzą. | in progress |
| 7 | Stworzenie aplikacji w której pokażemy działanie biblioteki. Np. generowanie mapy serwisu po stronie klienta. | open |
| 8 | | done |
| 9 | Biblioteka umożliwia podstawowe metody łamania zabezpieczeń | in progress |
| 10 | | open |
| 11 | | open |
| 12 | | open |

+ Main Wymagania [LINKI] Tutoriale i Dokumentacje

Podsumowanie

java-web-bot-library

To biblioteka mająca za zadanie testowanie i analizę rozwiązań webowych od protokołów komunikacyjnych, sprawdzania bezpieczeństwa aż po analizę stron www.

Wykonali:

Rafał Aleksander

Bartłomiej Bolek

Patryk Nizio
