

Przedmiot: **Zastosowania Technologii Informatycznych**
Prowadzący: **dr inż. Jarosław Bąk**

PROJEKT

OKE2018 CHALLENGE

Task 1: Focused Named Entity Identification and Linking



Autorzy:

Jacek Kmiecik

Michał Stancelewski

Spis treści

1. Opis projektu	3
1.1 Tematyka projektu	3
1.2 Zasoby	3
2. Użyte narzędzia i technologie	3
2.1 Java	3
2.2 Biblioteka NLP2RDF	3
2.3 Biblioteka Apache Jena	3
2.4 Biblioteka JavaFX	4
2.5 IDE IntelliJ Idea	4
3. Opis działania programu	4
3.1 Uruchomienie programu	4
3.2 Opis funkcjonalności i interfejsu	5
3.3 Instrukcja użytkowania	6
3.4 Obsługa błędów	6
4. Testy programu	7
4.1 Test nr 1	7
4.2 Test nr 2	8
4.3 Test nr 3	10

1. Opis projektu

1.1 Tematyka projektu

Tematem projektu jest program komputerowy dokonujący analizy tekstowej dowolnego łańcucha tekstowego. Następnie przeprowadzana jest identyfikacja konkretnych elementów z podanego łańcucha na podstawie danych z bazy DBpedia. Program identyfikuje zasoby z klas ontologicznych DBpedii: "Person", "Place", "Organization".

1.2 Zasoby

A. Link do projektu na GITHUB:

https://github.com/zerhaz/ZTI_2020_OKE2018_TASK1

B. Link do opisu zadania projektowego:

<https://project-hobbit.eu/challenges/oke2018-challenge-eswc-2018/tasks/#task1>

2. Użyte narzędzia i technologie

2.1 Java

Jako język programowania postanowiliśmy wybrać Java w wersji 1.8. Wybór ten podyktowany był szeroką dostępnością bibliotek oraz materiałów i informacji dostępnych w ramach tego bardzo popularnego języka. Jest też to język przystępny i wydajny w przypadku stosunkowo mało rozbudowanych projektów jak nasz.

2.2 Biblioteka NLP2RDF

Biblioteka Open Source do języka Java służąca do odczytywania i przetwarzania danych w formacie RDF i NIF (NLP Interchange Format).

2.3 Biblioteka Apache Jena

Apache Jena to bezpłatna biblioteka języka Java. Umożliwia ona przetwarzanie danych w formacie RDF. W przystępny sposób ułatwia budowanie aplikacji w oparciu o sieci semantyczne z wykorzystaniem m.in. zapytań SPARQL.

2.4 Biblioteka JavaFX

JavaFX to popularna biblioteka wykorzystywana do tworzenia aplikacji okienkowych w Javie. Nie jest tak zaawansowana jak nowoczesne rozwiązania front-endowe w postaci frameworków JavaScript (React, Angular, itd.), ale całkowicie wystarczająca, aby w kompleksowy i zarazem przejrzysty sposób stworzyć interfejs użytkownika. JavaFX upraszcza napisanie warstwy wizualnej aplikacji poprzez utworzenie struktury widoku w pliku FXML i ostylizowania go w pliku CSS.

2.5 IDE IntelliJ Idea

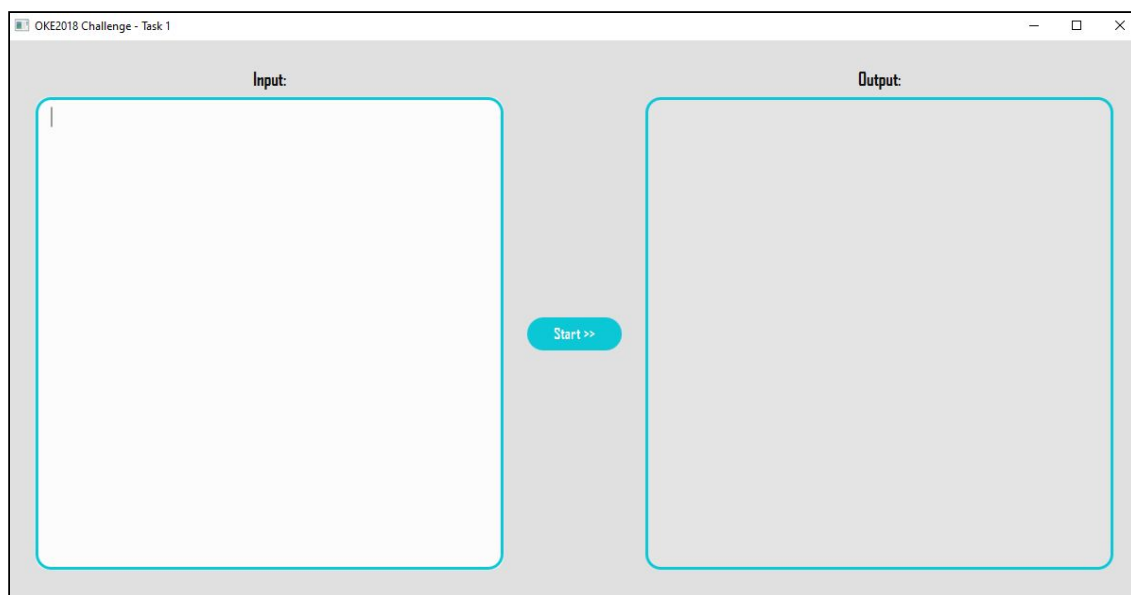
Do stworzenia projektu wybraliśmy IDE w postaci IntelliJ Idea. Na ten wybór zdecydowaliśmy się, gdyż jest to oprogramowanie, które w wersji bezpłatnej oferuje wszystkie konieczne do stworzenia projektu funkcje i obsługuje wszystkie wspomniane powyżej technologie..

3. Opis działania programu

3.1 Uruchomienie programu

Program można uruchomić z wykorzystaniem środowiska IntelliJ Idea pobierając kod źródłowy z GitHub. Do uruchomienia aplikacji należy użyć opcji "Run" na klasie MainApp.java.

Po poprawnym uruchomieniu programu powinno pojawić się okno jak na *rys. 1*.



rys. 1.

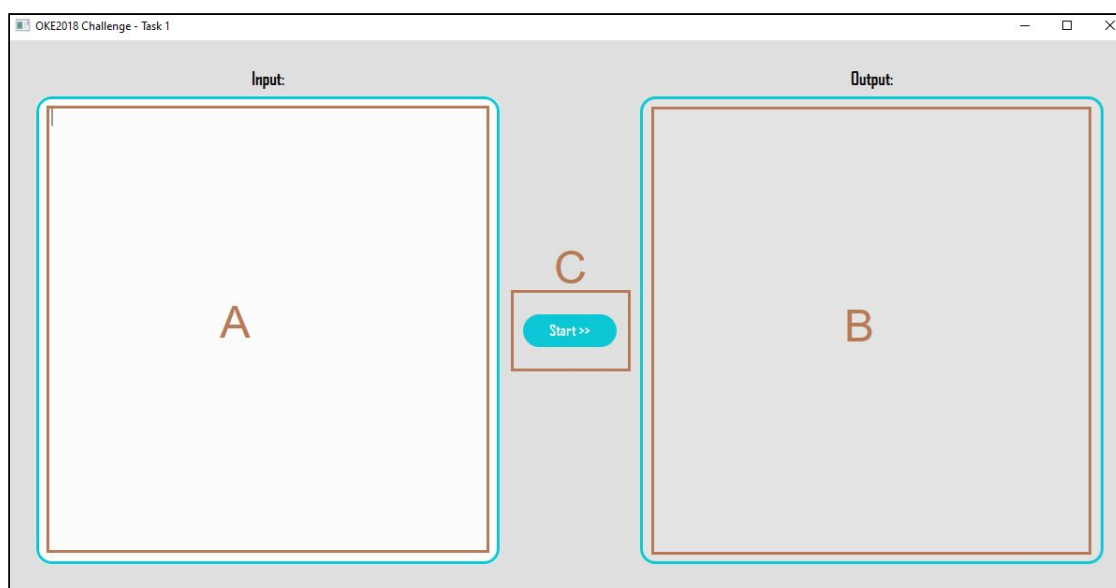
3.2 Opis funkcjonalności i interfejsu

Aplikacja okienkowa po uruchomieniu prezentuje następujące elementy widoczne na rys.2:

A - pole tekstowe na dane wejściowe

B - pole prezentujące dane wyjściowe

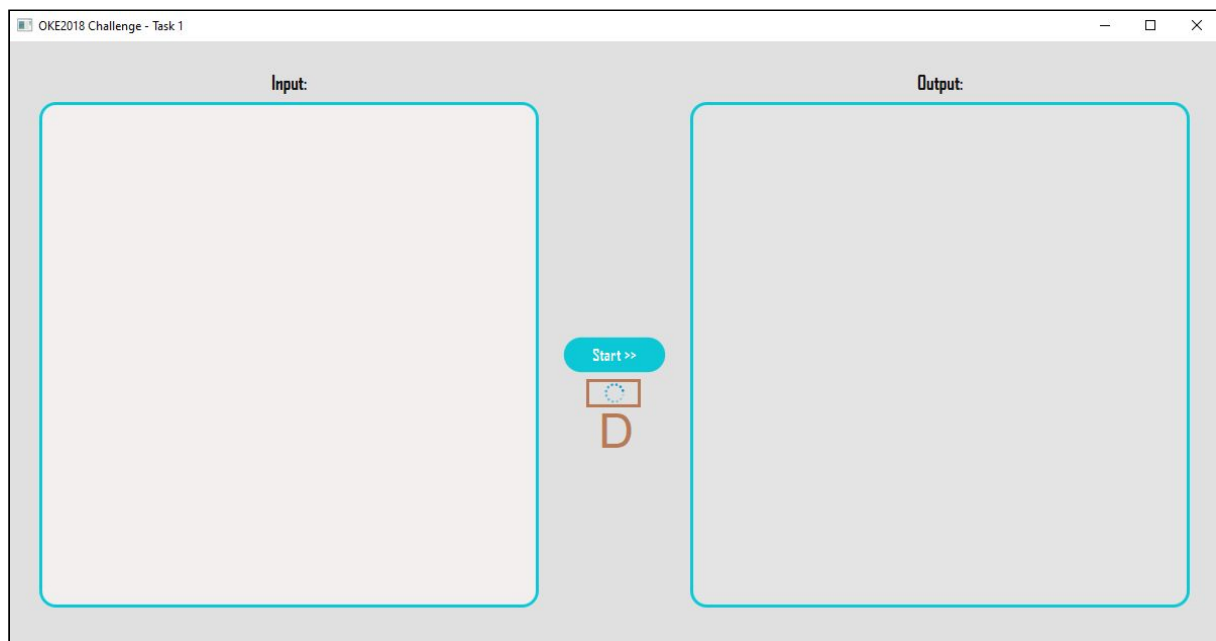
C - przycisk, który po wciśnięciu uruchamia analizę tekstu z pola A i włącza oczekiwanie na uzyskanie odpowiedzi w polu B



rys. 2.

3.3 Instrukcja użytkowania

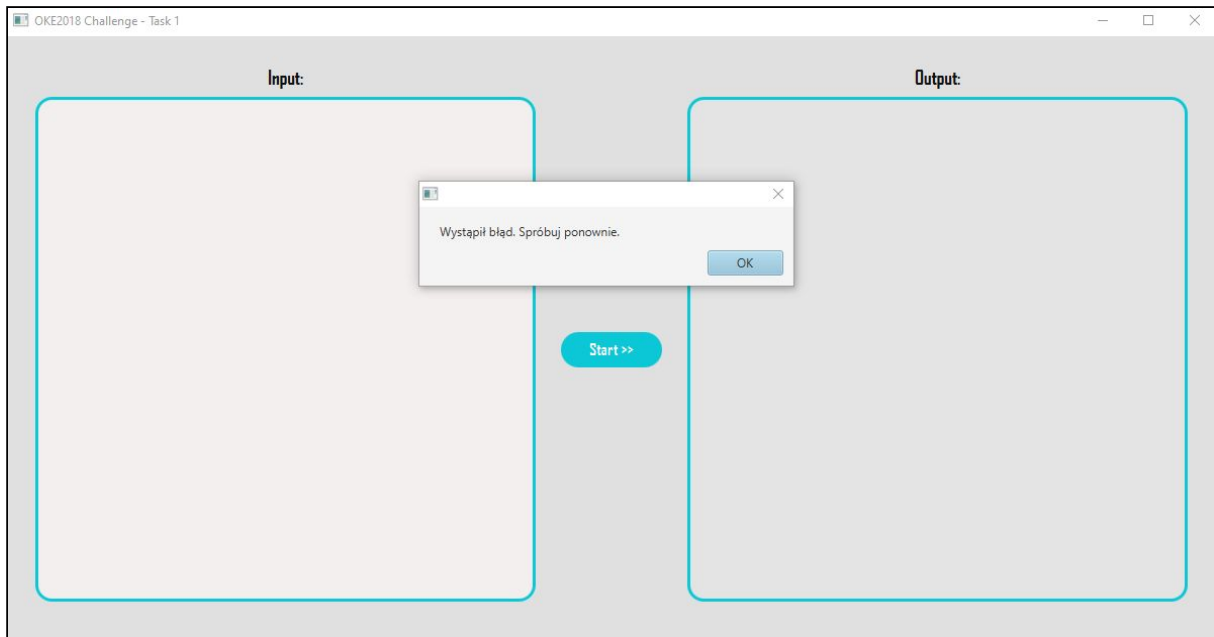
Użytkownik wprowadza w polu A z rys. 2. dane w odpowiednim formacie. Następnie aby przeprowadzić analizę wprowadzonego tekstu musi wcisnąć przycisk C. Pod przyciskiem pojawia się animowany element D (rys. 3.) świadczący, że program jest w trakcie przetwarzania zapytania. Jeśli przetwarzanie zakończy się sukcesem, to element D zniknie a w polu B pojawi się tekst odpowiedzi.



rys. 3.

3.4 Obsługa błędów

Jeśli program napotka na błąd zostanie wyświetlone okienko z powiadomieniem o wystąpieniu błędu, jak zaprezentowano na rys. 4. Błędem może być na przykład podanie w polu A z rys. 2. danych w formacie, których program nie rozpoznaje.



rys. 4.

4. Testy programu

4.1 Test nr 1

Kod wejściowy:

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#> .
<http://example.com/example-jb2#char=0,83>
  a nif:RFC5147String , nif:String , nif:Context ;
  nif:beginIndex "0"^^xsd:nonNegativeInteger ;
  nif:endIndex "83"^^xsd:nonNegativeInteger ;
  nif:isString "Donald Tusk was a polish prime minister and Barack Obama was an american
president."^^xsd:string .
```

Odpowiedź:

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#> .
<http://example.com/example-jb2#char=0,130>
```

```

a nif:RFC5147String , nif:String , nif:Context ;
nif:beginIndex "0"^^xsd:nonNegativeInteger ;
nif:endIndex "130"^^xsd:nonNegativeInteger ;
nif:isString "William Ian Lawther was a Northern Irish footballer who played for Sunderland
and five other clubs in the English Football League."^^xsd:string .
<http://example.com/example-jb2#char=106,129>

a nif:RFC5147String , nif:String , nif:Context ;
nif:anchorOf "English_Football_League"@en ;
nif:beginIndex "106"^^xsd:nonNegativeInteger ;
nif:endIndex "129"^^xsd:nonNegativeInteger ;
itsrdf:taIdentRef dbpedia:English_Football_League .
<http://example.com/example-jb2#char=8,19>

a nif:RFC5147String , nif:String , nif:Context ;
nif:anchorOf "Ian_Lawther"@en ;
nif:beginIndex "8"^^xsd:nonNegativeInteger ;
nif:endIndex "19"^^xsd:nonNegativeInteger ;
itsrdf:taIdentRef dbpedia:Ian_Lawther .

```

OKE2018 Challenge - Task 1

Input:

```

@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#> .
<http://example.com/example-jb2#char=0,130>
a nif:RFC5147String , nif:String , nif:Context ;
nif:beginIndex "0"^^xsd:nonNegativeInteger ;
nif:endIndex "130"^^xsd:nonNegativeInteger ;
nif:isString "William Ian Lawther was a Northern Irish footballer who played for Sunderland a

```

Start >>

Output:

```

@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#> .
<http://example.com/example-jb2#char=0,130>
a nif:RFC5147String , nif:String , nif:Context ;
nif:beginIndex "0"^^xsd:nonNegativeInteger ;
nif:endIndex "130"^^xsd:nonNegativeInteger ;
nif:isString "William Ian Lawther was a Northern Irish footballer who played for Sunderland a
<http://example.com/example-jb2#char=106,129>
a nif:RFC5147String , nif:String , nif:Context ;
nif:anchorOf "English_Football_League"@en ;
nif:beginIndex "106"^^xsd:nonNegativeInteger ;
nif:endIndex "129"^^xsd:nonNegativeInteger ;
itsrdf:taIdentRef dbpedia:English_Football_League .
<http://example.com/example-jb2#char=8,19>
a nif:RFC5147String , nif:String , nif:Context ;
nif:anchorOf "Ian_Lawther"@en ;
nif:beginIndex "8"^^xsd:nonNegativeInteger ;
nif:endIndex "19"^^xsd:nonNegativeInteger ;
itsrdf:taIdentRef dbpedia:Ian_Lawther .

```

rys. 5.

4.2 Test nr 2

Kod wejściowy:

```

@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .

```



```

@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#>.
<http://example.com/example-jb2#char=0,83>
a nif:RFC5147String , nif:String , nif:Context ;
nif:beginIndex "0"^^xsd:nonNegativeInteger ;
nif:endIndex "83"^^xsd:nonNegativeInteger ;
nif:isString "Donald Tusk was a polish prime minister and Barack Obama was an american
president."^^xsd:string .

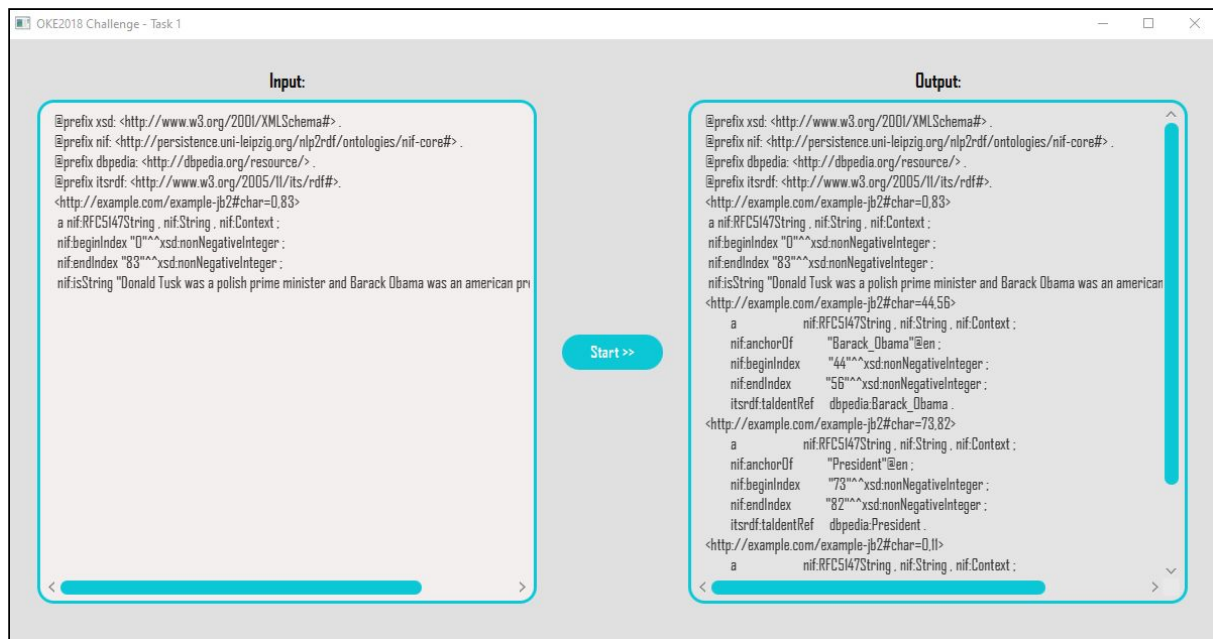
```

Odpowiedź:

```

@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#>.
<http://example.com/example-jb2#char=0,83>
a nif:RFC5147String , nif:String , nif:Context ;
nif:beginIndex "0"^^xsd:nonNegativeInteger ;
nif:endIndex "83"^^xsd:nonNegativeInteger ;
nif:isString "Donald Tusk was a polish prime minister and Barack Obama was an american
president."^^xsd:string .
<http://example.com/example-jb2#char=44,56>
a nif:RFC5147String , nif:String , nif:Context ;
nif:anchorOf "Barack_Obama"@en ;
nif:beginIndex "44"^^xsd:nonNegativeInteger ;
nif:endIndex "56"^^xsd:nonNegativeInteger ;
itsrdf:taIdentRef dbpedia:Barack_Obama .
<http://example.com/example-jb2#char=73,82>
a nif:RFC5147String , nif:String , nif:Context ;
nif:anchorOf "President"@en ;
nif:beginIndex "73"^^xsd:nonNegativeInteger ;
nif:endIndex "82"^^xsd:nonNegativeInteger ;
itsrdf:taIdentRef dbpedia:President .
<http://example.com/example-jb2#char=0,11>
a nif:RFC5147String , nif:String , nif:Context ;
nif:anchorOf "Donald_Tusk"@en ;
nif:beginIndex "0"^^xsd:nonNegativeInteger ;
nif:endIndex "11"^^xsd:nonNegativeInteger ;
itsrdf:taIdentRef dbpedia:Donald_Tusk .

```



rys. 6.

4.3 Test nr 3

Kod wejściowy:

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#> .
<http://example.com/example-task1#char=0,146>
a nif:RFC5147String, nif:String, nif:Context ;
nif:beginIndex "0"^^xsd:nonNegativeInteger ;
nif:endIndex "146"^^xsd:nonNegativeInteger ;
nif:isString "Florence May Harding studied at a school in Sydney, and with
Douglas Robert Dundas , but in effect had no formal training in either botany or art."@en
```

Odpowiedź:

```
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#> .
<http://example.com/example-task1#char=0,146>
a nif:RFC5147String, nif:String, nif:Context ;
nif:beginIndex "0"^^xsd:nonNegativeInteger ;
nif:endIndex "146"^^xsd:nonNegativeInteger ;
```

```

    nif:isString      "Florence May Harding studied at a school in Sydney, and with
Douglas Robert Dundas , but in effect had no formal training in either botany or art."@en .
<http://example.com/example-task1#char=0,20>
  a                  nif:RFC5147String , nif:String , nif:Context ;
  nif:anchorOf       "Florence_May_Harding"@en ;
  nif:beginIndex     "0"^^xsd:nonNegativeInteger ;
  nif:endIndex       "20"^^xsd:nonNegativeInteger ;
  itsrdf:taIdentRef  dbpedia:Florence_May_Harding .
<http://example.com/example-task1#char=44,50>
  a                  nif:RFC5147String , nif:String , nif:Context ;
  nif:anchorOf       "Sydney"@en ;
  nif:beginIndex     "44"^^xsd:nonNegativeInteger ;
  nif:endIndex       "50"^^xsd:nonNegativeInteger ;
  itsrdf:taIdentRef  dbpedia:Sydney .

```

OKE2018 Challenge - Task 1

Input:

```

@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#> .
<http://example.com/example-task1#char=0,146>
  a      nif:RFC5147String , nif:String , nif:Context ;
  nif:beginIndex     "0"^^xsd:nonNegativeInteger ;
  nif:endIndex       "146"^^xsd:nonNegativeInteger ;
  nif:isString       "Florence May Harding studied at a school in Sydney, and with Douglas f

```

Output:

```

@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix nif: <http://persistence.uni-leipzig.org/nlp2rdf/ontologies/nif-core#> .
@prefix dbpedia: <http://dbpedia.org/resource/> .
@prefix itsrdf: <http://www.w3.org/2005/11/its/rdf#> .
<http://example.com/example-task1#char=0,146>
  a      nif:RFC5147String , nif:String , nif:Context ;
  nif:beginIndex     "0"^^xsd:nonNegativeInteger ;
  nif:endIndex       "146"^^xsd:nonNegativeInteger ;
  nif:isString       "Florence May Harding studied at a school in Sydney, and with Douglas f
<http://example.com/example-task1#char=0,20>
  a      nif:RFC5147String , nif:String , nif:Context ;
  nif:anchorOf       "Florence_May_Harding"@en ;
  nif:beginIndex     "0"^^xsd:nonNegativeInteger ;
  nif:endIndex       "20"^^xsd:nonNegativeInteger ;
  itsrdf:taIdentRef  dbpedia:Florence_May_Harding .
<http://example.com/example-task1#char=44,50>
  a      nif:RFC5147String , nif:String , nif:Context ;
  nif:anchorOf       "Sydney"@en ;
  nif:beginIndex     "44"^^xsd:nonNegativeInteger ;
  nif:endIndex       "50"^^xsd:nonNegativeInteger ;
  itsrdf:taIdentRef  dbpedia:Sydney .

```

Start >>

rys. 7.