

Michał Marcińczuk
marcinczuk@gmail.com



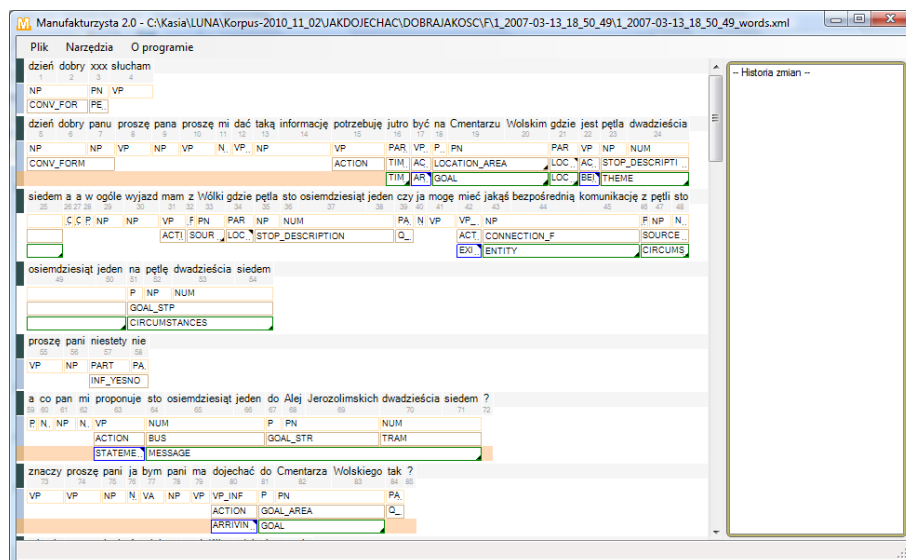
Manufakturyzsta Luna

Spis treści

1. Wstęp	2
2. Wymagania	2
3. Struktura projektu	2
4. Obsługa programu	2
4.1. Otwarcie projektu	3
4.2. Interfejs	3
4.3. Edycja słów	3
4.3.1. Wstawianie	3
4.3.2. Łączenie	4
4.3.3. Usuwanie	5
4.4. Poziomy anotacji	5
4.5. Operacje na anotacjach	5
4.5.1. Dodawanie	5
4.5.2. Edycja	7
4.5.3. Usuwanie	8
5. Schemat anotacji	8
5.1. Frazy i pojęcia	8
5.2. Ramy	8
6. Autorzy	9
A. Schemat fraz i pojęć <i>Luna.xml</i>	10
B. Schemat ram <i>frame_fe.txt</i>	11

1. Wstęp

Manufakturyzista Luna jest programem komputerowym do wizualizacji oraz edycji korpusu transkrypcji rozmów telefonicznych, powstałym w ramach projektu LUNA¹. Za pośrednictwem interfejsu graficznego możliwa jest edycja anotacji na poziomie fraz składniowych, pojęć, ram i nawiązań, a także przeglądanie informacji morfologicznych.



Rysunek 1. Interfejs aplikacji Manufakturyzista Luna.

2. Wymagania

Manufakturyzista Luna wymaga systemu **Windows** z zainstalowaną platformą **.NET Framework 3.0** lub nowszą jej wersją.

Program uruchamiany na plikach zapisanych na płycie DVD pozwala na edycję plików bez możliwości zapisu ich na płycie.

Zalecane jest skopiowanie całego katalogu zawierającego program na twardy dysk i uruchamianie go stamtąd.

3. Struktura projektu

Projekt jest zapisem pojedynczej rozmowy i składa się z sześciu plików reprezentujących różne poziomy anotacji:

- **nazwa_projektu.words.xml** — treść rozmowy z wydzielonymi słowami i analizą morfologiczno-syntaktyczną;
- **nazwa_projektu.turns.xml** — podział na wypowiedzi, czyli sekwencje słów wypowiedzianych przez poszczególnych rozmówców;
- **nazwa_projektu.chunks.xml** — podział na frazy składniowe;
- **nazwa_projektu.attvalue.xml** — opis pojęć i ich wartości (na bazie słów);
- **nazwa_projektu.frames.xml** — opis ram (na bazie pojęć);
- **nazwa_projektu.anaphoras.xml** — powiązania anaforyczne między pojęciami.

¹ Strona domowa projektu: <http://www.ist-luna.eu>.

4. Obsługa programu

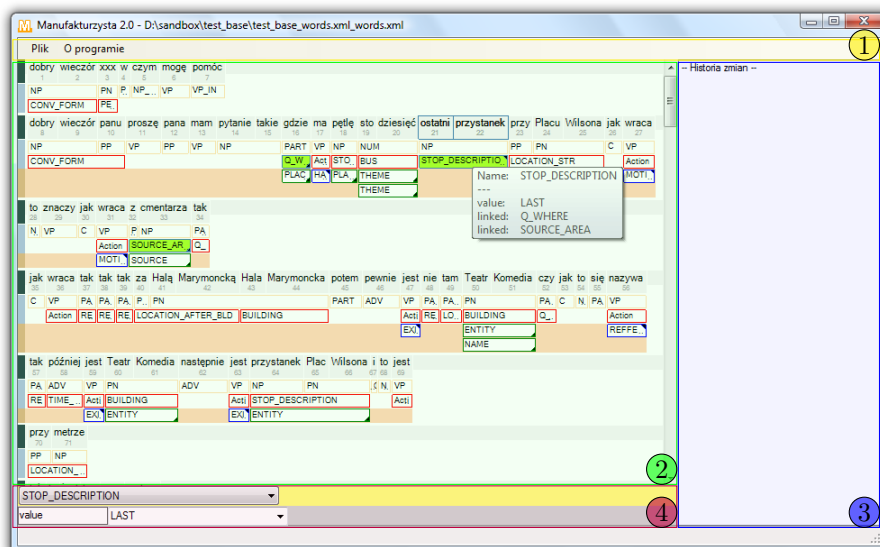
Program jest uruchamiany przez podwójne kliknięcie pliku Manufakturzysta.exe lub kliknięcie odpowiedniego klawisza okienka startowego płyty.

4.1. Otwarcie projektu

Aby otworzyć projekt, należy wybrać opcję **Projekt > Otwórz projekt Luna ...** lub skorzystać ze skrótu klawiaturowego **Ctrl + L** i wskazać plik ze słowami, tj. **nazwa_projektu.words.xml**.

4.2. Interfejs

Po otwarciu projektu w interfejsie użytkownika widać 4 główne elementy (widoczne na poniższym zdjęciu):



1. **Menu górne** zawierające opcje, które umożliwiają otwarcie i zamknięcie projektu, oraz wyświetlenie informacji o programie.
2. **Panel do anotacji** widoczny w centralnej części programu, na którym wyświetlana jest zawartość projektu i możliwa jest edycja anotacji na wszystkich dostępnych poziomach.
3. **Historia zmian** widoczna po prawej stronie *panelu do anotacji*. Historia umożliwia wycofanie zmian wprowadzonych od momentu otwarcia projektu poprzez kliknięcie na liście zmian tej akcji, do której chcemy cofnąć zmiany.
4. **Edytor anotacji** widoczny w dolnej części okna po zaznaczeniu frazy składniowej lub pojęcia. Edytor służy do zmiany atrybutów anotacji oraz ich wartości. W przypadku fraz możliwe jest również wskazanie elementu głównego (dla fraz NP i VP).

4.3. Edycja słów

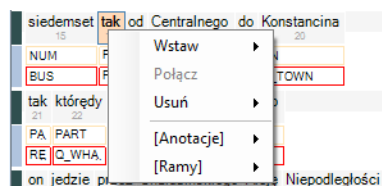
W programie możliwe są następujące operacje na słowach: wstawianie nowych słów, łączenie istniejących oraz usuwanie. Wymienione operacje opisane są szczegółowo w kolejnych podpunktach.

4.3.1. Wstawianie

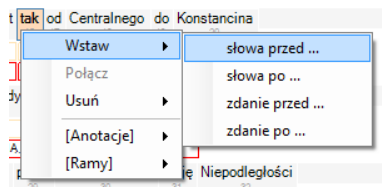
Możliwe jest wstawienie sekwencji słów przed wskazanym słowem lub po nim, oraz jako nową wypowiedź przed istniejącą lub po niej. Po dodaniu nowych słów program generuje odpowiednie znaczniki i aktualizuje indeksy słów we wszystkich plikach. Dalsza modyfikacja słów (forma bazowa, morfologia itd.) możliwa jest wyłącznie poprzez ręczną edycję pliku **nazwa_projektu.words.xml**.

Wstawianie słów (krok po kroku)

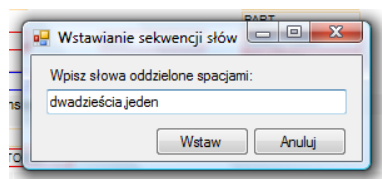
1. Klikamy prawym przyciskiem myszy anotację, przed którą lub po której ma być wstawiona sekwencja słów.



2. Z menu kontekstowego wybieramy **Wstaw** ▸ **słowa przed ...** lub **słowa po ...**.

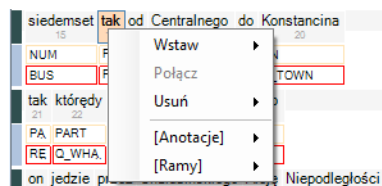


3. W oknie dialogowym wpisujemy sekwencję słów oddzielając każde słowo spacjami.

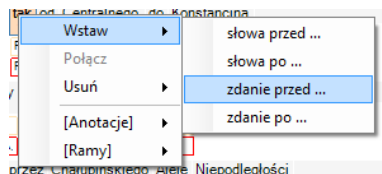


Wstawianie wypowiedzi (krok po kroku)

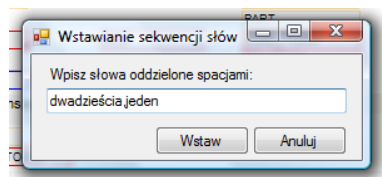
1. Klikamy prawym przyciskiem myszy dowolną anotację w zdaniu, przed którym lub po którym ma być wstawiona nowa wypowiedź.



2. Z menu kontekstowego wybieramy **Wstaw** ▸ **zdanie przed ...** lub **zdanie po ...**.



3. W oknie dialogowym wpisujemy sekwencję słów oddzielając każde słowo spacjami.

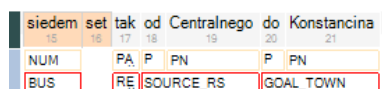


4.3.2. Łączenie

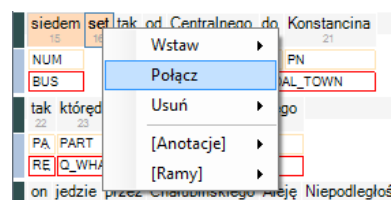
Opcja łączenia słów służy do zastąpienia sekwencji słów jednym elementem. Nowe słowo przyjmuje wszystkie atrybuty drugiego słowa z łączonej sekwencji. Wartość atrybutu *orth* jest równa konkatenacji atrybutów *orth* łączonych słów.

Łączenie słów (krok po kroku)

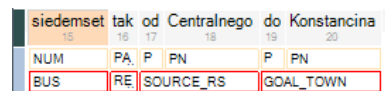
1. Zaznaczamy sekwencję słów.



2. Klikamy prawym przyciskiem myszy dowolne z zaznaczonych słów i z menu kontekstowego wybieramy opcję **Połącz**.



3. Nowe słowo zostanie podświetlone.



4.3.3. Usuwanie

Aby usunąć słowo, klikamy je prawym przyciskiem myszy i wybieramy z menu kontekstowego **Usuń** ▷ **słowo/a**.

4.4. Poziomy anotacji

Każda wypowiedź wyświetlana jest w nowej linii i składa się z trzech warstw.

Pierwsza warstwa, wyróżniona ciemnomorskim początkiem wierszy, zawiera sekwencję słów. Poniżej słów wyświetlone są ich indeksy liczone od początku dokumentu, zaczynając od 1. Analiza morfologiczno-syntaktyczna słów dostępna jest w dymku wyświetlanym po najechaniu kursorem na słowo.

Druga warstwa, wyróżniona jasnoniebieskim początkiem wierszy, zawiera frazy składniowe, pojęcia i nawiązania anaforyczne.

Trzecia warstwa o ciemnopiaskowym początku wierszy zawiera opis ramowy.

1 warstwa	to	znaczy	jak	wraca	z	cmientarza	tak	– słowa i morfologia;
2 warstwa	N	VP	C	VP	P	NP	PA	– frazy, pojęcia i nawiązania;
3 warstwa			Action	SOURCE_AR	Q			– ramy;

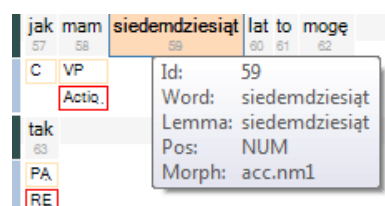
4.5. Operacje na anotacjach

Program umożliwia dodawanie, edycję i usuwanie *fraz składniowych*, *pojęć*, *ram* i *nawiązań anaforycznych*.

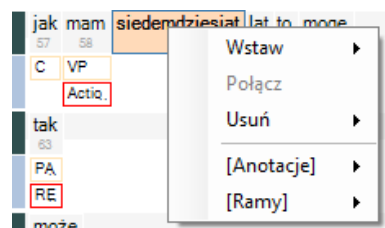
4.5.1. Dodawanie

Dodawanie fraz syntaktycznych (krok po kroku)

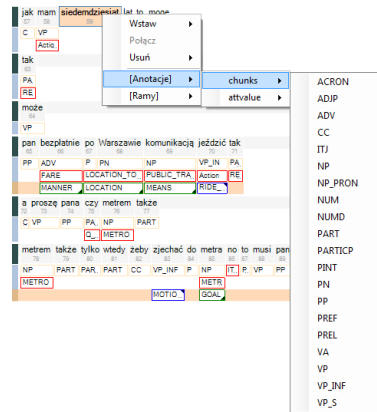
1. Zaznaczamy sekwencję słów klikając jeden koniec sekwencji i trzymając lewy przycisk myszy przesuwamy kursor na drugi jej koniec.



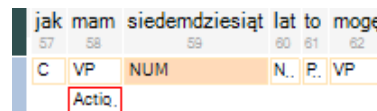
2. Klikamy prawym przyciskiem myszy dowolne słowo z zaznaczonej sekwencji.



3. Z menu kontekstowego wybieramy [Anotacje] ▷ chunks. Wyświetli się lista dostępnych anotacji.

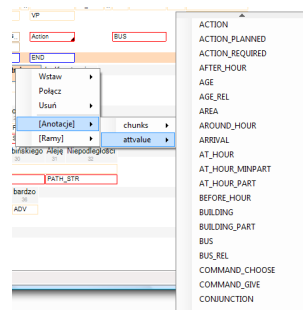
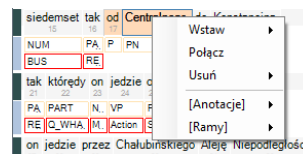
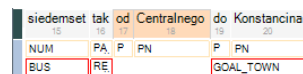


4. Wybieramy z listy jeden z typów fraz składniowych. W przykładzie po prawej dodano anotację NUM.

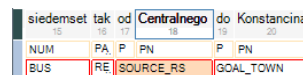


Dodawanie *pojęć* (krok po kroku)

1. Zaznaczamy sekwencje słów klikając jeden koniec sekwencji i trzymając lewy przycisk myszy przesuwamy kursor na drugi jej koniec.
2. Klikamy prawym przyciskiem myszy dowolne słowo z zaznaczonej sekwencji.
3. Z menu kontekstowego wybieramy [Anotacje] ▷ attvalue. Wyświetli się lista dostępnych pojęć.

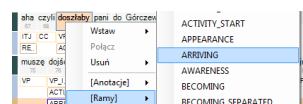
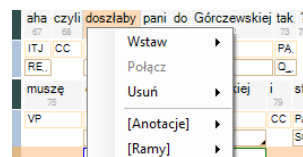


4. Wybieramy z listy jedno z pojęć. W przykładzie dodano pojęcie SOURCE_RS.

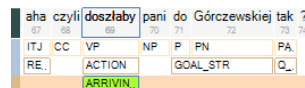


Dodawanie *ram* (krok po kroku)

1. Klikamy prawym przyciskiem myszy wybrane słowo.
2. Z menu kontekstowego wybieramy opcję [Ramy]. Wyświetli się pełna lista ram.

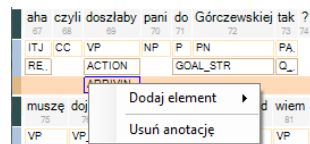


- Wybieramy z listy jedną z ram.
W przykładzie dodano ramę
ARRIVING.

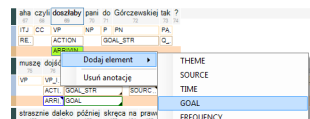


Dodawanie *elementów ram* (krok po kroku)

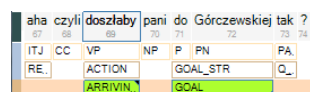
- Klikamy prawym przyciskiem myszy wybraną ramę.



- Z menu kontekstowego wybieramy **Dodaj element**. Wyświetli się lista dostępnych elementów ramy.



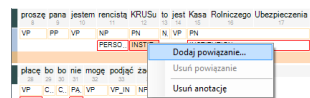
- Wybieramy z listy jeden z elementów ramy i wybieramy to pojęcie, które ma stać się elementem ramy.



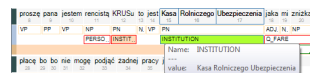
Nazwa ramy zaznaczona jest granatową otoczką i trójkątem w prawym górnym rogu, elementy ramy zaś zieloną otoczką i trójkątem w prawym dolnym rogu. Po najechaniu na predykat (element tworzący ramę) wszystkie elementy ramy zostaną podświetlone na zielono.

Dodawanie *nawiązań anaforycznych*

- Klikamy prawym przyciskiem wybrane pojęcie.



- Z menu kontekstowego wybieramy **Dodaj powiązanie** i klikamy drugie pojęcie, aby utworzyć dowiązanie.



Relacje anaforyczne między pojęciami zaznaczone są trójkątami w prawym dolnym (podstawa nawiązania) lub górnym rogu (człon nawiązany). Po najechaniu na człon nawiązany podświetla się na zielono człon nawiązany i odpowiednia podstawa nawiązania. W dymku widocznym przy członie nawiązanym podana jest nazwa pojęcia stanowiącego podstawę nawiązania.

4.5.2. Edycja

Wszystkie typy anotacji mogą być przesuwane i modyfikowane przy pomocy klawiszy i myszki. Po zaznaczeniu anotacji (kliknięciu anotacji lewym przyciskiem myszy) możliwe jest:

- przesuwanie anotacji w lewo i prawo w obrębie wypowiedzi przy pomocy strzałek,
- przenoszenie anotacji w obrębie wypowiedzi i między wypowiedziami przy pomocy myszy. W tym celu należy kliknąć anotację lewym przyciskiem myszy, wcisnąć klawisz *Shift* i przytrzymując oba przyciski przesunąć kursor.
- jeżeli po przesunięciu anotacja najdzie na inną anotację, to możliwe jest jej przesunięcie w górę lub w dół używając strzałek. Po przesunięciu anotacja przejdzie do istniejącego wiersza lub zostanie stworzony nowy wiersz. Pionowe ułożenie anotacji nie jest zapisywane do pliku. Po zapisie i ponownym wczytaniu pliku układ w pionie może ulec zmianie.

Granice anotacji mogą być zmieniane przy pomocy strzałek z wciśniętym klawiszem *Control* (lewa granica) lub *Alt* (prawa granica).

Lista skrótów klawiszowych:

←	—	przesunięcie <i>anotacji</i> w lewo
→	—	przesunięcie <i>anotacji</i> w prawo
↑	—	przesunięcie <i>anotacji</i> o wiersz w górę
↓	—	przesunięcie <i>anotacji</i> o wiersz w dół
<i>Control</i> i ←	—	przesunięcie <i>początku anotacji</i> w lewo (poszerzenie z lewej)
<i>Control</i> i →	—	przesunięcie <i>początku anotacji</i> w prawo (zwężenie z lewej)
<i>Alt</i> i ←	—	przesunięcie <i>końca anotacji</i> w lewo (zwężenie z prawej)
<i>Alt</i> i →	—	przesunięcie <i>końca anotacji</i> w prawo (poszerzenie z prawej)

4.5.3. Usuwanie

Aby usunąć anotację (*frazę składniową, pojęcie, ramę, element ramy*), zaznaczamy anotację i wciskamy klawisz *Delete*. Po usunięciu ramy usuwane są również wszystkie jej elementy. Można również kliknąć anotację prawym klawiszem myszy i wybrać z menu kontekstowego opcję **Usuń anotację**.

Usunięcie *nawiązania* odbywa się poprzez kliknięcie pojęcia prawym przyciskiem myszy, wyborze opcji **Usuń powiązanie** i kliknięciu wybranego dowiązania z rozwiniętej listy.

5. Schemat anotacji

5.1. Frazy i pojęcia

Schemat anotacji zapisany jest w pliku *Luna.xml* i obejmuje:

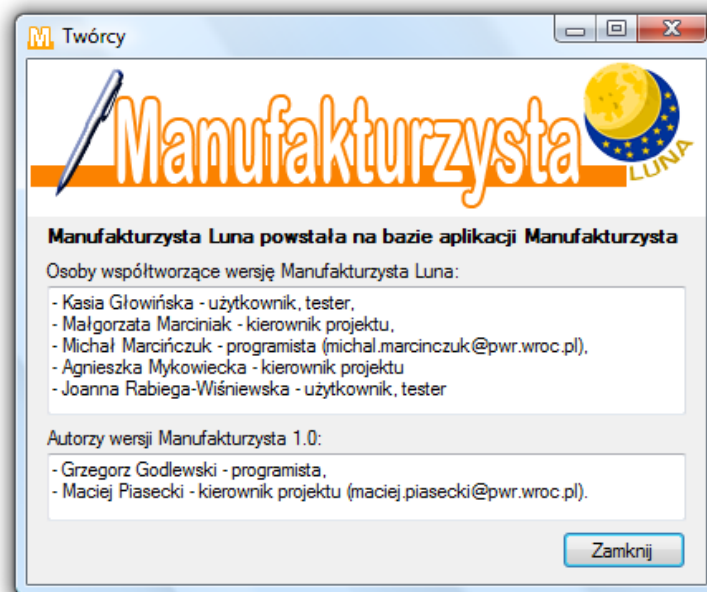
- listę typów anotacji z podziałem na grupy (frazy składniowe, pojęcia);
- definicję atrybutów anotacji na poziomie grupy i poszczególnych anotacji;
- definicję kolejności wyświetlania anotacji w edytorze (atrybut *priority*);
- kolor anotacji (atrybuty *color* i *bcolor*).

Elementy anotacji, które nie są opisane w schemacie, otoczone są czerwoną ramką. Fragment schematu anotacji *Luna.xml* znajduje się w załączniku A.

5.2. Ramy

Schemat ram znajduje się w pliku tekstowym *frame_fe.txt*. Każda rama opisana jest przez dwie linie — pierwsza zawiera nazwę ramy, druga listę elementów oddzielonych przecinkami. Fragment schematu ram znajduje się w załączniku B.

6. Autorzy



A. Schemat fraz i pojęć *Luna.xml*

Fragment schematu fraz i pojęć.

```
<?xml version='1.0' encoding='UTF-8'?>
<annotations version="2.0">
  <group name="chunks" level="1">
    <annotations>
      <annotation priority="95" color="Violet"
        bgcolor="Moccasin" name="ACRON"/>
      <annotation priority="95" color="Violet"
        bgcolor="Moccasin" name="ADJP"/>
      <annotation priority="95" color="Violet"
        bgcolor="Moccasin" name="ADV"/>
      <annotation priority="95" color="Violet"
        bgcolor="Moccasin" name="CC"/>
      <annotation priority="95" color="Violet"
        bgcolor="Moccasin" name="ITJ"/>
      <annotation priority="95" color="Violet"
        bgcolor="Moccasin" name="NP"/>
      <annotation priority="95" color="Violet"
        bgcolor="Moccasin" name="NP_PRON"/>
      <annotation priority="95" color="Violet"
        bgcolor="Moccasin" name="NUM"/>
      <annotation priority="95" color="Violet"
        bgcolor="Moccasin" name="NUMD"/>
      ...
    </annotations>
  </group>
  <group name="attvalue" level="2">
    <annotations>
      <annotation priority="90" color="Tan" name="AFTER_HOUR">
        <attribute name="value" type="string">
          <value>10 : 00</value>
          <value>12 : 30</value>
          <value>6 : 00 am</value>
        </attribute>
      </annotation>
      <annotation priority="90" color="Tan" name="AGE">
        <attribute name="value" type="string">
          <value>General</value>
          <value>7</value>
          <value>Above7</value>
          <value>Below7</value>
          <value>Below21</value>
          <value>26</value>
          <value>Above26</value>
          <value>Below26</value>
          <value>Above65</value>
          <value>70</value>
          <value>Above70</value>
          <value>Below70</value>
          <value>SchoolAge</value>
          <value>Irrelevant</value>
        </attribute>
      </annotation>
      ...
    </group>
  </annotations>
```

B. Schemat ram *frame_fe.txt*

Fragment schematu ram.

```
Frame = GIVE
FrameElements = DONOR, RECIPIENT, THEME, TIME, CIRCUMSTANCES

Frame = GIVING
FrameElements = THEME

Frame = HALT
FrameElements = THEME, PLACE, POSITION, TIME, PATH

Frame = INCLUDE
FrameElements = PART, TOTAL

Frame = INTEREST
FrameElements = EXPERIENCER, STIMULUS, TIME, MANNER

Frame = LASTING
FrameElements = EVENT, TIME_DISTANCE

Frame= LOCATING
FrameElements= GROUND, LOCATION, PERCEIVER, SOUGHT_ENTITY,
               CIRCUMSTANCES

Frame = MOTION
FrameElements = CARRIER, CONTAINIG_EVENT, DIRECTION, DISTANCE,
               DURATION, FREQUENCY, GOAL, MANNER, MEANS, PATH,
               PLACE, SOURCE, THEME, TIME
```