

Serviciul pentru procesare de text

Descrierea problemei

Serviciul pentru procesare de text va primi un text si va returna entitatile care se gasesc in text. Serviciul este structurat astfel incat sa permita adaugarea unei noi entitati si returnarea ei, entitatea care va fi implementata initial, va fi concept.

Principalii actori

TextAnalyzer: Isi va seta ce entitati is doreste sa indentifice si metoda de separare a textului, dupa care va prelua textul primit de creier si va returna o lista de entitati (concepte).

TextSeparator: Va primi un text si va separa textul in functie de metoda de implementare (Semantic parsing, Bag of words) intr-o lista de String-uri (cu o structura arborescena).

EntitiesExtractor: Va primi o lista de String-uri, o va filtra astfel incat sa preia doar elementele ce ar putea apartine entitatii respective (concept) si va cauta in memorie (ShortMemory, LongMemory) elementele respective.

EntityRegister: Va primi o lista de String-uri, ce au proprietatea ca ar putea apartine unei anumite entitati si le va adauga in memorie (ShortMemory, LongMemory).

ShortMemory: Va contine o lista de enitati si va returna o entitate ce are numele echivalent cu String-ul primit in cazul in care apare in lista.

LongMemory: Va face legatura cu baza de date si va cauta in baza de date o instanta a unuei entitati cu numele echivalet cu String-ul primit in cazul in care apare in baza de date.

Principalele scenarii de utilizare pentru fiecare actor

TextAnalyzer:

Set entities

- 1) Set entities

Set textSeparator

- 1) Set textSeparator

Analyze

- 1) Set textSeparator
- 2) Set entities
- 3) Analyze

Exceptii: fie modalitatea de separare a textului, fie entitatile ce vrea sa le recunoasca nu sunt setate

Set register

- 1) Set register

Add entity

- 1)Set textSeparator
- 2)Set entities
- 3)Set register
- 3)Analyze
- 4)add entity

Exceptii: fie modalitatea de separare a textului, fie entitatile ce vrea sa le recunoasca, fie registeri
Pentru fiecare entitate nu sunt setati

TextSeparator:

Create list

- 1) Create list

EntitiesExtractor:

Filter

- 1) Filter

Extract entities

- 1) Filtreer
- 2) Extract entities

EntityRegister:

Add

- 1) Add

ShortMemory:

Search

- 1) Search(lista)

LongMemory:

Search

- 1) Search(Interogari la baza de date)