В этом разделе я рассмотрела основные технологии, языка, библиотеки и инструментарий для работы с Natural Language Processing. Выполнила их сравнение с учетом недостатков и преимуществ.  
Этот этап является неотъемлемой частью для дальнейшего развития дипломного проекта. Ведь он позволяет адекватно произвести оценку и предусмотреть приблизительные результаты, учитывая недостатки и преимущества того или иного подхода.  
Именно в следующем разделе я собираюсь использовать полученные данные для дальнейшего внедрения функционала в дипомному проекте.

Аннататоры

Теориетичские ведомости

XMI

XML

UML

2.У цьому розділі описана специфіка і особливості реалізації програми парсера, технології, мови, та середовища розробки які були використані у процесі розробки. Був розглянутий механізм побудови UML діаграм, описані алгоритми для вилучення класів з дерева залежностей яке надає Core NLP. Також у цьму розділі були розроблені правила для конвертації частин мови у UML сутності. У розділі 2.3. розглядається робота з форматом XMI, впровадження та використання пакету javax.xml для запису XMI сутностей у файл. У розділі 2.4. описана логіка роботи програми та основні архітектурні особливостей парсеру, також у цьому розділі описано впровадження бібліотеки Core NLP та її використання. Також була наведена логіка побудови UML діаграми шляхом перетворення елементів проміжного графу у UML елементи.

Результатом виконаної роботи є розроблений програмний продукт для парсингу тексту, який використовує результат роботи бібліотеки CoreNLP, а саме - дерево залежностей, виконує його парсинг згідно розробленим правилам. Далі на основі отриманих даних, генерує UML модель та зберігає результат у форматі XMI.

Даний результат буде використовуватись у якості вхідних даних у дипломному проекті Олександра Василейко [2] з ціллю подальшого проведення їх аналізу, редагування та конвертації у формат OWL. Розроблене програмне забезпечення не є досконалим, тому що воно не вірішує всіх пролем з якими можна зіткнутися у процессі його роботи. Тому наступним кроком буде проведення експеременту та визначення всіх недоліків та переваг з метою покращення.

**3 Били взяты разние тексти**

**- одни спец подготовлени и отформатированы и отформатированы**

**- а другие нет**

* **Создание различных как положительных так и отрицательных условий для работы программного обеспечения**
* **Тестирование разработанного программного обеспечения моделируя различные события развития**
* **Анализ и оценка работы отдельного модуля**
* **Определение узких мест и возможности их оптимизации**

Підсумовуючи проведену роботу, можна стверджувати, що основна мета була досягнута, але це ніяк не означає що отриманий результат є ідеальним. Наступним у подальшому розвитку даного ПЗ буде, розширеня правил конвертації тектсу у UML діаграму, удосконалення та оптимізація алгоритму перетворення тексту у UML, а також розширення бібліотеки Stanford Core NLP.

**1**

Цей дипломний проект зосереджує у собі роботу з семантично насиченим коротким текстом, з використанням бібліотеки Stanford Core NLP, а також створення UML діаграм на основі цього тексту.

Практичне значення цієї дипломної роботи становить, отримання та використання UML діаграм у подальшому аналізі та редагуванні тексту, з метою конвертації у OWL формат (мова онтологій) — дипломна робота [2].

**2**

В первой главе работы рассмотрены краткие теоретические ведомости направления использования NLP и Обзор существующего инструментария

Одним из важнейших направлений использования NLP является разработка способности  
компьютерной программы "понимать" естественный язык используя классификацию  
  различных лингвистических единиц (текстов, слов, словосочетаний, предложений), реализуемой  
практически во всех приложениях лингвистического процессора, информационном поиске, машинном переводе,  
автоматическом реферирования и др.

**3.** Расммотреть та Огляд існуючого інструментарію   
Stanford Core NLP

Складність роботи з Natural Language Processing і недоліки існуючих бібліотек

**4** в частности были рассмотрена библиотека **Stanford Core NLP**

Стенфорд CoreNLP надає широкий набір інструментів для аналізу природної мови. Вона може надати базові форми слів, частини мови, абревіатури, власні назви, нормалізовані дати, час і числові величини. Також ця бібліотека може розпізнати структуру речення у термінах фраз та залежності слів. Вказує яка фраза відноситься до тієї ж сутності, вказує настрій, а також розуміє цитати [4].

Вона містіть такі модулі аналізу тексту (Анататори) Додаток Д, які підтримують наступні мови таблиця

**5** Аннататори с помощью них и происходи парсинг текста

я использовала такие такие чтоб получить такие результати

дерево вида

Теориетичские ведомотсии

XMI

XML

UML

У цьому розділі я розглянула основні технології, мови, бібліотеки та інструментарій для роботи з Natural Language Processing. Виконала їх порівняння з урахуванням недоліків та переваг.

Цей етап є невід’ємною частиною для подальшого розвитку дипломного проекту. Адже він дозволяє адекватно зробити оцінку та передбачити приблизні результати, враховуючи недоліки та переваги одного чи іншого підходу.

Саме у слідуючому розділі я збираюсь використовувати отримані дані з метою подальшого впровадження функціоналу у дипомному проекті.

В первой главе работы была рассмотрена следующая задача ...»,

«Получены следующие результаты решения задачи ...