

# Звіт №3. Специфікація вимог до програмної системи

Звіт №3. Специфікація вимог до програмної системи	1
<b>Введення</b>	<b>3</b>
<b>Глосарій</b>	<b>3</b>
<b>Специфікація функціональних користувацьких вимог</b>	<b>4</b>
U1. Надіслати файли на аналіз	4
U2. Завантажити документ зі сторонніх сервісів	4
U3. Отримати попереднє посилання на результат аналізу документів.	5
U4. Отримати відповідь на спробу переглянути результат раніше готовності системи	5
U5. Отримати відповідь на спробу отримати неіснуючий результат аналізу документів	5
U6. Отримати результати аналізу за посиланням	5
U7. Переглянути графіки	5
U8. Порівняти надіслані файли	5
<b>Специфікація функціональних системних вимог</b>	<b>6</b>
U1. Надсилання файлів на аналіз. Додаток № 1, 5.	6
U2. Інтеграція зі сторонніми сервісами зберігання документів	6
U3. Попереднє повідомлення про майбутнє місцезнаходження результату	6
U4. Постійне звітування про рівень готовності результатів аналізу документів	6
U5. Відповідь на спробу отримати результат раніше моменту готовності системи.	6
U6. Отримання результатів аналізу за посиланням. Додаток № 2, 3, 4.	7
U7. Перегляд графіків.	7
U8. Порівняння вихідних файлів	7
<b>Нефункціональні вимоги</b>	<b>8</b>
<b>Додаток 1</b>	<b>10</b>
<b>Додаток 2</b>	<b>11</b>
<b>Додаток 3</b>	<b>12</b>
<b>Додаток 4</b>	<b>13</b>

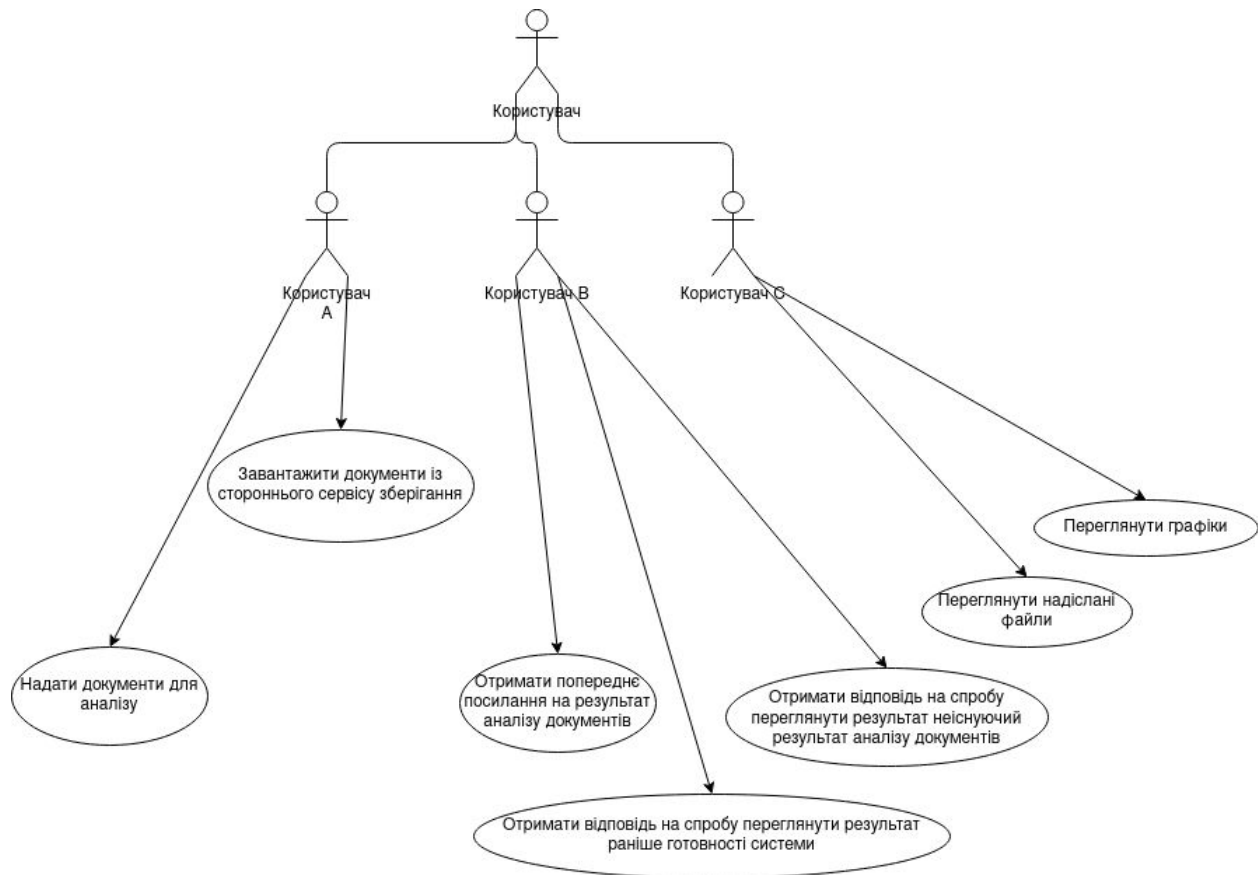
## Введення

Метою цього документу є означення функціоналу розроблюваної системи дослідження схожості текстових даних під назвою FOJ (Fair Online Judge) у частковій відповідності зі специфікацією вимог до програмної системи згідно зі стандартом IEEE 830 [1].

## Глосарій

Термін	Значення
<b>Fair Online Judge</b>	Назва розроблюваної системи.
<b>Документ</b>	Текстовий файл. Сукупність таких файлів обробляється та аналізується системою.
<b>Аналіз документів</b>	Процес обробки групи текстових файлів за допомогою засобів структурного та семантичного представлення тексту з метою з'ясувати схожість документів між собою.
<b>Користувач А</b>	Людина, яка завантажує файли до системи і таким чином надає системі ці файли для аналізу на схожість та для надання результатів аналізу за посиланням, строк дії якого обмежений.
<b>Користувач В</b>	Людина, яка очікує готовності результату аналізу текстових даних, надісланих на обробку до сервісу.
<b>Користувач С</b>	Людина, яка інтерактивно вивчає результати аналізу, проведеного системою над файлами, які завантажив <b>Користувач А</b> .

## Специфікація функціональних користувацьких вимог



Назва	Актор	Опис
U1. Надіслати файли на аналіз	Користувач А	Система має надавати користувачу можливість надсилати довільну кількість текстових файлів для аналізу.
U2. Завантажити документ зі сторонніх сервісів	Користувач А	Користувач має змогу надати доступ до сервісів зберігання простих для читання документів з метою дозволити FOJ прочитати і проаналізувати ці документи.

U3. Отримати попереднє посилання на результат аналізу документів.	Користувач В	За успішного надання групи текстових файлів система має надати користувачу посилання на результати виконаного над ними аналізу.
U4. Отримати відповідь на спробу переглянути результат раніше готовності системи	Користувач В	Потрібно підтримувати статус рівня готовності системи надати результат.
U5. Отримати відповідь на спробу отримати неіснуючий результат аналізу документів	Користувач В	Система має відповідати на спроби користувача переглянути результати аналізу, навіть якщо результати не можуть бути надані.
U6. Отримати результати аналізу за посиланням	Користувач В	При переході за згенерованим системою посиланням система має надати користувачу результати аналізу над надісланими <b>Користувачем А</b> даними.
U7. Переглянути графіки	Користувач С	Ознайомлення із результатами аналізу надісланих текстових даних включає в себе перегляд інформативного графіку.
U8. Порівняти надіслані файли	Користувач С	Ознайомлення із результатами аналізу надісланих текстових даних включає в себе порівняння надісланих файлів.

## Специфікація функціональних системних вимог

### U1. Надсилання файлів на аналіз. Додаток № 1, 5.

**Користувач А** має змогу за допомогою інтерфейсу веб-браузеру (Chrome, Firefox, Edge) на мобільному та настільному пристрої відправити на сервіс **FOJ** сумарно обсяг даних, що не перевищує **10 мегабайт**. Система має працювати із текстовими файлами, які закодовані у форматі UTF-8. У випадку надсилання будь-якого іншого типу даних система повідомляє користувача про відмову в обслуговуванні із інформативним повідомленням про помилку.

Якщо документи не було отримано сервісом або якщо було отримано лише один, система має повідомити про це **Користувача А** та відмовити йому в обслуговуванні.

### U2. Інтеграція зі сторонніми сервісами зберігання документів

**Користувач А** має змогу надати доступ до сервісів зберігання простих для читання документів (Dropbox, Google Docs) з метою дозволити FOJ прочитати і проаналізувати ці документи. Система працює виключно із документами текстового формату.

### U3. Попереднє повідомлення про майбутнє місцезнаходження результату

Система має надати **Користувачу В** посилання (URL) на ресурс. У посилання має міститися унікальний ідентифікатор користувача. Унікальний ідентифікатор має відповідати стандарту UUID4.

### U4. Постійне звітування про рівень готовності результатів аналізу документів

У випадку звернення **Користувача В** до системи за посиланням, яке було надано системою **Користувачу А** раніше, система має переадресувати користувача на веб-сторінку із повідомленням, яке у формі графічного елементу Status Bar показує рівень готовності системи надати результат.

### U5. Відповідь на спробу отримати результат раніше моменту готовності системи.

Система має обробляти запити до сторінки, призначеної для переглядання результату, протягом усього часу своєї роботи. У випадку, коли система отримує запит на перегляд такого ресурсу, але

він ще не готовий, вона має повідомити користувача про свою неготовність. Система має відповідати на спроби користувача отримати результати на неіснуючі результати аналізу документів.

U6. Отримання результатів аналізу за посиланням.  
Додаток № 2, 3, 4.

Результати аналізу мають бути отримані у вигляді матриці відстаней. Дані користувачів, які були завантажені у систему, мають бути доступні не більше ніж добу з моменту їх отримання на сервері.

U7. Перегляд графіків.

Зображення матриці відстаней відбувається у вигляді теплової карти (двовимірної сітки, клітини якої представляють відстань від одного файла до іншого за допомогою насиченості кольору).

U8. Порівняння вихідних файлів

Користувач обирає два файли та переходить до режиму їх перегляду у двох стилізованих елементах веб-сторінки, які слугують буферами, призначеними тільки для читання.

## Нефункціональні вимоги

Код	Вимога
NF1	Система має бути виконана як веб-сервіс, працювати із браузерами за протоколом HTTP, бути доступною для користувача мережі Інтернет.
NF2	<p>Необхідність швидкої роботи алгоритмів аналізу текстової інформації не є критичною, тому можна дозволити системі працювати у два етапи:</p> <ul style="list-style-type: none"><li>• прийом даних і відправлення їх на обробку у асинхронному режимі;</li><li>• обробка даних, у результаті якої ресурс, який обіцяний користувачу, стає доступним.</li></ul> <p>Очікування результату тривалістю у 2 хвилини вважається допустимим. Потрібно наголосити на потенційній розширюваності системи, оскільки передбачено використання системою часу на обробку інформації.</p>
NF3	Веб-сторінка має проходити тест на адаптивність, передбачений сервісом <a href="https://responsivedesignchecker.com/">https://responsivedesignchecker.com/</a> . Користувацький інтерфейс системи має бути доступним на мобільних та десктопних пристроях.
NF4	Система має будувати веб-інтерфейс на фреймворку Vue.js у зв'язку із простотою підтримуваності такого рішення і недостатнім рівнем готовності команди працювати із альтернативами – React, Angular.
NF5	Система має використовувати файлову систему ОС серверу та базу даних SQLite як сховище даних із мінімальним необхідним обсягом роботи для його підтримання та супроводу.
NF6	Результати виконаного аналізу над надісланими користувачем даними мають бути дійсними не більше ніж протягом однієї доби з моменту появи їх на сервері. Дотримання верхньої межі на обсяг інформації, що зберігається, зумовлене обмеженістю дискового простору серверу.
NF7	Має бути інтеграція із стороннім сервісом зберігання даних, яке працює за протоколом S3, який би дозволив не вирішувати проблему обмеженості дискового простору. Такий підхід



	зменшить обсяг проблем, які виникають за бажання горизонтально розширюватися.
NF8	Скористатися наступними рішеннями в області розробки застосунків NLP (Natural Language Processing): модель для представлення слів GloVe [3]. Використати наступні підходи при написанні програмного коду: cosine similarity (для структурної подібності), soft cosine similarity (для семантичної подібності), алгоритм Майерса [4, 5].
NF9	Систему потрібно реалізувати із урахуванням того, що значна частина логіки має перебувати на веб-сервісі, а фронтенд має спиратися на надаваний інтерфейс (REST API). Таким чином буде отриманий більший рівень модульності системи.

## Додаток 1

### **Прецедент U1: Надсилання документів**

#### 1.1 Прецедент U1: Надсилання документів

##### 1.1.1 *Короткий опис*

Система має надавати користувачу можливість надсилати довільну кількість текстових файлів для аналізу.

*Діюча особа:* **Користувач А**

#### 1.2 Потік подій

Прецедент починається, коли **Користувач А** звертається до сервісу у браузері.

##### 1.2.1 *Базовий потік - Вхід*

1. Користувач звертається до сервісу із запитом отримати стартову сторінку.

2. Користувач обирає спосіб надання документів на сервіс. У випадку надання за допомогою завантаження із локального пристрою переходить до п.3.

3. Сервіс перевіряє, чи не відправив користувач лише один файл, та чи не відправив користувач файл недопустимого формату. Якщо ні, система переходить до п.4.

4. Користувач отримує посилання від веб-сервісу на результати аналізу документів, які будуть доступні після завершення обчислень.

##### 1.2.2 *Альтернативні потоки*

1.2.2.1 Користувач обирає спосіб завантаження документів на сервер за допомогою стороннього сервісу зберігання документів. Користувач не надіслав більше одного документу або принаймні один з документів має недопустимий формат.

Якщо при виконанні п.2 користувач обрав спосіб надання документу за допомогою стороннього сервісу, то

1. Система надає користувачу можливість надати документ за допомогою такого сервісу. Підтримуються сервіси Google Drive, Dropbox.

Якщо при виконанні п.3 **Користувач А** не надіслав валідні для роботи сервісу дані, то

1. Система надає користувачу повідомлення про помилку та відмову в обслуговуванні.

#### 1.3 **Спеціальні вимоги**

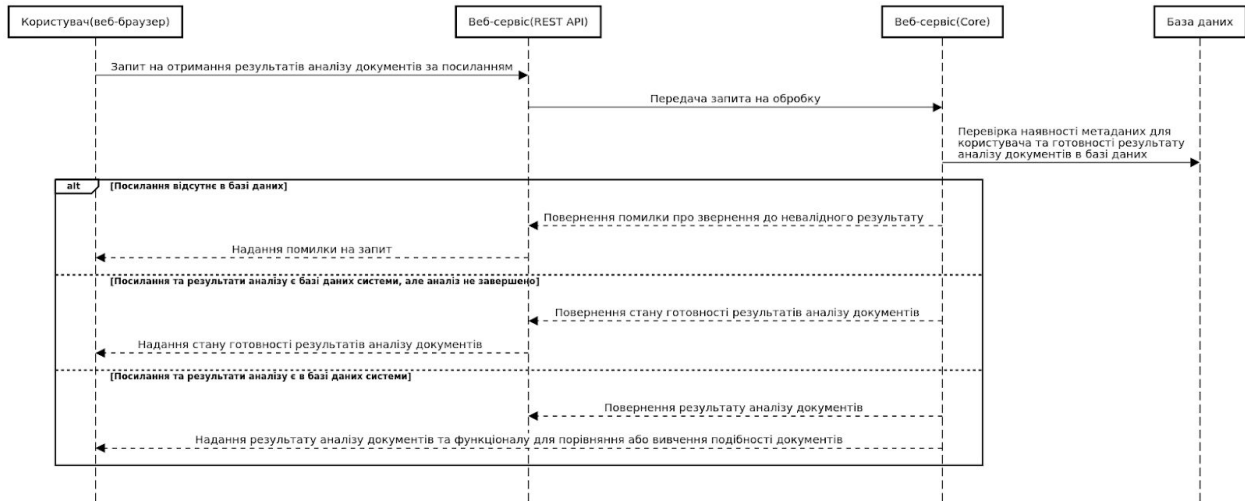
Відсутні.

#### 1.4 **Передумови**

Відсутні.

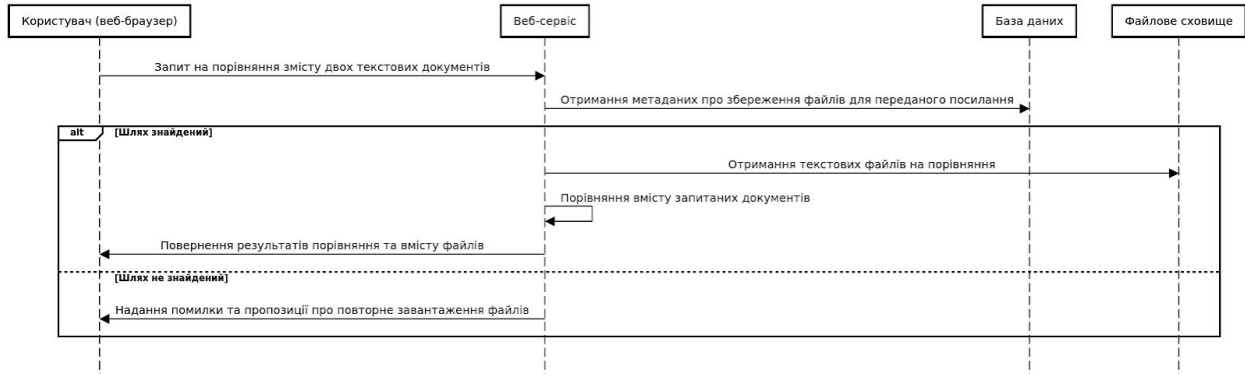
## Додаток 2

Процес звернення користувача до веб-сервісу з наміром дослідити результати аналізу документів



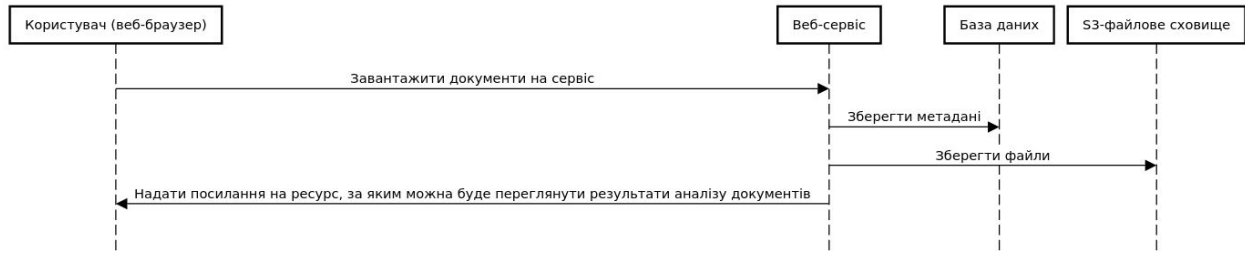
## Додаток 3

Процес запиту Користувача з метою порівняння вмісту двох документів



## Додаток 4

### Процес завантаження Користувачем документів на сервіс



#### Посилання

- [1] [https://en.wikipedia.org/wiki/Software\\_requirements\\_specification](https://en.wikipedia.org/wiki/Software_requirements_specification)
- [2] <https://tools.ietf.org/html/rfc4122>
- [3] [https://en.wikipedia.org/wiki/GloVe\\_\(machine\\_learning\)](https://en.wikipedia.org/wiki/GloVe_(machine_learning))
- [4] [https://en.wikipedia.org/wiki/Cosine\\_similarity](https://en.wikipedia.org/wiki/Cosine_similarity)
- [5] <https://neil.fraser.name/writing/diff/myers.pdf>