**Лабораторна робота №2**

**Попередня обробка тексту. Стоп-слова**

**Виконали:**

|  |  |
| --- | --- |
| **Група** | **ПІБ** |
| ПД-31 | Гапей Максим Юрійович |
| ПД-31 | Марковський Павло Павлович |

**Мета роботи:** отримати навички попередньої обробки текстів на прикладі задачі виділення стоп-слів.

**Завдання**

* Скласти для тексту із лабораторної роботи №1 список стоп-слів з урахуванням тематики предметної області. Дозволяється використання існуючих списків стоп-слів, які треба доповнити незначущими для даної предметної області словами.   
  *Зауваження.* *Не слід включати до стоп-слів ВСІ слова, які не мають безпосереднього відношення до тематики тексту. Крім того, при складанні списку стоп-слів слід пам’ятати, що список стоп-слів зазвичай містить нормальні форми слів, а в тексті їм можуть відповідати декілька словоформ.*
* Позначити безпосередньо в тексті знайдені стоп-слова (рекомендується різні стоп слова помітити різним кольором чи різною заливкою тексту).
* Створити звіт із описом визначених груп стоп-слів (назва групи стоп-слів, чому ця група виділена, опис важливих ознак групи). Результати можна звести у вигляді таблиці:

|  |  |  |  |
| --- | --- | --- | --- |
| **Назва групи стоп-слів** | **Чому ця група вважається стоп-словами** | **Опис важливих ознак групи** | **Приклади елементів групи** |
| цифри | в більшості випадків цифри не несуть смислового навантаження | прописні, арабські, римські цифри | 1, 2, три, чотири, V, VI |
| знаки пунктуації | пунктуація ігнорується майже на 99% | символи, які не являються цифрами та буквами | . , = + / ! " ; : % ? \* ( ) # $ < > |
| прийменники, займенники, дієприкметники, суфікси,  одиничні букви | зазвичай не змінюють семантику речення | не є іменником, дієсловом та прикметником | у, в, він, вона, був, була, як, та, більш, менш |
| часто вживані слова | подібні слова у тексті | часто повторюються та мають спілий корінь (основу слова) | Рік, рокам, роком, років, роки, роками |

**Текст 17**

**Zamrii I.V., Vyshnivskyi V.V.**

*State University of Telecommunications, Kyiv*

**STRUCTURE OF THE SHARED DATA ENVIRONMENT** **OF THE ENTERPRISE WITH CRITICAL INFRASTRUCTURE**

У роботах [1, 2] досліджуються властивості систем, які забезпечують можливості їхнього функціонування при змінах параметрів внутрішнього та зовнішнього середовища впродовж тривалих проміжків часу. Характер поведінки системи обирається відповідно до змін зовнішніх умов і з врахуванням функціонального інваріанту системи, який можна назвати внутрішньою метою її функціонування.

Аналізуючи останні наукові дослідження можна стверджувати, що методам та моделям побудови складних технічних систем присвячено наукові роботи О.А. Машкова [3], О.В. Барабаша [4], Г.А. Кучука [5] та ін.; проблемі відмовостійкості складних технічних систем присвячено роботи В.А. Машкова, В.А. Савченка [6] та ін.; питання стійкості систем відносно зовнішніх дестабілізуючих впливів досліджувалося в роботах О.Г. Додонова [7], І.В. Рубанова та ін.

Питанням управління складністю інформаційних систем підприємства та впровадженню інформаційних технологій присвячено роботи [8-10]. У роботах [11-16] досліджувалися питання організації єдиного інформаційного простору для різних підприємств та галузей. Зокрема, робота [17] описує методику створення єдиного інформаційного простору на сучасному виробничому підприємстві з функціонально стійким виробничим процесом. Сучасні промислові підприємства у великій мірі автоматизовані та здатні автономно виконувати багато технологічних процесів протягом наперед заданих часових інтервалів, випускаючи продукцію з високими стандартами споживчої та експлуатаційної якості [18-22].

В результаті накопиченого досвіду в даній сфері на сьогоднішній день розроблено велика кількість методик оперативно-виробничого планування та управління промисловим підприємством, але важливим і відкритим питанням залишається об’єднання всіх систем для оптимізації, автоматизації, зменшення трудомісткості, збільшення завадостійкості до внутрішніх та зовнішніх дестабілізуючих факторів.

Для інформаційного супроводу виробів виробничого центру підприємства з критичною інфраструктурою необхідне безперервне ведення і вдосконалення єдиного інформаційного простору. Це забезпечує ефективний впорядкований обмін супроводжувальною інформацією, що інтенсивно змінюється, незалежно від методів і засобів програмно-технічної реалізації та стимулює створення так званих географічно та організаційно розподілених виробничих систем.

Сontinuous acquisition and life cycle support (CALS) забезпечує інформаційну підтримку виробу протягом всіх етапів виробництва [20, 21]. Вона заснована на застосуванні інтегрованого інформаційного середовища при забезпеченні одноманітних способів управління процесами та взаємодії всіх учасників процесу. Серед основних цілей CALS є безперервне забезпечення якості та оптимізація об'єктів та процесів за допомогою зміни їхньої конфігурації (рис. 1), при цьому головним інструментом даних технологій буде єдиний відкритий інформаційний простір (SDE - shared data environment).

Програмні комплекси CALS-технологій у рамках єдиного інформаційного простору повинні відповідати наступним важливим вимогам:

* дотримання об'єктно-орієнтованого принципу побудови комп'ютерних систем;
* повна інтеграція систем між собою за наявності можливості передачі інформації між компонентами та модулями за рахунок своїх внутрішніх зв'язків;
* можливість функціонування систем як у спільному, так і в автономному режимах;
* наявність єдиного інтерфейсу;
* мережева взаємодія між активними користувачами даних програмних комплексів;
* єдині архітектури комп'ютерних баз даних та методів доступу до них.

CALS-технології передбачають безперервне вдосконалення та модернізацію (реінжиніринг) виробничих структур та бізнес-процесів (business processes reengineering) [16].

Останні на основі інформаційно-комунікаційних технологій та використання єдиного інформаційного простору забезпечують досягнення поставлених цілей та реалізацію сформульованої політики (рис. 2).