

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

ОДЕСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ імені І.І.МЕЧНИКОВА

(повне найменування вищого навчального закладу)

Факультет математики, фізики та інформаційних технологій

(повне найменування інституту, назва факультету (відділення))

Кафедра математичного забезпечення комп'ютерних систем

(повна назва кафедри (предметної, циклової комісії))

Дипломна робота

на здобуття освітньо-кваліфікаційного рівня «бакалавр»

(освітньо-кваліфікаційний рівень)

на тему Система аналізу відповідності коду документа універсальному
десятковому класифікатору

Analysis system the document code compliance with the universal decimal
classifier

Виконав: студент денної форми навчання

напряму підготовки 6.050102 – Комп'ютерна інженерія

(шифр і назва напряму підготовки, спеціальності)

Хоба Юрій Юрійович

(прізвище, ім'я, по-батькові)

Керівник ст. викладач Трубіна Н. Ф.

(науковий ступінь, вчене звання, прізвище та ініціали, підпис)

Рецензент доц. Шугайло Ю. Б.

(науковий ступінь, вчене звання, прізвище та ініціали)

Рекомендовано до захисту:

Протокол засідання кафедри

№ від «» 2019 р.

Завідувач кафедри

Захищено на засіданні ЕК №

протокол № від «» 2019 р.

Оцінка / /
(за національною шкалою, шкалою ECTS, бали)

Голова ЕК

Є.В. Малахов

(підпис)

(прізвище, ініціали)

О.О. Арсірій

(прізвище, ініціали)

Одеса - 2019

АНОТАЦІЯ

Метою роботи є створення електронної інтерактивної системи аналізу відповідності коду документа універсальному десятковому класифікатору. Система дає можливість перегляду усіх внесених до неї розділів УДК та ключових слів, а також дозволяє зручно керувати ними. В системі реалізований власний алгоритм, що дозволяє за анотацією або текстом документу визначити відповідність коду документу за універсальним десятковим класифікатором його вмісту по ключовим словам з бази даних.

Дані системи зберігаються в базі даних, створеній за допомогою PostgreSQL. Клієнтські додатки створені на мові C# за допомогою середовища розробки Microsoft Visual Studio 2017. Аналіз відповідності даних здійснюється за допомогою запропонованого алгоритму і стандартних інструментів PostgreSQL.

ABSTRACT

The aim is to create an electronic interactive analysis system for the document code compliance with the universal decimal classifier. The system allows user to view all UDC sections and keywords entered and to manage them conveniently. The system has its own algorithm, which allows given annotation or the text of a document to determine the document code compliance with the universal decimal classifier of its content by keywords from the database.

Data is stored in the system database created using PostgreSQL. Client applications were written in C# using Microsoft Visual Studio 2017 development environment. Data compliance analysis is performed by using the proposed algorithm and standard PostgreSQL tools.

АННОТАЦИЯ

Целью работы является создание электронной интерактивной системы анализа соответствия кода документа универсальному десятичному классификатору. Система дает возможность просмотра всех занесенных в нее разделов УДК и ключевых слов, а также позволяет удобно ими управлять. В системе реализован собственный алгоритм, позволяющий по аннотации или тексту документа определить соответствие кода документа универсальному десятичному классификатору его содержания по ключевым словам из базы данных.

Данные системы хранятся в базе данных, созданной при помощи PostgreSQL. Клиентские приложения написаны на языке C# при помощи среды разработки Microsoft Visual Studio 2017. Анализ соответствия данных совершается при помощи предложенного алгоритма и стандартных инструментов PostgreSQL.

ЗМІСТ

ВСТУП	6
1 УНІВЕРСАЛЬНА ДЕСЯТКОВА КЛАСИФІКАЦІЯ ДОКУМЕНТІВ	8
1.1 Огляд предметної області.....	8
1.2 Класифікація документу за ключовими словами	12
1.3 Web-сервіси для визначення УДК.....	13
1.4 Висновки та уточнення завдання.....	17
2 ПРОЕКТУВАННЯ СИСТЕМИ.....	19
2.1 Загальна архітектура системи	19
2.2 Вимоги до системи.....	20
2.3 Побудова моделі інформаційної системи.....	23
2.4 Вибір засобів розробки	27
2.5 Методи аналізу даних	29
3 РЕАЛІЗАЦІЯ СИСТЕМИ.....	35
3.1 Реалізація моделі даних	35
3.2 Реалізація адміністраторського додатку	35
3.3 Реалізація додатку користувача.....	46
4 ТЕСТУВАННЯ СИСТЕМИ	54
ВИСНОВКИ.....	56
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	57
ДОДАТОК А СЦЕНАРІЙ СТВОРЕННЯ БАЗИ ДАНИХ	60
ДОДАТОК Б СЦЕНАРІЙ СТВОРЕННЯ ЗБЕРЕЖЕНИХ ПРОЦЕДУР.....	61
ДОДАТОК В РЕЗУЛЬТАТИ ТЕСТУВАННЯ.....	62

ВСТУП

Універсальна десяткова класифікація була розроблена у зв'язку зі зростанням кількості наукових праць і друкованих видань, і використовується для швидкого та зручного пошуку розділу літератури, науки або мистецтва, до якого відноситься публікація.

УДК має 9 кореневих розділів, які повністю охоплюють усі тематичні напрямки. Однак через складність системи позначень розділів, автору публікації може бути складно правильно визначити УДК документу. Визначити УДК документу може людина, яка розбирається як в бібліографії, так і в проблематиці статті. Тому код, визначений неосвіченою в даній галузі людиною, може бути помилковим. У зв'язку з цим виникає потреба у створенні системи, що дозволяє провести аналіз відповідності вже заданого коду документа, універсальному десятковому класифікатору, визначеному цією системою. Тобто, система повинна на основі анотації чи ключових слів документу визначати, наскільки правильно користувач визначив УДК документу.

Поставлена задача є досить актуальною для різних видавництв, а також для інтернет-ресурсів, які розміщують на своїх серверах наукові статті та літературні твори. Ще більшої актуальності цієї задачі надає постанова Кабінету міністрів України «Про припинення використання Бібліотечно-бібліографічної класифікації та впровадження Універсальної десяткової класифікації» № 177 від 22.03.2017 р, що передбачає поступове переведення бібліотек на універсальну десяткову класифікацію як міжнародну систему.

Об'єктом дослідження є універсальний десятковий класифікатор. Класифікація – це поділ множини об'єктів на класифікаційні угруповання за їх схожості або відмінності на основі певних ознак відповідно до прийнятих правил. Ознаками у даному випадку є ключові слова документу. Правилами ж є завчасно накопичені дані, тобто відповідності ключових слів до УДК, зібрані з книг, статей або публікацій. Класифікація проводиться шляхом

встановлення закономірностей між ключовими словами і УДК, яким вони відповідають. Класифікація документів досить складна і має ієрархічну структуру.

Предметом дослідження є методи класифікації документу за ключовими словами або анотацією. Ключові слова, як і анотація, дають опис змісту текстового документа високого рівня, що дозволяє виявити його тематику. Кожен код може бути описаний за допомогою набору ключових слів. Таким чином, для визначення УДК документу, необхідно сперш за все сформувати базу даних з відповідностей ключових слів до УДК, на основі певних документів і статей.

Метою даної роботи є створення системи для оцінки правильності визначення коду документа. Дана система буде надавати можливість накопичувати знання, додаючи до неї нові коди УДК та ключові слова, що відповідають розділам УДК, а також на основі цих знань аналізувати правильність визначення УДК коду документа користувачем.

Для досягнення поставленої мети необхідно вирішити наступні завдання:

- дослідити предмету область;
- проаналізувати наявність допоміжних засобів визначення УДК;
- сформулювати вимоги до створюваної системи;
- запропонувати деяку математичну модель оцінки якості співвідношення документа до розділу УДК;
- провести проектування бази даних інформаційної системи;
- провести проектування прикладної програми;
- обрати інструментальне середовище розробки і виконати програмну реалізацію системи;
- провести тестування розробленої програми.

ВИСНОВКИ

В ході виконання дипломної роботи, була розроблена система, яка включає у себе базу даних і два настільних додатки, і дозволяє проводити аналіз відповідності коду документу універсальному десятичному класифікатору.

Була розглянена предметна область і недоліки існуючих онлайн-сервісів для визначення УДК, на основі чого були зроблені висновки і уточнення до завдання.

Були сформульовані вимоги до системи, спроектовані додатки і база даних, запропонована власна модель оцінки якості відповідності документу до розділу УДК, описані методи аналізу даних, на підставі чого були обрані засоби розробки, а потім розроблена уся система.

Методи розробки були обрані по причині їх переваг: швидкість і зручність розробки, наявність необхідних інструментів для виконання поставлених задач, безкоштовність.

У результаті була розроблена система, яка має простий і зручний дизайн. Вона дозволяє адміністратору накопичувати дані і керувати ними, а користувачу – аналізувати відповідність вибраних ключових слів до конкретного розділу УДК. Система також надає в процесі роботи користувачу різну інформацію, необхідну для правильного визначення відповідності ключових слів до УДК.

Напрямок подальшого розвитку системи є удосконалення аналізу відповідності документу до розділу УДК, шляхом обчислення додаткових коефіцієнтів ваги кожного ключового слова до розділів УДК, на основі аналізу частоти появи цих ключових слів у різних документах. Також необхідно поліпшити якість і корисність інформації, що надається користувачу в процесі роботи.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Універсальна десяткова класифікація – Потрібно Знати [Електронний ресурс]. – Режим доступу: http://nado.znate.ru/Универсальная_десятичная_классификация
2. УДК: структура, властивості та принципи [Електронний ресурс]. – Режим доступу: <http://www.spsl.nsc.ru/win/nelbib/udk/strukt.html>
3. Онлайн-довідник TeaCode [Електронний ресурс]. – Режим доступу: <https://www.teacode.com/online/udc/>
4. Онлайн-довідник Triumph [Електронний ресурс]. – Режим доступу: <https://www.triumph.ru/udk-klassifikator.html>
5. Онлайн-довідник НовыйФормат [Електронний ресурс]. – Режим доступу: <https://типография-новый-формат.рф/udk>
6. Онлайн-довідник UDCSummary [Електронний ресурс]. – Режим доступу: <http://www.udcsummary.info/php/index.php>
7. Postgres Pro Standard: Документація – Типи, призначені для текстового пошуку [Електронний ресурс]. – Режим доступу: <https://postgrespro.ru/docs/postgrespro/9.5/datatype-textsearch>
8. Міжнародний науково-теоретичний журнал «Кібернетика і системний аналіз» [Електронний ресурс]. – Режим доступу: <http://www.kibernetika.org/ContentsUA.html>
9. Міжнародний науковий журнал «УСиМ» [Електронний ресурс]. – Режим доступу: <http://usim.org.ua/?lang=uk>
10. Міжнародний наукометричний науковий журнал «Східно-Європейський журнал передових технологій» [Електронний ресурс]. – Режим доступу: <http://journals.uran.ua/eejet/issue/archive>

11. Алфавітно-предметні покажчики до основних розділів УДК [Електронний ресурс]. – Режим доступу: <http://www.spsl.nsc.ru/win/nelbib/UDK/part2.html>
12. Проблеми організації об'єктно-орієнтованого доступу до реляційних баз даних [Електронний ресурс]. – Режим доступу: http://window.edu.ru/resource/979/70979/files/isp_15_2008.pdf
13. Дефрагментація і статистика індексів SQL Server [Електронний ресурс]. – Режим доступу: <https://www.flenov.info/story/show/Defragmentaciya-i-statistika-indeksov-SQL-Server>
14. Інтегрована аналітика. Як отримати максимальну вигоду з ERP-систем [Електронний ресурс]. – Режим доступу: <http://citforum.ck.ua/database/articles/erp-system/>
15. Прикладні системи штучного інтелекту [Електронний ресурс]. – Режим доступу: https://studme.org/116358/matematika_himiya_fizik/prikladnye_sistemy_iskusstvennogo_intellekta
16. Де застосовують нейрокомп'ютери у фінансовій діяльності [Електронний ресурс]. – Режим доступу: <https://neuronus.com/stat/175-primeneniya-nejrokompyuterov-v-finansovoj-deyatelnosti.html>
17. Вивчення людського мозку за допомогою штучних нейронних мереж [Електронний ресурс]. – Режим доступу: <https://neuronus.com/stat/1269-izuchenie-chelovecheskogo-mozga-pri-pomoshchi-iskusstvennykh-nejronnykh-setej.html>
18. Прончев Г. Б., Об актуальности локальных вычислительных сетей в настоящее время / Г. Б. Прончев, В. В. Фесенко, В. В. Брутов, В. Г. Михасев, С. Г. Воробьев // Молодой ученый. 2010. №12. Т.1. – С. 53.
19. Методичні матеріали «Локальні мережі в освітніх установах» [Електронний ресурс]. – Режим доступу: <https://www.gyrnal.ru/library/ru/obshee/2/2/3/>

20. Ветелина Е., Исследование тенденций развития современных сетевых технологий на примере программно-конфигурируемых сетей // Новые информационные технологии в автоматизированных системах. 2015. – С. 502.
21. Атамкулова М.Т., Компьютерные вирусы и антивирусные программы / М. Т. Атамкулова, А. А. Саримсаков // Известия ошского технологического университета. 2016. Т.2. – С. 136.
22. Усенков Д. Ю., Компьютерные вирусы: «борьба брони и снаряда» // Компьютерные инструменты в образовании, 2004. №3. – С. 15.

ДОДАТОК А

СЦЕНАРІЙ СТВОРЕННЯ БАЗИ ДАНИХ

```
CREATE TABLE udc (  
    id SERIAL PRIMARY KEY,  
    code TEXT UNIQUE NOT NULL CHECK (code SIMILAR TO ('([0-  
9]|.|/|-|`)*')),  
    name TEXT NOT NULL CHECK (name NOT LIKE ''),  
    parent_id INTEGER REFERENCES udc NULL CHECK (parent_id <>  
id)  
);
```

```
CREATE TABLE keyword (  
    id SERIAL PRIMARY KEY,  
    name TEXT UNIQUE  
);
```

```
CREATE OR REPLACE FUNCTION check_key_word() RETURNS TRIGGER AS  
$$  
    DECLARE  
    BEGIN  
        NEW.name = lower(NEW.name);  
        RETURN NEW;  
    END;  
$$ LANGUAGE plpgsql;
```

```
CREATE TRIGGER keyword_trigger  
    BEFORE INSERT OR UPDATE ON keyword  
    FOR EACH ROW EXECUTE PROCEDURE check_key_word();
```

```
CREATE TABLE udc_keyword (  
    udc_id INTEGER REFERENCES udc,  
    keyword_id INT REFERENCES keyword,  
    PRIMARY KEY(udc_id, keyword_id)  
);
```

ДОДАТОК Б

СЦЕНАРІЙ СТВОРЕННЯ ЗБЕРЕЖЕНИХ ПРОЦЕДУР

```

CREATE OR REPLACE FUNCTION get_udk_parents_with_depth(INTEGER)
RETURNS TABLE (depth INTEGER, parent_id INTEGER)
AS $$
BEGIN
    RETURN QUERY
    WITH RECURSIVE
    rec_a (dpth, id, code, parent_id) AS
    (
        SELECT 0 AS dpth, udk.id, udk.code, udk.parent_id
    FROM udk WHERE id = $1
        UNION ALL
        SELECT dpth+1 AS depth, udk.id, udk.code,
    udk.parent_id FROM rec_a, udk WHERE udk.id = rec_a.parent_id
    )
    SELECT dpth, id FROM rec_a;-- WHERE id <> $1;
END;
$$ LANGUAGE plpgsql;

CREATE OR REPLACE FUNCTION get_udk_and_parents_keywords(integer)
RETURNS TABLE(id integer, name text, result boolean)
AS $$
BEGIN
    RETURN QUERY
    WITH udk_kws AS (
        SELECT k.id, k.name FROM keyword k JOIN test ts ON
    ts.keyword_id=k.id
        WHERE ts.udk_id = ANY (SELECT parent_id FROM
    get_udk_parents_with_depth($1))
        GROUP BY k.id, k.name ORDER BY 1 )
    SELECT results.id, results.name, res FROM (SELECT k.id,
    k.name, plainto_tsquery('russian', k.name) @@ t.tsv AS res
        FROM udk_kws k, t) AS results
    ORDER BY 1;
END;
$$ LANGUAGE plpgsql;

```

ДОДАТОК В

РЕЗУЛЬТАТИ ТЕСТУВАННЯ

Тестування для УДК «004.65 Бази даних та їх структура. Система керування базами даних».

Тест 1.

Текст анотації:

«Рассматривается задача нечеткого сравнения коллекций в приложениях реконсиляции. Задача возникает при оптимистической репликации структурированных данных и документов и имеет многочисленные приложения в таких актуальных областях, как управление конфигурацией программного обеспечения, управление мобильными базами данных, построение платформ и систем коллективной инженерии. В работе анализируются стандартные типы коллекций языков объектно-ориентированного моделирования, для которых описываются и обосновываются способы представления, журнализации, вычисления, принятия и согласования изменений. Для выделенных типов коллекций дается строгая, семантически содержательная интерпретация конфликтов и предлагаются конструктивные методы их идентификации и разрешения. Примеры иллюстрируют предлагаемые решения, а также служат мотивацией создания универсальной, основанной на модельном представлении оригинальной среды коллективной инженерии с развитыми возможностями семантически корректной и функционально содержательной реконсиляции дивергентных реплик данных.»

На рисунку В.1 приведений результат аналізу даного тексту. Система знайшла у цьому тексті 2 ключових слова, що відносяться до заданого розділу УДК (004.65). В додаток, система також знайшла у тексті 3 ключові слова, що відносяться до розділів УДК верхніх рівнів, але теж відносяться до заданого розділу УДК. Ці ключові слова описують більш загальні поняття,

тому мають менший вплив. Відповідно до формули 2.2, наведеної у розділі 2.5, був обчислений відсоток відповідності, що дорівнює 16%.

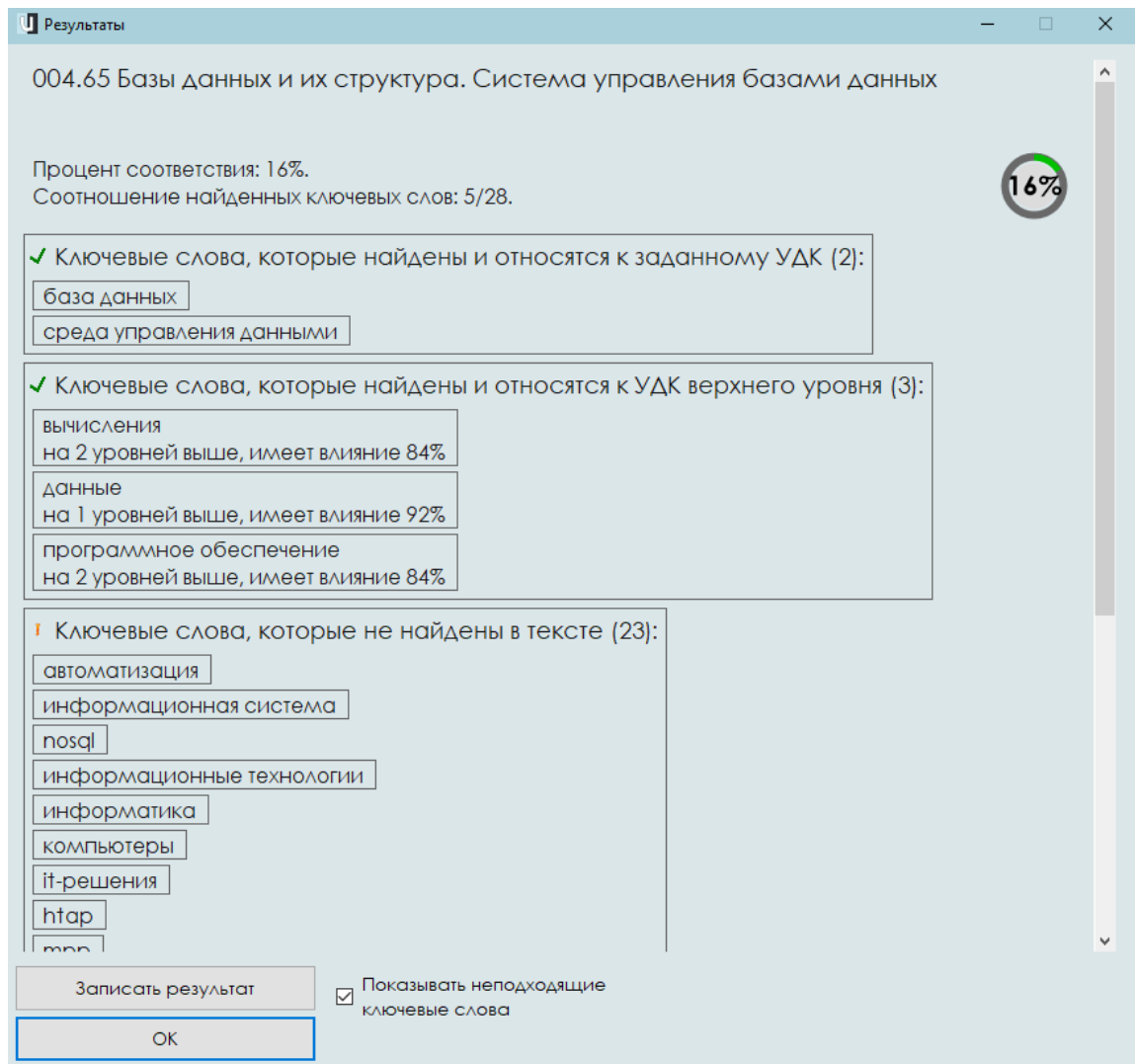


Рисунок В.1 – Результаты тесту 1

Система також може відобразити результат аналізу в іншому вигляді, показуючи користувачеві обрану гілку розділів УДК з усіма ключовими словами, які відносяться до цих розділів (рис. В.2). Тут же виділяються усі знайдені у тексті слова.

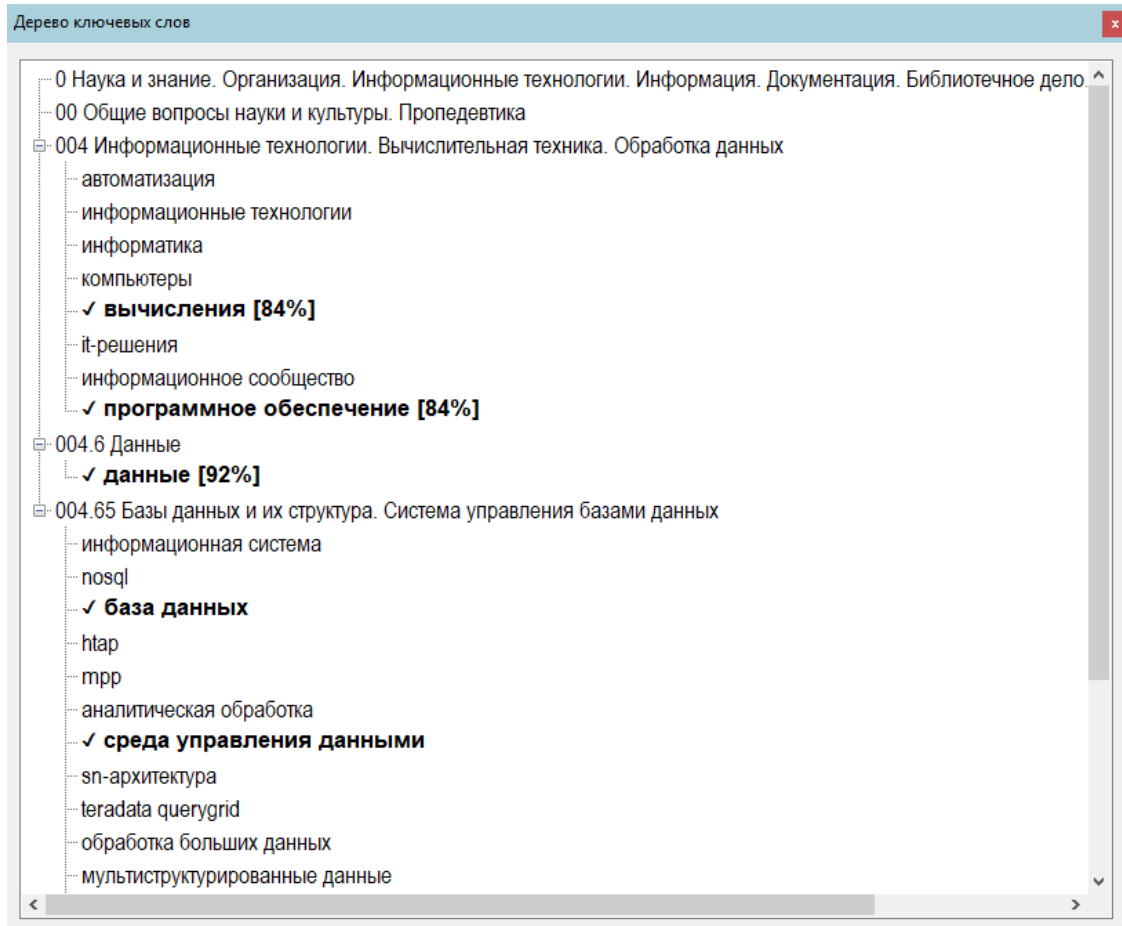


Рисунок В.2 – Додаткова інформація до тесту 1

Тест 2.

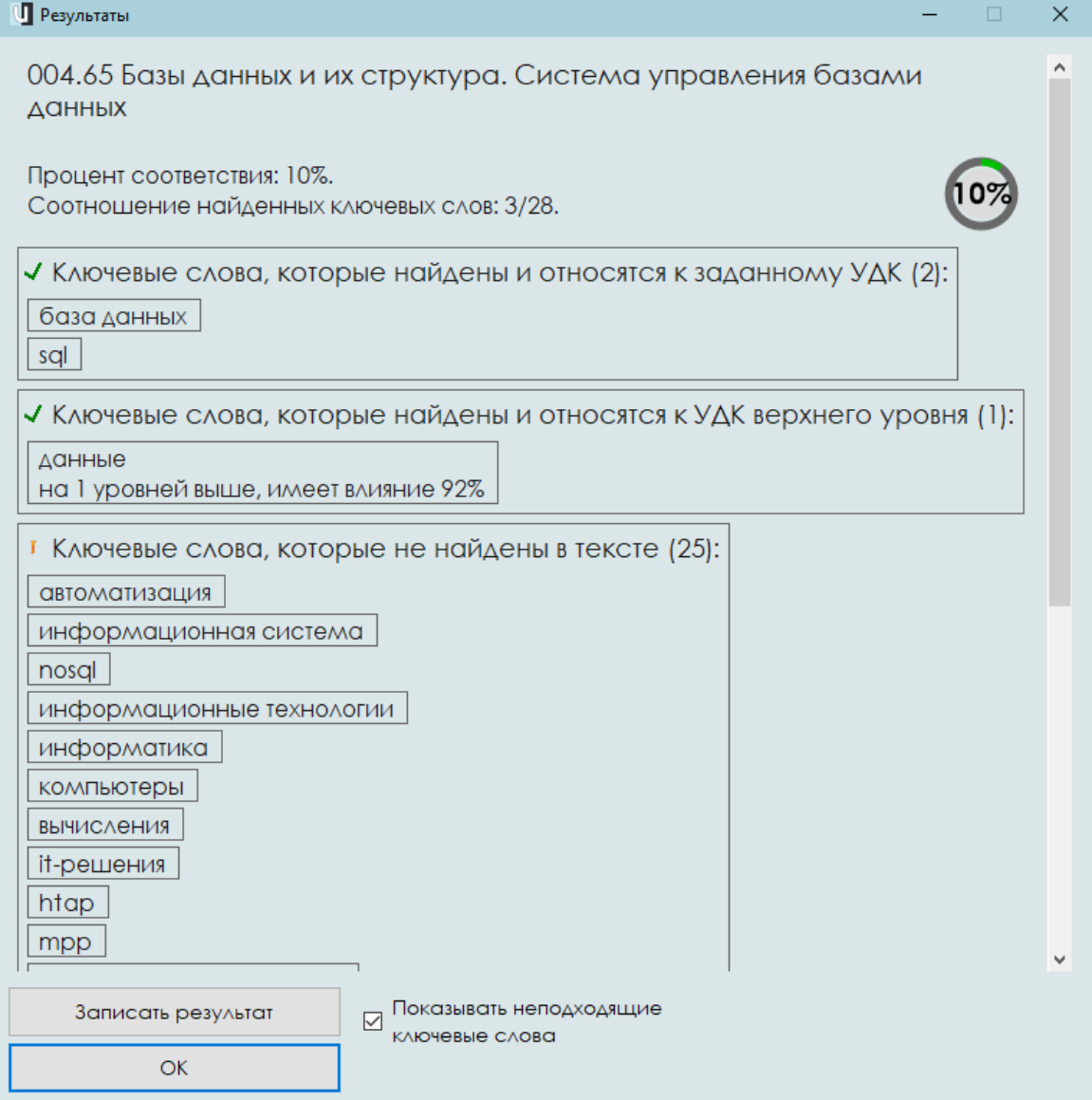
Текст статті:

«Продолжим знакомится с производительностью базы данных MS SQL Server и как ее улучшить и сегодня я решил рассказать про две примерно смежные темы - фрагментация индексов и статистика. Обе темы объединяют как раз индексы и они влияют на их производительность, поэтому я решил рассмотреть их одновременно.

Фрагментация индексов и некорректная статистика может повлиять на производительность базы данных и как SQL Server выполняет запросы. Сервер по умолчанию собирает статистику сам и нам не нужно заботиться о ней. При создании индексов можно указать – обновлять

статистику автоматически или нет и по умолчанию как раз все обновляется.»

Результат тесту 2 наведено на рисунку В.3.



Результаты

004.65 Базы данных и их структура. Система управления базами данных

Процент соответствия: 10%.
Соотношение найденных ключевых слов: 3/28.

10%

✓ Ключевые слова, которые найдены и относятся к заданному УДК (2):

база данных
sql

✓ Ключевые слова, которые найдены и относятся к УДК верхнего уровня (1):

данные
на 1 уровень выше, имеет влияние 92%

! Ключевые слова, которые не найдены в тексте (25):

автоматизация
информационная система
nosql
информационные технологии
информатика
компьютеры
вычисления
it-решения
htap
mpp

Записать результат

☒ Показывать неподходящие ключевые слова

OK

Рисунок В.3 – Результат тесту 2

Тест 3.

Текст статті:

«За последние десять лет крупные компании вкладывали большие средства, заменяя свои системы обработки транзакций ERP-системами

(Enterprise Resource Planning, ERP - системы Планирования ресурсов предприятия).

С другой стороны, большое значение уделялось разработке Хранилищ и/или витрин данных, позволяющих агрегировать и представлять данные и оказывающих существенную поддержку в принятии решений. Часто такие проекты выполнялись параллельно, однако при этом в ERP-системах никак не использовались возможности и преимущества Хранилищ данных.

В результате многие фирмы, потратившие миллионы долларов на ERP-решения, остались неудовлетворенными, так как не могут своевременно получить из систем обработки транзакций агрегированные данные. Следующее поколение ERP-систем должно учесть эту проблему.

В этой статье рассматривается развитие ERP-систем, роль Хранилищ данных в информационной ERP-архитектуре, а также перспективы создания интегрированных сред, использующих возможности этих двух технологий. Кроме того, дается оценка достижений двух крупнейших фирм (SAP и People Soft), производителей ERP-систем в области разработки Хранилищ данных и аналитических приложений.»

Результат тесту 3 наведено на рисунку В.4.

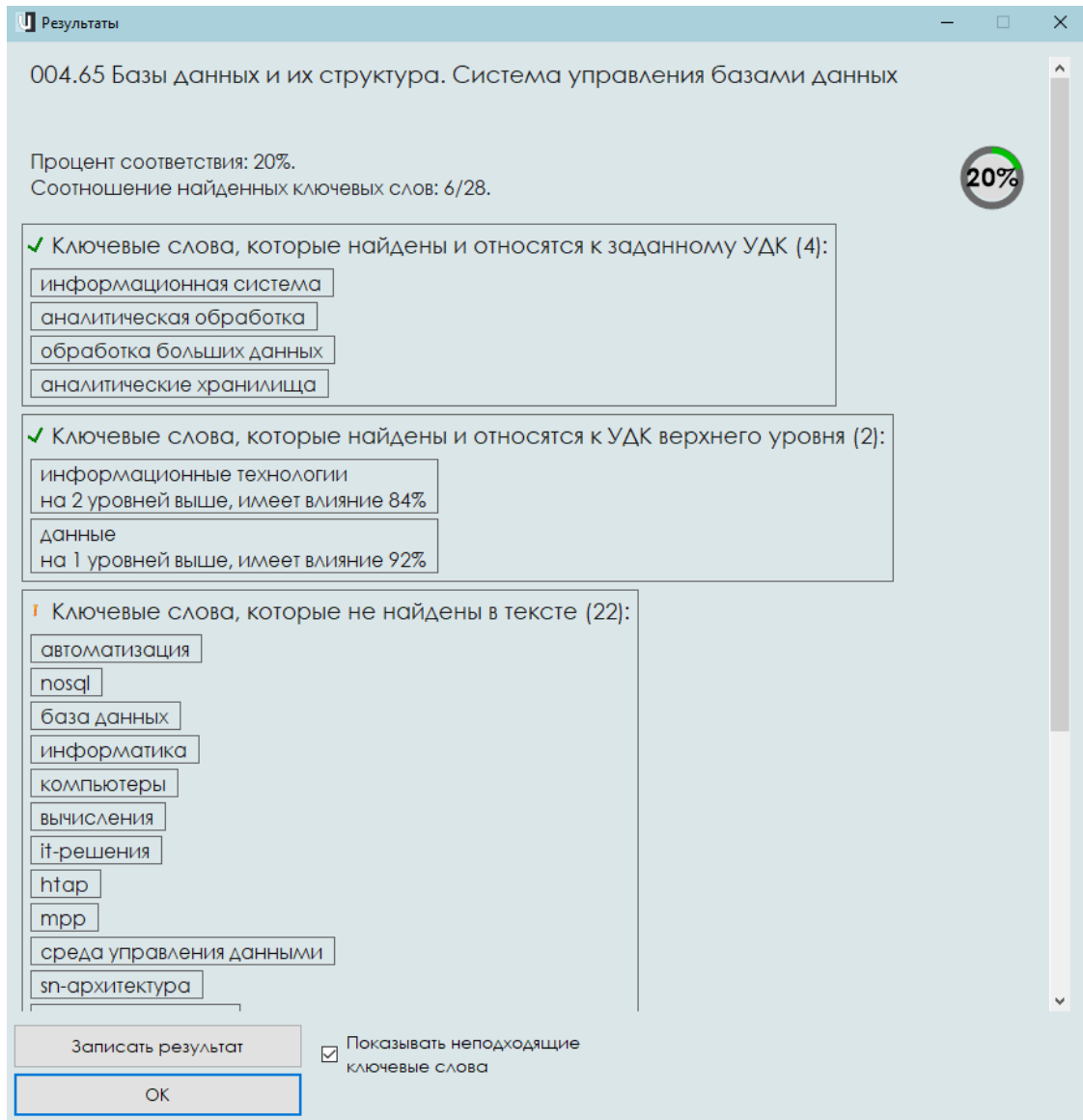


Рисунок В.4 – Результат тесту 3

Тестування для УДК «004.89 Прикладні системи штучного інтелекту. Інтелектуальні системи, засновані на використанні знань».

Тест 4.

Текст документу:

«Предметом этой книги является прикладной искусственный интеллект (ИИ), и здесь мы рассуждаем о практически реализуемых и полезных проектах, избегая пикировки в стиле пресловутого «голового петуха» Платона — Диогена [5]. Понятие «искусственный интеллект» созвучно понятию интеллекта естественного, но имеет с последним мало общего. Важно понять, что символический искусственный интеллект не является моделью человеческого мышления[6], так же как, например, резина не является химической моделью каучука, хотя обладает рядом схожих физических свойств. Действительно, программа с искусственным интеллектом использует некоторые алгоритмы, свойственные также и человеческому мышлению (в конце концов, любая программа написана людьми). Однако в реальной жизни человек будет решать ту же задачу иначе, чем это делает программа (производя другие операции, используя другие алгоритмы, задействуя структуры психики, пока что недостаточно изученные, а потому не поддающиеся переводу в математические формулы и программный код). Лобовой подход — изложение «человеческих» алгоритмов «машинным» языком — малоэффективен, как мы увидим в дальнейшем. Другими словами, если мы относим задачу к интеллектуальным, то любое ее решение, по нашему определению, свидетельствует о некотором интеллекте. Если решение добывается программным путем, то эта программа содержит элементы интеллекта.

Программа, умеющая решать задачи в предметной области, которую традиция относит к интеллектуальным, называется прикладной системой с элементами искусственного интеллекта.»

Результат тесту 4 наведено на рисунку В.5.

Результаты

004.89 Прикладные системы искусственного интеллекта.
Интеллектуальные системы, основанные на использовании знаний

Процент соответствия: 9%.
Соотношение найденных ключевых слов: 3/31.

9%

✓ Ключевые слова, которые найдены и относятся к заданному УДК (2):

- интеллектуальная система
- прикладные системы искусственного интеллекта

✓ Ключевые слова, которые найдены и относятся к УДК верхнего уровня (1):

- искусственный интеллект
на 1 уровень выше, имеет влияние 92%

⚠ Ключевые слова, которые не найдены в тексте (28):

- автоматизация
- информационные технологии
- информатика
- компьютеры
- вычисления
- it-решения
- дерево решений
- нейронная сеть
- генетический алгоритм
- обработка естественного языка
- построение семантических сетей
- модель знаний
- идентификация системы
- обучаемые связи
- нейроны
- нейронные слои
- программный комплекс
- семантическая разметка
- информационное сообщество
- знания

Записать результат

☒ Показывать неподходящие ключевые слова

OK

Рисунок В.5 – Результат тесту 4

Тест 5.

Текст анотації документу:

«В приложении описываются основные направления применения нейрокомпьютинга в финансовой деятельности. Под нейрокомпьютингом понимается использование как нейрокомпьютеров, так и нейросетевых алгоритмов в виде эмуляции нейронных сетей на персональных компьютерах. Приведен краткий перечень основных задач, где нейрокомпьютеры имеют эффективность, как отношение производительности к стоимости, значительно превышающую эффективность как традиционных методов регрессионного анализа, так и экспертных систем, основанных на построении формальной модели объекта или явления. Описываются принципы построения нейронных сетей и основные характеристики их функционирования. Приводится краткая характеристика состояния проблемы в мире. Перечень основных задач и характеристика проблемы в мире основаны на обзоре и анализе зарубежных публикаций, посвященных данной проблеме.

Для разработчиков, интересующихся конкретными направлениями применения нейрокомпьютеров в финансовой деятельности, в Научном центре нейрокомпьютеров имеется более подробный информационный материал.»

Результат тесту 5 наведено на рисунку В.6.

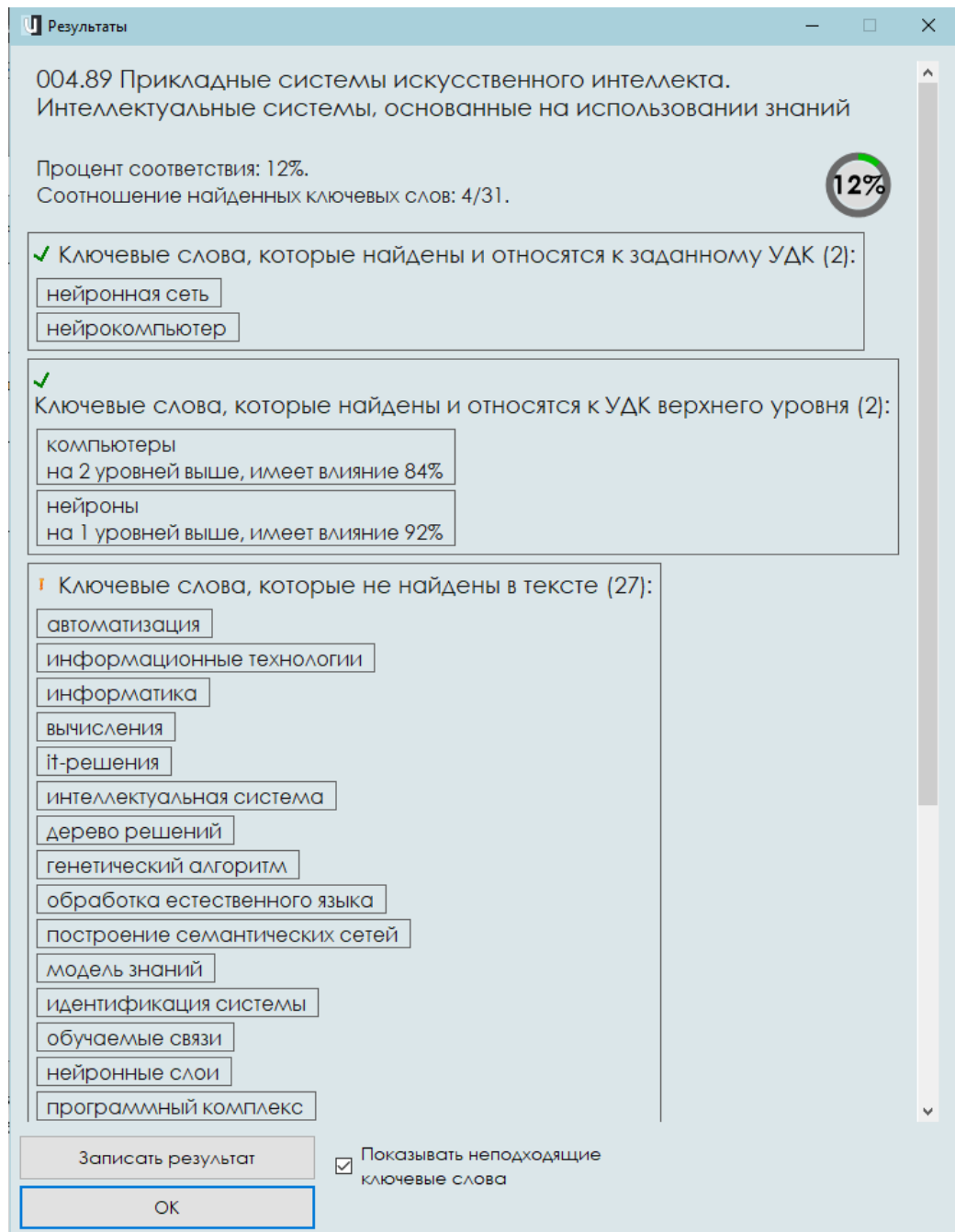


Рисунок В.6 – Результат тесту 5

Тест 6.

Текст документу:

«Как рождается сознание ? Исследователи подозревают, что ответ на этот вопрос лежит в связи между нейронами. Однако к сожалению, мало известно об нейронах головного мозга. Это связано также с проблемой

времени: отслеживание связей между нейронами в собранных данных потребует огромное количество человеко-часов. На сегодняшний день, ни один компьютер не смог идентифицировать нейронные контакты клеток достаточно хорошо. Ученые из института нейробиологии Макса Планка в Мартинсрид планируют изменить это с помощью искусственного интеллекта. Они обучили несколько искусственных нейронных сетей и тем самым позволили значительно ускорить анализ реальных нейронных цепей.»

Результат тесту 6 наведено на рисунку В.7.

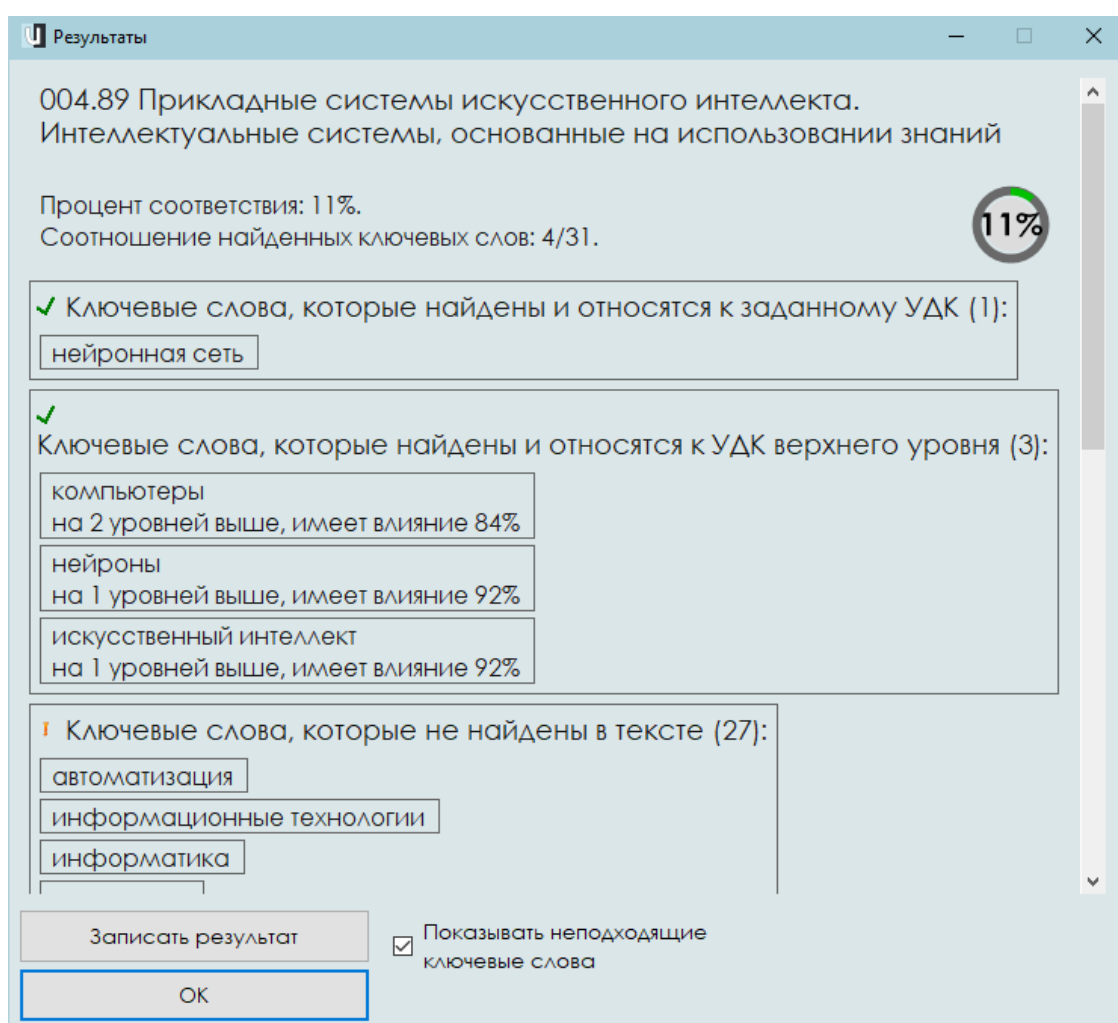


Рисунок В.7 – Результат тесту 6

Тестування для УДК «004.7 Зв'язок комп'ютерів. Мережі ЕОМ. обчислювальні мережі».

Тест 7.

Текст документу:

«На сегодняшний день процессы глобализации охватили все сферы жизни мирового сообщества. Глобальная вычислительная сеть Интернет является мощным средством развития глобальных процессов и ярчайшим проявлением социальной интеграции. Согласно данным ФОМ [1], полученным по итогам 2009 года, полугодовая аудитория Интернета в России составила 39,9 млн. человек или 35% населения старше 18 лет. Месячная аудитория Интернета составила 36,8 миллионов человек (32%), суточная – 21,3 млн. (18%). По сравнению с данными, опубликованными ранее RUметрикой [2], полугодовая аудитория выросла на 6,4%, месячная – на 5,4%, суточная – на 3,4%. С появлением новых сервисов Интернета возникают новые возможности для социальных коммуникаций, организации виртуальных социальных сред. Люди разделяются по интересам, принадлежности к учебному заведению, месту жительства и т.д. Однако, на этом фоне, в настоящее время в Интернете ярко проявляется феномен локализации пользователей [3, 4]. Если еще несколько лет назад пользователь, несмотря на малую пропускную способность каналов связи, за сеанс посещал более десятка различных сайтов, то в настоящее время, несмотря на значительный рост пропускной способности каналов связи, значительная часть пользователей ограничивается посещением только одного ресурса – социальной сети. Локализации пользователей также будет способствовать введение национальных алфавитов в доменное пространство Интернета. Дальнейшая стратификация пользователей может привести к тому, что через некоторое время в Интернете появятся слабосвязанные друг с другом сегменты [3, 4]. В результате произойдет вырождение глобальной сети Интернет в 11 (по числу введенных языков) региональных сетей. В связи с этим, на наш взгляд, будет возрастать

роль локальных вычислительных сетей как основных ячеек Интернета [5].
В настоящей работе будет проанализирована актуальность использования локальных вычислительных сетей, рассмотрены задачи, которые можно решать с помощью локальных вычислительных сетей в настоящее время.»

Результат тесту 7 наведено на рисунку В.8.

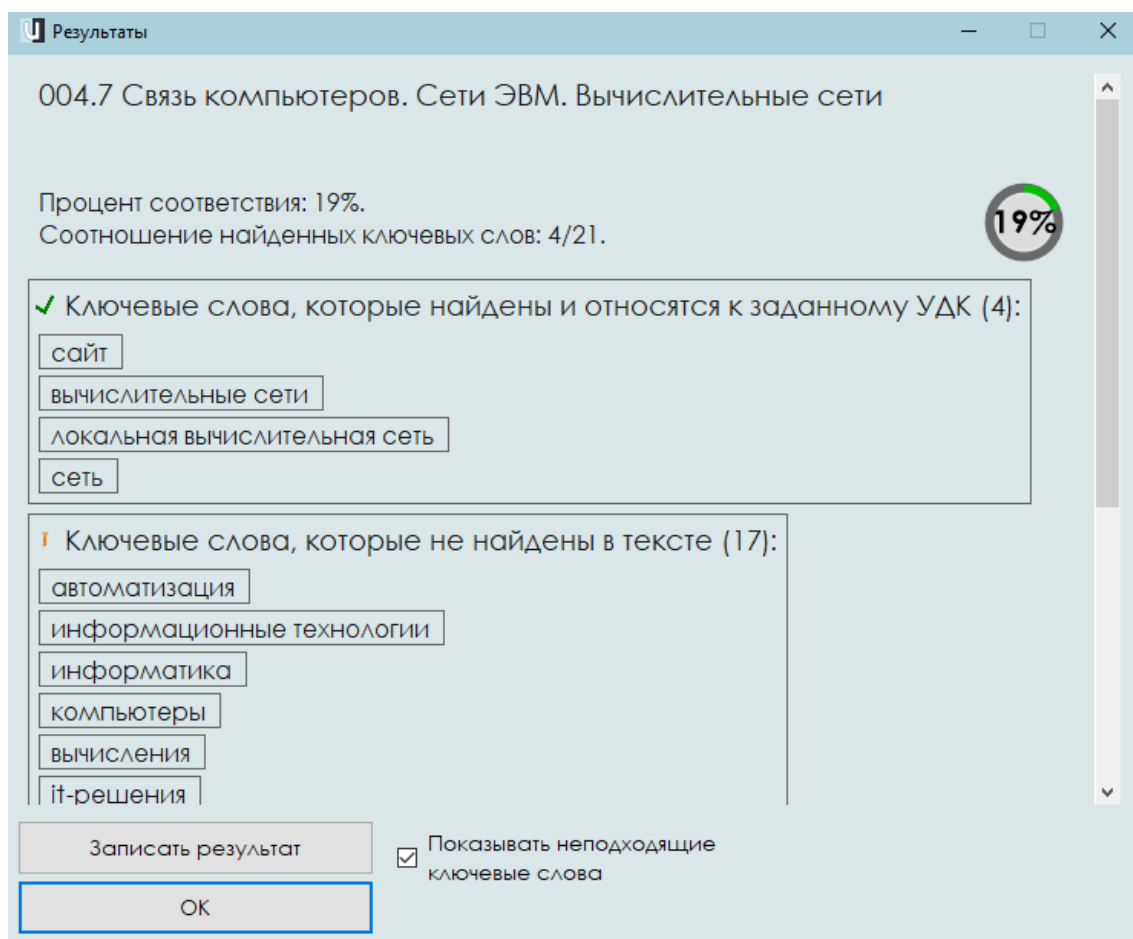


Рисунок В.8 – Результат тесту 7

Тест 8.

Тексту документу:

«Локальная вычислительная сеть (ЛВС или LAN) - это когда несколько компьютеров соединяются между собой проводами (или совсем недавно вошедший в жизнь вариант - без проводов, по радиоканалам), и

пользователь может обратиться к файлам на другом компьютере просто через "Мой компьютер".

Другой, не менее, а часто даже более используемой функцией локальных сетей является одновременная работа нескольких пользователей с одной программой. Например, при работе с системами "1С:Предприятие", локальная сеть позволит нескольким сотрудникам с разных компьютеров одновременно работать с одной базой. Так, в это же самое время, начальник может просматривать отчет по закупкам в текущем месяце, а бухгалтер рассчитывать зарплату сотрудникам.

Третьей значительной функцией локальных сетей является использование ограниченных ресурсов несколькими потребителями. Этот вариант даёт прямую экономию сразу же! К примеру, если в одной комнате работает три человека и всем нужно печатать документы, то установка локальной сети и одного принтера обойдётся дешевле, чем трех принтеров - по одному на каждый компьютер. А если компьютеров 4, то ещё дешевле... При этом, в качестве работы разницы не будет - работа с принтером подключенным по сети ничем не отличается от использования принтера подключенного к компьютеру. Другим примером служит интернет в офисе (а на сегодня, часто даже дома). Если у Вас есть компьютер, подключённый к интернету, то установка локальной сети позволит пользоваться интернетом всем компьютерам в сети.»

Результат тесту 8 наведено на рисунку В.9.

Результаты

004.7 Связь компьютеров. Сети ЭВМ. Вычислительные сети

Процент соответствия: 18%.
Соотношение найденных ключевых слов: 4/21.

18%

✓ Ключевые слова, которые найдены и относятся к заданному УДК (3):

вычислительные сети
локальная вычислительная сеть
сеть

✓ Ключевые слова, которые найдены и относятся к УДК верхнего уровня (1):

компьютеры
на 1 уровень выше, имеет влияние 92%

! Ключевые слова, которые не найдены в тексте (17):

автоматизация
информационные технологии
информатика
вычисления
it-решения
сайт

Записать результат

☒ Показывать неподходящие ключевые слова

ОК

Рисунок В.9 – Результат тесту 8

Тест 9.

Текст анотації документу:

«На сегодняшний день компьютерные сети переживают ряд проблем, возможные решения которых предлагает такая новая сетевая технология, как программно-конфигурируемых сети (ПКС). В данной работе будут рассмотрены основные недостатки современных компьютерных сетей, а также концепция, лежащая в основе ПКС, основные преимущества и слабые стороны данной сетевой технологии. В завершении будут приведены факты, иллюстрирующие тенденции будущего развития ПКС в России и мире.»

Результат тесту 9 наведено на рисунку В.10.

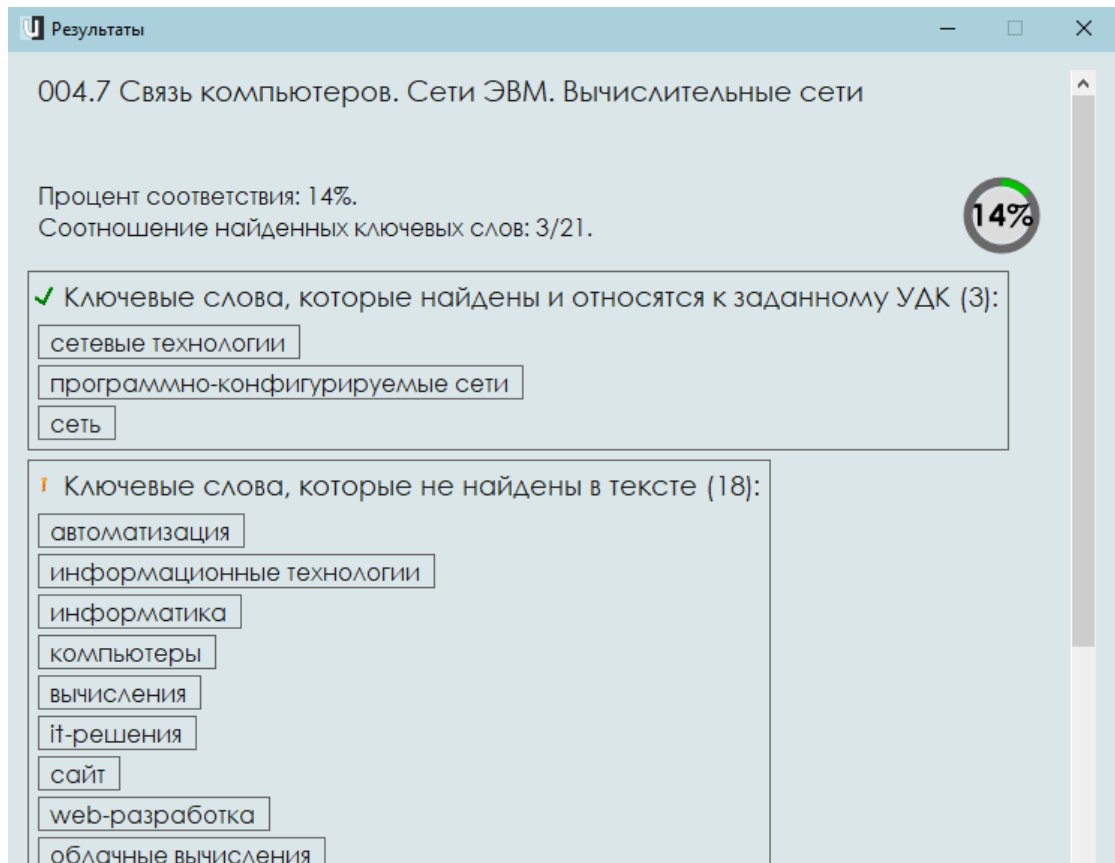


Рисунок В.10 – Результат тесту 9

Тестування для УДК «004.056.57 Захист від комп'ютерних вірусів».

Тест 10.

Текст анотації документу:

«Статья посвящена проблеме защиты компьютера от вирусов, проблеме информационной безопасности, так как проникающие в их сети вирусы могут нарушить целостность и сохранность информации. В статье рассматриваются разновидности компьютерных вирусов, признаки заражения, пути проникновения вирусов, методы защиты и типы антивирусных программ. Рекомендуются антивирусные программы, наиболее распространённые в странах СНГ.»

Результат тесту 10 наведено на рисунку В.11.

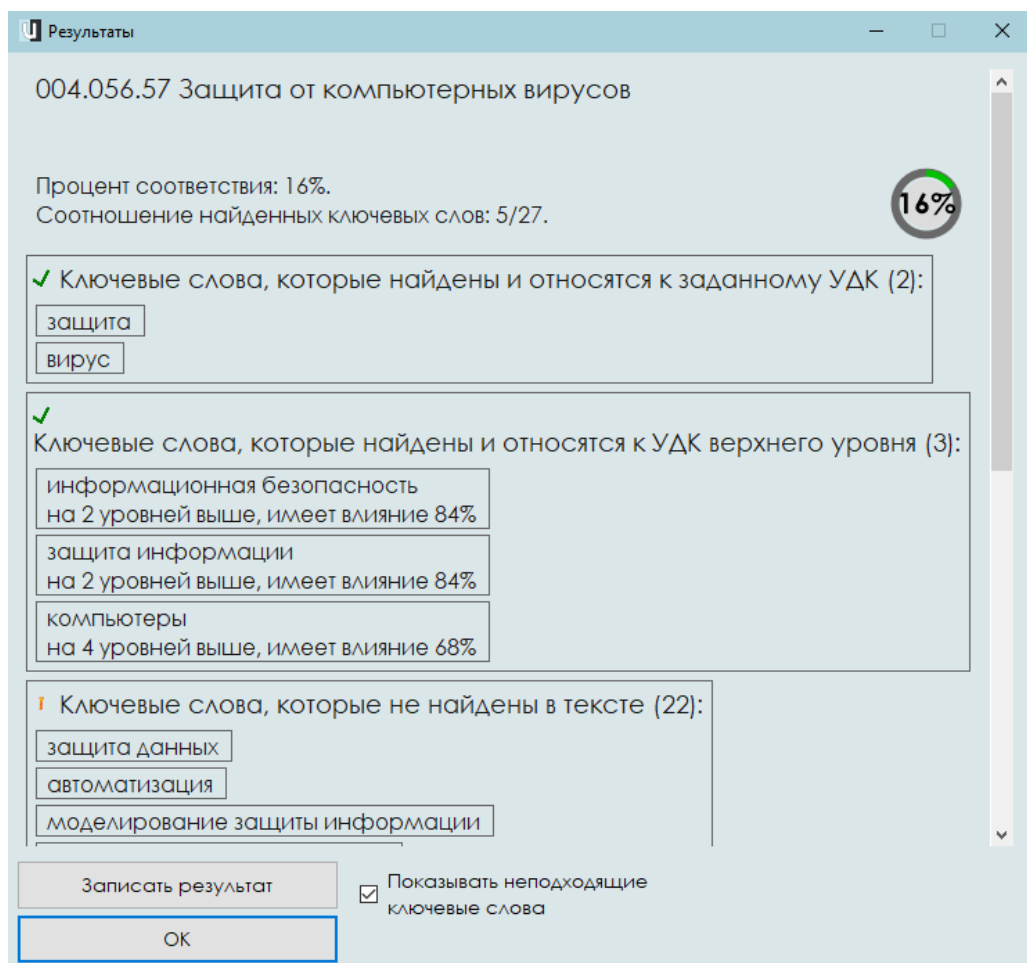


Рисунок В.11 – Результат тесту 10

Тест 11.

Текст анотації документу:

«Автор знакомит читателя с механизмами распространения компьютерных вирусов и воздействия вирусов на программное обеспечение компьютера. Рассматриваются средства противодействия вирусам, даются рекомендации с целью профилактики заражения компьютера. Технологическое образование.»

Результат тесту 11 наведено на рисунку В.12.

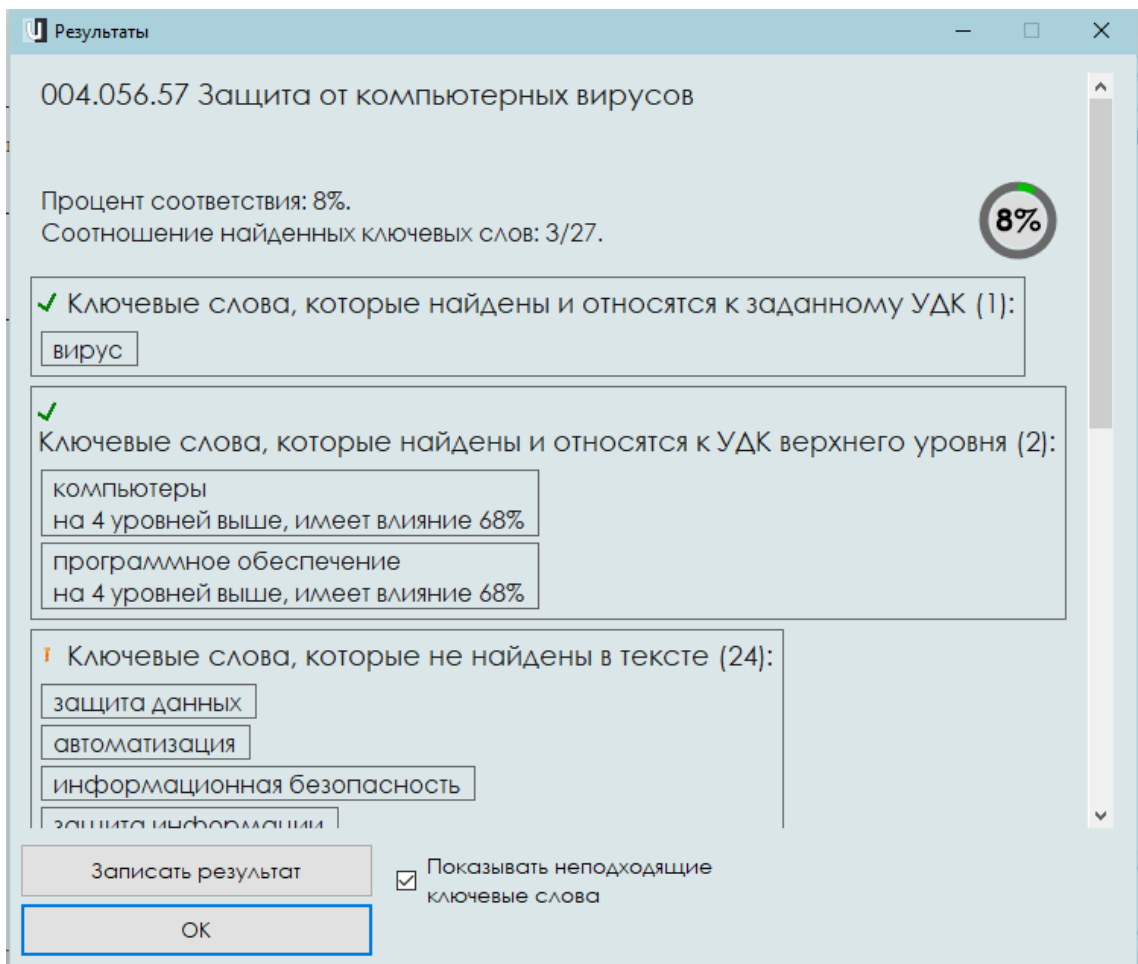


Рисунок В.12 – Результат тесту 11