

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ТЕРНОПІЛЬСЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
ІМЕНІ ІВАНА ПУЛЮЯ
ФАКУЛЬТЕТ КОМП'ЮТЕРНО-ІНФОРМАЦІЙНИХ СИСТЕМ І ПРОГРАМНОЇ
ІНЖЕНЕРІЇ
КАФЕДРА КОМП'ЮТЕРНИХ НАУК

БЕЗКОРОВАЙНИЙ ОЛЕГ РОМАНОВИЧ

УДК 004.9

**МЕТОДИКА ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ДЛЯ ПОШУКУ
СЛАБОФОРМАЛІЗОВАНИХ НАБОРІВ ДАНИХ**

122 «Комп'ютерні науки»

Автореферат

дипломної роботи на здобуття освітнього ступеня «магістр»

Тернопіль
2018

Роботу виконано на кафедрі комп'ютерних наук Тернопільського національного технічного університету імені Івана Пулюя Міністерства освіти і науки України

Керівник роботи: Кандидат економічних наук, доцент кафедри комп'ютерних наук
Матійчук Любомир Павлович,
Тернопільський національний технічний університет імені Івана Пулюя,

Рецензент: кандидат фізико-математичних наук,
доцент кафедри фізики
Крамар Олександр Іванович,
Тернопільський національний технічний університет імені Івана Пулюя,

Захист відбудеться 24 грудня 2018 р. о 9⁰⁰ годині на засіданні екзаменаційної комісії №30 у Тернопільському національному технічному університеті імені Івана Пулюя за адресою: 46001, м. Тернопіль, вул. Руська, 56, навчальний корпус №1, ауд. 702

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми роботи у зв'язку із великою популярністю web-аплікацій, тобто використанням мережевих ресурсів, зокрема поштових сервісів зростає необхідність у структуризації та формалізації даних

Мета роботи: розробка гнучкого програмного комплексу для обробки електронних листів та повідомлень, які не характеризуються конкретним адресатом

Об'єкт, методи та джерела дослідження.

Об'єктом дослідження є процеси обробки слабоформалізованої інформації та даних, що знаходяться у контенті поштових сервісів.

Предмет дослідження – методи та програмні засоби обробки слабоформалізованої інформації.

Теоретичні дослідження по розробці методів та програмного забезпечення обробки слабоформалізованої інформації ґрунтуються на застосуванні системного аналізу, методології функціонального моделювання, інженерії знань, семантичних мереж, теорії множин, теорії графів, теорії нечіткого виведення, теорії алгоритмів та об'єктно-орієнтованого проектування.

Наукова новизна отриманих результатів:

Запропоновано метод персоніфікації інтелектуальних корпоративних ресурсів компанії, який дозволяє досягнути більш ефективного виявлення і точності в порівнянні зі стандартними підходами за рахунок використання математичного апарату алгебри кінцевих предикатів.

Практичне значення отриманих результатів.

Запропоновані методи і програмні засоби забезпечили комплексне вирішення таких завдань як підвищення ефективності структурування інформації. Створений програмний продукт дозволяє будувати системи для обробки слабоформалізованої інформації.

Апробація. За результатами досліджень проведених в рамках магістерської роботи зроблена доповідь на VII міжнародній науково-технічній конференції молодих учених та студентів яка відбулась 28-29 листопада 2018 року у м. Тернополі на базі Тернопільського національного технічного університету імені Івана Пулюя.

Структура роботи. Робота складається з розрахунково-пояснювальної записки та графічної частини. Розрахунково-пояснювальна записка складається з вступу, 7 частин, висновків, переліку посилань та додатків. Обсяг роботи: розрахунково-пояснювальна записка – 107 арк. формату А4. графічна частина – 9 аркушів формату А1.

ОСНОВНИЙ ЗМІСТ РОБОТИ

У **вступі** проведено огляд сучасного стану програмного забезпечення для опрацювання неформалізованої, слабо формалізованої, не структуризованої та слабо структуризованої інформації та сформовано основні завдання, які необхідно вирішити.

В першому розділі розглянуто особливості роботи з погано формалізованою і слабо структурованою інформацією, основні проблеми інтелектуальної обробки інформації, здійснено аналіз моделей представлення знань в інтелектуальних інформаційних системах та сформовано задачі дослідження.

В другому розділі розглянуто модель вилучення глибинних знань для систем організаційного управління, метод персоніфікації інтелектуальних корпоративних ресурсів компанії.

В третьому розділі проведена розробка семантично-орієнтованого лінгвістичного процесора для добування нових знань з потоку документів корпоративної інформаційної системи та реалізація модулів системи.

В спеціальній частині здійснено проектування інтерфейсу веб-орієнтованої системи та тестування web-системи.

В розділі «Обґрунтування економічної ефективності» розраховано основні техніко-економічні показники проведених та доведено доцільність реалізації даного дослідження.

В частині «Охорона праці та безпека в надзвичайних ситуаціях» розглянуті аспекти охорони праці при використанні інформаційних технологій в освітньому середовищі та організація оповіщення і зв'язку у надзвичайних ситуаціях техногенного та природного характеру.

В частині «Екологія» досліджено роботу з банками екологічної інформації та основні джерела антропогенного забруднення навколишнього середовища.

У загальних висновках щодо дипломної роботи описано прийняті в проекті технічні рішення і організаційно-технічні заходи, які забезпечують виконання завдання на дипломне проектування; оригінальні технічні рішення, прийняті автором в процесі роботи; технічні рішення роботи, які можуть бути впроваджені на практиці;

В додатках до пояснювальної записки приведено ксерокопію тез доповідей, програмний код основних модулів системи.

В графічній частині подано тему, мету та завдання до дипломної роботи.

Діаграма варіантів використання головного меню. Алгоритм роботи системи. Діаграма діяльності функції «Про програму» та «Авторські права». Блок схему тестування системи. Автоматизований тест для працівника. Автоматизований тест для відношення слова-працівник. Автоматизований тест для доступу до даних. Висновки.

ВИСНОВКИ

В процесі виконання дипломної роботи в першому розділі:

- Проаналізовано поняття «слабкоформалізованої інформації» та її використання у мережі Інтернет.
- Проведено огляд та аналіз систем обробки інформації, здійснено порівняльну оцінку технологій на яких вони базуються.
- Здійснено порівняння моделей представлення знань та методів обробки інформації.

– Сформовано задачі дослідження для підвищення ефективності структурування слабкоформалізованої інформації.

В другому розділі дипломної роботи:

– Проведено теоретичні дослідження стосовно основних методів обробки слабкоформалізованої інформації.

– Розкрито основні підходи аналізу інформації, та можливості їх подальшого вдосконалення, а також розробки нових методів та засобів.

– Одним з ключових методів, який вибрано для аналізу та вдосконалення є метод персоніфікації інтелектуальних корпоративних ресурсів компанії.

В третьому розділі дипломної роботи:

– Проведено розробку семантично-орієнтованого лінгвістичного процесора для добування нових знань з потоку документів корпоративної інформаційної системи та реалізація модулів системи.

В четвертому розділі дипломної роботи проведено практичну реалізацію розроблених методів та алгоритмів, здійснено програмну реалізацію. Здійснено програмну реалізацію, проведено апробацію та тестування.

В п'ятому розділі дипломної роботи розраховано основні техніко-економічні показники проведених досліджень.

В шостому розділі дипломної роботи розглянуті аспекти охорони праці при використанні інформаційних технологій в освітньому середовищі та організація оповіщення і зв'язку у надзвичайних ситуаціях техногенного та природного характеру.

В сьомому розділі дипломної роботи досліджено роботу з банками екологічної інформації та основні джерела антропогенного забруднення навколишнього середовища.

СПИСОК ОПУБЛІКОВАНИХ АВТОРОМ ПРАЦЬ ЗА ТЕМОЮ РОБОТИ

АВТОМАТИЗАЦІЯ ПРОЦЕСІВ МЕНЕДЖМЕНТУ ТА ІНТЕЛЕКТУАЛЬНОЇ ОБРОБКИ СЛАБКОФОРМАЛІЗОВАНОЇ ІНФОРМАЦІЇ / О. Р. Безкоровайний, Т. В. Сергієнко. // Тези доповіді на VII міжнародній науково-технічній конференції молодих учених та студентів яка відбулась 28-29 листопада 2018 року у м. Тернополі на базі Тернопільського національного технічного університету імені Івана Пулюя. Том 2. – 2018. – С. 10-11.

АНОТАЦІЯ

Дипломна робота присвячена розробці гнучкого програмного комплексу для обробки електронних листів та повідомлень, які не характеризуються конкретним адресатом.

В першому розділі розглянуто особливості роботи з погано формалізованою і слабо структурованою інформацією, основні проблеми інтелектуальної обробки

інформації, здійснено аналіз моделей представлення знань в інтелектуальних інформаційних системах та сформовано задачі дослідження.

В другому розділі розглянуто модель вилучення глибинних знань для систем організаційного управління, метод персоніфікації інтелектуальних корпоративних ресурсів компанії.

В третьому розділі проведена розробка семантично-орієнтованого лінгвістичного процесора для добування нових знань з потоку документів корпоративної інформаційної системи та реалізація модулів системи.

В четвертому розділі дипломної роботи здійснено проектування інтерфейсу веб-орієнтованої системи та тестування web-системи.

В повному обсязі виконано розділи «Обґрунтування економічної ефективності», «Охорона праці та безпека в надзвичайних ситуаціях» та «Екологія».

Об'єктом дослідження є процеси обробки слабоформалізованої інформації та даних, що знаходяться у контенті поштових сервісів.

Предмет дослідження – методи та програмні засоби обробки слабоформалізованої інформації.

Мета роботи: розробка гнучкого програмного комплексу для обробки електронних листів та повідомлень, які не характеризуються конкретним адресатом.

Ключові слова: слабоформалізована інформація, онтологія, корпоративна інформаційна система.

ANNOTATION

The thesis is devoted to the development of a flexible software package for processing emails and messages that are not specific to the addressee.

The first section deals with the peculiarities of working with poorly formalized and poorly structured information, the main problems of intellectual processing of information, an analysis of the models of knowledge representation in intelligent information systems and the research tasks.

In the second section the model of the extraction of deep knowledge for organizational management systems, the method of personalization of intellectual corporate resources of the company is considered.

In the third section, the development of a semantically oriented linguistic processor for obtaining new knowledge on the flow of corporate information system documents and the implementation of system modules.

In the fourth section of the thesis, the design of the interface of a web-based system and testing of the web-system.

The sections "Justification of economic efficiency", "Labor protection and safety in emergencies" and "Ecology" are executed in full.

The object of the study is the processing of poorly formalized information and data contained in the content of postal services.

Subject of research - methods and software for the processing of poorly formalized information.

Purpose: development of a flexible software package for handling emails and messages that are not specific to the addressee.

Key words: weakly formalized information, ontology, corporate information system.