

АЛГОРИТМ ІНТЕЛЕКТУАЛЬНОГО ПОШУКУ ЗНАЬ ЯК МЕХАНІЗМ РЕАЛІЗАЦІЇ ІНФОРМАЦІЙНОЇ БЕЗПЕКИ

Одним із засобів реалізації інформаційної безпеки підприємства чи організації є розробка документів різноманітних рівнів призначення, орієнтовані на кожного співробітника [1]. Для цього необхідно суміщати технічне завдання з юридичними вимогами та рекомендаціями, що будуть зрозумілі для кожного користувача системи безпеки. Автоматична обробка таких електронних текстових документів на змістовному рівні передбачає застосування методів лінгвістичного аналізу, основних положень теорії тексту та сучасних інформаційних технологій [2].

Тому метою даного дослідження є розкриття особливостей автоматичної змістовної обробки електронних текстових документів на основі використання формальних моделей представлення знань.

Наукова новизна даного дослідження полягає у розробці алгоритму інтелектуального пошуку та вилучення знань з електронних текстових документів шляхом використання математичного апарату логіки предикатів [3].

Алгоритм інтелектуального пошуку та вилучення знань з електронних текстових документів, на основі якого формується формальна логіко-лінгвістична модель тексту, включає в себе такі кроки:

1. Аналіз та фіксація кількості складних синтаксичних частин, абзаців та речень в електронному текстовому документі.

2. Аналіз кожного, отриманого на попередньому кроці алгоритму, речення.

3. Визначення множини відношень, суб'єктів, об'єктів, предметів відношень та їх характеристик шляхом застосування бази даних, яка представляє собою набір реляційних таблиць,

кожна з яких відповідає за певну частину мови та відповідні їй граматичні характеристики.

4. Застосування основних принципів формування логіко-лінгвістичних моделей речень природної мови для створення логічних виразів, простих предикатів, кожен з яких інтерпретує структуру простого речення природної мови.

5. Створення логіко-лінгвістичних моделей речень природної мови на основі використання бази знань, що представляє собою продукційну модель з формалізованими правилами формування словосполучень та складних речень природної мови [4].

6. Застосування основних принципів синтезу логіко-лінгвістичних моделей для встановлення логічних зв'язків між логіко-лінгвістичними моделями речень природної мови, що сформовані в п.5 цього алгоритму.

7. Визначення характеристик кожного абзацу електронного текстового документу шляхом застосування абстрактних моделей, що інтерпретують можливі види тематичних та рематичних зв'язків у текстах.

8. Побудова логіко-лінгвістичної моделі тексту на основі принципів, сформульованих у вигляді тверджень [4].

Таким чином, запропонований алгоритм дозволяє автоматично формалізувати електронний текстовий документ, отримуючи в результаті його аналізу логіко-лінгвістичну модель, компоненти якої формуються на основі вилучених із текстової інформації знань.

Література

1. Родичев Ю. Нормативная база и стандарты в области информационной безопасности. Питер, 2017. 256 с.

2. Широков В. А. Феноменология лексикографических систем. Київ: Наукова думка, 2004. 326 с.

3. Вавиленкова А. И. Основные принципы синтеза логико-лингвистических моделей. Кибернетика и системный анализ. 2015. Т. 51, № 5. С. 176 – 185.

4. Вавиленкова А. І. Критерії аналізу логіко-лінгвістичних моделей речень природної мови. Вісник Національного технічного університету «Харківський політехнічний університет». Харків, 2017. № 7 (1229). С. 118 – 122.