

ВІДГУК

офіційного опонента на дисертаційну роботу

ШАБАЛИ Євгенії Євгенівни

«Інформаційна технологія ідентифікації змісту будівельних нормативів»,
представлену на здобуття наукового ступеня кандидата технічних наук
за спеціальністю 05.13.06 – інформаційні технології

В будівельній галузі для управління використовуються великі обсяги інформації. Співробітники у багатьох випадках уже не в змозі самостійно опрацювати всю необхідну інформацію, навіть з використанням автоматизованих систем управління. У зв'язку з цим широкого застосування набули комп'ютерні системи, що використовують методи автоматичного опрацювання природної мови. Створення великих сховищ даних робить актуальними задачі екстракції знань і формування коректно побудованих текстових документів, що вимагає створення засобів повнотекстового пошуку інформації, а також порівняння текстів за їх змістом. Одним із найкращих розв'язків цих задач є застосування інтроформаційного методу, який забезпечує розрахунок несилового впливу фрагментів речень на вибір ключових понять.

Більшість існуючих моделей або не дозволяють автоматизувати процедури класифікації, або дають високий рівень похибок і низьку точність при їх практичній реалізації. Це пов'язано з тим, що до теперішнього часу розробка систем автоматичної обробки текстів природною мовою відбувалася без використання смислового аналізу або з його мінімальним використанням. Таким чином виникла потреба у створенні інформаційної технології ідентифікації змісту текстів будівельної тематики, яка дозволить обробляти природномовні тексти будівельних нормативів.

У зв'язку з вище зазначеним, тематика дисертаційних досліджень, що включає розробку рефлексорних моделей і методів лінгвістичного аналізу змісту текстів будівельних нормативів, а також використання ідей штучного інтелекту для удосконалення ведення нормативної документації є актуальною.

Дисертаційна робота Є.Є. Шабали присвячена питанням обробки природомовних текстів та ідентифікації змісту текстів будівельних нормативів за рахунок розробки та впровадження нових моделей, методів та засобів лінгвістичної обробки текстів. Очевидно вирішення важливої науково-технічної задачі, якій присвячена дисертаційна робота Є.Є. Шабали, має безумовну актуальність і значимість. Підтвердженням цьому є те, що робота виконувалась в рамках досліджень науково-дослідної тематики «Створення системи перевірки ступеню унікальності наукових робіт» (номер державної реєстрації 0114U000126).

Основний зміст роботи. Робота складається зі вступу, чотирьох розділів, загальних висновків, списку використаних літературних джерел і трьох додатків. Повний обсяг дисертації складає 170 сторінок, 22 рисунка, 7 таблиць, список використаних джерел зі 130 найменувань.

У Вступі розглянуто актуальність теми, визначено мету та задачі дослідження, підкреслено наукову новизну та практичне значення роботи, розкрито апробація отриманих результатів та їх впровадження.

У першому розділі наведено основні теоретичні відомості, які стосуються особливостей побудови і архітектури інформаційних технологій ідентифікації змісту будівельних нормативів, лінгвістичного аналізу, розробки та ведення нормативної бази, аналіз структури існуючої нормативної бази в будівництві. Також виявлені основні принципи розвитку нормативного забезпечення будівництва та проведений аналіз існуючих підходів до розробки інтелектуальних систем. Виявлено, що побудова інформаційної технології ідентифікації змісту є затребуваною в сфері будівництва, є одною зі складових успішного управління нормативною базою та підготовки процесів будівництва.

Визначено, що побудова та впровадження інформаційної технології ідентифікації змісту будівельних нормативів є актуальним напрямком що може підвищити точність та якість визначення змісту нормативних документів, і як наслідок, скорити час на реферування природномовних текстів.

У *другому розділі* розглядається розробка науково-методологічного та понятійного базису управління процесом лінгвістичного аналізу текстів будівельних нормативів.

Показано, що основну увагу при проектуванні інформаційної технології ідентифікації змісту будівельних нормативів слід приділяти її здатності інтелектуальної обробки даних, а саме, можливості навчання інформаційної технології на основі досвіду її використання і можливості роботи із знаннями на основі формалізації тексту.

Розглянуто поняття інтроформації як інтелектуальної організації людини що формує семантичну складову текстів і є джерелом його прояву в будівельних нормативах та запропоновано використання інтроформаційного методу для визначення величини несилового впливу на процес визначення змісту. Також розроблено структуру інформаційного простору, який необхідно створити для ідентифікації змісту тексту будівельних нормативів та розроблено архітектуру системи інформаційного забезпечення процесу ідентифікації тексту будівельних нормативів.

У *третьому розділі* розроблено застосування інтроформаційного методу, який дозволяє реалізувати в штучних інтелектуальних системах рефлекторні принципи визначення змісту тексту будівельної тематики. Впевненість в правильності визначення змісту будівельних нормативів базується на статистичних даних, таких як об'єм тексту (кількість речень), які використовуються для розрахунку впливу на зміст через інтроформаційні міри – визначеність та інформованість. Для оцінки впливу фрагментів текстів будівельних нормативів на правильність визначення ключових понять в роботі використано інтроформаційний метод. Запропонована модель та метод дозволяють не тільки оцінити спільну умовну імовірність появи того чи іншого змісту в тексті будівельних нормативів, але й оцінити правильність визначення такої імовірності

В *четвертому розділі* розглянуто розробка засобів інформаційної технології ідентифікації змісту текстів будівельних нормативів. Обробка

експериментальних даних (результатів) інформаційної технології, розробка бази даних речень та ключових понять дозволяє проводити навчання системи на основі визначених ключових понять експертом в галузі будівництва. В даному розділі була сформульована мета впровадження інформаційної технології ідентифікації змісту будівельних нормативів, основні принципи побудови інформаційної технології, функціональні та нефункціональні вимоги до розробки та наведені критерії якості функціонування системи. Розроблено алгоритм визначення ключових понять тексту будівельних нормативів, наведений опис вхідних даних та принципи введення їх в систему та шляхи досягнення високих результатів роботи інформаційної технології. Описано структуру бази даних, визначено сутності, атрибути сутностей, їх ідентифікатори, розроблено логічну та фізичну моделі інформаційної технології ідентифікації змісту будівельних нормативів, наведені компоненти бази даних інформаційної технології ідентифікації змісту будівельних нормативів та зв'язки між ними.

В розділі наведено приклад ідентифікації змісту текстів будівельних нормативів розробленими засобами та виконано порівняння з іншими програмними засобами (TextAnalys), які показали менш ефективний результат, ніж розроблена інформаційна технологія.

Наукова новизна роботи обумовлена тим, що вперше вирішена науково-технічна задача розробки рефлексорних моделей і методів лінгвістичного аналізу змісту текстів будівельних нормативів.

У дисертаційній роботі автором отримані такі результати:

- вперше для забезпечення розрахунку несилового впливу фрагментів речень на вибір ключових понять у нормативній базі в будівельній галузі запропоновано застосування інтроформаційного методу, що дозволяє формально підійти до оцінки впливу на визначення змісту нормативних документів;
- вперше описано процес визначення впевненості в правильності

інтроформаційного змісту речення з нормативної бази, який пропонується реалізувати через оцінку сумісних умовних ймовірностей появи фрагментів тексту по частковим умовним ймовірностям, що дозволить створювати інформаційні технології ідентифікації змісту будівельних нормативів не на основі евристики, а використовуючи математичний апарат теорій імовірності та несилової взаємодії;

- вперше запропоновано визначення ключових понять тексту на основі рефлекторно-ймовірнісного підходу, що характеризується описом поведінки системи на основі рефлексів та дозволяє якісно обробляти нормативну документацію та скоротити час на пошук необхідних підрозділів будівельних нормативів;
- удосконалено міру впливу та методи оперування числовими значеннями впливів словосполучень речень на ідентифікацію змісту природномовних текстів через введення імовірнісних оцінок впливу;
- отримала подальший розвиток математична модель оперування мірою впливу природномовних текстів на зміст будівельних нормативів, що відрізняється від існуючих орієнтацією не тільки на виділення впливових слів, але і на виділення впливових фрагментів речень, що дозволяє застосувати інтроформаційний метод до вирішення задачі ідентифікації змісту будівельних нормативів.

Практичне значення одержаних результатів полягає в тому, що реалізована на їх основі інформаційна технологія дозволяє:

- оптимізувати процес ведення нормативної документації, структурувати базу будівельних нормативів;
- значно зменшити час на пошук необхідної інформації в будівельних нормативах;
- підвищити ефективність процесу обробки нормативної документації.

Отримані в роботі методи та моделі ідентифікації змісту будівельних нормативів можуть бути впроваджені в роботу проектно-будівельних підприємств.

Основні результати дисертаційної роботи достатньо повно викладені у 5 статтях, опублікованих у фахових наукових журналах та збірниках, 1 – у закордонному виданні та 6 в тезах доповідей науково-технічних конференцій.

Результати роботи є науково обґрунтовані. Проведені дослідження базуються на використанні теорії систем, теорії несилового взаємодії, методів системного аналізу, теорії моделювання, об'єктно-орієнтованого програмування і теорії проектування та реалізації автоматизованих інформаційних систем.

Зауваження до дисертації.

1. Підрозділ 1.1 дисертації присвячений сучасній системі ведення нормативної бази ресурсів в будівництві, однак він не має посилань на сучасні нормативні документи в цій галузі та, на мій погляд, замалий за обсягом – майже 2 сторінки. Аналогічне зауваження до підрозділу 2.1 – його обсяг 1,5 сторінки.

2. В першому розділі автором проведений детальний аналіз сучасних програмних засобів ведення нормативної бази в будівництві, але недостатньо розкриті результати аналізу моделей та методів, які використовуються сьогодні в інформаційних технологіях управління нормативною базою будівництва.

3. В другому розділі наведена така інформація, як структура будівельних державних норм (рис.2.1), підходи до створення інтелектуальних системи (рис. 2.5), яка стосується до аналізу і повинна бути в першому розділі.

4. Підрозділ 2.3 має назву «Інтроформаційний підхід до побудови інформаційної технології ідентифікації змісту будівельних нормативів», але визначення назви нового підходу автор в цьому підрозділі не надає, нова термінологія з'являється тільки в наступному підрозділі, що не є логічними.

5. На рис. 4.5 припущена помилка – в блок-схемі два блоки символізують кінець алгоритму.

6 Список використаних джерел на 56% складає літературу, випущену до 200 року, що є застарілими джерелами (є навіть публікації за 1945р.).

Оцінюючи роботу у цілому хотілось би відмітити наступне:

Роботу виконано за актуальною темою. Отримано нові наукові і практичні результати, які вирішують важливу науково-технічну задачу розробки моделей, методів та засобів обробки природомовних текстів та ідентифікації змісту текстів будівельних нормативів. Дисертація Є.Є.Шабали є завершеною науковою працею. Автореферат та опубліковані роботи в повній мірі відображають отримані результати та зміст дисертації, яка відповідає паспорту спеціальності 05.13.06 – інформаційні технології.


Робота «Інформаційна технологія ідентифікації змісту будівельних нормативів» відповідає вимогам пп.11, 12, 13 «Порядку присудження наукових ступенів і присвоєння вченого звання старшого наукового співробітника», затвердженого постановою Кабінету Міністрів України від 24.07.2013 року №567, які висуваються до кандидатських дисертацій. Таким чином, не зважаючи на наведені зауваження, вважаю, що Шабала Євгенія Євгенівна, заслуговує присудження наукового ступеня кандидата технічних наук за спеціальністю 05.13.06 – інформаційні технології.

Офіційний опонент, кандидат технічних наук, доцент, в.о. завідувача кафедрою бізнес адміністрування та управління проектами Університету економіки та права «КРОК»

О.Б. Данченко

Вчений секретар Вченої ради Університету економіки та права «КРОК» в.о. доцент

В.М. Лободаєв

Підписано:  Лободаєв В.М. засвідчую:
Директор департаменту та Вищого навчального закладу
Університету економіки та права "Крок"
Шабала Є.Є.
15.09.2015 р.