

«Комп'ютерна технологія порівняльного аналізу електронних текстів»

Основні наукові результати

Запропоновано уніфіковану форму логіко-лінгвістичної моделі текстової інформації, що здатна відображати синтаксичну структуру речень природної мови.

Побудовано алгоритм виявлення збігів в одномовних електронних текстах без врахування синонімів, розбіжностей синтаксичних структур текстів та порядків чергування слів у реченнях.

Запропоновано систему продукцій для встановлення характеристик структурних одиниць тексту.

Розроблено метод автоматичного анутовування електронних текстів, що базується на використанні сформованої системи продукцій для встановлення характеристик структурних одиниць тексту.

Розроблено проект комплексу технічних засобів спеціалізованого багатомовного пошукового серверу, застосування якого дозволить здійснювати порівняльний аналіз електронних документів, що знаходяться в мережі Інтернет, за змістом.

Сформульовано основні положення технології порівняльного аналізу електронних текстів.

Запропоновано тести для аналізу коректності розроблених алгоритмів порівняльного аналізу електронних текстів.

Наукова новизна отриманих результатів звітної етапу полягає в наступному:

- вперше запропонована уніфікована форма логіко-лінгвістичної моделі текстової інформації, що здатна відображати синтаксичну структуру речень природної мови;

- вперше розроблено проект комплексу технічних засобів спеціалізованого багатомовного пошукового серверу, застосування якого дозволить здійснювати порівняльний аналіз електронних документів, що знаходяться в мережі Інтернет, за змістом.

- одержав подальший розвиток метод порівняльного аналізу електронних текстів, що дозволяє враховувати синоніми, розбіжності синтаксичних структур текстів та порядки чергування слів у реченнях;

- удосконалена робота системи порівняльного аналізу електронних текстів за рахунок використання логіко-лінгвістичних моделей, як бази знань створеної аналітичної системи.

Практична цінність

Комп'ютерну систему порівняльного аналізу електронних текстів передбачається впровадити у:

- навчальних закладах України – для експертизи студентських письмових робіт (рефератів, домашніх завдань, контрольних робіт, курсових та дипломних проектів і робіт);

- у спеціалізованих радах з захисту дисертацій – для експертизи кандидатських і докторських дисертацій та їх авторефератів.

Впровадження системи в навчальних закладах України змусить студентів самостійно виконувати навчальні роботи, сприятиме підвищенню якості навчання і створенню відповідного іміджу навчальних закладів України серед провідних університетів ближнього та дальнього зарубіжжя, а також стане ефективним засобом захисту інтелектуальної власності. Крім того, система дозволить: автоматизувати функції обліку, збереження і архівації електронних форм письмових робіт, що виконуються студентами відповідно навчального плану протягом усього періоду навчання; надавати адміністрації навчальних закладів можливість оперативного контролю здачі письмових робіт, а також перегляду результатів їхньої перевірки в різних інформаційних розрізах:

- окремо взятого студента, групи, факультету;
- окремої дисципліни;
- виду письмової роботи;
- викладача, кафедри і т.д.

Впровадження комп'ютерної системи порівняльного аналізу електронних текстів у спеціалізованих радах з захисту дисертацій поставить заслін непорядним пошукувачам наукових ступенів, сприятиме підвищенню ефективності та результативності наукових досліджень, до-

зволить створити національній депозитарій кандидатських і докторських дисертацій, тобто електронну форму державного інтелектуального ресурсу України.

Практична цінність отриманих результатів полягає в тому, що вони дозволять у подальшому здійснити програмну реалізацію підсистеми попередньої обробки електронних документів, провести її тестування і розробити необхідну супроводжувальну документацію.

Інноваційні аспекти і переваги запропонованих методів та алгоритмів попередньої обробки електронних текстів полягають в:

- можливості автоматичного аналізу текстових документів;
- можливості визначати стилістику і авторство тексту;
- незалежності результатів обробки текстів від мови документу (в рамках обмежень: використання в іншій мові кирилиці або латинського алфавіту; наявність комп'ютерної програми-перекладача на українську мову; наявність відкритого для структуризації електронного синонімічного словника мови);
- високій релевантності результатів пошуку схожих текстових документів;
- скороченні тимчасових витрат для аналізу текстової інформації.

Перелік основних наукових публікацій, доповідей на конференціях, семінарах

1. Вавіленкова А.І. Метод автоматизованого формування логіко-лінгвістичних моделей текстової інформації як об'єднуюча ланка інформаційних технологій та лінгвістики на шляху розуміння комп'ютером природної мови /А.І. Вавіленкова // Прикладна лінгвістика та лінгвістичні технології MegaLing'2009: зб. наук. праць. – К.: Довіра, 2009. – С. 100–103.

2. Вавіленкова А.І. Розробка експертної системи на базі методу автоматизованого формування логіко-лінгвістичної моделі текстової інформації /А.І. Вавіленкова // Збірник тез II Міжнар. наук.-техн. конф. «Комп'ютерні системи та мережні технології» (CSNT–2009). – К.: НАУ, 2009. – С. 22.

3. Вавіленкова А.І. Логіко-лінгвістична модель як засіб відображення синтаксичних особливостей текстової інформації /А.І. Вавіленкова // Математичні машини та системи. – 2010. – № 2. – С. 134–137.

4. Литвиненко О.Є., Лезвінський О.С. Застосування програми конвертування текстових документів в комп'ютеризованій системі порівняльного аналізу електронних текстів. – Наукоємні технології. – 2009. – №2. – С. 56-58.

5. Литвиненко О.Є., Бурко Д.А. Інженерно-лінгвістичні принципи аналізу текстів. – Наукоємні технології. – 2009. – №3. – С. 60-63.

6. Литвиненко О.Є., Андрущенко А.А. Лінгвістичний аналізатор текстів на запозичення з використанням словників. – Наукоємні технології. – 2009. – №3. – С. 63-65.

7. Литвиненко О.Є., Бурко Д.А. Моделі семантичного аналізу текстів. – Наукоємні технології. – 2009. – №4. – С. 55-58.

8. Вавіленкова А.І. Система автоматизованого формування логіко-лінгвістичних моделей текстової інформації. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2009 – С. 14.

9. Литвиненко О.Є. Підсистема порівняльного аналізу електронних текстів як засіб підвищення якості підготовки фахівців. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2009 – С. 26.

10. Шепелева Т.В., Краліна Г.С. Дослідження і розробка ефективних алгоритмів токенизації. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2009 – С. 27.

11. Лупандін М.В., Безкровна І.А., Юр'єва І.М. Розробка алгоритмів семантичної структуризації даних. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2009 – С. 28.

12. Дишлюк О.М., Нечипорук О.П. Розробка методів та алгоритмів індексації електронних текстів. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2009 – С.29.

13. Вавіленкова А.І. Аналіз онтології як моделі представлення знань текстової інформації.

ції. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2010. – С. 16.

14. Масловський Б.Г., Тупота Є.В. Методи порівняльного аналізу електронних текстів. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2010. – С. 18.

15. Дишлюк О.М., Вавіленкова А.І. Метод автоматизованого формування логіко-лінгвістичних моделей електронних текстів. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2010. – С. 19.

16. Шепелева Т.В., Краліна Г.С. Основні підходи при пошуку збігів та протиріч в одномовних електронних текстах. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2010. – С. 20.

17. Юр'єва І.М., Сич М.Ю. Постановка задачі виявлення збігів в одномовних електронних текстах у випадку розбіжностей синтаксичних структур. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2010. – С. 21.

18. Артамонов Є.Б. Концепція візуально-лінгвістичного пошуку в електронних бібліотеках / Є.Б. Артамонов, А.В. Кузнецов. Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». - К.: НАУ, 2010. – С. 22.

19. Нечипорук Е.П. Программная система автоматического формирования текста. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2011. – С. 19.

20. Вавіленкова А.І. Семантичний аналіз як необхідний етап у вилученні змісту з текстової інформації. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2011. – С. 13. 15.

21. Сич М.Ю., Русаков О.І. Алгоритм анутовання текстів довільної структури. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2011. – С.

22. Юр'єва І.М. Анутовання як метод аналітичної обробки текстових документів. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2011. – С. 22.

23. Масловський Б.Г. Роль ключових слів у процесі пошуку інформації. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2011. – С. 23.

24. Дишлюк О.М. Основні механізми пошукових систем. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2011. – С.

25. Нечипорук О.П. Багатомовний пошуковий сервер. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2011. – С.

26. Краліна Г.С. Пошукові механізми в мережі Інтернет. Тези доповідей Міжнародної науково-технічної конференції «Інтелектуальні технології лінгвістичного аналізу». – К.: НАУ, 2011. – С.

27. Додонов А.Г., Литвиненко А.Е., Луцкий М.Г. Методы принятия решений в автоматизированной системе управления предполетной подготовкой летательных аппаратов. Монография. К.: НАУ, 2011 – С. 338.