Облікова картка НДДКР

Державний обліковий номер: 0211U000041

Державний реєстраційний номер: 0110U004136

Відкрита

Дата реєстрації: 11-01-2011

1. Етапи виконання

Номер етапу: 1

Назва етапу: Розроблення комп'ютеризованої інформаційно-аналітичної системи розуміння, інтерпретації, генерації текстової інформації та цифрової технології змістовної обробки текстової інформації.

Початок етапу: 01-2010

Закінчення етапу: 12-2010

Вид звітного документа: Остаточний звіт

2. Виконавець

Назва організації: Міжнародний науково-навчальний центр інформаційних технологій та систем НАН та МОН України

Код ЄДРПОУ/ІПН: 24741741

Підпорядкованість: Національна академія наук України

Адреса: 03680, ГСП, м.Київ-187, пр. Академіка Глушкова, 40

Телефон: 526-2549

Телефон: 526-1570

E-mail: vig@irtc.org.ua

3. Власник результатів НДДКР (продукції)

Назва організації: Національна академія наук України

Код ЄДРПОУ/ІПН: 00019270

Адреса: вул. Володимирська, 54, м. Київ, Київська обл., 01030, Україна

Підпорядкованість: Кабінет Міністрів України

Телефон: 380442350981

E-mail: prez@nas.gov.ua

WWW: http://nas.gov.ua

4. Джерела та напрями фінансування

Підстава для проведення робіт: 34 - договір (замовлення) з центральним органом виконавчої влади, академією наук (головними розпорядниками бюджетних коштів на проведення НДДКР)

КПКВК: 6541050

Напрям фінансування: 2.2 - прикладні дослідження і розробки

Джерела фінансування

Джерело фінансування: 7713 - кошти держбюджету

Фактичний обсяг фінансування за звітний етап: 259.52 тис. грн.

5. Науково-технічна робота

Назва роботи (укр)

Розроблення комп'ютеризованої інформаційно-аналітичної системи розуміння, інтерпретації, генерації текстової інформації та цифрової технології змістовної обробки текстової інформації

Назва роботи (англ)

To develop working project of the information and analytical system for understanding, interpretation, generation of the text information and digital technology of meaningful text information processing.

Реферат (укр)

Створено алгоритми семантичного аналізу текстів із застосуванням семантичних онтологічних баз знань; Побудовано мультилінгвістичну глобальну онтологічну базу знань, яка підтримує українську, російську та англійську мову, містить більше ніж 140 000 концептів . Створено технологію побудови інтелектуальних прикладних систем семантичної обробки текстів природною мовою; Розроблено ряд прикладних програм інтелектуальної обробки текстів - система "Рефератор", система "VitaminE", система смислової тематичної класифікації текстових корпусів та потоків, система смислового пошуку текстів, система семантичної фільтрації текстів, система визначення "запозичених" фрагментів тексту, система семантичної генерації оглядів корпусів текстів, система генерації мультиреферату корпусів текстів, система фільтрації Internet-повідомлень з використанням лінгвістичних методів аналізу текстів, система автоматизованої побудови природно-мовних інтерфейсів до баз даних та інших інформаційних систем.

Реферат (англ)

Research object - models, algorithms and software implementation of analysis, synthesis, semantic interpretation of natural language texts; algorithms for automatic translation, indexing, abstracting, best text models for solving: authorship definition and textual parallelism and plagiarism detection . The aim of work - creation within GTSTP "Pattern computer" a technology of intellectual linguistic natural language texts processing. This technology is the basis of a fundamentally new intelligent computer systems with developed modular architecture. Research methods - methods of computational linguistics, creating of mathematical models of morphology, syntax and semantics of natural language. Methods of graph theory, ontological semantics, associative-search networks, semantic text processing, machine learning and pattern recognition. To prove the correctness and effectiveness were created software systems of natural language text processing and their effectiveness was demonstrated. Main results - o The algorithm of semantic analysis using semantic ontological knowledge bases; o An ontological multilanguage global knowledge base that supports Ukrainian, Russian and English, contains more than 140 000 concepts. o Technology for building intelligent applications of the semantic processing of natural language texts; o A number of applications of intellectual text processing - System "Referator", system "VitaminE", semantic content classification system of text flows and corpuses, semantic text search system, semantic text filtering system, system for detection of "borrowed" text passages, system of semantic reviews generation, multidocument summarization system, Internet- messages filtering system using linguistic methods of text analysis, system of CAD of natural language interfaces to databases and other informational systems. The technology is unique in Ukraine and is at one level with the best foreign analogues.

Індекс УДК: 002.53/.55:001.814, 002.53/55.001.814

Коди тематичних рубрик НТІ: 20.19.19

6. Науково-технічна продукція (НТП)

НТП 1

Назва продукції (укр): Комп'ютеризована інформаційно-аналітична система

Назва продукції (англ): Computer information-fnflitic system

Очікувані результати:

Галузь застосування: Розроблена технологія може бути використана при створенні різноманітних багатофункціональних аналітичних систем автоматичної обробки текстів

Опис продукції (укр): Технологія смислової комп'ютерно - лінгвістичної обробки природномовних текстів, що базується на потужних лінгвістичних базах знань та алгоритмах інтелектуальної семантичної обробки текстів та взаємодіє з цими базами. Системи автоматичного смислового реферування та індексації (класифікації) текстів на природній мові, система семантичної обробки текстових документів на основі онтологічного семантичного аналізу, пошукова машина на основі семантичного онтологічного ядра бази знань та алгоритмів семантичного аналізу текстів та програмні модулі системи змістовного аналізу текстів, пошуку та визначення сегментів тексту.

Соціально-економічна спрямованість НТП:

Стадія завершеності НТП: Звіт по НДДКР

Впровадження НТП: Не впроваджено

Строки впровадження: 2011-2013

Виробник продукції: МННЦІТіС

Споживачі продукції: Державні, приватні установи, контролюючі органи

Перспективні ринки: Державні, приватні установи, контролюючі органи

Права інтелектуальної власності: За договорами

Форми та умови передачі продукції: Продаж продукції

7. Бібліографічний опис

1.Анисимов А.В., Тарануха В.Ю. Эвристические алгоритмы для определения канонических форм и грамматических характеристик слов // Кибернетика и системный анализ.- 2004. - №3.- С.6-17. 2.Анисимов А.В., Порхун О.В. Автоматична класифікація текстів з допомогою нейронної мережі Кохонена // Вісник Київського університету. - 2005.- № 1. 3.Марченко О.О. Методи автоматизації мовної локалізації розроблених онтологічних баз знань // Вісник Київського університету. - 2005. - №4. - С.120-1255.Марченко О.О., Дерев'янченко О.В. Моделювання семантичного контексту при аналізі текстів на природній мові // Вісник Київського університету. Сер. фіз.-мат. науки. -2006.- № 2.

8. Звітна документація

Кількість сторінок в звіті: 120

Мова звіту: Українська

Кількість файлів у звіті: 1

9. Заключні відомості

Перелік організацій-співвиконавців

Назва організації: Київський національний університет імені Т.Шевченка

Код ЄДРПОУ/ІПН: 02070944

Адреса: 01601, Київ-30, вул.Волоимирська,64

Підпорядкованість:

Перелік осіб-виконавців

Анісімов Анатолій Васильович

Луцкая Неля Митрофановна

Марченко Олександр Олександрович

Никоненко Андрій Олексійович

Тарануха Володимир Юрійович

Керівник організації:

Гриценко Володимир Ілліч (к. т. н., професор)

Керівники роботи:

Анісімов Анатолій Васильович

Керівник відділу реєстрації наукової діяльності УкрІНТЕІ

Юрченко Т.А.