**Щербак Сергій Сергійович. Методи та засоби формування баз онтологічних знань на підставі джерел табличних структур : Дис... канд. наук: 05.13.23 – 2007**

|  |  |
| --- | --- |
| |  | | --- | | Щербак С.С. Методи та засоби формування баз онтологічних знань на підставі джерел табличних структур. – Рукопис.  Дисертація на здобуття наукового ступеня кандидата технічних наук за спеціальністю 05.13.23 – системи та засоби штучного інтелекту. – Харківський національний університет радіоелектроніки, Харків, 2006.  Дисертацію присвячено поліпшенню організації електронного документообігу в Semantic Web шляхом розробки ефективних методів та засобів формування баз онтологічних знань на підставі джерел табличних структур. Розглянуті існуючі методи формування баз онтологічних знань та сформульовані їх недоліки. Вперше запропоновано метод формування баз онтологічних знань, що дає можливість на основі запитів у вигляді структурно-логічних схем додавати нові об’єкти предметної області, проводити цілеспрямований пошук онтологічних знань та поповнювати об’єкти новими екземплярами. Вдосконалено комплексну табличну модель мови гіпертексту HTML, що забезпечує умови для організації уніфікованого доступу при формуванні баз онтологічних знань. Одержав подальший розвиток метод аналізу джерел знань табличних структур, який забезпечує розширення вихідних даних за рахунок настроювання критеріїв інтерпретації структурно-логічних схем та формує запити на введення об’єктів предметної області до бази онтологічних знань. Вперше запропоновано метод формування екземплярів об’єктів предметної області на підставі джерел знань табличних структур, який відрізняється тим, що дані вводяться до баз онтологічних знань з урахуванням «спрямованої на властивості» реалізації об’єктів предметної області, завдяки чому забезпечується обробка наборів даних, які організовані за допомогою табличних структур. | |
| |  | | --- | | У дисертаційній роботі представлені результати, які у відповідності з поставленою метою є рішенням актуальної задачі підвищення ефективності електронного документообігу в Semantic Web шляхом розробки методів формування баз онтологічних знань на основі джерел табличних структур. Отримані результати мають важливе наукове й практичне значення як для розробки сучасних промислових технологій баз знань ініціативи Semantic Web, так і для рішення конкретних задач організації електронного документообігу підприємств на основі баз онтологічних знань із можливістю одержання нових знань у вигляді зв'язаних відношеннями об’єктів ПрО з джерел знань табличних структур різних форматів.  Проведені дослідження дозволили зробити наступні висновки.  1. В результаті аналізу сучасного стану рішення задачі формування баз онтологічних знань розглянуті особливості джерел знань табличних структур та відзначена орієнтація відомих систем і проектів на ручне введення знань про предметну область. Так широка різноманітність графічних способів структурування табличних даних та відсутність універсальних засобів їх аналізу становить серйозну перешкоду для ефективного застосування класичних методів аналізу джерел такого типу для створення об’єктів предметних областей, їх екземплярів і формування на їх основі баз онтологічних знань. Це й обумовило вибір напрямку досліджень, формулювання мети й задач досліджень.  2. Розроблено формальну модель баз онтологічних знань, що дозволяє забезпечити цикл робіт з обробки існуючих і додаванню нових знань до баз онтологічних знань у вигляді зв'язаних відношеннями об’єктів предметних областей шляхом введення необхідної функціональності до об’єктів предметних областей.  3. Запропоновано метод формування баз онтологічних знань, який на відміну від існуючих методів використовує адитивну ієрархічну організацію баз онтологічних знань, що дає можливість на основі запитів у вигляді структурно-логічних схем додавати нові об’єкти предметної області, проводити цілеспрямований пошук онтологічних знань та поповнювати об’єкти новими екземплярами.  4. Вдосконалено комплексну табличну модель мови гіпертексту HTML, орієнтовану на представлення довільних табличних структур, шляхом введення об’єктних характеристик для опису вмісту комірок, таких як «приналежність до об’єкта», «тип вмісту заголовної комірки», що забезпечує умови для організації уніфікованого доступу при формуванні баз онтологічних знань.  5. Одержав подальший розвиток метод аналізу джерел знань табличних структур, який забезпечує розширення вихідних даних за рахунок настроювання критеріїв інтерпретації структурно-логічних схем та формує запити на введення об’єктів предметної області до бази онтологічних знань, що дозволяє збільшити рівень розуміння вмісту джерел при автоматичній обробці.  6. Запропоновано метод формування екземплярів об’єктів предметної області на підставі джерел знань табличних структур, який відрізняється тим, що дані вводяться до баз онтологічних знань з урахуванням «спрямованої на властивості» реалізації об’єктів предметної області, завдяки чому забезпечується обробка наборів даних, які є організованими за допомогою табличних структур.  7. Здійснена розробка математичного забезпечення запропонованих у роботі методів, що характеризується високою швидкістю обходу як існуючих в базах онтологічних знань об’єктів предметних областей, так і комірок джерел знань табличних структур за рахунок переваг цілеспрямованого перебору, та реалізовано їх алгоритмічне забезпечення на прикладі формування бази онтологічних знань на підставі конкретного джерела табличних структур, що підтвердило їхню працездатність.  8. Проведено експериментальні дослідження з обчислення верхньої оцінки складності алгоритмічного забезпечення методу формування баз онтологічних знань на підставі цілеспрямованого перебору, що підтвердило його ефективність і лінійний характер зростання складності.  9. Розроблено архітектуру та здійснена реалізація комплексу програмних засобів формування баз онтологічних знань підсистеми аналізу табличних документів, що стала складовою частиною системи формування звітності про результати психологічних досліджень, які проходять в автоматизованих класах. Експлуатація цього комплексу підтвердила ефективність запропонованих в дисертаційній роботі моделей та методів, що засвідчено відповідним актом про впровадження. | |