**Храмова Ірина Олександрівна. Дослідження інформаційних інфраструктур підготовки даних для прийняття рішень : дис... канд. техн. наук: 05.13.06 / НАН України; Інститут проблем реєстрації інформації. — К., 2007. — 183арк. — Бібліогр.: арк. 156-161.**

|  |  |
| --- | --- |
| |  | | --- | | **Храмова І.О. Дослідження інформаційних інфраструктур підготовки даних для прийняття рішень.**- Рукопис.  Дисертація на здобуття наукового ступеню кандидата технічних наук за спеціальністю 05.13.06 - автоматизовані системи управління та прогресивні інформаційні технології. - Інститут проблем моделювання в енергетиці ім. Г.Є. Пухова Національної Академії наук України, Київ, 2007 р.  Дисертаційна робота присвячена розробці процесів та методів побудови інформаційної інфраструктури (ІнІ) підготовки даних для прийняття рішень у міжвідомчих організаційних структурах з різнорідних за ступенем структурованості та типом даних інформаційних ресурсів з роздільним управлінням.  Отримані нові наукові результати включають: вимоги та необхідні умови створення ІнІ визначеного типу; розроблений архітектурний підхід до синтезу ІнІ, інформаційну та функціональну структури допоміжної підсистеми управління ІнІ; математичну модель для аналізу характеристик продуктивності операційних вузлів ІнІ. | |
| |  | | --- | | 1. В дисертаційній роботі вирішене актуальне завдання – розробка процесів та методів побудови ІнІ підготовки даних для прийняття рішень у міжвідомчих організаційних структурах з різнорідних ІР з роздільним управлінням. Вирішеним є також завдання створення математичних моделей для аналізу функціонування операційних компонентів таких інфраструктур у сервіс-орієнтованому розподіленому середовищі. Сукупність наукових результатів дослідження дозволяє впровадити цілісний підхід до синтезу ІнІ корпоративних СППР на основі технологій СД, можливість інтеграції в яких досягається шляхом додавання семантичної складової до структури метаданих окремих ІС, виділення в окрему інформаційно-обчислювальну складову спеціалізованих інформаційних структур та процесів управління розподіленими метаданими, запровадження механізму семантичного узгодження інформаційних ресурсів, що долучаються до ІнІ СППР. Запропоновані аналітичні моделі дозволяють дослідити характеристики продуктивності операційних вузлів, на яких функціонують компоненти ІнІ, та проводити оцінку і співставлення різних підходів до організації в них обчислювальних процесів на пріоритетній основі. 2. Проведеним дослідженням технологій СД та їх застосування для створення ІнІ підтримки прийняття рішень виявлено, що хоча й існують певні технологічні передумови для утворення динамічних проблемно-орієнтованих конфігурацій інформаційних компонентів з децентралізованим управлінням елементами ІнІ, жодна з існуючих узагальнюючих структур СППР не включає необхідних для цього засобів. Показано, що хоча за формою організації та вимогами, що висуваються до даних, СППР найкраще реалізуються саме за допомогою технології СД існує істотна розбіжність між традиційним уявленням про функціональне призначення, критичність та місце технології СД і сучасними можливостями застосування цієї технології в інформаційному контурі СППР організаційного управління. Ці розбіжності стосууються як принципів проектування моделі даних, так і підходів до реалізації СД, здатних поєднуватися в межах корпоративних ІнІ. 3. Проведене дослідження способів та засобів побудови корпоративних СППР на основі мережі СД (вітрин даних) в умовах децентралізованого управління інформаційними структурами показало, що реалізація таких інфраструктур є досить складним технологічним завданням, конкретні рішення якого залежать від ряду окреслених у роботі факторів. Створення єдиної інформаційної архітектури, в якій контролюються всі інформаційні потоки, передбачає існування такого операційного середовища, яке складається з різноманітних інструментів і функцій, що реалізують безупинний процес створення, експлуатації, поступового розширення і зміни інформаційного складу без втрати його цілісності та семантичної узгодженості. Показано, що до основних компонентів такого середовища (програмних засобів доступу до джерел даних, інструментів очистки та перетворення оперативних і зовнішніх даних, сервера бази даних інформаційного сховища, програмних засобів управління сховищем) повинна бути додана спеціалізована інформаційна структура, яка утворила б керований інтегрований «шар метаданих». Окрім інформації, що стосується локалізації даних та механізмів доступу, яка надається за звичай репозиторіями або реєстрами, ця структура повинна утворювати і підтримувати також високорівневий опис семантики доступної в ІнІ інформації. 4. З метою підвищення зрілості ІнІ в напрямку залучення до використання в СППР спеціальних знань (про призначення, контекст та інші атрибути доступних інформаційних структур) запропоновано підхід до синтезу корпоративної ІнІ, який полягає у інтеграції семантичних метаданих складових ІC шляхом визначення єдиної онтології для спільного використання із подальшою синхронізацією понять окремих ІС з цією загальнодоступною онтологією. Запропонований підхід до побудови семантичної моделі ІнІ дозволяє підтримувати систему проекцій онтологій задіяних ІС на окремо розроблену стандартну онтологію, яка є загальнодоступним корпоративним ІР. Визначено стандартний життєвий цикл семантичної моделі інформаційних компонентів ІнІ СППР, який повинен включати щонайменше такі процеси: збір вимог, збір метаданих, розробка онтології (створення узагальненої інформаційної моделі), відбиття часткових схем даних на онтологію (мапування), публікація/розгортання онтології, використання онтології. 5. В роботі визначено типові задачі аналізу метаданих у ІнІ СППР, для рішення яких використовуються переважно методи розкриття даних: 1) класифікації; 2) прогнозування; 3) кластеризації; 4) знаходження виключень; 5) асоціювання. Для СППР в організаційному управлінні, де присутній значний (біля 80%) обсяг неструктурованої та напівструктурованої інформації, важливим самостійним напрямком інтелектуального аналізу даних стає розкриття текстів (Text Mining). Залучення до ІнІ СППР семантичних моделей складових ІС надає змогу розширити можливості пошукового механізму такими функціями інтелектуального аналізу, як нечіткий пошук, фільтрування (категоризація), персоналізація. 6. Головною рисою запропонованого підходу до побудови шару метаданих ІнІ СППР є відмова від фіксації сталого стану, він має змінюватися кожен раз із додаванням до ІнІ нового компоненту. Запропоноване в роботі практичне рішення доповнює функціональність мережі інформаційних ресурсів ІнІ веб-службою каталогізації (реєстрації) компонентів ІнІ, забезпечивши механізмом узгодження понять кожну успадковану систему чи службу реєстрації. Основні функціональні цілі запропонованих програмних компонентів полягають у тому, щоб дати можливість динамічно «вирівнювати» (співвідносити семантично) реєстри (каталоги, репозиторії) і підтримувати розподілене об'єднання метаданих, створених зв'язаними в рамках корпоративної мережі додатками. 7. Функціонування ІнІ СППР у відкритих архітектурах, орієнтованих на послуги, вимагає створення ефективного обчислювального середовища, що забезпечує високий рівень доступності загальних інформаційно-обчислювальних послуг і здатне адаптуватися до коливань в доступності ресурсів. Для дослідження продуктивності обчислювальних процесів в операційних вузлах ІнІ в роботі запропоновано модель СМО, що дозволяє моделювати різні варіанти організації обчислень на пріоритетній основі і може використовуватись для рішення задач аналізу та синтезу компонентів ІнІ. Сформульовано загальну математичну постановку задачі для визначення показників ефективності обслуговування в СД із адаптацією до непродуктивних відволікань ресурсів із врахуванням різних підходів до дисциплін обслуговування заявок у вузлі та дисциплін поновлення обслуговування після відмови вузла. Для різних комбінацій дисциплін обслуговування та поновлення обслуговування після відмови синтезовано відповідні моделі і отримані аналітичні вирази характеристик обслуговування в операційному вузлі ІнІ таких, як середній час очікування початку обслуговування заявок різних типів, середній час перебування заявок визначеного типу у вузлі, середнє число заявок визначеного типу в черзі на обслуговування, середнє число заявок визначеного типу в системі. Ці характеристики дозволяють обчислювати більш загальні показники ефективності функціонування компонентів ІнІ СППР, як то середні сумарні штрафи за перебування завдань у системі та їхню втрату або знецінення результату внаслідок перевищення припустимого директивного часу, фактична продуктивність вузла, втрати продуктивності внаслідок виникнення конфліктних ситуацій, виграш у еквівалентній продуктивності і т. ін. 8. Проведені дослідження стосуються переважно інтеграції систем підготовки даних для прийняття рішень, але запропоновані методики та підходи можуть бути використані і для інтеграції до корпоративних СППР систем видачі рекомендацій ОПР.   Подальші дослідження можуть розвиватися у бік створення теоретичних засад і практичних підходів до проектування, створення та впровадження нового класу інформаційних систем, які б відповідали вимогам змінюваності у часі як організаційної структури, так і архітектури задіяних у технологічних процесах наявних інформаційних та аналітичних ресурсів (віртуалізованих СППР). | |