

УДК 004.422

О.Ю. ЛИСЕНКО, кандидат технічних наук

О.Г. МАНОХІН

Л.В. МАНОХІНА

О.В. МАКСЮТА

А.О. НОВІКОВА

ДГЦУ

РОЗРОБКА МОДЕЛЕЙ КОМПЛЕКСНОЇ ІНТЕГРАЦІЇ ІНФОРМАЦІЙНИХ СИСТЕМ І РЕСУРСІВ ГЕМОЛОГІЧНОГО ПРОФІЛЮ

Частина перша

Актуальність теми

Прогрес у розвитку інформаційних систем (далі – ІС) і у вирішенні питань інтеграції різних ІС, а також розвиток в області мережевих інструментів і інтерфейсу користувача зумовлює зростаючий попит у світі як на спеціалізовані продукти, виконані для конкретного бізнес-процесу, так і універсальні інтеграційні ІС.

Відповідно до доручень Президента та Уряду України, а також виходячи з вимог наших клієнтів і ринку в цілому наша організація близько двох років тому поставила перед собою завдання щодо створення універсального інформаційного середовища – гемологічної інформаційної платформи (далі – ГП) нового покоління з метою подальшого розгортання на її базі різних підсистем та інформаційних ресурсів гемологічного профілю.

До основних підсистем, які складають розроблювану ГП, і завдань, що вони вирішують, можна віднести:

- Підсистеми інтеграції електронного документообігу і бізнес-процесів, орієнтованих на створення сценаріїв комплексної автоматизації бізнес-процесів (систем електронного документообігу – СЕД), для яких характерне тривале існування в часі, участь кінцевих користувачів і корпоративного програмного забезпечення. У рамках цієї підсистеми в ДГЦУ створюється комплексна автоматизована система електронного документообігу (КАСЕД).

- Портал – консолідація (користувальницького) доступу користувача (створення єдиної точки доступу) до інформаційних ресурсів.

- Підсистеми інтеграції даних – інтеграція (консолідація, синхронізація) корпоративних сховищ даних.

- Підсистеми черги повідомлень – забезпечення взаємодії (обмін даними,

виклик функції) корпоративних прикладних програм.

Зараз за планами НТР в ДГЦУ проводиться розробка КАСЕД, впровадження якої зробить можливим вирішення таких важливих завдань, як:

- автоматизація роботи з документами і бізнес-процесами;

- забезпечення спільної роботи з даними як у локальній мережі ДГЦУ, так і з використанням віддаленого доступу до інформаційних сервісів та ресурсів ДГЦУ;

- забезпечення зв'язку СЕД ДГЦУ з іншими інформаційними системами, у тому числі з ІС органів виконавчої влади;

- забезпечення безпеки і надійності збереження інформації.

Постановка завдань

На початку 2012 року Науково-технічною радою ДГЦУ прийнята програма здійснення комплексних заходів зі ство-

рення єдиної автоматизованої інформаційно-аналітичної системи з гемології за темою «Створення єдиної гемологічної інформаційної платформи засобів і продукції українського сегмента світового ринку товарів і послуг у сфері гемології», яка запланована на 2012–2016 рр. і виконується в три етапи.

Роботи першого етапу (2012 р.) були спрямовані на створення системи внутрішнього електронного документообігу ДГЦУ на платформі DocsVision, було впроваджено нову IT-інфраструктуру дата-центру, створено систему захисту інформації від несанкціонованого втручання зовні та забезпечення контрольованого доступу до ресурсів для організації віддаленої роботи експертів.

На другому етапі робіт (2013–2014 рр.) будуть розширені функціональні, сервісні та технічні можливості системи електронного документообігу. Для надання можливості самостійного доступу до інформаційних ресурсів ГП, як платним, так і безкоштовним, найважливішим атрибутом ГП є наявність клієнтських кабінетів (КК) на Порталі ГП. За допомогою КК клієнт зможе самостійно, зареєструвавшись у системі і сплативши за допомогою електронної системи платежів обрані ресурси, отримувати необхідні для нього послуги.

Зокрема, буде встановлена інтеграція з іншими зовнішніми інформаційними системами, що надають інтерфейси для обміну електронною інформацією, а також здійснюватиметься подальше створення на базі ГП різних інформаційно-довідкових ресурсів гемологічної спрямованості з можливостями їх розширення і модифікації.

У рамках другого етапу запланована робота на 2013 рік за темою «Розробка спеціалізованих інтерфейсів і протоколів обміну даними між різними системами». Планується розробити багатofункціональний інтерфейс між комплексною автоматизованою СЕД та бухгалтерською системою «ПАРУС», що дозволить працювати у двосторонньому напрямку з транспортними файлами, які забезпечать передачу даних між «ПАРУС» і СЕД, та вирішувати завдання контролю цілісності переданих даних і синхронізації баз даних. Реалізація цього напрямку дасть можливість включити фінансові документи бухгалтерського обліку в загальний документообіг комплексної СЕД організації.

На третьому етапі робіт (2015–2016 рр.) планується об'єднати на базі гемологічної платформи різні інформаційні ресурси та системи гемологічного профілю: КАСЕД, системи електронної взаємодії з органами виконавчої влади (СЕВ ОБВ), системи дистанційного навчання (СДН) та інші.

Необхідність об'єднати різнопланові ІС у єдиний інформаційний простір ставить певні вимоги до моделей можливої інтеграції. Пропоновані моделі повинні дозволяти реалізацію будь-яких складних систем з можливістю «розширення» їхнього технологічного базису. Одним із рішень цього завдання є використання портальних технологій, які дозволяють будувати ІС з готових блоків з наступним нарощуванням своєї структури.

Інформаційні системи і ресурси гемологічної платформи

На рисунку 1 представлена загальна схема ГП з описом інформаційних ресурсів і систем, які входять в неї, і зв'язків між ними.

Аналіз існуючих підходів до інтеграції ІС

Зараз на ринку існує велика кількість СЕД, що володіють, як правило, схожими базовими функціями. Істотної відмінності між ними з'являються лише в реалізації конкретних завдань.

Головною проблемою таких систем є недостатня розвиненість наявних засобів інтеграції з іншими системами. Ця проблема досить актуальна, оскільки в ДГЦУ вже існує кілька розрізнених інформаційних систем, поступово придбаних або розроблюваних у процесі роботи підприємства. Це викликано було тим, що не існує єдиної системи, яка б покрила всі функціональні потреби організації. У результаті було сформовано не єдине інформаційне середовище, а набір розрізнених систем.

Ці ІС здебільшого володіють різноманітним функціоналом, логікою, архітектурою і форматом зберігання даних. Причому більша частина таких систем створювалась різними розробниками для вирішення певних завдань, а значить, системи містять лише найпростіші (на рівні передачі інформації файлами певної структури) механізми інтеграції з іншими ІС. Ця ситуація викликає цілу

низку проблем, таких як багаторазове дублювання збереженої інформації, складність пошуку необхідних для синхронізації даних, низька надійність зберігання і невисока ефективність роботи з даними, а також складність підтримки цілісності і несуперечливості збережених даних. Для вирішення перерахованих проблем необхідна система, яка стане сполучною ланкою в разі об'єднання ІС в один інформаційний простір. Таке об'єднання буде найефективнішим, якщо системи підтримуватимуть декілька способів інтеграції [1]. Найпростіший з них полягає у можливості використання однієї і тієї самої інформації. Інший спосіб заснований на використанні стандартизованого опису переданих даних і надання системам набору сервісів для роботи з ним. Третій спосіб здійснює зв'язок ІС за рахунок створення в сполучній системі спеціальних бізнес-процесів, які дозволяють об'єднувати внутрішні бізнес-процеси цих систем. Як правило, СЕД містять в собі механізми інтеграції, спрямовані на вирішення конкретних завдань і підтримують лише деякі з перерахованих способів інтеграції.

В інтеграції ІС важливу роль відіграє импорт і експорт даних. Так як більшість СЕД засновані на описі документів з використанням метаданих, то під час импорту документів актуальним стає завдання автоматичного визначення метаданих. Відсутність у СЕД такого механізму є ще однією суттєвою проблемою.

Таким чином, актуальним завданням є розробка системи, яка дасть можливість на своїй основі різними способами інтегрувати існуючі в ДГЦУ ІС, а також забезпечить автоматизацію управління документами і бізнес-процесами. Очевидно, що така система повинна належати до класу СЕД.

Принципи інтеграції ІС

Незважаючи на те, що завдання інтеграції корпоративних прикладних програм не нове, воно як і раніше залишається одним з найсерйозніших, з яким доводиться стикатися багатьом підприємствам. Виконаний аналіз показав, що виокремлюється декілька рівнів інтеграції [2, 3]:

- *Інтеграція бізнес-процесів* заснована на визначенні, реалізації та керуванні процесами обміну інформацією між різними бізнес-системами.

- *Інтеграція прикладних програм* заснована на об'єднанні даних або функцій однієї програми з іншою, завдяки чому забезпечується інтеграція, близька до реального часу.

- *Інтеграція даних* заснована на ідентифікації та каталогізації даних з метою їх подальшого використання.

- *Інтеграція на основі стандартів* заснована на використанні стандартних форматів даних (наприклад, CORBA, JavaRMI, XML).

- *Інтеграція платформ* стосується процесів та інструментів, за допомогою яких різні системи можуть здійснювати безпечний і оптимальний обмін інформацією.

У літературі з інтеграції ІС виокремлено три принципи інтеграції [4]:

- *Інформаційно-орієнтований* заснований на використанні однієї інформації двома і більше системами. При цьому для надання можливості роботи з інформацією у кожній системі є набір відкритих сервісів. Цей принцип реалізується на рівні інтеграції програм.

- *Сервісно-орієнтований* заснований на використанні стандартизованого опису формату передачі даних. При цьому дані зберігаються в єдиній базі даних системи *middleware* і є набір сервісів для роботи з ними. Цей принцип реалізується на рівнях інтеграції даних, платформ та на рівні використання стандартів інтеграції.

- *Процесно-орієнтований* заснований на можливості приєднання до внутрішніх прикладних процесів кожної ІС таким чином, щоб не просто використовувати їх функції, а створити новий бізнес-процес, який би зв'язав ці ІС. Цей принцип реалізується на рівнях інтеграції бізнес-процесів, платформ та на рівні використання стандартів інтеграції.

Жоден з представлених принципів інтеграції не є універсальним і не існує загального способу вирішення завдання інтеграції ІС. Залежно від ситуації найбільш вдалим можуть виявитися різні способи інтеграції. Тому максимальна ефективність досягається в тому випадку, якщо інтеграція здійснюється на основі системи *middleware* і забезпечуються всі три принципи інтеграції. У ролі такої платформи пропо-

нується використовувати СЕД, що володіє низкою додаткових властивостей та здійснює інформаційне забезпечення трьох вищевикладених принципів. При цьому далеко не кожна СЕД може бути для цього використана. Основними причинами цього є:

- *Неможливість повного забезпечення сервісно-орієнтованого принципу інтеграції.* Для забезпечення цього принципу СЕД повинна надавати можливість зберігання і обробки як структурованих, так і слабоструктурованих даних, забезпечуючи їх стандартизованим структурованим описом і встановлюючи відносини між ними. Така СЕД повинна мати набір інтерфейсів, які забезпечать іншим ІС можливість роботи з цими даними.

- *Неможливість повного забезпечення процесно-орієнтованого принципу інтеграції.* Для забезпечення цього принципу система повинна мати можливість керування бізнес-процесами і надавати інтерфейс іншим ІС для створення в ній типових бізнес-процесів і керування ними у своїх цілях.

- *Відсутність можливості автоматичного визначення метаданих у текстовому документі,* що необхідно для здійснення ефективного імпорту слабоструктурованих даних в СЕД з інших ІС. Цей механізм також корисний під час потокового введення документів.

- *Відсутність можливості вибору оптимальної архітектури СЕД,* що в деяких випадках може призвести до обмеження функціональності системи в цілому.

Аналіз популярних СЕД щодо використання їх з метою інтеграції

Сьогодні на ринку існує багато систем, призначених для вирішення проблем автоматизації управління документами і бізнес-процесами (систем електронного документообігу). Значна їх частина має схожий базовий функціонал, властивий будь-яким системам такого типу. Кожна система має і характерні особливості, які є її «родзинкою». Спроби провести порівняльний аналіз цих систем і скласти зведену таблицю не мали успіху [5], оскільки

дуже важко сформуванати змістовну систему критеріїв так, щоб:

- критерії могли легко перевірятися,
- критерії були значимими з точки зору потенційного замовника,
- система критеріїв була повною,
- дійсно поділяла системи на різні групи.

На сьогодні у відкритому доступі можна знайти лише короткі порівняння різних СЕД, які не дають адекватної картини, бо порівняння, яке має практичну користь, повинне ґрунтуватися на великій кількості параметрів. Крім цього, кількість СЕД зростає експоненціально, що ще більше ускладнює завдання порівняння [6]. Тому було виокремлено тільки загальні критерії оцінки і далі проведено огляд найпопулярніших систем, які добре себе зарекомендували [7]. Для цього було проаналізовано публікації, присвячені питанням порівняння існуючих на ринку СЕД, і виділено ті з них, які мають максимальне число позитивних відгуків і лідирують у рейтингах [5, 8, 9, 10]: DocsVision, Albetty, DIRECTUM, ЕВФРАТ-Документооборот, OfficeMedia, ДЕЛО, PayDox, ЛЕТОГРАФ, Documentum.

Перед виконанням аналізу було визначено низку критеріїв для оцінки систем: призначення та основні об'єкти, функціонал, архітектура. Далі аналізу піддавався функціонал системи, зручність його використання і ефективність реалізації. Вивчення публікацій показало, що функції СЕД можна розділити на кілька груп [6, 11, 12]:

- Функції навігації і організації доступу користувачів до інформації.
- Функції роботи з документами.
- Функції маршрутизації документів і контролю їх стану.
- Функції роботи з архівом.
- Функції автоматизації бізнес-процесів.
- Функції організації групової роботи.
- Функції пошуку і керування даними.
- Функції з розширення функціональності.
- Функції, що забезпечують інтеграцію з іншими ІС.
- Функції, що забезпечують безпеку зберігання і передачі інформації.

¹Middleware – широко використовуваний термін, який означає шар або комплекс технологічного програмного забезпечення для забезпечення взаємодії між різними прикладними програмами, системами, компонентами.

У процесі аналізу здійснювалося вивчення архітектури системи. При цьому особлива увага приділялася особливостям реалізації системної платформи, сервера прикладних програм, сервера бази даних і клієнтської прикладної програми.

Аналіз систем, згідно з першим критерієм, показав, що всі вони поділяються на дві групи за тими ключовими об'єктами, з якими працюють, і, як наслідок, за своїм призначенням. До першої групи належать СЕД, ключовими об'єктами яких є тільки документи. Відповідно призначення таких систем – автоматизація управління документами. З розглянутих систем це Albetty і ДЕЛО.

До другої групи належать СЕД, основними об'єктами в яких виступають як документи (у найширшому їх розумінні – від звичайних паперових до електронних будь-якого формату і структури), так і бізнес-процеси, що визначають регламент руху документів та їх обробки. Такий підхід є одночасно конструктивним і більш універсальним, забезпечуючи автоматизацію документообігу та всіх бізнес-процесів підприємства в рамках єдиної концепції та єдиного програмного інструментарію. Такі системи деколи називають системами електронного документообігу нового покоління. Призначенням цих систем є автоматизація управління документами і бізнес-процесами. До систем другої групи належить більшість розглянутих систем: DocsVision, DIRECTUM, ЕВФРАТ-Документооборот, OfficeMedia, PayDox, Documentum, ЛЕТОГРАФ.

Підсумком аналізу функціоналу став повний список функцій систем із зазначенням особливостей їх реалізації для розглянутих СЕД.

Архітектура систем здебільшого є триланковою. Зазвичай в СЕД для зберігання даних використовують бази даних як власної розробки (наприклад, БД компаній «Гарант Інтернейшенл» і «Cognitive Technologies»), так і відомих розробників (наприклад, Oracle, MS SQL Server, Lotus Notes/ Domino, MS Access). Логіка систем найчастіше розташовується на сервері прикладних програм, часто об'єднаному з web-сервером, а в деяких випадках і з СУБД. Клієнтська частина, як правило, встановлюється на персональний комп'ютер користувача. Якщо надається можливість роботи зі стандартного web-браузера, то користувачеві доступ-

ний обмежений функціонал. Також існує ряд систем, орієнтованих виключно на web-клієнта. Під час впровадження таких систем користувач обирає архітектуру із заздалегідь визначеного списку доступних рішень.

Підбивши підсумок, в першу чергу видно, що в ролі основних об'єктів СЕД повинні виступати як документи, так і бізнес-процеси. За наявності відповідного функціоналу це дозволить вирішити завдання автоматизації управління документами і бізнес-процесами.

Аналіз списку функціоналу, необхідного для вирішення поставлених завдань, показав, що всі розглянуті системи мають схожі базові функції, а суттєві відмінності з'являються в реалізації більш серйозних механізмів. Це, наприклад, засоби розширення функціоналу систем, графічні редактори для створення бізнес-процесів, редактори звітів, розвинені засоби пошуку і т. ін.

До основних недоліків розглянутих систем можна віднести відсутність можливості автоматичного визначення метаданих зі змісту текстового документа і недостатню розвиненість наявних засобів інтеграції.

Усі розглянуті системи є комерційними і мають значну вартість, яка включає вартість самого програмного забезпечення, вартість ліцензії на кожне робоче місце, вартість використовуваної СУБД, вартість підтримки та оновлення системи. Значних витрат коштує безпосереднє впровадження системи, що включає настройку на існуючу специфіку організації, навчання персоналу роботи з системою. У результаті не кожна організація (і особливо ВНЗ) зможе дозволити собі такі витрати.

Здійснений аналіз СЕД показав, що жодна з розглянутих систем не може бути використана як платформа *middleware*, так, щоб повністю забезпечити всі три принципи інтеграції.

Основні принципи і вимоги до моделі комплексної інтеграції ІС

Відступаючи від конкретних систем реалізації, сформулюємо загальні вимоги, яким повинні задовольняти моделі, що дозволяють проектувати СЕД.

Моделі повинні забезпечувати:

- Можливість ефективного додавання, редагування, видалення і роботи з даними довільного (як структуровано-

го, так і слабоструктурованого) типу і розміру.

- Можливість використовувати одну і ту саму інформацію спільно з іншими ІС, що повинно здійснюватися за рахунок надання системам сервісів СЕД для роботи зі своєю частиною інформації та використання їх сервісів для роботи з даними цих ІС.

- Можливість інтеграції ІС за рахунок використання стандартизованого опису формату передачі даних і наявності у систем подібних сервісів роботи з цим форматом.

- Можливість об'єднання внутрішніх прикладних процесів кожної ІС за рахунок створення в СЕД типових бізнес-процесів.

- Можливість автоматичного визначення метаданих із змісту текстового документа, а також визначення метаданих з опису, представленого в спеціальному форматі (наприклад, в XML-форматі).

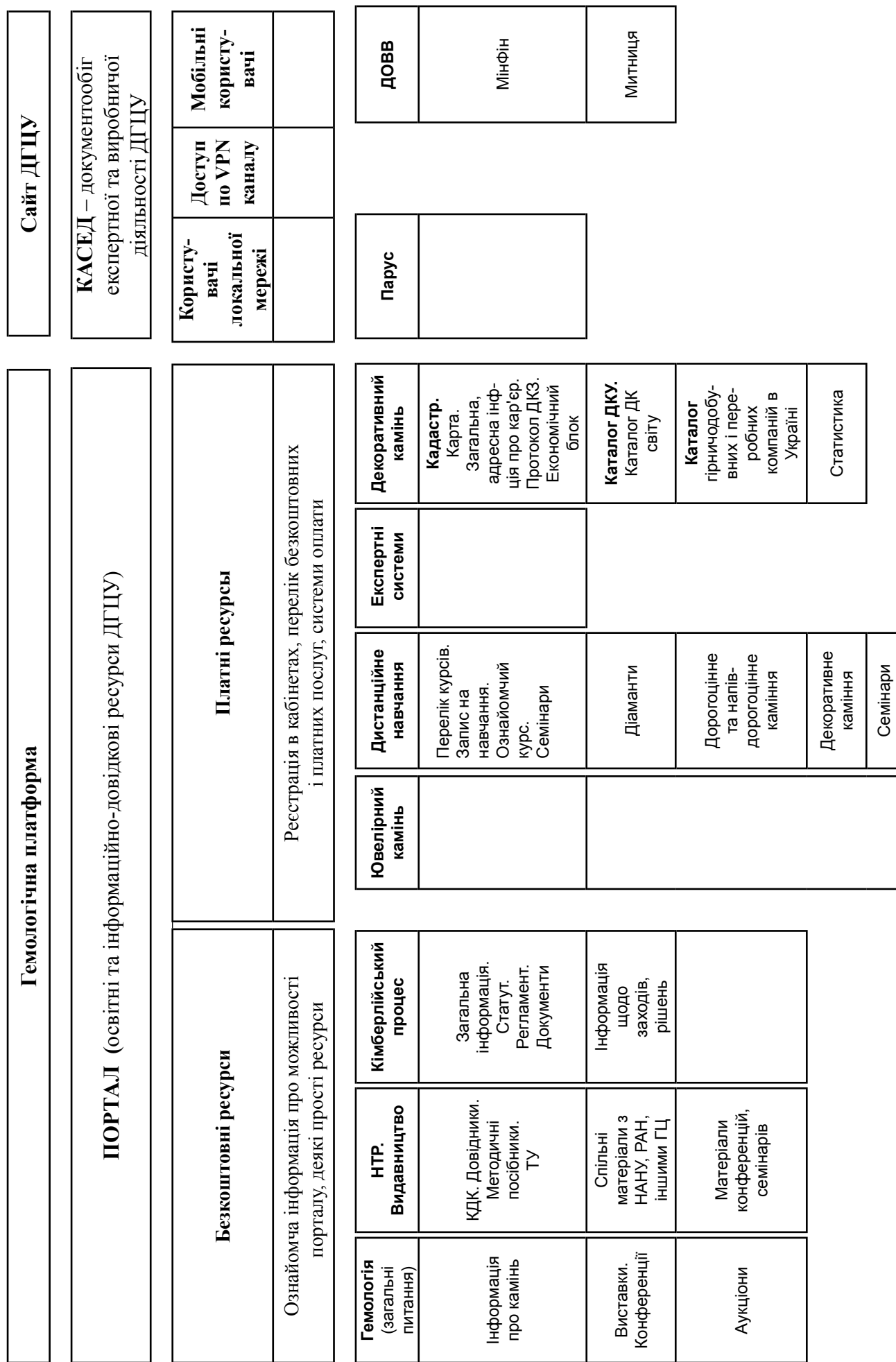
- Можливість управління документами (створення, редагування, видалення документів, ведення історії роботи з документами, підтримання можливості встановлення посилань між документами, автоматизацію заповнення частини метаданих, забезпечення можливості спільного доступу до даних, керування версіями документів, можливість зазначення посилань на документи, розташовані в зовнішніх ІС).

- Можливість маршрутизації (підтримка жорсткої і динамічної маршрутизації, моніторинг поточної діяльності користувачів з документами і завданнями, контроль поточного стану документів, автоматичне створення періодичних документів).

- Можливість керування бізнес-процесами, включаючи моделювання бізнес-процесів, виконання і моніторинг бізнес-процесів, наявність інтерфейсу для створення бізнес-процесів.

- Можливість реалізації зручних засобів навігації та організації доступу користувачів до інформації (наявність відносин порядку між об'єктами в системі за рахунок підтримки деревовидної структури представлення даних, забезпечення ефективного пошуку даних та керування ними, наявність закладки, що дозволяє зберігати важливі документи й контакти для подальшого швидкого доступу до них).

Рисунок 1. Загальна схема гемологічної платформи



- Безпека зберігання та передачі інформації (підтримка аутентифікації, авторизація, розмежування прав доступу до об'єктів СЕД, делегування прав доступу до документа від однієї особи іншій).

- Механізм налаштування СЕД на обрану архітектуру.

Висновки

1. На підставі аналізу літератури, присвяченої питанням інтеграції з інформаційними системами, здійснено класифікацію рівнів і принципів інтеграції. Показано, що максимальна ефективність досягається в тому випадку, якщо інтеграція здійснюється на основі системи middleware, і забезпечуються три принципи інтеграції: інформаційно-орієнтований, сервісно-орієнтований та процесно-орієнтований.

2. На підставі аналізу існуючих на ринку вітчизняних СЕД, що володіють максимальним числом позитивних відгуків і є лідерами рейтингів, визначено ряд вимог до моделей, за допомогою яких можна проектувати СЕД. Вимоги забезпечують як автоматизацію керування документами і бізнес-процесами, так й інтеграцію різних інформаційних систем. Одержано повний перелік функцій, необхідних СЕД для вирішення завдань автоматизації керування документами і бізнес-процесами.

3. На підставі вимог до моделей сформульовані вимоги безпосередньо до самої СЕД. Проведено аналіз на відповідність розглянутих СЕД пред'явленим вимогам до типової СЕД. Показано, що серед розглянутих систем, які повністю задовольняють вимогам, не існує.

4. Обґрунтовано необхідність розробки системи, яка задовольняє висунутим вимогам і забезпечує рішення завдань інтеграції, автоматизації керування документами і бізнес-процесами. Така система повинна вирішити завдання інформаційного забезпечення ДГЦУ.

Використана література

1. Гудов А.М. Информационные и математические модели, заложенные в систему электронного документооборота КемГУ [Текст] / А.М. Гудов, С.Ю. Завозкин // Вестник НГУ. – 2005. – Том 2, вып.1, серия «Информационные технологии в образовании». – С. 68 – 73.
2. Josey A. Интеграция корпоративных приложений: основные понятия [Электронный ресурс] / Перевод: Intersoft Lab. <http://citcity.ru/11132>.
3. Стрелкова Е. Интеграция данных предприятия [Электронный ресурс] / Е. Стрелкова // <http://www.osp.ru/os/2003/04/182921>. – Открытые системы. – 2003. – № 4.
4. Князева Т. Отечественные системы автоматизации делопроизводства [Текст] / Т. Князева // Компьютерра. – 2005.
5. Акопянц А. К выбору систем автоматизации документооборота [Электронный ресурс] / Акопянц, А. // <http://offline.computerra.ru/2003/523/31278>.
6. Андреев В. Приступая к созданию корпоративной системы автоматизации документооборота [Электронный ресурс] / В. Андреев // http://www.doc-online.ru/a_id/17.
7. Завозкин С.Ю. Система электронного документооборота ВУЗа [Текст] / А.М. Гудов, С.Ю. Завозкин // Труды VI Всероссийской научно-практической конференции «Системы автоматизации в образовании, науке и производстве». – Новокузнецк: СибГИУ, 2007. – С. 278–281.
8. Храмцовская Н. Стандарты СЭД: что подойдет России? [Электронный ресурс] // http://www.eos.ru/eos_delopr/eos_analytics/detail.php?ID=33191&SECTION_ID=767.
9. Сравнение лучших систем электронного документооборота 2010. [Электронный ресурс] / <http://ais.rissoft.ru/>.
10. Князева Т. Отечественные системы автоматизации делопроизводства [Текст] / Т. Князева // Компьютерра. – 2005.
11. Сенкевич В. Автоматизация хаоса, или зачем нужны системы электронного документооборота [Электронный ресурс] / В. Сенкевич // <http://www.silicontaiga.ru/home.asp?artId=3280>.
12. Типовые требования к автоматизированным системам электронного документооборота. Спецификация [Электронный ресурс] / <http://www.interface.ru/home.asp?artId=4286>.