

УДК 519.816

ЗАСТОСУВАННЯ СИСТЕМИ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ ДЛЯ ВИБОРУ КОНФІГУРАЦІЇ СИСТЕМИ ЕЛЕКТРОННОГО ДОКУМЕНТООБІГУ

О.В. Андрійчук, С.В. Каденко, В.В. Циганок

Інститут проблем реєстрації інформації НАН України

e-mail: andreychuck@ukr.net

1. Вступ. Стан проблеми

На перед-проектному (попередньому) та аналітичному етапах розробки систем електронного документообігу (Electronic record (document) management system або ERMS/EDMS, чи СЕД), коли формулюється технічне завдання, вкрай важливо подбати про те, щоб воно максимально точно відповідало вимогам клієнта (користувача). Формулювання технічного завдання покладається на проектну робочу групу, до складу якої мають входити аналітики, які представляють клієнта та аналітики, які представляють виконавця. Разом вони мають виробити технічне завдання, що забезпечить тривале та стає функціонування СЕД у майбутньому. Типові вимоги до систем управління електронними записами (electronic record management systems (ERMS)) (більш загальна категорія) та СЕД викладені у стандарті MoReq [1]. Більшість вимог, наведених у цьому документі, є функціональними, і можуть бути описані за допомогою булівських (наприклад, „система має містити засоби для управління вхідними чергами») чи кількісних (наприклад, максимальна кількість атрибутів метаданих, яку підтримує система) параметрів. Ці параметри (фактори або критерії) часто називають „відчутними” (tangible) [2]. Для опису не-функціональних вимог (представлених, здебільшого, у розділі 11 стандарту MoReq) необхідно формулювати „невідчутні” (intangible) фактори або критерії. Цифрові бібліотеки та архіви (що також є типами СЕД) мають довгі життєві цикли, тож не-функціональні вимоги мають критичне значення для цих типів СЕД, і іноді є важливішими за функціональні вимоги. Для оцінки рівня відповідності майбутньої СЕД не-функціональним вимогам виникає необхідність у залученні експертів до процесу розробки цієї СЕД. Чим адекватніше сформульоване технічне завдання (з урахуванням як функціональних, так і не-функціональних вимог), і чим точніше воно відповідає потребам клієнта, тим менше часу і коштів буде витрачено на реконструювання майбутньої системи під час наступних етапів розробки та впровадження. У [3] наводяться характеристики конкретних носіїв для архівного збереження даних. У даній доповіді пропонується, на основі комплексного цільового підходу, експертним шляхом оцінювати не окремі носії, а цілі конфігурації.

2. Опис підходу до розробки технічного завдання на основі експертних знань

У даному контексті аналітики (що представляють клієнта та виконавця) виступатимуть у ролі експертів. Ще раз підкреслимо, що використання експертних технологій дозволить усунути невідповідності майбутньої СЕД потребам клієнта, оскільки експерти є найбільш компетентними спеціалістами у заданій предметній області, здатними найбільш точно виявити такі невідповідності і врахувати їх під час формулювання вимог.

Не-функціональні вимоги являють собою слабо структуровану предметну область (проблематично визначити зв'язки між ними та кількісно описати їх). Отже, одним з найзручніших способів опису не функціональних вимог буде застосування ієрархічного підходу, тобто, побудова ієрархії вимог [2], до якої увійдуть як функціональні (відчутні), так і не-функціональні (невідчутні) вимоги. На відміну від підходу, викладеного у [4], в основу ієрархії критеріїв пропонується покласти готовий стандарт MoReq. Частина вимог можна взяти прямо зі стандарту, та сформулювати з них універсальний фрагмент ієрархії, адже, даний стандарт стосується усіх типів СЕД. Також слід зазначити, що вимоги,

викладені у MoReq, тісно пов'язані між собою і їх буде зручно представити у вигляді графу ієрархії. Водночас MoReq залишає простір для додавання нових вимог у випадку конкретних конфігурацій СЕД. Проектна група, звичайно, включатиме спеціалістів з більш вузьких галузей (експертів у предметній області, розробників, аналітиків, адміністраторів, тестерів, менеджерів тощо). Отже, кожен групу вимог, що відносяться до цих конкретних галузей мають формулювати відповідні спеціалісти зі складу проектною групи. Також саме ці спеціалісти мають оцінювати ступінь відповідності майбутньої конфігурації СЕД вимогам клієнта, що відносяться до відповідних галузей. Коли ієрархію побудовано, експерти мають присвоїти усім критеріям ієрархії відносні ваги. Потім усі можливі варіанти конфігурацій СЕД мають бути оцінені за усіма критеріями, після чого мають бути розраховані агреговані оцінки (показники відносної ефективності або рейтинги) варіантів конфігурацій за ступенем відповідності потребами клієнта.

3. Опис програми та гіпотетичний приклад

Технологія комплексного цільового експертного оцінювання, коротко описана вище, реалізована у серії систем підтримки прийняття рішень (СППР) “Солон”, розробленої у лабораторії СППР в Інституті проблем реєстрації інформації НАНУ (<http://www.dss-lab.org.ua/Main.html>). Підкреслимо, що СППР “Солон” призначена для забезпечення експертної підтримки прийняття рішень у будь-яких слабо-структурованих предметних областях. Багатокритеріальний експертний вибір найбільш прийнятної конфігурації СЕД є лише однією з задач, які можна розв'язувати за її допомогою. Пропонуємо розглянути роботу СППР на гіпотетичному прикладі, де оцінюються та порівнюються 5 альтернативних конфігурацій СЕД. У даному прикладі ієрархію вимог формулювали автори доповіді (на основі стандарту MoReq та загальних засад розробки СЕД), у дійсності ж передбачається, що її будують експерти у відповідній галузі. СЕД, варіанти конфігурації якої оцінюються в даному випадку, призначена для управління архівом гідрометеорологічної інформації. Архів призначений для накопичення та зберігання гідрометеорологічної інформації, та для інформаційної підтримки метеорологічних досліджень. Обсяги та типи даних наступні: 50 000 цифрових магнітних стрічок з даними про метеорологічну обстановку за кілька десятиліть, 1500 Тб супутникових даних на магнітооптичних носіях, 1000 Тб інших інформаційних матеріалів; 200 мільйонів різноманітних паперових документів форматів А0-А4, 400 мільйонів фотографій, у тому числі – документи на фотопапері, книги та рукописи. Загальний обсяг архіву – 6 Pb. Річний приріст обсягу даних – приблизно 10 Тб, у тому числі, 1 Тб даних наземних вимірювань та 8 Тб супутникових даних. Ключові завдання створення архіву: переведення до цифрового вигляду паперових носіїв; міграція даних на сучасні пристрої для зберігання; надання доступу до різних типів даних; впровадження веб-технологій та багаторівневого (ієрархічного) збереження даних. На основі цих завдань були сформульовані наступні варіанти конфігурацій. Конфігурація 1: стрічкова бібліотека IBM TS3500 (1 шт) (<http://www-03.ibm.com/systems/storage/tape/ts3500/>), дискова система зберігання даних IBM DS8300, Сервер управління на базі IBM system z9 BC під управлінням ОС z / vm і zLinux, ПЗ керування архівними даними IBM DB2 Content Manager OnDemand і IBM Tivoli Storage Manager (<http://www-01.ibm.com/software/tivoli/products/storage-mgr/productline/compare.html>). Конфігурація 2: Архівні оптичні накопичувачі ЭЛАР HCM 7000-BD (40 шт) (http://ncm.ru/nsm_bd.shtml), дискова система зберігання даних IBM DS8300, сервер управління на базі IBM system z10 EC (http://ru.wikipedia.org/wiki/IBM_System_z10) під управлінням ОС z / vm і zLinux, ПЗ системи архівації та управління документами Saperion (<http://www.saperion.ru/>). Конфігурація 3: Архітектура системи зберігання даних EMC Centera, що представляє надлишковий масив незалежних вузлів – RAIN (Redundant Array of Independent Nodes). Вузли доступу й вузли зберігання, що входять в архітектуру, представляють собою

сервери платформи Intel з АТА дисками. Сервера з'єднані між собою у внутрішню Ethernet LAN, а також мають Ethernet для зовнішнього підключення. Всі вузли працюють під управлінням модифікованої ОС Linux, що має назву CentraStar Operating Environment. Пошук у архіві EMC Centera здійснюється через ПЗ Centera Seek and Chargeback Reporter (<http://www.emc.com/products/detail/hardware/centera.htm>). Конфігурація 4: бібліотека на магнітній стрічці Quantum Scalar i2000 (10 шт.). (<http://www.quantum.com/ServiceandSupport/SoftwareandDocumentationDownloads/S2K/Index.aspx>); системи зберігання на магнітних дисках COPAN 400M Native MAID (1 шафа) від компанії SGI. (<http://www.sgi.com/products/storage/maid/400M/specifications.html>); сервер управління на базі IBM system z9 BC під управлінням ОС z / vm і zLinux. ПЗ QStar HSM. Конфігурація 5: Бібліотека на основі UDO або магнітооптичних носіїв Plasmon G638 (70 шт.) (http://www.dataarchivecorp.com/udo-plasmon_g-638.htm); системи зберігання на магнітних дисках COPAN 400M Native MAID (1 шафа) від компанії SGI ; сервер управління на базі IBM system z9 BC під управлінням ОС z / vm і zLinux. ПЗ системи архівації та управління документами Saperion.

Гіпотетична ієрархія критеріїв, за якими оцінюються альтернативні варіанти конфігурацій, будується наступним чином. На найвищому (0-му) рівні знаходиться лише головна ціль – „побудова ефективної системи архівного збереження даних”. Наступний, 1-й рівень включає її безпосередні під-цілі (або під-критерії) – “виконання вимог до системи збереження даних”, “низька собівартість”, “схвалення з боку вищого керівництва організації”. Ці цілі декомпонуються більш детально; 4-й та 5-й рівні ієрархії складаються з критеріїв, що відповідають вимогам стандарту MoReq (у тому числі, не-функціональні вимоги). Загалом, до ієрархії входить близько 90 критеріїв, пов'язаних між собою. Структура ієрархії наведена на Рисунку 1. Відносний вплив кожного критерію на свого „предка” у графі ієрархії оцінюється експертами. Коли усі впливи оцінені, обраховується відносна ефективність (рейтинги) усіх альтернативних варіантів конфігурацій. Відносні ваги (рейтинги) 5 наведених конфігурацій (що входять до ієрархії у якості проектів під номерами 87-91) показані у Таблиці 1. Повний перелік критеріїв (вимог), що входять до ієрархії, наведений у Таблиці 2. Як бачимо з Таблиці 1, Конфігурація 2 є найкращою за вказаними критеріями.

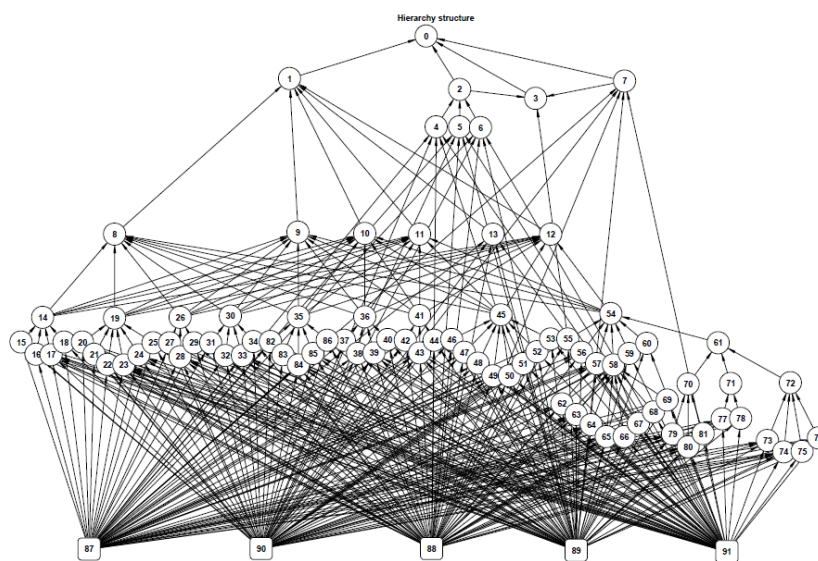


Рисунок 1. Структура ієрархії критеріїв

Таблиця 1. Рейтинги (відносна вагомість) альтернативних варіантів конфігурацій

Номер конфігурації	1	2	3	4	5
Відносна вагомість	0.2153	0.2361	0.1875	0.1667	0.1944
Ранг	2	1	4	5	3

Таблиця 2. Перелік критеріїв (під-цілей, або вимог)

№	Формулювання цілі	№	Формулювання цілі
0	Створення системи збереження даних архіву	46	Управління неелектронними документами
1	Задоволення вимог до системи збереження даних	47	Зберігання на протязі встановленого терміну й подальші дії з гібридними папками
2	Фінансові затрати на створення системи збереження даних	48	Управління інформаційними документами
3	Сприяння створенню системи з боку менеджерів вищої ланки	49	Управління робочими процесами
4	Наявність типових рішень	50	Електронний підпис
5	Наявність фахівців з попереднім досвідом розробки	51	Шифрування
6	Ступінь інтенсивності участі кожного фахівця	52	Електронні водяні знаки
7	Затрати часу на створення системи збереження даних	53	Інтероперабельність та відкритість
8	Принцип системної єдності	54	Нефункціональні вимоги
9	Принцип розвитку	55	Простота використання
10	Принцип комплексності	56	Продуктивність та масштабування
11	Принцип стандартизації	57	Доступність системи
12	Принцип універсальності	58	Технічні стандарти
13	Принцип нових задач	59	Законодавчі та нормативні вимоги
14	Схема класифікації	60	Аутсорсінг та управління даними третіми сторонами
15	Налаштування схеми класифікації	61	Довготермінове зберігання та боротьба з застаріванням технологій
16	Класи і папки	62	Апаратне забезпечення
17	Тома	63	Операційні системи
18	Обслуговування схеми класифікації	64	Промислові стандарти інтерфейсу користувача
19	Управління доступом й безпека	65	Реляційні СУБД
20	Доступ	66	Мережеві протоколи й операційні системи
21	Аудит/Системний журнал	67	Використання кодування на різних рівнях
22	Резервне копіювання й відновлення	68	Стандарти обміну
23	Управління переміщенням документів	69	Інтерфейс прикладного програмування й комплекти розробника
24	Автентичність	70	Деградація носіїв інформації
25	Категорії доступу	71	Застарівання обладнання
26	Порядок зберігання док-тів на протязі встановленого терміну та подальші дії	72	Застарівання форматів
27	Порядок зберігання	73	Міграція форматів
28	Експертиза цінності	74	Емуляція
29	Переміщення, експорт та знищення документів	75	Консервація технологій
30	Реєстрація документів	76	Зв'язування даних та програмного забезпечення
31	Реєстрація	77	Моніторинг стану обладнання

№	Формулювання цілі	№	Формулювання цілі
32	Масовий імпорт	78	Міграція інформації на нові сучасні носії
33	Типи документів	79	Дотримання відповідних умов зберігання, використання та обробки носіїв інформації
34	Управління електронною поштою	80	Забезпечити поточну заміну носіїв
35	Ідентифікація інформаційних об'єктів	81	Забезпечити збереження декількох копій документів та їх систематичне порівняння
36	Пошук, вибирання й представлення	82	Ідентифікація класу
37	Пошук й вибирання	83	Ідентифікація теки
38	Представлення: відображення на екрані	84	Ідентифікація тому
39	Представлення: друк	85	Ідентифікація документа
40	Представлення: інше	86	Ідентифікація витягу з документа
41	Адміністративні функції	87	Конфігурація 1
42	Загальне адміністрування	88	Конфігурація 3
43	Звітність	89	Конфігурація 4
44	Змінення, видалення, редагування й перегляд документів	90	Конфігурація 2
45	Інші функції	91	Конфігурація 5

4. Висновки

Запропоновано технологію експертної підтримки прийняття рішень для оцінки варіантів конфігурацій СЕД на етапі перед-проектного дослідження. Технологія є особливо актуальною у контексті оцінки варіантів конфігурацій щодо їхньої відповідності не-функціональним вимогам. Технологія дозволяє відобразити взаємні впливи та зв'язки між різними вимогами (представленими булівськими, кількісними та якісними критеріями). Рейтинги варіантів конфігурацій обраховуються на основі алгоритмічно побудованої функції агрегації оцінок, що включає булівські, чисельні вирази, а у випадку наявності в графі ієрархії зворотних зв'язків – передбачає ітераційний обрахунок рейтингів. Одним з можливих напрямків подальших досліджень є узагальнення підходу на інші стандарти розробки СЕД, окрім MoReq.

Література

1. MoReq Specification: Model Requirements for the Management of Electronic Records (2001). Електронний ресурс. Режим доступу: <http://www.cornwell.co.uk/moreq.html>.
2. Saaty, T. L. Relative Measurement and Its Generalization in Decision Making; Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors; The Analytic Hierarchy/Network Process. / T.L. Saaty // *RACSAM 102 (2)*, 2008, pp. 251-318.
3. Петров, І.В. Аналіз характеристик носіїв при побудові систем архівного зберігання інформації / І.В. Петров, Б.О. Березін, А.М. Стеценко, Н.В. Солоніна, В.О. Лесков // Реєстрація, зберігання і обробка даних, 2010, Т. 12, № 2 – С. 209-215.
4. Березін Б.О. Підтримка прийняття рішень при побудові систем довготермінового зберігання інформації / Б.О. Березін, П.Т. Качанов, В.В. Циганок, О.В. Андрійчук // Проблеми розвитку інформаційного суспільства: матеріали Міжнародного форуму. – 2009. – С. 145-153.