

І. Петров, В. Леснов,
ІПРІ НАН України

**ПРИНЦИПИ ТА ПІДХОДИ ДО ПОБУДОВИ
ЦИФРОВИХ СХОВИЩ ДАНИХ
ДОВГОСТРОКОВОГО ЗБЕРІГАННЯ
НА БАЗІ СУЧАСНИХ ТЕХНОЛОГІЙ**

Розглянуто питання методологічного забезпечення створення проектів цифрових сховищ даних довгострокового зберігання (ЦСДДЗ) для сучасних цифрових бібліотек та архівів. Запропонована узагальнена схема створення проекту цифрового сховища для зберігання інформаційних ресурсів наукової та культурної спадщини.

Ключові слова: створення цифрових сховищ даних, побудова систем довгострокового зберігання інформації, Digital Information Preservation.

Постановка проблеми

В умовах формування в Україні основ інформаційного суспільства особливої важливості набуває проблема збереження наукової та культурної спадщини загальнонаціонального значення, забезпечення вільного доступу громадян та фахівців до цих інформаційних ресурсів за допомогою комп'ютерних мереж та технологій як у теперішній час, так і у майбутньому.

Спроби реалізації деяких проектів зі створення цифрових бібліотек та архівів в Україні є спорадичними, малоефективними та нездатними задовольняти потреби суспільства на державному рівні. Переважну більшість цифрових інформаційних ресурсів створює та підтримує кожна інституція окремо, без належної взаємодії та координації з іншими установами. Галузеві бібліотеки та архіви створюють вузькоспеціалізовані сховища даних окремих видів інформації, які не завжди відповідають сучасним технологіям та стандартам.

Однією з ілюстрацій негативного стану справ у відповідній галузі є загальноукраїнська цифрова бібліотека дисертацій, яка знаходиться в Українському інституті науково-технічної та економічної інформації (УкрНТЕІ). До бібліотеки з сорока тисячами дисертацій доступ надається лише з 15 комп'ютерів, випущених до 1995 р. Вона не має офіційного статусу електронної державної наукової бібліотеки, а є лише підрозділом

відділу УкрНТЕІ. Для порівняння, у Росії дисертації з фондів «Електронної бібліотеки дисертацій» Російської державної бібліотеки (РДБ) можна прочитати в 250 віртуальних читальних залах, розташованих по всій Росії в багатьох вищих навчальних закладах та бібліотеках країни. У київських ВНЗ діє 8 віртуальних читальних зал РДБ та 5 в інших великих містах України. Отже, до поцифрованого російського наукового контенту українські науковці мають кращий доступ, ніж до українського національного [1].

Важливість практичних рішень поставлених завдань

Основу нормативно-правової бази для розвитку цифрової культури в Україні забезпечують закони та державні програми. Це Закони України «Про пріоритетні напрями інноваційної діяльності в Україні» (2003), «Про Національний архівний фонд і архівні установи» (2003), «Про бібліотеки і бібліотечну справу» (1995), «Про Основні засади розвитку інформаційного суспільства в Україні на 2007–2015 роки» (2007), «Про національну програму інформатизації» (1998), «Про музеї та музейну справу» (1995). Серед державних програм: «Державна Програма розвитку діяльності Національної бібліотеки України імені В. І. Вернадського на 2005–2010 роки» від 25 серпня 2004 р.; Державна Програма «Інформаційні та комунікаційні технології в освіті і науці» на 2006–2010 роки від 7 грудня 2005 р.; «Державна програма розвитку культури на період до 2007 року».

Огляд досліджень у галузі цифрового зберігання інформації

При створенні масштабних сховищ цифрової інформації ключовим моментом є вибір адекватної моделі зберігання архівного об'єкта (АО), що забезпечує відтворюваність і зрозумілість архівної інформації в майбутньому. Цей аспект включає як смислові та асоціативні зв'язки інформаційного об'єкта, так і елементарну можливість відтворення об'єкта в адекватному вигляді після, можливо, неодноразової зміни поточної версії прикладного ПЗ, у якому об'єкт спочатку створювався, або після того, як програма-технічна платформа, що застосовувалася при створенні об'єкта, припинила своє існування чи стала неприйнятною з найрізноманітніших причин (знята з виробництва, виробник належить іншій компанії тощо).

Одним з найбільш відомих сьогодні сучасних стандартів організації масивів цифрових даних є модель OAIS (Open Archival Information System), розроблена в кінці XX століття фахівцями NASA [2].

Ця модель чітко описує правила отримання первинної (еталонної) копії об'єкта (так званого тематичного змісту), доповнення об'єкта описовими компонентами (метаданими), подачу об'єкта у адміністративно-

виробничий цикл електронного архіву у вигляді сформованого особливим чином пакета даних (SIP – Submission Information Package), доповнення пакета низкою інших метаданих і «складування» об'єкта у вигляді АО у сховище.

На рівні сховища (архіву) передбачено низку процедур, що забезпечують міграцію об'єкта за необхідності з носія на носій, можливу конвертацію чи переробку об'єкта в більш досконалий формат, актуалізацію контекстних зв'язків об'єкта з іншими об'єктами (наприклад, сторінки книги, картина з галереї, відеозапис події тощо) в міру еволюції даного об'єкта і пов'язаних з ним інших об'єктів тощо.

Крім того, ця модель дозволяє при розробці управляючих і прикладних програмних комплексів конкретної установи жорстко регулювати правила видачі копій об'єкта за запитом. При цьому більшість об'єктів на архітектурному рівні експорту (DIP – Dissemination Information Package) матимуть кілька автентичних копій, що відповідають політиці інформаційної безпеки або технологічної доцільності експортних операцій: наприклад, ступінь дозволу растрової копії зображення, спрямованої у видавничий відділ для поліграфічного друку і виданої за запитом звичайного користувача електронного архіву будуть відрізнятися за обсягом, ступенем дозволу, глибиною кольору тощо.

2000 р. розпочався проект зі створення електронної бібліотеки в рамках Національної Королівської бібліотеки Нідерландів (НКБН). Застосування розглянутого вище підходу (використання моделі збереження даних OAIS, визначення адміністративно-виробничих процедур, що забезпечують гарантовану «упаковку» тематичних об'єктів і метаданих, переміщення пакетів АО у сховище, міграцію та еволюцію об'єктів зберігання з плином часу, тощо) дозволило до 2002 р. створити програмно-апаратне та адміністративно-виробниче рішення, що забезпечує виконання всього комплексу поставлених завдань. Головним розробником цього проекту виступила компанія IBM, виділивши спеціальну групу науковців, завданням яких було дослідження в галузі технологій, що дозволяють зберігати дані у вигляді пакетів, які забезпечують відновлення еталонної копії або цифрового оригіналу незалежно від актуальності програмно-апаратної платформи, на якій ці оригінали були створені. Рішення, розроблене IBM для НКБН отримало назву Digital Information Archival System (DIAS) [3].

Міжнародний проект CASPAR (Cultural, Artistic, and Scientific knowledge for Preservation, Access and Retrieval) [4], ініційований у квітні 2006 р. Європейським співтовариством і UNESCO, мав на меті розробку стандартів і механізмів довготривалого (необмеженого) терміну

збереження культурної, наукової та художньої спадщини світу. Він фінансується із загальноєвропейських джерел (наприклад, в рамках програми Sixth Framework Programme Priority IST-2005-2.5.10, «Access to and preservation of cultural and scientific resources»). В основі проекту – випробування і розвиток міжнародно визнаної, дієвої і вже апробованої моделі довготривалого зберігання електронних цифрових об'єктів OAIS Reference Model (ISO: 14721:2003). У рамках проекту CASPAR компанія IBM розробляє модель Preservation DataStore, що підтримує всі стратегії довготривалого зберігання архівних об'єктів.

Основні підходи та рішення при побудові цифрових сховищ даних довгострокового зберігання (ЦСДЗ)

При довготривалому зберіганні архівних об'єктів (АО) велике значення має вибір стратегії цього процесу.

Існують різні моделі збереження АО, призначені для забезпечення адекватного надання АО за запитом упродовж довготривалого терміну. У кожній з них є як свої переваги, так і недоліки. Комбінування різних моделей залежно від природи і призначення АО – основне завдання при розробці адміністративно-виробничих процедур електронного сховища.

Музейний тип: вихідні дані і засоби перегляду зберігаються в оригінальному вигляді та підтримуються в робочому стані. Перевага: гарантована можливість перегляду еталонної копії в первісному вигляді. Недолік: заборона на інтерпретацію даних, високі витрати на обслуговування програмно-апаратних засобів (необхідно зберігати та підтримувати швидко застаріваючі програмно-апаратні платформи).

Емулятор: вихідні дані зберігаються оригінальному вигляді за допомогою використання універсального віртуального комп'ютера IBM – UVC (Universal Virtual Computer) [5]. Перевага: значне скорочення витрат, пов'язаних з підтримкою та експлуатацією застарілих програмно-апаратних систем. Недолік: необхідність підтримки віртуальної платформи UVC.

Міграція: міграція основних компонентів тематичного змісту інформаційного масиву. Перевага: збереження основних характеристик ідентичності і цілісності об'єктів зберігання. Недолік: можлива поява «цифрового шуму».

Опис: створення описового змісту, що дозволяє відтворювати оригінал. Перевага: немає необхідності підтримки вихідних даних і оригінальних засобів перегляду. Недолік: втрата оригінального контенту та зміна його на похідний.

Серед великої кількості інформаційних технологій, задіяних у сучасних

системах зберігання даних, почали помітно виділятися ті, що підвищують ефективність функціонування при зниженні витрат на впровадження, експлуатацію та розвиток ЦСДЗ. Серед них на сьогодні найбільш актуальні:

Консолідація і віртуалізація. Об'єднання обчислювальних ресурсів розрізаних комп'ютерних систем і компонентів, що накопичуються у процесі довготривалого зберігання, в єдиний керований пул визначається як консолідація систем зберігання.

Реалізацію цього підходу забезпечує *технологія віртуалізації*. Ця технологія втілена в програмних компонентах верхнього рівня, які забезпечують управління і адміністрування систем зберігання. Компанії Cisco за останні чотири роки вдалося знизити вартість зберігання мегабайта даних майже в 17 разів за рахунок віртуалізації СЗД з мережевою архітектурою.

Гнучкий розподіл ресурсів середовища зберігання. Технологія тонкого розподілу (Thin Provisioning) втілюється в динамічному виділенні віртуалізованої місткості зберігання з єдиного резервного пулу середовища зберігання даних. Можливості цієї технології стають доступними тільки після впровадження технології віртуалізації. Це дозволяє уникати неекономного використання ресурсів середовища зберігання в процесі експлуатації системи та економії засобів при її побудові.

Стиснення і дедуплікація даних. Стиснення даних дозволяє знизити вимоги до місткості середовища зберігання за рахунок зменшення швидкості доступу до них. Технологія дедуплікації дозволяє усунути дублікати даних СЗД, що виникають в процесі довготривалого накопичення даних і скоротити місткість середовища зберігання, необхідну для зберігання інформації, в десятки разів (залежно від початкових даних).

Підвищення енергоефективності. Енергоспоживання СДЗД можна знизити за рахунок модифікації дискових масивів:

- застосування накопичувачів з меншою швидкістю обертання: енергоспоживання дискових масивів при цьому може бути зменшене майже на третину; переведення неактивних накопичувачів в режим очікування застосовується в деяких дискових масивах Dell і EMC.

- застосування накопичувачів з максимальною питомою щільністю зберігання даних; на поточний момент це 3.5” HDD пристрої, місткістю у 2 тб.

Нові компоненти середовища зберігання. Застосування в дискових масивах твердотілих накопичувачів (SSD). Популярними стають твердотілі накопичувачі (SSD), що в дискових масивах застосовуються як доповнення до HDD та в деяких моделях серверів HP замінюють жорсткі диски взагалі.

В цілому їх використання виправдане лише в системах корпоративного класу спільно з HDD через їх високу ціну.

Позитивні властивості:

- висока швидкість обміну даними;
- висока надійність у зв'язку з відсутністю механічно рухомих частин у накопичувачах;
- низьке енергоспоживання;

Недоліки:

- висока ціна;
- кінцева кількість циклів запису.

Уніфікація середовища зберігання СЗД. Уніфікації технічних характеристик компонентів СЗД різних виробників і забезпечення їх взаємозаміни шляхом відповідності відкритим галузевим стандартам дозволяє замовникам вибирати системи, які щонайкраще відповідають їх потребам. З цією метою виробники і користувачі мережевих систем зберігання заснували в кінці 1997 року Промислову Асоціацію Мережевих Систем Зберігання – Storage Networking Industry Association (SNIA) [6].

На сьогодні всі провідні виробники СЗД є членами цієї асоціації. Насамперед технічний комітет SNIA розробив і затвердив еталонну модель пам'яті (SNIA Shared Storage Model) [7], яка визначала варіанти існуючих архітектур і функціональність систем зберігання даних.

Загальна Модель Зберігання (SSM) описує набір практично можливих мережевих архітектур зберігання. Мережева архітектура задає специфічний функціональний розподіл сервісів через фізичні і логічні ресурси в середовищі зберігання. На практиці ця архітектура використовується для розроблення проектів систем зберігання, що включають вимоги до кількості і об'єму інформаційних ресурсів, і, можливо, вибір постачальників, версій, тощо.

Загальна Модель Зберігання є свідомо спрощена на перший погляд тому, що її мета – зробити модель легкою у використанні та достатньо універсальною, щоб охопити широкий спектр мереж зберігання, які вже існують або ще можуть бути розгорнуті.

Також – виділити фундаментальні структури середовища зберігання, які мають найбільший ефект при визначенні функціональності системи. Це включає пункти:

- Функції або послуги, які мережева архітектура зберігання може підтримувати.

Розподіл функціональної відповідальності серед компонентів системи. Взаємини між управлінням і потоками даних у системі.

- Межові умови при взаємодії між модулями архітектури, де

сумісність, ймовірно, є проблемою. Це включає місця, де має місце взаємодія (і потрібна сумісність), а також питання, які види сумісності повинні задовольнятися.

Забезпечення ефективного управління складним та багатокомпонентним середовищем зберігання великих обсягів даних може бути реалізованим лише за умови застосування типових стандартизованих методів, протоколів та інтерфейсів для всіх складників цього середовища. Саме такий підхід до управління будь-яким компонентом середовища зберігання було задекларовано при розробці специфікації SMI-S. Цю специфікацію розроблено 2002 р. асоціацією SNIA [6] на базі попередніх проектів у цій галузі. 2005 р. реліз SMI-S 1.1.0 був прийнятий ISO і став таким чином міжнародним стандартом [8].

У цьому стандарті реалізовано об'єктний підхід, що надає можливість інтерпретувати все обладнання мережевої пам'яті як сукупність об'єктів, абстрагуючись від їх фізичних характеристик, що підвищує потенціал адаптації та модифікації системи в цілому при появі нових технологій та пристроїв на ринку SAN систем. На сьогодні SMI-S стає загальноприйнятим стандартом побудови мережевої інфраструктури та середовища зберігання для ЦСДДЗ, і тому майже всі виробники технічних засобів SAN сертифікують свою продукцію на відповідність цій специфікації.

Основні вимоги до ЦСДДЗ:

- Централізація і структурованість: централізована багаторівнева структура, що виключає безпосередній доступ користувачів до еталонних копій документів страхового архівного фонду (САФ).
- Гарантоване збереження та автентичність об'єктів архівного зберігання: необмежений термін збереження з періодичною актуалізацією форматів, перенесення даних на сучасні носії із забезпеченням автентичності і, за необхідності, юридичної значимості електронних документів.
- Гнучкість: IT-архітектура відкрита для впровадження будь-яких прикладних задач і технологій, які можуть бути задіяні у майбутньому.
- Універсальність способів збереження фонду документів і використання сервісів організації власника фонду: використання апробованих міжнародних технологій і стандартів збереження є запорукою нормального функціонування ЦСДДЗ в майбутньому, і входження установи власника ЦСДДЗ (наприклад, бібліотеки чи музею) в національне або міжнародне професійне співтовариство.
- Масштабованість обсягів зберігання без зміни базових технологій:

застосовуються технологічні рішення, які повинні дозволяти нарощувати потужність без перерв і збоїв у роботі системи.

Ефективна модель управління ЦСДДЗ: наявність чітко визначених обов'язкових адміністративно-виробничих процедур (сервісів), що забезпечують гарантоване збереження фондів, наприклад, введення даних, комплектація, каталогізація, збереження та експорт (видача) матеріалів фондів.

Типізація проектного процесу. Застосування типізованих і стандартизованих моделей та інформаційних технологій дозволяє забезпечити вищу надійність процесу проектування, ніж у випадках використання інноваційних експериментальних моделей і технологій. Крім цього, визнані ефективними на практиці методи, моделі і технології оформлюються зазвичай у відкриті стандарти, розроблені і супроводжувані міжнародними комітетами і організаціями, що на сьогодні є суттєвою альтернативою закритим фірмовим технологіям, все ще розповсюдженим в інформаційній індустрії. Значною перевагою такої альтернативи є можливість безпроблемної зміни виробника або постачальника технічних засобів і рішень у процесі проектування і супроводу. У правовому та фінансовому аспекті відкриті стандарти звільняють від обмеження умов використання та ліцензійних відрахувань власнику закритих специфікацій.

Структурно-функціональний опис інформаційних систем, названий також архітектурою, сучасна теорія систем у більшості випадків визначає за допомогою блокових або багаторівневих еталонних моделей. У найзагальнішому випадку структуру інформаційних систем зберігання даних розділяють на 2 головних блоки – верхнього рівня, який визначає організацію інформаційних фондів системи і нижнього рівня, що визначає середовище зберігання даних. Більш детальний вигляд багаторівневої моделі сучасного ЦСДДЗ наведено на рисунку.

Визначення типової методики створення проекту ЦСДДЗ почнемо з розгляду загальної методології організації процесів практичної діяльності [9]. Проектування зазвичай розглядається у послідовних стадіях та етапах його проведення. Різними авторами їх склад і структура будується по-різному, але підпорядкувавши її загальній логіці організації процесу продуктивної практичної діяльності можна помітити, що структура вибудовується нібито в подвійній логіці одночасно: як послідовність дій проектування з одного боку, і за рівнями абстракції і конкретизації (зверху – вниз) – з іншого. Отже, процес проектування включає фази та стадії:

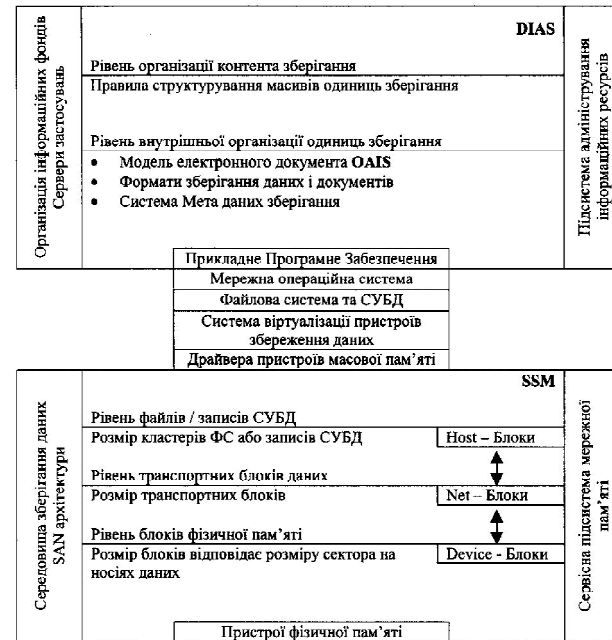


Рис. Багаторівнева модель сучасного ЦСДДЗ

Фаза зовнішнього проектування:

1. Концептуальна. Складається з етапів: виявлення протиріч; формулювання проблеми; визначення проблематики; визначення мети; вибір критеріїв.
2. Моделювання. Складається з етапів: побудова моделі; оптимізація; вибір (ухвалення рішення з технічного завдання – ТЗ на систему).

Фаза внутрішнього проектування:

1. Конструювання системи. Складається з етапів: декомпозиція; агрегація; дослідження умов реалізації проекту.
2. Технологічної підготовки розгортання та вводу в експлуатацію.

Типова методика проектування сучасних комп'ютерних інформаційних систем, різновидом яких є ЦСДДЗ, застосовується у відповідності із описаною вище методологією. Вона пропонує розподіл процесу створення проекту на декілька послідовних взаємопов'язаних стадій, основними з яких є стадії зовнішнього проектування, що завершується

розробкою технічного завдання (ТЗ) та повної специфікації системи (у загальній методології відповідають стадії 1 і 2), та стадії внутрішнього проектування, що має визначити всі складники системи та детальне узгодження їхньої взаємодії (у загальній методології – 3 і 4).

Фаза зовнішнього проектування ЦСДДЗ починається із визначення та формулювання прикладної задачі, пов'язаної зі збереженням даних. В першу чергу необхідно точно сформулювати цілі і завдання, що має вирішувати нова система зберігання даних, зробити попередню оцінку наслідків її повномасштабного впровадження. Головним орієнтиром на цьому етапі можуть бути такі базові складники системної концепції як:

- джерела інформаційного наповнення;
- тематика і типи даних фондів, що зберігаються;
- кількість і якість сервісів для користувачів системи.

Далі необхідно визначити кількісні показники, пов'язані з цими складниками, що є головними технічними характеристиками системи:

- приблизний об'єм сховища даних і кількість документів, що мають зберігатися, на початковому етапі експлуатації;
- динаміку приросту об'єму фондів, числа користувачів тощо, тобто провести оцінююче масштабування системи;
- продуктивність системи у разі одночасного обслуговуванні певного числа користувачів.

Всі вищезгадані фактори, а також організаційні, правові, фінансові, технічні та інші системоутворювальні аспекти у сукупності з можливостями сучасних інформаційних технологій надають вихідний матеріал для проведення системного аналізу завдання створення та застосування ЦСДДЗ. Результатом цих досліджень має стати концепція системи в цілому та технічні вимоги до ЦСДДЗ.

Концепція системи зберігання як конструктивна ідея є визначеною, якщо визначена еталонна модель архівного електронного документа як інформаційного об'єкта та модель або стратегія процесу його тривалого (понад 10 років) чи постійного зберігання в умовах змін програмно-технічних засобів, що використовуються для його створення та зберігання.

Технічні вимоги до ЦСДДЗ, крім забезпечення головних технічних характеристик, займаються такими питаннями:

- формати передавання і довготривалого зберігання документів та супровідна службова інформація (бібліографічні, технічні, архівні та інші групи метаданих);
- визначення критеріїв відбору та порядку передачі електронних інформаційних ресурсів;

- забезпечення автентичності та цілісності електронних інформаційних ресурсів;
- вибір та підтримка типових моделей структурування інформаційного контенту та архітектури середовища зберігання даних;
- забезпечення технологічного регламенту проведення конвертації та міграції інформаційних ресурсів у випадку необхідності зміни програмно-технічних засобів ЦСДДЗ;
- забезпечення порядку доступу до документів ЦСДДЗ та користування ними;
- впровадження комплексної системи захисту інформації ЦСДДЗ;
- гармонізація вимог до зберігання архівних і бібліотечних матеріалів, супровідних метаданих та технології цифрового підпису з міжнародними технічними стандартами і стандартами ЄС.

Після завершення стадії створення концепції сховища даних відбувається перехід до стадії конструювання моделі верхніх рівнів системи. Кількість базових компонентів системної моделі визначаємо через перелік основних завдань, які має розв'язувати система:

- організація збору Об'єктів Зберігання (ОЗ) та упаковка (можливо з конвертацією) у відповідні архівні контейнери, які визначені моделлю архівного електронного документа;
- структурування масивів ОЗ архівного фонду у середовищі зберігання;
- організація доступу до об'єктів зберігання зовнішніх користувачів;
- організація безпеки та надійності процесу зберігання;
- підтримка середовища зберігання та його адміністрування з урахуванням обраної стратегії зберігання.

Всі ці типові завдання вирішуються системами зберігання даних, побудованими на базі функціонально-структурної моделі IBM DIAS, що отримала широке застосування у міжнародній практиці створення великих проектів цифрових бібліотек та архівів. На базі цієї моделі провадиться узгодження та оптимізація взаємодії всіх застосованих у системі сучасних інформаційних технологій.

Завершується фаза зовнішнього проектування виданням ТЗ та повної специфікації ЦСДДЗ.

Фаза внутрішнього проектування призначена здійснити вибір апаратних та програмних компонентів, здатних реалізувати на практиці ТЗ, створене на стадіях зовнішнього проектування, також забезпечити етапи розгортання та вводу в експлуатацію ЦСДДЗ.

Світова практика побудови аналогічних систем свідчить, що типова

модель нижніх рівнів системи (рисунки), що визначає стандартну архітектуру середовища зберігання даних та принципи управління цим середовищем, суттєво спрощує етапи декомпозиції та агрегації на стадії конструювання системи. Найбільш відомою і, швидше за все, безальтернативною на теперішній час є модель SSM [7], що стандартизує нижні рівні систем зберігання, що можуть бути побудовані за типовими архітектурами SAN або NAS.

Завершення етапу декомпозиції та агрегації характеризується розробкою мережевої схеми системи згідно з обраною архітектурою та визначенням компонентів стека ПЗ, що забезпечить необхідну функціональність системи.

Базовий комплект пристроїв для побудови SAN-середовища:

- сервери застосувань та сервісних робіт;
- пристрої масової пам'яті для збереження фонду даних;
- комутатори мережевих каналів;
- каналоутворювальне обладнання.

Типовий стек ПЗ для забезпечення функціонування системи з архітектурою SAN:

- застосування та сервісне ПЗ;
- мережева операційна система;
- файлова система та СУБД;
- система віртуалізації пристроїв збереження даних;
- драйвери пристроїв масової пам'яті.

До кожного з цих компонентів системи повинні бути сформульовані окремі ТВ з урахуванням їхньої взаємодії.

Наступний етап загальної методології проектування систем — дослідження умов реалізації, призначений для вибору конкретних постачальників вищезгаданих компонентів на вітчизняному ринку інформаційних технологій згідно з вимогами, визначеним на попередньому етапі. На цьому етапі може знадобитись експериментальна перевірка окремих елементів, вузлів, серверів та мережевого устаткування на відповідність технічним вимогам до них у складі проєктованого ЦСДДЗ, оскільки іноді задекларовані у фірмових рекламних матеріалах характеристики продуктів не зовсім відповідають реальним, або в теоретичні розрахунки на попередніх етапах проєкту закралася помилка.

Остання стадія внутрішньої фази проектування систем — стадія технологічної підготовки розгортання та вводу в експлуатацію спроектованої системи. Вона полягає в підготовці робочих матеріалів, необхідних для експлуатації спроектованої системи: учбово-програмної

документації, інструкційно-програмного забезпечення тощо, а також, наприклад, посадових інструкцій виконавців, тобто всього комплексу документації, що супроводжує реалізацію складного проєкту. Після проведення технологічних прогонів та тестів настає черга здачі ЦСДДЗ в експлуатацію.

Висновки

Запропонована в дослідженні типова методика побудови цифрових сховищ даних довгострокового зберігання демонструє спроможність поступової формалізації процесу проектування цих систем на основі загальної методології організації процесів практичної діяльності. Це в свою чергу зумовлює можливість застосування часткової автоматизації цього процесу у вигляді комп'ютерної системи прийняття рішень.

В результаті реалізації методики побудови ЦСДДЗ в установах, що мають бути задіяні у програмах зі збереження культурної та наукової спадщини, очікується прискорення впровадження проєктних рішень, а також це дозволить:

- забезпечити належний рівень стандартизації та типізації архітектури, апаратно-програмних засобів, форматів даних для зберігання різнопланових документів та об'єктів у сховищах;
- зменшити людські та фінансові витрати за рахунок суттєвого зниження обсягів попередніх досліджень та вибору проєктних рішень.

Подальші дослідження щодо актуалізації, вдосконалення та використання типової методики побудови ЦСДДЗ для бібліотек та архівів мають бути спрямовані на слідкування за появою нових стандартизованих інформаційних технологій та засобів на світовому та вітчизняному ринках, а також на результати експлуатації впроваджених систем для корегування цієї методики за принципом зворотнього зв'язку. Методику слід також доповнити засобами забезпечення сертифікації спроектованих систем для контролю якості кінцевого результату.

Список використаних джерел

1. Пелагеша, Н. Європейська цифрова бібліотека : проєкт створення // Бібліотечний вісник. – 2008. – № 5. – С. 7.
2. ISO 14721:2003 http://www.iso.org/iso/catalogue_detail.htm?csnumber=24683
3. DIAS <http://www.ibm.com/nl/dias/>
4. CASPAR <http://www.casparpreserves.eu/>
5. IBM – UVC <http://www.ibm.com/nl/dias/resource/uvc.pdf>
6. SNIA <http://www.snia.org/standards/home>

7. SSM http://www.hpl.hp.com/research/papers/Shared_Storage_Model.pdf
8. ISO/IEC 24775:2007 http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=41545
9. Новиков, А. М., Новиков, Д. А. Методология // Российская академия наук, Институт проблем управления. – Москва, 2007. – С. 255–321.

УДК 025.17:025.13:004

Галина Ковальчук,
зав. відділу НБУВ,
д-р іст. наук, професор

КРИТЕРІЇ ВІДБОРУ КНИЖКОВИХ ПАМ'ЯТОК ДЛЯ ПЕРШОЧЕРГОВОГО ПОЦИФРУВАННЯ

У статті визначено певні засади, які мають враховуватися фондоутримувачами при виборі критеріїв для першочергового поцифрування книжкових пам'яток, а також охарактеризовано пріоритети, які обрав відділ стародруків та рідкісних видань НБУВ при формуванні тематичних електронних колекцій.

Ключеві слова: книжкові пам'ятки, унікальні примітки, першодруки.

Користувачі інформацією переконалися, що онлайнвий доступ до інформаційних ресурсів є зручним, швидким, привабливим економічно, коли ці ресурси надаються безкоштовно. Зрозуміло, що поцифрування накопиченої в книжках та інших видах друків інформації є невідворотним, і цей процес буде активно й швидко розвиватися. Відповідно постає питання відбору друкованих видань для першочергового поцифрування, адже нереально, принаймні найближчими роками, перевести у цифровий формат всі книги, журнали, газети, що були колись надруковані чи зберігаються, наприклад, у державних сховищах.

Критерії, за якими відбираються друки для першочергового поцифрування в установах-фондоутримувачах, відомі і обґрунтовані. Насамперед береться до уваги або актуальність тематики для сьогоденних споживачів (тобто потреба зробити текст доступним одразу значній кількості користувачів через мережу), або загрозливий фізичний стан примірника з метою зафіксувати інформацію на ньому, надати доступ до копії і вилучити оригінал з обігу, законсервувати його.

Ці основні підходи єдині як для сучасних масових видань, так і для книжкових пам'яток, переважно – історично за давних рідкісних і цінних видань (рукописні книги не розглядаємо, вони апріорі є пам'ятками історії та культури). Для книжкових пам'яток доцільно висунути більш детальні критерії, ніж вищеназвані. Безперечно, кожна бібліотека розробляє такі критерії для себе з урахуванням наявних фондів, профілю комплектування та обслуговування, певних традицій чи пріоритетів. Найголовніша вимога