

УДК 681.3.02

О. Г. Додонов¹, О. В. Нестеренко², А. В. Бойченко¹

¹Інститут проблем реєстрації інформації НАН України
вул. М. Шпака, 2, 03113 Київ, Україна

²Державний департамент з питань зв'язку та інформатизації України
вул. Хрещатик, 22, 01001 Київ, Україна

Методологія створення Національного реєстру електронних інформаційних ресурсів

Розглянуто питання забезпечення формування, інтеграції та використання електронних інформаційних ресурсів України в сучасних умовах. Запропоновано концептуальні рішення щодо створення Національного реєстру електронних інформаційних ресурсів.

Ключові слова: інформаційні ресурси, інтеграція, органи державної влади.

Постановка проблеми

В умовах формування в країні основ інформаційного суспільства особливої важливості набуває розбудова Системи національних електронних інформаційних ресурсів (СНІР) [1], метою якої є створення інструменту, здатного підвищити ефективність використання загальнонаціональних інформаційних ресурсів та істотно знизити витрати на їх утримання за рахунок використання метаданих щодо інформаційних масивів, баз даних, реєстрів, кадастрів та інших видів інформаційних ресурсів (ІР), у першу чергу тих, що утворюються за рахунок держбюджету в органах влади та інших державних установах та організаціях [2].

Проблематика формування СНІР пов'язана, перш за все, з тим, що наявні в державних органах автоматизовані інформаційні системи, а також бази даних є організаційно розпорошеними та функціонально роз'єднаними і не забезпечують інформаційну взаємодію державних та недержавних структур; створення та експлуатація баз даних, надання інформації, її використання та передача не забезпечені організаційно, не підкріплені належною нормативно-правовою базою; не існує єдиних правил та методів доступу до інформаційних ресурсів, які б забезпечували якісну інформаційну взаємодію органів влади, в тому числі з урахуванням вимог інформаційної безпеки. До цього слід додати відсутність єдиної системи класифікації та кодування інформації, певного рівня підготовки фахівців для роботи з інформаційними ресурсами, а також недосконалість правової та технологі-

чної бази надання інформації про діяльність органів державної влади. Як головний наслідок цих обставин можна назвати суттєве обмеження доступу громадян та підприємництва до інформаційних ресурсів, що є власністю держави.

Для вирішення вищезазначених проблем необхідним є створення інструменту, здатного забезпечити інтеграцію загальнонаціональних інформаційних ресурсів за рахунок використання метаданих щодо цих ресурсів та, як наслідок, підвищити ефективність їх використання.

Питанням розбудови національних інформаційних ресурсів останнім часом присвячується усе більше публікацій. Але слід зазначити, що хоча концептуально питання щодо інтеграційних процесів у цій сфері були поставлені, наприклад, в [3, 4], досліджені вони ще недостатньо. Також відсутня й методологія цих процесів. Невирішеними залишаються й задачі побудови ядра СНІР, яке має забезпечувати заходи щодо управління системою, захисту ресурсів, втілення методології створення, ведення та використання національних електронних інформаційних ресурсів.

Завданням цієї статті є проведення дослідження щодо методології створення Національного реєстру електронних інформаційних ресурсів (НРІР) — головного елемента СНІР.

Зв'язок із важливими науковими чи практичними завданнями

Створення НРІР тісно пов'язане з курсом України у бік розбудови інформаційного суспільства. Основу нормативно-правового забезпечення формування, використання та захисту національних ресурсів становлять Закони України «Про Національну програму інформатизації», «Про інформацію», «Про захист інформації в автоматизованих системах». Основні положення державної політики у сфері національних інформаційних ресурсів також визначає Концепція формування системи національних інформаційних ресурсів, затверджена розпорядженням Кабінету Міністрів України від 18.05.2003 р. № 259-р. Постановою КМУ від 17.03.2004 р. № 326 затверджено Положення про Національний реєстр електронних інформаційних ресурсів. Згідно з цим положенням до Національного реєстру включаються web-сайти, бази даних та реєстри в електронній формі.

Аналіз останніх досліджень і публікацій

Проблема створення, використання, захисту електронних інформаційних ресурсів стоїть сьогодні надзвичайно гостро. Стрімке зростання обсягів інформаційних ресурсів, які використовуються людством, спонукає наукові кола та виробників програмного забезпечення до активного пошуку вирішення задачі підвищення ефективності використання інформації. Головними напрямками пошуку є впорядкування інформаційних масивів, розробка нових методів і технологій зберігання й доступу до інформації та забезпечення захисту інформації.

Як засвідчує досвід США, зусилля щодо впорядкування електронних інформаційних ресурсів шляхом їх каталогізації дозволяє досягти інтеграції останніх до потужної системи національних бібліотечних каталогів. Тому більшість зусиль до каталогізації прикладають саме бібліотеки. Слід виділити проекти InterCAT, CORC, INFOMINE, CATRIONA [6].

OCLC (Online Computer Library Center) — міжнародна організація, яка об'єднує понад 53 тисячі бібліотек у 96 країнах і має на меті допомогти бібліотекам та іншим агенціям організувати інформаційні ресурси світу таким чином, щоб шукачі знання могли знайти їх задля створення нового знання. Для цього OCLC допомагає бібліотекам обмінюватися описами своїх колекцій та створює рамки співпраці для організації нового знання [7]. Основними напрямками діяльності цієї організації є: Cataloging and Metadata, Collection Management, Digitization and Preservation, Electronic content, Reference, Resource Sharing, DUNS (Data Universal Numbering System).

Більшість країн на урядовому рівні розробили і продовжують роботу над стандартами для створення, розвитку та використання інформаційних ресурсів [12]:

Великобританія: урядовий шлюз, урядовий Інтернет, e-GIF, e-GMS (e-Government Metadata) — стандарт, який задає елементи для опису метаданих та їх відображення Dublin Core, AGLS, NGDF, GILS та PRO;

Данія: інфраструктура Infostructurebase;

Швеція: Government Elink (Ge);

Австралія: FedLink — урядовий шлюз і захищений урядовий Інтранет;

Гонконг: Government System Architecture (GSA) і Electronic Service Delivery (ESD) Scheme;

США: Федеральна корпоративна архітектура інформаційних технологій державних організацій.

Особливий інтерес представляють підходи США до стандартизації у сфері використання інформаційних технологій у державних організаціях. Федеральний уряд США вважає, що критично важливу роль в сфері ініціатив з реформування й підвищення ефективності роботи державних організацій за рахунок впровадження інформаційних технологій відіграють дві складові:

— реалізація принципів електронного уряду;

— концепція корпоративної архітектури державної організації.

Останнім часом значних результатів досягнуто у галузі доступу до інформаційних ресурсів за допомогою web-сервісів.

Web-сервіси забезпечують практичну можливість будувати найперспективнішу сервісно-орієнтовану архітектуру (Service Oriented Architecture, SOA) розподілених інформаційних інфраструктур із використанням технологій web-сервісів на основі промислових продуктів (наприклад, BizTalk Server 2004 від Microsoft). Очікується, що з 2007–2008 рр. альтернатив підходу SOA не буде [8, 9].

02.02.2005 р. Асоціація OASIS затвердила як стандарт специфікацію UDDI 3.0, що регламентує реєстри доступних для виклику web-сервісів. На думку експертів, UDDI 3.0 — перша версія специфікації, що відповідає вимогам сервіс-орієнтованої архітектури. Стандарт дозволяє створювати розподілені реєстри інформаційних ресурсів. Однією з його основних особливостей є наявність механізму підтвердження автентичності інформації у реєстрі, реалізованого на основі цифрових підписів формату XML Digital Signatures. Модель безпеки, яка використовується в UDDI 3.0, побудована на політиках — за їх допомогою контролюються керування доступом, тиражування, підписка, делегування прав на пересилання даних і ключі UDDI — унікальні ідентифікатори, які привласнюються кожному

запису реєстру. UDDI 3.0 дозволяє створювати закриті, напівзакриті й відкриті реєстри з можливістю їхньої інтеграції на різних рівнях [11].

Специфікація UDDI підтримує наступні системи класифікації:

NAICS — система класифікації галузей промисловості;

UNSPSC — універсальна стандартна класифікація товарів та послуг;

ISO 3166 — стандарт стосується кодів країн, регіонів тощо.

Реєстр UDDI являє собою розподілену БД яка містить інформацію трьох категорій.

Білі сторінки — назва, адреса організації, контактна інформація, web-адреса, універсальний номер у системі DUNS.

Жовті сторінки — вид комерційної діяльності, місце дислокації та основна продукція, належність до різних класифікаторів (NAICS).

Зелені сторінки — специфікація комерційних послуг, порядок бізнес взаємодії, також WSDL файл.

Важливе місце у дослідженнях займають проблеми захисту інформаційних ресурсів. Слід виділити дослідження у галузі захисту інформації, які вже стали традиційними та зусилля, спрямовані на створення національних та наднаціональних захищених інформаційних структур. Заслужують на увагу підходи Китаю щодо контролю Internet та безпечна архітектура США щодо створення системи інформаційної безпеки шляхом реалізації доктрини «інформаційної парасольки».

Особлива увага до питань інформаційної безпеки обумовлюється також тим, що у більшості індустріально розвинутих країн проводяться дослідження та розробки нової інформаційної зброї, що потенційно загрожує безпосереднім контролем над інформаційними ресурсами держави. Інформаційні загрози реалізуються через порушення інфраструктури, вільного обігу інформації, неправомірні дії щодо використання інформації; через невідповідність інформаційної політики, засобів інформування громадськості життєво важливим інтересам суспільства. Широке використання маніпулятивних технологій, тенденційна модифікація інформаційних ресурсів, формування викривленої інформаційної реальності призводить до зруйнування глобального інформаційно-психологічного середовища, трансформації ціннісних орієнтацій суспільства, порушення фундаментальних прав і свобод громадян [10].

Глобального характеру набули інформаційні загрози в науково-технологічній сфері: від феномену транскордонного переміщення інтелектуальних ресурсів, тобто вивезення інформації унікального науково-технологічного характеру на біологічних носіях, до міжнародних систем спостереження, аналізу та прогнозування тенденцій науково-технологічного розвитку в різних країнах з метою доступу до конфіденційних баз і банків даних. Критичними для безпеки у сфері науки та технологій є структури накопичення науково-технічної інформації, інституції та структури фундаментальних і прикладних досліджень, об'єкти інтелектуальної власності, ноу-хау. Інформаційно-технологічний аспект безпеки зорієнтований на реалізацію системних заходів, спрямованих на максимальне вдосконалення науково-технологічної сфери, ефективний захист інтелектуальних ресурсів.

Особливості створення Національного реєстру електронних інформаційних ресурсів

Створення НРІР вимагає неперервного системного аналізу національних інформаційних ресурсів, їх взаємозв'язків та захищеності. При цьому НРІР слід розглядати як розподілену складну інформаційну систему, оскільки він наділений усіма властивостями складних систем, а саме: великою кількістю складових компонент, взаємодією з навколишнім середовищем, ієрархічною структурою та мінливістю у часі.

НРІР має поєднати у собі окремі реєстри інформаційних ресурсів, які забезпечуватимуть облік інформаційних ресурсів, згрупованих за певною спільною ознакою. Наприклад: за формою власності — Реєстр інформаційних ресурсів органів державної влади; за тематикою — Реєстр інформаційних ресурсів національної культури; за місцезнаходженням — регіональні реєстри інформаційних ресурсів тощо.

Взаємодія НРІР із навколишнім середовищем передбачає створення ряду захищених каналів для доступу широкого кола користувачів.

Ієрархічність НРІР полягає у підпорядкованості реєстрів, які поступово включатимуться до складу НРІР, та забезпеченні централізованого керування всією системою.

Під мінливістю НРІР у часі мається на увазі здатність до нарощування, модернізації та мобільної зміни конфігурації програмно-апаратного комплексу НРІР у залежності від стану національних інформаційних ресурсів.

Основні завдання та функції Національного реєстру електронних інформаційних ресурсів

При побудові НРІР необхідно дотримуватися наступних умов:

— зберігання інформації щодо зареєстрованих та знятих з реєстрації інформаційних ресурсів із забезпеченням конфіденційності, цілісності та доступності інформації;

— автоматизації процесів реєстрації та зняття з реєстрації інформаційних ресурсів як загального користування, так і спеціалізованих, довільних тематичних напрямків, з урахуванням обмежень на розпорядження та використання інформаційних ресурсів, що накладаються їх власником, а також законодавством України;

— забезпечення пошуку інформації українських інформаційних ресурсів, доступних для вільного розповсюдження, включаючи контекстний пошук, пошук за реквізитами та пошук по каталогу;

— інтеграції НРІР з Єдиним web-порталом органів виконавчої влади;

— формування інформаційних бюлетенів (каталогів) інформаційних ресурсів України;

— підготовки копій НРІР для формування зведеного Національного архівного фонду України.

У Національному реєстрі електронних інформаційних ресурсів реалізуються наступні групи функцій.

Обліково-реєстраційні функції:

— занесення інформації про ІР до Реєстру;

- редагування інформації про ІР в Реєстрі;
- видалення інформації про ІР з Реєстру;
- автоматизована індексація ІР з Інтернету.

Класифікаційні функції:

- експертиза тематичної належності ІР;
- створення системи рубрик (галузь знань, розташування, належність);
- редагування системи рубрик;
- автоматизована рубрикація даних.

Інформаційно-пошукові функції:

- пошук ІР за каталогами (тематичні рубрики, регіони, власники);
- пошук ІР за його реквізитами (ключові слова, дата реєстрації, місцезнаходження, власник, тематична рубрика).

Інформаційно-аналітичні функції:

- отримання статистичної та аналітичної інформації щодо ресурсів;
- визначення дублювання та неузгодженості ресурсів;
- визначення авторських прав на ІР або його частину.

Адміністративні функції:

- управління розподіленою базою даних;
- управління збереженням та відновленням даних НРІР;
- управління безпекою.

Web-портал Національних інформаційних ресурсів має забезпечити отримання інформації щодо певного інформаційного ресурсу, включаючи точну адресу.

Для вирішення задач СНІР доцільно використовувати найбільш важливі види класифікації:

- за типами об'єктів даних (документи, таблиці, колекції тощо);
- за структурними ознаками (файли, каталоги, розділи і т.д.);
- за методами формування (бази даних, web-сайти, реєстри, кадастри тощо).

Основні технологічні рішення при створенні Національного реєстру електронних інформаційних ресурсів

Для реалізації НРІР використано набір сервісів, згрупованих у функціональні підсистеми, до яких входять:

- розподілена база даних;
- підсистема індексації та пошуку інформації;
- комунікаційна підсистема;
- підсистема адміністрування;
- підсистема захисту інформації.

Користувачі НРІР мають взаємодіяти із сервісами за допомогою клієнтського програмного забезпечення, наприклад, системи прийому електронної пошти, програми доступу до даних, що розміщені у базі даних тощо.

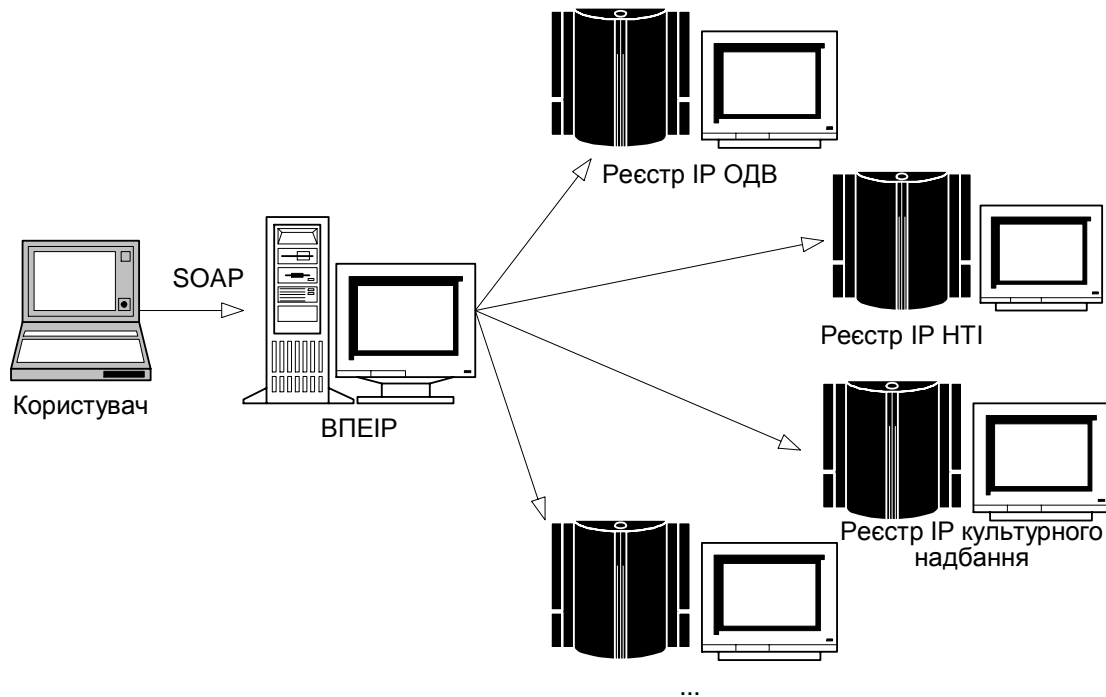
Програмно-апаратний комплекс НРІР включає наступні компоненти:

- сервери баз даних;
- сервери додатків;
- комунікаційні сервери;
- робочі місця адміністраторів й операторів;

- засоби резервування;
- комунікаційне обладнання для підключення до сучасних високошвидкісних мереж;
- периферійне обладнання, достатнє для ведення НРІР.

Інтеграція інформаційних ресурсів до НРІР передбачає створення розробниками ІР спеціальних програмних елементів — менеджерів ІР, які забезпечуватимуть інформаційну взаємодію відповідного ІР з іншими елементами СНІР. Для цього менеджер ІР реалізує інтерфейси доступу до відповідного ІР на основі стандартизованих в межах СНІР правил та механізмів. Інструментом для реалізації даного підходу є технологія створення реєстру UDDI, розроблена для реєстрації інформації щодо підприємств та їхніх web-сервісів.

Для занесення інформації до НРІР передбачене використання інтерфейсу SOAP (див. рисунок). Для забезпечення реєстрації та пошуку даних власники ІР надсилають WSDL-описи web-сервісів.



Доступ користувачів до відкритої частини розподіленої БД НРІР.

НРІР є розподіленою системою, що забезпечує можливість реєстрації ІР на будь-якому з реєстрів НРІР. З постійним проміжком часу інформація, зареєстрована в одному з вузлів РБД НРІР, переноситься до інших. Користувач при зверненні до будь-якого з вузлів НРІР може отримати інформацію щодо зареєстрованих ІР. Власник ІР має звертатися лише до того з вузлів НРІР, у якому його ІР було зареєстровано спочатку.

Опис ІР складається з ідентифікаційної та класифікаційної інформації.

Модель даних НРІР містить наступні групи структур даних:

- описи інформаційних ресурсів;

- класифікатори ІР;
- представники;
- канали доступу;
- взаємозв'язки між ресурсами.

Представники — описи організацій та фізичних осіб, які пов'язані з ІР на якомусь з етапів його життєвого циклу: власники, розробники, видавці.

Канали доступу — це шляхи отримання ІР: он-лайн доступ, отримання за запитом через електронну пошту та будь-які інші.

Опис інформаційних ресурсів

Враховуючи поточний стан та перспективи розвитку, в НРІР закладено можливість класифікації інформаційного ресурсу за наступними ознаками, такими як [5]:

- унікальний ідентифікатор інформаційного ресурсу;
- назва інформаційного ресурсу;
- дата реєстрації інформаційного ресурсу;
- реферат — короткий опис інформаційного ресурсу;
- джерело — до яких зареєстрованих ресурсів входить даний інформаційний ресурс;
- заголовок інформаційного ресурсу;
- зв'язки між інформаційними ресурсами;
- опис вимог щодо конфіденційності інформаційного ресурсу;
- створювач інформаційного ресурсу;
- видавець інформаційного ресурсу;
- короткий опис функцій інформаційного ресурсу;
- електронна адреса інформаційного ресурсу в мережі Internet.

Структура класифікаторів НРІР дозволяє визначати порядок розташування найменувань об'єктів класифікації, класифікаційних угруповань, їх кодових позначень, а також опис прийнятої системи класифікації та кодування. Класифікатори системи будуються за блочним принципом. При реєстрації ІР його власник може вказувати додаткову інформацію, яка дозволить додатково класифікувати ІР. Це дозволить значно полегшити користувачам пошук потрібних ІР. Передбачене подальше розширення системи класифікаторів шляхом введення в структуру бази даних НРІР можливості додавання нових класифікаторів без зміни логічної моделі даних.

У ході розвитку НРІР, зокрема реалізації аналітичних функцій, модель даних НРІР буде доповнюватися новими групами структур даних.

Організація інформаційної взаємодії

До стандартів інформаційної взаємодії у СНІР відносяться:

- стандарти даних;
- стандарти міжвідомчого обміну інформацією;
- стандарти метаданих (і пошуку інформації);
- стандарти безпеки.

Стандарти даних, в свою чергу, містять:

- файли опису структури моделі даних;

— файли з документацією для ІР; документацію щодо моделі даних містять файли з наступними форматами даних: ТХТ, HTML, XML, DOC;

— конфігураційні файли для опису типових структур об'єктів даних для web-інтерфейсу.

Модель даних окремого ІР включає описи інформаційних таблиць даного ІР, описи зв'язків між відношеннями (ключі та індекси) та опис тематичних областей (підмножини відношень з відповідними зв'язками).

Відповідно до стандарту № 19115, метадані представляються у вигляді сукупності тематичних блоків, кожний з яких містить інформацію щодо певних характеристик даних: якість, просторова прив'язка, часова прив'язка, актуальність, оновлюваність і т.д. У якості одного з основних форматів представлення метаданних ISO рекомендує формат XML, який забезпечує представлення різноманітних, складних за структурою і семантикою даних та не залежить від програмної платформи і може передаватися й оброблятися через гетерогенні локальні й глобальні обчислювальні середовища.

До стандартів обміну інформацією входять:

— опис механізмів публікації, пошук та використання ІР, представленого як web-сервіс: UDDI;

— опис інтерфейсів: WSDL;

— способи доступу до ІР: SOAP;

— опис формату обміну: XML;

— опис формату обміну геоінформаційних даних: GML.

Слід зазначити, що названі стандарти інформаційної взаємодії дозволять швидко інтегрувати ІР, доступ до яких може бути отриманий через Інтернет. Інтеграція ІР, які з різних причин не можуть бути опубліковані, вимагає розробки окремих методик, які будуть впровадженні на наступних етапах створення СНІР.

Висновки

У результаті реалізації НРІР очікується підвищення рівня інформаційної безпеки держави; прискорення впровадження «електронного урядування»; підвищення прозорості та відкритості роботи органів державної влади; подальший розвиток національної культури, освіти, науки за рахунок використання новітніх інформаційних технологій. Також передбачається розвиток структури та інфраструктури ресурсів, зростання обсягів послуг, наданих з використанням національних ресурсів; підвищення рівня інформаційної культури та загального освітнього рівня населення; зниження витрат, пов'язаних з використанням ресурсів, за рахунок їх систематизації.

НРІР дозволить поєднати у собі окремі реєстри інформаційних ресурсів, які забезпечуватимуть облік інформаційних ресурсів, згрупованих за певною спільною ознакою.

Для побудови НРІР застосована методологія, яка базується на використанні останніх досліджень у напрямку інтеграції даних.

Реалізація методології дозволить отримати наступні переваги:

— зменшити дублювання інформаційних систем;

— забезпечити більшу сумісність і можливості з інтеграції систем;

- зменшити вартість інформаційних систем;
- оптимізувати процеси виконання функцій організації;
- збільшити продуктивність та ефективність роботи;
- удосконалити керованість систем;
- удосконалити керованість організацій та підтримку процесів перетворень.

Перспективи подальших досліджень щодо формування та використання НРІР як найбільш важливого елемента СНІР є вдосконалення структури БД та пошукового механізму НРІР, а також методологія використання НРІР для забезпечення сертифікації національних ресурсів, розробка гармонізованих з міжнародними стандартами державних стандартів національних ресурсів та їх впровадження.

1. Петров В.В., Нестеренко О.В., Монастирецький М.Г., Шагалов В.Ю. Національні інформаційні ресурси. Проблеми формування, розвитку, управління і використання // Реєстрація, зберігання і оброб. даних. — 2001. — Т. 3, № 2. — С. 38–49.
2. Додонов О.Г., Нестеренко О.В., Бойченко А.В., Бойченко О.А. Формування, інтеграція та використання інформаційних ресурсів органів державної влади // Реєстрація, зберігання і оброб. даних. — 2002. — Т. 4, № 3. — С. 69–75.
3. Нестеренко О.В. Технології інтеграції інформаційних ресурсів інформаційно-аналітичних систем органів державної влади // Науково-технічна інформація. — 2001. — № 4. — С. 3–6.
4. Нестеренко О.В. Єдина державна система електронних інформаційних ресурсів // Науково-технічна інформація. — 2003. — № 4. — С. 3–9.
5. Бойченко О.А. Підхід до організації реєстрації розподілених інформаційних ресурсів // Реєстрація, зберігання і оброб. даних. — 2003. — Т. 5, № 3. — С. 74–80.
6. Еременко Т.В. Каталогизация ресурсов Интернета (Опыт библиотек США) // Науч. и техн. б-ки. — 2002. — № 5. — С. 53–67.
7. <http://www.oclc.org>
8. Дубова Н. SOA: подходы к реализации // Открытые системы. — 2004. — № 6. — С. 37–41.
9. Ньюкомер Э. Web-сервисы. Для профессионалов. — СПб.: Питер, 2003. — 256 с.
10. Макаренко Є.А. Політичні доктрини глобальної інформаційної безпеки // Сб. науч. тр. «Информационные технологии и безопасность». Вып. 1–2. Институт проблем регистрации информации НАН Украины. — К., 2002.
11. http://www.oasis-open.org/news/oasis_news_02_03_05.php
12. Данілін А.В. Стандарти і єдина архітектура інформаційних технологій. — Microsoft Press, 2003. <http://www.microsoft.com/Ukraine/Government/Analytics/IntegrationTechnologies/Standards.msp>

Надійшла до редакції 16.08.2005