

УДК 681.32(075)

Маслянюк П.П., Лісов П.М.

ІНФОРМАЦІЙНО-КОМУНІКАЦІЙНІ СИСТЕМИ ТА ТЕХНОЛОГІЇ ОБРОБКИ ІНФОРМАЦІЙНИХ РЕСУРСІВ

З розвитком інформаційних технологій у всьому світі зростає розуміння переваг, які може надати організаційній структурі інформаційно-комунікаційна система. Тому питання створення інформаційно-комунікаційних систем, які б ефективно працювали з інформаційними ресурсами, є надзвичайно актуальним на сьогодні.

Інформаційно-комунікаційні системи

Тут під „системою” ми розуміємо множину сутностей інформатизації та взаємозв'язків між ними [1]. Згідно чинного законодавства України, інформаційно-телекомунікаційна система – це сукупність інформаційних та телекомунікаційних систем, які у процесі обробки інформації діють як одне ціле [2].

Українське законодавство передбачає декілька визначень інформаційної системи:

- організаційно-технічна система, в якій реалізується технологія обробки інформації з використанням технічних і програмних засобів [2].
- система оброблення даних засобами накопичення, зберігання, оновлення та їх пошуку і відображення [3].

Ці визначення не протирічать, а доповнюють одне-одного. Так друге визначення уточнює, яку саме обробку інформації може проводити інформаційна система.

Основне призначення інформаційно-комунікаційної системи (ІКС) – забезпечити власне комунікації та обробку інформаційних ресурсів організаційної структури.

Класифікацію ІКС можна проводити за багатьма критеріями.

Класифікація за критерієм функціональної повноти виглядає так: [1]

1. ІКС, які задовольняють інформаційні потреби – інформаційні системи. Такі системи забезпечують доступ користувачів (персоналу, клієнтів, тощо) до інформації. Вони повинні забезпечувати зберігання інформації, пошук, швидкий та захищений доступ.
2. ІКС, призначені для підтримки бізнес-процесів (інформаційно-аналітичні системи). Прикладом таких бізнес-процесів можуть бути банківські транзакції.
3. ІКС, призначені для управління організаційною структурою – автоматизовані системи управління. Вони повинні забезпечувати автоматизацію комунікацій, документообігу, контролю за виконанням наказів та розпоряджень, тощо.
4. Інтелектуалізовані системи, до яких відносяться експертні системи, системи оцінки ефективності та прогнозування результатів управлінських рішень, тощо. Такі системи забезпечують виконання окремих інтелектуальних функцій людини.

В ІКС можна виділити такі складові:

- апаратне забезпечення;
- програмне забезпечення;
- інформаційне забезпечення;
- організаційне забезпечення – сукупність методів і засобів роботи персоналу, який здійснює експлуатацію системи;
- кадрове забезпечення;
- правове забезпечення – сукупність норм права, які визначають юридичний статус системи.

Класифікацію ІКС можна також проводити за територіальною ознакою:

1. локальні (системи для малого бізнесу);
2. середні інтегровані;
3. великі інтегровані (глобальні) системи. [4]

За способом організації ІКС діляться на:

1. файл-серверні;
2. клієнт-серверні;
3. на основі інтернет-технологій.

Інформаційні ресурси

Однією з основних складових частин інформаційно-комунікаційної системи є інформаційні ресурси. Дані, інформація, та знання являють собою абстрактні об'єкти [5]. Для роботи з ними необхідна їх матеріалізація у вигляді інформаційних ресурсів. Згідно законодавства України інформаційний ресурс – сукупність документів у інформаційних системах (бібліотеках, архівах, банках даних тощо) [6].

Документ – це упорядкована сукупність даних, інформації та знань, яка надає можливості доступу, передачі, обробки, тощо. Прикладом документа може бути паперовий документ, фільм, комп'ютерний файл, тощо [1].

Середовищем зберігання документів є інформаційно-комунікаційна система, яка забезпечує доступ, обмін інформацією та її обробку. Така система не обов'язково повинна бути комп'ютеризована.

Інформаційні ресурси можна класифікувати за такими ознаками [5]:

1. Приналежністю ресурсу до певної організаційно-технологічної системи (наприклад, бібліотечної мережі, ЗМІ, корпоративної системи);
2. Способом виділення об'єктів обліку (твори, документи, видання, бази даних, інтернет-сторінки, сайти, тощо);
3. Призначенням ресурсу (масова інформація, освіта, бізнес, особиста переписка, тощо)
4. Змістом ресурсу:
 - тематичним,
 - об'єктним,
 - функціональним;

5. Видовим складом ресурсу (видами документів);
6. Джерелом інформації:
 - національне або закордонне,
 - офіційне або неофіційне,
 - тощо;
7. Правовим статусом ресурсу (публічні документи, об'єкти інтелектуальної власності, спам, таємні документи, тощо);
8. Структурним типом ресурсу, що включає:
 - можливість відділення даних від програм та представлення,
 - формати,
 - кодування,
 - інше;
9. Відкритістю ресурсу (відкритий або з обмеженим доступом);
10. Рівнем структурованості:
 - структуровані;
 - неструктуровані.
11. Способом розповсюдження і носієм;
12. Мовою ресурсу;

Крім того, інформаційні ресурси мають ряд характеристик:

1. Характеристики продуктивності:
 - пропускна спроможність,
 - час реакції,
 - час затримки;
2. Характеристики надійності;
3. Характеристики маштабованості;
4. Характеристики розширюваності;
5. Характеристики прозорості;
6. Повна вартість володіння.

Повна вартість володіння інформаційного ресурсу визначається не тільки (і не стільки) вартістю використаних при його створенні апаратних і програмних засобів, а і вартістю інформації, яка в нього закладена.

Комп'ютеризовані інформаційні ресурси можна розділити на такі основні види:

1. файлові системи;
2. бази даних;
3. інформаційні сховища;
4. інформаційні колектори;
5. веб-ресурси.

Файлові системи є найпростішим і найбільш розповсюдженим типом інформаційних ресурсів. Вони дозволяють зберігати дані, інформацію та знання довільного типу та довільної структури.

База даних – це інформаційний ресурс, що дозволяє впорядковано зберігати дані по групі об'єктів, які мають однаковий набір властивостей. Приклади баз даних: FoxPro, Microsoft SQL Server. Класифікація баз даних проводиться за такими критеріями:

1. За характером інформації, яка зберігається:
 - фактографічні (картотеки),
 - документальні (архіви);
2. За способом зберігання даних:
 - централізовані (зберігаються на одному комп'ютері,
 - розподілені (використовуються в локальних і глобальних комп'ютерних мережах);
3. За структурою організації даних:
 - ієрархічні,
 - табличні (реляційні),
 - об'єктні.

Варто зазначити, що при розробці сучасних систем інколи доречно використовувати БД із змішаною структурою [7].

Інформаційні сховища використовуються для зберігання інформації – даних з визначеними взаємозв'язками між ними. Вони представляють собою системи, які, спираючись на бази даних (або інші ресурси) надають користувачам підготовану в певній формі інформацію. Таким чином, інформаційні сховища слід використовувати там, де необхідно отримувати не окремі дані в невеликій кількості а розв'язувати аналітичні задачі, пов'язані з обробкою великої кількості різнорідних даних з різних джерел.

Інформаційні колектори. Використовуються для зберігання знань. Прикладом інформаційного колектора може бути бібліотечна система GreenStone, розроблена в Новозеландському університеті Вайкато у рамках Проекту створення цифрових бібліотек [8]. Такі системи призначені для зберігання неструктурованої інформації, що призводить до ряду їх особливостей.

Веб-ресурси представляють собою ресурси, які зберігаються на різних елементах, розподілених у гетерогенній мережі. Прикладом такого ресурсу є мережа Інтернет. Перевагою даного типу ресурсу є надійність та забезпечення зручного доступу. Недоліком є складність організації управління і пошуку.

Технології обробки інформаційних ресурсів

При розробці ІКС важливою задачею є організація пошуку у інформаційному ресурсі. Можна виділити такі методи пошуку:

1. Повний перебір
2. Індексація
3. Класифікація
4. Тегування

Метод повного перебору є єдиним методом, який можна застосувати для пошуку у неструктурованих інформаційних ресурсах, таких як файлова система. Він полягає в перегляді усіх документів. Недоліком методу є великий час пошуку. Ефективність даного методу в першу чергу залежить від того, наскільки упорядковано зберігає інформацію користувач (наскільки він знає, де

слід шукати). Перевагою методу є те, що його застосування не вимагає попередньої обробки інформації.

Метод індексації полягає в присвоєнні документам (або атрибутам документів) певних індексів. Такі індекси можуть бути впорядковані, що дозволяє швидко знаходити необхідний документ. Даний метод використовується для структурованих інформаційних ресурсів.

Метод пошуку за допомогою класифікації полягає у визначенні певної ієрархічної структури класів документів. Кожен документ відноситься до певного класу. Прикладом може бути універсальний десятковий класифікатор (УДК). Недоліком такого підходу є те, що для пошуку за класифікатором необхідно вміти точно визначити, до якого класу віднести певну інформацію. Крім того, дуже часто інформація може бути віднесена до декількох областей знань. Для вирішення цієї проблеми в системі Yahoo запропоновано наступний підхід – вказівники на підрозділи класифікатора додаються до інших підрозділів. Однак в результаті ми отримуємо хаотичну систему [9].

Тегування полягає в тому, що кожному документу автор ставить у відповідність набір тегів (*ключових слів*), які характеризують його зміст. Недоліком такого підходу є неоднозначність визначення ключових слів. Модифікацією підходу, яку запропонував Ширкі [9] є те, що теги визначають користувачі системи. В такому випадку декілька найбільш часто використаних тегів повинні коректно описувати документ. Однак це не позбавляє тегування його основного недоліку – різні люди схильні використовувати різні терміни для опису одних понять.

Останнім часом почалася розробка засобів для оптимізації пошуку по файловій системі ПК, які дозволяють замінити повний перебір на швидший метод пошуку (Google Desktop Search, „Персональний Поисковик „Яндекса”). Принцип роботи таких засобів полягає в індексації файлів на жорсткому диску. Таким чином, файлова система певним чином структурується. Недоліком таких засобів є те, що вони підтримують обмежену кількість форматів, так як для того, щоб індексувати лише ті значення, за якими доречно вести пошук (і

відповідно отримати список індексів прийнятної розміру), необхідно знати структуру файлів.

Пошук інформації досить часто організовується ієрархічно, тобто ведеться пошук у результатах попереднього пошуку. У таких випадках пошук буде проводитися швидше, якщо зберігати проміжні результати пошуку. Таким чином можна не починати пошук спочатку. Однак такий підхід складно застосувати у деяких пошукових системах, так як він вимагає додаткових ресурсів. Тому досить часто ієрархічний пошук реалізується простим додаванням нової умови до попередніх (Наприклад, у пошуковій системі Google).

Використання індексації, класифікації та тегування вимагає первинної обробки інформації. Така обробка включає аналіз інформації з метою її класифікації. Таким чином неструктуровану інформацію можна певним чином структурувати. Аналіз документів може проводитися за змістом або за діями користувачів. Аналіз за діями користувачів дозволяє з'ясувати, наскільки документ відповідає визначеним для нього ключовим словам чи класам. Прикладом може бути індексація веб-ресурсів роботами пошукових систем. Такі роботи оцінюють релевантність сторінок. Для фільтрування „небажаних” сторінок можна використовувати сервіси, які дозволять користувачам відмічати їх, так що в подальшому вони не будуть включатися в пошук.

Висновки

Таким чином, ми можемо однозначно стверджувати що класифікація інформаційних ресурсів та ІКС надає розробнику можливості для обґрунтування вибору необхідних інформаційних ресурсів та класу ІКС для кожного конкретного проекту інформатизації організаційної структури.

У відповідності із визначеним класом бізнес-процесів можна визначити і засоби та технології обробки інформаційних ресурсів.

Чіткий перелік засобів інформатизації дозволяє вже на етапі створення інформаційно-комунікаційної системи організаційної структури визначити не

тільки значення функціональних характеристик а і повну вартість володіння інформаційно-комунікаційної системи. Це дозволяє визначитись, чи варто вкладати певні фінансові ресурси в проект інформатизації, чи ні.

Література

1. *Маслянюк П.П.* Концепція інформатизації корпоративних структур // Наукові вісті НТУУ „КПІ”. 2003, №3 – с 510-525.
2. *Закон України «Про захист інформації в інформаційно-телекомунікаційних системах»* (від 31.052005р. М2594-ІУ).
3. *Постанова Кабінету Міністрів України від 20.01.1997 р. № 40 “Про затвердження Концепції створення Єдиної державної автоматизованої паспортної системи”.*
4. *Черненко М., Слепцов С.* Принципы классификации управленческих информационных систем // Корпоративные системы – 2004 №1
5. *Маслянюк П.П. , Ліссов П.М.* Проблеми і технології продукування інформаційних ресурсів // Всеукраїнська науково-практична конференція «Сучасні тенденції розвитку інформаційних технологій і науці, освіті, економіці», Луганськ, 11-14 грудня 2006.
6. *Закон України “Про Національну програму інформатизації”.*
7. *Сиротюк О.* Особенности проектирования современных баз данных / www.computerworld.com.ua
8. *Loots M., Camarzan D., Witten I.* Greenstone Digital Library. Developer’s guide.
9. *Clay Shirky* Ontology is Overrated: Categories, Links, and Tags