

# АРХІТЕКТУРА ТА КОМПОНЕНТИ КОМП'ЮТЕРНИХ СИСТЕМ

УДК 004.9

П. Жежнич, О. Сопрунюк,  
Національний університет "Львівська політехніка",  
кафедра інформаційних систем та мереж

## КОНСОЛІДАЦІЯ ВІДКРИТИХ ІНФОРМАЦІЙНИХ РЕСУРСІВ У ТУРИСТИЧНІЙ СФЕРІ

© Жежнич П., Сопрунюк О., 2013

Зосереджено увагу на ролі консолідації даних у туристичній сфері, наголошено на необхідності використання наявних програмних засобів для обробки туристичної інформації, описано проблему структурування даних та зведення їх до уніфікованого формату, а також виділено етапи формування туристичної документації.

**Ключові слова:** туристична документація, інтеграція даних, консолідована інформація, структурованість джерел, туристична сфера, відкриті веб-ресурси

**The article describes the role of data consolidation in the tourism sphere, the need to use existing software tools to handle tourist information is highlighted, the problem of data resources structuring and unified format creation are emphasized, and the scheme of the formation of tourism documentation is presented.**

**Key words:** tourism documentation, integration, consolidated information, information resources structure, tourism sphere, opened web-resources.

### Вступ

Розвиток туризму сприяє збільшенню прибутків країни, а отже, її економічному розвитку. У зв'язку з цим надання якісних туристичних послуг в умовах сьогодення є однією з актуальних проблем туристичного бізнесу. Постійне ж зростання попиту на якісний туристичний продукт потребує удосконалення та підвищення його якості шляхом формування туристичної документації з урахуванням міжнародних стандартів її оформлення, наприклад, міжнародного стандарту ISO 9001-2001.

Важливою є діагностика та дослідження джерел формування туристичної документації, зважаючи на їх структурованість. Адже від типу джерела залежить подальший вибір алгоритмів аналізу даних.

### Постановка проблеми

Сьогодні ведення бізнесу в сфері туризму ґрунтується переважно на використанні відповідних інформаційних систем, що забезпечують оперативну обробку різноманітної інформації, необхідної у процесі прийняття рішень.

Характерною особливістю опрацювання туристичної інформації та відповідно формування туристичної документації є необхідність надання користувачу компактної та достовірної інформації про різноманітні послуги на основі взаємодії між туристичною фірмою та споживачем туристичного продукту у реальному часі.

Постійне зростання попиту на туристичний продукт вимагає удосконалення туристичної документації з урахуванням міжнародних стандартів. Важливими є питання структуризації та консолідації туристичної інформації, що дає можливість оперативно оптимізувати процеси прийняття рішень про надання користувачам широкого спектра туристичних послуг.

Тому роль консолідації даних для розвитку сучасної туристичної сфери є актуальною задачею інформаційного забезпечення бізнес-процесів.

### **Аналіз наукових досліджень**

Сьогодні дослідники велику увагу приділяють особливостям використання інформаційних технологій та програмних засобів у діяльності підприємств туристичної сфери [18]. Існує багато праць, в яких описано проблеми аналізу інформації різного типу, проблеми структурованості даних та зведення їх до єдиного формату [9, 10, 15, 16]. Автори [20] досліджують проблеми інтеграції даних з використанням ETL-технологій (extraction – екстракція, transformation – перетворення, loading – завантаження).

У працях [11, 15, 20, 21] описано три типи інтеграції: консолідація, федералізація і тиражування. Консолідація даних полягає у формуванні єдиної інтегрованої бази даних (БД) з різноманітних початкових джерел. За допомогою цього методу можна обробити і перетворити значні обсяги розрізної інформації, провести так звану реструктуризацію ресурсів. Прикладом консолідації є технології ETL та управління корпоративними даними ECM (enterprise content management).

У випадку федералізації даних існує віртуальна картина первинних джерел інформації, яка приймає згенерований бізнес-додатками запит і відбирає дані з первинних джерел, інтегруючи їх так, щоб вони відповідали запиту, і потім дані передаються до бізнес-додатків. Доступ до початкових джерел здійснюється через метадані. Прикладом є технологія інтеграції корпоративних даних EIT (enterprise information integration) [11, 15, 20, 21].

Якщо здійснюється лише копіювання джерел даних з урахуванням оновлення інформаційного наповнення і відповідно перенесення їх в інше місце внаслідок дублювання початкового джерела, то відбувається тиражування даних. Прикладом є технології інтеграції корпоративних додатків EAI (enterprise application integration) та тиражування корпоративних даних EDR (enterprise data replication).

На сучасному етапі тотальної інформатизації спостерігається посилення впливу інформаційних процесів на суспільство. Це сприяє розвитку потужного інформаційного середовища. Як результат, туристичні організації накопичують у процесі своєї діяльності великі обсяги даних і стикаються з проблемами їх оперативної обробки [9, 15, 16, 21].

У працях [2, 7] описано проблему неоднорідності форматів даних, неповноти отриманої інформації та запропоновано метод використання концептуальних моделей, що одержують інформацію з контексту для прийняття рішень із обслуговування туриста в умовах неповноти інформації. Крім того, досліджують процеси створення документації, використання DFD-діаграм для опису туристичних процесів, відбору даних та їх оброблення за допомогою відповідних програм [3, 11].

Однак немає узагальненої схеми формування туристичної документації, не визначено правил та параметрів аналізу структурованих та неструктурованих джерел.

### **Цілі статті**

Інтенсивний розвиток туристичної галузі, затребуваність туристичного продукту вимагають розвитку інформаційних технологій збору й обробки інформації з метою оперативного і надійного надання туристичних послуг користувачам.

Структурування і консолідація інформації істотно впливають на ефективність і оперативність задоволення потреб користувачів і дають змогу оптимізувати процеси туристичного бізнесу.

Метою роботи є:

- встановити роль структурування та консолідації інформації для розвитку туристичної сфери;
- обґрунтувати вибір відкритих веб-ресурсів для формування туристичної документації;
- з'ясувати необхідність автоматизації обробки інформації та вибору правил і параметрів аналізу структурованих та неструктурованих джерел;
- визначити етапи формування туристичної документації.

### **Використання відкритих веб-ресурсів для консолідації туристичної інформації**

В умовах постійного зростання обсягів інформації, розрізненості її джерел, наявності різних систем обробки даних з метою їх подальшої узгодженості, динамічності, доступності на основі

створення єдиного інформаційного простору, інтеграція даних відіграє важливу роль для розвитку різних галузей, зокрема туристичної сфери.

Для ефективної діяльності та отримання якісних результатів туристична установа повинна створити необхідні умови для оперативної обробки всієї доступної інформації з метою приведення її до єдиного стандартного вигляду, щоб надалі мати змогу обробляти інформацію щодо надання відповідних туристичних послуг. Для цього потрібно враховувати як процеси відбору даних, так і процеси їх перевірки, очищення та перетворення для подальшої інтеграції з використанням різних методів й інструментів. Обґрунтований вибір технології інтеграції даних і реалізація її за допомогою програмних засобів нададуть фірмі можливість адекватного контролю інформації [9].

Федералізацію даних недоцільно використовувати в сфері туризму, оскільки доволі складно створити віртуальну базу з великою охоплюваністю інформації, і це може спричинити технічні проблеми з обслуговування великої кількості користувачів.

Інтеграція даних за допомогою тиражування створює проблеми з підвищенням авторитетності ресурсу, тому що під час копіювання створюються ідентичні або подібні ресурси, що може спричинити певне розсіювання інформації, а не зосередження її в єдиному цілісному ресурсі. До того ж тиражування даних придатне для однорідних джерел інформації, тому використовувати його в процесі формування туристичної документації на основі відкритих веб-ресурсів недоцільно.

Існує також гібридний підхід до інтеграції даних, який поєднує методи консолідації, федералізації та тиражування (рис. 1). Прикладом застосування гібридного типу можна назвати інтеграцію даних про клієнта (customer data integration, скор. CDI) з метою надання повної інформації про нього. Такий підхід також використовують в туризмі. До прикладу, загальні дані про туристів зберігають в єдиній консолідованій базі даних, а дані про замовлення туру можуть бути федералізовані. Якщо ж споживач туристичного продукту оновлює певні дані про себе, то таке оновлення внаслідок розподілення даних відправляється в інші системи [11, 15, 20, 21]. Однак, враховуючи недоліки федералізації та тиражування, гібридний метод не варто використовувати.

Отже, у процесах формування туристичної документації на основі відкритих веб-ресурсів доцільно використовувати консолідацію даних, оскільки інформаційні веб-джерела характеризуються значною розрізненістю даних і є неоднорідними за структурою, а за допомогою механізму консолідації можна сформувати готовий цілісний туристичний ресурс (рис. 1).

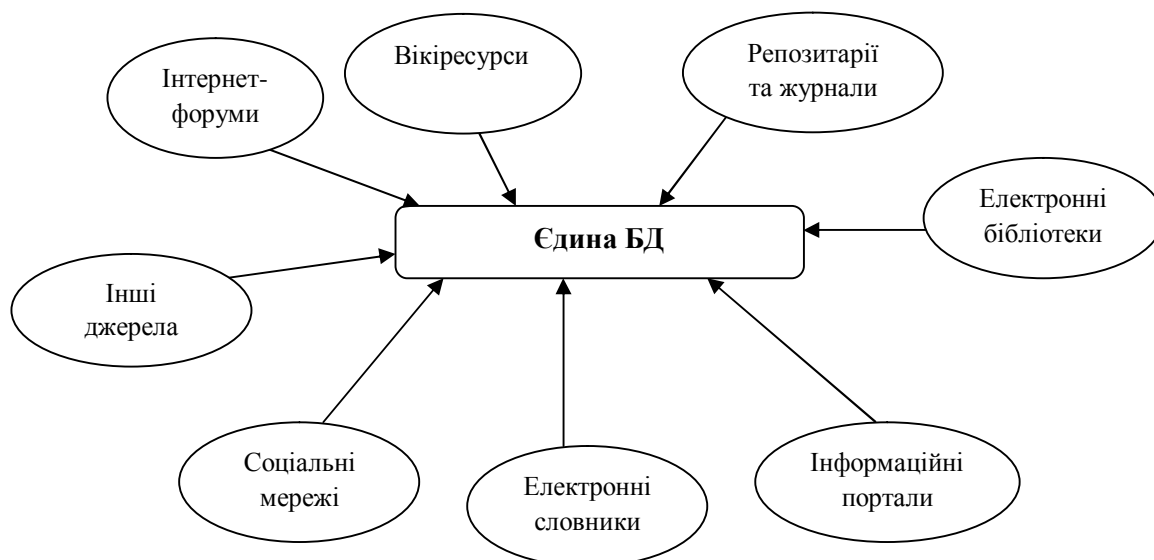


Рис. 1. Загальна схема консолідації джерел туристичної інформації

У процесах формування туристичної документації доцільно використовувати відкриті веб-ресурси. Доступність необхідної інформації відіграє важливу роль для оперативності її обробки. Вільний доступ до різноманітних джерел сприяє більшій достовірності отриманих результатів, а

відкриті джерела – швидшому поширенню інформації, можливості редагування, характеризуються вищим ступенем надійності та предметною охопленістю. Відкриті веб-ресурси сприяють швидкому доступу до різних джерел даних. За допомогою пошукових систем можна зібрати велику кількість різноманітної туристичної інформації.

Отже, сьогодні всі туристичні організації оперують великими обсягами даних, які необхідно консолідувати. Консолідація інформації потребує використання єдиного формату даних. Механізм роботи з даними можна подати так:

- спочатку потрібно проаналізувати наявні джерела інформації, поділити їх на структуровані й неструктуровані, згрупувати, вибрати джерела даних;

- вибрати необхідну інформацію за допомогою інструментів пошуку;

- відібрані дані зберегти, виконати їх структурування, звести до єдиного формату з метою подальшої автоматизованої обробки;

- збережені дані опрацювати (для цього необхідно мати повний набір можливостей, включаючи пошук, перегляд, редагування й перетворення), сформувані набір правил та параметрів аналізу як структурованих, так і неструктурованих джерел і застосувати їх в роботі, процес можна задокументувати;

- для обміну даними повинні використовуватися різні механізми, наприклад, архівування та передавання мережами, публікація тощо [8, 13, 19, 22].

Проте веб-дані різного формату, отримані з різнорідних джерел, зберігаються в розрізних системах, і для виконання широкого спектра дій їх необхідно інтегрувати, виконуючи консолідацію одержаних результатів, зокрема, наприклад, за допомогою XML-технологій.

#### **Аналіз веб-джерел для формування туристичної документації**

Отже, одним з основних джерел інформації є Інтернет. Це інформаційне середовище, основним базисом якого є відкриті веб-ресурси, доступне для широкого загалу користувачів. Звичайно, велика частина цих ресурсів не є цінною з погляду вірогідності та достовірності інформації, тому потрібно завжди здійснювати велику кількість інформаційно-аналітичної роботи, щоб відфільтрувати отримані дані й уникнути дезінформації – поширення запевнено неправдивих даних з метою введення користувачів у оману.

Тому необхідно завжди аналізувати отримані дані із урахуванням всіх правил аналізу відповідно до типу джерела інформації, зіставляти інформацію з різних джерел, переглядати як позитивні, так і негативні відгуки, орієнтуватися на престижність форуму, враховувати суб'єктивні характеристики джерела, ідентифікувати рівень зацікавленості інформатора в поширенні інформації.

Для того щоб правильно проаналізувати дані, потрібно передусім розділити структуровані та неструктуровані джерела і визначити параметри та правила їх аналізу, створити сховища структурованих та неструктурованих джерел туристичної інформації, здійснити структурування неструктурованих даних і в результаті агрегування туристичних даних сформувані консолідований інформаційний ресурс туристичної документації (рис. 2).

Гетерогенність даних створює необхідність їх поділу на структуровані (наприклад, формалізовані тексти, інформація з баз даних тощо) та неструктуровані (наприклад, факси, електронна пошта, голосова пошта, різноманітні файлові документи, презентації тощо), з метою подальшого вибору адекватного набору правил обробки різнорідної інформації [19].

Дуже велика кількість інформації є неструктурованою, тому потрібно вибирати ефективні інструменти її обробки, зводячи інформацію до узгодженого вигляду для подальшого аналізу та систематизації.

Проблема структурування даних є актуальною та складною, тому її потрібно вирішувати комплексно, застосовуючи різні підходи, концепції, технології та інструменти обробки.

Будь-яке джерело даних пізніше обробляється, над ним здійснюються маніпуляції й інформація набуває певного вигляду, підлаштовується під формат і далі бере участь в обробці та дослідженнях, зберігаючись в певному структурованому вигляді.

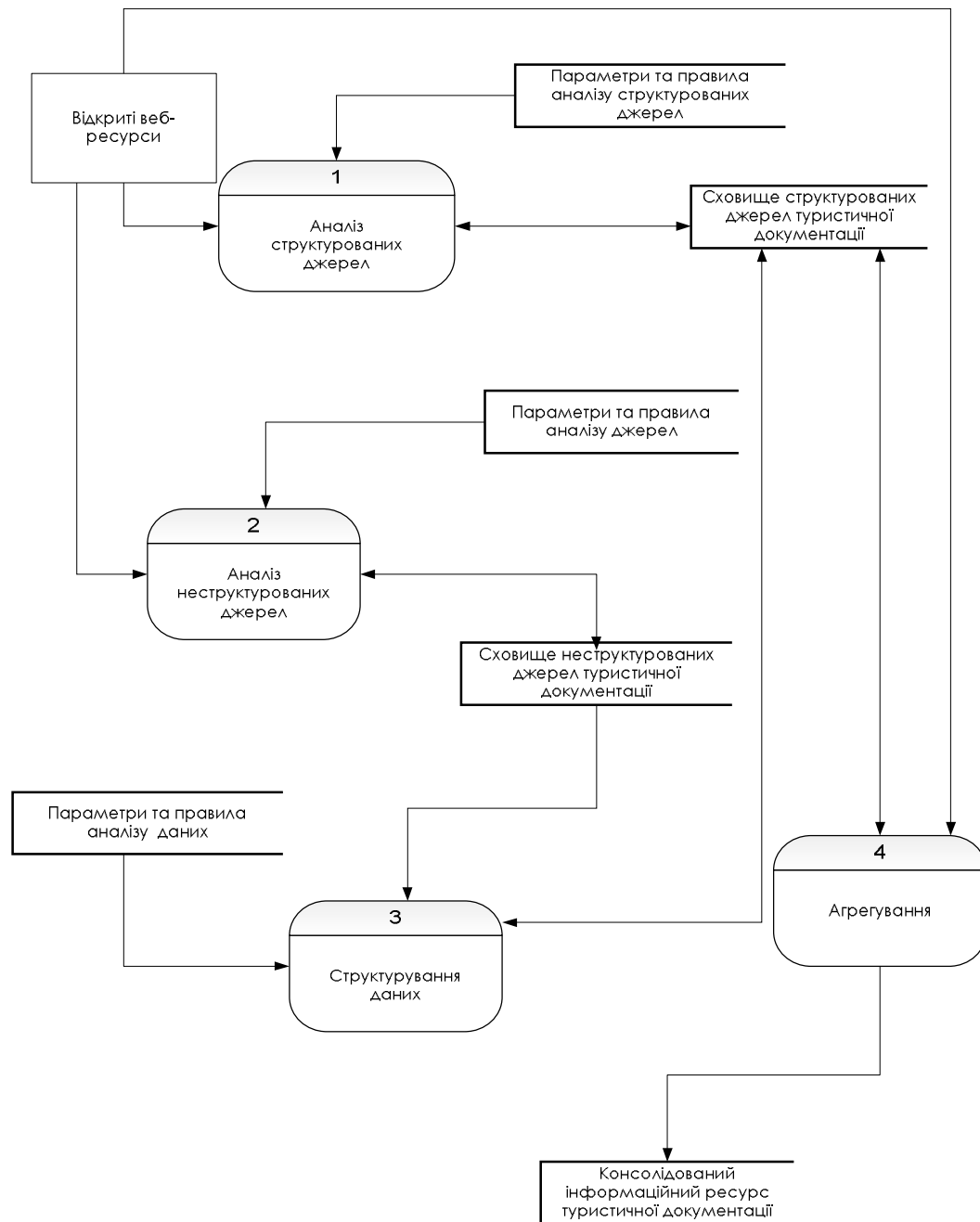


Рис. 2. Схема аналізу джерел інформації для формування туристичної документації

Видобувають інформацію за допомогою певних інструментів екстракції даних, які дають змогу подати її уніфіковано, в єдиному прийнятому форматі (рис. 3) [22].

В будь-якій туристичній організації важливою складовою діяльності є обробка інформації, яка повинна здійснюватися автоматизовано, щоб пришвидшити процеси маніпуляції даними. Для цього дані повинні бути зведені до уніфікованого формату.

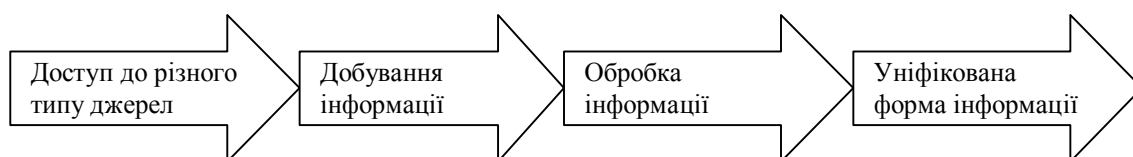


Рис. 3. Технологія формування туристичної документації

Сьогодні для екстракції даних можна використовувати спеціальні інформаційно-аналітичні системи, які перед початком обробки інформації можливо налаштувати на певні джерела даних. Відповідно до типу джерела формують певний набір правил обробки даних відповідно до структури і задають їх системі. Для того щоб обробити неструктуровані джерела, необхідно використовувати інформаційні тезауруси, логічні умови, словники предметної області [8]. Обробка неструктурованих джерел потребує більше часу, оскільки їх необхідно спершу звести до структурованого вигляду.

Для обробки потрібно також вказувати критерії відбору даних. Потім інформацію із структурованих джерел треба звести до вигляду внутрішньої структури інформаційних об'єктів вибраної інформаційно-аналітичної системи і зберегти в базі даних.

Потрібно вибирати такі програмні засоби, які дадуть змогу консолідувати інформацію з різних джерел, обробити її, перетворивши в єдиний формат для візуалізації, збереження та подальшої маніпуляції, створення певної моделі даних для аналізу.

Для того щоб згрупувати джерела за структурованістю, потрібно чітко окреслити ті сфери, інформація про які важлива для туриста. Це допоможе виділити основні тематичні напрями вилучення інформації і надалі систематизувати основні правила аналізу даних – як структурованих, так і неструктурованих.

З огляду на це мають бути задокументовані такі основні напрями, як: умови проживання, туристичні об'єкти, харчування, розваги, екскурсійні маршрути, бронювання, транспортні перевезення, інформація про місце відпочинку, страхування, туристичні формальності, графік туру, готельні послуги тощо. Повна та адекватна інформація про кожен з напрямів забезпечить належну якість надання послуг.

### Етапи формування туристичної документації

Будь-який документ має свій життєвий цикл, який складається з таких етапів, як: народження, встановлення, публікація, архівування. На основі цього формування туристичної документації складається з таких етапів: аналітичний, пошуковий, обробка інформації, створення документації та маніпуляція документацією (рис. 4).



Рис. 4. Етапи формування туристичної документації

На аналітичному етапі необхідно здійснити інформаційно-аналітичний огляд стану проблеми, щоб зрозуміти, якої інформації не вистачає, якого типу документацію потрібно створити і розробити проект туристичної документації. Для цього необхідно виявити прогалини в туристичній інформації на основі аналізу думок, запитів, скарг та пропозицій споживачів туристичного продукту, проаналізувати вимоги сучасного туристичного ринку, а також сформулювати критерії створення якісної туристичної документації.

На пошуковому етапі, для того щоб знайти потрібну туристичну інформацію, необхідно вибрати системи пошуку інформації для формування інформаційного запиту, простежити, чи знайдена інформація відповідає параметрам запиту, і здійснити відбір інформації, перед тим вибравши технологію, методи та засоби екстракції інформації.

На наступному етапі розробляють шаблони документації, наповнюючи їх обробленою інформацією і переводячи створену документацію в електронну форму.

Створену документацію затверджують після тестування та верифікації на придатність і вводять в експлуатацію.

Формування туристичної документації – це складний процес, який потребує затрат часу, залучення висококваліфікованих фахівців, упровадження електронних інформаційних технологій обробки інформації, засобів її групування, ідентифікації джерел туристичної інформації, логічного аналізу матеріалів, вибору методів екстракції даних, зведення даних до єдиного доступного для споживача туристичного продукту формату за допомогою консолідації обробленої інформації.

### **Висновок**

Консолідація як один із типів інтеграції дає змогу перетворювати розрізнені початкові дані на єдиний ресурс, що забезпечує цілісність, узгодженість та надійність інформаційного наповнення.

Одним із основних джерел інформації в туризмі є відкриті веб-ресурси з вільним доступом користувача.

Важливою складовою діяльності будь-якої туристичної організації є автоматизована обробка інформації, яка потребує зведення даних до уніфікованого формату.

Актуальною і водночас складною є проблема структурування отриманих даних. Для перетворення неструктурованих текстів на інформацію, придатну для формування туристичної документації, слід застосовувати відповідні комп'ютерні алгоритми і програми обробки даних. Необхідно сформулювати набір правил та параметрів аналізу як структурованих, так і неструктурованих джерел з метою адекватної обробки інформації.

Визначено етапи формування туристичної документації, що допоможе оптимізувати туристичну діяльність.

1. Баласанян В. Автоматизация делопроизводства и документооборота: введение в проблему / В. Баласанян // Рынок ценных бумаг. – 1998. – № 16. – С. 25–27. 2. Буров Є.В. Опрацювання контекстних даних в інтелектуальному туристичному порталі / Буров Є.В., Королук Ю.В. / Вісник Нац. ун-ту “Львівська політехніка” “Комп’ютерні науки та інформаційні технології”. – Львів: Вид-во Львівської політехніки, 2010. – № 672. – С. 228–236. 3. Вендров А. М. CASE-технологии. Современные методы и средства проектирования информационных систем. – М: Финансы и статистика, 1998. – 176 с. 4. Гавриш С. Система автоматизації діловодства та електронного документообігу / С. Гавриш // Довідник секретаря та офіс-менеджера. – 2007. – № 2. – С. 25–31. 5. Жежнич П.І. Аналіз сучасних засобів автоматизації роботи з електронною документацією / П.І. Жежнич, О.О. Сопрунюк, О. Марчик // Вісник Національного університету “Львівська політехніка”: Комп’ютерні науки та інформаційні технології. – Львів: Вид-во Нац. ун-ту “Львівська політехніка”, 2012. – № 732. – С. 395–398. 6. Жежнич П. І. Основні підходи до організації електронного організаційно-розпорядчого документообігу / П.І. Жежнич, О.О. Сопрунюк, О. Марчик // Інформація, комунікація, суспільство (ІКС-2012): матеріали I Міжнародної

наукової конференції ІКС-2012. – Львів: Вид-во Нац. ун-ту “Львівська політехніка”, 2012. – С. 34–36.

7. Захарова К.В. Використання сучасних інформаційних технологій в туристичній галузі Криму / К.В. Захарова // Економіка Криму / наук.-пр. журнал. – Сімферополь: Таврійський національний університет ім. В.І. Вернадського, 2009. – № 27. – С. 55–58.

8. Извлечение знаний из корпоративных источников информации [Электронный ресурс]. Технологии разведки для бизнеса, 2011. – Электрон. дан. – Режим доступа: <http://www.it2b.ru/blog/arhiv/511.html>.

9. Кунанець Н.Е. Вступ до спеціальності “Консолідована інформація”: навч. посіб. / Н.Е. Кунанець, В.В. Пасічник. – Львів: Вид-во Нац. ун-ту “Львівська політехніка”, 2010. – 196 с. – (Серія “Консолідована інформація”)

10. Ланде Д.В. Системи моніторингу, витягу фактів, побудови зв'язків на основі аналізу неструктурованих текстів / Д.В. Ланде, В.М. Фурашев // Правова інформатика. – К.: Науково-дослідний центр правової інформатики, 2010. – № 2(26). – С. 3–9.

11. Литвин В.В. Проектування інформаційних систем: навч. посіб. / В. В. Литвин, Н.Б. Шаховська. – Львів: Вид-во Нац. ун-ту “Львівська політехніка”, 2010. – 210 с.

12. Мельниченко С.В. Інформаційні технології в туризмі: теорія, методологія, практика: монографія. – К.: Київ. нац. торг.-екон. ун-т, 2008. – 493 с.

13. Неподнятая целина. Как использовать неструктурированную информацию в бизнесе [Электронный ресурс]. – Проект iBusiness, 2012. – Режим доступа: <http://i-business.ru/blogs/17349>

14. Орлов П.І. Інформація та інформатизація: Нормативно-правове забезпечення: наук.-практ. посіб. – Х.: Вид-во Ун-ту внутр. справ, 2003. – 724 с.

15. Пасічник В.В., Шаховська Н.Б. Сховища даних: навч. посіб. – Львів: Магнолія 2006, 2007.

16. Пелецишин А.М. Пошук, консолідація та аналіз даних у глобальній системі World Wide Web / А.М. Пелецишин // Вісник Нац. ун-ту “Львівська політехніка”: [збірник наукових праць] – Львів: Вид-во Нац. ун-ту “Львівська політехніка”. – 2003. – № 489: Інформаційні системи та мережі. – С. 233–247.

17. Пелецишин А.М. Роль прогалін у документації в процесах виготовлення та експлуатації продукції / А.М. Пелецишин, О. Ю. Тимовчак-Максимець, О.О. Сопрунок // Вісник Нац. ун-ту “Львівська політехніка”: Комп'ютерні науки та інформаційні технології. – Львів: Вид-во Нац. ун-ту “Львівська політехніка”, 2011. – № 719. – С. 17–22.

18. Скопень М.М. Комп'ютерні інформаційні технології в туризмі / М.М. Скопень.– К.: Кондор, 2005. – 301

19. Управление неструктурированной информацией [Электронный ресурс]. Intelligent enterprise, 2004. – Электрон. дан. – Режим доступа: <http://www.ietag.ru/opinions/detail.php?ID=17860>

20. Шаховська Н.Б. Аналіз інформаційних систем оброблення даних туристичної сфери / Н.Б. Шаховська, Д.І. Угрин // Науковий вісник НЛТУ України: зб. наук.-техн. праць. – Львів : НЛТУ України. – 2008. – Вип. 18.10. – С. 258–263.

21. Шаховська Н. Б., Пасічник В. В. Сховища та простори даних: монографія.– Львів: Вид-во Львівської політехніки, 2009. – 244 с.

22. Clarabridge Document Analytics – инновационный инструмент для ситуационного анализа национальных проектов [Электронный ресурс]. C-new ,2006. [Электронный ресурс]. – Режим доступа:<http://www.cnews.ru>.