

Міністерство освіти і науки України  
Національний університет “Львівська політехніка”

На правах рукопису

Марків Оксана Олександрівна

УДК 81'322.2:[004.774+004.773.7+004.624]

**Лінгвістичне забезпечення формування туристичної  
документації на основі відкритих веб-ресурсів**

Спеціальність 10.02.21 – Структурна, прикладна і математична лінгвістика

Дисертація на здобуття наукового ступеня

кандидата технічних наук

Науковий керівник:  
д.т.н., проф. кафедри соціальних  
комунікацій та інформаційної діяльності  
Жежнич Павло Іванович

Львів - 2017

## Зміст

<b>Зміст .....</b>	<b>2</b>
<b>Перелік умовних скорочень .....</b>	<b>4</b>
<b>Вступ .....</b>	<b>5</b>
<b>Розділ 1. Аналіз особливостей туристичної діяльності.....</b>	<b>14</b>
1.1. Аналіз тенденцій розвитку туристичної галузі в сучасних умовах. 14	
1.2. Аналіз використання відкритих веб-ресурсів у туристичній діяльності .....	28
1.3. Аналіз використання XML-технологій для формування туристичної документації.....	35
1.4. Обґрунтування вибору методу оцінювання якості туристичної документації.....	39
1.5. Висновки до розділу .....	47
<b>Розділ 2. Розробка інформаційної моделі туристичної документації .48</b>	
2.1. Розробка загальної моделі туристичної документації .....	48
2.2. Розробка моделі елементарних одиниць інформації у туристичній документації.....	58
2.3. Класифікація туристичної документації за особливостями загальної моделі.....	62
2.4. Розробка моделі інформаційних джерел формування туристичної документації.....	68
2.5. Розробка методу лінгвістичного співставлення атомарної ситуації з елементами туристичної документації .....	73
2.6. Висновки до розділу .....	85
<b>Розділ 3. Розробка процедур комп'ютерно-лінгвістичного формування туристичної документації.....</b>	<b>87</b>
3.1. Структуризація процесу формування туристичної документації з використанням відкритих веб-ресурсів .....	87

3.2. Розробка процедури вибору відкритих веб-ресурсів для формування туристичної документації.....	89
3.3. Розробка процедур лінгвістичного співставлення вмісту відкритих веб-ресурсів з елементами туристичної документації .....	97
3.4. Розробка процедур виявлення прогалин у туристичній документації	112
3.5. Розробка процедури виявлення некоректної інформації у туристичній документації.....	119
3.6. Висновки до розділу .....	124
<b>Розділ 4. Побудова системи комп'ютерно-лінгвістичного формування туристичної документації .....</b>	<b>126</b>
4.1. Розробка показників якості туристичної документації .....	126
4.2. Розробка системи комп'ютерно-лінгвістичного формування туристичної документації.....	132
4.3. Аналіз практичного впровадження результатів досліджень.....	142
4.4. Висновки до розділу .....	147
<b>Висновки .....</b>	<b>149</b>
<b>Література.....</b>	<b>151</b>
<b>Додаток А .....</b>	<b>170</b>
<b>Додаток Б .....</b>	<b>178</b>
<b>Додаток В .....</b>	<b>182</b>

## Перелік умовних скорочень

Скорочення, термін, позначення	Пояснення
БД	База даних
ВР	веб-ресурс
ВВР	Відкритий веб-ресурс
ДТО	Дія над туристичним об'єктом
ІН	Інформаційне наповнення
ІП	Інформаційний продукт
ІР	Інформаційний ресурс
ІС	Інформаційна система
ІСТП	Інтерес споживача туристичного продукту
ПЗ	Програмне забезпечення
СТП	Споживач туристичного продукту
ТД	Туристична документація
ТІ	Туристична інформація
ТО	Туристичний об'єкт
ТП	Туристична послуга
ISO	Міжнародна організація із стандартизації
XML	e <u>X</u> tensible <u>M</u> arkup <u>L</u> anguage – мова розширюваної розмітки
XSLT	<u>X</u> SL <u>T</u> ransformations – перетворення XML-документів
XSL	e <u>X</u> tensible Style Language – мова шаблонів розмітки
WWW	<u>W</u> orld <u>W</u> ide <u>W</u> eb – всесвітня павутина

## Вступ

**Актуальність теми.** На сьогодні туризм є одним із перспективних напрямів як внутрішньо-, так і зовнішньоекономічної діяльності країни та її регіонів. Аналіз туристичної діяльності показує, що за останні 25 років кількість міжнародних туристів у світі зросла майже втричі. В Україні динаміка росту міжнародних туристів до 2014 була ще стрімкішою.

Основним джерелом інформації про туристичну діяльність є звичайно Всесвітня павутина WWW. Експоненціальний ріст обсягів інформації, що є доступною у WWW, приводить до того, з кожним роком туристична інформація потребує все більших затрат часу та людських ресурсів на її опрацювання. Тому, очевидно, актуальними є створення підходів до автоматизованого опрацювання інформації про туристичну діяльність, яка доступна у WWW.

Дослідження щодо використання новітніх інформаційних технологій у діяльності підприємств туристичної сфери і туризмі в цілому здійснювали С.Арімов, В.Балута, П.Бурцев, А.Галиновський, О.Губанова, В.Гуляєв, А.Демаш, М.Єфремова, І.Зорін, І.Калашников, В.Квартальнов, Д.Купінський, Ю.Міронов, М.Морозов, Н.Морозова, Т.Новгородцев, Н.Плотнікова, М.Окороков, М.Робсон, М.Скопєнь, О.Тоффлер, Ф.Уллах, М.Хайдеггер, А.Шлевков, С.Чупров, Є.А.Фалько та інші. Проблемам загального управління туристичними підприємствами (в т.ч. з використанням інформаційних технологій) присвячено праці таких вчених як М.І.Кабушкін, С.В.Мельниченко, О.В.Апілат, Л.М.Гопкало та інших, а проблеми логістики туристичних послуг досліджували такі науковці як Г.І.Михайліченко, Н.В.Задорожна, О.Я.Кобилух, Н.Й.Коніщева, І.Г.Смирнов, В.В.Іванов та інші. Моделюванням об'єктів туристичної галузі за допомогою інформаційних технологій присвячено роботи Я.І.Виклюка [17]. Однак ці дослідження не розглядають на науковому рівні питання використання інформаційного наповнення веб-сайтів, побудованих за

концепцією Веб 2.0, для розвитку туристичного бізнесу. Це приводить до того, що туристичні компанії не належним чином використовують інформаційні ресурси WWW, які формуються як надавачами туристичних чи супутніх послуг, так і їх споживачами.

Основна проблема масового використання інформаційного наповнення веб-сайтів для підтримки туристичної діяльності пов'язана з великою неоднорідністю структури веб-сторінок і слабкою структурованістю їхніх фрагментів (кількість структурованої інформації оцінюється до 20%) [15, 68]. Тому акцент у першу чергу необхідно робити не на механістичному опрацюванні інформаційного наповнення веб-сторінок, а на лінгвістичних особливостях цього інформаційного наповнення з урахуванням таких чинників:

- інформація про туристичні послуги є розпорошеною на різних веб-сайтах;
- споживачі туристичних послуг надають окрему необхідну туристичну інформацію навіть краще, ніж надавачі туристичних послуг;
- туристична інформація вимагає постійної підтримки щодо повноти та достовірності.

Фактично туристична діяльність потребує якісного інформаційного забезпечення, яке формується як певний інформаційний продукт, який повинен задовольняти потреби споживача [70, 140]. Цей інформаційний продукт становить туристичну документацію, що містить інформацію з різних веб-ресурсів, згруповану за розділами.

Туристична документація характеризується значною інформаційною охоплюваністю. Щоб її наповнити, потрібно опрацювати велику кількість джерел туристичної інформації. Оскільки джерела туристичної інформації у відкритому доступі характеризуються розрізненістю форматів, потрібно обрати оптимальний метод інтеграції туристичної інформації для створення єдиного цілісного туристичного ресурсу. Це дозволить покращити технологію формування туристичної документації, а, отже, забезпечити якісне надання

інформації про туристичну послугу.

Доцільність використання відкритих веб-ресурсів в процесах формування туристичної документації зумовлена доступністю необхідної інформації, що відіграє важливу роль для оперативності її опрацювання. Відкриті джерела сприяють більш швидкому поширенню інформації, характеризуються більшим ступенем предметної охопленості. Однак вони вимагають особливого контролю щодо якості інформації

Таким чином, сучасні туристичні підприємства потребують якісної документації для інформаційної підтримки надання широкого спектру туристичних послуг, яка б формувалася шляхом лінгвістичного аналізу інформаційного наповнення великої кількості веб-сайтів з відкритим доступом у WWW. Це підтверджує актуальність наукової задачі розроблення методів і засобів формування туристичної документації шляхом комп'ютерно-лінгвістичного аналізу інформаційного наповнення відкритих веб-ресурсів.

#### **Зв'язок роботи з науковими програмами, планами, темами.**

Дисертаційні дослідження виконані у межах пріоритетного наукового напрямку розвитку науки і техніки “Інформаційні та комунікаційні технології”, визначеного Законом України “Про пріоритетні напрями розвитку науки і техніки” від 05.12.2012 р. Тема дисертаційної роботи відповідає науковому напрямку кафедри інформаційних систем та мереж Національного університету “Львівська політехніка”. Робота виконана в межах науково-дослідних робіт “Методи та засоби консолідації баз даних в інформаційних системах електронного урядування”, 2010-2012 рр., № держреєстрації 0110U005022 та “Лінгвістичне забезпечення консолідації відкритих інформаційних ресурсів”, 2013 р., № держреєстрації 0113U005274 (автором розроблено нові моделі та методи збору та верифікації туристичної інформації на основі комп'ютерно-лінгвістичного аналізу інформаційного наповнення неоднорідних відкритих веб-ресурсів).

#### **Мета і завдання дослідження.** Метою дослідження є розроблення

методів і засобів формування туристичної документації за допомогою комп'ютерно-лінгвістичного аналізу інформаційного наповнення відкритих веб-ресурсів.

Для досягнення поставленої мети необхідно розв'язати такі задачі:

- проаналізувати особливості формування туристичної документації на основі методів структурування інформаційного наповнення відкритих веб-ресурсів;
- розробити інформаційну модель туристичної документації та інформаційного джерела з урахуванням лінгвістичних особливостей їх елементів;
- розробити метод лінгвістичного співставлення інформаційного наповнення відкритих веб-ресурсів з елементами туристичної документації;
- розробити процедури формування та виявлення недоліків туристичної документації на основі лінгвістичного співставлення інформаційного наповнення відкритих веб-ресурсів з елементами туристичної документації;
- розробити метод оцінювання якості туристичної документації;
- провести експериментальне дослідження та здійснити апробацію отриманих результатів.

**Об'єктом дослідження** є процес формування туристичної документації

**Предметом дослідження** є методи і засоби формування туристичної документації за допомогою комп'ютерно-лінгвістичного аналізу та опрацювання інформаційного наповнення, що міститься у відкритих веб-ресурсах.

**Методи дослідження.** Для інформаційного моделювання туристичної документації та відкритих веб-ресурсів використано теоретико-множинні підходи, апарат теорії реляційних баз даних, теорію графів. Виявлення вербальних характеристик інформаційного наповнення відкритих веб-ресурсів



проведено за допомогою методів прикладної та математичної лінгвістики. Методи співставлення інформаційного наповнення відкритих веб-ресурсів з структурними елементами туристичної документації за допомогою лінгвістичних індикативних ознак побудовано із застосуванням апарату нечітких множин. В основі процедури вибору відкритих веб-ресурсів для формування туристичної документації лежать сучасні методи інформаційного пошуку в Інтернет. Для розроблення процедур лінгвістичного співставлення інформаційного наповнення відкритих веб-ресурсів з елементами туристичної документації використано методи інтелектуального аналізу текстів та методи побудови штучних мов маркування текстів. Розроблені процедури виявлення та усунення недоліків у туристичній документації ґрунтуються на теорії множин, частотному аналізі, методах визначення подібності текстів. Для розроблення методів оцінювання якості туристичної документації використано метрики оцінювання якості програмних продуктів за стандартом ISO/IEC-25010. Проектування прототипу системи автоматизованого формування туристичної документації здійснено за допомогою діаграмних засобів структурного моделювання.

**Наукова новизна одержаних результатів** полягає науковому обґрунтуванні та розробленні методів і засобів комп'ютерно-лінгвістичного аналізу інформаційного наповнення відкритих веб-ресурсів для формування туристичної документації. При цьому отримано такі наукові результати:

- уперше побудовано формальну модель туристичної документації на основі множин туристичних послуг, відгуків споживачів туристичних послуг та класифікаторів з урахуванням відповідних індикативних ознак їх елементів, що стало основою для побудови процедур комп'ютерно-лінгвістичного формування туристичної документації і показників її якості туристичної документації;
- удосконалено формальну модель інформаційного джерела для формування туристичної документації шляхом виокремлення блоків і

атомарних ситуації туристичної інформації на веб-сторінках, що стало основою для побудови процедури вибору відкритих веб-ресурсів, процедур лінгвістичного співставлення вмісту відкритих веб-ресурсів з елементами туристичної документації;

- набув подальшого розвитку метод лінгвістичного виявлення інформації у веб-ресурсах шляхом співставлення атомарних ситуацій з елементами туристичної документації на основі модифікованих індикативних ознак, що дозволило побудувати відповідні процедури лінгвістичного співставлення вмісту відкритих веб-ресурсів з елементами туристичної документації;
- уперше розроблено метод оцінювання якості туристичної документації на основі функціональних показників якості, показників якості придатності до використання та показників якості супроводжуваності стандарту ISO/IEC-25010, що забезпечило практичне оцінювання результатів виконання процедур формування та перевірки туристичної документації.

**Практичне значення одержаних результатів** дисертаційної роботи зумовлено тим, що вони дають змогу автоматизовано формувати туристичну документацію, проводити її перевірку та оцінювати її якість із врахуванням потреб споживача. Зокрема, практично цінними є такі результати:

- розроблено процедури комп'ютеро-лінгвістичного формування туристичної документації на основі співставлення вмісту відкритих веб-ресурсів з елементами за допомогою індикативних ознак, а також процедури виявлення прогалин та некоректної інформації у туристичній документації на основі аналізу її структури та вмісту, які дозволяють покращити показники якості туристичної документації і є необхідними для побудови системи автоматизованого формування туристичної документації;
- розроблено показники якості туристичної документації на основі

функціональних показників якості, показників якості придатності до використання та показників якості супроводжуваності стандарту ISO/IEC-25010, що забезпечило практичне оцінювання результатів виконання процедур формування та перевірки туристичної документації;

- розроблено архітектуру та функціональну структуру системи комп'ютерно-лінгвістичного формування туристичної документації на основі процедур формування та показників якості туристичної документації, що дає можливість на практиці реалізувати і впровадити цю інформаційну систему.

Практичне значення дисертаційної роботи підтверджується позитивним ефектом від застосування розроблених методів і засобів комп'ютерно-лінгвістичного аналізу інформаційного наповнення відкритих веб-ресурсів для формування туристичної документації. Результати дисертаційного дослідження впроваджено у туристичних компаніях “Королівські подорожі” та “Тур де Львів”, а також у Національному університеті “Львівська політехніка”. Використання результатів дисертаційної роботи підтверджено актами про впровадження.

Результати дисертаційної роботи використовують у навчальному процесі в Національному університеті «Львівська політехніка» у лекційних курсах дисциплін «Методи та засоби інженерії даних та знань», «Методи та засоби інтеграції даних» під час підготовки студентів за спеціальністю 122 «Комп'ютерні науки та інформаційні технології», що підтверджено відповідними актами.

**Особистий внесок здобувача.** Усі наукові результати дисертаційної роботи отримані автором самостійно. У друкованих працях, опублікованих у співавторстві, автору належать: [12] – аналіз особливостей генерування запитів у веб-системах, [36] – аналіз особливостей автоматизації опрацювання електронної документації у сфері надання туристичних послуг, [37, 166, 169] –

метод автоматизованого формування туристичної документації на основі інтелектуального аналізу текстів, [39, 168] – компоненти якості туристичної документації, [40] – класифікація та систематизація документів за обов'язковими атрибутами, [41, 165] – метод формування та моніторингу якості туристичної документації, [42] – лінгвістичний аналіз структурних елементів туристичної документації та метод формування туристичної документації з урахуванням прогалин, [59, 134] – множина соціокультурних виразів, [96] – аналіз процесу усунення прогалин у документації, [144] – підхід семантичного маркування слабо структурованих даних, [155, 156, 157] – підхід до інформаційного моделювання за допомогою редукції графів, [167] – технологія формування туристичної документації на основі слабо структурованих текстів, [170] – метод формування туристичної документації на основі відкритих веб-ресурсів.

**Апробація результатів дисертації.** Основні результати наукових досліджень неодноразово доповідалися на ряді міжнародних та всеукраїнських наукових конференціях, зокрема:

- 1) the 4-th, 7-th, 8-th, 11-th International Conferences «Computer Science and Information Technologies» (Lviv, CSIT-2009, CSIT-2012, CSIT-2013, CSIT-2016);
- 2) the XXI Polish-Ukrainian Conference “CAD in Machinery Design – Implementation and Educational Problems” (Poland, Łańcut: CADMD-2013);
- 3) Human. Computer. Communication: the 1st International Scientific Conference HCC-2013 (Lviv, 2013);
- 4) V Міжнародна науково-практична конференція “Математика. Інформаційні технології. Освіта” (Луцьк-Світязь, 2016);
- 5) II, III, IV, V Міжнародні наукові конференції “Інформація, комунікація, суспільство” (Львів-Славське, 2013, 2014, 2015, 2016); IV, V-th International Conferences of Young Scientists «Computer Science and

Engineering» (Lviv, CSE-2010, CSE-2011).

Про результати дисертаційних досліджень автор регулярно доповідав на наукових семінарах кафедри інформаційних систем та мереж Національного університету “Львівська політехніка” (2011 – 2016).

Практичне значення дисертаційної роботи підтверджується застосуванням розроблених теоретичних засад в підприємствах туристичної сфери України.

**Публікації.** основні результати дисертаційного дослідження опубліковано у 25 наукових працях: 9 – опубліковано у фахових наукових виданнях України; 1 – в закордонному періодичному виданні; 15 – у матеріалах наукових конференцій, серед яких 1 – у закордонному виданні, 1 – в наукометричній базі Scopus.

## **Розділ 1. Аналіз особливостей туристичної діяльності**

В умовах інтенсивної глобалізації і значного розвитку інформаційного суспільства спостерігається стрімке зростання кількості туристичних послуг та швидкий розвиток туристичної індустрії. При цьому істотно збільшується кількість інформаційних ресурсів у відкритому доступі, де містить інформація про туристичну діяльність.

Для ефективної туристичної діяльності необхідно створити відповідні умови для оперативної обробки всієї доступної інформації з метою приведення її до єдиного стандартного вигляду, щоб в подальшому мати можливість обробляти й використовувати інформацію залежно від потреб надання 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 та інструментів.

Опрацювання великих масивів інформації, розміщеної на відкритих веб-ресурсах, вимагає розробки певних підходів до її автоматизованої консолідації, класифікації, структурування, переробки і використання. При цьому в межах побудови відповідних інформаційних систем підтримки туристичної діяльності важливе місце займає лінгвістичне забезпечення як сукупність мовних засобів опрацювання інформації.

Основні результати розділу опубліковані автором у працях [36, 37, 40, 41, 42, 166, 167, 169, 170].

### **1.1. Аналіз тенденцій розвитку туристичної галузі в сучасних умовах**

Туризм – одна з галузей сучасної економіки, яка динамічно розвивається протягом тривалого часу. За даними Всесвітньої туристичної організації (The World Tourism Organization– UNWTO) за останні 20 років загальна кількість іноземних туристів у світі подвоїлася [159], досягнувши показника 1 186,3 осіб у

2015 році (у 2012 році кількість іноземних туристів досягла рубежу в 1 млрд.). Така ж динаміка спостерігається у Європі з показником 607,7 туристів у 2015 році (рис. 1.1).

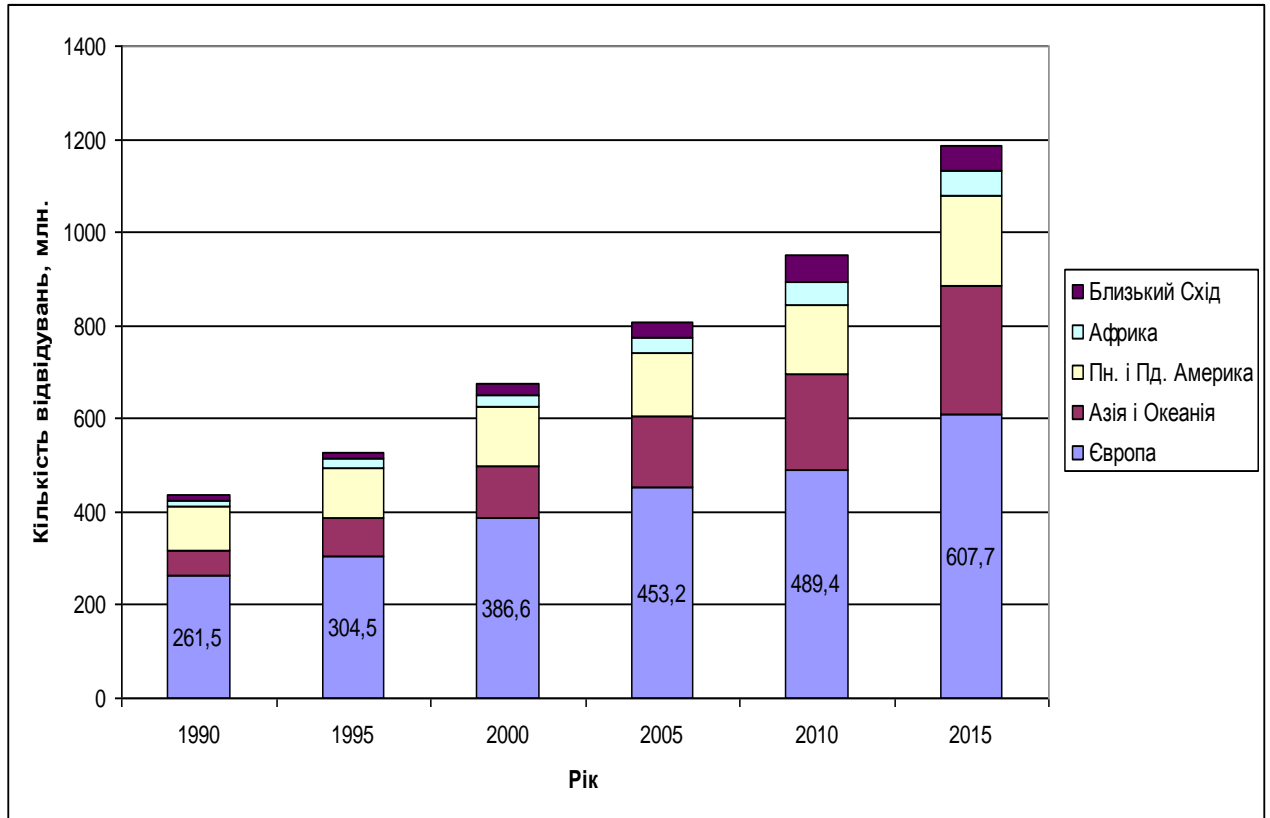


Рис. 1.1. Динаміка міжнародних туристичних відвідувань у світі

В Україні до 2014 року також спостерігалось стрімке зростання міжнародних туристичних відвідувань, зокрема за період 2000-2013 кількість іноземних туристів збільшилася майже в 3 рази [33, 99]. Незважаючи на значне скорочення іноземних туристів протягом 2014-2015 років в силу об'єктивних обставин в Україні і надалі прогнозується сталий ріст кількості міжнародних туристичних відвідувань [33, 158] (рис. 1.2).

Розвиток туристичної галузі України має інші тенденції: починаючи з 2000 року кількість туристів, які обслуговуються суб'єктами туристичної діяльності стійко коливається відносно позначки понад 2 млн. туристів в рік. При цьому частка внутрішніх туристів стрімко падає, тоді як частка громадян

України, які виїжджали за кордон, стрімко зростає [33] (рис. 1.3). Частка іноземних туристів до 2014 року демонструвала незначне спадання.

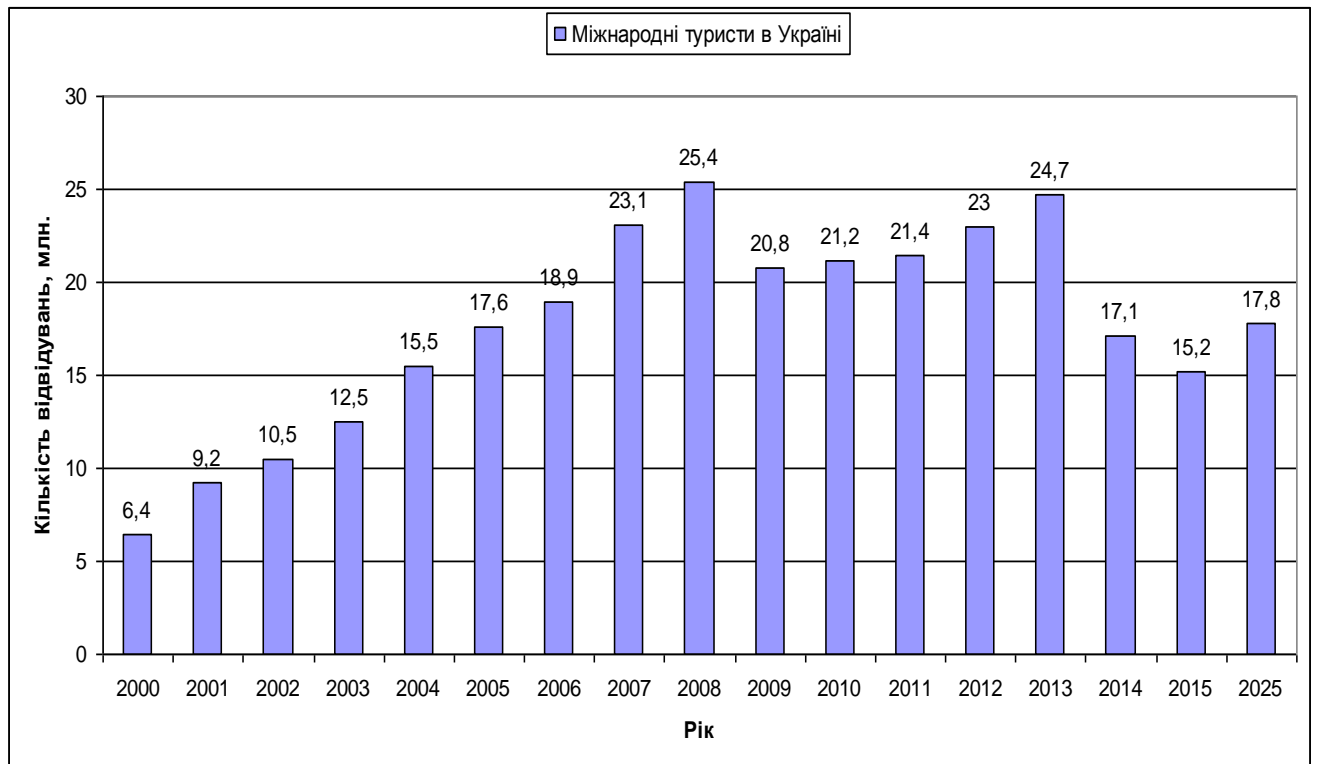


Рис. 1.2. Динаміка міжнародних туристичних відвідувань в Україні

Такі тенденції показують на наявність певних проблем в туристичній галузі України, які доцільно пов'язувати з:

- фінансовим станом туристів України;
- якістю надання туристичних послуг в Україні;
- якістю надання інформаційних послуг в Україні.

Власне ця робота направлена на розв'язування задач збору інформації про туристичну діяльність у межах розв'язування останньої проблеми.

Туристична діяльність в Україні регламентується на законодавчому рівні, зокрема Закон України "Про туризм" [48] регулює відносини між суб'єктами туристичної діяльності. Цей закон дає такі основні визначення в межах туристичної діяльності.



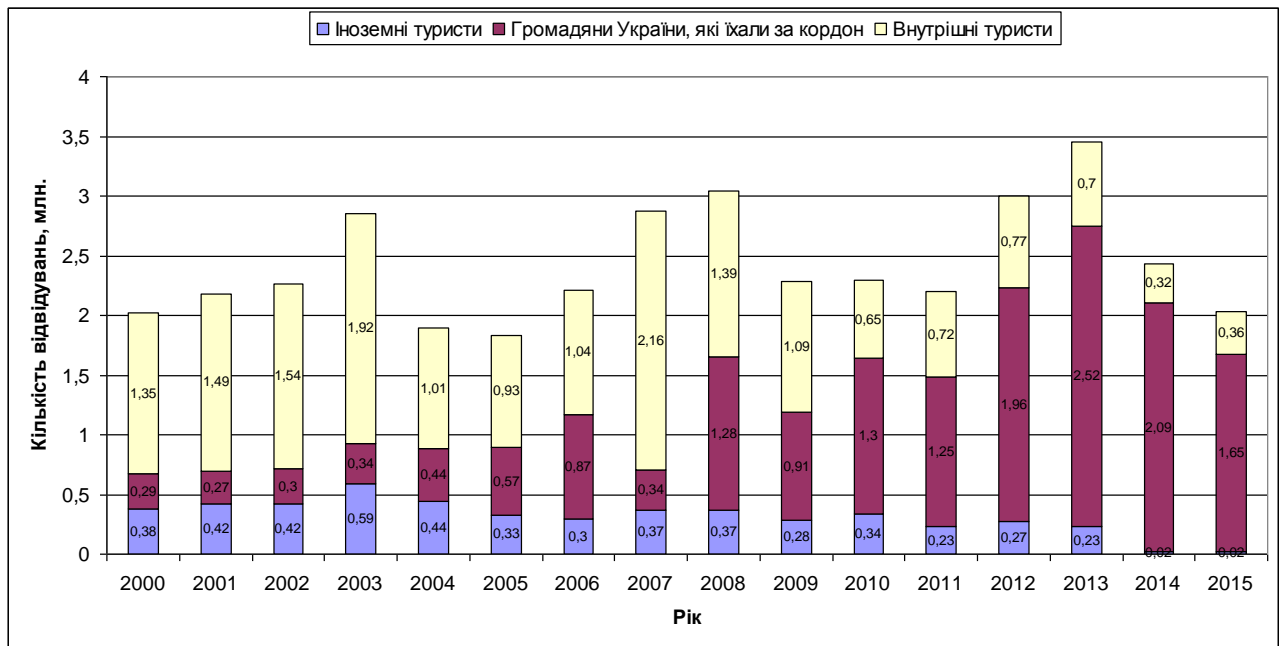


Рис. 1.3. Динаміка обслуговування туристів суб'єктами туристичної діяльності України

**Туризм** – тимчасовий виїзд особи з місця постійного проживання в оздоровчих, пізнавальних, професійно-ділових чи інших цілях без здійснення оплачуваної діяльності в місці перебування

**Туристичний продукт** – попередньо розроблений комплекс туристичних послуг, який поєднує не менше ніж дві такі послуги, що реалізується або пропонується для реалізації за визначеною ціною

**Туристична послуга (ТП)** – послуга суб'єкта туристичної діяльності щодо розміщення, харчування, транспортного, інформаційно-рекламного обслуговування, а також послуги закладів культури, спорту, побуту, розваг тощо, спрямована на задоволення потреб туристів.

Ринок ТП містить відносини між [56, 60, 63, 119]:

- виробником (надавачем) туристичної послуги і споживачем;
- виробниками туристичної послуги;
- споживачами туристичної послуги;

Зв'язок між виробником та споживачем туристичної послуги базується на тому що споживач може бути виробником і навпаки виробник споживачем.

Споживачем ТП можуть бути туристи, місцеві жителі, туристичні фірми, туристичні агентства, туристичні оператори, суб'єкти підприємницької діяльності.

**Туристична діяльність** – діяльність з надання різноманітних туристичних послуг.

Туристична діяльність має такі характерні особливості [57, 58, 73, 78, 79, 81, 82, 102, 108, 126, 131, 147]:

- міжнародний характер функціонування;
- сезонні коливання попиту на туристичні послуги;
- кожен суб'єкт може бути як споживачем туристичного продукту так і виробником;
- велика кількість суб'єктів туристичної діяльності;
- реалізація та споживання туристичного продукту не збігається у часі та просторі;
- важливість сервісного обслуговування як базової конкурентної переваги;
- рішення про покупку приймається індивідуально з використанням різноманітних джерел інформації;
- покупці туристичного продукту як правило не є спеціалістами у даній сфері;
- процес купівлі/продажу на ринку як правило потребує великих формальностей (особливо, у випадку зарубіжного туризму);
- необхідність використання інформаційного забезпечення для візуалізації туристичної послуги з огляду на її невідчутність.

Важливим елементом туристичної діяльності є туристична інформація у вигляді потоків даних, послуг і платежів, яка забезпечує зв'язки між виробниками туристичних послуг та їх споживачами.

Під **туристичною інформацією (ТІ)** будемо розуміти відомості про

туристичні об'єкти, суб'єкти туристичної діяльності та туристичну діяльність загалом.

Основним призначенням туристичної інформації є забезпечення взаємодії між споживачем і виробником туристичного продукту у вигляді двосторонньої комунікації на основі взаємного спілкування. Тобто туристична інформація продукується, обробляється, зберігається і передається в межах певної інформаційної системи (ІС) [44, 89, 100]. В сучасних умовах така ІС повинна будуватися з використанням моделі В2С (Business-to-Consumer) – надання послуг кінцевому споживачу на основі онлайн спілкування між суб'єктами туристичної діяльності [136].

З погляду надавача туристичних послуг модель В2С передбачає формування наперед підготовленого масиву інформації, в основі якого лежить використання систем електронного документообігу [1, 5, 6, 7, 11, 21, 32, 40, 62, 76, 80, 97, 101, 116]. Такі системи забезпечують ефективний обмін інформацією в межах туристичної установи, а також дозволяють ефективно взаємодіяти з іншими туристичними організаціями та оперативно задовольняють запити користувачів.

З погляду споживача туристичних послуг функціонування моделі В2С пов'язане з такими особливостями [2, 3]:

- обмін інформацією між виробником та споживачем туристичного продукту здійснюється неперервно, переважно через Інтернет;
- надання туристичних послуг має персоналізований характер з урахуванням потреб споживача туристичного продукту;
- надання туристичних послуг має обмежені часові рамки, оскільки послугу можна отримати у заданому місці у визначений час, тобто споживання послуги відбувається у місці пропозиції з одночасним її формуванням без можливості її збереження;
- якість надання туристичних послуг має суб'єктивні оцінки, оскільки задоволення послугою залежить від вподобань окремого туриста,

природних факторів в момент отримання послуги, а також відгуків інших клієнтів;

- якість інформаційного забезпечення залежить від якості інформаційного наповнення туристичного продукту та ефективності процесів його формування.

Також з погляду споживача туристичних послуг важливі відомості щодо психологічних особливостей контактування туристів з представниками місцевих громад, які можуть змінюватися від хороших стосунків до конфліктних ситуацій [90].

### **1.1.1. Аналіз інформаційного забезпечення туристичної діяльності**

Туристична діяльність пов'язана з ефективним опрацюванням та наданням необхідної інформації. Виділяють такі основні проблеми, які виникають під час роботи з інформацією в туристичній галузі [50]:

- відсутність або неактуальність веб-сайту туристичного підприємства;
- не достатньо повна інформація про туристичні послуги;
- використання туристичних послуг третіх осіб без належного інформування про це.

Такі проблеми на підприємствах у першу чергу пов'язані з відсутністю комплексного підходу щодо формування інформаційного забезпечення туристичної діяльності, хоч існує чітке розуміння неможливості функціонування туристичної галузі без широкого використання інформаційних технологій [115].

**Інформаційне забезпечення** – це сукупність форм документів, нормативної бази та реалізованих рішень про обсяги, розміщення та форми організації інформації, що циркулює в організації в межах ІС [44]. Інформаційне забезпечення включає єдині системи класифікації і кодування інформації, уніфіковані системи документації, схеми інформаційних потоків, методології побудови баз даних та ін.

На сьогодні основним середовищем формування інформаційного забезпечення у туристичній галузі є Інтернет. Власне на основі інформації, отриманої через Інтернет, споживач приймає рішення щодо придбання туристичного продукту (послуги).

З погляду надавача туристичних послуг в Інтернет проводяться такі основні види діяльності [77, 92, 121, 130]:

- моніторинг ринку туристичних послуг (дослідження потреб споживачів, аналіз діяльності конкурентів, аналіз відгуків тощо) через пошукові системи, веб-спільноти, тематичні сайти і портали, електронні каталоги, власний веб-сайт, е-пошту;
- маркетинг (інформування про діяльність підприємства, реклама, PR, стимулювання збуту тощо) через власний Веб-сайт, електронні каталоги, веб-спільноти, банерну та контекстну рекламу, перехресні посилання, е-пошту;
- підключення до Інтернет-сервісів, які надають туристичні послуги (бронювання квитків на транспорт, резервування житла тощо), фінансові послуги (платіжні системи) або інформаційні послуги (отримання додаткової інформації про країни, оперативний прогноз погоди в різних країнах світу, оперативна інформація про ціни на проживання, харчування та інші туристичні послуги тощо);
- отримання професійної та ділової інформації через пошукові системи, веб-спільноти, тематичні сайти і портали, каталоги, конференції, розсилки;
- підтримання ділового зв'язку через власний веб-сайт, е-пошту.

З погляду споживача туристичних послуг Інтернет виступає середовищем для таких видів діяльності [77, 121]:

- отримання інформацію про туристичні місця і маршрути, туристичні підприємства та їхні послуги;

- отримання консультацій та ознайомлення з відгуками інших споживачів;
- підключення до Інтернет-сервісів, які надають туристичні та фінансові послуги.

Важливо відзначити, що значна частина туристів (84% у 2013 році) свої рішення щодо подорожі приймають на основі інформації про думки та відгуки, отриманою з веб-спільнот [45]. Така довіра до інформації про туризм ґрунтується тому, що від 40% до 75% споживачів туристичних послуг діляться своїми відгуками.

### **1.1.2. Аналіз особливостей формування туристичної документації**

Документація – це один з видів інформаційного забезпечення ІС. ДСТУ 2392-94 визначає документацію як множину документів, підібраних зі спеціальною метою [26]. За цим же стандартом документ – це записана інформація, яка може розглядатися як одиниця в інформаційній діяльності (збирання, опрацювання, збереження та поширення інформації). За ДСТУ 2732:2004 документ – це інформація, зафіксована на матеріальному носії, основною функцією якого є зберігати та передавати її в часі та просторі [27]. Важливою особливістю будь-якого документа є його реквізити, які дозволяють ідентифікувати інформацію, що міститься в документі [44, 112, 127].

Документація містить дві складові: інформаційну та матеріальну. До інформаційної складової відносяться: наявність структурних блоків, мова подання, наявність прогалін, достовірність, чітка структура. До матеріальної складової належать форма документа, форма носія інформації, спосіб документування чи запису.

Під час побудови функціонування ІС використовуються спеціальний вид документації – технічна документація. Зокрема, технічна документація описує проектування, розробку та використання програмних продуктів, баз даних, інформаційних сервісів тощо.

За ДСТУ 3278-95 **технічна документація** – це система графічних і текстових документів, необхідних і достатніх для безпосереднього використання на усіх стадіях життєвого циклу продукції.

Створення технічної документації передбачає використання стандартизованої термінології, виконання визначених положень та вимог щодо складу і змісту уніфікованих форм документів, а також дотримання порядку їх виконання [85, 87, 98].

Туристична діяльність пов'язана з наданням або окремих послуг, або цілісних туристичних продуктів. Тоді технічна документація до туристичних послуг і/або продуктів надає споживачеві необхідну інформацію про туристичні послуги і/або продукти

**Визначення 1.1. Туристична документація (ТД)** – це технічна документація про туристичні послуги і/або продукти, яка необхідна для споживача туристичної послуги і/або продукту.

Туристична документація (як і будь-який інший вид документації) може формуватися на матеріальному носії або електронно. Якщо інформація документів, що утворюють документацію, зафіксована у вигляді електронних даних включно з обов'язковими реквізитами документа, то такі документи є електронними [47, 101, 106], і відповідна документація теж вважається електронною.

Надалі, використовуючи визначення 1.1, будемо говорити виключно про електронну туристичну документацію.

Направленість на споживача дозволяє розглядати ТД як інформаційний продукт (ІП), оскільки ІП – це документована інформація, підготовлена відповідно до певних вимог для її поширення [44]. Тому від якості ТД як ІП залежить задоволеність потреб споживачів ТД.

Формування якісної ТД пов'язане з досягненням таких цілей в інтересах користувача [19, 20, 107, 110, 111, 120]:

- підвищення привабливості ТП в очах споживача (потенційного туриста), що може підвищити конкурентоспроможність ТП;
- надання усієї інформації, потрібної для розуміння особливостей надання ТП, що дозволяє захистити надавача ТП від можливих претензій, які можуть виникнути під час неналежного споживання ТП.

Під **інформаційною потребою** споживача ТП будемо розуміти необхідність (потребу) в отриманні певної інформації про ТП для задоволення своїх інтересів.

**Туристичний інформаційний продукт (ТІП)** – задокументована згідно з вимогами ТІ, що задовольняє інформаційні потреби споживача ТП і подана у вигляді ТП.

**Споживач туристичного продукту (СТП)** – суб'єкт туристичної діяльності.

**Туристична послуга** – послуга, що надається СТП відповідно до його інформаційних потреб у вигляді інформаційного продукту.

Систематизація ТІ та формування ТІП надає такі переваги:

- зосередження ТІ в одному цілісному інформаційному ресурсі, що має єдине форматування та середовище зберігання, уніфіковані структуру і подання, спільні методи та засоби формування документів;
- доступність в будь-який момент часу будь-якому споживачеві ТІП при мінімальних затратах;
- зменшення витрат часу на отримання необхідної інформації, оскільки при якісному формуванні ТД не потрібно постійно звертатися до експертів;
- підвищення точності інформації, оскільки ретельна перевірка інформації експертами та використання великої кількості різних джерел з подальшою консолідацією зменшує її суб'єктивність.

Можливі такі методи формування ТІП:

- 1) за допомогою фахівців/експертів.



2) автоматизовано за допомогою певної інформаційної системи;

3) комбіновано за допомогою попередніх методів;

Перший метод має такі істотні недоліки:

- використання обмеженої кількості ресурсів може привести до викривлення інформації про ТП, якщо не буде усунуто заангажованості авторів інформації;
- використання обмеженої кількості експертів може привести до лінгвістичних помилок в орфографії, синтаксисі, нечіткості визначень та помилкових класифікацій за категоріями тощо;
- існування прогалів в ТД, недостатня деталізація ТП.

Другий метод покликаний максимально усунути ці недоліки, хоча на практиці його застосування неможливе без експертних оцінок.

### **1.1.3. Аналіз особливостей формування туристичного інформаційного продукту**

З визначення ТПП випливає, що туристичний інформаційний продукт – це туристична документація, орієнтована на споживача. Тобто ТПП – це ТД, що розглядається як ІП.

На рис. 1.4 подано класифікацію ТД. За цією класифікацією документи ТД згруповані за такими категоріями:

- документи нормативно-правової бази регулювання ТД;
- документи на оформлення взаємин;
- документи на просування продукту;
- документи для забезпечення моніторингу якості надання туристичних послуг.

Існує ряд нормативно-правових документів, які регламентують розвиток сфери туризму: нормативно-правові акти та стандарти державного та міждержавного характеру. Серед державних нормативно-правових актів в Україні виділяють закони України [46, 47, 48], акти Президента України, акти

Кабінету Міністрів України, акти центральних органів виконавчої влади.



Рис. 1.4. Класифікація документів, що входять в туристичну документацію

Стандарти стосуються управління документами, що створені або отримані організацією в процесі її діяльності, визначають процеси документування, надають інформацію щодо розробки методики управління документацією в організації та пов'язаних з нею процедур, систем і процесів, а також щодо розробки та впровадження систем документації [26, 27, 30, 113, 139].

Документацію оформлення взаємин умовно поділяють на:

- документи для замовлення (наприклад, заявка на бронювання, лист бронювання, підтвердження бронювання, розрахункові чеки тощо);
- документи для клієнта (наприклад, договір, путівка, ваучер, страховий поліс, квиток на транспорт, супровід гіда, віза, медична документація тощо);
- документи на підтвердження особи туриста (наприклад, паспорт, доручення на дітей, медична довідка, документи про несудимість та ін.);
- документи про умови надання послуг (наприклад, технологічна карта туристичної подорожі, графік завантаження туристичного підприємства, інформаційний листок до туристичної путівки, опис екскурсії та ін.).

В контексті дослідження цієї роботи з орієнтацією на споживача ТПІ надалі будемо брати до уваги такі типи документації:

- документи для клієнта;
- документи на підтвердження особи туриста;
- документи промоційної кампанії;
- документи про туристичну привабливість;
- документи про екологічну ситуацію;
- документи про зручності пересування.

Для формування ТПІ розглянемо основні сфери документування з погляду туриста. Це допоможе виділити основні тематичні напрямки вилучення інформації і в подальшому дозволить систематизувати основні правила аналізу даних як структурованих, так і неструктурованих.

Для туриста важливо отримати повну інформацію про майбутню поїздку. Туристи оцінюють загальну туристичну привабливість міста, історичні та культурні пам'ятки, атмосферу міста, зручність бронювання подорожі, якість життя, зручність пересування по місту, екологічну ситуацію, атрактивність та

ін. [4, 10, 22, 24, 34, 54, 132, 133, 160]. Тому повинні бути задокументовані основні напрями зацікавлення споживача ТП (рис. 1.5):



Рис. 1.5. Узагальнена класифікація розділів туристичної документації в інтересах споживача туристичного продукту

Класифікація розділів ТД за тематиками необхідна для чіткого наповнення ТП інформацією. В кожному розділі інформація має бути подана в єдиному вигляді як шаблон з впорядкованими інформаційними блоками та формою подання та мовним кодом.

## 1.2. Аналіз використання відкритих веб-ресурсів у туристичній діяльності

Режим доступу до інформації – це передбачений правовими нормами порядок одержання, використання, поширення і зберігання інформації [46]. За режимом доступу інформація поділяється на відкриту інформацію та інформацію з обмеженим доступом. Доступ до відкритої інформації забезпечується шляхом:

- систематичної публікації її в офіційних друкованих виданнях (булетенях, збірниках);

- поширення її засобами масової комунікації;
- безпосереднього її надання зацікавленим громадянам, державним органам та юридичним особам.

**Відкриті веб-ресурси (ВВР)** – ресурси WWW, які знаходяться у відкритому доступі для усіх користувачів Інтернет.

WWW є одним з основних джерел ТІ в Інтернет. ВВР – це та частина WWW, яка доступна для широкого загалу користувачів – потенційних туристів.

### **1.2.1. Переваги використання відкритих веб-ресурсів для формування туристичної документації**

Усі особливості та переваги використання ВВР для формування туристичної документації розділимо на дві групи:

1) за доступністю для кінцевого користувача:

- велика територіальна охоплюваність;
- орієнтованість на різні цільові аудиторії;
- необмежена часова доступність;
- відсутність обмежень за мовним критерієм;
- невеликі фінансові витрати.

2) за взаємодією між виробником та споживачем туристичного продукту:

- простота відслідковування відвідувачів сайту, визначення популярності джерела, моніторингу відгуків;
- взаємодія та надання інформацією в майже режимі реального часу;
- можливість самостійного ознайомлення з інформацією споживачем без суб'єктивності зацікавлених фізичних та юридичних осіб;
- можливість спілкування без візуального контакту;
- оперативне отримання зворотного зв'язку.

До джерел відкритого доступу, які доцільно використовувати для потреб формування ТД, віднесемо такі основні веб-ресурси [23, 51, 95, 123]:

- веб-спільноти – веб-форуми та соціальні мережі (у публічній компоненті);
- довідкові веб-ресурси – вікі-ресурси, електронні словники;
- веб-сховища – репозитарії та журнали, електронні бібліотеки;
- спеціалізовані веб-сайти – інформаційні портали, спеціалізовані туристичні сайти та каталоги.

### **1.2.2. Аналіз особливостей використання відкритих веб-ресурсів для формування туристичної документації**

Формування ТД з використанням різних відкритих веб-ресурсів можливе через інтеграцію інформації, яку можна отримати з цих ресурсів. Виділяють три методи інтеграції інформації (даних): тиражування, федералізація, консолідація [38, 69, 129]. Ці методи допомагають одержати цілісну картину про обрану предметну область.

Тиражування даних – це спосіб інтеграції, який полягає в копіюванні джерел даних з огляду на зміни в інформаційному наповненні і відповідно створення нового ідентичного чи подібного джерела внаслідок дублювання початкового джерела. Формування ТД через тиражування даних може здійснюватися за схемою, зображеною на рис. 1.6.

Під час тиражування відбувається копіювання ТІ з N джерел і створюється єдине сховище ТІ. В результаті опрацювання споживач ТП отримує ТД, в якій міститься інформація про різні ТП з цих N джерел.

Формування ТД через тиражування даних придатне лише для однорідних джерел інформації, що практично неможливо забезпечити для ВВР.

Федералізація даних – це спосіб інтеграції даних, який полягає у створенні віртуальної картини з різних інформаційних джерел із здійсненням доступу до інформації через метадані. Узагальнена схема федералізації ТІ подана на рис. 1.7.

Як видно із схеми, федералізація ТІ може використовуватися під час обробки неоднорідних ВВР. Однак при цьому можуть виникнути істотні

проблеми з продуктивністю опрацювання ТД, якщо працювати з великою кількістю джерел ТІ.

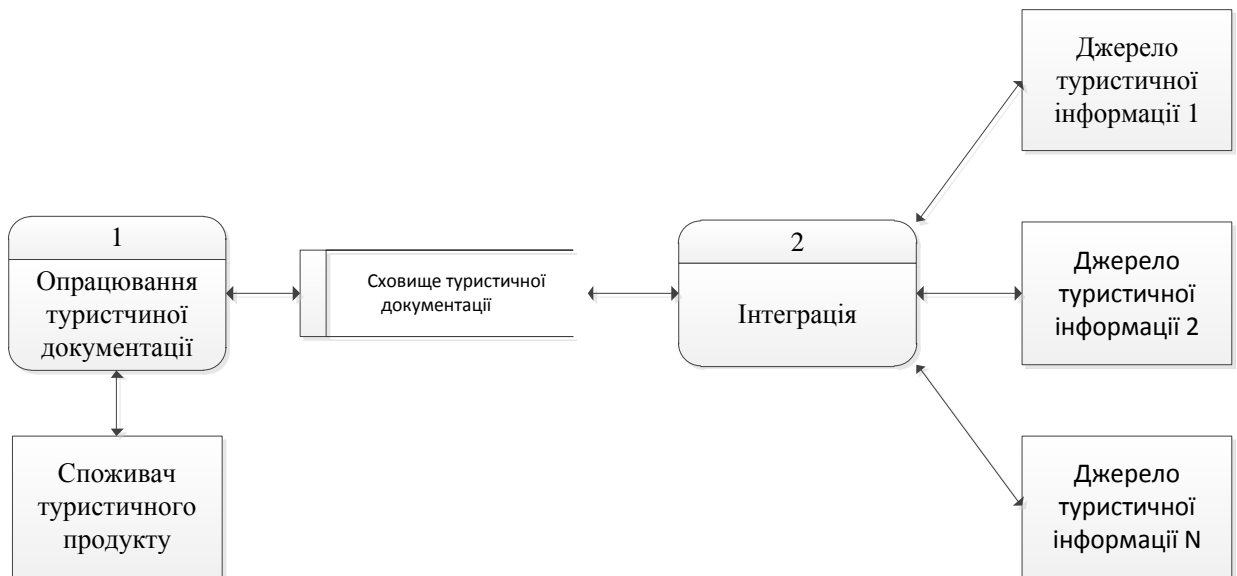


Рис. 1.6. Схема формування ТД через тиражування даних

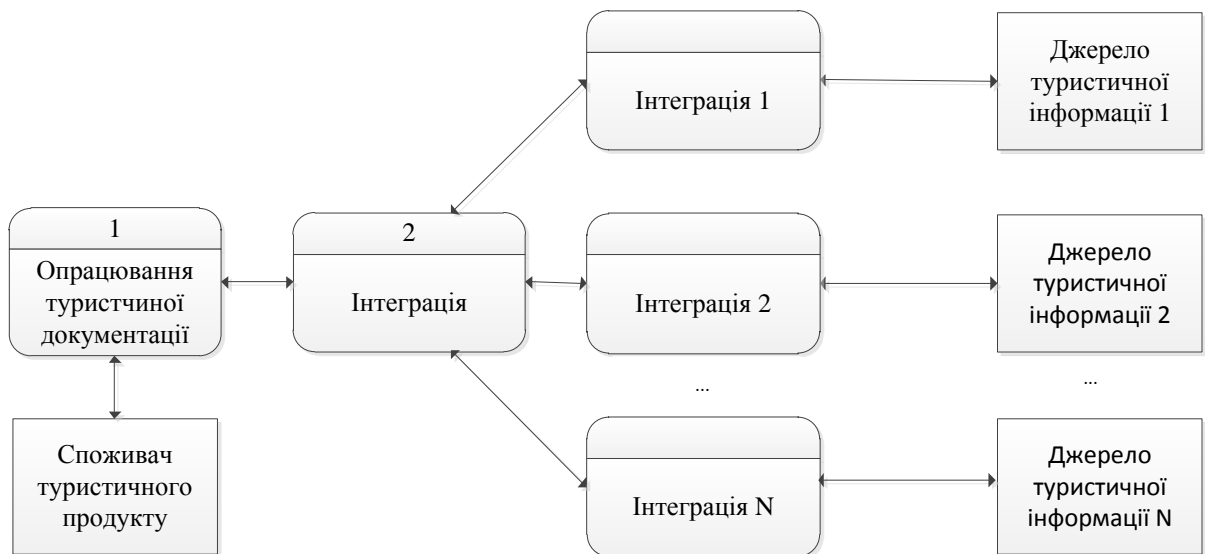


Рис. 1.7. Схема формування ТД через федералізацію даних

Консолідація даних полягає у формуванні єдиного сховища ТД з різнорідних початкових джерел ТІ. Узагальнена схема консолідації ТІ з джерел різного формату подана на рис. 1.8.

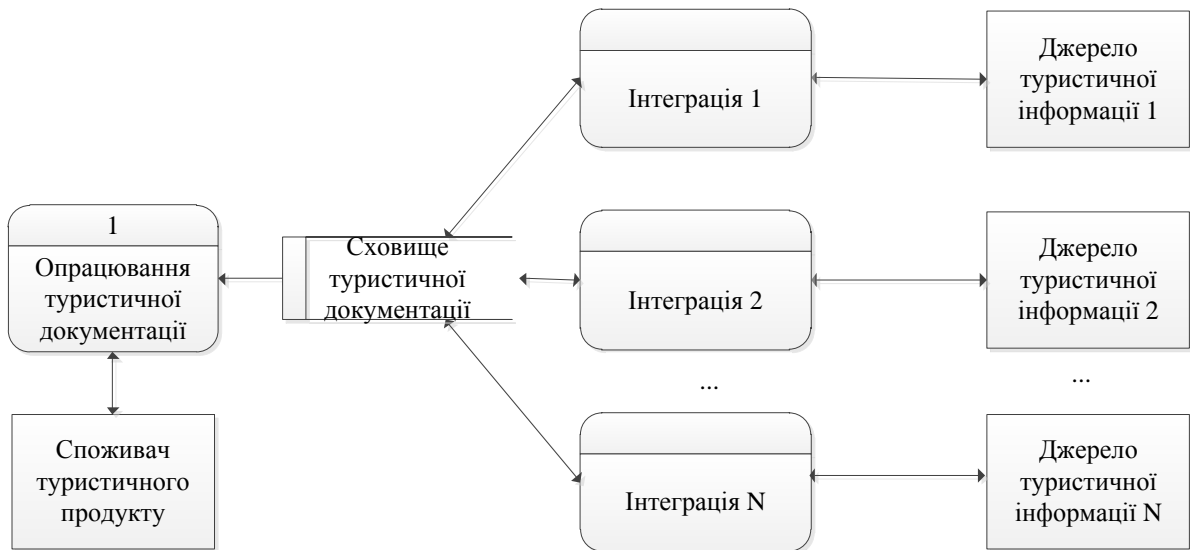


Рис. 1.8. Схема формування ТД через консолідацію даних

Консолідація ТІ дозволяє усунути недоліки, пов'язані з формування ТД через тиражування або федералізацію. Консолідація ТІ дозволяє обробити і перетворити значні обсяги розрізненої інформації, яка зберігається у ВВР, провести так звану реструктуризацію ресурсів, що дозволить створити відповідну ТД.

Консолідування інформації з багатьох різномірних ВВР в єдиний цілісний туристичний ресурс, що становить ТД і орієнтований на споживача ТП, включає такі етапи:

- 1) визначення цілей інформаційного ресурсу;
- 2) пошук та збір даних;
- 3) попереднє опрацювання та структурування даних;
- 4) аналіз і синтез інформації;
- 5) формування інформаційного продукту що задовольняє потреби споживача туристичного продукту.

Використання ВВР для консолідування інформації в ТД надає такі можливості:

- опрацювання різних інформаційних джерел;
- відкритість доступу в будь-який час для будь-якого споживача ТП при мінімальних затратах;



- можливість обміну досвідом та враженнями за допомогою коментування та залишення відгуків про ТП;
- цілісне подання консолідованої ТІ з різних джерел.

### 1.2.3. Аналіз підходів до моделювання відкритих веб-ресурсів

Питаннями моделювання інформаційних ресурсів у WWW присвячено ряд наукових робіт. Особливості такого подання залежать від тих задач, які ставляться перед автоматизованих опрацюванням ВР.

Загалом ВВР розглядається як веб-сайт, інформаційне наповнення (ІН) якого подається ієрархічно на трьох рівнях: веб-сайту, веб-сторінки, допису [8] (рис. 1.9). Веб-сайт розглядається як множина веб-сторінок, між якими існують посилання.

В [104] моделюється спеціалізований ВР – веб-галерея. Особливістю Веб-галереї є поділ веб-сторінок на декілька типів які утворюють ієрархію (дерево): експозиція, кімната, експонат. Аналогічний підхід до моделювання ВР зустрічається в [109] для веб-спільнот, де розглядається така ієрархія Веб-сторінок: форум, розділ форуму, дискусія.

У роботах [8, 75, 109, 118, 124] структура сторінок веб-спільнот подається як множина дописів окремих авторів. Кожен автор може створити один або декілька дописів на веб-сторінках ВР. В [104] структура сторінок веб-галереї подається як набір інформаційних блоків про експонати, куди входять:

- характеристики, пов'язані з початковим виявленням експонату;
- характеристики, пов'язані з транспортним переміщенням експонату;
- характеристики, пов'язані з дослідженням історії експонату;
- характеристики, пов'язані із зберіганням та відображенням експонату.

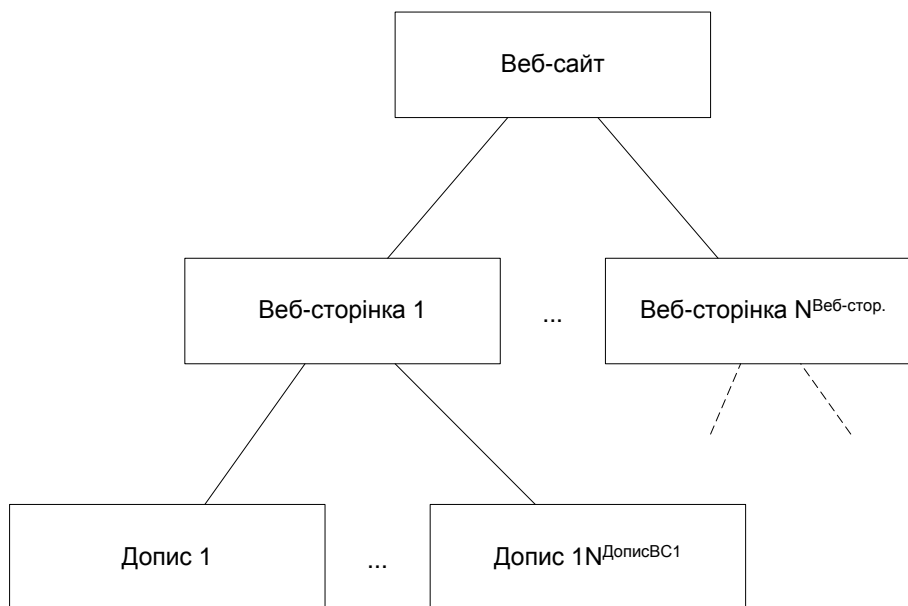


Рис. 1.9. Структура інформаційного наповнення веб-сайту

При цьому кожен інформаційний блок містить один або декілька записів відповідного змісту.

В [109] дописи дискусії веб-форуму складаються з таких основних складових:

- заголовок повідомлення;
- автор повідомлення;
- дата створення повідомлення;
- текст повідомлення.

В [8] дописи доповнено службовою інформацією, зокрема:

- час внесення останніх змін;
- вказівник на інший допис;
- скорочений варіант (анонс) повідомлення.

В [118] дописи класифікуються за певними тематичними ситуаціями споживацького досвіду, тобто в межах однієї дискусії можуть виникати декілька тематичних ситуацій (під-дискусій). Крім того дописи можуть розкладатися на окремі частини – атомарні ситуації споживацького досвіду.

Описані підходи до моделювання веб-спільнот та веб-галереї загалом

застосовні для будь-яких ВВР. З погляду формування ТД модель ВВР доцільно розглядати як ієрархічне подання таких складових:

- веб-сайту;
- веб-сторінки;
- інформаційного блоку;
- інформаційного допису;
- атомарної ситуації (частини) допису.

Характерною особливістю сучасних ВВР, які передбачають наповнення веб-сайту різними авторами (наприклад, учасниками веб-спільнот), є можливість пов'язування дописів з авторами. При цьому реагування на допис певного автора залежить від рівня довіри до цього автора [75].

Підсумовуючи, виділимо додаткові характеристики, які повинні мати складові ВВР з погляду формування ТД:

- для веб-сайту встановлюється рейтинг, який відображає авторитетність ВВР [75];
- дописи і атомарні ситуації пов'язуються з аспектами певної теми через маркери (ключові слова і фрази) [118, 124];
- автори дописів характеризуються рівнем довіри [75], який ґрунтується на мотивації дій автора [118], його інтересах [104], соціально-демографічних характеристиках [124], рівня персоніфікованості [8] тощо.

### **1.3. Аналіз використання XML-технологій для формування туристичної документації**

Задача інтеграції ТІ з ВВР є достатньо складною, оскільки практично неможливо забезпечити єдине подання цієї інформації на різних веб-сайтах. Це пов'язано з тим, що залежно від підходів до побудови веб-сайту використовуються різні підходи до структурування його вмісту, тобто використовуються різні метадані. Власне ефективне формування ТД практично

неможливе без стандартизації опису метаданих і приведення різних метаданих до єдиного формату.

### **1.3.1. Аналіз підходів до опису метаданих туристичної документації**

Метадані мають важливе значення для забезпечення ефективності пошуку інформації, опису слабо структурованої інформації, створення архівів із записами електронних документів. Для уніфікації метаданих для опису різноманітних веб-ресурсів використовують стандарт Dublin Core Metadata Initiative (скорочено Dublin Core – Дублінське ядро), що містить словник відповідних основних понять [103].

Дублінське ядро становить мінімально простий набір елементів для опису метаданих цифрових ІР. Семантика цих елементів базується на узгодженості позицій фахівців з різних галузей: бібліотечної науки, музейної справи і комп'ютерних технологій [13, 14, 18, 49, 83, 84]. При цьому враховано також і лінгвістичні проблеми, які притаманні для WWW, зокрема мова наповнення ІР [13, 18, 84].

Для реалізації метаданих Дублінського ядра та забезпечення сумісності метаданих різних Веб-ресурсів розроблене спеціальне середовище опису ресурсів RDF (Resource Description Framework), яке дозволяє описати спільну семантику, структуру та синтаксис подання ІР [148]. Фактично RDF – це технологія семантичної павутини (Semantic Web), яка передбачає надання інформації чітко визначеного значення, яке дозволяє ефективно використовувати її у спільній роботі різних систем. Оскільки використання RDF пов'язане з питаннями розробки ІС, то надалі ми не будемо торкатися проблематики застосування RDF.

В табл. 1.1 подано відповідність між елементами Дублінського ядра і ТД.

Таблиця 1.1.

Елементи опису метаданих Дублінського ядра в туристичній документації.

Аспект опису	Метадані	Зміст
Вміст ресурсу	Title – Назва	Назва елемента ТД
	Subject – Предмет	Розділ ТД
	Description – Опис	Опис елемента ТД
	Type – Тип	Тип елемента
	Source – Джерело	Інформація про ВВР, звідки отримано опис елемента ТД
	Relation – Відношення	Зв'язок елемента ТД з іншими елементами
Інтелектуальна власність	Creator – Автор	Автор допису на ВВР
	Publisher – Публікатор	Власник ВВР
	Contributor – Співавтор	Модератор (експерт), який здійснює опрацювання ТД
	Rights management – Права	Інформація про авторські права
Стан	Data – Дата	Дата створення чи редагування
	Format – Формат	Формат джерела
	Identifier – Ідентифікатор	URL сторінки, звідки отримано опис елемента ТД
	Language – Мова	Мова опису елемента ТД

### 1.3.2. Аналіз XML-засобів структурування інформації туристичної документації

XML (Extensible Markup Language) – мова розширюваної розмітки – це стандарт для створення мов розмітки ієрархічно структурованих даних для обміну між різними ІС в розподілених середовищах [137]. Фактично XML – це технологія або метамова для опису XML-мов за допомогою визначає набору базових лексичних та синтаксичних правил структурування інформації шляхом

застосування простих тегів. HTML (HyperText Markup Language) – це одна з багатьох XML-мов. Іншими прикладами XML-мов є: XSLT, RSS, Atom, XML Schema та ін.

Проаналізуємо ряд підходів, які практично використовуються для структурування інформації на веб-ресурсах і які впливатимуть на процедури формування ТД:

- 1) структурування інформації за допомогою XML-даних;
- 2) структурування інформації за допомогою стандартних XML-мов;
- 3) структурування інформації за допомогою HTML-тегів.

Перший підхід ґрунтується на тому, що будь-який веб-документ можна подати за допомогою трьох складових [38, 153]:

- структуровані дані документа, які подаються у форматі XML;
- логіко-графічне подання (розмітка) у вигляді заголовків, списків, таблиць, тощо, яке описується за допомогою шаблонів розмітки XSL (Extensible Style Language);
- зовнішній вигляд (стиль) – шрифти, розміри, відступи, кольори тощо, що описуються за допомогою таблиці стилів CSS (Cascade Style Sheet).

ВВР, який побудований за допомогою цього підходу, є «ідеальним» для консолідації даних, оскільки уся необхідна інформація вже є подана у вигляді структурованих XML-даних. Однак практично нема таких спеціалізованих ВВР, корисних для формування ТД.

Другий підхід передбачає експорт вмісту веб-сайту в спеціалізовані XML-формати, зокрема RSS і Atom.

RSS (Rich Site Summary) використовується для публікації потокової інформації (блоги, новини тощо), що часто змінюється [141]. RSS-документ містить повну або скорочену текстову інформацію, а також метадані як дата публікації чи ім'я автора. RSS дозволяє автоматизовано консолідувати дані з різних джерел [43].

Atom (або Atom Syndication Format) розроблений як альтернатива до RSS [146]. Цей формат широко використовується в різних сферах, включаючи журналістику, маркетинг тощо. Деякі веб-сайти надають можливість вибору формату експорту інформації між RSS- і Atom- форматами.

Незважаючи на відмінності між форматами RSS і Atom, які в основному стосуються переліку метаданих [149], вони однаково придатні для консолідації ТІ, оскільки подають інформацію в XML-форматі. Важливим обмеженням щодо їхнього застосування є сама наявність відповідного експорту на ВВР. Крім того використання цих форматів потребує додаткового дослідження на предмет достатності потокової інформації порівняно з тією, що відображається на веб-сайті.

Третій підхід полягає в аналізі структури HTML-сторінок і виявлення необхідних тегів та їх атрибутів, за допомогою яких маркуються дані. Типовими способами маркування даних є використання:

- тегів заголовків (h1, h2, h3), блоків (div, span), списків та їх елементів (ul, nl, li), таблиць та їх елементів (table, tr, td);
- атрибутів ідентифікаторів (id) і класів (class).

Аналіз структурованої інформації за допомогою HTML-тегів є універсальним способом, який доцільно використовувати для формування ТД з ВВР.

#### **1.4. Обґрунтування вибору методу оцінювання якості туристичної документації**

**Якість туристичної документації** – це категорія, яка відображає ступінь відповідності власних характеристик та властивостей обумовленим чи передбачуваним потребам та вимогам споживача туристичного продукту.

На сьогодні стандартизація туристичних послуг в Україні забезпечена такими основними стандартами, введеними в дію з 1 січня 1997 року:

- 1) міждержавний стандарт ГОСТ 28681.1-95 “Проектування туристичних послуг”, який встановлює порядок розробки документації при проектуванні туристичних послуг. Він призначений для підприємств, організацій різних організаційно-правових форм і громадян-підприємців, котрі надають туристичні послуги. Стандарт, який стосується проектування туристичних послуг, надає можливості розробляти й укомплектовувати туристичні подорожі як фізичним, так і юридичним особам. Адже, згідно з ним розробником проекту туристичних послуг може виступати як підприємство, так і громадянин-підприємець. Однак, з іншого боку, такі дії обмежуються Законом України “Про туризм”. Відповідно до нього на території нашої держави право щодо організації та забезпечення створення туристичного продукту мають туроператори;
- 2) міждержавний стандарт ГОСТ 28681.2-95 “Загальні вимоги до туристично-екскурсійного обслуговування”. Ним встановлені обов’язкові і рекомендаційні вимоги до якості туристичних послуг, які надають туристичні підприємства незалежно від їх відомчої приналежності і організаційно-правової форми;
- 3) міждержавний стандарт ГОСТ 28681.3-95 “Вимоги щодо забезпечення безпеки туристів і екскурсантів”. Цим стандартом установлені вимоги до туристичних та екскурсійних послуг, які забезпечують безпеку життя і здоров’я туристів та екскурсантів, методи їх контролю і потрібний для цілей обов’язкової сертифікації туристичних послуг. Цей стандарт призначений для підприємств усіх організаційно-правових форм і громадян-підприємців, котрі надають туристичні послуги населенню;
- 4) міждержавний стандарт ГОСТ 28681.4-95 “Класифікація готелів”. Даний стандарт встановлює класифікацію готелів (мотелів) різних організаційно-правових форм з кількістю номерів не менше 10.



Класифікація не розповсюджується на будинки і приміщення, які здаються в оренду для проживання. Стандарт придатний для використання з метою сертифікації готелів (мотелів).

Ці стандарти загалом встановлюють вимоги до туристичної діяльності, однак не порушують проблеми оцінювання якості ТД як технічної документації.

#### **1.4.1. Аналіз методів оцінювання якості, придатних для туристичної документації**

Неякісна ТД може спричинити неякісне надання туристичних послуг внаслідок неправильного інформування і як наслідок незадоволення туристів, негативний досвід яких поширюється і послаблює імідж туристичних установ. Побудова якісної ТД вимагає здійснення постійного моніторингу її якості, проведення її аналізу та формування сучасних вимог до якості, щоб ТД була доступною у користуванні і була максимально інформативною. Саме тому на сьогодні проблема побудови якісної ТД є нагальною для представників туристичного бізнесу, оскільки її розв'язання впливає на покращення якості туристичних послуг.

Якість ТД включає (рис. 1.10):

- якість інформаційного наповнення – текстова інформація повинна бути точною, повною, відповідати всім правилам граматики, орфографії та лексичним нормам;
- якість оформлення – візуальне подання інформації і структура не повинні відхилятися від норм та правил;
- якість подання інформації – інформація повинна бути подана з урахуванням рівня обізнаності та компетентності потенційного споживача.



Рис. 1.10. Компоненти якості документації

У цій роботі розглядаються лише питання якості інформаційного наповнення, тобто контенту ТД.

Якість інформації – сукупність властивостей, що зумовлюють та забезпечують можливості її використання для задоволення визначених потреб [35, 53, 65, 66, 71, 88, 114, 128, 135, 138, 161]. Згідно з національним стандартом України (відповідно до ISO 9000:2005), якість туристичної інформації – це категорія, яка відображає ступінь відповідності власних характеристик та властивостей обумовленим чи передбачуваним потребам та вимогам споживача ТП.

Виділимо кілька методів аналізу інформаційного наповнення ТД:

- звичайний;
- юридичний;
- соціально-психологічний
- автоматизований.

*Звичайний аналіз*, як зовнішній, так і внутрішній, полягає у вивченні суті матеріалу та логіки його викладення. Це звичайний аналіз на рівні розуміння текстової інформації з виявленням таких інформаційних прогалин, які не відповідають критеріям оформлення документів.

В результаті зовнішнього аналізу визначають вид документа, його автора, час і місце появи, цілі та ініціатора створення, а також рівень достовірності інформаційного наповнення.

На етапі внутрішнього аналізу досліджують зміст документа і виявляють рівень компетентності автора, від якого в першу чергу залежить стан якості створеного туристичного продукту. Адже лише компетентний розробник формує документацію відповідно до вимог та стандартів.

Для аналізу правових офіційних документів застосовують *юридичний аналіз*, в результаті якого виявляють правильність вживання юридичної документації та її відповідність критеріям оформлення.

*Соціально-психологічний аналіз* застосовують під час написання туристичної документації з урахуванням особистісних рис клієнта. Наприклад, це важливо під час створення індивідуальних туристичних маршрутів. Оскільки для ефективного забезпечення якості і конкурентоспроможності послуг важливу роль відіграє вивчення потреб клієнтів, а також способи задоволення потреб, доцільно досліджувати поведінку клієнтів, щоб реалізувати найбільш ефективне управління якістю послуг [25].

В умовах інтенсивного розвитку інформаційного суспільства та використання інтернет-технологій широкого застосування набув *метод автоматизованого аналізу* туристичної документації шляхом використання сучасних програмних засобів. Цей метод будемо застосовувати у цій роботі.

Побудова ТД направлена на споживача ТІ, тобто з погляду споживача ТД є ІІ. Тому до оцінювання якості ТД будемо застосовувати ті підходи, які вже є відомі для оцінювання якості таких ІІ як програмне забезпечення (ПЗ). Ефективність застосування таких підходів підтверджується результатами досліджень у суміжних галузях. Зокрема в [104] застосовано методи оцінювання якості ПЗ для ІІ у вигляді Веб-галерей, в [109] – до ІІ у вигляді Веб-форумів, а в [38] – до ІІ у вигляді баз даних.

Оскільки на сьогодні не відомі підходи щодо оцінювання якості ТД, то проаналізуємо придатність існуючих стандартів щодо оцінювання якості ПЗ для оцінювання якості ТД.

В Україні чинними є такі державні стандарти, пов'язані оцінюванням

якості ПЗ:

- ДСТУ 2844-94 [31];
- ДСТУ 2850-94 [29];
- ДСТУ ISO/IEC 12119:2003 [28].

Недоліком цих стандартів є те, що вони оперують не усіма показниками якості, які зустрічаються в міжнародному стандарті ISO/IEC-25010 [154] як продовженні стандарту ISO/IEC-9126 [150, 151].

Вагомий внесок у дослідження якості ПЗ, зокрема у вирішення задачі забезпечення якості життєвого циклу ПЗ, у своїх працях зробили Харченко В.С., Андон Ф.І., Липаєв В.В. та інші. Однак ми не будемо використовувати результати цих науковців, оскільки питання життєвого циклу ТД у цій роботі не розглядаються.

#### **1.4.2. Вибір методу оцінювання якості туристичної документації**

Метод оцінювання полягає у застосуванні показників якості стандарту ISO/IEC-25010 до оцінювання якості складових ТД. При цьому серед усіх показників якості необхідно вибрати ті, які відповідають вимогам побудови ТД як ІІІ.

ТД як ІІІ з погляду якості повинна володіти такими основними властивостями [55, 61, 72, 162, 163, 164]:

- забезпечувати певну множину структурних елементів, які наповнюються інформацією про туристичну діяльність;
- бути зрозумілою для споживача;
- простою та зручною у використанні;
- не містити прогалин в наповненні структурних елементів;
- містити актуальну інформацію;
- задовольняти потреби споживача.

Стандарт ISO/IEC-25010 містить такі характеристики якості ПЗ з погляду зовнішнього оцінювання [154]:

- 1) функціональність (Functional suitability) – відповідність функціональних можливостей ПЗ набору вимог користувача щодо функціональності за такими під-характеристиками:
  - відповідність (Appropriateness);
  - точність (Accuracy).
- 2) надійність (Reliability) – здатність ПЗ зберігати свій рівень якості функціонування у встановлених умовах за певний інтервал часу за такими під-характеристиками:
  - доступність (Availability);
  - відмовостійкість (Fault tolerance);
  - відновлюваність (Recoverability).
- 3) ефективність продуктивності (Performance efficiency) – співвідношення між рівнем якості функціонування ПЗ і обсягом використаних ресурсів при заданих умовах за такими під-характеристиками:
  - часова ефективність (Time behaviour);
  - використання ресурсів (Resource utilisation).
- 4) придатність до використання (Operability) – здатність ПЗ бути зрозумілим, вивченим і привабливим для користувача у заданих умовах використання за такими під-характеристиками:
  - відповідність розпізнавання (Appropriateness recognisability);
  - придатність до вивчення (Learnability);
  - простота використання (Ease of use);
  - рівень допомоги (Helpfulness);
  - привабливість (Attractiveness);
  - технічна доступність (Technical accessibility).
- 5) захищеність (Security) – рівень захисту ПЗ від випадкового або зловмисного доступу, використання, модифікації, знищення або оприлюднення за такими під-характеристиками:
  - конфіденційність (Confidentiality);

- цілісність (Integrity);
  - безвідмовність (Non-repudiation);
  - обліковність (Accountability);
  - аутентифікованість (Authenticity).
- 6) сумісність (Compatibility) – здатність ПЗ обмінюватися інформацією і/або виконувати необхідні функції в певних програмних і апаратних середовищах за такими під-характеристиками:
- замінність (Replaceability);
  - співіснування (Co-existence);
  - здатність до взаємодії (Interoperability).
- 7) супроводжуваність (Maintainability) – рівень обсягу робіт, необхідних для проведення змін (модифікацій) за такими під-характеристиками:
- модульність (Modularity);
  - можливість повторного використання (Reusability);
  - аналізованість (Analysability);
  - змінюваність (Changeability);
  - стійкість до змін (Modification stability);
  - тестованість (Testability).
- 8) переміщуваність (Transferability) – здатності ПЗ бути перенесеними з одного середовища в інше.
- мобільність (Portability);
  - адаптованість (Adaptability);
  - простота встановлення (Installability).

Аналіз метрик, що відповідають характеристикам/під-характеристикам якості ПЗ за стандартом ISO/IEC-25010, показує, що для оцінювання якості ТД придатними є такі характеристики/під-характеристики:

- 1) функціональність (Functional suitability):
  - відповідність (Appropriateness);
- 2) придатність до використання (Operability):

- відповідність розпізнавання (Appropriateness recognisability);
  - рівень допомоги (Helpfulness);
  - привабливість (Attractiveness);
- 3) супроводжуваність (Maintainability):
- модульність (Modularity);
  - стійкість до змін (Modification stability);

### **1.5. Висновки до розділу**

1. Проведено аналіз тенденцій розвитку туристичної галузі в умовах інформаційного суспільства, який підтвердив актуальність цієї роботи і дозволив визначити мету і задачі роботи.
2. Визначено особливості опрацювання туристичної інформації у вигляді документації, які стали основою для побудови інформаційної моделі туристичної документації.
3. Визначено підходи до використання відкритих веб-ресурсів для формування туристичної документації шляхом аналізу методів інтеграції інформаційних ресурсів і моделювання відкритих веб-ресурсів, що стало основою моделі інформаційних джерел формування туристичної документації.
4. Проаналізовано підходи до опису метаданих та структурування інформації в XML-технологіях, що дозволило побудувати інформаційну модель туристичної документації та інформаційного джерела з використанням індикативних ознак на основі лінгвістичних маркерів.
5. Обґрунтовано вибір методу оцінювання якості туристичної документації на основі стандарту ISO/IEC-25010, який є основою для розробки методів і процедур, що застосовуються під час формування туристичної документації.

## **Розділ 2. Розробка інформаційної моделі туристичної документації**

Характерною особливістю опрацювання туристичної інформації і 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.

В першому розділі відзначено, що туристична діяльність характеризується значною інформаційною насиченістю. Це пов'язано з тим, що в сучасних умовах дуже часто замовлення туристичних послуг здійснюється на основі відгуків інших споживачів туристичного продукту. Саме тому формування відповідного інформаційного забезпечення туристичної діяльності є необхідним і дієвим інструментом для покращення якості надання туристичних послуг.

Для організації процесу надання туристичних послуг необхідно перш за все привести туристичну інформацію до однорідного вигляду у вигляді документації. Для цього необхідно розробити інформаційну модель туристичної документації та методи її формування.

Основні результати розділу опубліковані автором у працях [142, 143, 155, 156, 157].

### **2.1. Розробка загальної моделі туристичної документації**

За визначенням 1.1 ТД розглядається як множина взаємопов'язаних електронних документів про туристичні послуги і/або продукти. При цьому обов'язковою складовою цих документів є текстова (числово-символьна) частина. З іншого боку орієнтованість на споживача передбачає розгляд ТД як інформації про певні туристичні об'єкти у вигляді фактів чи подій.

Робота з ТД з декількох ракурсів вимагає встановлення відповідних



логічних одиниць інформації. В ІС та БД виділяють три основні види логічної організації даних, які відрізняються між собою за логічною одиницею даних, якою оперують користувачі ІС і/або БД [38]:

- 1) фактографічний – логічною одиницею даних є факти чи події.
- 2) об'єктний – інформаційний об'єкт.
- 3) документальний – інформаційний документ.

Таким чином ТД повинна дозволяти одночасне подання ТІ або у вигляді окремих фактів чи подій, або у вигляді туристичних об'єктів, або у вигляді туристичних документів.

### **2.1.1. Побудова загальної структури туристичної документації**

Факти та події, подані як ТІ – це первинна інформація, яка входить в ТД. Тобто з погляду формування ТД вони як форми логічної організації даних вважаються неподільними. **Факти та події** – це елементарні (атомарні) відомості про туристичні послуги і туристичні об'єкти, які відповідно не мають та мають прив'язку до часу [38].

Джерелом фактів та подій є:

- інформація, подана на веб-сайтах надавачами ТП;
- інформація, яка відображена у відгуках споживачів ТП.

**Туристичний об'єкт (ТО)** – це сукупність фактів та подій, які цілісно відображають відомості про реальний об'єкт, цікавий для туристів. Важливою особливістю туристичних об'єктів є можливість здійснювати певну туристичну діяльність над ними (наприклад, відвідування, огляд, проходження тощо).

**Дії над туристичним об'єктом (ДТО)** – це сукупність фактів та подій, які цілісно відображають відомості про реальну діяльність над відповідним реальним об'єктом. Таким чином туристичні об'єкти та дії над ними описуються за однаковим принципом.

**Туристичний документ** – це сукупність фактів та подій або інших туристичних документів, які описують відомості про реальні об'єкти (один або декілька), дії над ними, туристичні послуги тощо з метою використання під час

туристичної діяльності.

Таким чином, туристичним документом слід вважати усю ТД або її частину, що має визначену мету використання. З поняття туристичного документа також впливає ієрархічна структура ТД.

На рис. 2.1 зображено загальну структуру туристичної документації у вигляді ієрархії (дерева) чотирьох рівнів:

- 1-ий рівень – туристична документація;
- 2-ий рівень – туристичні послуги та відгуки споживачів;
- 3-ий рівень – туристичні об'єкти і дії над ними, які стосуються ТП і відгуків;
- 4-ий рівень – факти та події, які стосуються ТО і ДТО.

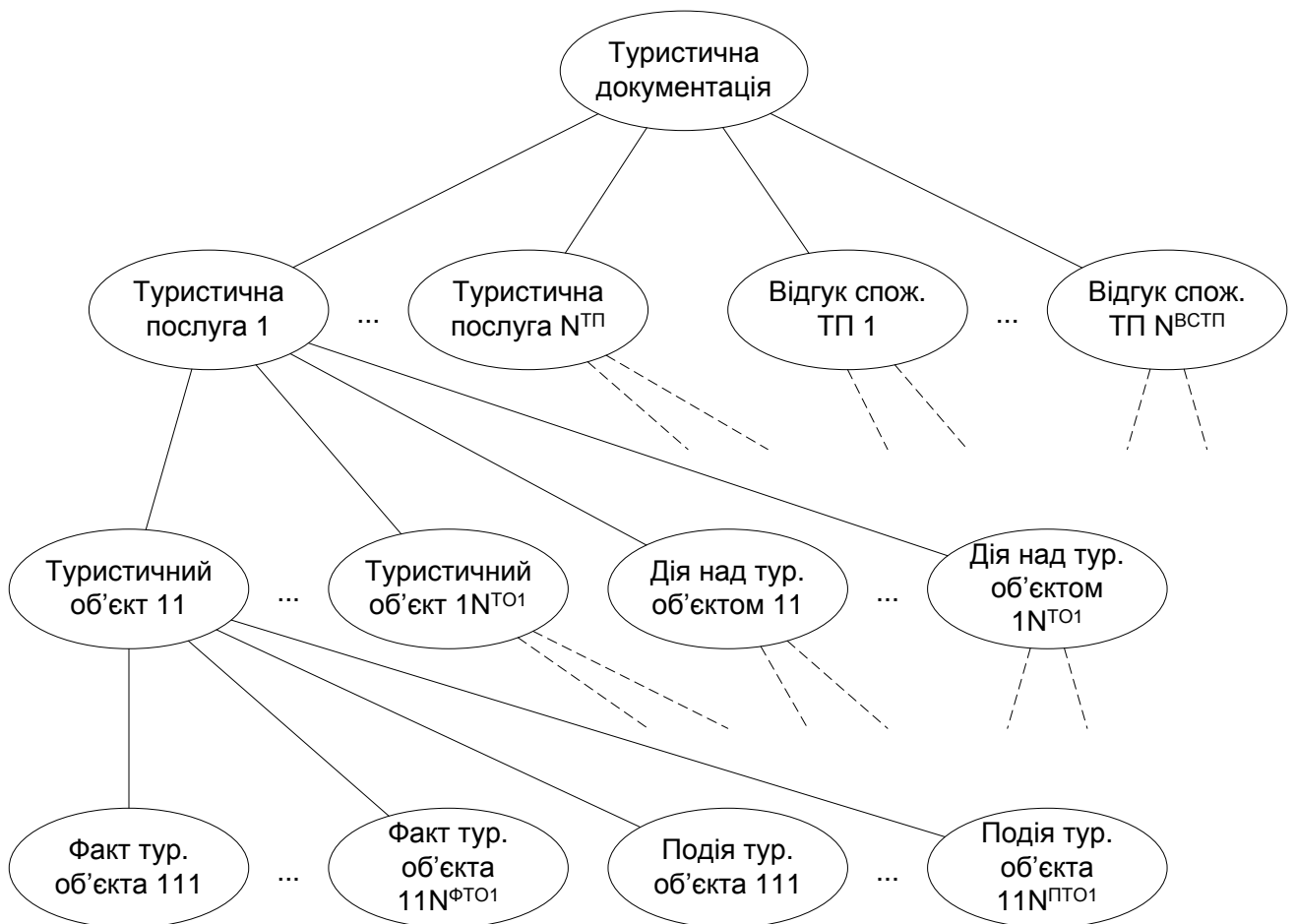


Рис. 2.1. Загальна структура туристичної документації

Отже, туристична документація складається з множини ТП та множини відгуків споживачів ТП. При цьому структура відгуків споживачів ТП є аналогічною до структури ТП, а структура дій над ТО є аналогічною до структури ТО. Елементарною одиницею інформації в структурі ТД є факти та події, які описують ТО та ДТО як в структурі ТП, так і в структурі відгуків споживачів ТП.

Зауважимо, що одні і ті ж ТО та ДТО можуть зустрічатися в різних ТП та відгуках споживачів ТП. Тобто насправді ТД подається у вигляді мережі, а не дерева, де кожен ТО та ДТО мав би пов'язуватися з усіма ТП та відгуками споживачів ТП. Однак така складніша структура ТД необхідна у першу чергу для побудови певної ІС, призначеної для формування, опрацювання та використання. Зокрема, це стосується проектування БД та побудови програмних інтерфейсів, що детально не розглядається у цій роботі. Спрощення мережної моделі до моделі у вигляді дерева використовується у методах редукції графів [155, 156, 157]. І власне цей підхід застосовано при побудові формальної моделі ТД, оскільки такої моделі є достатньо для застосування лінгвістичних підходів до автоматизованого формування ТД.

### 2.1.2. Формалізація загальної структури туристичної документації

На основі загальної структури ТД (рис. 2.1) побудуємо формальну модель ТД.

Подамо ТД у вигляді кортежа:

$$TD = \langle TS, TE \rangle, \quad (2.1)$$

де  $TD$  – туристична документація,  $TS$  – множина туристичних послуг,  $TE$  – множина відгуків споживачів туристичних послуг.

Якщо  $N^{(TS)}$  – кількість туристичних послуг в ТД, то множина туристичних послуг – це:

$$TS = \{TS_i\}_{i=1}^{N^{(TS)}}, \quad (2.2)$$

де  $TS_i$  –  $i$ -а туристична послуга в ТД.

Подамо ТП у вигляді кортежа:

$$TS_i = \langle TSO_i, TSA_i \rangle, \quad (2.3)$$

де  $TSO_i$  – множина об'єктів ТП (тобто ТО),  $TSA_i$  – множина дій над об'єктами ТП (тобто ДТО).

Нехай  $Object$  – множина усіх ТО, які можуть зустрічатися в ТД, а  $Action(Object)$  – множина відповідних ДТО, яка є підмножиною усіх дій в ТД:

$$Action(Object) = \{Action(Object)_i\}_{i=1}^{N^{(Action(Object))}}, \quad (2.4)$$

де  $Action(Object)_i$  –  $i$ -а дія,  $N^{(Action(Object))}$  – загальна кількість дій.

Тоді якщо  $N^{(TSO_i)}$  – кількість об'єктів ТП, то множина ТО – це:

$$TSO = \{TSO_{ij}\}_{j=1}^{N^{(TSO_i)}}, \quad (2.5)$$

де  $TSO_{ij} \in Object$  –  $j$ -й об'єкт ТП.

Аналогічно, якщо  $N^{(TSA_i)}$  – кількість дій над об'єктами ТП, то множина ДТО – це:

$$TSA = \{TSA_{ij}\}_{j=1}^{N^{(TSA_i)}}, \quad (2.6)$$

де  $TSA_{ij} \in Action(Object)$  –  $j$ -а дія над об'єктами ТП.

Подібно до (2.2), (2.3), (2.5), (2.6) формалізуємо поняття відгуків споживачів ТП та їх складових.

Якщо  $N^{(TE)}$  – кількість відгуків споживачів ТП в ТД, то множина відгуків споживачів туристичних послуг – це:

$$TE = \{TE_i\}_{i=1}^{N^{(TE)}}, \quad (2.7)$$

де  $TE_i$  –  $i$ -а відгук споживачів ТП в ТД.

Відгук споживачів ТП подамо у вигляді кортежа:

$$TE_i = \langle TEO_i, TEA_i \rangle, \quad (2.8)$$

де  $TEO_i$  – множина відгуків про ТО,  $TEA_i$  – множина відгуків про ДТО.

Нехай  $Experience$  – множина усіх відгуків про ТО або ДТО, які зустрічаються в ТД. Тобто:

$$Experience = Experience(Object) \cup Experience(Action(Object)), \quad (2.9)$$

де  $Experience(Object)$  і  $Experience(Action(Object))$  – відповідно відгуки про усі ТО і ДТО.

Якщо  $N^{(TEO_i)}$  – кількість відгуків про об'єкти ТП, то множина відгуків про ТО – це:

$$TEO = \{TEO_{ij}\}_{j=1}^{N^{(TEO_i)}}, \quad (2.10)$$

де  $TEO_{ij} \in Experience$  –  $j$ -й відгуків про ТО.

Аналогічно, якщо  $N^{(TEA_i)}$  – кількість відгуків про дії над об'єктами ТП, то множина відгуків про ДТО – це:

$$TEA = \{TEA_{ij}\}_{j=1}^{N^{(TEA_i)}}, \quad (2.11)$$

де  $TEA_{ij} \in Experience$  –  $j$ -й відгук про ДТО.

### 2.1.3. Формалізація структури туристичних об'єктів та дій над ними

Кожен ТО з множини  $Object$  містить інформацію про факти, які відображають загальну інформацію про цей об'єкт, і інформацію про події, які трапилися з цим об'єктом в певні моменти чи терміни часу.

Нехай  $Fact$  і  $Event$  – відповідно множини усіх фактів та подій, які містяться в ТД.

Якщо  $N^{(Object)}$  – загальна кількість елементів множини  $Object$ , то:

$$Object = \{Object_i\}_{i=1}^{N(Object)}. \quad (2.12)$$

Тоді кожен ТО  $Object_i$  з множини усіх можливих об'єктів ТД – це кортеж такого вигляду:

$$Object_i = \langle ODefinition_i, OFact_i, OEvent_i \rangle, \quad (2.13)$$

де  $ODefinition_i$  – множина атрибутів ТО,  $OFact_i$  – множина фактів про ТО,  $OEvent_i$  – множина подій про ТО.

Якщо  $N^{(OFact_i)}$  – кількість фактів про ТО, то множина фактів про ТО – це:

$$OFact_i = \{OFact_{ij}\}_{j=1}^{N^{(OFact_i)}}, \quad (2.14)$$

де  $OFact_{ij} \in Fact$  –  $j$ -й факт про ТО.

Якщо  $N^{(OEvent_i)}$  – кількість подій про ТО, то множина подій про ТО – це:

$$OEvent_i = \{OEvent_{ij}\}_{j=1}^{N^{(OEvent_i)}}, \quad (2.15)$$

де  $OEvent_{ij} \in Event$  –  $j$ -а подія про ТО.

З (2.12) випливає, що уся множина ДТО  $Action(Object)$  визначається через дії по кожному ТО з  $Object$ :

$$Action(Object) = \{Action(Object_i)\}_{i=1}^{N(Object)}, \quad (2.16)$$

де  $Action(Object_i)$  – дії над  $i$ -м ТО.

Нехай  $N^{(Action(Object_i))}$  – кількість дій над  $i$ -тим ТО. Тоді дії над  $i$ -м ТО – це така множина:

$$Action(Object_i) = \{Action(Object_i)_j\}_{j=1}^{N^{(Action(Object_i))}}, \quad (2.17)$$

де  $Action(Object_i)_j$  –  $j$ -а дія над  $i$ -м ТО.

Кожну дію над ТО подамо у вигляді кортежу:

$$Action(Object_i)_j = \langle ADefinition_{ij}, AFact_{ij}, AEvent_{ij} \rangle, \quad (2.18)$$

де  $ADefinition_{ij}$  – множина атрибутів  $j$ -тої дії над ТО,  $AFact_{ij}$  – множина фактів про  $j$ -ту дію над ТО,  $AEvent_{ij}$  – множина подій про  $j$ -ту дію над ТО.

За аналогією до (2.14), якщо  $N^{(AFact_{ij})}$  – кількість фактів про  $j$ -ту дію над ТО, то відповідна множина фактів про ДТО – це:

$$AFact_{ij} = \{AFact_{ijk}\}_{k=1}^{N^{(AFact_{ij})}}, \quad (2.19)$$

де  $AFact_{ijk} \in Fact$  –  $j$ -й факт про ДТО.

За аналогією до (2.15), якщо  $N^{(AEvent_{ij})}$  – кількість подій про  $j$ -у дію над ТО, то множина подій про відповідну ДТО – це:

$$AEvent_{ij} = \{AEvent_{ijk}\}_{k=1}^{N^{(AEvent_{ij})}}, \quad (2.20)$$

де  $AEvent_{ijk} \in Event$  –  $j$ -а подія про ДТО.

З (2.14), (2.15), (2.19), (2.20) впливають такі співвідношення між множинами фактів та подій ТД:

$$OFact_i \subseteq Fact, AFact_{ij} \subseteq Fact.$$

$$OEvent_i \subseteq Event, AEvent_{ij} \subseteq Event.$$

#### 2.1.4. Формалізація структури відгуків споживачів

Розглянемо відгук споживача про ТО або ДТО як набір фактів, які відображають загальну думку споживача про цей об'єкт чи дію, і подій, які трапилися з цим споживачем і пов'язані з цим об'єктом чи дією. Будемо вважати, що факти і події, що відображають відгук споживача, також належать відповідно до множин  $Fact$  і  $Event$ .

На основі (2.9) та (2.12) подамо відгуки про усі ТО як множину відгуків по кожному окремому ТО:

$$Experience(Object) = \{Experience(Object_i)\}_{i=1}^{N^{(Object)}}, \quad (2.21)$$

де  $Experience(Object_i)$  – відгуки до  $i$ -го ТО.

Нехай  $N^{(Experience(Object_i))}$  – кількість відгуків до  $i$ -го ТО. Тоді усі відгуки до  $i$ -го ТО становлять таку множину:

$$Experience(Object_i) = \left\{ Experience(Object_i)_j \right\}_{j=1}^{N^{(Experience(Object_i))}}, \quad (2.22)$$

де  $Experience(Object_i)_j$  –  $j$ -й відгук до  $i$ -го ТО.

Кожен відгук до ТО подамо у вигляді кортежу:

$$Experience(Object_i)_j = \langle EODefinition_{ij}, EOFact_{ij}, EOEvent_{ij} \rangle, \quad (2.23)$$

де  $EODefinition_{ij}$  – множина атрибутів  $j$ -го відгуку про ТО,  $EOFact_{ij}$  – множина фактів у  $j$ -му відгуку про ТО,  $EOEvent_{ij}$  – множина подій у  $j$ -му відгуку про ТО.

Якщо  $N^{(EOFact_{ij})}$  – кількість фактів у  $j$ -му відгуку про ТО, то множина фактів у відгуку про ТО – це:

$$EOFact_{ij} = \left\{ EOFact_{ijk} \right\}_{k=1}^{N^{(EOFact_{ij})}}, \quad (2.24)$$

де  $EOFact_{ijk} \in Fact$  –  $k$ -тий факт у відгуку про ТО.

Якщо  $N^{(EOEvent_{ij})}$  – кількість подій у  $j$ -му відгуку про ТО, то множина подій у відгуку про ТО – це:

$$EOEvent_{ij} = \left\{ EOEvent_{ijk} \right\}_{k=1}^{N^{(EOEvent_{ij})}}, \quad (2.25)$$

де  $EOEvent_{ijk} \in Event$  –  $k$ -а подія у відгуку про ТО.

На основі (2.9), (2.16) та (2.17) подамо відгуки про ДТО як множину відгуків до дій над кожним ТО:

$$Experience(Action(Object)) = \left\{ \left\{ Experience(Action(Object)_i)_j \right\}_{j=1}^{N^{(Action(Object)_i)}} \right\}_{i=1}^{N^{(Object)}}, \quad (2.26)$$

де  $Experience(Action(Object)_i)_j$  – відгуки до  $j$ -ї дії над  $i$ -м ТО.



Нехай  $N^{(Experience(Action(Object_i)_j))}$  – кількість відгуків до  $j$ -ї дії над  $i$ -м ТО. Тоді усі відгуки до цієї ДТО становлять таку множину:

$$Experience(Action(Object_i)_j) = \left\{ Experience(Action(Object_i)_j)_k \right\}_{k=1}^{N^{(Experience(Action(Object_i)_j))}}, \quad (2.27)$$

де  $Experience(Action(Object_i)_j)_k$  –  $k$ -й відгук до  $j$ -ї дії над  $i$ -м ТО.

Кожен відгук до дії над ТО подамо у вигляді кортежу:

$$Experience(Action(Object_i)_j)_k = \langle EADefinition_{ijk}, EAFact_{ijk}, EAEvent_{ijk} \rangle, \quad (2.28)$$

де  $EADefinition_{ijk}$  – множина атрибутів  $k$ -го відгуку про ДТО,  $EAFact_{ijk}$  – множина фактів у  $k$ -му відгуку про ДТО,  $EAEvent_{ijk}$  – множина подій у  $k$ -му відгуку про ДТО.

За аналогією до (2.24), якщо  $N^{(EAFact_{ijk})}$  – кількість фактів у  $k$ -му відгуку про  $j$ -у дію над ТО, то відповідна множина фактів у відгуку про ДТО – це:

$$EAFact_{ijk} = \left\{ EAFact_{ijkl} \right\}_{l=1}^{N^{(EAFact_{ijk})}}, \quad (2.29)$$

де  $EAFact_{ijkl} \in Fact$  –  $l$ -й факт у відгуку про ДТО.

За аналогією до (2.25), якщо  $N^{(EAEvent_{ijk})}$  – кількість подій у  $k$ -му відгуку на  $j$ -у дію над ТО, то відповідна множина подій у відгуку про відповідну ДТО – це:

$$EAEvent_{ijk} = \left\{ EAEvent_{ijkl} \right\}_{l=1}^{N^{(EAEvent_{ijk})}}, \quad (2.30)$$

де  $EAEvent_{ijkl} \in Event$  –  $l$ -а подія у відгуку про ДТО.

З (2.24), (2.25), (2.29), (2.30) випливають такі співвідношення між множинами фактів та подій у відгуках ТД:

$$EOFact_{ij} \subseteq Fact, EAFact_{ijk} \subseteq Fact.$$

$$EOEvent_{ij} \subseteq Event, EAEvent_{ijk} \subseteq Event.$$

## 2.2. Розробка моделі елементарних одиниць інформації у туристичній документації

Елементарними одиницями інформації у ТД є факти та події. З фактів та подій складаються такі складові ТД як туристичні об'єкти (2.13), дії над туристичними об'єктами (2.18), відгуки до туристичних об'єктів (2.23), відгуки до дій над туристичними об'єктами (2.28). З іншого боку кожна елементарна одиниця інформації ТД доцільно розглядати як окремий документ, який фіксує відомості про ТО, ДТО, відгуки до ТО та ДТО на основі інформації, що міститься у відповідному ВВР. Таким чином опис кожного факту і події повинен ґрунтуватися на елементах опису метаданих Дублінського ядра.

### 2.2.1. Формалізація фактів у туристичній документації

Будемо вважати, що факт – це встановлення значень характеристик (властивостей) певного поняття в ТД (ТО, ДТО, відгуки до ТО та ДТО) без прив'язки до того, що ці значення встановлені в певний момент часу в дійсності. Іншими словами факт – це набір значень атрибутів, які описують цей факт.

Усі факти ТД поділимо на дві групи:

- факти визначення понять (ТО, ДТО, відгуків до ТО та ДТО);
- факти, що описують інші характеристики понять.

Факти визначення понять описуються так:

- $ODefinition_i$  – атрибутами факту визначення  $i$ -того ТО (2.13);
- $ADefinition_{ij}$  – атрибутами факту визначення  $j$ -ї дії над  $i$ -м ТО (2.18);
- $EODefinition_{ij}$  – атрибутами факту визначення  $j$ -го відгуку про  $i$ -й ТО (2.23);
- $EADefinition_{ijk}$  – атрибутами факту визначення  $k$ -го відгуку про  $j$ -у дію над  $i$ -м ТО (2.28).

Нехай *Classifier* – множина класифікаторів ТД. Тобто:

$$Classifier = \{Classifier_i\}_{i=1}^{N^{(Classifier)}}, \quad (2.31)$$

де  $N^{(Classifier)}$  – загальна кількість класифікаторів.

Кожен класифікатор подамо як словник термінів у вигляді кортежу термінів та зв'язків між ними:

$$Classifier_i = \langle Term_i, TermRels_i \rangle, \quad (2.32)$$

де  $Term_i = \{Term_{ij}\}_{j=1}^{N^{(Term_i)}}$  - множина термінів  $i$ -го класифікатора,  $N^{(Term_i)}$  - кількість термінів  $i$ -го класифікатора,  $TermRels_i \subseteq \{\{Term_{ij}, Term_{ik}\}\}_{j,k=1}^{N^{(Term_i)}} \equiv Term_i \times Term_i$  - множина зв'язків між термінами, яка є підмножиною декартового добутку двох множин термінів  $i$ -го класифікатора.

Фактично класифікатор – це певний граф, вершинами якого є терміни, а дуги – зв'язки між термінами. Найчастіше на практиці класифікатор подається у вигляді дерева або лісу.

На основі проведеного аналізу туристичної діяльності та розробленої загальної моделі ТД виділимо основні класифікатори, які повинні бути присутніми у будь-якій ТД:

- типи поняття – типи ТО (місто, будівля, пам'ятник, замок, вулиця, урочище, річка, канал тощо), типи ДТО (відвідування, проходження, оглядання, пропливання тощо), типи відгуків про ТО, типи відгуків (особиста думка, переказ, позитив, негатив тощо) про ДТО (особиста участь, переказ, позитив, негатив тощо);
- розділи ТД – умови проживання, харчування, розміщення, визначні місця, розваги, транспорт, інформація про персонал та ін.;
- тип документа – документи для клієнта, документи на підтвердження особи туриста, документи промоційної кампанії, документи про туристичну привабливість, документи про екологічну ситуацію, документи про зручності пересування та ін.;

Подамо атрибути кожного факту визначення як кортеж обов'язкових атрибутів і термінів класифікаторів:

$$O\text{Definition}_i = \langle \text{MandatoryProperties}(\text{Object}_i), \text{Classifier}(\text{Object}_i) \rangle, \quad (2.33)$$

$$A\text{Definition}_{ij} = \left\langle \begin{array}{l} \text{MandatoryProperties}(\text{Action}(\text{Object}_i)_j) \\ \text{Classifier}(\text{Action}(\text{Object}_i)_j) \end{array} \right\rangle, \quad (2.34)$$

$$EO\text{Definition}_{ij} = \left\langle \begin{array}{l} \text{MandatoryProperties}(\text{Experience}(\text{Object}_i)_j) \\ \text{Classifier}(\text{Experience}(\text{Object}_i)_j) \end{array} \right\rangle, \quad (2.35)$$

$$EO\text{Definition}_{ij} = \left\langle \begin{array}{l} \text{MandatoryProperties}(\text{Experience}(\text{Action}(\text{Object}_i)_j)_k) \\ \text{Classifier}(\text{Experience}(\text{Action}(\text{Object}_i)_j)_k) \end{array} \right\rangle, \quad (2.36)$$

де  $\text{MandatoryProperties}(\dots)$  – обов'язкові атрибути поняття як Ідентифікатор, Заголовок/Назва, Опис,  $\text{Classifier}(\dots)$  – множина термінів класифікаторів, до яких відноситься поняття.

Факти в (2.14), (2.19), (2.24), (2.29), що описують інші характеристики відповідно ТО, ДТО, відгуків до ТО та ДТО, окрім обов'язкових атрибутів містять також атрибути Дублінського ядра, пов'язані з інформацією про джерело (тобто ВВР), автора, мову публікації:

$$O\text{Fact}_{ij} = \left\langle \begin{array}{l} \text{MandatoryProperties}(O\text{Fact}_{ij}), \text{Classifier}(O\text{Fact}_{ij}) \\ \text{Source}(O\text{Fact}_{ij}), \text{Author}(O\text{Fact}_{ij}), \text{Language}(O\text{Fact}_{ij}) \end{array} \right\rangle, \quad (2.37)$$

$$A\text{Fact}_{ijk} = \left\langle \begin{array}{l} \text{MandatoryProperties}(A\text{Fact}_{ijk}), \text{Classifier}(A\text{Fact}_{ijk}) \\ \text{Source}(A\text{Fact}_{ijk}), \text{Author}(A\text{Fact}_{ijk}), \text{Language}(A\text{Fact}_{ijk}) \end{array} \right\rangle, \quad (2.38)$$

$$EO\text{Fact}_{ijk} = \left\langle \begin{array}{l} \text{MandatoryProperties}(EO\text{Fact}_{ijk}), \text{Classifier}(EO\text{Fact}_{ijk}) \\ \text{Source}(EO\text{Fact}_{ijk}), \text{Author}(EO\text{Fact}_{ijk}), \text{Language}(EO\text{Fact}_{ijk}) \end{array} \right\rangle, \quad (2.39)$$

$$EAFact_{ijkl} = \left\langle \begin{array}{l} MandatoryProperties(EAFact_{ijkl}), Classifier(EAFact_{ijkl}), \\ Source(EAFact_{ijkl}), Author(EAFact_{ijkl}), Language(EAFact_{ijkl}) \end{array} \right\rangle, \quad (2.40)$$

де  $MandatoryProperties(...)$  – обов'язкові атрибути відповідного факту,  $Classifier(...)$  – множина термінів класифікаторів, до яких відноситься факт,  $Source(...)$  – атрибути джерела інформації як Видавець, URL-адреса,  $Author(...)$  – атрибути про автора поданої інформації,  $Language(...)$  – атрибути мови інформації.

### 2.2.2. Формалізація подій у туристичній документації

Будемо вважати, що подія – це встановлення значень характеристик (властивостей) певного поняття в ТД (ТО, ДТО, відгуки до ТО та ДТО) з обов'язковою прив'язкою до моменту часу в дійсності, коли ці значення були встановлені. Подія визначається набором значень атрибутів включно з часовими атрибутами, які описують цю подію.

Подамо атрибути кожної події в (2.15), (2.20), (2.25), (2.30), які описують відповідно ТО, ДТО, відгуків до ТО та ДТО, як кортеж обов'язкових атрибутів, термінів класифікаторів, атрибутів джерела, автора, мови публікації, часу публікації:

$$OEvent_{ij} = \left\langle \begin{array}{l} MandatoryProperties(OEvent_{ij}), Classifier(OEvent_{ij}), \\ Source(OEvent_{ij}), Author(OEvent_{ij}), Language(OEvent_{ij}), \\ Time(OEvent_{ij}) \end{array} \right\rangle, \quad (2.41)$$

$$AEvent_{ijk} = \left\langle \begin{array}{l} MandatoryProperties(AEvent_{ijk}), Classifier(AEvent_{ijk}), \\ Source(AEvent_{ijk}), Author(AEvent_{ijk}), Language(AEvent_{ijk}), \\ Time(AEvent_{ijk}) \end{array} \right\rangle, \quad (2.42)$$

$$EOEvent_{ijk} = \left\langle \begin{array}{l} MandatoryProperties(EOEvent_{ijk}), Classifier(EOEvent_{ijk}), \\ Source(EOEvent_{ijk}), Author(EOEvent_{ijk}), Language(EOEvent_{ijk}), \\ Time(EOEvent_{ijk}) \end{array} \right\rangle, \quad (2.43)$$

$$EAEvent_{ijkl} = \left\langle \begin{array}{l} MandatoryProperties(EAEvent_{ijkl}), Classifier(EAEvent_{ijkl}), \\ Source(EAEvent_{ijkl}), Author(EAEvent_{ijkl}), Language(EAEvent_{ijkl}), \\ Time(EAEvent_{ijkl}) \end{array} \right\rangle, (2.44)$$

де  $MandatoryProperties(...)$  – обов'язкові атрибути відповідної події,  $Classifier(...)$  – множина термінів класифікаторів, до яких відноситься подія,  $Source(...)$  – атрибути джерела,  $Author(...)$  – атрибути про автора,  $Language(...)$  – атрибути мови,  $Time(...)$  – атрибути, пов'язані з часовими властивостями публікації як Дійсний час події, Час публікації на сайті джерела.

### 2.3. Класифікація туристичної документації за особливостями загальної моделі

Туристична документація складається з інформації про туристичні послуги та задокументовані відгуки споживачів цих ТП. Рис. 2.1 відображає загальне подання ТД, а вирази (2.1)-(2.30) описують відповідну формальну модель ТД та її складових. Крім загального випадку ТД, коли наявні усі її складові, очевидно можливі випадки, коли деякі частини ТД відсутні. Це може бути пов'язано або з певними етапами формування ТД, або з метою формування ТД. Виділення особливостей ТД дозволяє обрати відповідні методи та засоби щодо її доповнення або переструктурування, а також визначити способи її подання для споживачів ТП.

Розглянемо ці часткові випадки ТД детальніше і проведемо відповідну їхню класифікацію. Виділимо такі основні варіанти сформованої ТД:

- візитка про туристичні послуги;
- база знань споживачів туристичних послуг.

#### 2.3.1. Аналіз особливостей туристичної документації у вигляді візитки

ТД, яка подається у вигляді візитки, очевидно відображає лише інформацію про туристичну діяльність і не відображає інформацію про відгуки споживачів. Тобто вираз (2.1) спрощується:

$$TD = \langle TS \rangle. \quad (2.45)$$

Під час формування ТД у вигляді візитки доцільно окремо виділити такі етапи:

1) ТД містить лише загальну інформацію про ТО;

ТД включає лише загальну інформацію про ТО, але не містить інформації про окремі факти і події, пов'язані з ТО. Тобто (2.3), (2.13) подаються відповідно так:

$$TS_i = \langle TSO_i \rangle, \quad (2.46)$$

$$Object_i = \langle ODefinition_i \rangle, \quad (2.47)$$

Така модель ТД може застосовуватися для створення списку об'єктів, словників, які подаються споживачам у вигляді брошури з переліком назв ТО (рис. 2.2).

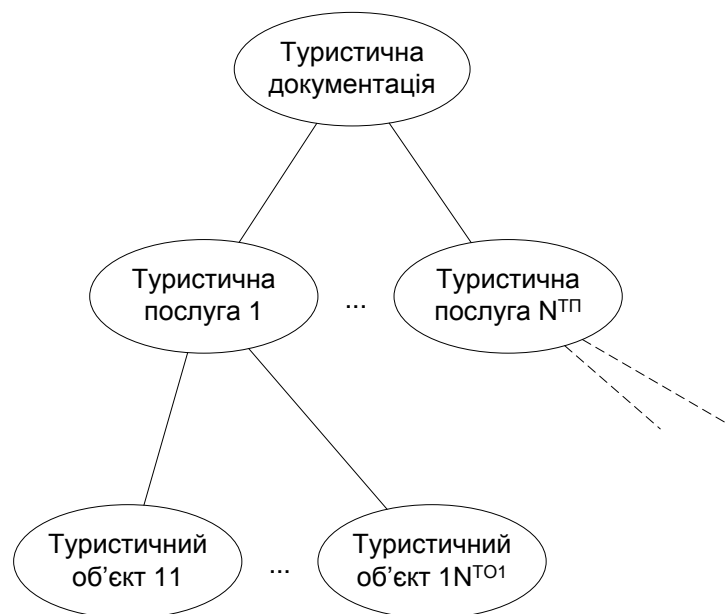


Рис. 2.2. Загальна структура туристичної документації у вигляді словника

Приклад 2.1. Прикладом ТД у вигляді словника є перелік доступних доступних екскурсій:

- вечірній Львів (Оперний театр, Ратуша,...);

- середньовічний Львів (Успенська церква, Підземелля,...).

2) ТД містить загальну інформацію про ТО і ДТО;

Цей етап доповнює ТД попереднього етапу загальною інформацією про ДТО без інформації про окремі факти і події, пов'язані з ДТО. Тобто (2.3) не вироджується в (2.46), (2.13) вироджується в (2.47), а (2.18) подається так:

$$Action(Object_i)_j = \langle ADefinition_{ij} \rangle. \quad (2.48)$$

Ця модель ТД використовується на етапі інформування про ТП, наприклад, під час роботи інформаційного бюро, коли надаються загальні інформаційні послуги про ТП, але без детальної специфікації того, що відбувається під час надання послуги (рис. 2.3).

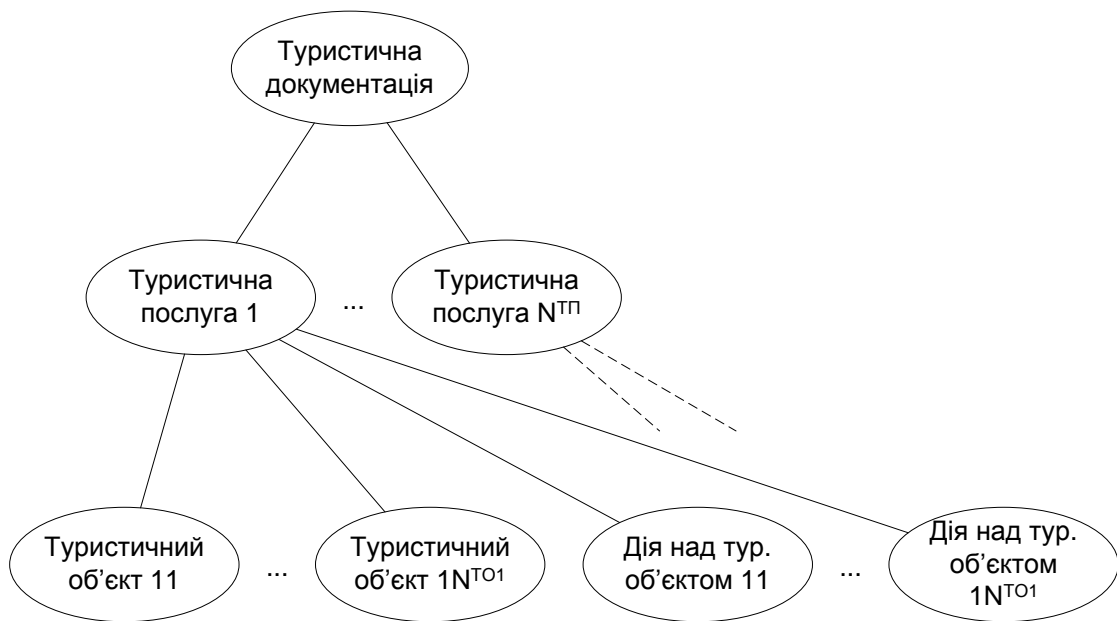


Рис. 2.3. Загальна структура туристичної документації для інформаційного бюро

3) ТД містить статичну інформацію про ТО і ДТО.

На цьому етапі ТД попереднього етапу містить загальну інформацію про ТО та ДТО, яка доповнена додатковими фактами. Тобто (2.3) не вироджується в (2.46), а (2.13) і (2.18) подаються так:

$$Object_i = \langle ODefinition_i, OFact_i \rangle, \quad (2.49)$$



$$Action(Object_i)_j = \langle ADefinition_{ij}, AFact_{ij} \rangle. \quad (2.50)$$

Особливістю цієї моделі є відображення лише статичної інформації про ТП без врахування того, що ТО мають певну історію, ДТО можуть мати певну тривалість. Такий вид ТД використовують для формування розширених словників (рис. 2.4).



Рис. 2.4. Загальна структура туристичної документації у вигляді розширеного словника

### 2.3.2. Аналіз особливостей туристичної документації у вигляді бази знань споживачів

Під базою знань на веб-ресурсах, призначених для функціонування спільнот, найчастіше розуміють збірку статей, які створюються для узагальнення і накопичення досвіду користувачів [52]. Така база знань відображається на загальнодоступному інформаційно-довідковому веб-ресурсі.

Будемо вважати, що ТД побудована у вигляді бази знань, якщо вона містить факти і події лише у відгуках споживачів. Тобто, вираз (2.1) і (2.3) не вироджуються у відповідно (2.45) і (2.46), але мають місце вирази (2.47) і (2.48).

За аналогією до ТД у вигляді візитки розглянемо етапи формування ТД у вигляді бази знань споживачів:

1) відгуки споживачів ТД містять лише загальну інформацію про ТО;

ТД включає лише загальну інформацію про ТО у відгуках споживачів, але не містить інформації про окремі факти і події в цих відгуках. Тобто (2.8) і (2.23) подаються відповідно так:

$$TE_i = \langle TEO_i \rangle, \quad (2.51)$$

$$Experience(Object_i)_j = \langle EODefinition_{ij} \rangle, \quad (2.52)$$

Така модель ТД може служити для створення словника споживача, що містить перелік назв ТО, які зустрічаються у відгуках (рис. 2.5).

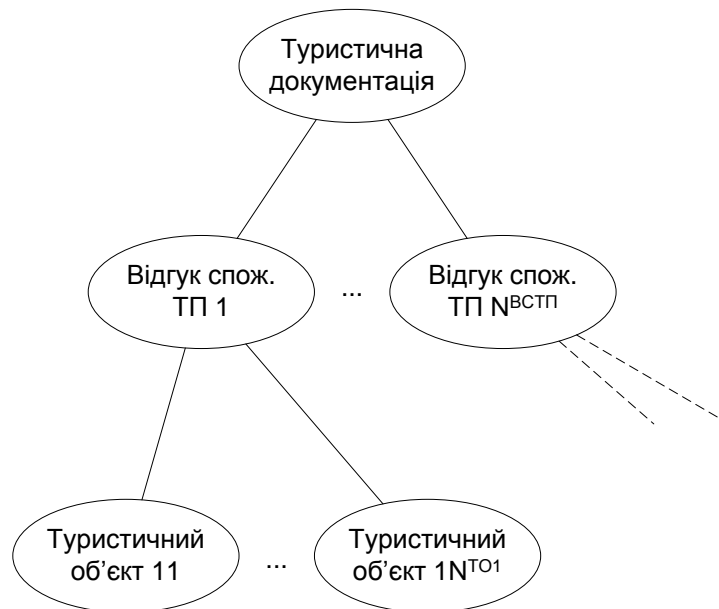


Рис. 2.5. Загальна структура туристичної документації у вигляді словника споживача.

2) відгуки споживачів ТД містять загальну інформацію про ТО і ДТО;

На цьому етапі ТД доповнюється загальною інформацією про ДТО без інформації про окремі факти і події, пов'язані з ДТО. Тобто (2.8) не вироджується в (2.51), (2.23) вироджується в (2.52), а (2.28) подається так:

$$Experience\left>Action\left(Object_i\right)_j\right)_k = \langle EADefinition_{ijk} \rangle. \quad (2.53)$$

Ця модель ТД фактично містить каркас (зміст) бази знань споживачів, але не включає детальних відгуків (рис. 2.6).



Рис. 2.6. Загальна структура туристичної документації у вигляді каркасу бази знань споживачів.

3) Відгуки споживачів ТД містять статичну інформацію про ТО і ДТО;

На цьому етапі ТД з попереднього етапу доповнюється фактами у відгуках споживачів про ТО та ДТО. Тобто (2.8) не вироджується в (2.51), а (2.23) і (2.28) подаються так:

$$Experience\left(Object_i\right)_j = \langle EODefinition_{ij}, EOFact_{ij} \rangle, \quad (2.54)$$

$$Experience\left>Action\left(Object_i\right)_j\right)_k = \langle EADefinition_{ijk}, EAFact_{ijk} \rangle, \quad (2.55)$$

Ця модель ТД відображає лише статичну інформацію у відгуках споживачів без врахування того, що ця інформація може мати певні прив'язки до моментів і термінів часу (рис. 2.7).

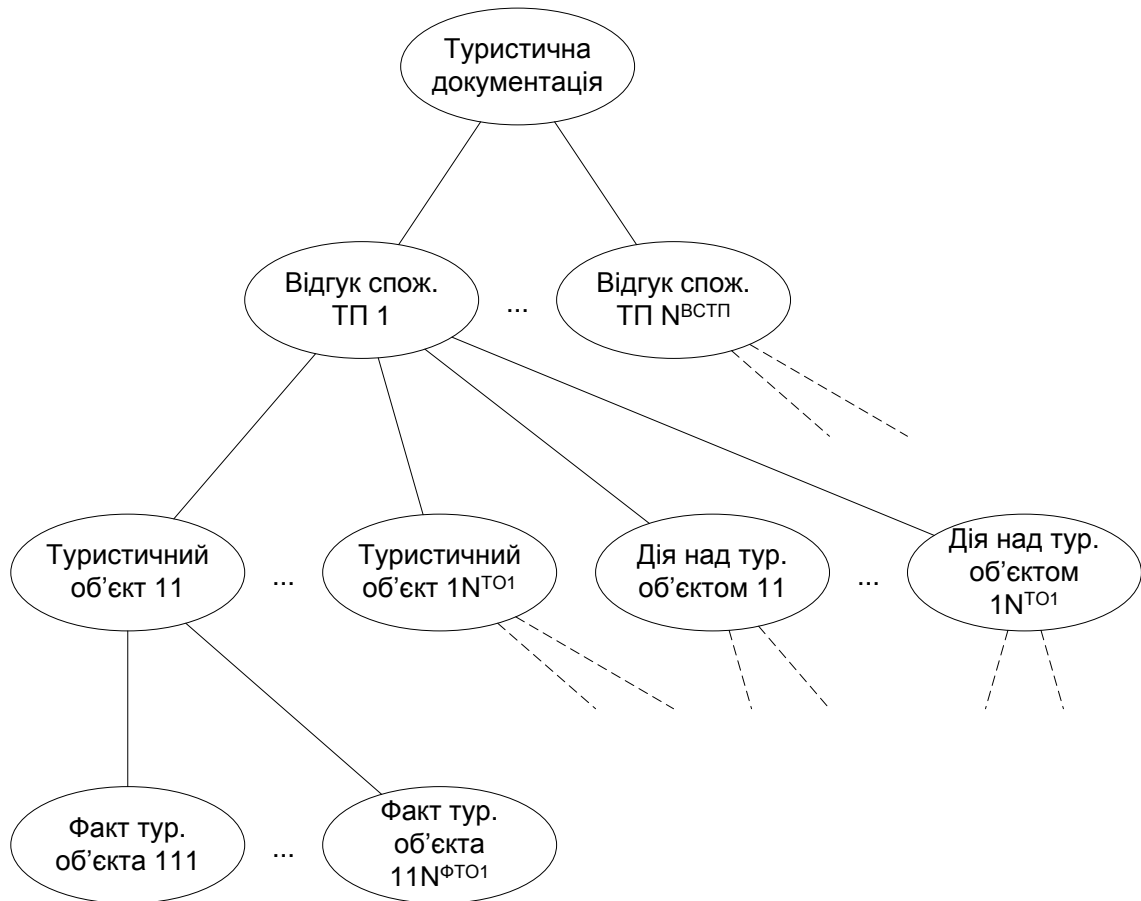


Рис. 2.7. Загальна структура туристичної документації у вигляді статичної бази знань споживачів

## 2.4. Розробка моделі інформаційних джерел формування туристичної документації

У цій роботі ми розглядаємо формування ТД лише на основі відкритих Веб-ресурсів, які доступні для використання без процедур авторизації (тобто анонімно). До таких ВВР відносять веб-спільноти, довідкові веб-ресурси, веб-сховища, спеціалізовані веб-сайти. Аналіз XML-засобів структурування інформації туристичної документації показав, що незалежно від виду ВВР існує

три основні методи структурування інформації: за допомогою XML-даних, стандартних XML-мов та HTML-тегів.

#### 2.4.1. Розробка загальної моделі інформаційного джерела туристичної документації

На основі проведеного аналізу підходів до моделювання ВВР побудуємо загальну модель інформаційного джерела ТД. Нехай *Site* – це ВВР, поданий у вигляді веб-сайту. Будемо вважати, що цей сайт подається так:

$$Site = \langle MandatoryProperties(Site), Page, PageRels \rangle, \quad (2.56)$$

де  $MandatoryProperties(Site)$  – обов'язкові атрибути сайту (адреса, назва, гасло тощо),  $Page = \{Page_i\}_{i=1}^{N^{(Page)}}$  – множина сторінок сайту,  $N^{(Page)}$  – кількість сторінок сайту,  $PageRels \subseteq \{\langle Page_i, Page_j \rangle\}_{i,j=1}^{N^{(Page)}} \equiv Page \times Page$  – множина зв'язків (посилань) між сторінками. При цьому  $Page_1$  – головна сторінка сайту.

В загальному випадку (2.56) відображає орієнтований граф, вершинами якого є веб-сторінки, а дуги подаються посиланнями. Таке подання є важливе для побудови навігації між сторінками на веб-сайті, однак воно є надлишковим для задачі опрацювання ІН веб-сайту. З погляду формування ТД важливою є одностороння-зв'язність графа веб-сайту, у який входять усі вершини  $Page$ , але не обов'язково входять усі дуги  $PageRels$ .

Під інформаційним джерелом ТД, яке подане на ВВР, будемо розуміти редукцію веб-сайту (2.56) у вигляді дерева [155, 156, 157]:

$$Source = Reduction(Site) = \langle MandatoryProperties(Site), Page, RPageRels \rangle, \quad (2.57)$$

де  $RPageRels \subseteq PageRels$  – множина посилань між сторінками, яка забезпечує односторонню зв'язність від головної сторінки  $Page_1$ , і достатня для того, щоб цей граф не містив циклів.

Кожна сторінка – це сукупність певних атрибутів, інформаційних блоків та співвідношень між цими блоками:

$$Page_i = \langle MandatoryProperties(Page_i), PageBlock_i, PageBlockRels_i \rangle, \quad (2.58)$$

де  $MandatoryProperties(Page_i)$  – обов'язкові атрибути  $i$ -ї сторінки (адреса, заголовок, мова тощо),  $PageBlock_i = \{PageBlock_{ij}\}_{j=1}^{N^{(PageBlock_i)}}$  – множина інформаційних блоків  $i$ -ї сторінки,  $N^{(PageBlock_i)}$  – кількість інформаційних блоків  $i$ -ї сторінки,

$$PageBlockRels_i \subseteq \left\{ \langle PageBlock_{ij}, PageBlock_{ik} \rangle \right\}_{j,k=1}^{N^{(PageBlock_i)}} \equiv PageBlock_i \times PageBlock_i -$$

множина співвідношень між інформаційними блоками  $i$ -ї сторінки, яка відображає вкладеність (ієрархію) блоків.

Інформаційний блок визначається множиною інформаційних дописів і описується так:

$$PageBlock_{ij} = \langle MandatoryProperties(PageBlock_{ij}), Post_{ij}, PostRels_{ij} \rangle, \quad (2.59)$$

де  $MandatoryProperties(PageBlock_{ij})$  – обов'язкові атрибути  $j$ -го інформаційного блоку (зокрема назва),  $Post_{ij} = \{Post_{ijk}\}_{k=1}^{N^{(Post_{ij})}}$  – множина дописів  $j$ -го інформаційного блоку,  $N^{(Post_{ij})}$  – кількість дописів  $j$ -того інформаційного блоку,  $PostRels_i \subseteq \left\{ \langle Post_{ijk}, Post_{ijl} \rangle \right\}_{k,l=1}^{N^{(Post_{ij})}} \equiv Post_{ij} \times Post_{ij}$  – множина співвідношень між інформаційними дописами  $j$ -го інформаційного блоку, яка відображає вкладеність (ієрархію) дописів.

Допис – це інформаційна одиниця, яку генерує певний автор на Веб-сайті:

$$Post_{ijk} = \langle MandatoryProperties(Post_{ijk}), Author(Post_{ijk}), Text_{ijk} \rangle, \quad (2.60)$$

де  $MandatoryProperties(Post_{ijk})$  – обов'язкові атрибути  $k$ -того інформаційного допису (заголовок, час тощо),  $Author(Post_{ijk})$  – автор  $k$ -того допису,  $Text_{ijk}$  – текст допису.

Відзначимо, що текст допису може містити розмітку (наприклад, HTML-

розмітку), яка позначає (виділяє, структурує) окремі його складові.

#### 2.4.2. Розробка моделі атомарної ситуації в інформаційному джерелі туристичної документації

Незалежно від наявності розмітки у тексті в дописах будемо вирізняти певні частини (атомарні ситуації), які відображають відомості про окремі факти або події:

$$Part(Post_{ijk}) = \{PostPart_{ijkl}\}_{l=1}^{N^{(Part(Post_{ijk}))}}, \quad (2.61)$$

де  $PostPart_{ijkl}$  –  $l$ -та частина  $k$ -того допису,  $N^{(Part(Post_{ijk}))}$  – кількість частин  $k$ -того допису.

Слід зауважити, що частинами (атомарними ситуаціями) можуть бути окремі абзаци, групи речень або окремі речення, які містяться в тексті допису.

Для визначення механізмів виявлення атомарних ситуацій у дописах введемо поняття маркера атомарної ситуації – фраза, слово або частина слова, яка вказує (маркує) роль речення чи групи речень.

Введемо такі позначення залежно від типів маркерів:

- $ObjectMarker = \{ObjectMarker_i\}_{i=1}^{N^{(ObjectMarker)}}$  – множина маркерів ТО (такі фрази як ‘Львів’, ‘Карпати’, ‘Ратуша’ тощо),  $N^{(ObjectMarker)}$  – кількість маркерів ТО;
- $TermMarker = \{TermMarker_i\}_{i=1}^{N^{(TermMarker)}}$  – множина маркерів термінів класифікаторів ТД (такі фрази як ‘харчування’, ‘проживання’ тощо),  $N^{(TermMarker)}$  – кількість маркерів класифікаторів ТД;
- $ActionMarker = \{ActionMarker_i\}_{i=1}^{N^{(ActionMarker)}}$  – множина маркерів дій над ТО (такі фрази як ‘відвідування’, ‘огляд’, ‘проходження’ тощо),  $N^{(ActionMarker)}$  – кількість маркерів дій над ТО;
- $ExperienceMarker = \{ExperienceMarker_i\}_{i=1}^{N^{(ExperienceMarker)}}$  – множина маркерів відгуків споживачів (такі фрази як ‘я був’, ‘бачив’, ‘проживав’ тощо),  $N^{(ExperienceMarker)}$  – кількість маркерів відгуків споживачів;

- $TimeMarker = \{TimeMarker_i\}_{i=1}^{N(TimeMarker)}$  – множина маркерів часових складових у дописах (час публікації, точні дати, такі фрази як ‘вчора’, ‘місяць тому’ тощо),  $N(ObjectMarker)$  – кількість маркерів ТО;

Також будемо вважати, що загальна множина маркерів ТД – це об’єднання:

$$Marker = ObjectMarker \cup TermMarker \cup ActionMarker \cup ExperienceMarker \cup TimeMarker \quad (2.62)$$

Кожен маркер подається у вигляді фіксованої фрази або символічного шаблону, наприклад з використанням таких спеціальних символів маски:

- ‘\*’ (зірочка) – довільна кількість будь-яких символів;
- ‘?’ (знак питання) – точно один будь-який символ.

Приклад 2.2. Якщо застосовуються символи маски ‘\*’ і ‘?’, то прикладами маркерів ТД можуть бути: ‘\*Льв?в\*’, ‘\*опер\*театр\*’, ‘<div class="price"\*</div>’ та ін.

Тоді частина допису (атомарна ситуація) – це такий кортеж:

$$PostPart_{ijkl} = \left\langle \begin{array}{l} Part(Text_{ijk})_l, ObjectMarker(PostPart_{ijkl}), \\ TermMarker(PostPart_{ijkl}), ActionMarker(PostPart_{ijkl}), \\ ExperienceMarker(PostPart_{ijkl}), TimeMarker(PostPart_{ijkl}) \end{array} \right\rangle, \quad (2.63)$$

де  $Part(Text_{ijk})_l$  – частина тексту допису  $Text_{ijk}$  у вигляді речення чи групи речень,  $ObjectMarker(PostPart_{ijkl}) \in ObjectMarker$  – маркер ТО для заданої атомарної ситуації,  $TermMarker(PostPart_{ijkl}) \in TermMarker$  – маркер терміну класифікатора ТД,  $ActionMarker(PostPart_{ijkl}) \in ActionMarker$  – маркер дії над ТО,  $ExperienceMarker(PostPart_{ijkl}) \in ExperienceMarker$  – маркер відгуку споживача,  $TimeMarker(PostPart_{ijkl}) \in TimeMarker$  – маркер часової складової.

Будемо вважати, що маркер ТО для заданої атомарної ситуації завжди



має бути визначений, тоді як інші маркери можуть невизначені, тобто можуть приймати невизначене значення NULL.

Очевидно, що декілька маркерів можуть стосуватися одного ТО, терміну класифікатора ТД чи відгуку споживача, одної дії над ТО чи часової складової. Таким чином в дописі допускається наявність декількох атомарних ситуацій, які містять однакову частину тексту допису  $Text_{ijk}$ , але відрізняються хоча б одним маркером. Такі атомарні ситуації будемо називати **семантично тотожними**.

Пов'язування частини допису з певними подіями чи фактами через маркери дає можливість здійснити їхню прив'язку до елементів ТД.

## **2.5. Розробка методу лінгвістичного співставлення атомарної ситуації з елементами туристичної документації**

Кожна атомарна ситуація повинна бути проаналізована на предмет відповідності до одного з елементів ТД, а саме:

- досвіду споживача чи описової частини ТО і ДТО;
- дії над об'єктом чи об'єкта;
- події чи факту;
- класифікатора ТД.

При цьому слід врахувати такі фактори як рейтинг ВВР (включно з його сторінками) і рівень довіри до автора допису, тобто його авторитет з погляду формування ТД.

Нехай  $Rating(Source) \in [0,1]$  – рейтинг ВВР (2.57). Один з підходів до визначення рейтингу ВВР описано в [75]. Також для його визначення можна застосовувати рейтинги пошукових машин, індекси зворотних посилань тощо.

Будемо вважати, що рейтинг головної сторінки визначається як рейтинг ВВР, тобто:

$$Rating(Page_1) = Rating(Source). \quad (2.64)$$

Тоді рейтинг будь-якої сторінки визначається як деяка функція від рейтингу головної сторінки:

$$Rating(Page_i) = f(Rating(Page_1)). \quad (2.65)$$

Приклад 2.3. Функція  $f$  в (2.65) може відображати залежність рейтингу сторінки від кількості переходів від головної сторінки:

$$Rating(Page_i) = k^{h_i} Rating(Page_1),$$

де  $k \in [0,1]$  – коефіцієнт зменшення рейтингу,  $h_i$  - відстань (тобто кількість переходів) від  $i$ -тої сторінки до головної в дереві (2.57).

Позначимо  $Trust(Author(Post_{ijk})) \in [0,1]$  – рівень довіри до автора допису (2.60). Рівень довіри може визначатися комплексно або одним з методів, описаних [75]. Також визначення рівень довіри до автора можна враховувати мотивацію його дій [118], його інтереси [104], соціально-демографічні характеристики [124], рівень персоніфікованості [8] тощо. Тобто рівень довіри залежить від багатьох чинників і його визначення може відрізнятися залежно від потреб формування ТД.

Побудуємо методи встановлення відповідності атомарних ситуацій до елементів ТД шляхом пов'язування цих елементів з нечіткими множинами лінгвістичних маркерів. За основу цього підходу покладено метод визначення рекреаційної атриктивності на основі нечіткої логіки [160].

### **2.5.1. Розробка методу лінгвістичного співставлення атомарної ситуації з об'єктами туристичної документації**

Для співставлення атомарної ситуації допису (2.63) з об'єктами ТД необхідно також встановити до якого розділу ТД ця атомарна ситуація максимально відповідає. Тут можливі такі варіанти відповідності:

- факт ТО в описовій частині ТД;

- подія ТО в описовій частині ТД;
- факт ТО у відгуках споживачів ТД;
- подія ТО у відгуках споживачів ТД.

Для того, щоб виявити варіант, найбільш притаманний до атомарної ситуації допису, скористаємося поняттям індикативної ознаки.

**Індикативна ознака** – сукупність маркерів, яким притаманні певні спільні лінгвістичні особливості [124].

Для ТД індикативні ознаки можуть, зокрема, описувати об'єкт, відгук, часову складову.

Для кожного об'єкта ТД  $Object_i$  з (2.13) визначимо індикативну ознаку так:

$$ObjectIndicator(Object_i) = \left\{ \left\langle Marker(Object_i)_j, \mu(Object_i)_j \right\rangle \right\}_{j=1}^{N(ObjectIndicator(Object_i))}, \quad (2.66)$$

де  $Marker(Object_i)_j \in Marker$  –  $j$ -й маркер для об'єкта,  $\mu(Object_i)_j \in [0,1]$  – міра відповідності  $j$ -го маркера до  $i$ -го об'єкта,  $N^{(ObjectIndicator(Object_i))}$  – кількість маркерів, що відповідають об'єкту.

Для визначення, чи відображає атомарна ситуація відгук споживача, введемо таку індикативну ознаку:

$$ExperienceIndicator = \left\{ \left\langle Marker(ExperienceIndicator)_i, \mu(ExperienceIndicator)_i \right\rangle \right\}_{i=1}^{N(ExperienceIndicator)}, \quad (2.67)$$

де  $Marker(ExperienceIndicator)_i \in Marker$  –  $i$ -й маркер для відображення відгуку споживача,  $\mu(ExperienceIndicator)_i \in [0,1]$  – міра відповідності  $i$ -го маркера до відгуку споживача,  $N^{(ExperienceIndicator)}$  – кількість маркерів, що відображають відгук споживача.

Аналогічно визначимо індикативну ознаку часової складової в атомарній ситуації:

$$TimeIndicator = \left\{ \left\langle Marker(TimeIndicator)_i, \mu(TimeIndicator)_i \right\rangle \right\}_{i=1}^{N^{(TimeIndicator)}} \quad (2.68)$$

де  $Marker(TimeIndicator)_i \in Marker$  -  $i$ -й маркер для відображення часової складової,  $\mu(TimeIndicator)_i \in [0,1]$  – міра відповідності  $i$ -го маркера до часової складової,  $N^{(ExperienceAdequacy)}$  – кількість маркерів, що відображають часову складову

Якщо  $Object_p \in Object$  –  $p$ -й об'єкт ТД, то з урахуванням індикативної ознаки (2.66) відповідність маркера ТО атомарної ситуації (2.63) до об'єкта  $Object_p$  – це така величина

$$ObjectAdequacy(Object_p, ObjectMarker(PostPart_{ijkl})) = Rating(Page_i) * Trust(Author(Post_{ijk})) * \mu(Object_p, ObjectMarker(PostPart_{ijkl})), \quad (2.69)$$

де для деякого  $q$  виконується:

$$\mu(Object_p, ObjectMarker(PostPart_{ijkl})) = \begin{cases} \mu(Object_p)_q, & \text{якщо } ObjectMarker(PostPart_{ijkl}) = Marker(Object_p)_q; \\ 0, & \text{якщо } ObjectMarker(PostPart_{ijkl}) \neq Marker(Object_p)_q. \end{cases}$$

На відміну від маркерів ТО маркери відгуку споживача показують, чи можна вважати атомарну ситуацію дійсно думкою споживача про ТО. Тому відповідність атомарної ситуації (2.63) до відгуку споживача з урахуванням індикативної ознаки (2.67) – це така величина:

$$ExperienceAdequacy(ExperienceMarker(PostPart_{ijkl})) = Rating(Page_i) * Trust(Author(Post_{ijk})) * \mu(ExperienceMarker(PostPart_{ijkl})), \quad (2.70)$$

де для деякого  $p$  виконується:

$$\mu(ExperienceMarker(PostPart_{ijkl})) = \begin{cases} \mu(ExperienceIndicator)_p, & \text{якщо } ExperienceMarker(PostPart_{ijkl}) = Marker(ExperienceIndicator)_p; \\ 0, & \text{якщо } ExperienceMarker(PostPart_{ijkl}) \neq Marker(ExperienceIndicator)_p. \end{cases}$$

Аналогічно визначимо відповідність атомарної ситуації (2.63) до часової складової з урахуванням індикативної ознаки (2.68):

$$\begin{aligned} & \text{TimeAdecuacy}(\text{TimeMarker}(\text{PostPart}_{ijkl})) = \\ & \text{Rating}(\text{Page}_i) * \text{Trust}(\text{Author}(\text{Post}_{ijk})) * \mu(\text{TimeMarker}(\text{PostPart}_{ijkl})), \end{aligned} \quad (2.71)$$

де для деякого  $p$  виконується:

$$\begin{aligned} & \mu(\text{TimeMarker}(\text{PostPart}_{ijkl})) = \\ & \begin{cases} \mu(\text{TimeIndicator})_p, \text{ якщо } \text{TimeMarker}(\text{PostPart}_{ijkl}) = \text{Marker}(\text{TimeIndicator})_p; \\ 0, \text{ якщо } \text{TimeMarker}(\text{PostPart}_{ijkl}) \neq \text{Marker}(\text{TimeIndicator})_p. \end{cases} \end{aligned}$$

Введемо мінімально допустимі значення для відповідностей (2.69), (2.70), (2.71):  $\alpha^{(\text{ObjectAdecuacy})} \in (0,1]$ ,  $\alpha^{(\text{ExperienceAdecuacy})} \in (0,1]$ ,  $\alpha^{(\text{TimeAdecuacy})} \in (0,1]$ . Тоді, якщо задовольняється критерій:

$$\text{ObjectAdecuacy}(\text{Object}_p, \text{ObjectMarker}(\text{PostPart}_{ijkl})) \geq \alpha^{(\text{ObjectAdecuacy})}, \quad (2.72)$$

то вважається, що атомарна ситуація (2.63) відноситься до ТО  $\text{Object}_p$ . В протилежному випадку атомарна ситуація (2.63) не відноситься до ТО  $\text{Object}_p$ .

Якщо задовольняється критерій:

$$\text{ExperienceAdecuacy}(\text{ExperienceMarker}(\text{PostPart}_{ijkl})) \geq \alpha^{(\text{ExperienceAdecuacy})}, \quad (2.73)$$

то вважається, що атомарна ситуація (2.63) відноситься до частини відгуків споживачів ТД. В протилежному випадку атомарна ситуація (2.63) відноситься до описової частини ТО і ДТО.

І нарешті, якщо задовольняється критерій:

$$\text{TimeAdecuacy}(\text{TimeMarker}(\text{PostPart}_{ijkl})) \geq \alpha^{(\text{TimeAdecuacy})}, \quad (2.74)$$

то вважається, що атомарна ситуація (2.63) відноситься до події, оскільки містить часову складову. В протилежному випадку атомарна ситуація (2.63) відноситься до факту, бо не містить часової складової.

### 2.5.2. Розробка методу лінгвістичного співставлення атомарної ситуації з діями над об'єктами туристичної документації

Цей метод будується аналогічно до попереднього. Визначимо індикативну ознаку дії над ТО (2.4) так:

$$ActionIndicator(Action(Object_i)_j) = \left\{ \left\langle \left\langle \begin{array}{l} /Marker(Action(Object_i)_j)_k \\ \mu(Action(Object_i)_j)_k \end{array} \right\rangle \right\rangle_{k=1}^{N^{(ActionIndicator(Action(Object_i)_j))}} \right\}, \quad (2.75)$$

де  $Marker(Action(Object_i)_j)_k \in Marker$  –  $k$ -й маркер для ДТО,  $\mu(Action(Object_i)_j)_k \in [0,1]$  – міра відповідності  $k$ -го маркера до цієї ДТО,  $N^{(ActionIndicator(Action(Object_i)_j))}$  – кількість маркерів, що відповідають ДТО.

Якщо  $Action(Object_p)_q \in Action(Object_p)$  –  $p$ -а дія над  $q$ -м об'єктом ТД, то з урахуванням індикативної ознаки (2.75) відповідність маркера ДТО атомарної ситуації (2.63) до дії  $Action(Object_p)_q$  – це така величина:

$$ActionAdequacy(Action(Object_p)_q, ActionMarker(PostPart_{ijkl})) = Rating(Page_i) * Trust(Author(Post_{ijk})) * \mu(Action(Object_p)_q, ActionMarker(PostPart_{ijkl})), \quad (2.76)$$

де для деякого  $r$  виконується:

$$\mu(Action(Object_p)_q, ActionMarker(PostPart_{ijkl})) = \begin{cases} \mu(Action(Object_p)_q)_r, & \text{якщо } ActionMarker(PostPart_{ijkl}) = Marker(Action(Object_p)_q)_r; \\ 0, & \text{якщо } ActionMarker(PostPart_{ijkl}) \neq Marker(Action(Object_p)_q)_r. \end{cases}$$

Введемо мінімально допустиме значення для відповідності (2.76):  $\alpha^{(ActionAdequacy)} \in (0,1]$ . Тоді, якщо задовольняється критерій:

$$ActionAdequacy(Action(Object_p)_q, ActionMarker(PostPart_{ijkl})) \geq \alpha^{(ActionAdequacy)}, \quad (2.77)$$

то вважається, що атомарна ситуація (2.63) відноситься до ДТО  $Action(Object_p)_q$ . В протилежному випадку атомарна ситуація (2.63) не відноситься до  $Action(Object_p)_q$ .

Для визначення відношення атомарної ситуації (2.63) до досвіду споживача, події чи факту ТД застосовуються критерії (2.73), (2.74).

### 2.5.3. Розробка методу лінгвістичного співставлення атомарної ситуації з класифікаторам туристичної документації

Для співставлення атомарної ситуації допису (2.63) з класифікаторами ТД необхідно встановити ті терміни класифікаторів, які максимально відповідають цій атомарній ситуації.

Визначимо індикативну ознаку для терміну класифікаторів ТД з (2.32) так:

$$TermIndicator(Term_{ij}) = \left\{ \left\langle Marker(Term_{ij})_k, \mu(Term_{ij})_k \right\rangle \right\}_{k=1}^{N^{(Indication(Term_{ij}))}}, \quad (2.78)$$

де  $Marker(Term_{ij})_k \in Marker$  –  $k$ -й маркер для терміну,  $\mu(Term_{ij})_k \in [0,1]$  – міра відповідності  $k$ -того маркера до терміну,  $N^{(Indication(Term_{ij}))}$  – кількість маркерів, що відповідають терміну класифікатора.

Нехай  $Term_{pq} \in Term_p$  –  $q$ -й термін  $p$ -го класифікатора. Тоді з урахуванням індикативної ознаки (2.78) відповідність маркера терміну класифікатора ТД атомарної ситуації (2.63) до терміну  $Term_{pq}$   $p$ -го класифікатора – це така величина:

$$TermAdequacy(Term_{pq}, TermMarker(PostPart_{ijkl})) = Rating(Page_i) * Trust(Author(Post_{ijk})) * \mu(Term_{pq}, TermMarker(PostPart_{ijkl})), \quad (2.79)$$

де для деякого  $r$  виконується:

$$\mu(\text{Term}_{pq}, \text{TermMarker}(\text{PostPart}_{ijkl})) = \begin{cases} \mu(\text{Term}_{pq})_r, & \text{якщо} \\ \text{TermMarker}(\text{PostPart}_{ijkl}) = \text{Marker}(\text{Term}_{pq})_r; \\ 0, & \text{якщо} \\ \text{TermMarker}(\text{PostPart}_{ijkl}) \neq \text{Marker}(\text{Term}_{pq})_r. \end{cases}$$

Якщо відповідність (2.66) є не меншою ніж задане значення, тобто вона задовольняє критерій:

$$\text{TermAdequacy}(\text{Term}_{pq}, \text{TermMarker}(\text{PostPart}_{ijkl})) \geq \alpha^{(\text{TermAdequacy})}, \quad (2.80)$$

де  $\alpha^{(\text{TermAdequacy})} \in (0, 1]$  – мінімально допустиме значення відповідності маркера до терміну класифікатора ТД, то вважається, що атомарна ситуація (2.63) класифікується терміном  $\text{Term}_{pq}$ .

#### 2.5.4. Розробка методу лінгвістичного співставлення атомарної ситуації з атрибутами фактів і подій туристичної документації

Співставлення атомарної ситуації з фактами визначення ТО (2.33), (2.35) і ДТО (2.34), (2.36) не відбувається, оскільки ці факти визначаються експертом, який відповідає за формування ТД). Тому метод співставлення атомарної ситуації з атрибутами стосується фактів і подій, описаних відповідними інформаційними моделями (2.37), (2.38), (2.39), (2.40) та (2.41), (2.42), (2.43), (2.44).

Пов'яжемо атрибути фактів і подій з лінгвістичними маркерами. Нехай у фактах і подія ТД застосовується множина атрибутів:

$$\text{Attribute} = \{ \text{Attribute}_i \}_{i=1}^{N^{(\text{Attribute})}}, \quad (2.81)$$

де  $\text{Attribute}_i$  –  $i$ -й атрибут,  $N^{(\text{Attribute})}$  – загальна кількість атрибутів.

Для кожного атрибута  $\text{Attribute}_i$  з (2.81) визначимо індикативну ознаку так:



$$AttributeIndicator(Attribute_i) = \left\{ \left\langle PageLevel(Attribute_i)_j, Marker(Attribute_i)_j, \mu(Attribute_i)_j \right\rangle \right\}_{j=1}^{N^{(AttributeIndicator(Attribute_i))}}, \quad (2.82)$$

де  $PageLevel(Attribute_i)_j$  – рівень сторінки (сторінка, інформаційний блок, допис, атомарна ситуація), на якому необхідно виявляти маркер,  $Marker(Attribute_i)_j \in Marker$  –  $j$ -й маркер для атрибута,  $\mu(Attribute_i)_j \in [0,1]$  – міра відповідності  $j$ -го маркера до  $i$ -го атрибута,  $N^{(AttributeIndicator(Attribute_i))}$  – кількість маркерів атрибута.

Зауважимо, що якщо значення атрибута маркується тегом, то доцільно встановити  $\mu(Attribute_i)_j = 1$ .

Якщо  $Attribute_p \in Attribute$  –  $p$ -тий атрибут, то з урахуванням індикативної ознаки (2.82) відповідність маркера до цього атрибута визначається так:

$$AttributeAdequacy(Attribute_p, AttributeMarker(PostPart_{ijkl})) = Rating(Page_i) * Trust(Author(Post_{ijk})) * \mu(Attribute_p, AttributeMarker(PostPart_{ijkl})), \quad (2.83)$$

де  $AttributeMarker(PostPart_{ijkl})$  – маркер атрибут, виявлений в атомарній ситуації або у відповідному дописі, або у відповідному інформаційному блоці, або на відповідній сторінці, і для деякого  $q$  виконується:

$$\mu(Attribute_p, AttributeMarker(PostPart_{ijkl})) = \begin{cases} \mu(Attribute_p)_q, & \text{якщо } AttributeMarker(PostPart_{ijkl}) = Marker(Attribute_p)_q \\ i PageLevel(AttributeMarker(PostPart_{ijkl})) = PageLevel(Attribute_p)_q; \\ 0, & \text{якщо } AttributeMarker(PostPart_{ijkl}) \neq Marker(Attribute_p)_q \\ \text{або } PageLevel(AttributeMarker(PostPart_{ijkl})) \neq PageLevel(Attribute_p)_q. \end{cases}$$

де  $PageLevel(AttributeMarker(PostPart_{ijkl}))$  – рівень сторінки (сторінка,

інформаційний блок, допис, атомарна ситуація), на якому виявлено маркер  $AttributeMarker(PostPart_{ijkl})$ .

Якщо відповідність (2.69) задовольняє критерій:

$$AttributeAdequacy(Attribute_p, AttributeMarker(PostPart_{ijkl})) \geq \alpha^{(AttributeAdequacy)}, \quad (2.84)$$

де  $\alpha^{(AttributeAdequacy)} \in (0,1]$  – мінімально допустиме значення міри відповідності маркера до атрибута, то вважається, що виявлений маркер  $AttributeMarker(PostPart_{ijkl})$  відповідає атрибуту  $Attribute_p$ .

### 2.5.5. Розробка методу лінгвістичного співставлення сукупності атомарних ситуацій з елементами туристичної документації

Критерії (2.72), (2.73), (2.74), (2.77), (2.80), (2.84) показують відповідність атомарної ситуації до одного з елементів ТД за виявленим маркером. При цьому не враховується такий варіант, коли з частиною тексту допису, що міститься в атомарній ситуації, пов'язуються декілька маркерів, які пов'язуються з одним елементом ТД.

Для усунення зазначеного недоліку побудуємо критерій відповідності сукупності семантично тотожних атомарних ситуацій до одного з елементів ТД.

Нехай маємо дві семантично тотожні атомарні ситуації одного допису  $PostPart_{ijkl_1}$  і  $PostPart_{ijkl_2}$ , тобто їхні текстові частини в (2.63) рівні:  $Part(Text_{ijk})_{l_1} = Part(Text_{ijk})_{l_2}$ . Якщо виконується:

$$\mu(Object_p, ObjectMarker(PostPart_{ijkl_1})) \geq \mu(Object_p, ObjectMarker(PostPart_{ijkl_2})),$$

то сукупна міра відповідності цих двох атомарних ситуацій до об'єкта  $Object_p \in Object$  визначається так:

$$\begin{aligned} & \mu(Object_p, ObjectMarker(PostPart_{ijkl_1}), ObjectMarker(PostPart_{ijkl_2})) = \\ & \mu(Object_p, ObjectMarker(PostPart_{ijkl_1})) + \\ & (1 - \mu(Object_p, ObjectMarker(PostPart_{ijkl_1}))) \mu(Object_p, ObjectMarker(PostPart_{ijkl_2})) \end{aligned}$$

Тоді відповідність маркерів ТО сукупності семантично тотожних атомарних ситуацій до об'єкта  $Object_p$  – це така величина:

$$ObjectAdequacy \left( \begin{array}{l} Object_p, ObjectMarker(PostPart_{ijkl_1}), \\ ObjectMarker(PostPart_{ijkl_2}) \end{array} \right) = \\ Rating(Page_i) * Trust(Author(Post_{ijk})) * \\ \mu(Object_p, ObjectMarker(PostPart_{ijkl_1}), ObjectMarker(PostPart_{ijkl_2})) \quad . \quad (2.85)$$

Аналогічно будуються інші відповідності маркерів сукупності семантично тотожних атомарних ситуацій до елементів ТД:

- відповідність до відгуків споживачів:

$$ExperienceAdequacy \left( \begin{array}{l} ExperienceMarker(PostPart_{ijkl_1}), \\ ExperienceMarker(PostPart_{ijkl_2}) \end{array} \right) = \\ Rating(Page_i) * Trust(Author(Post_{ijk})) * \\ \mu(ExperienceMarker(PostPart_{ijkl_1}), ExperienceMarker(PostPart_{ijkl_2})) \quad , \quad (2.86)$$

де

$$\mu(ExperienceMarker(PostPart_{ijkl_1}), ExperienceMarker(PostPart_{ijkl_2})) = \\ \mu(ExperienceMarker(PostPart_{ijkl_1})) + \\ (1 - \mu(ExperienceMarker(PostPart_{ijkl_1}))) \mu(ExperienceMarker(PostPart_{ijkl_2})) \quad .$$

- відповідність до часових складових:

$$TimeAdequacy(TimeMarker(PostPart_{ijkl_1}), TimeMarker(PostPart_{ijkl_2})) = \\ Rating(Page_i) * Trust(Author(Post_{ijk})) * \\ \mu(TimeMarker(PostPart_{ijkl_1}), TimeMarker(PostPart_{ijkl_2})) \quad , \quad (2.87)$$

де

$$\mu(TimeMarker(PostPart_{ijkl_1}), TimeMarker(PostPart_{ijkl_2})) = \\ \mu(TimeMarker(PostPart_{ijkl_1})) + \\ (1 - \mu(TimeMarker(PostPart_{ijkl_1}))) \mu(TimeMarker(PostPart_{ijkl_2})) \quad .$$

- відповідність до дій над ТО:

$$\begin{aligned} & \text{ActionAdequacy} \left( \begin{array}{l} \text{Action}(\text{Object}_p)_q, \text{ActionMarker}(\text{PostPart}_{ijk_1}) \\ \text{ActionMarker}(\text{PostPart}_{ijk_2}) \end{array} \right) = \\ & \text{Rating}(\text{Page}_i) * \text{Trust}(\text{Author}(\text{Post}_{ijk})) * \\ & \mu(\text{Action}(\text{Object}_p)_q, \text{ActionMarker}(\text{PostPart}_{ijk_1}), \text{ActionMarker}(\text{PostPart}_{ijk_2})) \end{aligned} \quad , (2.88)$$

де

$$\begin{aligned} & \mu(\text{Action}(\text{Object}_p)_q, \text{ActionMarker}(\text{PostPart}_{ijk_1}), \text{ActionMarker}(\text{PostPart}_{ijk_2})) = \\ & \mu(\text{ActionMarker}(\text{PostPart}_{ijk_1})) + \\ & (1 - \mu(\text{ActionMarker}(\text{PostPart}_{ijk_1}))) \mu(\text{ActionMarker}(\text{PostPart}_{ijk_2})) \end{aligned} \quad .$$

- відповідність до термінів класифікаторів:

$$\begin{aligned} & \text{TermAdequacy}(\text{Term}_{pq}, \text{TermMarker}(\text{PostPart}_{ijk_1}), \text{TermMarker}(\text{PostPart}_{ijk_2})) = \\ & \text{Rating}(\text{Page}_i) * \text{Trust}(\text{Author}(\text{Post}_{ijk})) * \\ & \mu(\text{Term}_{pq}, \text{TermMarker}(\text{PostPart}_{ijk_1}), \text{TermMarker}(\text{PostPart}_{ijk_2})) \end{aligned} \quad , (2.89)$$

де

$$\begin{aligned} & \mu(\text{Term}_{pq}, \text{TermMarker}(\text{PostPart}_{ijk_1}), \text{TermMarker}(\text{PostPart}_{ijk_2})) = \\ & \mu(\text{TermMarker}(\text{PostPart}_{ijk_1})) + \\ & (1 - \mu(\text{TermMarker}(\text{PostPart}_{ijk_1}))) \mu(\text{TermMarker}(\text{PostPart}_{ijk_2})) \end{aligned} \quad .$$

- відповідність до атрибутів:

$$\begin{aligned} & \text{AttributeAdequacy} \left( \begin{array}{l} \text{Attribute}_p, \text{AttributeMarker}(\text{PostPart}_{ijk_1}) \\ \text{AttributeMarker}(\text{PostPart}_{ijk_2}) \end{array} \right) = \\ & \text{Rating}(\text{Page}_i) * \text{Trust}(\text{Author}(\text{Post}_{ijk})) * \\ & \mu(\text{Attribute}_p, \text{AttributeMarker}(\text{PostPart}_{ijk_1}), \text{AttributeMarker}(\text{PostPart}_{ijk_2})) \end{aligned} \quad , (2.90)$$

де

$$\mu(\text{Attribute}_p, \text{AttributeMarker}(\text{PostPart}_{ijkl_1}), \text{AttributeMarker}(\text{PostPart}_{ijkl_2})) = \mu(\text{AttributeMarker}(\text{PostPart}_{ijkl_1})) + (1 - \mu(\text{AttributeMarker}(\text{PostPart}_{ijkl_1})))\mu(\text{AttributeMarker}(\text{PostPart}_{ijkl_2}))$$

## 2.6. Висновки до розділу

1. Вперше побудовано формальну модель туристичної документації на основі множин туристичних послуг та відгуків споживачів туристичних послуг, які включають туристичні об'єкти і дії над туристичними об'єктами, а також множини класифікаторів, що стало основою для побудови процедур комп'ютерно-лінгвістичного формування туристичної документації через індикативні ознаки і показники якості туристичної документації.
2. Побудовано моделі елементарних одиниць інформації у туристичній документації на основі логічних організацій даних шляхом виокремлення обов'язкових атрибутів, множини термінів класифікаторів, атрибутів джерела, атрибутів автора, атрибутів мови та атрибутів часових властивостей, що дало змогу побудувати процедури лінгвістичного співставлення інформації відкритих веб-ресурсів з елементами туристичної документації, процедур виявлення прогалин і некоректної інформації у туристичній документації.
3. Проведено класифікацію туристичної документації за особливостями загальної моделі, що дозволило розробити показники якості придатності до використання туристичної документації.
4. Удосконалено формальну модель інформаційного джерела для формування туристичної документації шляхом виокремлення блоків і атомарних ситуацій туристичної інформації на веб-сторінках, що стало основою для побудови процедури вибору відкритих веб-ресурсів, процедур лінгвістичного співставлення вмісту відкритих веб-ресурсів з елементами туристичної документації

5. Розроблено метод лінгвістичного співставлення атомарної ситуації зі елементами туристичної документації на основі модифікованих індикативних ознак соціально-демографічних характеристик учасників віртуальних спільнот, що дозволило побудувати відповідні процедури лінгвістичного співставлення вмісту відкритих веб-ресурсів з елементами туристичної документації.

## **Розділ 3. Розробка процедур комп'ютерно-лінгвістичного формування туристичної документації**

Формування туристичної документації – це складний процес, який потребує затрат часу, залучення висококваліфікованих фахівців, впровадження інформаційних технологій обробки інформації, засобів її групування, ідентифікацію джерел туристичної інформації, логічного аналізу матеріалів, вибору методів екстракції даних, зведення даних до єдиного доступного для споживача туристичного продукту формату шляхом консолідації обробленої інформації. Ефективне формування туристичної документації практично неможливе без розроблення відповідних програмних компонент програмно-алгоритмічного комплексу підтримки туристичної діяльності.

У цьому розділі розглядаються процедури комп'ютерно-лінгвістичного формування туристичної документації, які є необхідні для реалізації та впровадження програмних компонент збору, опрацювання та збереження туристичної інформації, отриманої з відкритих Веб-ресурсів.

Основні результати розділу опубліковані автором у працях [74, 96, 134, 143, 144, 152, 155, 156, 157].

### **3.1. Структуризація процесу формування туристичної документації з використанням відкритих Веб-ресурсів**

На основі проведеного аналізу особливостей формування ТД і використання відкритих веб-ресурсів у туристичній діяльності проведемо структуризацію формування ТД.

Процес формування туристичної документації складається з таких етапів як (рис. 3.1):

- аналітичний;

- пошуковий;
- обробка інформації;
- створення документації;
- маніпуляція документацією.

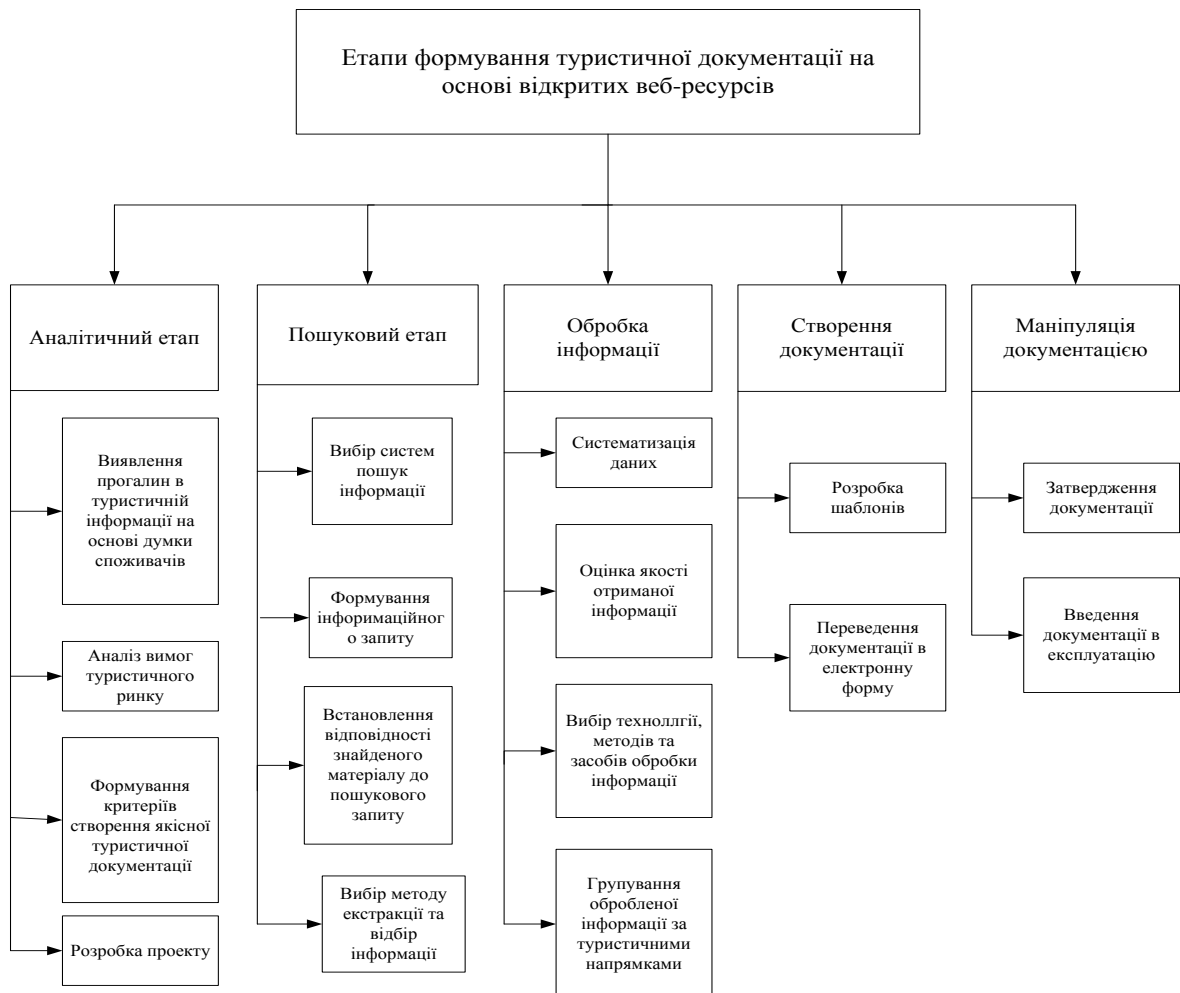


Рис. 3.1. Етапи формування туристичної документації з використанням відкритих веб-ресурсів

На аналітичному етапі необхідно здійснити інформаційно-аналітичний огляд стану проблеми, щоб зрозуміти, якої інформації не вистачає, якого типу документацію потрібно створити і розробити відповідний проект туристичної документації. Для цього необхідно виявити можливі прогалини в туристичній інформації на основі аналізу думок, запитів, скарг та пропозицій споживачів



туристичного продукту, проаналізувати вимоги сучасного туристичного ринку, а також сформувані критерії створення якісної туристичної документації.

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

На наступному етапі відбувається розробка шаблонів документації шляхом наповнення їх обробленою інформацією і переведення її в електронну форму.

Створену документацію затверджують після тестування згідно діючих стандартів у т.ч. відповідності її ISO та верифікації на придатність і вводять в експлуатацію.

### **3.2. Розробка процедури вибору відкритих веб-ресурсів для формування туристичної документації**

Від належного вибору відкритих веб-ресурсів, придатних для формування ТД, залежить як ефективність наповнення ТД, так і якість самої ТД. Тому спочатку необхідно сформулювати критерії придатності ВВР до формування ТД, а потім за результатами пошуку ВВР перевіряти їх на відповідність цим критеріям.

#### **3.2.1. Визначення критеріїв відповідності відкритих веб-ресурсів до формування туристичної документації**

Розроблені інформаційна модель туристичної документації та методи співставлення атомарної ситуації зі складовими ТД дозволяють визначити критерії відповідності ВВР до формування ТД.

1) авторитетність відкритих веб-ресурсів;

Визначення загальної авторитетності ВВР є доволі складною задачею з багатьма факторами [75]. З погляду формування ТД авторитетність ВВР

дозволяє відкинути ті веб-сайти, надають некоректну, неперевірену, неактуальну, недостовірну інформацію.

$$Rating(Source) \geq \alpha^{(Rating(Source))}, \quad (3.1)$$

де  $\alpha^{(Rating(Source))} \in (0,1]$  – мінімально допустиме значення рейтингу ВВР.

2) відповідність до тематики;

Тематика ТД визначається такими складовими:

- ТО – показують географію ведення туристичної діяльності;
- ДТО – відображають зміст послуг в межах туристичної діяльності;
- класифікатори ТД – вказують на окремі теми і ключові слова, з якими пов'язуються розділи ТД.

Відповідність ВВР до тематики ТД можна встановити за допомогою експертних оцінок або через визначення релевантності ВВР.

У випадку експертного оцінювання необхідно зібрати оцінки експертів щодо відповідності ВВР до усіх об'єктів, дій над об'єктами і класифікаторів, знайти агреговане значення експертної оцінки ВВР (наприклад, за методом Дельфі) і перевірити таку умову:

$$ExpertThemeEval(Source) \geq \alpha^{(ExpertThemeEval(Source))}, \quad (3.2)$$

де  $\alpha^{(ExpertThemeEval(Source))} \in (0,1]$  – мінімально допустиме значення експертної оцінки відповідності ВВР до тематики ТД.

**Релевантність** позначає міру відповідності результатів пошуку завданню, поставленому в запиті до пошукової машини [145]. Використання релевантності дозволяє автоматизовано визначити міру відповідності ВВР до тематики ТД.

Модифікуємо метод виявлення релевантних веб-форумів [118]. Для цього побудуємо шаблон розширеного параметризованого запиту з уточненнями для виявлення релевантних ВВР до ТО, ДТО і класифікаторів ТД (рис. 3.2).

Уточнення в цьому шаблоні полягають у заданні параметрів розширеного пошуку через оператор аналізу і відповідний специфікатор. Оператор аналізу – це один з варіантів задання ключових слів, фраз, альтернативних слів, діапазону чисел, адреси сайту тощо (оператори аналізу для пошукової машини Google описано тут [105]). Специфікатор – це послідовність символів (слова, фрази, числа тощо).



Рис. 3.2. Шаблон розширеного параметризованого запиту з уточненнями для виявлення релевантних відкритих веб-ресурсів

Операція виявлення ВВР застосовується у випадку, коли наперед відомо адресу або заголовки веб-сайту ВВР. Операції встановлення релевантності до ТО, ДТО чи класифікатора ТД застосовуються для визначення відповідності ВВР до окремих розділів ТД.

Результатом виконання розширеного параметризованого пошукового запиту є множина веб-сторінок з URL-адресами, які повертає пошукова машина. Пронумеруємо позицію кожної веб-сторінки, що підлягає розгляду, від 1 до  $N^{(RelevanceThemeEval)}$ . веб-сторінки з позиціями більшими за  $N^{(RelevanceThemeEval)}$  не будемо приймати до уваги. Нехай  $N^{(RelevanceThemeEval(Source))}$  – позиція веб-сторінки джерела в результаті виконання розширеного параметризованого пошукового запиту. Тоді критерій відповідності ВВР за релевантністю до тематики, заданою в розширеному параметризованому пошуковому запиті, має вигляд:

$$RelevanceThemeEval(Source) = 1 - \frac{N^{(RelevanceThemeEval(Source))} - 1}{N^{(RelevanceThemeEval)} - 1} \geq 0. \quad (3.3)$$

3) достовірність дописів;

В [8] показано, що достовірність (як властивість інформації бути правильно сприйнятою) прямо залежить від рівня персоніфікованості автора. В загальному будемо вважати, що достовірність допису прямо залежить від рівня довіри до його автора.

Нехай  $Author(Source) = \{Author(Source)_i\}_{i=1}^{N(Author(Source))}$  – множина усіх авторів ВВР, відповідно  $Trust(Author(Source)_i) \in [0,1]$  рівень довіри до  $i$ -го автора, який може визначатися за підходами, описаними в [8, 75, 104, 118, 124]. Тоді сумарний рівень довіри до авторів ВВР – це:

$$Trust(Author(Source)) = \frac{\sum_{i=1}^{N(Author(Source))} Trust(Author(Source)_i)}{N(Author(Source))}. \quad (3.4)$$

Тоді критерій достовірності дописів ВВР формулюється так:

$$Trust(Author(Source)) \geq \alpha^{(Trust(Author(Source)))}, \quad (3.5)$$

де  $\alpha^{(Trust(Author(Source)))} \in (0,1]$  – мінімально допустиме значення сумарного рівня довіри до авторів ВВР.

4) актуальність дописів;

Актуальність дописів визначається часом їхньої публікації допису. Для допису  $Post_{ijk}$  з (2.60) позначимо  $PostTime(Post_{ijk})$  – час його публікації. Нехай  $T_0$  – момент часу початку відліку актуальності (тобто дописи, які опубліковані раніше за  $T_0$  вважаються неактуальними),  $T_{now}$  – поточний момент часу. Тоді актуальність допису – це:

$$Actuality(Post_{ijk}) = \begin{cases} \frac{(T_{now} - T_0) - (T_{now} - PostTime(Post_{ijk}))}{T_{now} - T_0}, & PostTime(Post_{ijk}) \geq T_0 \\ 0, & PostTime(Post_{ijk}) < T_0 \end{cases}. \quad (3.6)$$

Тоді на основі (2.56), (2.58), (2.59) актуальність ВВР визначається через актуальність кожного допису:

$$Actuality(Source) = \frac{\sum_{i=1}^{N^{(Page)}} \sum_{j=1}^{N^{(PageBlock)}} \sum_{k=1}^{N^{(Post_{ij})}} Actuality(Post_{ijk})}{\sum_{i=1}^{N^{(Page)}} \sum_{j=1}^{N^{(PageBlock)}} \sum_{k=1}^{N^{(Post_{ij})}} (1)}. \quad (3.7)$$

Критерій актуальності дописів ВВР формулюється так:

$$Actuality(Source) \geq \alpha^{(Actuality(Source))}, \quad (3.8)$$

де  $\alpha^{(Actuality(Source))} \in (0,1]$  — мінімально допустиме значення сумарного рівня актуальності дописів ВВР.

### 3.2.2. Розробка процедури вибору джерела для формування туристичної документації

Вибір джерела для формування ТД відбувається у три етапи:

- 1) формування множини джерел ВВР;
- 2) перевірка джерела на відповідність критеріям формування ТД;
- 3) вибір веб-сторінок джерела, які є корисними для формування ТД.

Блок-схема алгоритму знаходження джерела ВВР подано на рис. 3.3. Цей алгоритм на основі вхідних ключових слів *Keywords* здійснює запит до пошукової машини *Search(Keywords)*, результатом якого є множина сайтів *Sites*. Послідовно кожен сайт  $Site = \text{NEXT}(Sites)$  з множини *Sites*, який задовольняє критерій релевантості (3.3), перевіряється експертом на предмет відповідності тематики ТД. Якщо виконується критерій (3.2), то цей сайт додається до множини джерел *Sources* ВВР.

Перевірка на відповідність джерел, отриманих на попередньому етапі, критеріям формування ТД складається з таких кроків:

- 1) для заданого джерела *Source* визначити: рейтинг (авторитетність)  $Rating(Source)$ , рівень довіри до авторів (достовірність дописів)  $Trust(Author(Source))$ , рівень актуальності дописів  $Actuality(Source)$ ;
- 2) якщо джерело *Source* задовольняє критеріям (3.1), (3.5), (3.8), то воно приймається для формування ТД;

3) якщо джерело  $Source$  не задовольняє хоча б одному критерію (3.1), (3.5), (3.8), то воно відхиляється для формування ТД.

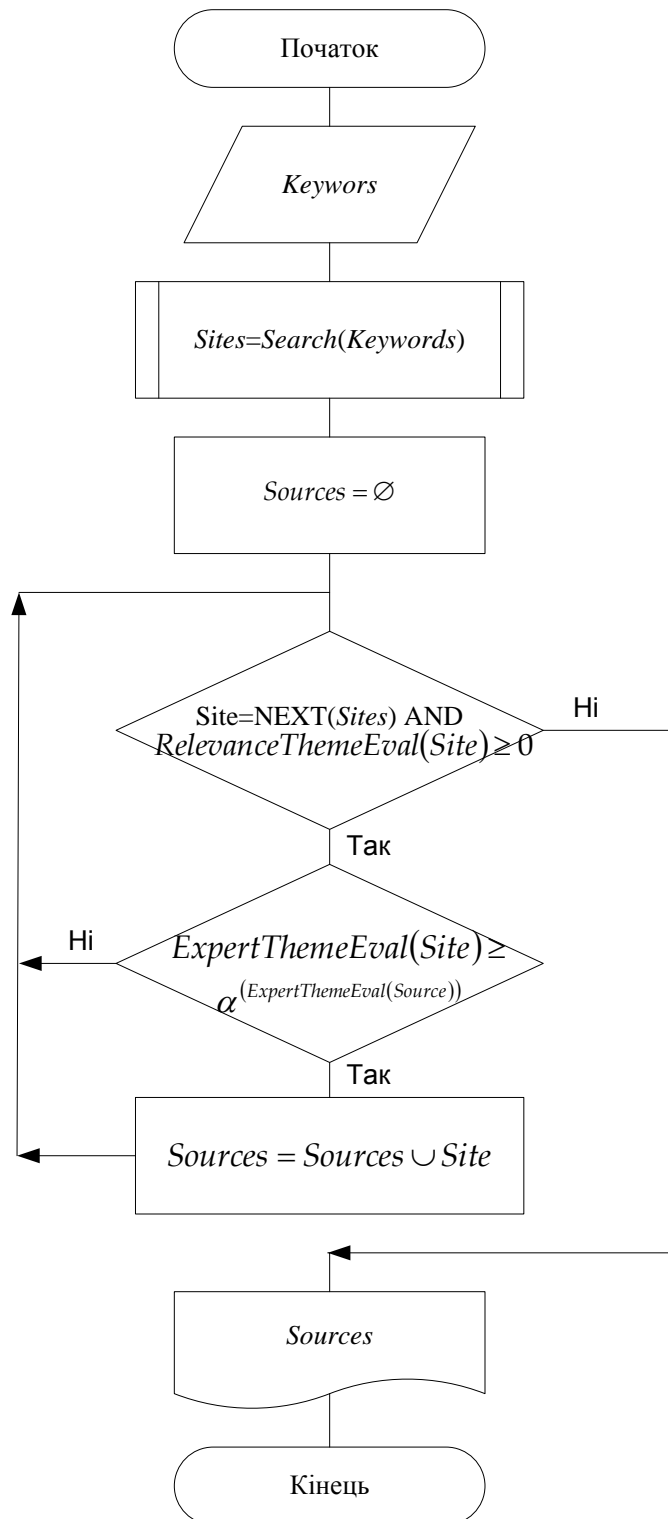


Рис. 3.3. Алгоритм формування множини джерел ВВР

В межах джерела ВВР, яке відповідає критеріям формування ТД на 2 етапі, можуть існувати як веб-сторінки, які містять необхідну інформацію для ТД, так і Веб-сторінки, що не містять такої інформації. Очевидно, на третьому етапі необхідно виділити множину веб-сторінок, що містять корисну для ТД інформацію.

Будемо вважати, що оскільки кожна сторінка подається як множина інформаційних блоків, а кожен блок – як множина допис, у які входять атомарні ситуації, то на основі (2.63) можна визначити множини маркерів для сторінки  $Page_i$ :

- $ObjectMarker(Page_i) = \{ObjectMarker(Page_i)_j\}_{j=1}^{N(ObjectMarker(Page_i))} \subset ObjectMarker$   
– маркери ТО;
- $ActionMarker(Page_i) = \{ActionMarker(Page_i)_j\}_{j=1}^{N(ActionMarker(Page_i))} \subset ActionMarker$   
– маркери дій над ТО;
- $ExperienceMarker(Page_i) \subset ExperienceMarker$  – маркери відгуків споживачів;
- $TimeMarker(Page_i) = \{TimeMarker(Page_i)_j\}_{j=1}^{N(TimeMarker(Page_i))} \subset TimeMarker$  – маркери часової складової.
- $TermMarker(Page_i) = \{TermMarker(Page_i)_j\}_{j=1}^{M(TermMarker(Page_i))} \subset TermMarker$  – маркери термінів класифікаторів ТД.

Визначимо максимальні значення міри відповідності (2.69), (2.70), (2.71), (2.76) для веб-сторінки:

$$MaxObjectAdequacy(Page_i) = \max_{p=1}^{N(Object)} \max_{j=1}^{N(ObjectMarker(Page_i))} ObjectAdequacy(Object_p, ObjectMarker(Page_i)_j). \quad (3.9)$$

$$MaxExperienceAdequacy(Page_i) = \max_{j=1}^{N(ExperienceMarker(Page_i))} ExperienceAdequacy(ExperienceMarker(Page_i)_j). \quad (3.10)$$

$$\begin{aligned} \text{MaxTimeAdecuacy}(Page_i) = \\ \max_{j=1}^{N(\text{TimeMarker}(Page_i))} \text{TimeAdecuacy}(\text{TimeMarker}(Page_i)_j) \end{aligned} \quad (3.11)$$

$$\begin{aligned} \text{MaxActionAdecuacy}(Page_i) = \\ \max_{p=1}^{N(\text{Object})} \max_{q=1}^{N(\text{Action}(\text{Object}_p))} \max_{j=1}^{N(\text{ActionMarker}(Page_i))} \text{ActionAdecuacy} \left( \begin{array}{l} \text{Action}(\text{Object}_p)_q, \\ \text{ActionMarker}(Page_i)_j \end{array} \right). \end{aligned} \quad (3.12)$$

$$\begin{aligned} \text{MaxTermAdecuacy}(Page_i) = \\ \max_{j=1}^{N(\text{TermMarker}(Page_i))} \text{TermAdecuacy}(\text{TermMarker}(Page_i)_j) \end{aligned} \quad (3.13)$$

Тоді визначимо корисність сторінки як середньо зважене від значень мір відповідності (3.9), (3.10), (3.11), (3.12), (3.13):

$$\begin{aligned} \text{Utility}(Page_i) = & w^{(\text{ObjectAdecuacy})} \text{MaxObjectAdecuacy}(Page_i) + \\ & w^{(\text{ExperienceAdecuacy})} \text{MaxExperienceAdecuacy}(Page_i) + \\ & w^{(\text{TimeAdecuacy})} \text{MaxTimeAdecuacy}(Page_i) + \\ & w^{(\text{ActionAdecuacy})} \text{MaxActionAdecuacy}(Page_i) + \\ & w^{(\text{TermAdecuacy})} \text{MaxTermAdecuacy}(Page_i) \end{aligned}, \quad (3.14)$$

де  $w^{(\text{ObjectAdecuacy})}$ ,  $w^{(\text{ExperienceAdecuacy})}$ ,  $w^{(\text{TimeAdecuacy})}$ ,  $w^{(\text{ActionAdecuacy})}$ ,  $w^{(\text{TermAdecuacy})}$  – вагові коефіцієнти кожної з мір відповідності з області  $[0,1]$ , для яких виконується умова  $w^{(\text{ObjectAdecuacy})} + w^{(\text{ExperienceAdecuacy})} + w^{(\text{TimeAdecuacy})} + w^{(\text{ActionAdecuacy})} + w^{(\text{TermAdecuacy})} = 1$ .

Отже, веб-сторінка містить корисну для ТД інформацію, якщо виконується критерій:

$$\text{Utility}(Page_i) = \alpha^{(\text{Utility})}, \quad (3.15)$$

де  $\alpha^{(\text{Utility})} \in (0,1]$  – мінімально допустиме значення корисності сторінок ВВР.



### **3.3. Розробка процедур лінгвістичного співставлення вмісту відкритих веб-ресурсів з елементами туристичної документації**

Для співставлення вмісту ВВР з елементами ТД спочатку необхідно провести підготовчі дії:

- 1) розкласти веб-сторінки ВВР на інформаційні блоки;
- 2) розкласти інформаційні блоки на окремі дописи;
- 3) розкласти дописи на атомарні ситуації.

Співставлення вмісту ВВР з елементами ТД здійснюється із застосуванням інтелектуального аналізу текстів (Text Mining [67]) атомарних ситуацій. При цьому встановлюється відповідність атомарних ситуацій до ТО ДТО, відгуків споживачів, часової складової і класифікаторів.

#### **3.3.1. Розробка процедури розкладання веб-сторінок на окремі дописи**

Виявлення окремих інформаційних блоків і дописів на веб-сторінці ВВР залежить від формату, у якому подається інформація на сторінці. Проведений аналіз XML-засобів структурування інформації ТД показав, що основними підходами до структурування інформації на веб-ресурсах є використання XML-даних, стандартних XML-мов і HTML-тегів. Серед цих підходів тільки використання стандартних XML-мов (як RSS, Atom) дозволяє повністю автоматизувати процес розкладання веб-сторінок на окремі дописи, оскільки наперед відомі правила розмітки цих сторінок.

Зупинимося детальніше на структуруванні інформації на веб-ресурсах з використанням XML-даних і HTML-тегів. Особливістю цих підходів є те, що залежно від ВВР і окремих сторінок на ВВР може використовуватися свій набір тегів і атрибутів, за допомогою яких маркується текст.

Кожен інформаційний блок чи допис є обмежений ззовні певним набором вкладених тегів з атрибутами чи без них. При цьому будемо відрізняти інформаційні блоки від дописів за такою ознакою: інформаційний блок

обмежений таким набором тегів і атрибутів, який не є притаманний для інших частин сторінки, тоді як певні множини дописів обмежуються однаковим набором тегів і атрибутів.

Алгоритм пошуку інформаційних блоків на веб-сторінці складається з таких кроків:

1. сторінка синтаксично розбирається (parsing) на окремі вкладені теги ' $\langle \text{tag}(i) \rangle$ ' і атрибути ' $\text{attribute}(i,j)=\text{value}(i,j)$ ' у вигляді дерева з коренем, що містить головний тег сторінки ' $\langle \text{tag}(0) \rangle$ '. Вершини впорядковані таким чином, що вершина зліва передує вершині справа. При цьому вершинами графа є теги і атрибути, а дуги позначають вкладеність тегів або належність атрибута до тега;
2. визначається послідовність міток вершин інформаційного блоку як  $\text{BlockNode}=(\langle \text{tag}(i) \rangle)$ , де  $i=0$ ;
3. якщо існує мітка вершини ' $\langle \text{tag}(i+1) \rangle$ ', яка підпорядкована останньому тегу з  $\text{BlockNode}$ , та:
  - 3.1. вона не повторюється в межах дерева, то переходимо на п.2 з  $i=i+1$ ;
  - 3.2. існує ще одна мітка вершини ' $\langle \text{tag}(i+1) \rangle$ ', яка підпорядкована останньому тегу з  $\text{BlockNode}$ , то переходимо на кінець алгоритму;
  - 3.3. існує ще одна мітка вершини  $\langle \text{tag}(i+1) \rangle$ , яка не підпорядкована останньому тегу з  $\text{BlockNode}$ , то  $\text{BlockNode}=(\text{BlockNode}, \langle \text{tag}(i+1) \rangle)$  і переходимо на п.3 з  $i=i+1$ ;
4. якщо не існує мітка вершини ' $\langle \text{tag}(i+1) \rangle$ ', яка підпорядкована останньому тегу з  $\text{BlockNode}$ , але існує ' $\text{attribute}(i,j)=\text{value}(i,j)$ ', що підпорядковується останньому тегу з  $\text{BlockNode}$ , де  $j=1$ , та:
  - 4.1. послідовність  $(\text{BlockNode}, \text{attribute}(i,j)=\text{value}(i,j))$  зустрічається в іншій частині дерева, то за стратегією єдиного атрибута переходимо на п.4 з  $j=j+1$  або за стратегією нагромадження атрибутів переходимо на п.4 з  $\text{BlockNode}=(\text{BlockNode}, \text{attribute}(i,j)=\text{value}(i,j))$  і  $j=j+1$ ;

- 4.2. послідовність (*BlockNode*, 'attribute(*i,j*)=value(*i,j*)') не зустрічається в іншій частині дерева, то  $BlockNode=(BlockNode, 'attribute(*i,j*)=value(*i,j*)')$  і переходимо на кінець алгоритму;
5. якщо не існує мітка вершини '<tag(*i+1*)>', яка підпорядкована останньому тегу з *BlockNode*, але не існує 'attribute(*i,j*)=value(*i,j*)', що підпорядковується останньому тегу з *BlockNode*, то переходимо на п.3 з  $i=i+1$ .

Зауважимо, що в п.4.1 умовний перехід на п.4 означає, що можуть застосовуватися різні стратегії врахування атрибутів. У випадку використання комбінованої стратегії, коли одночасно застосовуються стратегії єдиного атрибута і нагромадження атрибутів, описаний алгоритм в п.4.1 перетворюється в паралельний алгоритм, для ефективною реалізації якого доцільно використовувати відповідні моделі оптимальної редукції [155, 156, 157].

Результатом роботи цього алгоритму є послідовність вершин *BlockNode*, яка легко перетворюється на XML/HTML шаблон інформаційного блоку.

Приклад 3.1. Нехай  $BlockNode=(\langle div \rangle, 'class="content"', \langle h1 \rangle)$ , тоді шаблон інформаційного блоку має вигляд ('...' позначають вміст інформаційного блоку):

```
<div class="content">
  <h1>...</h1>
</div>
```

Для пошуку інших інформаційних блоків за наведеним алгоритмом з дерева тегів і атрибутів веб-сторінки вилучається усе під-дерево, коренем якого є перший тег в послідовності *BlockNode*.

Алгоритм пошуку дописів інформаційного блоку на веб-сторінці доповнює попередній алгоритм і складається з таких кроків:

1. *BlockNode* – вхідна послідовність міток вершин інформаційного блоку, *PostNode*=() – послідовність міток вершин допису;
2. якщо існує мітка вершини '<tag(*i+1*)>', яка підпорядкована останньому тегу з (*BlockNode*, *PostNode*), та:

- 2.1. існує ще одна мітка вершини ' $\langle \text{tag}(i+1) \rangle$ ', яка підпорядкована останньому тегу з  $(\text{BlockNode}, \text{PostNode})$ , то  $\text{PostNode}=(\text{PostNode}, \langle \text{tag}(i+1) \rangle)$  і переходимо на кінець алгоритму;
- 2.2. не існує ще одна мітка вершини ' $\langle \text{tag}(i+1) \rangle$ ', яка підпорядкована останньому тегу з  $(\text{BlockNode}, \text{PostNode})$ , то  $\text{PostNode}=(\text{PostNode}, \langle \text{tag}(i+1) \rangle)$  і переходимо на п.2 з  $i=i+1$ ;
3. якщо не існує мітка вершини ' $\langle \text{tag}(i+1) \rangle$ ', яка підпорядкована останньому тегу з  $(\text{BlockNode}, \text{PostNode})$ , то з  $\text{PostNode}$  вилучаємо усі мітки вершин крім першої і переходимо на кінець алгоритму.

Результатом роботи цього алгоритму є послідовність вершин  $\text{PostNode}$ , яка легко перетворюється на XML/HTML шаблон допису в межах інформаційного блоку. Зауважимо, що якщо  $\text{PostNode}=(\text{PostNode}, \langle \text{tag}(i+1) \rangle)$ , то вміст інформаційного блоку безпосередньо становить один допис. Також, якщо в ході виконання алгоритму здійснився перехід на п.3, то інформаційний блок складається з одного допису.

### **3.3.2. Розробка процедури розкладання дописів веб-сторінок на атомарні ситуації**

Відповідно до інформаційної моделі кожному допису ставиться у відповідність множина атомарних ситуацій (2.61), кожна з яких стосується певної частини тексту допису і характеризується набором маркерів (2.63): ТО, терміну класифікатора ТД, дії над ТО, відгуку споживача, часової складової. При цьому маркер ТО завжди має бути визначений, а інші маркери можуть приймати невизначене значення NULL.

Побудуємо алгоритми розкладання дописів на атомарні ситуації за двома методами:

- методом декомпозиції;
- методом синтезу.

При цьому будемо вважати, що текст допису складається з атомарних

фрагментів – **синтаксичних структур**, якими виступають або окремі речення, або марковані тегами фрагменти тексту.

За методом декомпозиції від тексту допису відкидаються синтаксичні структури доти, доки результуючий фрагмент не буде стосуватися одного маркера ТО.

Алгоритм розкладання дописів на атомарні ситуації за методом декомпозиції складається з таких кроків:

- 1) текст допису  $Text$  синтаксично розбирається (parsing) на окремі вкладені теги ' $\langle tag(i) \rangle$ ', параграфи ' $paragraph(i)$ ' і речення ' $sentence(i)$ ' у вигляді дерева  $Part$ . Вершинами графа є теги, параграфи і речення, а дуги позначають їхню вкладеність. Вершини впорядковані таким чином, що вершина зліва передує вершині справа. При цьому допускаються такі варіанти вкладень:
  - 1.1. Тег вкладається в тег або речення;
  - 1.2. Параграф вкладається в тег;
  - 1.3. Речення вкладається в параграф.
- 2) частина тексту допису, що відповідає дереву  $Part$ ,  $Text(Part)=Text$ ;
- 3)  $ObjectMarker(Part)$  - множина усіх маркерів ТО, які знаходяться в  $Text(Part)$ ;
- 4) якщо  $ObjectMarker(Part)=\emptyset$ , то переходимо на кінець алгоритму;
- 5) якщо  $ObjectMarker(Part)$  складається з одного елемента, то переходимо на п.9;
- 6) якщо  $Part$  містить одну листкову вершину, мітка якої є або ' $\langle tag(i) \rangle$ ', або ' $sentence(i)$ ', то переходимо на п.9;
- 7) вилучаємо з  $Part$  праву листкову вершину, що має мітку ' $\langle tag(i) \rangle$ ' або ' $sentence(i)$ ';
- 8) якщо  $Part$  після вилучення містить листкову вершину з міткою ' $paragraph(i-1)$ ', то її також вилучаємо і переходимо на п.2;

- 9) для  $Text(Part)$  знаходимо множини маркерів  $TermMarker(Part)$ ,  $ActionMarker(Part)$ ,  $ExperienceMarker(Part)$ ,  $TimeMarker(Part)$  відповідно для класифікаторів ТД, дій над ТО, відгуків споживачів, часових складових;
- 10) За (2.63) будуються кортежі усіх атомарних ситуацій, що відповідають елементам декартового добутку  $ObjectMarker(Part) \times TermMarker(Part) \times ActionMarker(Part) \times ExperienceMarker(Part) \times TimeMarker(Part)$ , і переходимо на кінець алгоритму.

Відзначимо, що якщо  $ObjectMarker(Part) = \emptyset$ , то допис не містить інформації, яка підходить для формування ТД.

Для виявлення інших атомарних ситуацій з тексту допису  $Text$  вилучається фрагмент  $Text(Part)$  для дерева  $Part$ , отриманого в кінці алгоритму, і для цього скороченого тексту знову застосовується алгоритм з п.1.

За методом синтезу синтаксична структура допису, яка стосується одного маркера ТО, доповнюється наступними синтаксичними структурами поти, поки результуючий фрагмент буде стосуватися одного ТО.

Алгоритм розкладання дописів на атомарні ситуації за методом синтезу складається з таких кроків:

- 1) текст допису синтаксично розбирається (parsing) на окремі вкладені теги ' $\langle tag(i) \rangle$ ', параграфи ' $paragraph(i)$ ' і речення ' $sentence(i)$ ' у вигляді дерева  $Part$  аналогічно, як в п.1 алгоритму за методом декомпозиції;
- 2) в  $Part'$  додається ліва листкова вершина, що має мітку ' $\langle tag(i) \rangle$ ' або ' $sentence(i)$ ';
- 3)  $Text(Part')$  - частина тексту допису, що відповідає дереву  $Part'$ ;
- 4)  $ObjectMarker(Part')$  - множина усіх маркерів ТО, які знаходяться в  $Text(Part')$ , та:
  - 4.1. якщо  $ObjectMarker(Part') = \emptyset$ , то переходимо на п.5.
  - 4.2. якщо  $ObjectMarker(Part') \neq \emptyset$ , то переходимо на п.7.

5) якщо існує листкова вершина  $node(i+1)$  з міткою ' $\langle tag(i+1) \rangle$ ' або ' $sentence(i+1)$ ', що не належить  $Part'$ , та:

5.1. підпорядковується вершині з  $Part'$ , то ця листкова вершина додається до  $Part'$ , тобто  $Part' = ADD(Part', node(i+1))$ , і переходимо на п.3;

5.2. не підпорядковується вершині з  $Part'$ , то ця листкова вершина додається до  $Part'$ , а також до  $Part'$  додаються ті вершини  $node(i-1)$ ,  $node(i-2)$ ,... вищих рівнів відносно цієї листкової вершини і під-дерева  $Part'$ , тобто  $Part' = ADD(Part', node(i+1), node(i-1), node(i-2), \dots)$ , що в результаті отримуємо нове під-дерево  $Part'$  з галуженням в корені, і переходимо на п.3.

б) якщо не існує листкова вершина з міткою ' $\langle tag(i+1) \rangle$ ' або ' $sentence(i+1)$ ', що не належить  $Part'$ , то переходимо на кінець алгоритму;

7) якщо існує листкова вершина  $node(i+1)$  з міткою ' $\langle tag(i+1) \rangle$ ' або ' $sentence(i+1)$ ', що не належить  $Part'$ , та:

7.1. підпорядковується вершині з  $Part'$ , то:

7.1.1. якщо  $ObjectMarker(Part') = ObjectMarker(ADD(Part', node(i+1)))$ , то  $Part' = ADD(Part', node(i+1))$  і переходимо на п.7;

7.1.2. якщо  $ObjectMarker(Part') \neq ObjectMarker(ADD(Part', node(i+1)))$ , то переходимо на кінець алгоритму;

7.2. не підпорядковується вершині з  $Part'$ , то:

7.2.1. якщо  $ObjectMarker(Part') = ObjectMarker(ADD(Part', node(i+1), node(i-1), node(i-2), \dots))$ , то  $Part' = ADD(Part', node(i+1), node(i-1), node(i-2), \dots)$  і переходимо на п.7;

7.2.2. якщо  $ObjectMarker(Part') \neq ObjectMarker(ADD(Part', node(i+1), node(i-1), node(i-2), \dots))$ , то переходимо на кінець алгоритму.

Для виявлення інших атомарних ситуацій з тексту допису  $Text$  вилучається фрагмент  $Text(Part')$  для дерева  $Part$ , отриманого в кінці алгоритму, і для цього скороченого тексту знову застосовується алгоритм з п.1.

### 3.3.3. Розробка процедури лінгвістичного співставлення атомарних ситуацій дописів веб-сторінок з елементами туристичної документації

За допомогою інтелектуального аналізу текстів атомарні ситуації дописів відносяться до ТО, ДТО, відгуків споживачів, часової складової і класифікаторів. Віднесення атомарних ситуацій дописів відбувається на основі критеріїв (2.72), (2.73), (2.74), (2.77) і (2.80) за такими кроками:

- 1)  $PostPart_{ijkl}$  - атомарна ситуація (2.63), отримана за одним з алгоритмів розкладання дописів на атомарні ситуації;
- 2)  $FoundObjects(PostPart_{ijkl})$  – множина ТО, отримана за допомогою алгоритму виявлення ТО, які відповідають атомарній ситуації допису (блок-схема цього алгоритму подана на рис. 3.4);
- 3) якщо  $FoundObjects(PostPart_{ijkl}) = \emptyset$  - порожня множина, то переходимо на кінець алгоритму;
- 4) якщо не існує
 
$$FoundObjects(PostPart_{ijkl})_p = NEXT(FoundObjects(PostPart_{ijkl})),$$
 то переходимо на кінець алгоритму (NEXT – операція витягування наступного елемента множини);
- 5) якщо не виконується (2.72) для ТО  $FoundObjects(PostPart_{ijkl})_p$ , то переходимо на п.4;
- 6)  $FoundActions(FoundObjects(PostPart_{ijkl}))$  – множина ДТО, отримана за допомогою алгоритму виявлення ДТО, які відповідають атомарній ситуації допису (блок-схема цього алгоритму подана на рис. 3.5) ;



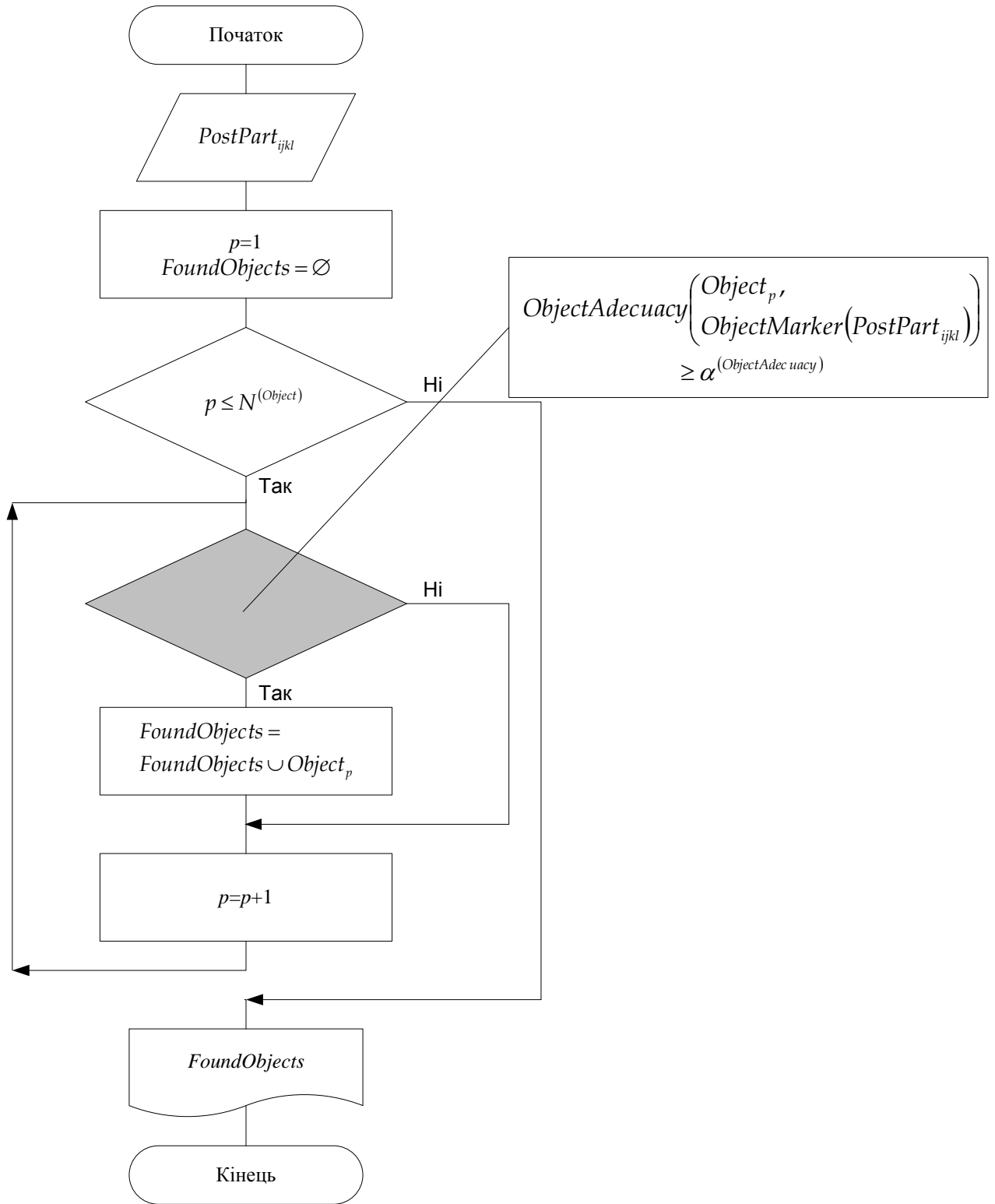


Рис. 3.4. Алгоритм виявлення туристичних об'єктів, які відповідають атомарній ситуації допису

7) якщо не існує  $FoundActions(FoundObjects(PostPart_{ijkl})_p)_q = NEXT(FoundActions(FoundObjects(PostPart_{ijkl})))$ , то

переходимо на п.4;

- 8) якщо  $FoundActions(FoundObjects(PostPart_{ijkl})_p)_q = \emptyset$ , то для ТО не виявлено дій в атомарній ситуації і переходимо на п.13;
- 9) якщо виконується (2.73), то атомарна ситуація  $PostPart_{ijkl}$  відноситься до відгуків споживачів, в протилежному випадку – до описової частини ТД;
- 10) якщо виконується (2.74), то атомарна ситуація  $PostPart_{ijkl}$  відноситься до подій, в протилежному випадку – до фактів;
- 11) атомарна ситуація  $PostPart_{ijkl}$  відноситься до множини термінів класифікаторів, отриманих за допомогою алгоритму виявлення термінів класифікаторів (рис. 3.6) ;
- 12) атомарна ситуація  $PostPart_{ijkl}$  зберігається як факт чи подія за п.10 до ДТО  $FoundActions(FoundObjects(PostPart_{ijkl})_p)_q$  в описовій частині ТД чи у відгуках споживачів за п.9, з класифікацією за п.11, і переходимо на п.7;
- 13) якщо виконується (2.73), то атомарна ситуація  $PostPart_{ijkl}$  відноситься до відгуків споживачів, в протилежному випадку – до описової частини ТД;
- 14) якщо виконується (2.74), то атомарна ситуація  $PostPart_{ijkl}$  відноситься до подій, в протилежному випадку – до фактів;
- 15) атомарна ситуація  $PostPart_{ijkl}$  відноситься до множини класифікаторів, отриманих за допомогою алгоритму виявлення класифікаторів (блок-схема цього алгоритму подана на рис. 3.6) ;
- 16) в атомарній ситуації  $PostPart_{ijkl}$  з урахуванням її належності до допису, інформаційного блоку і сторінки за алгоритмом виявлення атрибутів фактів і подій визначаються відповідні атрибути (блок-схема цього алгоритму подана на рис. 3.7) ;

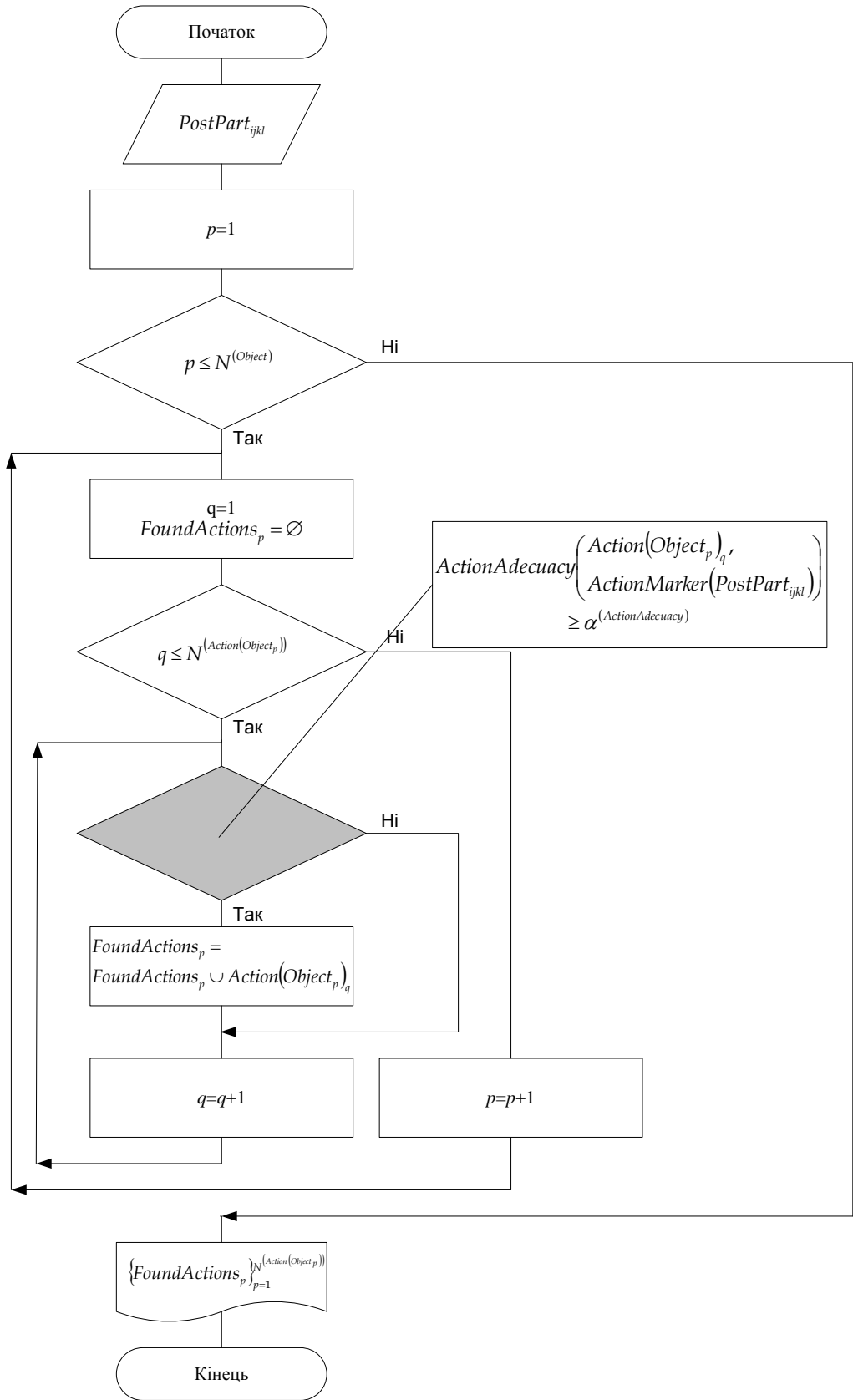


Рис. 3.5. Алгоритм виявлення дій туристичних об'єктів, які відповідають атомарній ситуації допису



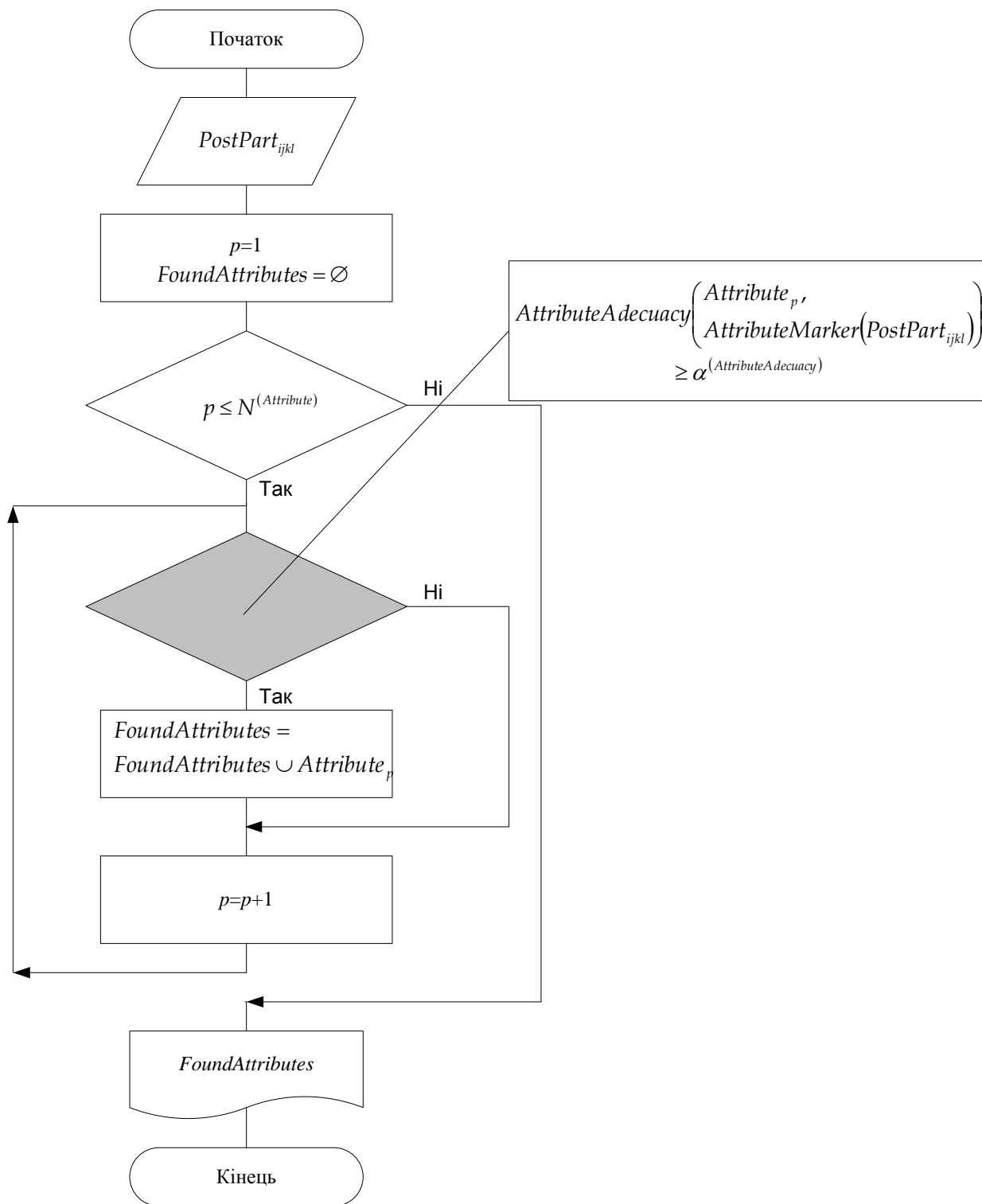


Рис. 3.7. Алгоритм виявлення атрибутів фактів і подій, які відповідають атомарній ситуації допису

17) на основі знайдених атрибутів за п.16 визначаються значення

атрибутів факту чи події, що відповідають атомарній ситуації  $PostPart_{ijkl}$ ;

- 18) атомарна ситуація  $PostPart_{ijkl}$  зберігається як факт чи подія за п.14 до ТО  $FoundObjects(PostPart_{ijkl})_p$  в описовій частині ТД чи у відгуках споживачів за п.13, з класифікацією за п.15, з значеннями атрибутів за п.17, і переходимо на п.7.

Для того, щоб відносити сукупності семантично тотожних атомарних ситуацій дописів до ТО, ДТО, відгуків споживачів, часової складової і класифікаторів за критеріями (2.85), (2.86), (2.87), (2.88) і (2.89), необхідно внести такі зміни в процедурі:

- в кроках 9, 13 використати критерій (2.86), а в кроках 10, 14 використати критерій (2.87);
- в алгоритмах на рис. 3.4, 3.5, 3.6 необхідно застосувати відповідно критерії (2.85), (2.88) і (2.89);
- в алгоритмі рис. 3.7 необхідно застосувати критерій (2.90).

Розроблена процедура співставлення атомарних ситуацій дописів Веб-сторінок з елементами туристичної документації у випадку виконання усіх необхідних критеріїв ставить у відповідність атомарну ситуацію певному новому факту чи новій події, що з'являється у відповідній частині ТД.

Відзначимо, що за цією процедурою у фактах чи подіях ТД зберігається не тільки ІН атомарної ситуації, але й заповнюються атрибути, що максимально точно відображають інформацію про заголовок допису, джерело, автора, час тощо – атрибути Дублінського ядра опису метаданих. В інформаційних моделях фактів (2.37), (2.38), (2.39), (2.40) та подій (2.41), (2.42), (2.43), (2.44) атрибути Дублінського ядра згруповані так:

- *MandatoryProperties(...)* – обов'язкові атрибути (як Заголовок);
- *Source(...)* – атрибути джерела інформації (як Видавець, URL-адреса);
- *Author(...)* - атрибути про автора (як Ім'я, Псевдо);

- *Language(...)* – атрибути мови інформації.
- *Time(...)* – атрибути, пов'язані з часовими властивостями публікації (як Дійсний час події, Час публікації на сайті джерела).

Більшість з цих атрибутів маркуються за допомогою відповідних тегів в XML- чи HTML-файлах веб-сторінок незалежно від адреси веб-сайту, способу його реалізації чи маркування тексту. Тому для наочності на рис. 3.8 подано рівні сторінки RSS-файлу, на яких можуть маркуватися атрибути за допомогою тегів (у Додатку А наведено приклад повного тексту RSS-файлу).

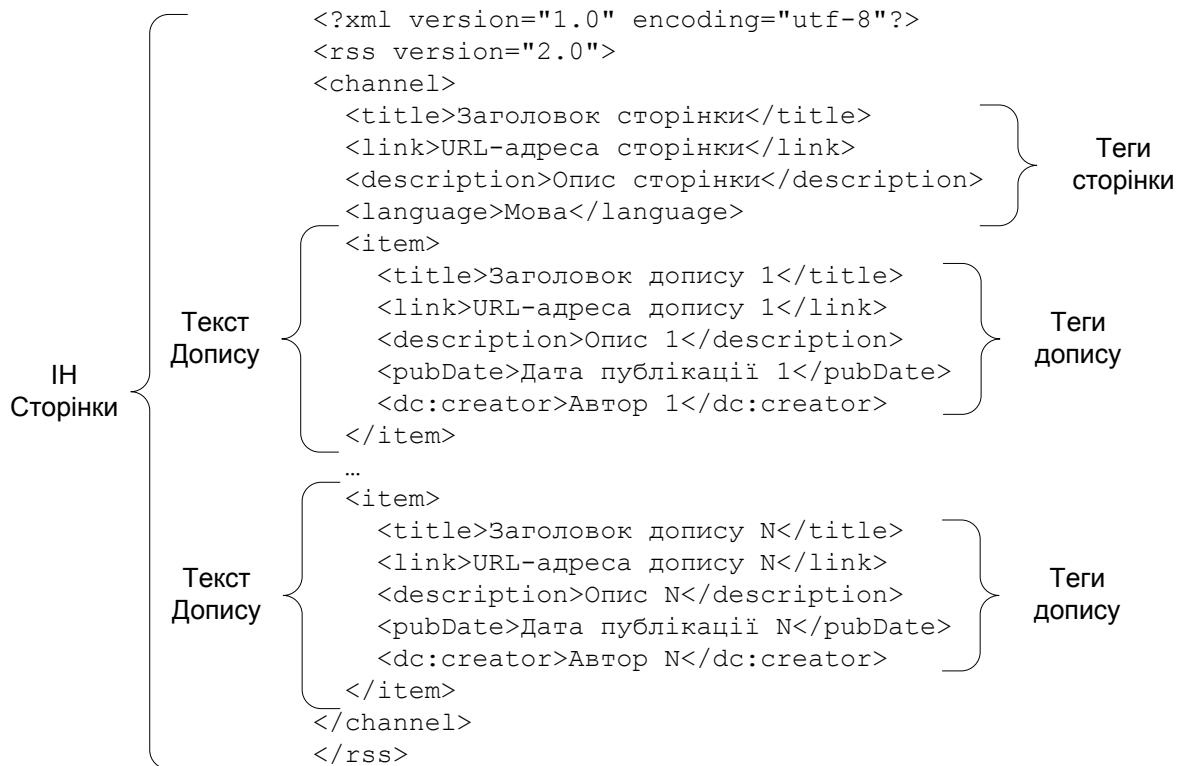


Рис. 3.8. Рівні тегів у структурі RSS-файлу

На цьому рисунку не подано атрибути, які можуть з'являтися на рівні атомарних ситуацій в межах допису. Однак із структури RSS-файлу випливає, що Опис допису, маркований тегом `<description>`, може містити текст з HTML-розміткою, у якому якраз і можуть виявлятися теги рівня атомарної ситуації.

Під час співставлення атомарних ситуацій з атрибутами ми не розрізняли,

чи дані атрибутів фактів та подій на веб-сторінці маркуються за допомогою тегів, чи їхню позицію можна визначити за певними символічними ознаками в тексті сторінки чи допису.

### **3.4. Розробка процедур виявлення прогалин у туристичній документації**

Розроблені алгоритми співставлення атомарних ситуацій дописів веб-сторінок з елементами ТД дозволяють наповнювати ТД консолідованою інформацією з ВВР. Однак ці алгоритми не вказують на можливу неточність і некоректність як структури ТД, так і її вмісту у вигляді фактів і подій. Одним із методів, який дозволяє знайти і усунути недоліки ТД, є метод виявлення та заповнення прогалин.

Поняття «прогалина» використовується в багатьох галузях, зокрема в законодавстві прогалина – це відсутність потрібної норми в законодавчих актах. У цій роботі під **прогалиною** в ТД будемо розуміти певну частину ТД, яка не наповнена інформацією про факти чи події.

Побудована загальна модель ТД дозволяє виділити два види прогалин:

1. Структурні прогалини в ТД.
2. Прогалини у вмісті ТД.

#### **3.4.1. Розробка процедур виявлення структурних прогалин у туристичній документації**

Структурні прогалини в ТД виникають тоді, коли структура ТД не відповідає наперед заданим вимогам, які формулюються як надавачем, так і споживачем ТП.

Загалом наявність структурної прогалини можна сформулювати так: множина однотипних елементів ТД містить менше елементів, ніж відповідна множина цих елементів у вимогах до ТД. Тобто, якщо  $Element^{(TD)}$  і



$Element^{(Requirement)}$  – відповідно множини елементів в ТД і у вимогах до ТД, то в загальному випадку має місце структурна прогалина, якщо:

$$Element^{(Gap)} = Element^{(Requirement)} \setminus Element^{(TD)} \neq \emptyset. \quad (3.16)$$

До елементів ТД, які можуть спричиняти структурні прогалини, віднесемо ТО, ДТО, класифікатори і терміни класифікаторів. Для усунення цих прогалин необхідно доповнити новими елементами відповідні множини.

Доповнення множин ТО і ДТО новими елементами відбувається за алгоритмом з такими кроками:

- 1) до множини  $Object$  додати новий елемент  $Object_i^{(Gap)}$  відповідно до вимог ТД, вказавши атрибути факту визначення цього об'єкта за (2.13) і (2.33), включаючи класифікатори;
- 2) визначити індикативну ознаку об'єкта  $ObjectIndicator(Object_i^{(Gap)})$  за (2.66);
- 3) включити об'єкт  $Object_i^{(Gap)}$  в туристичні послуги і відгуки споживачів відповідно до вимог ТД;
- 4) до множини дій  $Action(Object_i^{(Gap)})$  над ТО додати новий елемент  $Action(Object_i^{(Gap)})_j$  відповідно до вимог ТД, вказавши атрибути факту визначення цієї дії за (2.18) і (2.34), включаючи класифікатори;
- 5) визначити індикативну ознаку ДТО  $ActionIndicator(Action(Object_i^{(Gap)})_j)$  за (2.75);
- 6) включити ДТО  $Action(Object_i^{(Gap)})_j$  в туристичні послуги і відгуки споживачів відповідно до вимог ТД;

Зауважимо, що процедуру додавання нової дії над наявним в ТД об'єктом необхідно починати з п.4.

Доповнення множин класифікаторів і термінів класифікаторів ТД новими елементами відбувається за алгоритмом з такими кроками:

- 1) до множини *Classifier* додати новий елемент  $Classifier_i^{(Gap)}$  відповідно до вимог ТД;
- 2) до множини термінів  $Term_i^{(Gap)}$  класифікатора  $Classifier_i^{(Gap)}$  додати новий елемент  $Term_{ij}^{(Gap)}$  відповідно до вимог ТД за (2.32);
- 3) до множини зв'язків між термінами  $TermRels_i^{(Gap)}$  додати зв'язки нового елемента  $Term_{ij}^{(Gap)}$  з іншими термінами класифікатора  $Classifier_i^{(Gap)}$ ;
- 4) визначити індикативну ознаку терміну  $TermIndicator(Term_{ij}^{(Gap)})$  за (2.78).

Зауважимо, що процедуру додавання нового терміну для наявного класифікатора ТД необхідно починати з п.2.

### 3.4.2. Розробка процедур виявлення прогалин вмісту в туристичній документації

Прогалини вмісту в ТД виникають через відсутність інформації про факти і події, пов'язаної з певними структурними елементами ТД – ТО, ДТО та термінами класифікаторів.

Загалом наявність прогалини вмісту можна сформулювати так: для структурного елемента ТД не усі необхідні факти та події з ним пов'язані. Будемо розрізняти такі типи прогалини вмісту:

1. факти і факти визначення ТО і ДТО не пов'язуються з термінами класифікаторів;
2. факти не пов'язуються з ТО і ДТО;
3. події не пов'язуються з ТО і ДТО;

Якщо  $Term_{ij}$  – термін класифікатора,  $ElementDefiniton^{(TD)}(Term_{ij})$  – множина усіх фактів визначення ТО і ДТО, пов'язаних з цим терміном в ТД,  $ElementDefiniton^{(Expected)}(Term_{ij})$  – множина усіх фактів визначення ТО і ДТО, які

очікувано мали б бути пов'язані з цим терміном, то має місце прогалина вмісту за класифікацією фактів визначення, якщо:

$$\begin{aligned} &ElementDefiniton^{(Gap)}(Term_{ij})= \\ &ElementDefiniton^{(Expected)}(Term_{ij}) \setminus ElementDefiniton^{(TD)}(Term_{ij}) \neq \emptyset \end{aligned} \quad (3.17)$$

Якщо  $ElementFact^{(TD)}(Term_{ij})$  – множина усіх фактів ТО і ДТО, пов'язаних з терміном  $Term_{ij}$  в ТД,  $ElementFact^{(Expected)}(Term_{ij})$  – множина усіх фактів ТО і ДТО, які очікувано мали б бути пов'язані з цим терміном, то має місце прогалина вмісту за класифікацією фактів, якщо:

$$\begin{aligned} &ElementFact^{(Gap)}(Term_{ij}) \\ &= ElementFact^{(Expected)}(Term_{ij}) \setminus ElementFact^{(TD)}(Term_{ij}) \neq \emptyset \end{aligned} \quad (3.18)$$

Якщо  $ElementEvent^{(TD)}(Term_{ij})$  – множина усіх подій ТО і ДТО, пов'язаних з терміном  $Term_{ij}$  в ТД,  $ElementEvent^{(Expected)}(Term_{ij})$  – множина усіх подій ТО і ДТО, які очікувано мали б бути пов'язані з цим терміном, то має місце прогалина вмісту за класифікацією подій, якщо:

$$\begin{aligned} &ElementEvent^{(Gap)}(Term_{ij})= \\ &ElementEvent^{(Expected)}(Term_{ij}) \setminus ElementEvent^{(TD)}(Term_{ij}) \neq \emptyset \end{aligned} \quad (3.19)$$

Якщо  $Element_i$  – ТО або ДТО,  $ElementFact^{(TD)}(Element_i)$  – множина усіх фактів цього ТО або ДТО в ТД,  $ElementFact^{(Expected)}(Element_i)$  – очікувана множина усіх фактів ТО або ДТО, то має місце прогалина вмісту фактів, якщо:

$$\begin{aligned} &ElementFact^{(Gap)}(Element_i)= \\ &ElementFact^{(Expected)}(Element_i) \setminus ElementFact^{(TD)}(Element_i) \neq \emptyset \end{aligned} \quad (3.20)$$

Якщо  $ElementEvent^{(TD)}(Element_i)$  – множина усіх подій ТО або ДТО  $Element_i$  в ТД,  $ElementEvent^{(Expected)}(Element_i)$  – очікувана множина усіх подій ТО або ДТО, то має місце прогалина вмісту подій, якщо:

$$\begin{aligned} &ElementEvent^{(Gap)}(Element_i) = \\ &ElementEvent^{(Expected)}(Element_i) \setminus ElementEvent^{(TD)}(Element_i) \neq \emptyset \end{aligned} \quad (3.21)$$

Прогалини вмісту за класифікацією фактів визначення (3.17) усувається експертом, який повинен визначити множину термінів, які пов'язуються з ТО або ДТО.

Прогалини вмісту за класифікацією фактів і подій (3.18), (3.19), та прогалини вмісту фактів і подій (3.20), (3.21) усуваються такими способами:

- 1) використання нових ВВР, які краще відповідають наявним індикативним ознакам;
- 2) уточнення індикативних ознак ТО (2.66), досвіду (2.67), часової складової (2.68), ДТО (2.75) і терміну класифікатора (2.78) методом частотного аналізу.

Для виявлення нових ВВР, які краще відповідають наявним індикативним ознакам, необхідно застосувати алгоритм формування множини джерел ВВР рис. 3.3 за ключовими словами *Keywords*, які задані в маркерах індикативним ознак, і перевірити корисність Веб-сторінок цих ВВР за критерієм (3.15).

### **3.4.3. Частотний аналіз маркерів та індикативних ознак для усунення прогалин вмісту**

Частотний аналіз широко застосовується в криптографії і ґрунтується на наявності статистичного розподілу окремих символів і їх послідовностей у тексті повідомлення. Тобто, частотний аналіз передбачає, що частота появи заданого символу в досить довгих текстах одна і та ж для окремих текстів повідомлень.

Нехай *Element* – певний елемент ТД, який відображає або певний ТО, ДТО чи термін класифікатора, або певну ознаку досвіду чи часової складової. Тоді, відповідно на основі індикативних ознак ТО (2.66), ДТО (2.75), терміну класифікатора (2.78), досвіду (2.67), часової складової (2.68) узагальнимо поняття індикативної ознаки для *Element*:

$$ElementIndicator = \left\{ \left\langle \begin{array}{l} Marker(ElementIndicator)_i, \\ \mu(ElementIndicator)_i \end{array} \right\rangle \right\}_{i=1}^{N^{(ElementIndicator)}}, \quad (3.22)$$

де  $Marker(ElementIndicator)_i \in Marker$  –  $i$ -й маркер для відображення елемента ТД,  $\mu(ElementIndicator)_i \in [0,1]$  – міра відповідності  $i$ -го маркера до елемента ТД,  $N^{(ElementIndicator)}$  – кількість маркерів, що відображають елемента ТД.

Для застосування частотного аналізу маркерів для усунення прогалин вмісту в ТД необхідно проаналізувати частоту появи відповідних маркерів під час співставлення атомарної ситуації зі складовими туристичної документації.

Припустимо, що в результаті опрацювання ВВР виявлено  $N^{(PostPart,Processing)}$  атомарних ситуацій, і з них  $N^{(PostPart,Marker(ElementIndicator)_i)}$  – співставлені з індикативними ознаками, що містять маркер  $Marker(ElementIndicator)_i$ . Тоді частота появи маркера  $Marker(ElementIndicator)_i$  в атомарних ситуаціях – це така величина:

$$MarkerFrequency(Marker(ElementIndicator)_i) = \frac{N^{(PostPart,Marker(ElementIndicator)_i)}}{N^{(PostPart,Processing)}}. \quad (3.23)$$

Для оцінки частоти появи маркера введемо критерій:

$$MarkerFrequency(Marker(ElementIndicator)_i) \geq \alpha^{(MarkerFrequency)}, \quad (3.24)$$

де  $\alpha^{(MarkerFrequency)} \in (0,1]$  – мінімально допустиме значення для частоти появи маркера.

Наявність надто великої кількості маркерів, що не задовольняють критерій (3.24), може бути причиною наявності прогалин вмісту. Прогалини на основі частотного аналізу маркерів усуваються такими способами:

- 1) використання нових ВВР, які краще відповідають маркерам, що не задовольняють критерій (3.24);

- 2) здійснення переоцінки міри відповідності  $\mu(\text{ElementIndicator})_i$  маркера  $\text{Marker}(\text{ElementIndicator})_i$  для кожної індикативної ознаки (3.22), де зустрічається цей маркер;
- 3) вилучення маркера  $\text{Marker}(\text{ElementIndicator})_i$  з множини усіх маркерів  $\text{Marker}$ , і доповнення цієї множини новими маркерами.

Використання частотного аналізу індикативних ознак для усунення прогалин вмісту в ТД ґрунтується на аналізі частоти співставлення атомарних ситуацій з індикативними ознаками.

Припустимо, що в результаті опрацювання ВВР  $N^{(\text{PostPart}, \text{ElementIndicator})}$  атомарних ситуацій співставлені з індикативною ознакою (3.22). Тоді частота появи цієї індикативної ознаки в атомарних ситуаціях – це така величина:

$$\text{IndicatorFrequency}(\text{ElementIndicator}) = \frac{N^{(\text{PostPart}, \text{ElementIndicator})}}{N^{(\text{PostPart}, \text{Processing})}}. \quad (3.25)$$

Відповідний критерій для частоти появи індикативної ознаки має вигляд:

$$\text{IndicatorFrequency}(\text{ElementIndicator}) \geq \alpha^{(\text{IndicatorFrequency})}, \quad (3.26)$$

де  $\alpha^{(\text{IndicatorFrequency})} \in (0, 1]$  – мінімально допустиме значення для частоти появи індикативної ознаки.

Якщо індикативна ознака не задовольняє критерій (3.26) то вона може бути причиною наявності прогалин вмісту відповідного елемента ТД. Прогалини на основі частотного аналізу появи індикативних ознак усуваються такими способами:

- 1) використання нових ВВР, які краще відповідають індикативним ознакам, що не задовольняють критерій (3.26);
- 2) здійснення переоцінки мір відповідності  $\mu(\text{ElementIndicator})_i$  для індикативної ознаки (3.22);
- 3) вилучення індикативної ознаки, якщо елементом ТД є ТО, ДТО або термін класифікатора;

4) доповнення множини маркерів *Marker(ElementIndicator)*<sub>i</sub> новими маркерами.

### **3.5. Розробка процедури виявлення некоректної інформації у туристичній документації**

Під час наповнення ТД або її використання можуть виникати певні неточності ІН. Зокрема, це [86, 91, 93, 94]:

- наявність надлишкової інформації;
- наявність неактуальної інформації;
- неточність інформації через структурні прогалини.

Надлишковість інформації в ТД пов'язана з додаванням нової інформації про факти та події, яка ідентична або семантично близька до вже існуючої в ТД інформації. Наявність повторів в ІН приводить до поганого сприйняття інформації ТД.

Неактуальність інформації в ТД пов'язана з відсутністю її оновлень в часі. Поняття неактуальної інформації є відносним [38].

Приклад 3.2. Дані про вартість ТП, наявність місць проживання тощо вимагає частого періодичного оновлення. Тоді як загальна інформація про ТО чи ДТО може не змінюватись досить довго. Відгуки споживачів ТП також мають певний термін актуальності, який не такий малий як для вартості ТП, але не такий великий, як для загальної інформації про ТО.

Неточність інформації через структурні прогалини з'являється під час необхідності доповнити ТД новими структурними елементами: ТО, ДТО, класифікаторами і термінами. Тобто на момент внесення інформації про факт чи подію відбулося співставлення з одними структурними елементами ТД, тоді як поява нових структурних елементів може спричинити інші результати процедур співставлення. Це пов'язано з тим, що індикативні ознаки нових структурних елементів ТД можуть містити маркери, які раніше не застосовувалися.

Ми не будемо детально зупинятися на проблемах неточності ТД через структурні прогалини, оскільки вони усуваються шляхом повторного застосування процедур співставлення до ТД. При цьому усі факти і події, які містяться в ТД, розглядаються як дописи ВВР.

### 3.5.1. Розробка процедури виявлення надлишкової інформації у туристичній документації

Для усунення надлишкових фактів та подій в ТД необхідно задати правила їхнього порівняння. Це дозволить чітко встановити ситуації семантичної подібності фактів та подій.

Основним методом визначення подібності фактів та подій будемо вважати порівняння подібності їхнього текстового ІН. Існує ряд підходів до визначення подібності текстів. Однак усі вони базуються на певному числовому вираженні подібності, якими, зокрема, є [122]:

- кількість однакових шинглів (ланцюжків слів);
- кількість однакових описових (важливих) слів;
- відстань перестановки символів;
- структурна (графова) міра подібності.

У цій роботі не розглядаються питання подібності текстів. З погляду усунення надлишкової інформації з ТД важливою є можливість обчислення міри подібності між двома текстами. Тобто, якщо  $Text_1$  і  $Text_2$  – два фрагменти тексту, то міра подібності між ними:

$$Similarity(Text_1, Text_2) \in [0, 1], \quad (3.27)$$

де значення 1 позначає повну подібність (тотожність) текстів, а 0 – їхню повну взаємну унікальність.

Надлишковість інформації не стосується визначень (2.33), (2.34), (2.35), (2.36), оскільки доцільність додавання нових ТО і ДТО визначає експерт з побудови ТД. Тому будемо розглядати подібність в контексті ІН фактів (2.37), (2.38), (2.39), (2.40) і подій (2.41), (2.42), (2.43), (2.44).



Нехай  $ElementFact$  і  $ElementEvent$  – відповідно подія або факт, які відносяться до ТО або ДТО. І нехай  $Description(ElementFact) \in MandatoryProperties(ElementFact)$  і  $Description(ElementEvent) \in MandatoryProperties(ElementEvent)$  – один з обов'язкових атрибутів факту чи події, який відображає їхній текстовий опис (вміст). Тоді за атрибутом  $Description$  можна визначати подібність фактів чи подій (3.27). Тобто для двох фактів  $ElementFact_1$  і  $ElementFact_2$  та подій  $ElementEvent_1$   $ElementEvent_2$  мають місце вирази:

$$\begin{aligned} Similarity(ElementFact_1, ElementFact_2) = \\ Similarity(Description(ElementFact_1), Description(ElementFact_2)), \end{aligned} \quad (3.28)$$

$$\begin{aligned} Similarity(ElementEvent_1, ElementEvent_2) = \\ Similarity(Description(ElementEvent_1), Description(ElementEvent_2)). \end{aligned} \quad (3.29)$$

Критерії подібності формулюються так:

$$Similarity(ElementFact_1, ElementFact_2) \geq \alpha^{(Similarity)}, \quad (3.30)$$

$$Similarity(ElementEvent_1, ElementEvent_2) \geq \alpha^{(Similarity)}. \quad (3.31)$$

При виявленні двох подібних елементів (фактів чи подій) необхідно здійснити процедуру усунення надлишковості шляхом об'єднання елементів. При цьому треба визначити, що робити з атрибутами і класифікаторами елементів. Тут можливі такі стратегії:

- стратегія переваги – віддається перевага значенню атрибута чи класифікатора першого або другого елемента;
- стратегія об'єднання – здійснюється спроба об'єднати значення атрибута чи класифікатора двох елементів.

Використання цих стратегій залежить від особливостей характеристики елемента. Наприклад, для числових чи часових характеристик можливо скористатися лише стратегією переваги, тоді як для класифікаторів і текстових

характеристик можна скористатися і стратегією об'єднання. При цьому доцільно враховувати рейтинг джерела. Тобто якщо

$$Rating(Source(ElementFact_1)) > Rating(Source(ElementFact_2)), \quad (3.32)$$

то перевага віддається першому факту.

Аналогічно формулюється критерій переваги для другого факту, а також для подій.

У випадку, якщо рейтинги джерел спів мірні або однакові:

$$Rating(Source(ElementFact_1)) \approx Rating(Source(ElementFact_2)), \quad (3.33)$$

то доцільно застосовувати стратегію об'єднання для текстових характеристик і класифікаторів.

Процедура усунення надлишкової інформації в ТД складається з таких кроків:

- 1)  $Element_1$  і  $Element_2$  - два елементи ТД (факти чи події), які стосуються одного ТО чи ДТО і які треба перевірити на подібність;
- 2) якщо не виконується критерій  $Similarity(Element_1, Element_2) \geq \alpha^{(Similarity)}$ , то переходимо на кінець алгоритму;
- 3) для кожної пари атрибутів  $MandatoryProperties(Element_1)_i$  і  $MandatoryProperties(Element_2)_i$  визначається стратегія усунення надлишковості, за якою формується значення нового атрибута  $MandatoryProperties(Element_{NEW})_i$  (перелік атрибутів для двох елементів є однаковий, оскільки вони стосуються одного ТО або ДТО);
- 4) для класифікаторів  $Classifier(Element_1)$  і  $Classifier(Element_2)$  за визначеною стратегією усунення надлишковості формується множина  $Classifier(Element_{NEW})$ ;

- 5) аналогічно до п.3 і п.4 визначаються  $Source(Element_{NEW})$  – атрибути джерела,  $Author(Element_{NEW})$  – атрибути про автора,  $Language(Element_{NEW})$  – атрибути мови,  $Time(Element_{NEW})$ ;
- б) сформований новий елемент  $Element_{NEW}$ , побудований за допомогою  $MandatoryProperties(Element_{NEW})_i$ ,  $Classifier(Element_{NEW})$ ,  $Source(Element_{NEW})$ ,  $Author(Element_{NEW})$ ,  $Language(Element_{NEW})$  і  $Time(Element_{NEW})$  (лише для подій) замінює в ТД наявні елементи  $Element_1$  і  $Element_2$ .

В результаті виконання цього алгоритму замість двох семантично подібних елементів  $Element_1$  і  $Element_2$  (фактів чи подій, що відносяться до одного ТО або ДТО), в ТД з'являється один елемент  $Element_{NEW}$ .

### 3.5.2. Розробка процедури усунення неактуальної інформації у туристичній документації

**Актуальність інформації** – це ступінь відповідності (важливості, значимості, цінності тощо) інформації до поточного моменту часу.

Очевидно, що в ТД про актуальність доцільно говорити лише в контексті подій і усіх структурних елементів ТД, з якими пов'язуються події. Нехай  $ElementEvent$  – подія в ТД,  $Time(ElementEvent)$  – відповідна множина значень часових атрибутів цієї події. Тоді подія  $ElementEvent$  є актуальною відносно  $i$ -того атрибута  $Time(ElementEvent)_i$ , якщо

$$T_{NOW} - Time(ElementEvent)_i \leq T^{(Actuality)}, \quad (3.34)$$

де  $T_{NOW}$  – поточний момент часу,  $T^{(Actuality)}$  – термін (часовий відрізок) актуальності.

Процедура усунення неактуальної інформації в ТД складається з таких кроків:

- 1)  $OutdatedEvent = \emptyset$  - множина неактуальних подій ТД;

2) для кожної події  $ElementEvent_i$  з ТД визначається її актуальність по кожному часовому атрибуту з  $Time(ElementEvent_i)$  (3.34), і:

2.1. якщо хоча б по одному часовому атрибуту подія  $ElementEvent_i$  є неактуальною, то  $OutdatedEvent = OutdatedEvent \cup ElementEvent_i$ .

3) для нової ТД, яка будується з заданої ТД шляхом вилучення усіх подій з  $OutdatedEvent$  здійснюються процедури усунення прогалин вмісту;

4) нова ТД замінює задану ТД.

Очевидно, що під час виконання п.3 цього алгоритму у першу чергу доцільно використовувати ті ВВР, з яких були отримані події з  $OutdatedEvent$ .

### 3.6. Висновки до розділу

1. Проведено структурування процесу формування туристичної документації з використанням відкритих веб-ресурсів шляхом виділення відповідних етапів, що дало змогу розробити процедури формування туристичної документації і архітектуру системи автоматизованого формування туристичної документації.
2. Розроблено процедуру вибору відкритих веб-ресурсів для формування туристичної документації на основі критеріїв відповідності, що дало можливість оперувати лише релевантною множиною веб-сторінок для формування туристичної документації.
3. Розроблено процедури лінгвістичного співставлення вмісту відкритих веб-ресурсів з елементами туристичної документації за допомогою індикативних ознак, які є необхідними для побудови системи автоматизованого формування туристичної документації.
4. Розроблено процедури виявлення прогалин у туристичній документації шляхом аналізу структури туристичної документації, її вмісту, частотного аналізу маркерів та індикативних ознак, які дозволяють покращити функціональні показники якості, показники якості придатності до

використання та показників якості супроводжуваності туристичної документації.

5. Розроблено процедури виявлення некоректної інформації у туристичній документації які дозволяють покращити показники якості придатності до використання та показників якості супроводжуваності туристичної документації.

## **Розділ 4. Побудова системи комп'ютерно-лінгвістичного формування туристичної документації**

Розроблені процедури комп'ютерно-лінгвістичного формування туристичної документації є достатньо громіздкі і вимагають здійснення достатньо великих обчислювальних дій. Тому їхнє застосування на практиці вимагає побудови відповідної системи комп'ютерно-лінгвістичного формування туристичної документації. Основними функціями цієї системи є не тільки збір, обробка та збереження інформації у туристичній документації, але і постійний моніторинг та аналіз якості туристичної документації. Це пов'язано з достатньо динамічними змінами у вимогах до туристичної інформації, яка необхідна для надання якісних туристичних послуг.

В основі методу оцінювання якості туристичної документації покладено показники стандарту ISO/IEC-25010, які задовольняють вимоги до оцінювання якості туристичної документації. Таким чином, система комп'ютерно-лінгвістичного формування туристичної документації направлена на забезпечення належного рівня цих показників якості. Перевірку функціонування такої системи на практиці здійснено в туристичному агентстві «Королівські подорожі» та туристичній компанії «Тур де Львів».

Основні результати розділу опубліковані автором у працях [12, 39, 59, 117, 165, 168].

### **4.1. Розробка показників якості туристичної документації**

ТД як ІІІ впливає на впливає на якість надання туристичних послуг. Тому основним завданням виробника туристичного продукту є створення відповідних умов для введення в обіг такої документації, яка б в першу чергу задовольняла потреби споживачів, відповідала вимогам суспільства, стандартам

та технічним умовам. Такий підхід забезпечує конкурентоспроможність туристичних установ, передумовою якої є якісна ТД.

Якість ТД визначається за набором показників стандарту ISO/IEC-25010, які задовольняють метод оцінювання якості ТД.

#### 4.1.1. Розробка функціональних показників якості туристичної документації

Серед функціональних показників якості вимоги щодо оцінювання якості ТД задовольняє під-характеристика Відповідність (Appropriateness) – міра забезпечення ПЗ тих функцій, які специфіковані для цілей задач і користувачів.

Відповідність будемо визначати за функціональною повнотою так:

$$Quality^{(FS)} = 1 - \frac{A^{(FS)}}{B^{(FS)}}, \quad (4.1)$$

де  $A^{(FS)}$  – кількість функцій з виявленими проблемами,  $B^{(FS)}$  – загальна кількість функцій.

Виберемо в якості функцій наявність відповідних структурних елементів. Тобто якість ТД за функціональною відповідністю будемо оцінювати окремо за наявністю ТО, ДТО, класифікаторів і термінів класифікаторів. Тоді якщо  $Element^{(Requirement\ t)}$  – множина одного з структурних елементів ТД за вимогами,  $|Element^{(Requirement\ t)}|$  – кількість елементів цієї множини,  $Element^{(Gap)}$  – прогалина в структурних елементах, то за (3.16)

- $A^{(FS)} = |Element^{(Gap)}|$ ,  $B^{(FS)} = |Element^{(Requirement\ t)}|$ .

Таким чином відповідність за функціональною повнотою показує рівень присутності структурних прогалин в ТД за ТО, ДТО, класифікаторами і термінами класифікаторів.

#### 4.1.2. Розробка показників якості придатності до використання туристичної документації

Вимоги щодо оцінювання якості ТД задовольняють такі під-характеристики придатності до використання:

- відповідність розпізнавання (Appropriateness recognisability);
- рівень допомоги (Helpfulness);
- привабливість (Attractiveness).

Відповідність розпізнавання – це міра можливостей користувача розпізнати чи ПЗ підходить до його потреб.

Відповідність розпізнавання будемо визначати за повнотою описів так:

$$Quality^{(O.AR)} = \frac{A^{(O.AR)}}{B^{(O.AR)}}, \quad (4.2)$$

де  $A^{(O.AR)}$  – кількість зрозуміло описаних функцій,  $B^{(O.AR)}$  – загальна кількість функцій.

Виберемо в якості функцій наявність фактів чи подій окремо для відповідних структурних елементів ТД (ТО, ДТО та термінів класифікаторів). Тоді якщо  $Element^{(Present)}$  – множина наявних структурних елементів ТД,  $Element^{(DefinitionGap)}$  – множина структурних елементів ТД, що містять прогалини вмісту, визначені за (3.17),  $Element^{(ContentGap)}$  – множина структурних елементів ТД, що містять прогалини вмісту, визначені за (3.18) або (3.20), або (3.19) або (3.21), то

- $A^{(O.AR)} = |Element^{(DefinitionGap)}|$ ,  $B^{(O.AR)} = |Element^{(Present)}|$  – для термінів класифікаторів, з якими пов'язуються факти визначення ТО і ДТО;
- $A^{(O.AR)} = |Element^{(ContentGap)}|$ ,  $B^{(O.AR)} = |Element^{(Present)}|$  – для усіх можливих структурних елементів ТД, з якими пов'язуються факти і події ТО і ДТО.



Таким чином придатність до використання за відповідністю розпізнавання показує рівень присутності прогалин вмісту в ТД за ТО, ДТО, термінами класифікаторів.

Також цей показник якості модифікуємо для відображення прогалин вмісту за достовірністю та актуальністю фактів і подій (факти визначення тут не розглядаються). Для цього позначимо  $Element^{(TrustGap)}$  – множина елементів, що містять факти чи події, автори яких мають рівень довіри нижчий ніж допустимий за (3.5), і  $Element^{(ActualityGap)}$  – множина елементів, що містять неактуальні факти чи події за (3.34). Тоді якість ТД за повнотою достовірних і актуальних описів визначається за (4.2), якщо:

- $A^{(O.AR)} = |Element^{(TrustGap)}|$ ,  $B^{(O.AR)} = |Element^{(Present)}|$  – для достовірності;
- $A^{(O.AR)} = |Element^{(ActualityGap)}|$ ,  $B^{(O.AR)} = |Element^{(Present)}|$  – для актуальності.

Зауважимо, що достовірність змінюється в часі, оскільки довіра до автора допису, на основі якого згенеровано факт чи подію в ТД, може змінюватися залежно від його поведінки в часі.

Рівень допомоги – це міра забезпеченості користувача допомогою.

За доступністю допомоги рівень допомоги визначається так:

$$Quality^{(O.H)} = \frac{A^{(O.H)}}{B^{(O.H)}}, \quad (4.3)$$

де  $A^{(O.H)}$  – кількість задач з наявною описаною допомогою,  $B^{(O.H)}$  – загальна кількість тестованих задач.

Виберемо в якості задач, що тестуються, використання інформації про ТО або ДТО. Вважаємо, що ТО або ДТО пройшли тестування, якщо в ТД містяться відгуки споживачів до цих ТО або ДТО. Тоді якщо  $Element^{(Experience)}$  – множина ТО або ДТО, до яких є відгуки споживачів,  $Element^{(Experience Content)}$  – ТО або ДТО, до яких є відгуки споживачів, і які не мають прогалин вмісту, визначені за (3.18) або (3.20), (3.19) або (3.21), то

$$\bullet A^{(O.H)} = |Element^{(Experience\ Content)}|, B^{(O.H)} = |Element^{(Experience)}|.$$

Таким чином придатність до використання за рівнем допомоги показує частку ТО або ДТО, які мають вміст у вигляді фактів і подій, і які викликають інтерес у споживачів ТП, що відображається їхніми відгуками.

Привабливість – це рівень впливу (вражень) ПЗ на користувача. Привабливість визначається так:

$$Quality^{(O.A)} = \frac{A^{(O.A)}}{B^{(O.A)}}, \quad (4.4)$$

де  $A^{(O.A)}$  – кількість елементів інтерфейсу, пристосованих до задоволення потреб користувача,  $B^{(O.A)}$  – загальна кількість елементів інтерфейсу.

Під елементами інтерфейсу будемо розуміти ТО і ДТО, які є в описовій частині ТД, а пристосованими до задоволення потреб користувача будемо вважати ті ТО і ДТО, до яких є відгуки споживачів. Тобто якщо  $Element^{(Present)}$  – множина наявних ТО або ДТО в ТД, а  $Element^{(Experience)}$  – множина ТО або ДТО, до яких є відгуки споживачів, то:

$$\bullet A^{(O.A)} = |Element^{(Experience)}|, B^{(O.A)} = |Element^{(Present)}|.$$

Таким чином придатність до використання за привабливістю показує частку ТО або ДТО, які викликали інтерес в споживачів ТП, що відображено у їхніх відгуках.

#### **4.1.3. Розробка показників якості супроводжуваності туристичної документації**

Вимоги щодо оцінювання якості ТД задовольняють такі підхарактеристики супроводжуваності:

- модульність (Modularity);
- стійкість до змін (Modification stability).

Модульність показує на скільки ПЗ є скомпоноване з окремих достатньо незалежних компонент. Модульність будемо визначати так:

$$Quality^{(M.M)} = \frac{A^{(M.M)}}{B^{(M.M)}}, \quad (4.5)$$

де  $A^{(M.M)}$  – кількість компонент, які правильно структуровані,  $B^{(O.A)}$  – загальна кількість компонент.

Модульність ТД впливає на сприйняття та швидкість пошуку інформації про ТО та ДТО в ТП і відгуках споживачів. Використаємо таку особливість В- або В<sup>+</sup>-дерев [125] для визначення якості ТД за модульністю: кількість нащадків будь-якого внутрішнього (некореневого) вузла знаходиться в межах від  $t$  до  $2t$ , де  $t$  – фіксоване число. Нехай  $N_{\min}^{(Element)}$  – мінімальна кількість ТО або ДТО в ТП або відгуках споживачів. Тоді для модульності ТД, яка визначається за (4.5):

- $A^{(M.M)}$  – кількість ТП або відгуків споживачів, які містять від  $N_{\min}^{(Element)}$  до  $2N_{\min}^{(Element)}$  ТО і ДТО;
- $B^{(M.M)}$  – загальна кількість ТП або відгуків споживачів.

Супроводжуваність ТД за модульністю показує рівень структурованості ТО і ДТО у вигляді ТП або відгуків споживачів. Зауважимо, що якщо  $N_{\min}^{(Element)} = 0$ , то вважається  $A^{(M.M)} = 0$ , тобто якість ТД також дорівнює 0.

Стійкість до змін показує рівень можливостей ПЗ уникати неочікуваного ефекту від модифікацій цього ПЗ. За рівнем успішних змін стійкість визначається так:

$$Quality^{(M.MS)} = \frac{N_2^{(M.M)} / T_2^{(M.M)}}{N_1^{(M.M)} / T_1^{(M.M)}}, \quad (4.6)$$

де  $N_1^{(M.M)}$  – кількість недоліків, виявлених до зміни ПЗ,  $N_2^{(M.M)}$  – кількість недоліків, виявлених після зміни ПЗ,  $T_1^{(M.M)}$  – інтервал часу, протягом якого виявлялися недоліки до зміни ПЗ,  $T_2^{(M.M)}$  – інтервал часу, протягом якого виявлялися недоліки після зміни ПЗ.

Якщо інтервали часу до зміни і після ПЗ рівні, тобто  $T_1^{(M.M)} = T_2^{(M.M)}$ , то (4.6) спрощується до такого вигляду:

$$Quality^{(M.MS)} = \frac{N_2^{(M.M)}}{N_1^{(M.M)}}. \quad (4.7)$$

Для ТД недоліками є усі невідповідності, які зустрічаються в попередніх показниках якості, а саме:

- структурні прогалини – для відповідності ТД за функціональною повнотою;
- структурні елементи з прогалинами вмісту, з недостовірним і неактуальним вмістом – для придатності ТД до використання за відповідністю розпізнавання;
- ТО або ДТО, до яких є відгуки споживачів, і які не мають прогалин вмісту – для придатності ТД до використання за рівнем допомоги;
- ТО або ДТО, до яких нема відгуків споживачів – для придатності ТД до використання за привабливістю;
- ТП або відгуки споживачів, які містять занадто багато (більше за  $2N_{\min}^{(Element)}$ ) ТО і ДТО – для супроводжуваності ТД за модульністю.

Таким чином супроводжуваність ТД за стійкістю до змін показує як усуваються недоліки ТД після проведення відповідних процедур зміни структури і наповнення.

## **4.2. Розробка системи комп'ютерно-лінгвістичного формування туристичної документації**

Розроблені процедури вибору відкритих Веб-ресурсів, співставлення вмісту відкритих Веб-ресурсів з елементами туристичної документації, виявлення некоректної інформації у туристичній документації, а також сформовані показники якості туристичної документації для практичної реалізації потребують створення спеціалізованого ПЗ та застосування

необхідних інформаційних технологій. Фактично формування якісної ТД на основі ВВР неможливе без побудови відповідної інформаційної системи, яка забезпечує ефективне керування інформацією для потреб туристичної установи та для оперативного задоволення запитів користувачів [9, 64].

#### 4.2.1. Розробка технології автоматизованого формування туристичної документації

Враховуючи етапи формування ТД (рис. 3.1) на рис. 4.1 подано узагальнену технологію її формування на основі відкритих веб-ресурсів.

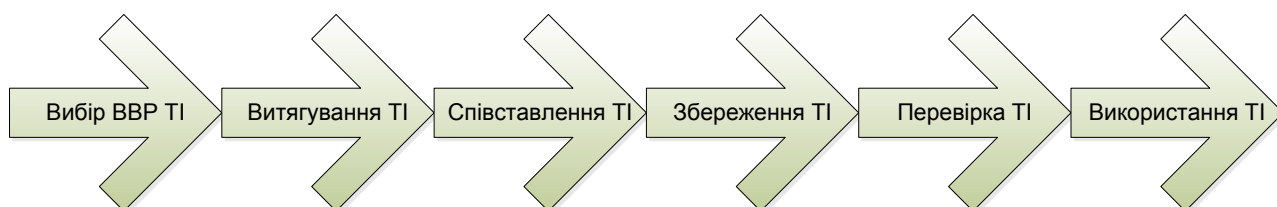


Рис. 4.1. Технологія формування туристичної документації на основі відкритих веб-ресурсів

Вибір ВВР для формування ТД здійснюється на основі таких розроблених критеріїв, як:

- авторитетність відкритих веб-ресурсів;
- відповідність до тематики;
- достовірність дописів;
- актуальність дописів.

Витягування ТІ з вибраних ВВР проводиться автоматизовано, в результаті чого отримані веб-сторінки розкладаються на:

- інформаційні блоки;
- дописи окремих авторів;
- атомарні ситуації.

Для співставлення ТІ виконуються автоматизовані процедури порівняння з елементами ТД, результатом яких є поява нових фактів і подій, пов'язаних з

ТО, ДТО та термінами класифікаторів. Нові факти і події зберігаються в ТД і є доступними для використання.

Перевірка ТІ, що зберігається в ТД, проводиться на основі розроблених процедур виявлення структурних прогалин і прогалин вмісту в ТД, частотного аналізу маркерів та індикативних ознак, процедур виявлення некоректної і неактуальної інформації в ТД. Результати перевірки ТІ відображаються у зміні таких показників якості ТД:

- функціональні показники якості ТД;
- показники якості придатності до використання ТД;
- показники якості супроводжуваності ТД.

Сформована ТД допускає такі основні напрями використання:

- аналіз ТІ надавачем ТП;
- створення веб-ресурсу для споживачів ТП;
- ІР для інших систем та сервісів.

#### **4.2.2. Розробка архітектури системи автоматизованого формування туристичної документації**

Розроблені процедури формування ТД та технологія автоматизованого формування ТД дозволяють спроектувати комплексну ІС автоматизованого формування ТД. Така комплексна ІС повинна забезпечувати формування якісної ТД на основі ВВР відповідно до розроблених показників якості.

Особливістю ІС автоматизованого формування ТД є її функціонування на трьох архітектурних рівнях (рис. 4.2):

- 1) рівень локальної ІС.
- 2) рівень агентів формування ТД.
- 3) рівень зовнішніх інформаційних сервісів та ресурсів.

На першому рівні здійснюється робота з ТД в межах локальної ІС, яка забезпечує:

- формування структурних елементів ТД;

- визначення індикативних ознак;
- моніторинг якості ТД;
- задання параметрів вибору ВВР, веб-сторінок та дописів.

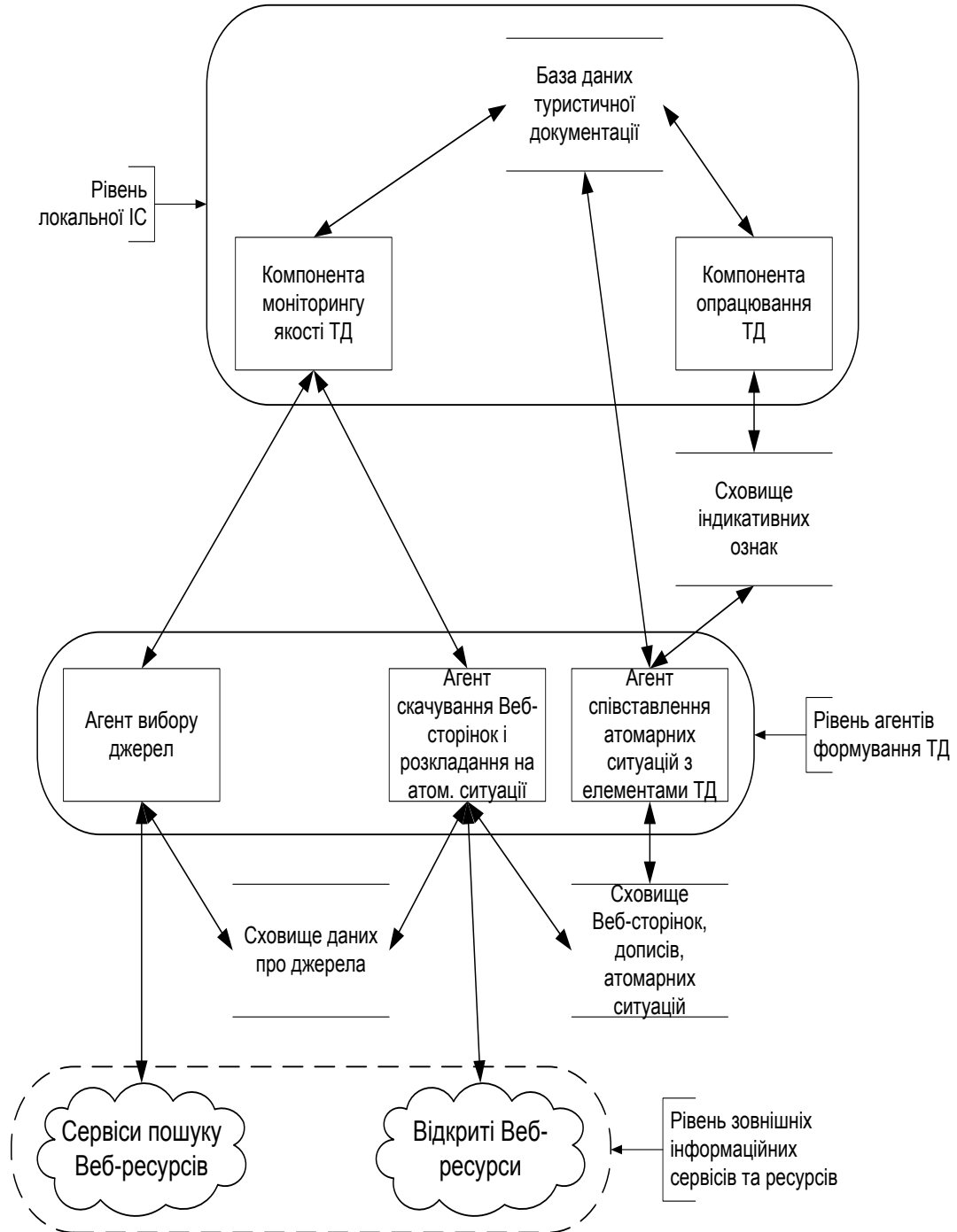


Рис. 4.2. Архітектура комплексної інформаційної системи автоматизованого формування туристичної документації

На другому рівні функціонують програмні агенти формування ТД, які призначені для:

- вибору ВВР;
- скачування веб-сторінок з ВВР;
- розкладання веб-сторінок на дописи та атомарні ситуації;
- співставлення атомарних ситуацій з елементами ТД;
- збереження співставлених атомарних ситуацій.

На третьому рівні для формування ТД використовуються зовнішні інформаційні сервіси та ресурсів:

- сервіси пошуку веб-ресурсів (пошукові машини);
- ВВР, які виступають джерелами інформації для ТД.

#### **4.2.3. Розробка функціональної структури системи автоматизованого формування туристичної документації**

На основі побудованої архітектура комплексної ІС автоматизованого формування ТД (рис. 4.2) розробимо функціональну структуру цієї системи [16].

На рис. 4.3 подано контекстну діаграму комплексної ІС автоматизованого формування ТД.

На рис. 4.4 подано діаграму потоків даних комплексної ІС автоматизованого формування ТД, в якій виділено такі основні функції:

- агент вибору джерел;
- агент скачування веб-сторінок;
- агент співставлення атомарних ситуацій;
- опрацювання ТД;
- моніторинг якості ТД.

Відповідно до процедури вибору джерела для формування ТД Агент вибору джерел складається з таких функцій (рис. 4.5):

- формування ключових слів;



- перевірка на відповідність джерела;
- визначення корисних веб-сторінок.

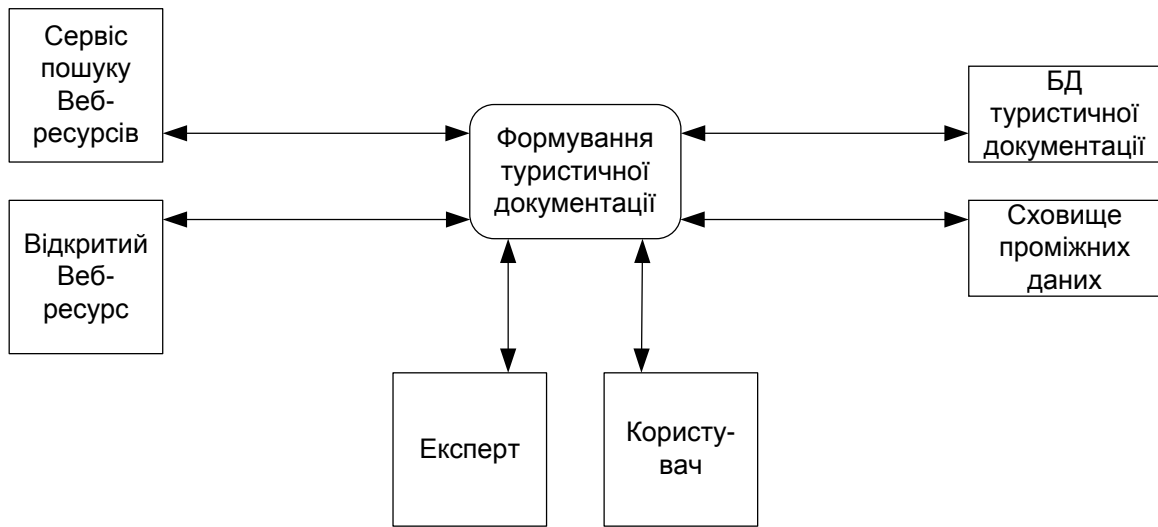


Рис. 4.3. Контекстна діаграма комплексної інформаційної системи автоматизованого формування туристичної документації

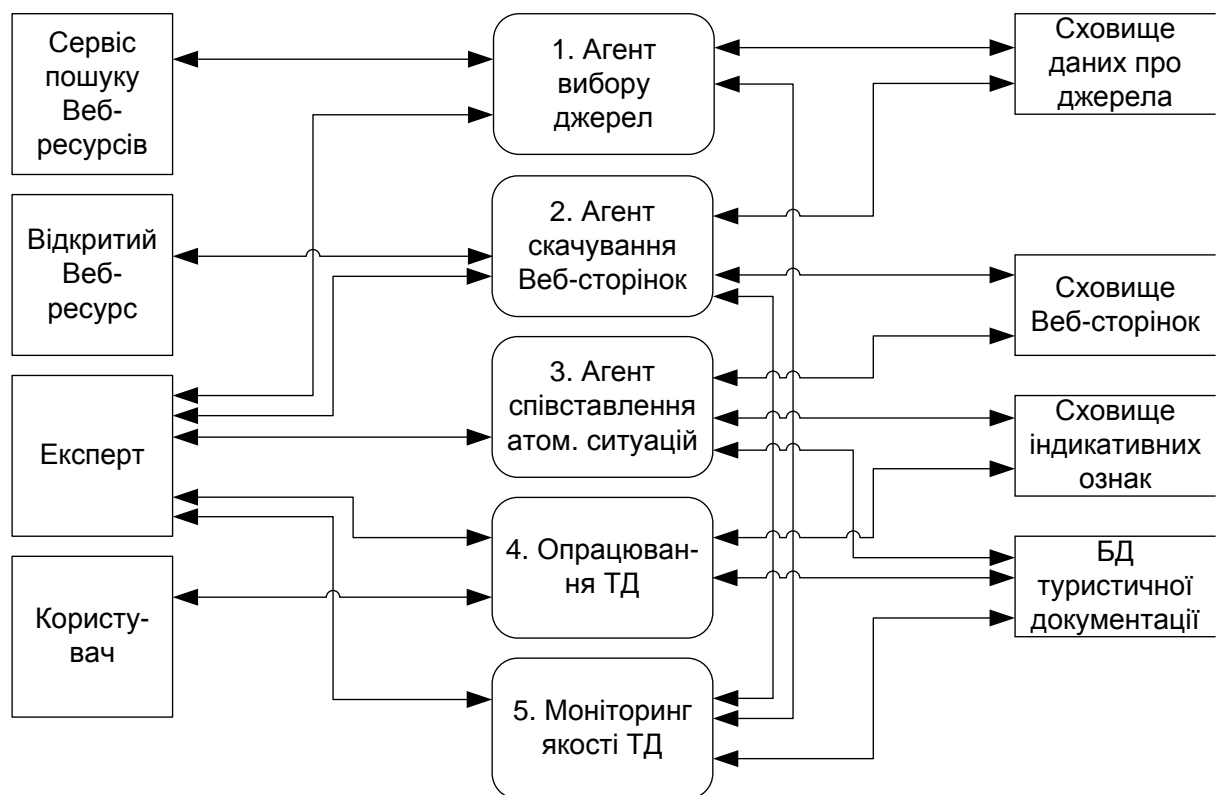


Рис. 4.4. Діаграма потоків даних комплексної інформаційної системи автоматизованого формування туристичної документації

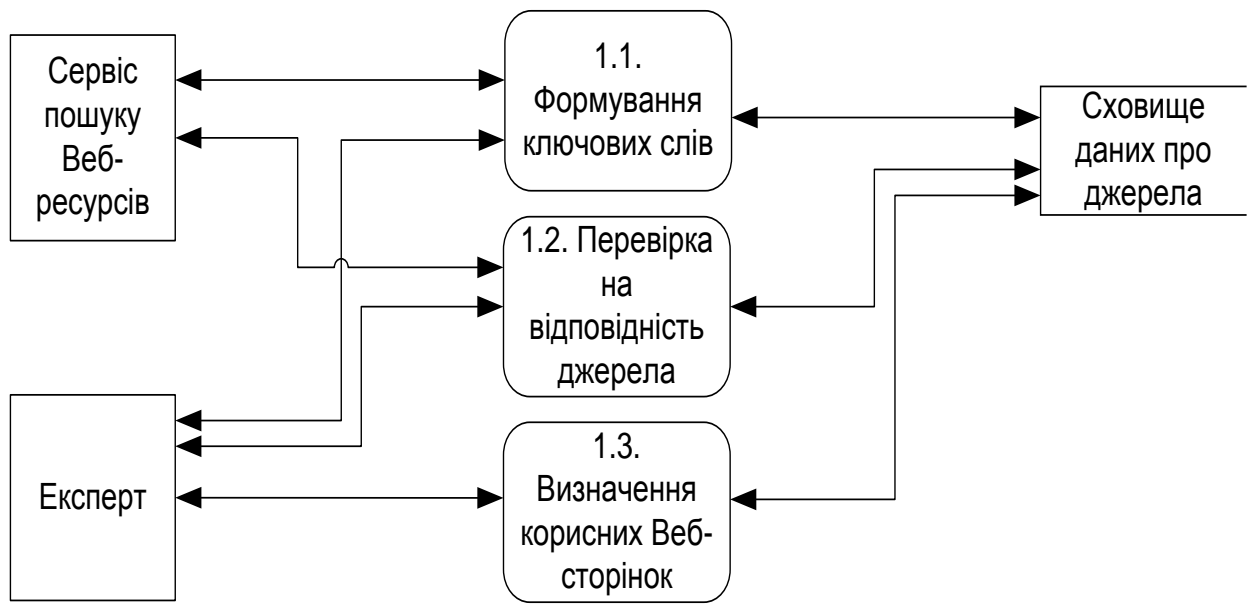


Рис. 4.5. Діаграма потоків даних Агента вибору джерел

З урахуванням процедури розкладання веб-сторінок на окремі дописи Агент скачування веб-сторінок включає такі функції (рис. 4.6):

- скачування веб-сторінки;
- пошук інформаційних блоків веб-сторінки;
- пошук дописів Веб-сторінки.

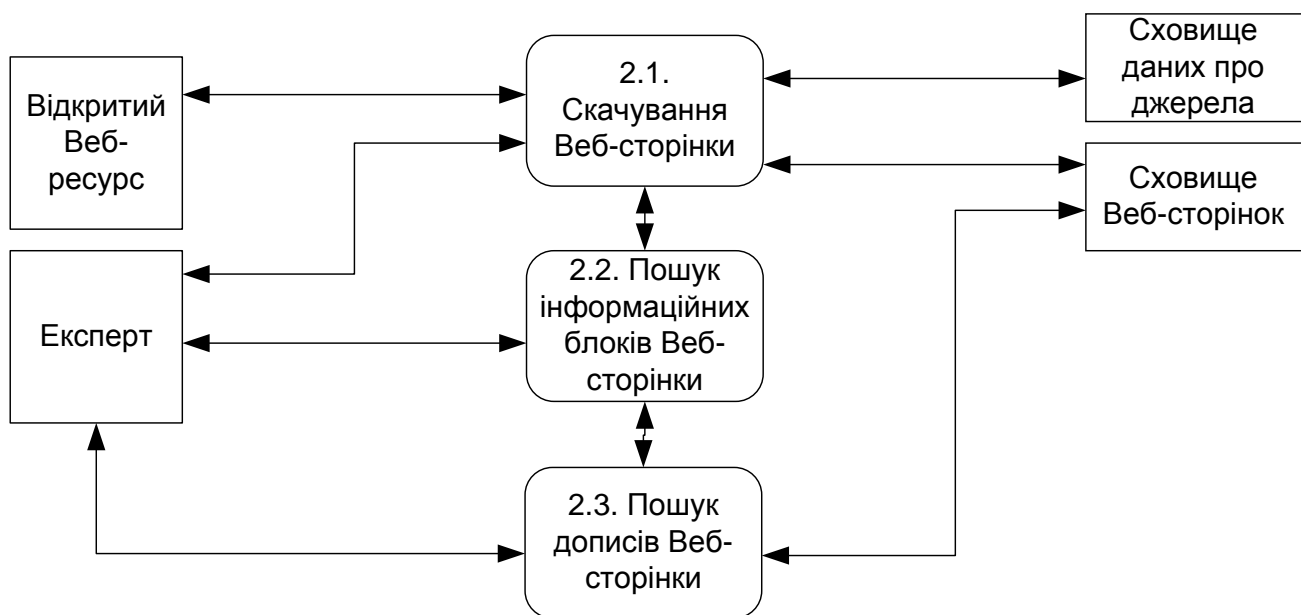


Рис. 4.6. Діаграма потоків даних Агента скачування веб-сторінок

Агент співставлення атомарних ситуацій діє відповідно до процедури розкладання дописів веб-сторінок на атомарні ситуації та процедури лінгвістичного співставлення атомарних ситуацій дописів веб-сторінок з елементами ТД. Він містить такі функції (рис. 4.7):

- розкладання дописів на атомарні ситуації;
- співставлення атомарних ситуацій з елементами ТД;
- збереження атомарної ситуації в ТД.

Опрацювання ТД складається з таких функцій (рис. 4.8):

- налаштування індикативних ознак;
- модерація ТД;
- перегляд ТД;
- експорт ТД.

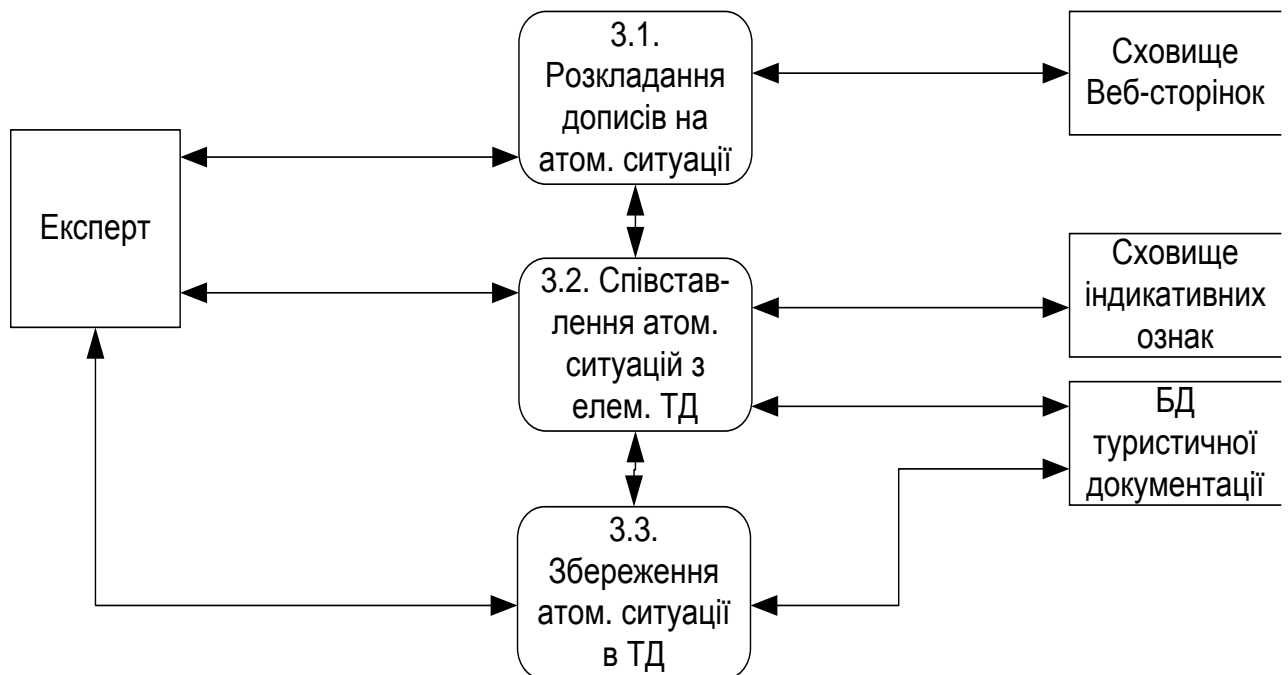


Рис. 4.7. Діаграма потоків даних Агент співставлення атомарних ситуацій

Моніторинг якості ТД функціонує на основі процедур виявлення прогалин у туристичній документації і процедур виявлення некоректної інформації з урахуванням розроблених показників якості ТД (рис. 4.9).

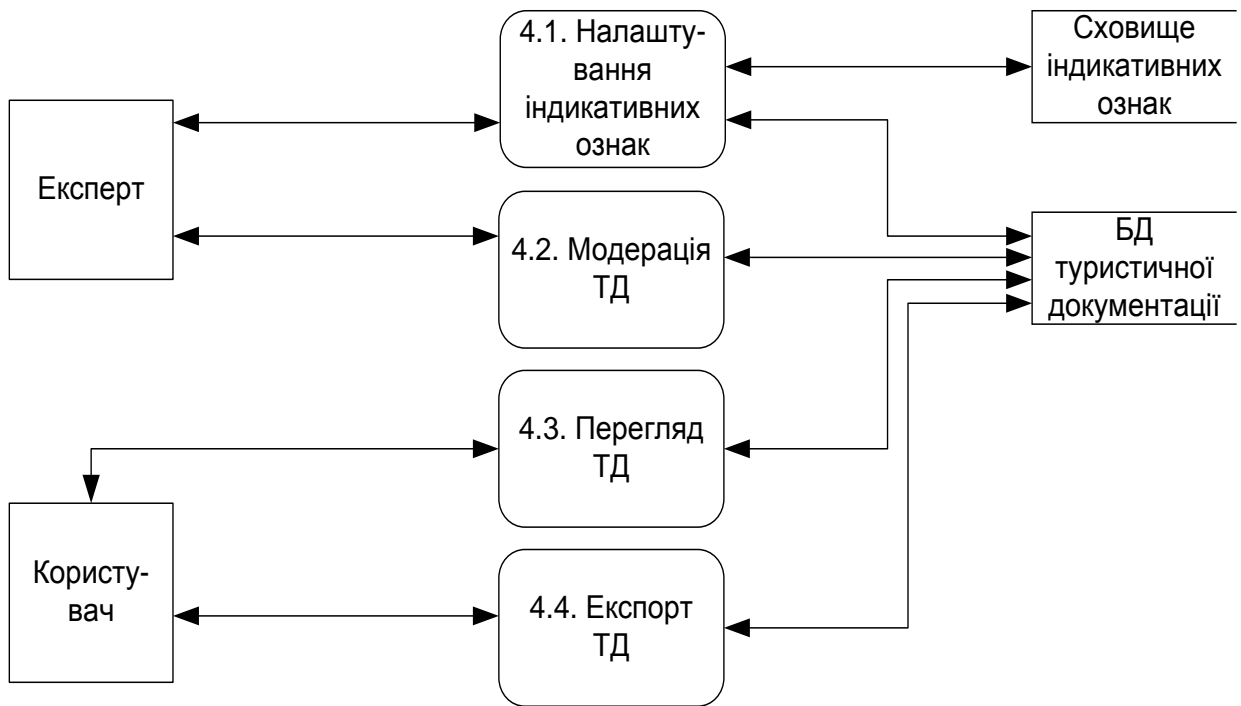


Рис. 4.8. Діаграма потоків даних Опрацювання туристичної документації

Моніторинг якості ТД включає такі функції:

- аналіз структурних прогалин ТД;
- аналіз прогалин вмісту ТД;
- аналіз достовірності ТД;
- аналіз актуальності ТД;
- аналіз рівня допомоги ТД;
- аналіз привабливості ТД;
- аналіз модульності ТД;
- аналіз усунення недоліків ТД.

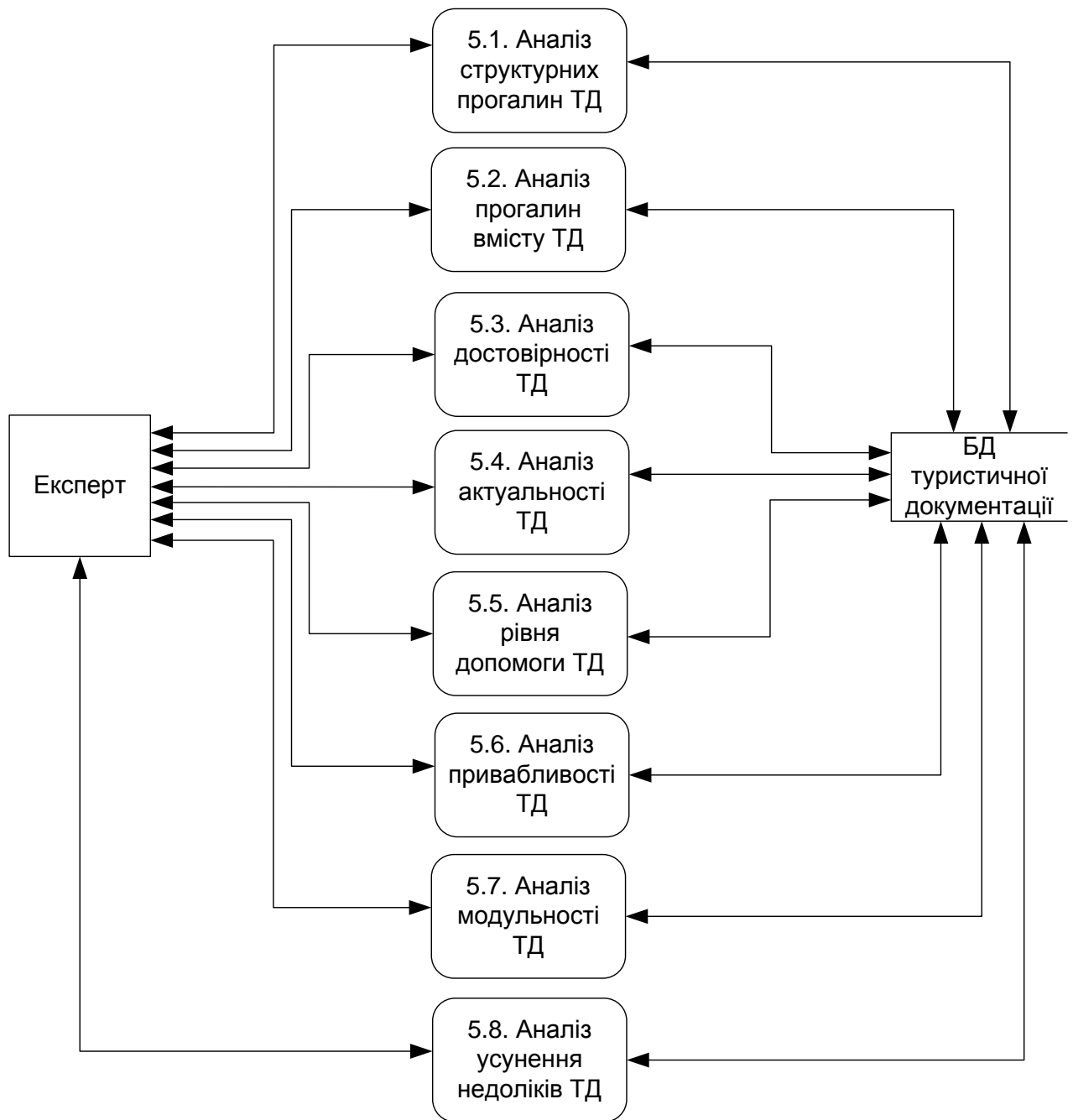


Рис. 4.9. Діаграма потоків даних Моніторингу якості туристичної документації

### 4.3. Аналіз практичного впровадження результатів досліджень

#### 4.3.1. Аналіз впровадження результатів досліджень у туристичному агентстві «Королівські подорожі»

«Королівські Подорожі» – туристична компанія, яка надає послуги з організації туристичних подорожей та відпочинку. Формування ТД цієї компанії за допомогою комп'ютерно-лінгвістичних процедур розпочато у вересні 2014 року і тривало 25 місяців. На кінець дослідного терміну ТД містила:

- 235 туристичних послуг (динаміку зміни кількості ТП подано на рис. 4.10);
- 1545 туристичних об'єктів (динаміку зміни кількості ТО подано на рис. 4.11).

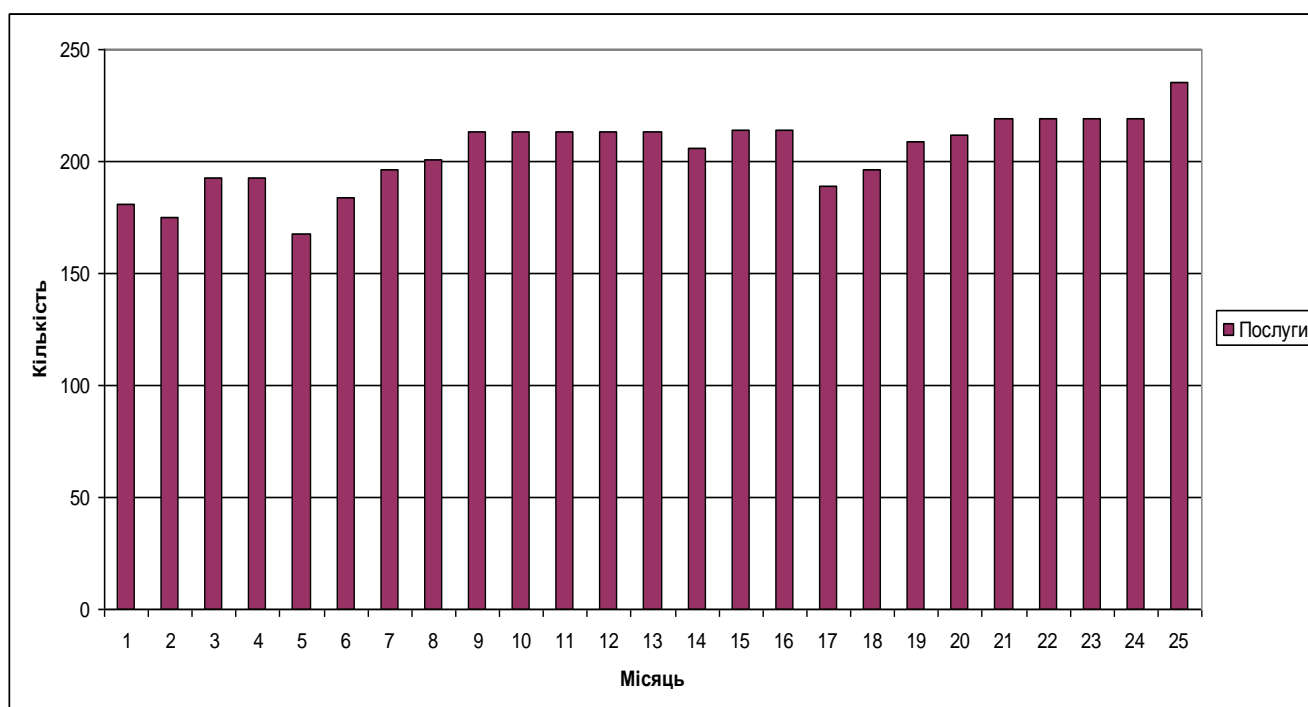


Рис. 4.10. Динаміка зміни кількості послуг в туристичній документації агентства «Королівські подорожі»

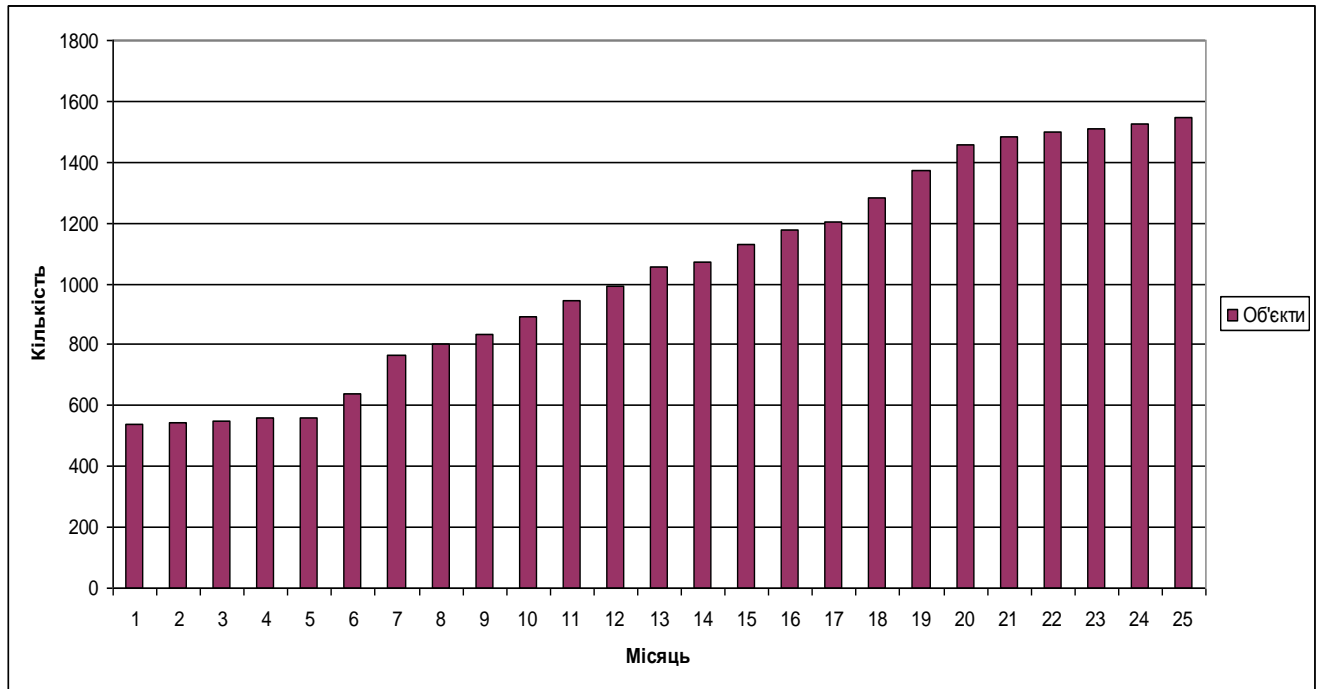


Рис. 4.11. Динаміка зміни кількості об'єктів в туристичній документації агентства «Королівські подорожі»

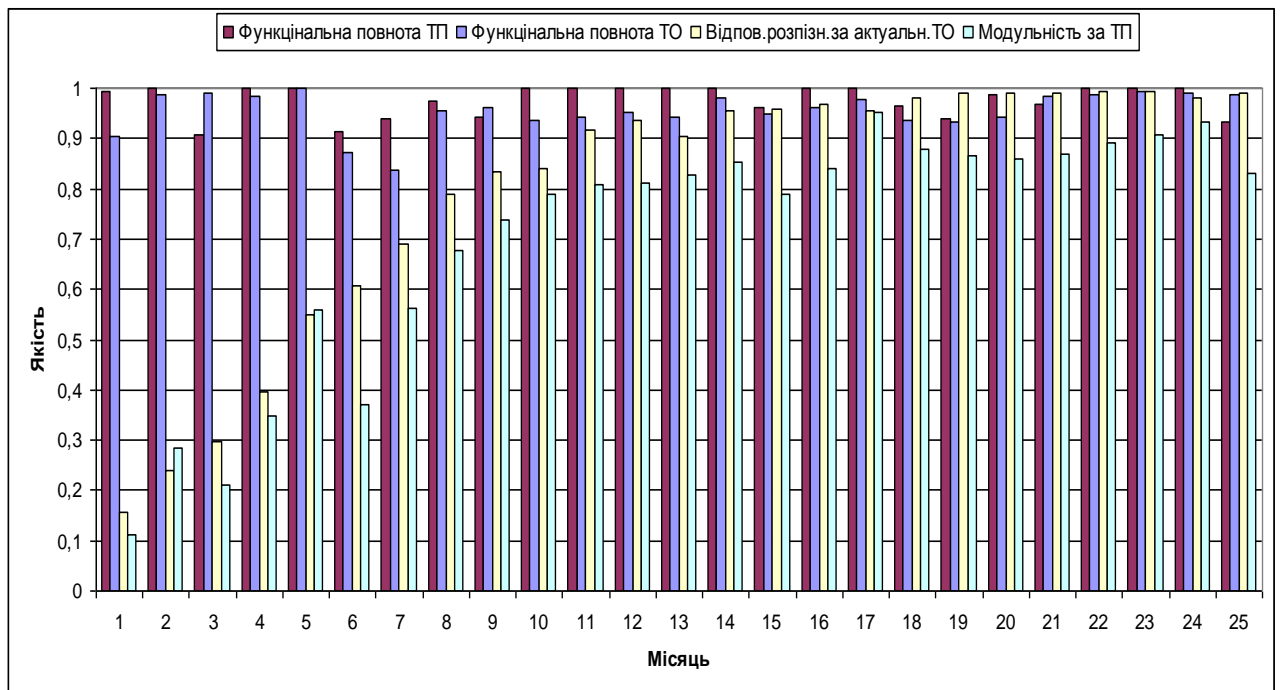


Рис. 4.12. Динаміка зміни показників якості туристичної документації агентства «Королівські подорожі»

Моніторинг якості ТД здійснювався за такими показниками (рис. 4.12):

- функціональна повнота ТП;
  - функціональна повнота ТО;
  - відповідність розпізнавання за актуальністю повноти описів ТО;
- модульність за ТП.

Рис. 4.12 показує, що в агентстві «Королівські подорожі» після першого року формування і підтримки ТД в актуальному стані вдалося повністю стабілізувати рівень якості ТД на достатньо високому рівні.

#### **4.3.2. Аналіз впровадження результатів досліджень в туристичній компанії «Тур де Львів»**

Туристична компанія «Тур де Львів» надає послуги з організації екскурсій у місті Львові, в Україні та за кордоном. Формування ТД цієї компанії за допомогою комп'ютерно-лінгвістичних процедур розпочато у липні 2015 року і тривало 15 місяців. На кінець дослідного терміну ТД містила:

- 210 туристичних послуг (динаміку зміни кількості ТП подано на рис. 4.13);
- 210 відгуків споживачів (динаміку зміни кількості відгуків подано на рис. 4.14);
- 1248 туристичних об'єктів (динаміку зміни кількості ТО подано на рис. 4.15);
- 2053 дії над туристичними об'єктами (динаміку зміни кількості ДТО подано на рис. 4.16).

Моніторинг якості ТД здійснювався за такими показниками (рис. 4.12):

- функціональна повнота ТП і відгуків споживачів;
- функціональна повнота ТО і ДТО;
- відповідність розпізнавання за актуальністю повноти описів ТО і ДТО;
- рівень допомоги за доступністю ТО і ДТО;
- привабливість ТО і ДТО;
- модульність за ТП.



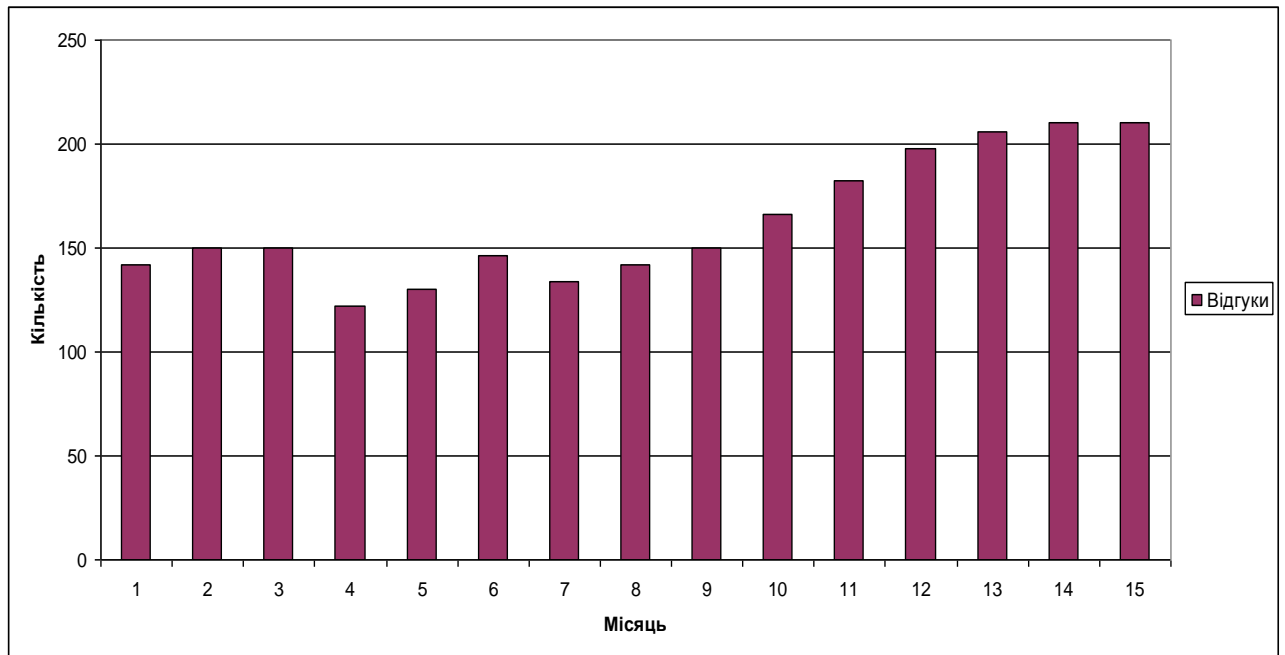


Рис. 4.13. Динаміка зміни кількості послуг в туристичній документації туристичної компанії «Тур де Львів»

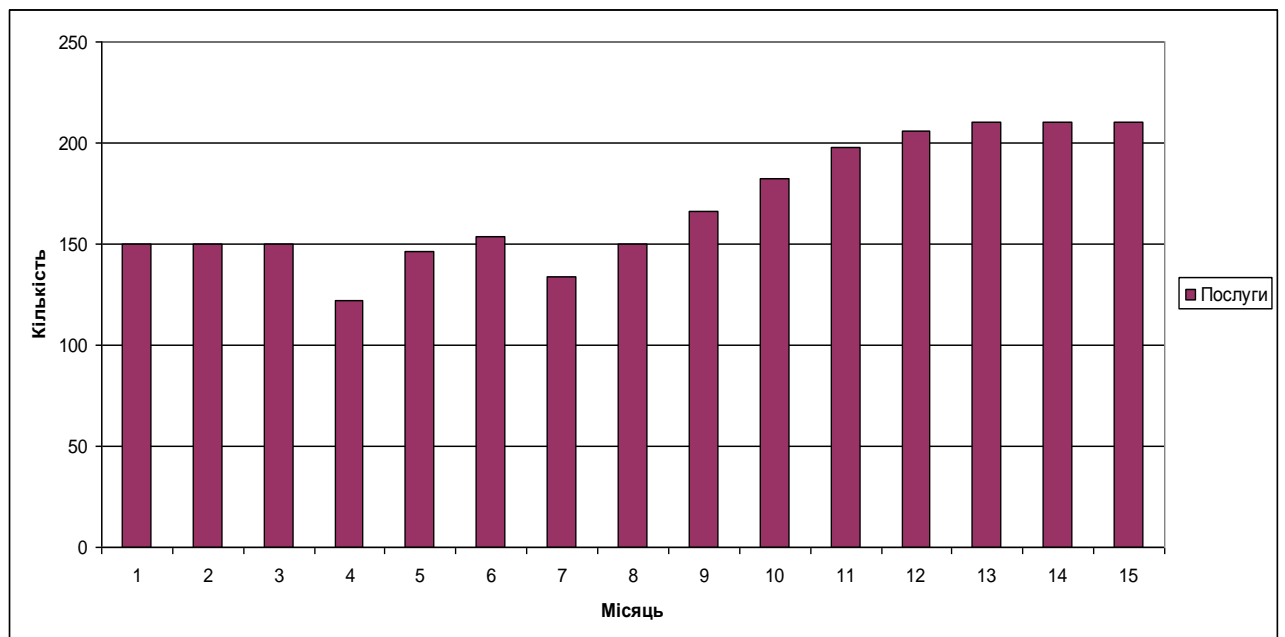


Рис. 4.14. Динаміка зміни кількості відгуків споживачів у туристичній документації туристичної компанії «Тур де Львів»

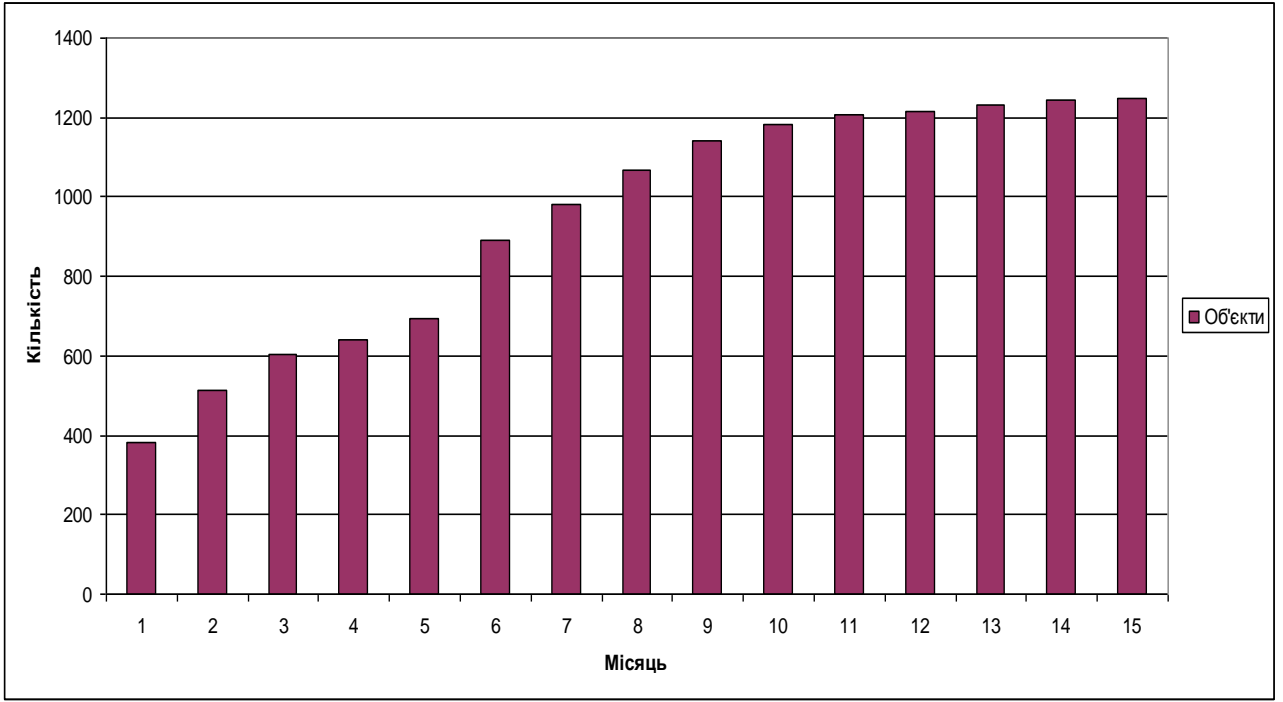


Рис. 4.15. Динаміка зміни кількості об'єктів в туристичній документації туристичної компанії «Тур де Львів»

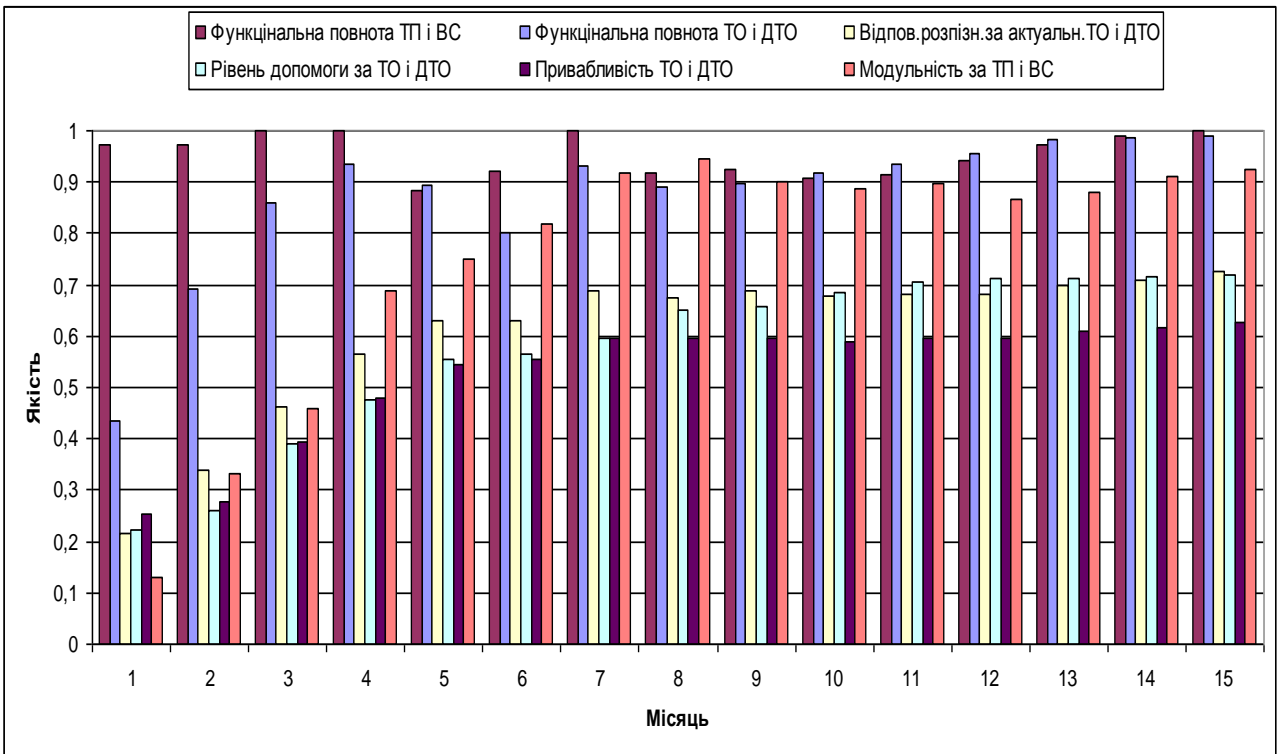


Рис. 4.16. Динаміка зміни кількості дій над об'єктами в туристичній документації туристичної компанії «Тур де Львів»

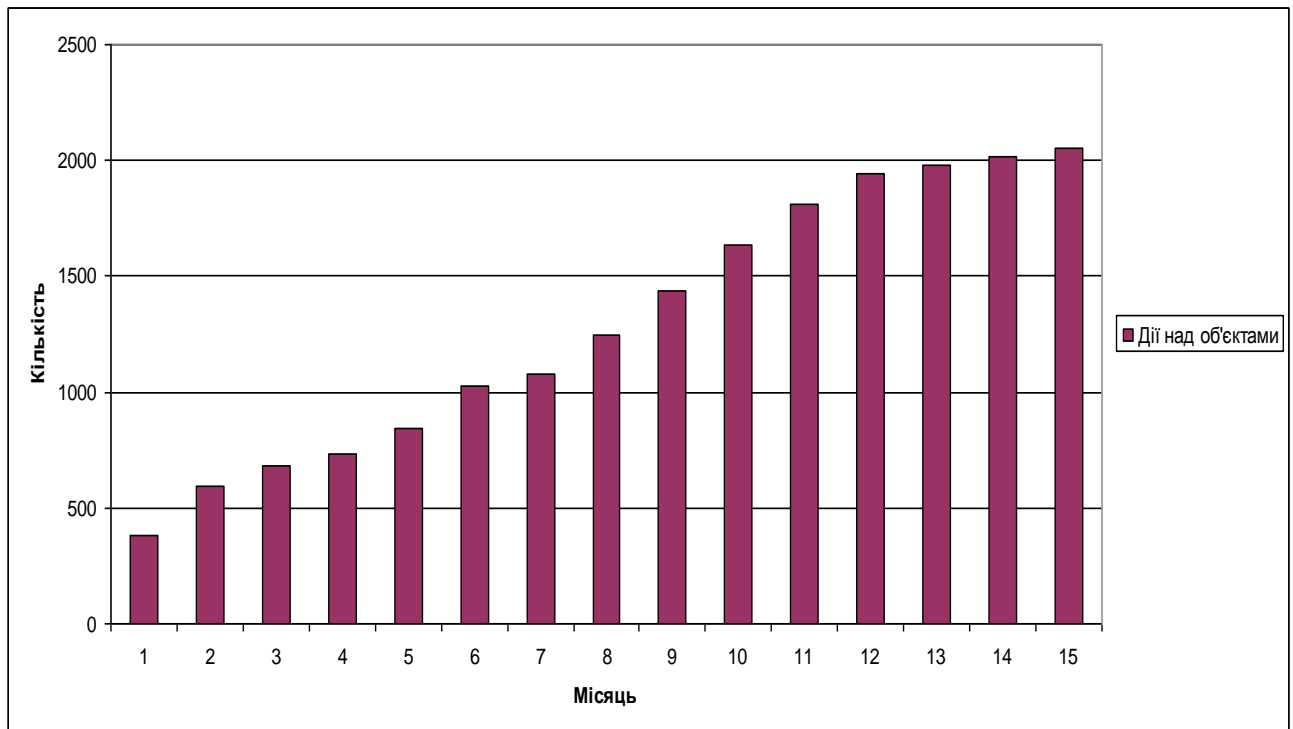


Рис. 4.17. Динаміка зміни показників якості туристичної документації туристичної компанії «Тур де Львів»

Рис. 4.17 показує, що в туристичній компанії «Тур де Львів» після першого року формування і підтримки ТД в актуальному стані вдалося повністю стабілізувати рівень якості ТД на достатньо високому рівні.

#### 4.4. Висновки до розділу

1. Уперше розроблено метод оцінювання якості туристичної документації на основі функціональних показників якості, показників якості придатності до використання та показників якості супроводжуваності стандарту ISO/IEC-25010, що забезпечило практичне оцінювання результатів виконання процедур формування та перевірки туристичної документації.
2. Розроблено архітектуру та функціональну структуру системи комп'ютерно-лінгвістичного формування туристичної документації на основі процедур формування та показників якості туристичної документації, що дає можливість на практиці реалізувати і впровадити цю інформаційну систему.

3. Аналіз впровадження результатів досліджень на практиці показав ефективність розроблених методів і засобів формування туристичної документації за допомогою комп'ютерно-лінгвістичного аналізу інформаційного наповнення відкритих веб-ресурсів, а також показує істотний ріст (в 1,5-3 рази) кількості структурних елементів з можливістю їхньої належної підтримки з показниками якості інформації про них від 0,6 до 1.

## Висновки

У дисертаційній роботі вирішено важливу наукову задачу розроблення методів і засобів формування туристичної документації шляхом комп'ютерно-лінгвістичного аналізу інформаційного наповнення відкритих веб-ресурсів.

Основними науково-практичними результатами є:

1. Визначено основні підходи до формування туристичної документації на основі відкритих веб-ресурсів шляхом аналізу методів структурування та інтеграції інформаційного наповнення веб-ресурсів, що стало основою для моделей туристичної документації та інформаційних джерел.
2. Уперше побудовано формальну модель туристичної документації на основі множин туристичних послуг, відгуків споживачів туристичних послуг та класифікаторів з урахуванням відповідних індикативних ознак їх елементів, що стало основою для побудови процедур комп'ютерно-лінгвістичного формування туристичної документації і показників її якості туристичної документації.
3. Удосконалено формальну модель інформаційного джерела для формування туристичної документації шляхом виокремлення блоків і атомарних ситуацій туристичної інформації на веб-сторінках, що стало основою для побудови процедури вибору відкритих веб-ресурсів, процедур лінгвістичного співставлення вмісту відкритих веб-ресурсів з елементами туристичної документації.
4. Набув подальшого розвитку метод лінгвістичного виявлення інформації у веб-ресурсах шляхом співставлення атомарних ситуацій з елементами туристичної документації на основі модифікованих індикативних ознак, що дозволило побудувати відповідні процедури лінгвістичного співставлення вмісту відкритих веб-ресурсів з елементами туристичної документації.
5. Розроблено процедури комп'ютерно-лінгвістичного формування туристичної документації на основі співставлення вмісту відкритих веб-ресурсів з

елементами за допомогою індикативних ознак, а також процедури виявлення прогалин та некоректної інформації у туристичній документації на основі аналізу її структури та вмісту, які дозволяють покращити показники якості туристичної документації і є необхідними для побудови системи автоматизованого формування туристичної документації.

6. Уперше розроблено метод оцінювання якості туристичної документації на основі функціональних показників якості, показників якості придатності до використання та показників якості супроводжуваності стандарту ISO/IEC-25010, що забезпечило практичне оцінювання результатів виконання процедур формування та перевірки туристичної документації.
7. Розроблено архітектуру та функціональну структуру системи комп'ютерно-лінгвістичного формування туристичної документації на основі процедур формування та показників якості туристичної документації, що на практиці дозволило істотно збільшити (в 1,5-3 рази) кількість структурних елементів з можливістю їхньої належної підтримки з показниками якості інформації про них від 0,6 до 1.

## Література

1. Алексенцев А.І. Автоматизація діловодства/ А.І. Алексенцев. – М.: ЗАТ Бізнес–школа «Інтел–синтез», 2004 р. – 240 с. – ISBN 240–345–136–1
2. Апилат Е.В. Построение регрессионной модели управления качеством услуг в туризме / Е.В. Апилат // Вісник Харків. держ. політехн. ун-ту. – Х., 2000. – № 5. – С. 41–45.
3. Апилат Е.В. Формирование индустрии туризма – основное направление повышения качества обслуживания туристов / Е.В. Апилат // Вісник Харків. держ. екон. ун-ту. – Х., 2000. – № 5. – С. 32–33.
4. Артеменко О.І. Інформаційна технологія моделювання просторового розвитку об'єктів туристичної інфраструктури: Дисертація на здобуття наукового ступеня к.т.н. – Львів: Національний університет "Львівська політехніка", 2013. – 137с.
5. Асеев Г. Управління сучасним документообігом: теорія, структура, методи/ Г. Асеев // Вісник Книжкової палати. – 2004. – № 5. – С. 32 – 36.
6. Баласаян В. Автоматизация делопроизводства и документооборота: введение в проблему / В. Баласаян// Рынок ценных бумаг. – 1998. – № 16. – С. 25–27.
7. Баласаян В. Электронный документооборот – основа эффективного управления современным предприятием/ В. Баласаян// Управление персоналом. – 2002. – № 2. – С. 18–21.
8. Березко О.Л. Методи та засоби персоніфікації інформаційного наповнення у глобальній системі World Wide Web: Дисертація на здобуття наукового ступеня к.т.н. – Львів: Національний університет "Львівська політехніка", 2011. - 196с.
9. Білинський Й.Й. Електронні системи/ Й.Й. Білинський, К.В. Огороднік, М.Й. Юкиш. – Вінниця: ВНТУ, 2011. – 208 с.

10. Блистів Т.В. Теорія і практика туристичних походів: навч. посіб. / Т.В. Блистів, І.І. Горбацьо, О.В. Нужний. – Л.: НВФ“Українські технології”, 2006. – 132 с.
11. Бобылева М.П. Вопросы использования элементов электронного документооборота внутри организации / М.П. Бобылева// Делопроизводство. – 2003. – № 2. – С. 14–18.
12. Бойко Н.І. Приклад використання систем масового обслуговування в організації роботи web-систем / Н.І.Бойко, О.О.Марків // Матеріали V Міжнародної науково-практичної конференції Математика. Інформаційні технології. Освіта. – Луцьк, 2016. – С. 12-14.
13. Бричковский В. И. Создание программного обеспечения управления информационными ресурсами Web портала виртуальной библиотеки/ В.И.Бричковский, А.М.Авраменко // Менеджмент библиотек учреждений образования: Материалы IV Международной конференции. – Витебск. – 2003. – С.13-18.
14. Бричковский В.И.Технология использования метаданных при создании электронных библиотек/ В.И.Бричковский, А.Г.Корсаков, Д.М.Гурский // Менеджмент библиотек учреждений образования: Материалы IV Международной конференции. – Витебск. – 2003. – С. 19-24.
15. Буров Є.В. Опрацювання контекстних даних в інтелектуальному туристичному порталі / Буров Є.В., Королюк Ю.В. / Вісник Нац. ун-ту “Львівська політехніка” “Комп’ютерні науки та інформаційні технології”. – Львів: Вид-во Львівської політехніки, 2010. – № 672. – С. 228–236.
16. Вендров А. М. CASE-технологии. Современные методы и средства проектирования информационных систем. – М: Финансы и статистика, 1998. – 176 с.
17. Виклюк Я.І. Розвиток методів та засобів математичного моделювання об’єктів туристичної галузі: Дисертація на здобуття наукового ступеня



- д.т.н. – Львів: Національний університет "Львівська політехніка", 2011. – 390с.
18. Волохін О.М. Каталогізація цифрових ресурсів Інтернет. Дублінське ядро метаданих: Посібник / О. М. Волохін. – Кіровоград: Відродження, 2003. – 72 с.
  19. Гаврилюк С.П. Зовнішні фактори впливу на конкурентоспроможність туристичних підприємств / С.П. Гаврилюк // Стратегія розвитку туристичної індустрії та громадського харчування: матер. міжнар. наук.-практ. конф. (25–26 жовтня 2000 р., Київ). – К., 2000. – С. 186–189.
  20. Гаврилюк С.П. Проблеми підвищення якості послуг туристичних підприємств в умовах конкуренції / С.П. Гаврилюк // Проблеми якості у громадському харчуванні, готельному господарстві та туризмі: матер. міжнар. наук.-практ. конф., 28–29 жовт. 1998 р. – К.: КДТЕУ. – С. 52–54.
  21. Гавриш С. Система автоматизації діловодства та електронного документообігу / С. Гавриш // Довідник секретаря та офіс-менеджера. – 2007. – № 2. – С. 25–31.
  22. Герасименко В.Г. Основы туристического бизнеса: учеб. пособие / В.Г. Герасименко. – Одесса: Черноморье, 1997. – 160 с.
  23. Готинян В.С. Можливості WEB– та ГІС/ДЗЗ–технологій в туризмі / Готинян В.С., Семененко А.В., Томченко О.В. [Електронний ресурс: <http://www.pryroda.gov.ua/ua/index.php?newsid=945>].
  24. Гриценко Н.І. Сервіс в туризмі: підручн. (Тема 7: Туроператори і турагенти на міжнародному ринку) / Н.І. Гриценко, В.А. Бурдейного, О.Є. Кутузової та ін. – Одеса : Талісман. 2006. – 160 с.
  25. Деминг Е.В. Выход из кризиса / Е.В. Деминг. – Тверь: Альба, 1994. – 497 с.
  26. ДСТУ 2392-94 (Інформація та документація, базові поняття, терміни та визначення). – Держстандарт України, 1994. – [http://normativ.ucoz.org/\\_ld/1/100\\_2392-94.pdf](http://normativ.ucoz.org/_ld/1/100_2392-94.pdf).

27. ДСТУ 2732:2004 (Діловодство й архівна справа, терміни та визначення). – Держспоживстандарт України, 2004. – <http://zakon.sop.com.ua/regulations/10637/478668/>.
28. ДСТУ 2844-94 (Програмні засоби ЕОМ. Забезпечення якості. Терміни та визначення): Національний стандарт України. – К. : Держстандарт України, 1995. – 22 с.
29. ДСТУ 2850-94 (Програмні засоби ЕОМ. Забезпечення якості. Показники та методи оцінювання якості програмного забезпечення): Національний стандарт України. – К. : Держстандарт України, 1996. – 42 с.
30. ДСТУ 3278-95 (Система розроблення та поставлення продукції на виробництво, основні терміни та визначення). – Держстандарт України, 1995.
31. ДСТУ ISO/IEC 12119-2003 (Інформаційні технології. Пакети програм. Тестування і вимоги до якості. ISO/IEC 12119:1994, IDT): Національний стандарт України. – К. : Держспоживстандарт України, 2004. – 26 с.
32. Діденко А.Н. Сучасне діловодство: посібник / А. Н. Діденко. – К.: Либідь, 2010. – 6-те вид., переробл. і доповн. – 480 с.
33. Економічна статистика. Економічна діяльність. Туризм. Туристичні потоки: [Електронний ресурс] // Державна служба статистики України. – [http://www.ukrstat.gov.ua/operativ/operativ2007/tyr/tyr\\_u/potoki2006\\_u.htm](http://www.ukrstat.gov.ua/operativ/operativ2007/tyr/tyr_u/potoki2006_u.htm).
34. Євдокименко В.К. До питання про моделювання розвитку рекреаційно-туристичних атракцій / В.К.Євдокименко, Я.І.Виклюк, В.Ф.Андрієнко // Науковий вісник Буковинського державного фінансово-економічного університету: Економічні науки. – Чернівці, 2013. – № 2. - С.241-256.
35. Жадан О.В. Основи управління якістю: [навч.-метод. посіб.] / О.В. Жадан, А.В. Крєтова, Г.М. Сичов. – Донецьк: «АПЕКС», 2004. – 99 с.
36. Жежнич П. Аналіз сучасних засобів автоматизації роботи з електронною документацією / П.Жежнич, О.Сопрунюк, О.Марчик // Комп'ютерні науки та інформаційні технології: Вісник Національного університету “Львівська політехніка”. – Львів, 2012. – №732. - С.395-398.

37. Жежнич П. Консолідація відкритих інформаційних ресурсів в туристичній сфері / П.Жежнич, О.Сопрунюк // Комп'ютерні науки та інформаційні технології: Вісник Національного університету "Львівська політехніка", - Львів, 2013. – №771. – С.3-11.
38. Жежнич П. І. Консолідовані інформаційні ресурси баз даних та знань: навчальний посібник/ П.І. Жежнич. – Львів: видавництво Національного університету "Львівська політехніка", 2010. –212 с.
39. Жежнич П. Основні підходи до моніторингу та аналізу якості туристичної документації / П.Жежнич, О.Сопрунюк // Інформаційні системи та мережі: Вісник Національного університету "Львівська політехніка". - Львів, 2012. - №743. – С.190-199.
40. Жежнич П. Основні підходи до організації електронного організаційно-розпорядчого документообігу / П.Жежнич, О.Марчик, О.Сопрунюк // Матеріали I Міжнародної наукової конференції "Інформація, комунікація, суспільство (ІКС-2012)". – Славське-Львів: Видавництво НУ «Львівська політехніка», 2012. – С.34-36.
41. Жежнич П. Особливості формування електронної документації в сфері туризму / П.Жежнич, О.Сопрунюк // Інформаційні системи та мережі: Вісник Національного університету "Львівська політехніка". – Львів, 2013. – №770. – С.101-108.
42. Жежнич П. Особливості формування інформаційного забезпечення туристичної діяльності / П.Жежнич, О.Сопрунюк // Інформаційні системи та мережі: Вісник Національного університету "Львівська політехніка". – Львів, 2014. – №783. – С.336-343.
43. Жежнич П.І. Система агрегації та аналізу RSS-каналів / Жежнич П.І., Пелешишин А.М., Серов Ю.О., Тарасов Д.О. // Вісник Національного університету "Львівська політехніка": Інформаційні системи та мережі. – Львів, 2007. – №589. – С.91-98.

44. Жежнич П.І. Технології інформаційного менеджменту: Навчальний посібник – Львів: Видавництво Національного університету “Львівська політехніка”, 2010. – 260 с.
45. Забалдіна Ю.Б. Роль соціальних мереж у маркетингу туристських дестинацій: [Електронне видання] // "Туризм: реалії та перспективи сталого розвитку": Матеріали Міжнародної науково-практичної конференції. – Київ: Київський національний торговельно-економічний університет, 2014. – С.128-129. – <https://knteu.kiev.ua/file/MTc=/344519a49a74e01ba891b5467f7c2bc7.PDF>
46. Закон України "Про інформацію": [Електронний ресурс] // Відомості Верховної Ради України. – №2658-ХІІ від 02.10.92 у редакції від 02.06.2016. – <http://zakon3.rada.gov.ua/laws/show/2657-12>.
47. Закон України "Про електронні документи та електронний документообіг": [Електронний ресурс] // Відомості Верховної Ради України. – №851-IV від 22.05.2003 у редакції від 06.11.2014. – <http://zakon3.rada.gov.ua/laws/show/851-15>.
48. Закон України "Про туризм": [Електронний ресурс] // Відомості Верховної Ради України. – №325/95-ВР від 15.09.95 у редакції №1282-IV від 18.11.2003 із змінами 2009-2015 років. – <http://zakon2.rada.gov.ua/laws/show/324/95-%D0%B2%D1%80>.
49. Захарова Е.Г. Опис наукових електронних ресурсів метаданими дублінського ядра / Е.Г.Захарова, О.В. Захарова, В.А. Резніченко // Проблеми програмування. — 2008. — № 2–3. Спец.випуск. — С. 507–514.
50. 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.

51. Игнатова И.Г. Подходы к созданию системы поддержки функционирования репозитория портала // Интернет-порталы: содержание и технологии: Сб. науч. ст. Вып. 1. – М.: Просвещение, 2003.
52. 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>
53. Ильенкова С.Д. Управление качеством: учебник для студентов вузов / С.Д. Ильенкова. – М.: Юнити-Дана, 2007. – 352 с.
54. Ільїн Л. Атрактивність озерних комплексів Шацького національного природного парку / Л.Ільїн, А.Гринасюк // Науковий вісник Східноєвропейського національного університету імені Лесі Українки. Географічні науки. – 2014. – № 11. – С. 25-29. – Режим доступу: [http://nbuv.gov.ua/UJRN/Nvvnug\\_2014\\_11\\_6](http://nbuv.gov.ua/UJRN/Nvvnug_2014_11_6)
55. Кардаш В.Я. Товарна інноваційна політика: підручник/ В.Я. Кардаш, І.А. Павленко, О.К. Шафалюк.— К.: КНЕУ, 2002. — 266 с.
56. Квартальнов В. А. Туризм / В.А. Квартальнов. – М.: Финансы и статистика, 2004. – 320 с.
57. Квартальнов В.А. Туризм: теория и практика: избранные труды: в 5 т. / В. А. Квартально. – М.: Финансы и статистика, 1998. – Т. 2. – 256 с; Т. 3. – 384 с.
58. Кифяк В. Ф. Організація туристичної діяльності в Україні / Кифяк В.Ф. – Чернівці : Книги – ХХІ, 2003. – 300 с.
59. Кісь Я.П. Інтернет-меми як засіб розповсюдження інформації / Я.П. Кісь, О.О.Сопрунюк // Матеріали IV Міжнародної конференції молодих вчених “Комп’ютерні науки та інженерія (CSE-2010)”. - Львів, 2010. - С.396-397.
60. Кляп М.П. М.П. Сучасні різновиди туризму: навч. посіб. / М.П. Кляп, Ф.Ф. Шандор. – К.: Знання, 2011. – 334 с.
61. Круглова Н.Ю. Система качества продукции: новые подходы (вопросы и ответы) : справоч. пособие / Н.Ю. Круглова. – М.: Легпромбытстрой, 1991. – 175 с.

62. Круковский М.Ю. Оценка эффективности систем Workflow / М.Ю. Круковский // Корпоративные системы. – 2007. – № 1. – С. 14 – 18.
63. Кудла Н.Є. Менеджмент туристичного підприємства: підручник / Н.Є. Кудла. – К.: Знання, 2012. – 343 с.
64. 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.
65. Лебединец В.А. Анализ актуальности интегрированных систем менеджмента в фармации / В.А. Лебединец, А.В. Бурсаков // Провизор. – 2008. – № 17. – С. 26–28.
66. Лебединец В. А. Анализ тенденций и прогнозирование перспектив внедрения систем менеджмента качества в Украине / В.А. Лебединец, С.Н. Коваленко, Е.М. Проскурня // Вісник НТУ «ХПІ» «Економічний прогрес та ефективність виробництва». – 2008. – № 2. – С. 127–136.
67. Костенко Н. Досвід контент-аналізу: моделі та практики: Монографія / Н.Костенко, В.Іванов. - К.: Центр вільної преси, 2003. - 62с.
68. Кугель Р.Д. Управление неструктурированной информацией [Электронный ресурс]. Intelligent enterprise, 2004. – Электрон. дан. – Режим доступа: <http://www.iemag.ru/opinions/detail.php?ID=17860>
69. 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.
70. Любіцева О.О. Методика розробки турів: навч. посіб./ О.О. Любіцева. – К.: Альтерпрес, 2003. – 104 с. 15. Маликова М.В. Управление документацией системы менеджмента качества / М.В. Маликова // Технологии качества жизни. – 2002. – № 1. – С. 45–64.
71. Мазур И. И. Управление качеством: учеб. пособие для студентов вузов / И. И. Мазур, В. Д. Шапиро; под общ.ред. И. И. Мазура. – М.: Омега-Л, 2005. – 256 с.

72. Маликова М. В. Управление документацией системы менеджмента качества / М. В. Маликова // Технологии качества жизни. — 2002. — № 1. — С. 45–64.
73. Мальська М. Я. Основи туристичного бізнесу: навч. посіб. / М. Я. Мальська, В. В. Худо, В. І. Цибух— Львів: Видавничий центр ЛНУ імені Івана Франка, 2003. — 360 с.
74. Марків О. Формування туристичної документації на основі виявлення та заповнення прогалин // Інформація, комунікація, суспільство (ІКС-2015): матеріали I Міжнародної наукової конференції ІКС-2015. — Славське-Львів: Видавництво НУ «Львівська політехніка», 2015. — С.230-231.
75. Марковець О.В. Математичне та програмне забезпечення організації взаємодії громадян з органами влади в гетерогенних Веб-середовищах: Дисертація на здобуття наукового ступеня к.т.н. — Львів: Національний університет "Львівська політехніка", 2015. - 167с.
76. Матвієнко О. Основи організації електронного документообігу: навчальний посібник / О. Матвієнко, М. Цивін. — К.: Центр навчальної літератури, 2008. — 112 с.
77. Мельниченко С.В. Інформаційні технології в туризмі: теоретичні та практичні аспекти / С.В. Мельниченко // Вісник Запорізьк. нац.. ун-ту. — 2010. — № 2(6). — С. 129–138. — [http://www.nbuuv.gov.ua/old\\_jrn/Natural/Vznu/eco/2010\\_2/129-138.pdf](http://www.nbuuv.gov.ua/old_jrn/Natural/Vznu/eco/2010_2/129-138.pdf).
78. 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://ibusiness.ru/blogs/17349>
79. Мельниченко С.В. Стандартизація та сертифікація як основа забезпечення якісного обслуговування у підприємствах готельного господарства та туризму: зб. наук. праць / С.В. Мельниченко, С.П. Гаврилюк — К.: КДТЕУ,

2000. – С. 21–26. (Особистий внесок: розкрито систему стандартизації та сертифікації туристських послуг).
80. Мешков Д.Ю. Електронний документ: обіг, зберігання, використання, законодавчі аспекти/ Д.Ю. Мешков // Архіви України. – 2001. – № 6. – С. 73–87.
81. Моисеева Н. К. Стратегическое управление туристской фирмой/Н.К. Моисеева. – М.: Финансы и статистика, 2000.
82. Морозов, М. А. Информационные технологии в социально–культурном сервисе и туризме. Оргтехника [Текст]: учебник / М. А. Морозов, Н. С. Морозова. – 2–е изд., стереотип. – М.: Изд. центр «Академия», 2004. – 240 с.
83. Лапо П.М., Соколов А.В. Введение в электронные библиотеки/ П.М.Лапо, А.В. Соколов // [Электронный ресурс]. – Москва, – 2005. - 92 с.
84. Негуляев Е.А. Универсальные схемы метаданных и задача описания веб-ресурсов // Библиосфера. – 2006. – №1. – с. 43–54.
85. Нефедова М.Г. Документация в информационном обществе / М.Г. Нефедова. – М: РОЯ, 200. – 238 с.
86. Неточности в документации ATF [Электронный ресурс]. – Режим доступа: <http://www.transaq.ru/forum/index.php?topic=369.0>
87. О документации [Электронный ресурс]. – Режим доступа: [http://www.gnuman.ru/IT\\_by\\_yourself/19\\_o\\_dokumentatsii/](http://www.gnuman.ru/IT_by_yourself/19_o_dokumentatsii/)
88. Окрепилов В.В. Управление качеством : учебник / В.В. Окрепилов. – М.: Экономика, 2000. – 911 с.
89. Орлов П.І. Інформація та інформатизація: Нормативно-правове забезпечення / П. І. Орлов //Науково-практичний посібник. – Х.: Вид-во Ун-ту внутр. справ, 2003. – 724с.
90. Осипенко К. В. Тенденції розвитку світового ринку туристичних послуг в умовах посилення глобалізаційних процесів / К. В. Осипенко. // Економіка. Управління. Інновації. – 2013. – № 1. – Режим доступу: [http://nbuv.gov.ua/UJRN/eui\\_2013\\_1\\_44](http://nbuv.gov.ua/UJRN/eui_2013_1_44)



91. Основні причини виникнення неякісної первинної інформації [Електронний ресурс]. – Режим доступу: [http://pidruchniki.com.ua/17910211/finansi/osnovni\\_prichini\\_viniknennya\\_neyakisnoyi\\_pervinnoyi\\_informatsiyi](http://pidruchniki.com.ua/17910211/finansi/osnovni_prichini_viniknennya_neyakisnoyi_pervinnoyi_informatsiyi)
92. Офіційний сайт Асоціації лідерів турбізнесу України (АЛТУ). Режим доступу : [www.altu.com.ua](http://www.altu.com.ua).
93. Ошибки и неточности в документации и примерах [Електронний ресурс]. – Режим доступу: <http://www.softwaresecurity.org/forum/viewtopic.php?f=5&t=93>
94. Ошибки/опечатки/неточности в существующей документации [Електронний ресурс]. – Режим доступу: <http://forum.bgbilling.ru/viewtopic.php?f=1&t=4435>
95. 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.
96. Пелецишин А. Роль прогалин у документації в процесах виготовлення та експлуатації продукції / А.Пелецишин, О.Сопрунюк, О.Тимовчак-Максимець // Комп’ютерні науки та інформаційні технології: Вісник Національного університету “Львівська політехніка”. - Львів, 2011. - №719. - С.17-22.
97. Перехрест Г. Впровадження електронного документообігу: огляд вітчизняних систем / Г. Перехрест// Довідник секретаря та офіс-менеджера. – 2007. – № 6. – С. 38–44.
98. Петрова Г. М. Моделирование технических текстов инженерного профиля как методологическая проблема: Автореф. дис. канд. пед. наук. – М., 2002.- 23 с.
99. Письменний О. А. Аналіз ринку міжнародного туризму України // Економічний аналіз: зб. наук. праць / Тернопільський національний

- економічний університет. – Тернопіль : Видавничо-поліграфічний центр Тернопільського національного економічного університету «Економічна думка», 2014. - Том 15. - № 1. - С. 144-153. - ISSN 1993-0259.
100. Пінчук Н.С. Інформаційні системи і технології в маркетингу / Н.С. Пінчук, Г.П. Галузинський, Н.С. Орленко. – К., 2005. – 352 с.
  101. Про електронні документи та електронний документообіг: Закон України // Вісник Держ. Комітету архівів України. – 2003. – Вип. 2 (14). – С. 15–22.
  102. Пуцентейло П.Р. Економіка і організація туристично–готельного підприємництва навчальний посібник. – К.: Центр учб. літ., 2007. – 344 с.
  103. Резниченко В.А. Застосування стандарту Dublin Core Metadata Initiative для опису наукових електронних ресурсів / В.А. Резниченко, О.В. Захарова, Е.Г. Захарова// Ін-т програмних систем НАН України. – К.: 2008. – 36 с.
  104. Ришковець Ю.В. Методи та засоби побудови Веб-галерей з урахуванням інтересів користувача: Дисертація на здобуття наукового ступеня к.т.н. - Львів: Національний університет "Львівська політехніка", 2013. - 180с.
  105. Розширений пошук Google: [Електронний ресурс] // Google Inc. - 2016. - [https://www.google.com.ua/advanced\\_search](https://www.google.com.ua/advanced_search).
  106. Рудюк В.В. Класифікація електронних документів / В.В. Рудюк. – М., 2002. – С. 25–27.
  107. Салухіна Н.Г. Стандартизація та сертифікація товарів і послуг: підручник / Н.Г. Салухіна, О.М. Язвінська. — К.: Центр учбової літератури, 2010. — 336 с.
  108. Сапрунова В.Б. Туризм: еволюція, структура, маркетинг / В.Б. Сапрунова – М.: Осць–89, 1997. – 160 с.
  109. Серов Ю. О. Методи та засоби побудови ефективних віртуальних спільнот на основі Веб-форумів: Дисертація на здобуття наукового ступеня к.т.н. - Львів: Національний університет "Львівська політехніка", 2010. - 165с.
  110. Сидоренко І.О. Критерій якості в конкурентоспроможності туристичних підприємств // Вісник ДІТБ: Серія "Менеджмент та маркетинг" / наук. ред. Н.Й. Коніщева / І.О. Сидоренко. – Донецьк, 2002. – № 6. – С. 144–148.

111. Сидоренко І.О. Якість – гарантія успіху та конкурентоспроможності / І.О. Сидоренко // Економічні науки: Науковий вісник Чернівецького торговельно–економічного інституту КНТЕУ. – Чернівці: АНТ Лтд, 2005. – Вип. III. – С. 458–461.
112. Силков С.В. Нормативная база и государственное регулирование электронного документооборота / С.В. Силков. – М., 2007. – С. 146–151.
113. Системи управління якістю. Вимоги : ISO 9001–2001.– [Чинний від 2001–104–01]. – К.: Держстандарт України, 2001. – 23 с.
114. Системи управління якістю. Настанови щодо поліпшення діяльності. Вимоги :ISO 9004–2001. – [Чинний від 2001–01–01]. – К.: Держстандарт України, 2001. – 44 с.
115. Скопень М.М. Комп'ютерні інформаційні технології в туризмі / М.М. Скопень.– К.: Кондор, 2005. – 301
116. Советов Б. Я. Информационные технологии [Текст] / Б. Я Советов: учеб. для студ. вузов. – 2–е изд., стер. – М.: Высш. шк., 2005. – 264 с.
117. Сопрунюк О.О. Метод оцінювання якості туристичної документації // Інформаційні системи та мережі: Вісник Національного університету “Львівська політехніка”. – Львів, 2014. – №805. – С.333-343.
118. Тимовчак-Максимець О.Ю. Комп'ютерно-лінгвістичні методи та засоби виявлення споживацького досвіду на веб-форумах: Дисертація на здобуття наукового ступеня к.т.н. – Львів: Національний університет "Львівська політехніка", 2013. – 190с.
119. Ткаченко Т.І. Управління якістю готельних послуг: монографія / Т.І. Ткаченко, Мельниченко С.В., Новак М.В. – К.: Київ. нац. торг.-екон. ун-т, 2006. – 234 с.
120. Ткаченко Т.І., Гаврилюк С.П. Аналіз конкуренції на туристичному ринку України.// Вісник Донецького інституту туристичного бізнесу. – 2000. – №2. – С. 113–121.

121. Третяков О.В. Інформаційно-комунікаційна трансформація туристичних підприємств: [Електронне видання] // "Туризм: реалії та перспективи сталого розвитку": Матеріали Міжнародної науково-практичної конференції. – Київ: Київський національний торговельно-економічний університет, 2014. – С.58-59. – <https://knteu.kiev.ua/file/MTc=/344519a49a74e01ba891b5467f7c2bc7.PDF>.
122. Тушницький Р. Алгоритм ідентифікації подібності складно-структурованих даних на основі семантичних мереж / Р.Б.Тушницький В.М.Макар // Східно-Європейський журнал передових технологій. - 2015. - №5/2 (77). – С.38-44. – <http://cyberleninka.ru/article/n/algoritm-identifikatsiyi-podibnosti-skladno-strukturovanih-danih-na-osnovi-semantichnih-merezh.pdf>.
123. Тюття Л.Т. Соціальна робота : теорія і практика : навч. посіб. / Л. Т. Тюття, І. Б. Іванова. – [2-ге вид.] – К.: Знання, 2008. – 574 с.
124. Федущко С.С. Методи та засоби комп'ютерно-лінгвістичного аналізу достовірності соціально-демографічних характеристик учасників віртуальних спільнот: Дисертація на здобуття наукового ступеня к.т.н. – Львів: Національний університет "Львівська політехніка", 2015. – 205с.
125. Цегелик Г.Г. Организация и поиск информации в базах данных – Львов: ВШ, 1987. - 176с.
126. Чорненька Н.В. Організація туристичної індустрії: навч. посіб. / Чорненька Н.В. – К.: Атіка, 2009. – 392 с.
127. Чуковенков А.Ю. Правила оформления документов / А.Ю.Чуковенков, В.Ф. Янковая. – М.: Проспект, 2004. – 210 с.
128. Шаповал М.І. Менеджмент якості : підручник / М.І. Шаповал. – К.: КОО «Знання», 2007. – 457 с.
129. 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.

130. Шеєнкова К.А. Інтернет-технології у просуванні туристичного продукту: [Електронне видання] // "Туризм: реалії та перспективи сталого розвитку": Матеріали Міжнародної науково-практичної конференції. – Київ: Київський національний торговельно-економічний університет, 2014. – С.416-417. – <https://knteu.kiev.ua/file/MTc=/344519a49a74e01ba891b5467f7c2bc7.PDF>.
131. Школа І.М. Менеджмент туристичної індустрії: навч. посіб. / за ред. проф. І.М. Школи. – Чернівці, 2003. – 596 с.
132. Явкін В.Г. Атрактивність рекреаційних ресурсів за ціновими пропозиціями в Чернівецькій області / В.Г.Явкін, Я.В.Кирпушко, А.В.Морар // Географія та туризм. – 2012. – Вип. 19. – С. 65-71. – Режим доступу: [http://nbuv.gov.ua/UJRN/gt\\_2012\\_19\\_12](http://nbuv.gov.ua/UJRN/gt_2012_19_12)
133. Ястремська О.О. Визначення атрактивності туристичних компаній у міжнародному туризмі // Економічний часопис - XXI. - К.: Інститут Трансформації Суспільства, 2014. – №11-12'2014. - С.124-127. - [http://soskin.info/userfiles/file/2014/11-12\\_2014/Iastremaska-O-O.pdf](http://soskin.info/userfiles/file/2014/11-12_2014/Iastremaska-O-O.pdf).
134. Andreychuk N.I. The contrastive analysis of the notion of Homo Socialis in English and Ukrainian / N.I.Andreychuk, O.Soprnyuk // Матеріали IV Міжнародної науково-технічної конференції CSIT'09. – Львів: Видавництво ПП “Вежа і Ко”, 2009. – Р.465-468.
135. Crosby P.B. Quality Is Free: The Art of Making Quality Certain / P.B. Crosby. – М.: Mentor Books, 1992. – 272 p.
136. Chaffey D. E-Business and E-Commerce Management (strategy, implementation, and practice): Fourth edition. - Edinburgh, : Pearson Education, 2009. - 764p.: ISBN 978-0-273-71960-1. – <https://geomart25.files.wordpress.com/2014/05/dave-chaffey-e-business-and-e-commerce-management-strategies-4th-ed-qwerty80.pdf>.
137. Extensible Markup Language (XML): [Електронний ресурс] // W3C. - 11 October 2016. – <https://www.w3.org/XML/>.

138. Juran J.M. Quality Control Handbook. /J.M. Juran J.M., F. Gryna. – M.: McGraw–Hill, 1998. – 1774 p.
139. Healy S. ISO 15489 Records management – its development and significance/S. Healy // Records Management Journal. – December 2001. – Vol. 11. – № – P. 133–142.
140. Larry Shillito M. Acquiring, processing, and deploying voice of the customer / M. Larry Shillito. – St. Lucie Press, 2001. –279 p.
141. Libby D. RSS 0.91 Spec, revision 3: [Електронний ресурс] // Netscape Communications. – 10 July 1999. – <https://web.archive.org/web/20001204093600/http://my.netscape.com/publish/formats/rss-spec-0.91.html>.
142. Markiv O. Graphical presentation of tourism documentation structure // “INFORMATION, COMMUNICATION, SOCIETY” (ICS-2016). – Slavske-LVIV, 2016.
143. Markiv O.O. Peculiarities of the dublin core use in the formation process of tourism documentation. // Вісник Комп’ютерні науки та інформаційні технології. – Львів: Видавництво Львівської політехніки, 2016. – №843. – С. 104-107.
144. Melnykova N. Semantic approach to personalization of medical data. / Melnykova N., Markiv O. // Матеріали XI Міжнародної науково-технічної конференції Комп’ютерні науки та інформаційні технології CSIT 2016. – Львів: Видавництво Львівської політехніки, 2016. – С. 59-62.
145. Mizzaro S. Relevance: The Whole History // Journal of the American Society for Information Science. - #48. - P.810 - 832.
146. Nottingham M. The Atom Syndication Format: [Електронний ресурс] / M.Nottingham, R.Sayre // Network Working Group. - December 2005. - <https://tools.ietf.org/html/rfc4287>.
147. Prideaux B. Managing Tourism and hospitality services: theory and international applications/B. Prideaux.,G. Moscardo, E. Laws. – London – 2006. – 337 p.

148. Resource Description Framework (RDF): [Електронний ресурс] // W3C. - 15 March 2014. - <https://www.w3.org/RDF/>.
149. RSS 2.0 and Atom 1.0 Compared // Atom Wiki. - 4 December 2007. - <http://www.intertwingly.net/wiki/pie/Rss20AndAtom10Compared>.
150. Software engineering – Product quality – Part 1: Quality model [Electronic resource] / ISO/IEC 9126-1:2001. – Mode of access : WWW/URL : <http://www.iso.org/iso/en/CatalogueDetailPage.CatalogueDetail?CSNUMBER=22749&ICS1=35&ICS2=80&ICS3>. – 15.06.2011. – Title from the screen.
151. Software engineering – Product quality – Part 2: External metrics [Electronic resource] / ISO/IEC TR 9126-2:2003. – Mode of access : WWW/URL : [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=22750](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=22750). – 20.06.2011. – Title from the screen.
152. Soprnyuk O. Lexical representation of human statuses in Tudor times: Tudor court titles // Матеріали IV Міжнародної конференції молодих вчених CSE-2011. - Львів: Видавництво НУ «Львівська політехніка», 2011.
153. Steegman B. XML for DB2 Information Integration / B. Steegman, R. Bourret, O.Cline, O. Guyennet, S. Priestly, V. Sylenko, U. Wahli // IBM International Technical Support Organization. – July, 2004. – First edition. – 712p.
154. Systems and software engineering – Systems and software Quality Requirements and Evaluation (SQuaRE) – System and software quality models / ISO/IEC 25010.2:2008. – Mode of access : [http://sa.inceptum.eu/sites/sa.inceptum.eu/files/Content/ISO\\_25010.pdf](http://sa.inceptum.eu/sites/sa.inceptum.eu/files/Content/ISO_25010.pdf).
155. Tkachenko S. Fast formation of binary trees as parallel (optimal) reduction process models / S.Tkachenko, O.Soprnyuk, V.Tkachenko, N.Yatsyk // CAD in Machinery Design – Implementation and Educational Problems: materials of the XXI Polish-Ukrainian Conference – CADMD 2013. - Poland, Łańcut: Warsaw University of Technology, 2013. - P.49.
156. Tkachenko S. Efficiency Enhancement of Optimal Reduction Method by Strengthening Parallelism of Structural Models Formation () / S.Tkachenko,

- O.Soprnyuk, V.Tkachenko, I.Solomko // Machine Dynamics Research. – Poland, 2013. – Vol.3, No. 2. – P.85-90.
157. Tkachenko S. Models of optimal reduction in decomposition problems solution and structures modeling / S.Tkachenko, O.Soprnyuk // Комп'ютерні системи проектування. Теорія і практика: Вісник Національного університету "Львівська політехніка". – Львів, 2013. – №777. - P.12-15.
158. Travel & Tourism Economic Impact 2015 Ukraine [Електронний ресурс]. – <https://www.wttc.org/-/media/files/reports/economic%20impact%20research/countries%202015/ukraine2015.pdf>
159. UNWTO Tourism Highlights (2016 Edition) [Електронний ресурс]. – <http://www.e-unwto.org/doi/pdf/10.18111/9789284418145>
160. Vykylyuk Y. Calculation of the territory recreation attractiveness using fuzzy logic / Y.Vykylyuk, O.Artemenko // International Journal of Computing. - 2014. - Vol.8, No.2. – P.109-118. – <http://computingonline.net/index.php/computing/article/view/673/635>.
161. Wang R. Information Quality (Advances in Management Information Systems) / R. Wang. – London, 2005. – 265p.
162. Weiss Edmond H. How to write usable user documentation / H. Weiss Edmond. – London: Oryx Press, 1991. –280 p.
163. Whitehouse F. Documentation: how to organise and control information processes in business and industry / F. Whitehouse. – L.: Business Books, 1971. – 181 p.
164. Whitehouse F. Systems Documentation: Techniques of Persuasion in Large Organisations/ F. Whitehouse. – L.: Business Books, 1973. – 191 p.
165. Zhezhnych P. Analysis of the tourism documentation quality improvement / P.Zhezhnych, O.Soprnyuk // Materials of the 7th International Scientific and Technical Conference "Computer science and information technology (CSIT'2012)". – Lviv: Publishing house "Vezha and Co", 2012. – P.34-36.



166. Zhezhnych P. Formation of the consolidated data resource of tourism information / P.Zhezhnych, O.Soprunityuk // VIII Міжнародна науково-технічна конференція “Computer Sciences and Information Technologies” (“Комп’ютерні науки та інформаційні технології”) (CSIT’2013). - 2013.
167. Zhezhnych P. Formation peculiarities of tourism documentation / P.Zhezhnych, O.Soprunityuk // Materials of the Second International Academic Conference IKS-2013. – Lviv: Lviv Polytechnic National University publisher, 2013. – P.194-196.
168. Zhezhnych P. Importance of the tourism information quality monitoring / P.Zhezhnych, O.Soprunityuk // Human. Computer. Communication: materials of the 1st International Scientific Conference HCC-2013. – Lviv: Lviv Polytechnic National University publisher, 2013. - P.39.
169. Zhezhnych P. Justification of Use Expedience of Tourism Data Consolidation / P.Zhezhnych, O.Soprunityuk // Міжнародна наукова конференція «Інформація. Комунікація. Суспільство» (ICS’2014). – Славське-Львів, 2014.
170. Zhezhnych P. Peculiarities of the open web-resources use in the formation process of tourism documentation / P.Zhezhnych, O.Soprunityuk // VIII Міжнародна науково-технічна конференція “Computer Sciences and Information Technologies” (“Комп’ютерні науки та інформаційні технології”) (CSIT’2013). – Львів, 2013.

## Додаток А

Приклад RSS-каналу відкритого веб-ресурсу <http://lviv.ridne.net>

```

<?xml version="1.0" encoding="utf-8" ?><rss version="2.0" xml:base="http://lviv.ridne.net/taxonomy/term/43/all"
xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:georss="http://www.georss.org/georss">
<channel>
<title>Площа Ринок</title>
<link>http://lviv.ridne.net/taxonomy/term/43/all</link>
<description></description>
<language>uk</language>
<item>
<title>Галерея Площі Ринок</title>
<link>http://lviv.ridne.net/node/642</link>
<description>&lt;p&gt;&lt;p&gt;&lt;p&gt;&lt;p&gt;&lt;a
href=&quot;http://lviv.ridne.net/map/prominentplaces/rynok&quot;&gt;Фотографії галереї на мапі (карті)&lt;/a&gt;&lt;p&gt;
&lt;p&gt;&lt;strong&gt;Площа Ринок&lt;/strong&gt; — площа у Львові, характерне явище для середньовічної архітектури
європейських міст.&lt;/p&gt;
&lt;p&gt;&lt;a href=&quot;http://lviv.ridne.net/node/642&quot;
target=&quot;_blank&quot;&gt;докладніше&lt;/a&gt;&lt;/p&gt;</description>
<comments>http://lviv.ridne.net/node/642#comments</comments>
<category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
<pubDate>Fri, 02 Mar 2007 21:57:27 +0000</pubDate>
<dc:creator>Pavlo Zhezhnych</dc:creator>
<guid isPermaLink="false">642 at http://lviv.ridne.net</guid>
</item>
<item>
<title>Галерея площі Ринок та навколишніх вулиць</title>
<link>http://lviv.ridne.net/node/2279</link>
<description>&lt;div class=&quot;field field-type-filefield field-field-image&quot;&gt;
&lt;div class=&quot;field-items&quot;&gt;
&lt;div class=&quot;field-item odd&quot;&gt;
&lt;img src=&quot;http://lviv.ridne.net/files/imagecache/thumbnail/images/IMG_0025_1.jpg&quot; alt=&quot;&quot;
title=&quot;&quot; class=&quot;imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default&quot;
width=&quot;50&quot; height=&quot;75&quot; /&gt; &lt;/div&gt;
&lt;div class=&quot;field-item even&quot;&gt;
&lt;img src=&quot;http://lviv.ridne.net/files/imagecache/thumbnail/images/IMG_0030_1.jpg&quot; alt=&quot;&quot;
title=&quot;&quot; class=&quot;imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default&quot;
width=&quot;100&quot; height=&quot;67&quot; /&gt; &lt;/div&gt;
&lt;div class=&quot;field-item odd&quot;&gt;
&lt;img src=&quot;http://lviv.ridne.net/files/imagecache/thumbnail/images/IMG_0060_1.jpg&quot; alt=&quot;&quot;
title=&quot;&quot; class=&quot;imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default&quot;
width=&quot;50&quot; height=&quot;75&quot; /&gt; &lt;/div&gt;
&lt;div class=&quot;field-item even&quot;&gt;
&lt;img src=&quot;http://lviv.ridne.net/files/imagecache/thumbnail/images/IMG_0065_1.jpg&quot; alt=&quot;&quot;
title=&quot;&quot; class=&quot;imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default&quot;
width=&quot;50&quot; height=&quot;75&quot; /&gt; &lt;/div&gt;
&lt;div class=&quot;field-item odd&quot;&gt;
&lt;img src=&quot;http://lviv.ridne.net/files/imagecache/thumbnail/images/IMG_0125_1.jpg&quot; alt=&quot;&quot;
title=&quot;&quot; class=&quot;imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default&quot;
width=&quot;50&quot; height=&quot;75&quot; /&gt; &lt;/div&gt;
&lt;/div&gt;
&lt;/div&gt;
&lt;p&gt;&lt;p&gt;&lt;p&gt;Площа виникла в XIV ст. внаслідок регулярного планування забудови за взірцем
середньовічних західноєвропейських міст, як центр торговельного та суспільного життя міста. З кожного кута площі
виходять дві взаємоперпендикулярні вулиці, котрі закінчувалися біля міських фортифікацій, що опоясували середньовічне
місто. Площу розміром 142 x 129 м утворюють 44 будинки і ратуша в центрі. Від початкової готичної забудови, яка
загинула під час пожежі 1527 р., збереглися підвали і фрагменти стін з готичними деталями. Дбайливо відновлюючи вцілілі
мури, нові будинки після руйнівної пожежі міщани будують тільки з каменю та цегли і називають їх кам'яницями.&lt;/p&gt;
&lt;p&gt;&lt;a href=&quot;http://lviv.ridne.net/node/2279&quot;
target=&quot;_blank&quot;&gt;докладніше&lt;/a&gt;&lt;/p&gt;</description>
<comments>http://lviv.ridne.net/node/2279#comments</comments>
<category domain="http://lviv.ridne.net/taxonomy/term/111">Світлини Львова</category>
<category domain="http://lviv.ridne.net/kinds/buildings">Будівлі, вулиці</category>

```



```

<category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
<georss:point>49.841515 24.031800</georss:point>
<pubDate>Fri, 08 Jan 2010 21:00:16 +0000</pubDate>
<dc:creator>Andriy Peleschyshyn</dc:creator>
<guid isPermaLink="false">2243 at http://lviv.ridne.net</guid>
</item>
<item>
<title>Львівські піцерії Pizza Calcio</title>
<link>http://lviv.ridne.net/node/2232</link>
<description>&lt;div class="field field-type-filefield field-field-image"&gt;
&lt;div class="field-items"&gt;
&lt;div class="field-item odd"&gt;
&lt;img src="http://lviv.ridne.net/files/imagecache/thumbnail/images/P1010184_resize.JPG"
alt=" " title=" " class="imagecache imagecache-thumbnail imagecache-default imagecache-
thumbnail_default" width="100" height="75" /&gt;
&lt;/div&gt;
&lt;/div&gt;
&lt;p&gt;&lt;p&gt;&lt;p&gt;Це зображення було завантажено з дописом &lt;a href="/node/2224"&gt;Львівські
піцерії&lt;/a&gt;. &lt;/p&gt;
&lt;div class="field field-type-nodereference field-field-reference"&gt;
&lt;div class="field-items"&gt;
&lt;div class="field-item odd"&gt;
&lt;div class="field-label-inline-first"&gt;
Див.також: &nbsp;&lt;/div&gt;
&lt;a href="/node/2224"&gt;Львівські піцерії&lt;/a&gt;
&lt;/div&gt;
&lt;/div&gt;
&lt;/description>
<comments>http://lviv.ridne.net/node/2232#comments</comments>
<category domain="http://lviv.ridne.net/kinds/closeup">Крупним планом</category>
<category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий
Ринок</category>
<category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
<category domain="http://lviv.ridne.net/galleries/thematic/linguistic">Лінгвістичний пейзаж Львова</category>
<georss:point>49.842767 24.032834</georss:point>
<pubDate>Fri, 27 Nov 2009 18:30:51 +0000</pubDate>
<dc:creator>Sonyar1</dc:creator>
<guid isPermaLink="false">2232 at http://lviv.ridne.net</guid>
</item>
<item>
<title>Коник на даху будинку. Площа Ринок</title>
<link>http://lviv.ridne.net/node/2152</link>
<description>&lt;div class="field field-type-filefield field-field-image"&gt;
&lt;div class="field-items"&gt;
&lt;div class="field-item odd"&gt;
&lt;img src="http://lviv.ridne.net/files/imagecache/thumbnail/images/IMG_0467-crop.jpg"
alt=" " title=" " class="imagecache imagecache-thumbnail imagecache-default imagecache-
thumbnail_default" width="100" height="75" /&gt;
&lt;/div&gt;
&lt;/div&gt;
&lt;p&gt;&lt;p&gt;&lt;p&gt;Якщо дивитися зверху, видно симпатичного коника, якого хтось намалював на даху
будинку.&lt;/p&gt;
&lt;div class="field field-type-date field-field-date"&gt;
&lt;div class="field-items"&gt;
&lt;div class="field-item odd"&gt;
&lt;div class="field-label-inline-first"&gt;
Дата: &nbsp;&lt;/div&gt;
&lt;span class="date-display-single"&gt;Нед, 09.08.2009&lt;/span&gt;
&lt;/div&gt;
&lt;/div&gt;
&lt;/description>
<comments>http://lviv.ridne.net/node/2152#comments</comments>
<category domain="http://lviv.ridne.net/kinds/панорама">Панорама</category>
<category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий
Ринок</category>
<category domain="http://lviv.ridne.net/microrayons/ancient">Старий Львів</category>
<category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
<georss:point>49.841486 24.030530</georss:point>

```

<pubDate>Sun, 09 Aug 2009 19:27:14 +0000</pubDate>  
 <dc:creator>Andriy Peleschshyn</dc:creator>  
 <guid isPermaLink="false">2152 at http://lviv.ridne.net</guid>  
 </item>  
 <item>  
 <title>Дзвони на Рагуші вибивають 5-ту вечора</title>  
 <link>http://lviv.ridne.net/node/2151</link>  
 <description>&lt;p&gt;&lt;p&gt;&lt;p&gt;Маленьке відео, що зняте мильничкою-телефоном. Якість поганенька, але звук ніби переданий добре.&lt;br /&gt;  
 Насолоджуйтесь львівськими дзвонами!&lt;br /&gt;  
 &lt;iframe src="http://www.youtube.com/embed/Es-cdwa79Ic?rel=1&amp;autoplay=0&quot; width="320&quot; height="205&quot; frameborder="0&quot;&gt;&lt;/iframe&gt;&lt;/p&gt;  
 </description>  
 <comments>http://lviv.ridne.net/node/2151#comments</comments>  
 <category domain="http://lviv.ridne.net/seasons/summer">Літо</category>  
 <category domain="http://lviv.ridne.net/dayperiods/day">День</category>  
 <category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий Ринок</category>  
 <category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>  
 <category domain="http://lviv.ridne.net/galleries/thematic/video">Відеоматеріали</category>  
 <pubDate>Sat, 08 Aug 2009 17:50:19 +0000</pubDate>  
 <dc:creator>Andriy Peleschshyn</dc:creator>  
 <guid isPermaLink="false">2151 at http://lviv.ridne.net</guid>  
 </item>  
 <item>  
 <title>Опис пожежі міста Львова 1527 р.</title>  
 <link>http://lviv.ridne.net/node/2093</link>  
 <description>&lt;p&gt;&lt;p&gt;&lt;p&gt;&lt;span style="font-weight:bold"&gt;&gt;1527 р., червня 3.— ОПИС ПОЖЕЖИ МІСТА ЛЬВОВА НЕВІДОМОГО АВТОРА&lt;/span&gt;&lt;p&gt;&lt;p&gt;Описую сумну та злочасну подію нашого прекрасного славнозвісного міста, жалогідну й навіть довговічну Із-за болю руйну, яку треба оплакувати через безпошадність цієї пожежі. Адже тисяча п'ятсот двадцять сьомого року в понеділок після свята Вознесення господнього, о двадцять третій годині, несподівано з броваря дому якогось-то Григорія Шолтиса з Клепарова, що стояв напроти монастиря францисканів, на вулиці Пекарів \*, яка лежала поруч, вибухла величезна й раптова пожежа.&lt;p&gt;&lt;p&gt;&lt;a href="http://lviv.ridne.net/node/2093" target="\_blank"&gt;докладніше&lt;/a&gt;&lt;/p&gt;&lt;/description>  
 <comments>http://lviv.ridne.net/node/2093#comments</comments>  
 <category domain="http://lviv.ridne.net/taxonomy/term/117">Історія Львова</category>  
 <category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий Ринок</category>  
 <category domain="http://lviv.ridne.net/galleries/commonstreets/virmenska">Вірменська, Театральна</category>  
 <category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>  
 <category domain="http://lviv.ridne.net/taxonomy/term/319">Латинський катедральний собор</category>  
 <category domain="http://lviv.ridne.net/taxonomy/term/318">Успенська церква</category>  
 <category domain="http://lviv.ridne.net/historicaldocs/richpospolyta">Річ Посполита – 16 ст.</category>  
 <pubDate>Tue, 21 Apr 2009 07:22:02 +0000</pubDate>  
 <dc:creator>Pavlo Zhezhnych</dc:creator>  
 <guid isPermaLink="false">2093 at http://lviv.ridne.net</guid>  
 </item>  
 <item>  
 <title>Про обмеження прав міщан русинів (українців)</title>  
 <link>http://lviv.ridne.net/node/2092</link>  
 <description>&lt;p&gt;&lt;p&gt;&lt;p&gt;&lt;span style="font-weight:bold"&gt;&gt;1525 р., лютого 18. — ДЕКРЕТ ПОЛЬСЬКОГО КОРОЛЯ СІГІЗМУНДА І ПРО ОБМЕЖЕННЯ ПРАВ МІЩАН РУСИНІВ (УКРАЇНЦІВ)&lt;/span&gt;&lt;p&gt;&lt;p&gt;... \* Ми, Сігізмунд... \*\*, сповіщаємо цим декретом усім, кому це потрібно, що минулого року, коли ми були у Львові, міщани руської віри цього нашого міста Львова, яких тут чимало, а саме ті, що сидять на тому ж праві, і ті самі, які нарівні з іншими міщанами, жителями цього міста, несуть міські тягари, скаржилися нам, що зазнають з боку бурмистра і райців згаданого міста перешкод у купуванні й набуванні будинків, розташованих поза їхньою вулицею, в триманні корчм і в торгівлі та шинкаруванні вином і пивом, у продажу інших подібних речей у відтинанні на лікті тканин, а також, що забороняють їм виконувати ремесло і належати до цехів і не допускають їх до них з великою їхньою втратою, шкодою і приниженням \*\*\*.&lt;p&gt;&lt;p&gt;&lt;a href="http://lviv.ridne.net/node/2092" target="\_blank"&gt;докладніше&lt;/a&gt;&lt;/p&gt;&lt;/description>  
 <comments>http://lviv.ridne.net/node/2092#comments</comments>  
 <category domain="http://lviv.ridne.net/taxonomy/term/117">Історія Львова</category>

```

<category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий Ринок</category>
<category domain="http://lviv.ridne.net/taxonomy/term/85">Підзамче</category>
<category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
<category domain="http://lviv.ridne.net/historicaldocs/richprospolyta">Річ Посполита – 16 ст.</category>
<pubDate>Mon, 20 Apr 2009 08:52:11 +0000</pubDate>
<dc:creator>Pavlo Zhezhnych</dc:creator>
<guid isPermaLink="false">2092 at http://lviv.ridne.net</guid>
</item>
<item>
<title>Пл. Ринок Художник В. Стопачинський</title>
<link>http://lviv.ridne.net/node/2027</link>
<description>&lt;div class="field field-type-filefield field-field-image"&gt;
&lt;div class="field-items"&gt;
&lt;div class="field-item odd"&gt;
&lt;img src="http://lviv.ridne.net/files/imagecache/thumbnail/images/159 (25) Пл. Ринок.JPG" alt=" " title=" " class="imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default" width="53" height="75" /&gt;
&lt;/div>
&lt;/div>
&lt;p>&lt;/p>
&lt;p>Це зображення було завантажено з дописом &a href="http://lviv.ridne.net/node/2026">Пл. Ринок Художник В. Стопачинський&a;
&lt;/p>
&lt;div class="field field-type-nodereference field-field-reference"&gt;
&lt;div class="field-items"&gt;
&lt;div class="field-item odd"&gt;
&lt;div class="field-label-inline-first"&gt;
Див. також:
&lt;/div>
&lt;a href="http://lviv.ridne.net/node/2026">Пл. Ринок Художник В. Стопачинський&a;
&lt;/div>
&lt;/div>
&lt;/description>
<comments>http://lviv.ridne.net/node/2027#comments</comments>
<category domain="http://lviv.ridne.net/kinds/buildings">Будівлі, вулиці</category>
<category domain="http://lviv.ridne.net/seasons/summer">Літо</category>
<category domain="http://lviv.ridne.net/dayperiods/day">День</category>
<category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий Ринок</category>
<category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
<category domain="http://lviv.ridne.net/taxonomy/term/65">Храми Львова</category>
<category domain="http://lviv.ridne.net/galleries/thematic/art">Мистецтво Львова</category>
<pubDate>Sat, 17 Jan 2009 23:52:27 +0000</pubDate>
<dc:creator>elektron</dc:creator>
<guid isPermaLink="false">2027 at http://lviv.ridne.net</guid>
</item>
<item>
<title>Пл. Ринок Художник В. Стопачинський</title>
<link>http://lviv.ridne.net/node/2028</link>
<description>&lt;div class="field field-type-filefield field-field-image"&gt;
&lt;div class="field-items"&gt;
&lt;div class="field-item odd"&gt;
&lt;img src="http://lviv.ridne.net/files/imagecache/thumbnail/images/159 (98) Пл. Ринок.JPG" alt=" " title=" " class="imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default" width="100" height="75" /&gt;
&lt;/div>
&lt;/div>
&lt;p>&lt;/p>
&lt;p>Це зображення було завантажено з дописом &a href="http://lviv.ridne.net/node/2026">Пл. Ринок Художник В. Стопачинський&a;
&lt;/p>
&lt;div class="field field-type-nodereference field-field-reference"&gt;
&lt;div class="field-items"&gt;
&lt;div class="field-item odd"&gt;
&lt;div class="field-label-inline-first"&gt;
Див. також:
&lt;/div>
&lt;a href="http://lviv.ridne.net/node/2026">Пл. Ринок Художник В. Стопачинський&a;
&lt;/div>
&lt;/div>
&lt;/description>
<comments>http://lviv.ridne.net/node/2028#comments</comments>
<category domain="http://lviv.ridne.net/kinds/buildings">Будівлі, вулиці</category>

```

```

<category domain="http://lviv.ridne.net/seasons/summer">Літо</category>
<category domain="http://lviv.ridne.net/dayperiods/day">День</category>
<category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий Ринок</category>
<category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
<category domain="http://lviv.ridne.net/galleries/thematic/art">Мистецтво Львова</category>
<pubDate>Sat, 17 Jan 2009 23:52:27 +0000</pubDate>
<dc:creator>elektron</dc:creator>
<guid isPermaLink="false">2028 at http://lviv.ridne.net</guid>
</item>
<item>
<title>Пл. Ринок Художник В. Стопачинський</title>
<link>http://lviv.ridne.net/node/2029</link>
<description>&lt;div class="field field-type-filefield field-field-image"
&lt;div class="field-items"
&lt;div class="field-item odd"
&lt;img src="http://lviv.ridne.net/files/imagecache/thumbnail/images/159 (99) Пл. Ринок.JPG"
alt=""; title=""; class="imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default"; width="92"; height="75"; /&gt; &lt;/div&gt;
&lt;/div&gt;
&lt;p&gt;&lt;p&gt;&lt;p&gt;Це зображення було завантажено з дописом &lt;a href="/node/2026"
&lt;p&gt;Пл. Ринок Художник В. Стопачинський&lt;/p&gt;
&lt;div class="field field-type-nodereference field-field-reference"
&lt;div class="field-items"
&lt;div class="field-item odd"
&lt;div class="field-label-inline-first"
Див.також: &nbsp;&lt;/div&gt;
&lt;a href="/node/2026"
&lt;/div&gt;
&lt;/div&gt;
&lt;/div&gt;
</description>
<comments>http://lviv.ridne.net/node/2029#comments</comments>
<category domain="http://lviv.ridne.net/kinds/buildings">Будівлі, вулиці</category>
<category domain="http://lviv.ridne.net/seasons/summer">Літо</category>
<category domain="http://lviv.ridne.net/dayperiods/day">День</category>
<category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий Ринок</category>
<category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
<category domain="http://lviv.ridne.net/taxonomy/term/65">Храми Львова</category>
<category domain="http://lviv.ridne.net/galleries/thematic/art">Мистецтво Львова</category>
<pubDate>Sat, 17 Jan 2009 23:52:27 +0000</pubDate>
<dc:creator>elektron</dc:creator>
<guid isPermaLink="false">2029 at http://lviv.ridne.net</guid>
</item>
<item>
<title>Пл. Ринок Художник В. Стопачинський</title>
<link>http://lviv.ridne.net/node/2026</link>
<description>&lt;div class="field field-type-filefield field-field-image"
&lt;div class="field-items"
&lt;div class="field-item odd"
&lt;img src="http://lviv.ridne.net/files/imagecache/thumbnail/images/159 (6) Пл. Ринок.jpg"
alt=""; title=""; class="imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default"; width="90"; height="75"; /&gt; &lt;/div&gt;
&lt;div class="field-item even"
&lt;img src="http://lviv.ridne.net/files/imagecache/thumbnail/images/159 (25) Пл. Ринок.JPG"
alt=""; title=""; class="imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default"; width="53"; height="75"; /&gt; &lt;/div&gt;
&lt;div class="field-item odd"
&lt;img src="http://lviv.ridne.net/files/imagecache/thumbnail/images/159 (98) Пл. Ринок.JPG"
alt=""; title=""; class="imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default"; width="100"; height="75"; /&gt; &lt;/div&gt;
&lt;div class="field-item even"
&lt;img src="http://lviv.ridne.net/files/imagecache/thumbnail/images/159 (99) Пл. Ринок.JPG"
alt=""; title=""; class="imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default"; width="92"; height="75"; /&gt; &lt;/div&gt;
&lt;/div&gt;

```

```

</div>>
</p>></p>></div class="field field-type-nodereference field-field-reference">>
  <div class="field-items">>
    <div class="field-item odd">>
      <div class="field-label-inline-first">>
        Див.також:&nbsp;&nbsp;&nbsp;</div>>
        <a href="/node/2027">>Пл. Ринок Художник В. Стопачинський</a>> </div>>
      <div class="field-item even">>
        <div class="field-label-inline">>
          Див.також:&nbsp;&nbsp;&nbsp;</div>>
          <a href="/node/2028">>Пл. Ринок Художник В. Стопачинський</a>> </div>>
        <div class="field-item odd">>
          <div class="field-label-inline">>
            Див.також:&nbsp;&nbsp;&nbsp;</div>>
            <a href="/node/2029">>Пл. Ринок Художник В. Стопачинський</a>> </div>>
          </div>>
        </div>>
      </div>>
    </description>
    <comments>http://lviv.ridne.net/node/2026#comments</comments>
    <category domain="http://lviv.ridne.net/kinds/buildings">Будівлі, вулиці</category>
    <category domain="http://lviv.ridne.net/seasons/summer">Літо</category>
    <category domain="http://lviv.ridne.net/dayperiods/day">День</category>
    <category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий Ринок</category>
    <category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
    <category domain="http://lviv.ridne.net/galleries/thematic/art">Мистецтво Львова</category>
    <pubDate>Sat, 17 Jan 2009 23:52:25 +0000</pubDate>
    <dc:creator>elektron</dc:creator>
    <guid isPermaLink="false">2026 at http://lviv.ridne.net</guid>
  </item>
  <item>
    <title>Пл. Ринок </title>
    <link>http://lviv.ridne.net/node/1941</link>
    <description>&lt;div class="field field-type-filefield field-field-image">>
      &lt;div class="field-items">>
        &lt;div class="field-item odd">>
          &lt;img src="http://lviv.ridne.net/files/imagecache/thumbnail/images/Пивоварчук Пл. Ринок Староєврейська 2 сору.jpg" alt="" title="" class="imagecache imagecache-thumbnail imagecache-default imagecache-thumbnail_default" width="42" height="75" /&gt; </div>>
        &lt;/div>>
      &lt;/div>>
      &lt;p>&gt;&lt;p>&gt;&lt;p>&gt;Це зображення було завантажено з дописом &lt;a href="/node/1940">>Пл. Ринок</a>&gt;.&lt;/p>&gt;
      &lt;div class="field field-type-nodereference field-field-reference">>
        &lt;div class="field-items">>
          &lt;div class="field-item odd">>
            &lt;div class="field-label-inline-first">>
              Див.також:&nbsp;&nbsp;&nbsp;</div>>
              &lt;a href="/node/1940">>Пл. Ринок</a>&gt; </div>>
            &lt;/div>>
          &lt;/div>>
        &lt;/div>>
      &lt;/div>>
    </description>
    <comments>http://lviv.ridne.net/node/1941#comments</comments>
    <category domain="http://lviv.ridne.net/kinds/buildings">Будівлі, вулиці</category>
    <category domain="http://lviv.ridne.net/seasons/summer">Літо</category>
    <category domain="http://lviv.ridne.net/dayperiods/evening">Вечір</category>
    <category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий Ринок</category>
    <category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
    <category domain="http://lviv.ridne.net/galleries/thematic/mywindow">Львів з мого вікна</category>
    <pubDate>Wed, 15 Oct 2008 20:56:00 +0000</pubDate>
    <dc:creator>elektron</dc:creator>
    <guid isPermaLink="false">1941 at http://lviv.ridne.net</guid>
  </item>
  <item>
    <title>Пл. Ринок </title>
    <link>http://lviv.ridne.net/node/1942</link>

```



```

<description>&lt;div class=&quot;field field-type-filefield field-field-image&quot;&gt;
&lt;div class=&quot;field-items&quot;&gt;
  &lt;div class=&quot;field-item odd&quot;&gt;
    &lt;img src=&quot;http://lviv.ridne.net/files/imagecache/thumbnail/images/ Пивоварчук Пл. Ринок Староєврейська
3 copy.jpg&quot; alt=&quot;&quot; title=&quot;&quot; class=&quot;imagecache imagecache-thumbnail imagecache-default
imagecache-thumbnail_default&quot; width=&quot;100&quot; height=&quot;56&quot; /&gt;    &lt;/div&gt;
  &lt;/div&gt;
&lt;/div&gt;
&lt;p&gt;&lt;p&gt;&lt;p&gt;Це зображення було завантажене з дописом &lt;a href=&quot;/node/1940&quot;&gt;&lt;/a&gt;&lt;/p&gt;
Ринок&lt;/a&gt;. &lt;/p&gt;
&lt;div class=&quot;field field-type-nodereference field-field-reference&quot;&gt;
  &lt;div class=&quot;field-items&quot;&gt;
    &lt;div class=&quot;field-item odd&quot;&gt;
      &lt;div class=&quot;field-label-inline-first&quot;&gt;
        Див.також:&nbsp;&nbsp;&lt;/div&gt;
        &lt;a href=&quot;/node/1940&quot;&gt;&lt;/a&gt;Пл. Ринок&lt;/a&gt;    &lt;/div&gt;
      &lt;/div&gt;
&lt;/div&gt;
&lt;/div&gt;
</description>
  <comments>http://lviv.ridne.net/node/1942#comments</comments>
  <category domain="http://lviv.ridne.net/kinds/buildings">Будівлі, вулиці</category>
  <category domain="http://lviv.ridne.net/seasons/summer">Літо</category>
  <category domain="http://lviv.ridne.net/dayperiods/evening">Вечір</category>
  <category domain="http://lviv.ridne.net/taxonomy/term/45">Центр, Площа Ринок, Проспект Свободи, Валова, Старий
Ринок</category>
  <category domain="http://lviv.ridne.net/taxonomy/term/43">Площа Ринок</category>
  <category domain="http://lviv.ridne.net/galleries/thematic/mywindow">Львів з мого вікна</category>
  <pubDate>Wed, 15 Oct 2008 20:56:00 +0000</pubDate>
  <dc:creator>elektron</dc:creator>
  <guid isPermaLink="false">1942 at http://lviv.ridne.net</guid>
  </item>
</channel>
</rss>

```

## Додаток Б

Приклад індикативної ознаки об'єкта туристичної документації 'Ратуша'

№	Маркер	Міра відповідності до об'єкта
1.	'*Льв?в*Ратуш*'	1
2.	'*Льв?в*веж*центр*'	0,7
3.	'*Льв?в*міськ*веж*'	0,9
4.	'*Льв?в*веж*міська*рад*'	0,95
5.	'*культ*столиц*Україн*Ратуш*'	0,5
6.	'*Ратуш*площ*Ринок*'	0,5
7.	'*Ратуш*кам'ян*веж*'	0,4
8.	'*Ратуш*кам'ян*веж* площ*Ринок*'	0,5
9.	'*Льв?в*центр*'	0,3

Приклад індикативної ознаки об'єкта туристичної документації 'Оперний театр'

№	Маркер	Міра відповідності до об'єкта
1.	'*Льв?в*Опер*театр*'	1
2.	'*Льв?в*Опер*'	0,9
3.	'*Льв?в*театр*Крушельницьк*'	0,8
4.	'*Льв?в*театр*опер*балет*'	0,95
5.	'*культ*столиц*Україн*оперн* театр*'	0,5
6.	'*театр*проспект*Свобод*'	0,95
7.	'*Льв?і*національн*академічн*театр*опер*га*балет*ім*Соломі*Крушельницько*'	1
8.	'*театр*столиц*Галичин*'	0,4
9.	'*Льв?в*центр*'	0,3

Приклад індикативних ознак термінів класифікатора туристичної документації 'Рубрики'

№	Маркер	Міра відповідності до терміну 'Визначні місця'
1.	'*ЮНЕСКО*'	0,5
2.	'*центр*Світов*спадщин*'	0,5
3.	'*головн*годинник*'	0,7
4.	'*оглядов*майданчик*'	0,7
5.	'*міжнар*організаці*«Орега Еуропа»*'	0,5
6.	'*театр*перлин*архітектур*'	0,7
7.	'*дзеркальн*зал*'	0,5

№	Маркер	Міра відповідності до терміну 'Харчування'
1.	'*кафе*'	0,9
2.	'*обід*'	0,5
3.	'*перекус*'	0,7
4.	'*відпочин*к*'	0,5
5.	'*вечер*в*підземел*театр*'	0,7
6.	'*Лів*бере*'	0,5

№	Маркер	Міра відповідності до терміну 'Розваги'
1.	'*фестиваль*'	0,7
2.	'*концерт*'	0,7
3.	'*екскурсі*'	0,7
4.	'*оперет*'	0,7
5.	'*опер*'	0,7
6.	'*балет*'	0,7
7.	'*конкурс*'	0,7

№	Маркер	Міра відповідності до терміну 'Інформація про персонал'
1.	'*касир*'	0,9
2.	'*мер*'	0,5
3.	'*екскурсовод*'	0,9
4.	'*сурмач*'	0,5
5.	'*актор*'	0,9
6.	'*диригент*'	0,5

Приклад індикативної ознаки дії над об'єктом туристичної документації 'Вихід на Ратушу'

№	Маркер	Міра відповідності до дії над об'єктом
1.	'*спогляд*оглядов*майданчик*'	0,8
2.	'*підніма*сход*'	1
3.	'*фотограф*огляд*майданчик*'	0,8
4.	'*загад*бажан*огляд*майданчик*'	0,5

Приклад індикативної ознаки дії над об'єктом туристичної документації 'Відвідування оперного театру'

№	Маркер	Міра відповідності до дії над об'єктом
1.	'*відвід*дзеркал*зал*'	0,9
2.	'*фотограф*'	0,5
3.	'*відвід*оперет*'	0,8
4.	'*екскурсі*'	0,5

Приклад індикативної ознаки досвіду споживача туристичної документації

№	Маркер	Міра відповідності до досвіду споживача
1.	'*я*бу*'	0,7
2.	'*ми*спогляд*краєвид*'	0,7
3.	'*задовол*'	0,5
4.	'*відвід*'	0,5
5.	'*краєвид*прекрасн*'	0,3
6.	'*дуже цікав*'	0,3
7.	'*ми*відвід*'	0,5
8.	'*вони*милу*'	0,5
9.	'*сподобал*'	0,5
10.	'*слух*'	0,5
11.	'*чудов*інтер'єр*'	0,6

Приклад індикативної ознаки часових складових туристичної документації

№	Маркер	Міра відповідності до часових складових
1.	'*минуло*ро*'	0,7
2.	'*вчора*'	0,7
3.	'*місяц*тому*'	0,7
4.	'*ввечері*'	0,3
5.	'*минуло*тиж*'	0,7
6.	'*сьогодні*'	0,7
7.	'*рік*тому*'	0,7
8.	'*зранку*'	0,3

Приклад індикативних ознак атрибутів елементів туристичної документації

№	Маркер	Міра відповідності до атрибута 'Заголовок'
1.	'<title>*</title>'	1

<b>№</b>	<b>Маркер</b>	<b>Міра відповідності до атрибута ‘Посилання’</b>
1.	‘<link>*/</link>’	1

<b>№</b>	<b>Маркер</b>	<b>Міра відповідності до атрибута ‘Опис’</b>
1.	‘<description>*/</description>’	1

<b>№</b>	<b>Маркер</b>	<b>Міра відповідності до атрибута ‘Мова’</b>
1.	‘<language>*/</language>’	1

<b>№</b>	<b>Маркер</b>	<b>Міра відповідності до атрибута ‘Час публікації’</b>
1.	‘<pubDate>*/</pubDate>’	1

<b>№</b>	<b>Маркер</b>	<b>Міра відповідності до атрибута ‘Автор’</b>
1.	‘<dc:creator>*/</dc:creator>’	1

## **Додаток В**

**Акти про впровадження результатів дисертаційних досліджень**

“ЗАТВЕРДЖУЮ”



*Терещук Н.С.*

« 15 » грудня 2016 р.

### АКТ

**про використання результатів дисертаційних досліджень**

**Марків Оксани Олександрівни**

**«Лінгвістичне забезпечення формування туристичної документації на основі  
відкритих Веб-ресурсів»**

Цей акт складений про те, що теоретичні результати дисертаційних досліджень Марків О.О. з формування туристичної документації за допомогою комп'ютерно-лінгвістичного аналізу інформаційного наповнення відкритих Веб-ресурсів, а саме:

- інформаційна модель туристичної документації та інформаційного джерела, що враховують лінгвістичні особливості їхніх елементів;
- методи комп'ютерно-лінгвістичного формування та верифікації туристичної документації за допомогою індикативних ознак;
- метод оцінювання якості туристичної документації на основі функціональних показників якості, показників якості придатності до використання та показників якості супроводжуваності за стандартом ISO/IEC-25010;


є використаними для:

- розробки процедур автоматизованого формування та верифікації туристичної документації шляхом комп'ютерно-лінгвістичного аналізу відкритих Веб-ресурсів;
- визначення напрямків підвищення якості туристичної документації як інформаційного продукту.

Це дозволило застосувати розроблені методи та засоби для автоматизованого формування каталогу туристичних послуг туристичної компанії «Королівські Подорожі». Зокрема це забезпечило:

- побудову системи моніторингу в реальному часі декількох сотень Веб-ресурсів з інформацією про туристичні об'єкти;
- налагодження системи підтримки актуальності та коректності інформації каталогу туристичних послуг, що дозволило оптимізувати використання людських ресурсів під час роботи з цим каталогом;
- створення системи контролю за якістю каталогу туристичних послуг.

Експерти:

Турівецький Назар 



“ЗАТВЕРДЖУЮ”

СВЯТОДІМІТРИЇВСЬКОЇ ЛІНГВІСТИЧНОЇ ШКОЛИ  
ОКСАНА  
ГРИГОРІВНА

558707840

“ 15 ” грудня 2016 р.

**АКТ**  
**про використання результатів дисертаційної роботи**  
**Марків Оксани Олександрівни**  
**«Лінгвістичне забезпечення формування туристичної документації на**  
**основі відкритих веб-ресурсів»**

Цей акт складений про те, що теоретичні результати дисертаційної роботи Марків О.О. «Лінгвістичне забезпечення формування туристичної документації на основі відкритих веб-ресурсів» у частині формальної моделі туристичної документації, формальної моделі інформаційного джерела для формування туристичної документації, методів лінгвістичного співставлення інформаційного наповнення відкритих веб-ресурсів з елементами туристичної документації та метод оцінювання якості туристичної документації є використаними під час функціонування сервісу надання туристичних послуг компанією «Тур де Львів». Використання теоретичних результатів дисертаційної роботи Марків О.О. дало змогу:

- збільшити обсяг корисної інформації про туристичні об'єкти у 2-3 рази залежно від видів туристичних послуг;
- організувати комплекс заходів щодо збору туристичної інформації, що дозволило досягнути значного скорочення затрат на консолідацію туристичної інформації з великої кількості веб-ресурсів;
- впровадити спеціальні засоби поглибленого аналізу відгуків споживачів туристичних послуг для планування розвитку туристичних послуг.

Експерти:

Мухомецька Трина



"Затверджую"  
Проректор з наукової роботи  
Національного університету  
"Львівська політехніка"  
проф. Чухрай Н.І.  
2016 р.



**А К Т**

про використання результатів дисертаційної роботи "Лінгвістичне забезпечення формування туристичної документації на основі відкритих веб-ресурсів" асистента кафедри інформаційних систем та мереж Марків Оксани Олександрівни, представленої на здобуття наукового ступеня кандидата технічних наук, при виконанні науково-дослідних робіт Національного університету "Львівська політехніка"

Ми, що нижче підписалися, начальник НДЧ, к.т.н., доц. Жук Л.В. та члени комісії: завідувач відділу науково-організаційного супроводу наукових досліджень, к.т.н. Лазько Г.В., заступник начальника планово-фінансового відділу Чулой Т.М. та зав. кафедри інформаційних систем та мереж, д.т.н., проф. Литвин В.В. цим актом підтверджуємо, що результати дисертаційної роботи асистента кафедри інформаційних систем та мереж Марків О.О. використано під час виконання науково-дослідних робіт кафедри інформаційних систем та мереж Національного університету "Львівська політехніка": "Методи та засоби консолідації баз даних в інформаційних системах електронного урядування" (2010/2012, № держреєстрації 0110U005022) та "Лінгвістичне забезпечення консолідації відкритих інформаційних ресурсів" (2013, № держреєстрації 0113U005274).

Марків О.О. розробила нові моделі та методи збору та верифікації туристичної інформації на основі комп'ютерно-лінгвістичного аналізу інформаційного наповнення з неоднорідних відкритих веб-ресурсів. Розроблені методи дали можливість автоматизовано формувати якісну туристичну документацію, проводити її перевірку і оцінювати її якість із врахуванням потреб споживача туристичного продукту.

Начальник НДЧ,  
канд. техн. наук, доц.

Л.В. Жук

Члени комісії:  
Зав. відділу НОСНД,  
канд. техн. наук

Г.В. Лазько

Заст. нач. ПФВ

Т.М. Чулой

Зав. кафедри  
Інформаційних систем та мереж,  
д.т.н. проф.

В.В. Литвин





"Затверджую"

Проректор з науково-педагогічної роботи

Національного університету "Львівська політехніка"

Давидчак О.Р.

2016 р.

## А К Т

про впровадження результатів кандидатської дисертаційної роботи асистента Марків Оксани Олександрівни на тему "Лінгвістичне забезпечення формування туристичної документації на основі відкритих веб-ресурсів" у навчальний процес на кафедрі інформаційних систем та мереж Національного університету "Львівська політехніка"

Даним актом засвідчується, що наукові та практичні результати дисертаційної роботи асистента кафедри інформаційних систем та мереж Марків О.О. на тему "Лінгвістичне забезпечення формування туристичної документації на основі відкритих веб-ресурсів" впроваджено в навчальний процес кафедри інформаційних систем та мереж Національного університету "Львівська політехніка", а саме:

- підходи до структурування інформаційного наповнення відкритих веб-ресурсів, методи інтелектуального аналізу текстів та методи побудови штучних мов маркування текстів для розроблення процедур лінгвістичного співставлення інформаційного наповнення відкритих веб-ресурсів з елементами туристичної документації використано в лекційному курсі навчальної дисципліни "Методи та засоби інженерії даних та знань" для студентів 1-го курсу освітньо-кваліфікаційного рівня "магістр" (спеціальність 122 – "Комп'ютерні науки та інформаційні технології", спеціалізація – "Системи штучного інтелекту") у таких темах:
  - тема № 1: Методи видобування знань з даних та текстів. Теоретичні аспекти видобування знань. Психологічний аспект. Лінгвістичний аспект. Гносеологічний аспект видобування знань. Методи структурування. Еволюція систем одержання знань;
  - тема № 4: Методи класифікації і систематизації знань. Теоретичні аспекти структурування знань. Ієрархічний підхід. Традиційні методології структуризації. Об'єктно-структурний підхід;
- підходи до використання відкритих веб-ресурсів для формування туристичної документації шляхом аналізу методів інтеграції інформаційних ресурсів і моделювання відкритих веб-ресурсів, а також аналіз підходів до опису метаданих та структурування інформації в XML-технологіях використано в лекційному курсі навчальної дисципліни "Методи та засоби інтеграції даних" для студентів 1-го курсу освітньо-кваліфікаційного рівня "магістр" (спеціальність 122 – "Комп'ютерні науки та інформаційні технології", спеціалізація – "Системи штучного інтелекту") у таких темах:
  - тема № 2: Технології та методи інтеграції даних (Інтеграція даних: глобальний (Global As View, GAV) та локальний (Local As View, LAV) методи. Підходи до інтеграції даних (консолідація, федералізація, поширення, гібридна та сервісна);
  - тема № 3: Технологія ETL (Extract, Transform and Load) (Структура процесу перевантаження даних (процес, фаза, крок, група процесів, стадії завантаження джерела даних, витягання даних, отримання (вивантаження) даних,

структуризація даних, опрацювання даних, опрацювання помилок, очищення даних). Підпроцеси STER та STAC. Переваги та недоліки ETL).

Розроблені Марків О.О. підходи до обробки туристичної інформації у відкритому доступі за допомогою інтелектуального аналізу текстів веб-ресурсів, а також застосування підходів до опису метаданих дають студентам можливість: оволодіти основними підходами до інтеграції відкритих веб-ресурсів; вивчити принципи формування моделей інформаційних джерел; ознайомитися з основними підходами до опису метаданих та структурування інформації.

Директор  
Інституту комп'ютерних наук  
та інформаційних технологій,  
д.т.н. проф.



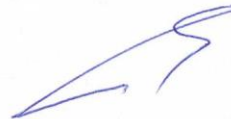
М.О. Медиковський

Зав. кафедри  
Інформаційних систем та мереж,  
д.т.н. проф.



В.В. Литвин

Лектор навчальної дисципліни  
“Методи та засоби інженерії  
даних та знань”  
д. т. н, проф.



В.В. Литвин

Лектор навчальної дисципліни  
“Методи та засоби інтеграції даних”  
к. т. н, доц.



Т.М. Басюк