

КОМП'ЮТЕРНА ІНЖЕНЕРІЯ

УДК 004.42.519.8

DOI <https://doi.org/10.32782/2521-6643-2024-1-67.9>

Пашковський Б. В., кандидат технічних наук,
доцент кафедри комп'ютерних систем і мереж
Івано-Франківського національного технічного
університету нафти і газу
ORCID: 0000-0003-1082-6837

РОЗРОБКА СИСТЕМИ ОЦИФРУВАННЯ ТА КАТАЛОГІЗАЦІЇ ПРОТОКОЛІВ КОМІСІЇ З БЕЗПЕКИ ДОРОЖНЬОГО РУХУ ІВАНО-ФРАНКІВСЬКОЇ МІСЬКОЇ ТЕРИТОРІАЛЬНОЇ ГРОМАДИ

Стаття присвячена розробці системи оцифрування та каталогізації протоколів та рішень комісії з безпеки руху Івано-Франківської міської територіальної громади (КБДР).

КБДР розглядає питання щодо реалізації заходів для забезпечення робіт з благоустрою, організації дорожнього руху та підвищення безпеки всіх учасників руху на території Івано-Франківської міської територіальної громади відповідно до генерального плану, проєктів детального планування та забудови населених пунктів, проєктів капітального ремонту, реконструкції та нового будівництва об'єктів вулично-дорожньої мережі, впровадження автоматизованих систем керування дорожнім рухом, комплексних транспортних схем і схем організації дорожнього руху, зон паркування з врахуванням принципів сталого розвитку та дотримання правил благоустрою.

З 2016 по 2024 роки комісією проведено більше ста засідань, де розглянуто більше 3150 питань. Протоколи комісії оформляються у вигляді PDF-файлів та завантажуються на сервер. Існує гостра необхідність у пошуку рішень комісії за певною адресою, розташуванням, тощо. Звернення громадян, муніципальної інспекції потребують формування відповідей та витягів чи було певне рішення прийняте чи ні. У статті розглянуто розробку системи оцифрування протоколів комісії та створення каталогу (веб сайту із базою даних) із функцією швидкого пошуку. Забезпечено можливість завантаження оригінального протоколу для формування офіційного витягу.

Спроектовано та реалізовано базу даних засобами Microsoft SQL Server з пакетом Entity Framework Core для зберігання рішень КБДР.

Розроблено систему оцифрування PDF-файлів рішення комісії та імпорту їх до бази даних.

*Побудовано веб застосунок із функцією пошуку з використанням технологій ASP.NET Core MVC та Razor Views.
Ключові слова: оцифрування, веб-застосунок, база даних.*

Pashkovskiy B. V. Development of a system of digitization and cataloging of protocols of the traffic safety commission of the Ivano-Frankivsk city-territorial community

The article is devoted to the development of a system of digitization and cataloging of protocols and decisions of the Traffic Safety Commission of the Ivano-Frankivsk City Territorial Community (KBDR).

KBDR considers the issue of implementation of measures to ensure improvement works, organization of road traffic and increase the safety of all road users on the territory of Ivano-Frankivsk urban territorial community in accordance with the master plan, projects of detailed planning and construction of settlements, projects of capital repair, reconstruction and new construction of objects of the street and road network, implementation of automated traffic control systems, complex transport schemes and traffic organization schemes, parking zones taking into account the principles of sustainable development and compliance with the rules of landscaping.

From 2016 to 2024, the commission held more than a hundred meetings where more than 3,150 issues were considered. Commission protocols are drawn up in the form of PDF files and uploaded to the server. There is an urgent need to find decisions of the commission at a certain address, location, etc. Appeals from citizens, municipal inspection requires the formation of answers and excerpts whether a certain decision was made or not. The article deals with the development of a system for digitizing commission protocols and creating a catalog (a website with a database) with a quick search function. It is possible to download the original protocol to form an official extract.

The database was designed and implemented using Microsoft SQL Server with the Entity Framework Core package for storing KBDR solutions.

A system for digitizing PDF files of the commission's decision and importing them into the database has been developed. A web application with a search function was built using ASP.NET Core MVC technology and Razor Views.

Key words: digitization, web application, database.

Постановка проблеми. Як було зазначено, протоколи засідань комісії з безпеки руху оформляються у вигляді PDF-файлів та завантажуються на сервер. Оскільки PDF є неструктурованим документом, існують проблеми пошуку рішень прийнятих комісією:

Для пошуку рішення, що містить певну адресу чи інший текст, необхідно відкрити кожен файл (більше 100 файлів) та провести пошук всередині файлу, що займає великий об'єм часу.

Цифрові документи мають безліч переваг, особливо в порівнянні з традиційними паперовими документами.

Цифрові документи можна зберігати на електронних пристроях, таких як комп'ютери, хмарні сервіси, флеш-накопичувачі і т.д. Це дозволяє зберігати велику кількість даних на невеликому просторі і швидко знаходити необхідну інформацію завдяки можливості пошуку. Цифрові документи можна легко редагувати без потреби друкувати або використовувати складні процеси. Це дозволяє зручно вносити зміни до документів і забезпечує їх актуальність. З цифрових документів можна отримати доступ у будь-який час і з будь-якого місця за допомогою Інтернету або зберігаючи їх на портативних пристроях. Це забезпечує швидкий і зручний доступ до інформації, що може бути важливим у сучасному світі. Використання цифрових документів сприяє зменшенню споживання паперу та інших ресурсів, що допомагає зберегти навколишнє середовище. Цифрові документи можна легко обмінювати електронно через електронну пошту, файлообмінні сервіси, месенджери тощо. Це зменшує час, потрібний для передачі інформації, і спрощує комунікацію між сторонами.

Аналіз останніх досліджень та публікацій. У роботі [1] автори проводять порівняння різного програмного забезпечення, що використовує технологію OCR (Optical Character Recognition – технологія оптичного розпізнавання символів). Ці програми дозволяють за декілька хвилин розпізнати відскановане зображення, PDF-документ чи цифрову фотографію, вилучити інформацію з них, та перетворити до формату, який буде зручним для редагування даних.

Авторка роботи [2] розглядає особливості збереження книжкових бібліотечних видань. Зокрема звертається увага на основні принципи й методи оцифрування друкованих документів, визначаються типові функції електронного фонду Запорізької обласної універсальної наукової бібліотеки.

Робота [3] є комплексним дослідженням особливостей оцифрування фондів бібліотек України. Запропоновано стратегічні напрями та базові організаційно методичні засади вдосконалення організації оцифрування та формування цифрових колекцій у бібліотеках України. Розроблено інноваційні моделі організації оцифрування документної спадщини України. Здійснено компаративний аналіз закордонного та українського досвіду з реалізації проєктів оцифрування документної спадщини.

Автори [4] розглядають парсинг PDF-документів та пропонують рішення, обмеження синтаксису, задля уникнення типових помилок, і пропонують використання формальних граматик. Також пояснюється, як узгодженість даних можна перевірити на більш детальному рівні за допомогою спеціального засобу перевірки типів. Підхід валідується на наборі реальних файлів.

Метою роботи є створення системи для швидкого пошуку по протоколах КБДР. Для досягнення мети необхідно вирішити наступні завдання:

- створити модель бази даних для збереження протоколів та питань, що розглядалися;
- створити програмне забезпечення для оцифрування PDF-файлів та занесення інформації до бази даних;
- створити програмне забезпечення (веб сайт) для відображення протоколів, питань, що розглядалися та реалізації функції пошуку по них.

Виклад основного матеріалу. Розглянемо типову структуру PDF-документу, що містить протокол засідання КБДР (рис. 1).

Як видно із рисунку, зазвичай питання поділяється на дві частини: «слухали» та «вирішили», а номер протоколу записаний у колонтитулах.

Для зберігання протоколу та його питань змодельємо дві сутності: Protocol та Item, які будуть реалізовувати відношення один до багатьох.

```
public class Protocol
{
    public int Id { get; set; }

    public string Number { get; set; }

    public DateTime Date { get; set; }

    public List<Item> Items = new();
}

public class Item
{
```

Розгляд питань:

1. СЛУХАЛИ: Питання щодо аварійності на вулицях міста Івано-Франківська та дорогах ОТГ за період з 2018 по 2023 рік за даними Управління патрульної поліції. Інфографіка, онлайн карта та статистичні дані підготовлені Богданом Пашковським.

ВИРШИЛИ: Прийняти звіт. Доручити Департаменту інфраструктури, житлової та комунальної політики міської ради врахувати подану інформацію та підготувати план заходів щодо вдосконалення дорожніх умов, організації дорожнього руху, ліквідації причин виникнення ДТП, забезпечення безпеки дорожнього руху, а також надати інформацію про аварійно-небезпечні ділянки та місця концентрації ДТП у відділ безпеки дорожнього руху Управління патрульної поліції до 15 березня 2024 року.

2. СЛУХАЛИ: Звернення Директора Головного сервісного центру МВС Олега Шпека щодо заміни дорожнього знаку 3.3 «Рух вантажного транспорту заборонено» із зазначенням максимально допустимої маси 20 т.

ВИРШИЛИ: погодити демонтаж даного знаку. Виконання рішення доручити КП «Муніципальна дорожня компанія».

3. СЛУХАЛИ: Звернення Директора «Тойота-Центр» Угрицького В. П. щодо прийняття рішення про облаштування світлофорного об'єкту на перехресті вулиць Надрічної та Богдана Хмельницького.

ВИРШИЛИ: Зобов'язати компанію «Тойота-Центр» замовити розробку схеми організації дорожнього руху з облаштуванням світлофорного об'єкту, та затвердити її в Управлінні патрульної поліції в установленому порядку.

4. СЛУХАЛИ: Звернення ТзОВ ФРАНКОФ ТРЕЙД Чорномаза В. Р. щодо внесення змін до схеми ОДР по вул. Надрічній з наданням дозволу на здійснення лівого повороту при заїзді на прилеглу територію торгового центру, і при виїзді з нього.

ВИРШИЛИ: Доручити Департаменту інфраструктури, житлової та комунальної політики Івано-Франківської міської ради врахувати дану пропозицію при виготовленні нової схеми організації дорожнього руху на вул. Надрічній в м. Івано-Франківську. ТзОВ «ФРАНКОФ ТРЕЙД» зобов'язати розробити схему організації руху та паркування на прилеглій території, та затвердити її в установленому порядку у Відділі безпеки дорожнього руху в Управлінні патрульної поліції в Івано-Франківській області.

Виконавчий комітет Івано-Франківської міської ради
вул. М.Сидоренка, 21, м. Івано-Франківськ, 76004, тел. 556515, 552038, e-mail: info@mk.if.ua, www.mk.if.ua
ПРОТОКОЛ №1 від 09.02.2024 року
засідання комісії з безпеки дорожнього руху
Івано-Франківської міської територіальної громади
сторінка 5 з 10

Рис. 1. Частина типового протоколу КБДР

```
public int Id { get; set; }  
public int? Number { get; set; }  
public int OrderNumber { get; set; }  
public string Heard { get; set; }  
public string Decided { get; set; }  
public Protocol Protocol { get; set; }  
}
```

Поле OrderNumber відповідатиме за порядок розгляду питання (інколи питання розглядаються в іншому порядку, ніж були представлені в порядку денному).

Розглянемо таблиці та діаграму бази даних (рисунки 2-4).

Column Name	Data Type	Allow Nulls
Id	int	<input type="checkbox"/>
Number	nvarchar(MAX)	<input type="checkbox"/>
Date	datetime2(7)	<input type="checkbox"/>

Рис. 2. Дизайн таблиці Protocols

Column Name	Data Type	Allow Nulls
Id	int	<input type="checkbox"/>
Number	int	<input checked="" type="checkbox"/>
Heard	nvarchar(MAX)	<input type="checkbox"/>
Decided	nvarchar(MAX)	<input type="checkbox"/>
ProtocolId	int	<input type="checkbox"/>
OrderNumber	int	<input type="checkbox"/>

Рис. 3. Дизайн таблиці Items

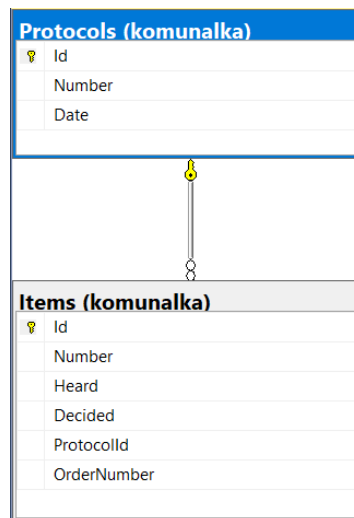


Рис. 4. Діаграма бази даних

Для визначення номеру і дати протоколу використаємо регулярні вирази.

В програмуванні, регулярний вираз (від англ. regular expression, скорочено regex або regexp, а іноді ще й називають rational expression)— це рядок, що описує або збігається з множиною рядків, відповідно до набору спеціальних синтаксичних правил. Вони використовуються в багатьох текстових редакторах та допоміжних інструментах для пошуку та зміни тексту на основі заданих шаблонів. Багато мов програмування підтримують регулярні вирази для роботи з рядками.

Використаємо наступний регулярний вираз, який застосуємо до першої сторінки: «ПРОТОКОЛ \u2116\s*(\d+)\s*від (\d+\.\d+\.\d+)\s*року».

Для визначення питань у протоколі слід спершу видалити колонтитули з кожної сторінки. Для цього використовується регулярний вираз «(?:_+|Виконавчий) .*? \d+ з \d+».

Будемо використовувати слова «слухали» та «вирішили» як орієнтири для визначення питань у протоколі.

Для кращого розуміння алгоритму роботи, лістинг коду, що відповідає за даних функціонал, надамо повністю.

```

public static List<Item> GetProtocolItems(string filePath)
{
    using var pdf = PdfDocument.Open(filePath);
    var text = string.Empty;

    foreach (var page in pdf.GetPages())
    {
        text += page.Text;
    }

    var firstPage = pdf.GetPages().FirstOrDefault();
    text += ContentOrderTextExtractor.GetText(firstPage);

    text = RemoveFooters(text);

    var itemRegex = new
    Regex(«СІУХАЛ?И?:?\s*(.*?)\s*\d*\.?ВИРШИЛ?И?:?\s*(.*?)\s*\d*\.*\s*?(?=СІУ-
    ХАЛ|Голов|Заступ)»);

    var itemMatches = itemRegex.Matches(text);

    var items = new List<Item>();

    for (int i = 0; i < itemMatches.Count; i++)
    {
        var item = new Item
        {
            Number = i + 1,
            Heard = itemMatches[i].Groups[1].Value,
            Decided = itemMatches[i].Groups[2].Value
        };

        items.Add(item);
    }

    return items;
}

```

Для можливості відображення протоколів побудуємо сайт з використанням ASP .NET MVC та Razor Views. Головна сторінка міститиме перелік протоколів, відсортованих від найновішого до найстарішого, із номером, датою, кількістю розглянутих питань і посиланням на перелік питань (рис. 5).

Протоколи КБДР

Протокол №	Дата	Кількість питань	
1	09.02.2024	27	Перейти до питань
10	10.11.2023	18	Перейти до питань
9	03.11.2023	38	Перейти до питань
8	08.09.2023	36	Перейти до питань
7	25.08.2023	27	Перейти до питань
6	23.06.2023	26	Перейти до питань
4	12.05.2023	24	Перейти до питань
5	06.05.2023	22	Перейти до питань

Рис. 5. Головна сторінка сайту

При переході за посиланням «перейти до питань» ми потрапляємо на сторінку із деталями протоколу, де бачимо номер, дату протоколу, посилання на завантаження вихідного PDF-файлу та перелік питань: номер, слухали, та вирішили (рис. 6).

Протоколи КБДР

Протокол №1 від 09.02.2024 року
[Завантажити вихідний файл протоколу](#)

Номер	Слухали	Слухали
1	Питання щодо аварійності на вулицях міста Івано-Франківська та дорогах ОТГ за період з 2018 по 2023 рік за даними Управління патрульної поліції. Інфографіка, онлайн карта та статистичні дані підготовлені Богданом Пашковським.	Прийняти звіт. Доручити Департаменту інфраструктури, житлової та комунальної політики міської ради врахувати подану інформацію та підготувати план заходів щодо вдосконалення дорожніх умов, організації дорожнього руху, ліквідації причин виникнення ДТП, забезпечення безпеки дорожнього руху, а також надати інформацію про аварійно-небезпечні ділянки та місця концентрації ДТП у відділ безпеки дорожнього руху Управління патрульної поліції до 15 березня 2024 року
2	Звернення Директора Головного сервісного центру МВС Олега Шпека щодо заміни дорожнього знаку 3.3 «Рух вантажного транспорту заборонено» із зазначенням максимально допустимої маси 20 т.	погодити демонтаж даного знаку. Виконання рішення доручити КП «Муніципальна дорожня компанія»
3	Звернення Директора	Зобов'язати компанію «Тойота-Центр»

Рис. 6. Сторінка деталей протоколу

Використавши форму для пошуку, можемо знайти усі питання, які містять часткове входження шуканого слова, протоколи у результаті пошуку відсортовані від найновішого до найстарішого (рис. 7).

Протоколи КБДР

Знайдено: 185 питань

Протокол №	Номер	Слухали	Вирішили
Протокол №1 від 09.02.2024	9	Звернення жителів будинку щодо дозволу на встановлення автоматичного шлагбауму на вул. Незалежності , 81 (в'їзд із вул. Лепкого).	залишити без розгляду
Протокол №1 від 09.02.2024	26	Звернення Муніципальної Інспекції "Добродій" про демонтаж незаконно встановлених ТМ, МАФ та інших елементів.	Доручити КП "Муніципальна інспекція "Добродій" демонтувати самовільно встановленні засоби обмеження руху у відповідності до актів: № 308 – вул. Мазепи, 146 (біля Парк Готелю) № 143- вул. Незалежності , 160 № 358 – вул. Грушевського, 22А № 414 – вул. Крива, 2 № 583 – вул. Довженка, 1А-3А № 4933 - вул. Короля Данила, 20, 20А № 5347 – вул. Мазепи, 35 Б № 3613 – вул. Мазепи. 42 № 4494- вул.

Рис. 7. Сторінка пошуку

Висновки. Розроблено систему оцифрування протоколів комісії та створення каталогу (вебсайту із базою даних) із функцією швидкого пошуку. Забезпечено можливість завантаження оригінального протоколу для формування офіційного витягу.

Створено з використанням технології ASP .NET Core MVC та Razor Views, як база даних використано Microsoft SQL Server з пакетом Entity Framework Core.

Сайт доступний за веб-адресою <https://kodlo.if.ua/> та широко використовується для швидкого пошуку рішень КБДР.

Список використаних джерел:

1. Дольнев, Р. О. Проблема оцифрування паперових документів. Тези доповідей Міжнародної науково-практичної конференції молодих учених, аспірантів та студентів «Інформаційні технології в сучасному світі: дослідження молодих вчених» 20 – 21 лютого 2020 р. Харківський національний університет радіоелектроніки, Харків. 2020. С. 31.

2. Горбенко, І. Оцифрування як засіб збереження книжкових видань бібліотек. Історія та сучасні тенденції розвитку видавничої справи. 2015. №8. С. 129-133

3. Шевченко, М. О. Оцифрування фондів бібліотек України. Дисертація на здобуття наукового ступеня доктора філософії. 2022. С. 2-4.

4. Endignoux, G., Levillain, O., & Migeon, J. Y.. Caradoc: A pragmatic approach to pdf parsing and validation. Security and Privacy Workshops (SPW). IEEE. 2016 ст. 126-139.

5. Вікіпедія: сайт URL [https://uk.wikipedia.org/wiki/Регулярний вираз](https://uk.wikipedia.org/wiki/Регулярний_вираз)

References:

1. Dolnev, R. O (2020, February). The problem of digitizing paper documents. Abstracts of reports of the International scientific and practical conference of young scientists, postgraduates and students «Information technologies in the modern world: research of young scientists» February 20 - 21, 2020. Kharkiv National University of Radio Electronics, Kharkiv. 31.

2. Horbenko, I. (2015). Digitization as a means of preserving book editions of libraries. History and modern trends in the development of publishing. 8, 129-133

3. Shevchenko, M. O. (2022). Digitization of library collections of Ukraine. Dissertation for the Doctor of Philosophy degree. 2-4.

4. Endignoux, G., Levillain, O., & Migeon, J. Y. (2016, May). Caradoc: A pragmatic approach to pdf parsing and validation. In 2016 IEEE Security and Privacy Workshops (SPW) (pp. 126-139). IEEE.

5. Regular expression. https://en.wikipedia.org/wiki/Regular_expression