

ПІДХІД ДО ЗБОРУ ТА ВІЗУАЛІЗАЦІЇ ДАНИХ ДЕРЖАВНОГО МОНІТОРИНГУ СТАНУ АТМОСФЕРНОГО ПОВІТРЯ МІСТА ВІННИЦІ

¹Вінницький національний технічний університет

Анотація

У роботі представлено результати розроблення системи збору та візуалізації даних державного моніторингу стану атмосферного повітря міста Вінниці. Запропоновано рішення з використанням технологій оптичного розпізнавання тексту для набору зображень, що містять дані моніторингу за 2025 рік. Розроблено веб-систему. із модулем перетворення зображень, що містять таблицю показників забрудненості повітря, у набір даних, придатних до подальшої обробки. Також розроблено автоматизований скрапер, що дозволяє зчитувати дані моніторингу із офіційного сайту, де вони розташовуються тимчасово для оприлюднення; зчитані дані моніторингу зберігаються та накопичуються в єдиній базі даних системи.

Ключові слова: екологічний моніторинг, якість повітря, інформаційна система, аналіз зображень.

Abstract

This work presents the results of the development of a system for collecting and visualizing data for state monitoring of the state of atmospheric air in the city of Vinnytsia. A solution using optical text recognition technologies is proposed for a set of images containing monitoring data for 2025. A web system has been developed. with a module for converting images containing a table of air pollution indicators into a set of data suitable for further processing. An automated scraper has also been developed that allows reading monitoring data from the official website, where they are temporarily located for publication; the read monitoring data is stored and accumulated in a single database of the system.

Keywords: ecological monitoring, air quality, information system, image processing.

Вступ

Сучасні інформаційні системи екологічного моніторингу забезпечують доступ до поточних та історичних даних про стан навколишнього середовища. Є і залишається актуальним автоматизований збір і обробка даних про забруднення атмосферного повітря, оскільки така інформація є необхідною для органів влади, науковців та населення. Водночас значна частина відкритих екологічних даних публікується у вигляді табличної інформації на веб-сайтах або у форматі зображень, що ускладнює їх подальше машинне опрацювання та аналіз.

Однією з практичних задач є видобування структурованих даних із зображень таблиць, отриманих шляхом знімків екрана з офіційних веб-ресурсів. Такий підхід часто використовується у випадках, коли відсутні програмні інтерфейси доступу до даних або коли формат подання інформації не передбачає прямого експорту. Додаткові труднощі виникають через похибки оптичного розпізнавання тексту, різні формати значень, а також неточності у написанні.

Метою даної роботи є розробка інформаційної системи для автоматизованого збору, обробки, зберігання та візуалізації даних державного моніторингу атмосферного повітря на основі зображень таблиць та відкритих веб-джерел. Для досягнення поставленої мети необхідно розробити алгоритм обробки табличних зображень, а також спроектувати структуру бази даних для збереження результатів вимірювань. Окрему увагу приділено питанням подальшої візуалізації отриманих даних у зручному для користувача вигляді.

Результати дослідження

У якості вхідних даних отримано набір знімків екрана із сайту Вінницького обласного центру з гідрометеорології [1]. У таблиці усі клітинки мають форму прямокутника, а також відсутні їх об'єднання. Подібні зображення обробляються за наступним алгоритмом:

1. Визначити ділянки зображення, які містять клітинки таблиці.
2. Скопіювати та зберегти кожну ділянку у вигляді масиву пікселів.
3. Для кожної збереженої ділянки отримати рядок тексту, який вона містить.
4. Отримані дані структурувати у вигляді двовимірного масиву рядків. Текст, який не вдалося розпізнати позначити відповідним чином.

Контроль забруднення повітря у місті Вінниці							
Середні концентрації забруднювальних речовин в повітрі міста Вінниці в кратності до гранично-допустимих концентрацій (ГДК):							
Дата	Завислі речовини	Діоксид сірки	Оксид вуглецю	Діоксид азоту	Фтористий водень	Аміак	Формальдегід
11.10.2025	0,0	0,1	0,3	1,4	0,7	0,1	0,0
10.10.2025	0,0	0,0	0,3	1,4	1,8	0,2	0,0
09.10.2025	0,0	0,1	0,4	1,9	0,0	0,1	0,0
08.10.2025	0,0	0,0	0,3	1,2	3,0	0,0	4,0
07.10.2025	0,0	0,0	0,3	1,1	0,8	0,2	0,0
06.10.2025	0,5	0,1	0,4	1,5	1,4	0,4	0,7
04.10.2025	0,0	0,0	0,3	1,5	1,9	0,4	1,3
03.10.2025	0,3	0,0	0,3	0,9	0,0	0,7	1,3
02.10.2025	0,0	0,0	0,4	1,2	0,1	0,7	0,0
01.10.2025	0,0	0,0	0,3	2,7	1,9	0,3	1,3

Рис. 1. Приклад вхідного зображення.

Якщо результат розпізнавання тексту виявиться неточним, порівняння рядків на точну відповідність дасть хибний результат. Є підстави застосувати наближену відповідність рядків для пошуку речовини, назва якої найкраще відповідає заголовку стовпця. На етапі перетворення тексту у дату може виникнути проблема різних форматів запису, наприклад ДСТУ 4163, ISO 8601 тощо. Для кожного випадку доцільно створити шаблон та перевіряти запис на відповідність кожному способу. Аналогічно можна працювати із числами з рухомою комою. Вони можуть мати кому або крапку у якості десяткового розділювача (рис. 1). Аналогічним способом можна обробляти дані безпосередньо з сайту, якщо періодично завантажувати <http> сторінку та видобувати потрібні значення. У такому випадку оптичне розпізнавання тексту та меж таблиць не потрібне, але все одно залишається проблема наближеного порівняння рядків [2].

Середньодобові показники (Оксид вуглецю), мг/м^3

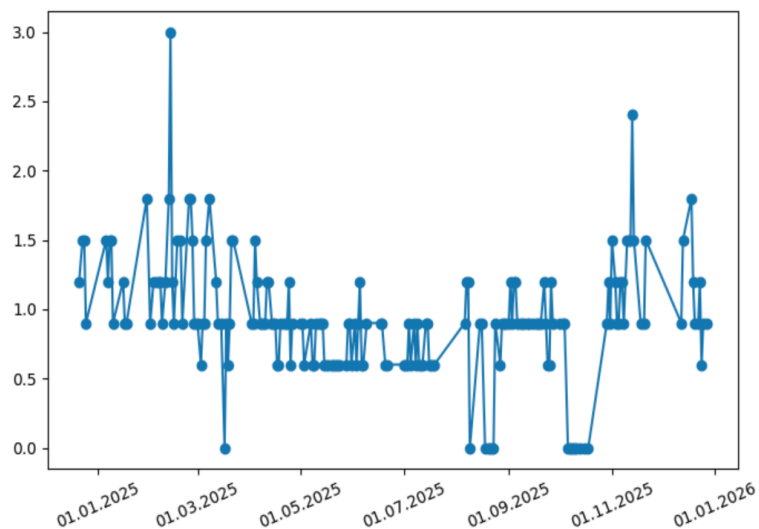


Рис. 2. Приклад візуалізації даних моніторингу стану атмосфери за показником – оксид вуглецю.

Середньодобові показники (Формальдегід), мг/м³

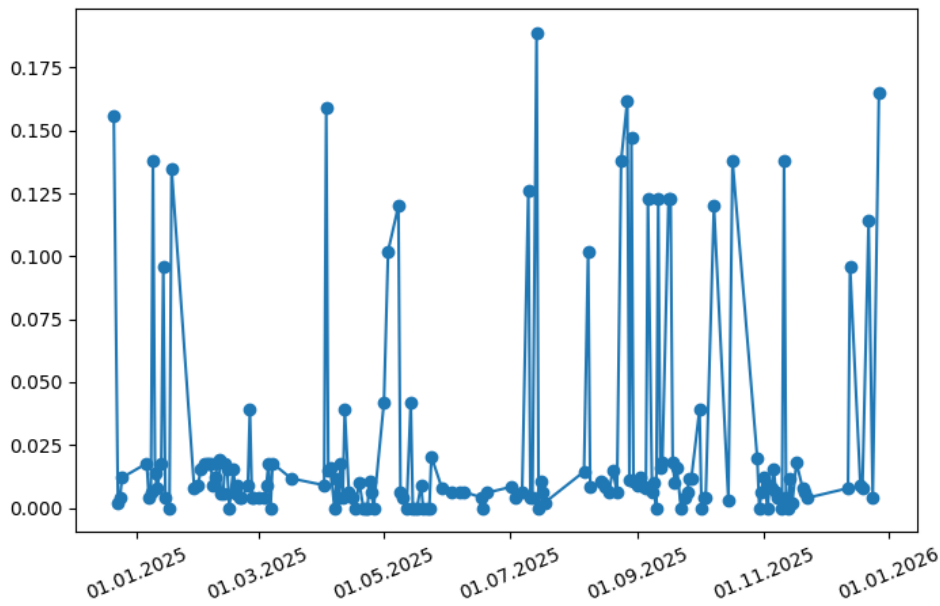


Рис 3. Приклад візуалізації даних моніторингу стану атмосфери за показником – формальдегід.

Інформаційна система повинна мати змогу збирати, обробляти, зберігати та візуалізувати дані, отримані від користувачів (рис. 2). Для збору даних користувачів, доцільно надати інтерфейс, який матиме кінцеві точки отримання даних у вигляді csv-таблиць та JSON-об’єктів. Виходячи із змісту (рис. 1), доцільно створити таблицю відомих речовин, де будуть зберігатися назви відомих забруднювачів, та таблицю вимірів, де будуть дати та значення середніх концентрацій. Оскільки дані у таблиці наведені у кратності до гранично допустимих концентрацій, варто створити таблицю ГДК відповідних речовин та дату настання чинності відповідного закону або постанови. При отриманні даних із бази, користувачу буде надіслано добуток відносних показників вимірювань на відповідні ГДК, які були чинні на момент вимірювань (рис. 3). Дані показників у мг/м³ також придатні до візуалізації за допомогою графіків та не потребують знань місцевих стандартів.

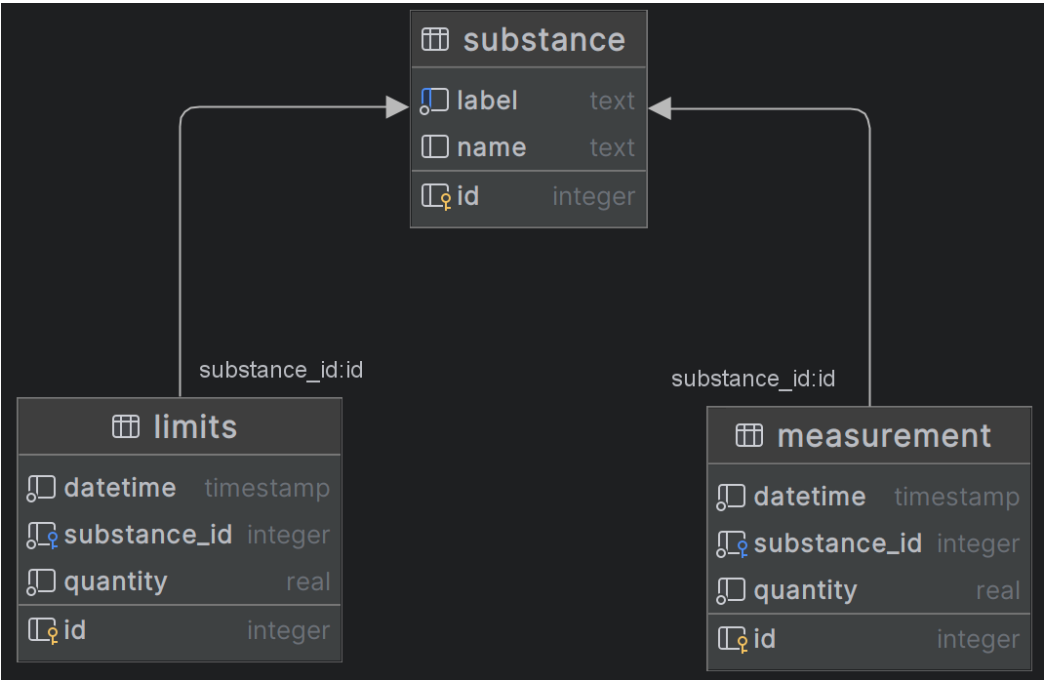


Рис 3. Структура таблиць бази даних системи

Реалізацію алгоритму виконано за допомогою мови Python [3]. Поділ на клітинки, оптичне розпізнавання тексту, наближену відповідність рядків та візуалізацію даних реалізовано з використанням відповідних бібліотек [4-7]. Розроблення веб-системи здійснено за допомогою фреймворка Flask та СУБД SQLite. Після обробки усіх вхідних зображень, усі зібрані дані завантажено до системи та розміщено за посиланням: <https://somerandomsymbols.pythonanywhere.com>.

Висновки

Розроблена інформаційна система дозволяє виконувати автоматичний збір, збереження та даних державного моніторингу стану атмосферного повітря міста Вінниці. Також, розроблено автоматизований скрапер, що дозволяє зчитувати дані моніторингу із офіційного сайту, де вони розташовуються тимчасово для оприлюднення; зчитані дані моніторингу зберігаються та накопичуються в єдиній базі даних системи.

Подальший розвиток системи може включати розширення функціоналу, наприклад: визначення та вилучення статистичних викидів [9], спричинених некоректними даними зі станцій моніторингу або помилками оптичного розпізнавання тексту, автоматичний збір та збереження даних з інших відкритих джерел, розроблення мобільного додатку для оперативного повідомлення про стан повітря.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Вінницький обласний центр з гідрометеорології [Електронний ресурс] // Офіційний сайт Вінницького обласного центру з гідрометеорології. – Режим доступу: <https://meteo.vn.ua> (дата звернення: 28.01.2026). – Назва з екрана.
2. Контроль забруднення повітря у місті Вінниця [Електронний ресурс] // Офіційний сайт Вінницького обласного центру з гідрометеорології. – Режим доступу: <https://meteo.vn.ua/activity/pollution/> (дата звернення: 28.01.2026). – Назва з екрана.
3. Мокін Б. І. Методи та засоби комп'ютерних обчислень : навч. посіб. / Борис Іванович Мокін, Віталій Борисович Мокін, Олександр Борисович Мокін. – Вінниця : ВНТУ, 2024. – 155 с.
4. img2table [Електронний ресурс] // PyPI. – Режим доступу: <https://pypi.org/project/img2table/> (дата звернення: 28.01.2026). – Назва з екрана.
5. easyocr [Електронний ресурс] // PyPI. – Режим доступу: <https://pypi.org/project/easyocr/> (дата звернення: 28.01.2026). – Назва з екрана.
6. RapidFuzz [Електронний ресурс] // PyPI. – Режим доступу: <https://pypi.org/project/RapidFuzz/> (дата звернення: 28.01.2026). – Назва з екрана.
7. matplotlib [Електронний ресурс] // PyPI. – Режим доступу: <https://pypi.org/project/matplotlib/> (дата звернення: 28.01.2026). – Назва з екрана.
8. Повітря у м. Вінниця [Електронний ресурс]. – Режим доступу: <https://somerandomsymbols.pythonanywhere.com/> (дата звернення: 28.01.2026). – Назва з екрана.
9. Системний аналіз стану атмосферного повітря регіону, з урахуванням впливу аномалій [Текст] : моногр. / Д. О. Шмундяк, В. Б. Мокін, Є. М. Крижановський ; Вінниц. нац. техн. ун-т. - Електронний ресурс. - Вінниця : ВНТУ, 2025. - 169 с.

Швець Михайло Владиславович – аспірант кафедри системного аналізу та інформаційних технологій, факультет інтелектуальних інформаційних технологій та автоматизації, Вінницький національний технічний університет, Вінниця, e-mail: misha.shvets02@gmail.com

Євгеній Миколайович Крижановський – канд. техн. наук, доцент кафедри системного аналізу та інформаційних технологій, Вінницький національний технічний університет, Вінниця, e-mail: kruzhan@gmail.com

Shvets Mykhailo V. – postgraduate student of the Department of System Analysis and Information Technologies, Faculty of Intelligent Information Technologies and Automation, Vinnytsia National Technical University, Vinnytsia, e-mail: misha.shvets02@gmail.com

Kryzhanovsky, Evgeniy M. – candidate of technical sciences, associate professor of the Department of System Analysis and Information Technologies, Vinnytsia National Technical University, Vinnytsia, e-mail: kruzhan@gmail.com