

## ДОСЛІДЖЕННЯ СТЕКУ ТЕХНОЛОГІЙ ДЛЯ ПОБУДОВИ ІНФОРМАЦІЙНОЇ СИСТЕМИ ДЛЯ НАКОПИЧЕННЯ ТА АНАЛІЗУ ДАНИХ ГАЗОСПОЖИВАННЯ

UDC 621.326

I. Osiichuk, O. Nazarevych, Ph.D.; Assoc. Prof.

## RESEARCH OF STACK TECHNOLOGIES FOR BUILDING AN INFORMATION SYSTEM FOR THE ACCUMULATION AND ANALYSIS OF GAS CONSUMPTION DATA

**Ключові слова:** стек технологій, газоспоживання, аналіз даних, випадковий процес.

**Key words:** technology stack, gas consumption, data analysis, random process.

Аналіз різноманітних даних щодо життєдіяльності людей завжди був важливим процесом. Після виконання своїх первинних потреб homo sapiens для підтримки життя людина рано чи пізно захоче дізнатись ці процеси зсередини, їх закономірності та особливості.

Основним завданням цього дослідження є подання та обґрунтування можливого переліку технологій, із допомогою яких можна ефективно проаналізувати різноманітні дані, зокрема, у якості прикладу, взято процеси газоспоживання на рівні міста. Така інформація може бути корисна як для постачальника, так і споживачів. У якості прикладу концепції роботи системи аналізу наведено діаграму Data Flow на рисунку 1.

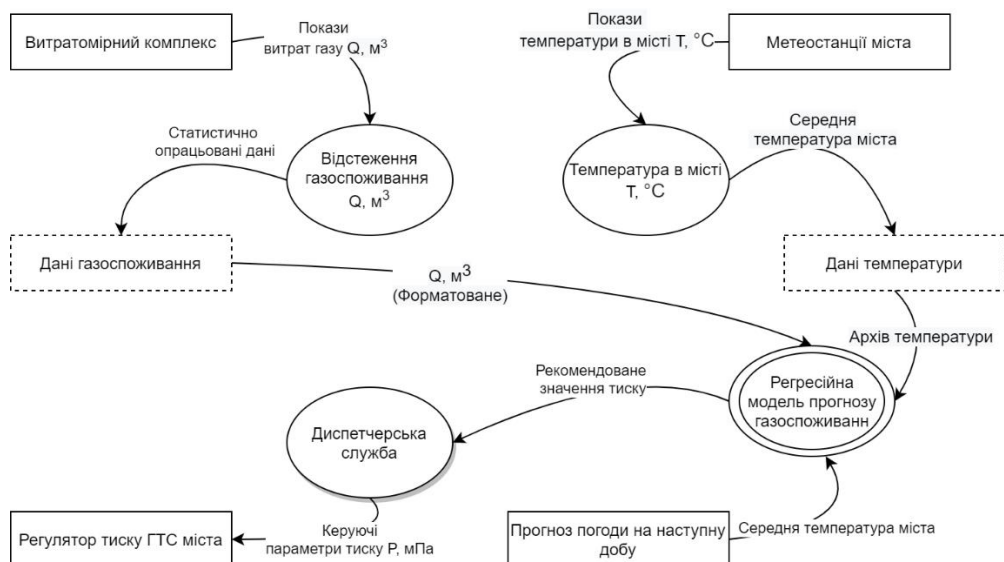


Рисунок 1. Діаграма Data Flow для системи аналізу газоспоживання

Інформаційною технологією, яка, мабуть, стала класичною для наукової роботи з даними, є Python і його стек бібліотек, зокрема numpy, pandas, matplotlib. Перевагою цієї мови програмування, завдяки чому само її використовують, є висока гнучкість і відсутність необхідності розбиратись з особливостями більш складних інструментів.

Якраз вищеописана гнучкість дозволяє писати короткі, лаконічні алгоритми, які легко читати, адже синтаксис Python вкрай близький до англійської мови з вкрапленнями ключових

слів, для механізмів програмування. Єдиною значущою проблемою цієї технології є її низька швидкість виконання, що витікає побічною дією від зручності та ще особливостей інтерпретації в мову С. Якраз цю проблему покликаний приборати numpy – бібліотека, що написана на С, але має інтерфейс у Python, відповідно, поєднує гнучкість та швидкість, назва розшифровується як Numeric Python, що натякає на основний вектор бібліотеки, а також особливу зручність при роботі з числовими масивами.

Друга наведена технологія – Pandas побудована на основі numpy, але ще більше покращує роботу з таблицями/матрицями. У межах бібліотеки фігурує поняття DataFrame – уособлення фрагменту таблиці з даними, це дозволяє маніпулювати самими даними, зокрема окремими стовпцями або рядками, змінювати індекси по осях, і нарешті, зручно виводити дані.

Звичайно, також вкрай зручним і фактично обов'язковим способом аналізу даних є їх візуалізація. Якраз для цих цілей і створена matplotlib. Загалом немало її функціоналу перенесено з Matlab, тому бібліотека буде зручною і для фахівців, що працювали з цією технологією. Для статистики ж є seaborn, що є надбудовою над matplotlib і дає більше зручностей для статистичних графіків, що дозволяє ще більш зручно зображати дані.

В результаті проведеного аналізу, нами запропоновано вище перелічений стек техноогій з написання відповідних методів для backend у вигляді API та frontend у вигляді вебсайту з відображенням і можливостями імпорту, накопичення даних вимірювання, візуалізації та їх статистичного опрацювання.

#### **Література.**

1. Назаревич О.Б. Інформаційна технологія моніторингу газоспоживання міста. Дисертація на здобуття наукового ступеня кандидата технічних наук: 05.13.06 – інформаційні технології / О.Б. Назаревич – Тернопіль, 2015. – 160 с. Nazarevych O.B. Information technology for monitoring of city gas consumption. Thesis for the degree of candidate of technical sciences: 05.13.06 – Information Technology / O.B. Nazarevych — Ternopil, 2015. — 160 p.
2. SciPy and NumPy / Eli Bressert, 2012 – 57 p.
3. Багаторівнева інформаційна система екомоніторингу та керування клімат контролем smart growing box / О. Б. Назаревич, А. О. Волоха, О. Г. Зимницький. // МАТЕРІАЛИ XVI науково-технічної конференції студентів, аспірантів, докторантів та молодих учених. – 2019. – С. 227–229.