Slipchenko, Volodymyr; Poliahushko, Liubov; Krush, Olha

Article

Developing an automated system for collecting various information in the system of complex eco-energy-economic monitoring

This Version is available at:
http://hdl.handle.net/11159/8993

Volodymyr Slipchenko,
Liubov Poliahushko,
Olha Krush

DEVELOPING AN AUTOMATED SYSTEM FOR COLLECTING VARIOUS INFORMATION IN THE SYSTEM OF COMPLEX ECO-ENERGY-ECONOMIC MONITORING

The study is devoted to the development of a system for the automated collection of primary information coming from different sources and in different formats. The object of the study is the properties of primary information and the processes that occur during its collection in systems of complex monitoring in various directions. One of the most problematic places is that information about the environmental, energy, economic and social state of the region can come from different sources and in different formats. This requires a significant investment of resources and time for the development of blocks in integrated monitoring systems.

In the course of the study, a systematic approach was used when analyzing the structure of sources and formats of information of various types, as well as an analysis method when developing a set of rules for collecting information.

The authors analyzed the structure of different sources and formats of primary information in the systems of environmental, energy, economic and medical monitoring. It has been determined that there is no single database and for conducting a comprehensive analysis it is necessary to collect data from various sources and in different formats, which requires a significant investment of resources and time. Therefore, in the course of the study, an automated system for collecting primary information was developed, which allows creating your own sets of rules for importing data from files of various formats. The proposed system can be integrated into large monitoring systems instead of an information collection unit. For example, the module for collecting primary information of the integrated eco-energy-economic monitoring system successfully uses the developed automated system to collect information that is necessary for analysis, modeling and forecasting of the territory by specialists of various profiles, namely: environmentalists, power engineers, economists and medical workers. This made it possible to reduce the total time for the development of a comprehensive eco-energy-economic monitoring system, as well as to increase the accuracy and timeliness of obtaining information during the operation of the monitoring system.

***Keywords:** automated system, information collection, set of rules, expert system, complex eco-energy-economic monitoring.*

1. Introduction

One of the priority areas for the development of any country in the world is to strike a balance between meeting the modern needs of mankind and protecting the interests of future generations, including their need for a safe and healthy environment. This corresponds to the general concept of sustainable development of the countries of the world, and its implementation is impossible without a comprehensive analysis of the environmental, economic and social parameters of the development of states, in particular Ukraine. To solve this problem, the specialists of the National Technical University «Igor Sikorsky Ukraine Kyiv Polytechnic Institute» in the framework of the implementation of the Decree of the President of Ukraine dated January 17, 1995 No. 53/95 «On the system for predicting the genetic risk of introducing new technologies and environmental pollution» [1], a complex eco-energy-economic monitoring system (CEEEM) was developed [2] for effective management decision-making to ensure a high quality of life and health of the population. The CEEEM system allows monitoring the region, district and city by specialists of various profiles, namely: an ecologist, a power engineer, an economist, a doctor, a lawyer and an analyst. Such a set of specialists includes all important aspects of the environmental, economic and social development of the population.

The architecture of the CEEEM system (Fig. 1) consists of three levels: the level of information collection, the level of information storage, the level of application software.
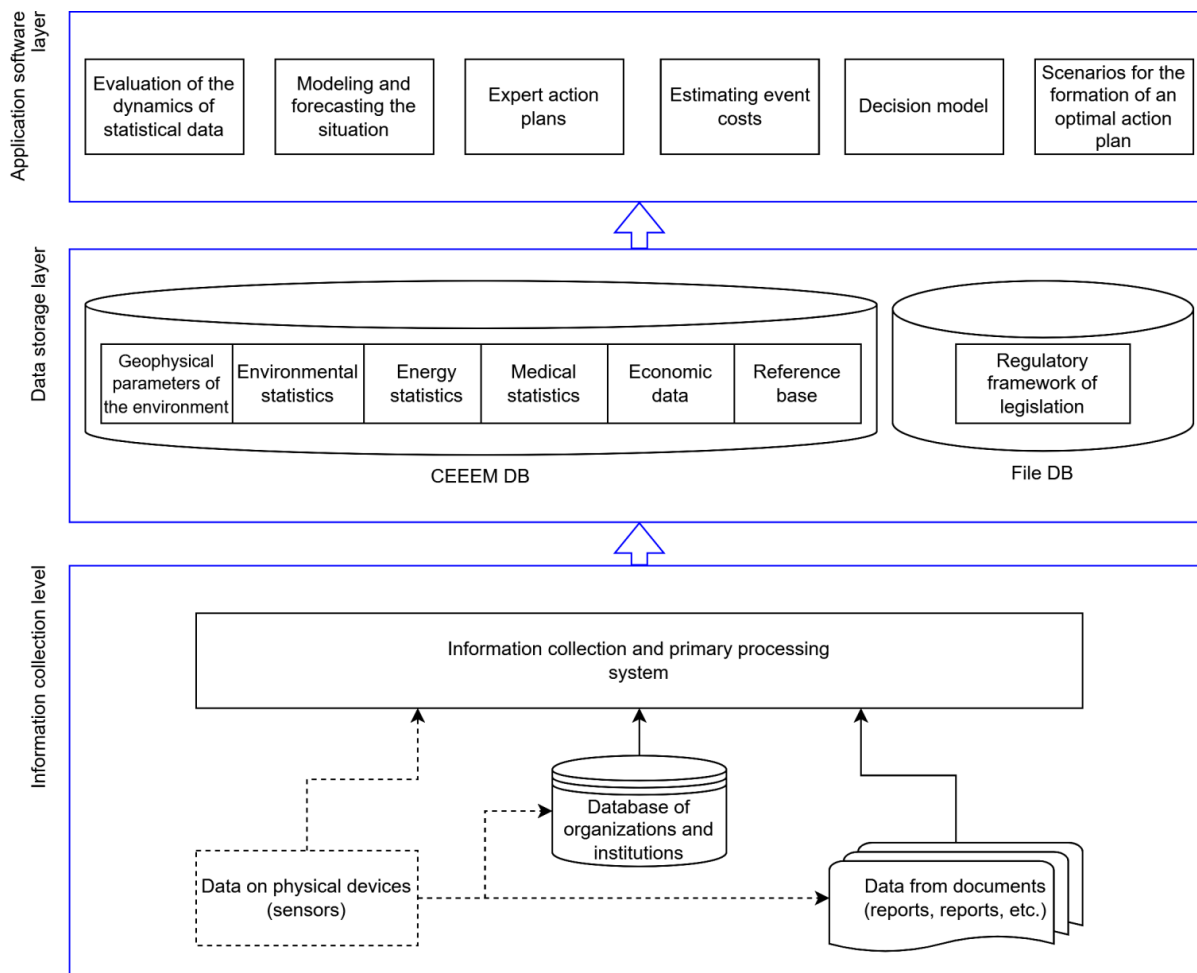
Fig. 1. Architecture of the CEEEM system

The collection level is used to collect data:
– from physical devices in real time, for example, emissions from hazardous industries, production or use of electricity;
– statistical information from documents (reports, reports, etc.);
– from databanks of companies, departments and other institutions.

The level of data storage provides saving and working with the database of the system, where the geophysical parameters of the environment and objects are stored, statistical data of experts, regulatory reference information and a file system to save the legal base.

The level of application software allows assessing, modeling and predicting the situation in the selected zone, providing recommendations for improving the situation, making management decisions and forming an optimal plan of countermeasures to ensure the sustainable development of the country.

The collection of information is a very important stage in the operation of the CEEEM system, since the speed of response in emergency situations, the systematic and reliability of the conducted studies of the state of the area depend on the correct and timely receipt of data. Considering that in the CEEEM system information comes from different sources and in different formats, it is relevant to automate the data collection process and develop an automated system for collecting heterogeneous information that could be integrated into the CEEEM.

2. The object of research and its technological audit

The object of research is the properties of primary information and the processes that occur when it is collected in systems of complex monitoring in various directions. Information about the environmental, energy, economic and social situation of the region can come from different sources and in different formats, which requires significant resources and time to develop blocks in integrated monitoring systems.

3. The aim and objectives of research

The aim of research is to develop an automated information collection system that does not depend on the structure and sources of information. The created automated information collection system should be compatible with the CEEEM system.

To achieve the aim, the following objectives are distinguished:

1. Analyze the features of sources and formats of primary information in monitoring systems.

2. Form a set of rules and commands for the primary information collection system and test it on real data.

3. Integrate the developed system for collecting primary information into the CEEEM system.

4. Research of existing solutions to the problem

An analysis of literary sources and systems of environmental, energy, economic and medical monitoring showed that the main methods of collecting information are:

1. Manual data entry is a fairly common method, especially in energy monitoring, land monitoring and bioecological monitoring systems. For example, the Automated System Energy Monitoring (ASEM) «Kyiv region is energy efficient» [3].

2. Automated data collection from sensors and automated posts. It is used in energy monitoring and air monitoring systems. For example systems:
– SaveEcoBot [4];
– ASEM [5];
– ICE [6].

3. Import data from files. For example systems:
– VDV [7];
– AirQ+ [8];
– FFI [9].

Automated data collection and import is usually carried out using separate information collection subsystems, which are part of large software systems and are usually designed for a specific method and data format. In such systems, adding new sources and formats of information requires refining the structure of the system and adding new blocks.

Data can be collected in electronic storages (databases) for further processing and analysis. Electronic repositories can be divided into two categories:

1) highly specialized repositories used by a specific software system for its own needs;

2) state data banks that combine environmental information from different regions of the country [10].

Systems that serve electronic storage of information and provide access to them can be divided into the following types:
– management systems for electronic data banks;
– systems that provide access to data through API, through electronic reports and through a software graphical interface.

Management systems for electronic data banks, access to which is granted by special permission and is carried out through internal data exchange mechanisms. In this case, developers know the format of the data and have programmatic access to it. This is the most convenient and efficient way, but the absence of a single state data bank for all indicators implies interaction with a large number of institutions that must provide permissions and information on ways to access data [11].

Systems that provide access to data via API. Such systems avoid many of the problems with obtaining permits, but require special converters to read the data. For example, the site saveecobot.com, which contains information about the state of the air in Ukraine and some countries of the world and provides access to data through ARI.

Systems that provide access to data in the form of electronic reports of various formats. An example is the open data portal [12], which provides data in pdf, jpeg, html, xls (xlsx), csv, doc (docx), json formats. For automated collection, pdf and jpeg formats are not considered as input data, since they require complex image recognition technologies that give too many errors, and html, xls (xlsx), csv, doc (docx), json formats can be converted programmatically using the appropriate libraries and enter the data into the database.

Systems that provide access to information only through a software graphical interface. They require parsing of the html page (in the case of web-based systems) or manual data entry into the monitoring system. For example, the web-based system «Open Environment» [13] and interactive soil maps [14]. This type of system is not considered as a source of information for an automated system for collecting primary information.

5. Methods of research

The following methods were used during the study:
– a systematic approach – when analyzing the structure of sources and formats of information of various types;
– method of analysis – in the development of rules for collecting information.

6. Research results

It is proposed to implement an automated system for collecting primary information using the principle of a set of rules of an interpretive expert system. The advantage of interpretive expert systems is that they are able to process various types of data [15].

An automated system for collecting information based on an expert system has the following structure (Fig. 2):
– rule base, which is a base of files containing sets of rules;
– database of the system in which information is to be stored;
– rule building module designed to create new sets of rules;
– an interpretation module designed to execute a set of commands using sets of rules and an information source file (the information obtained is stored in the target database).

The rules are formed by the rule building module once for a file of a specific structure. The rules file can be in the format of a txt or json text file.
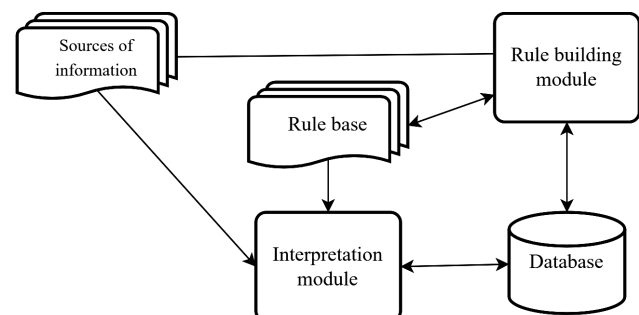
Fig. 2. Generalized scheme of the structure of an automated system for collecting information

Operation of the automated data collection system in the CEEEM system.

The CEEEM system collects data from different sources of information and different formats.

For example, information about the state of the atmosphere enters the system from the saveecobot.com website (Fig. 3) [5]. Through the ARI site, it is possible to get data in json format (Fig. 4).
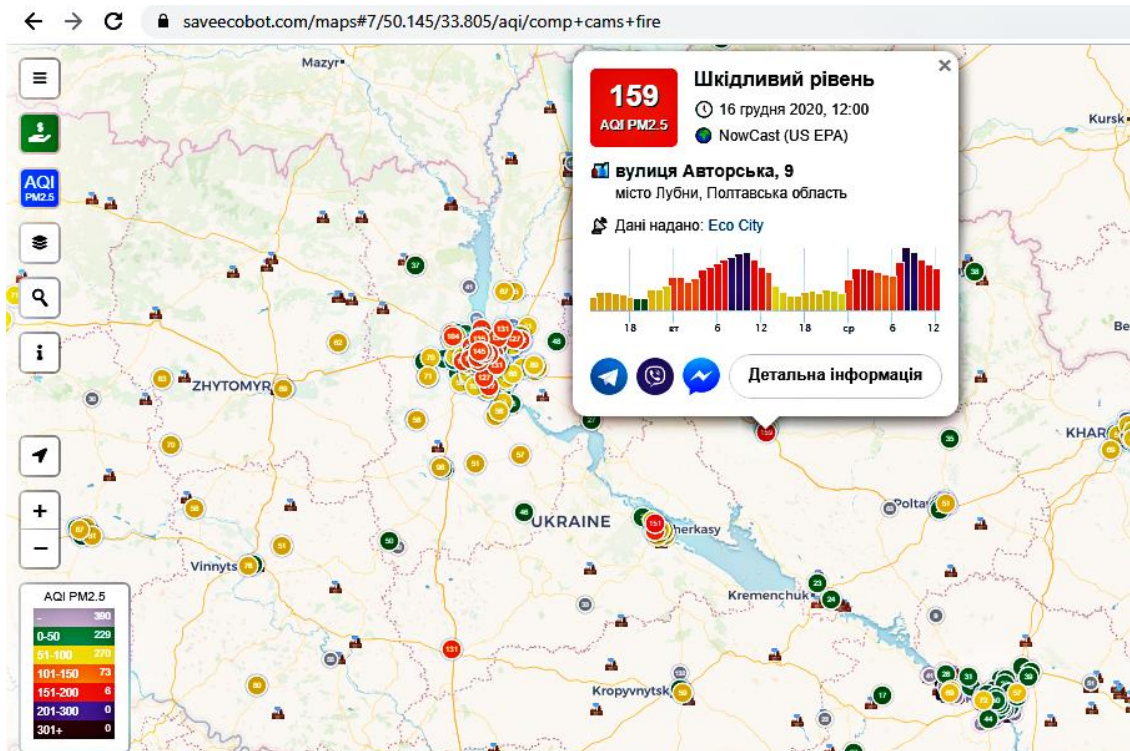
Fig. 3. Interactive map on saveecobot.com

```
{
    "id": "SAVEDNIPRO_001",
    "cityName": "Dnipro",
    "stationName": "Geroev Avenue, 40",
    "localName": "Save Dnipro #1",
    "timezone": "+0200",
    "latitude": "48.408944",
    "longitude": "35.072505",
    "pollutants": [
        {
            "pol": "Humidity",
            "unit": "%",
            "time": "2021-01-27 15:40:30",
            "value": 100,
            "averaging": "2 minutes"
        }
    ]
}
```

Fig. 4. An array of data on air quality at observation points
in json format

This format has a set of property names and their value as the name of the property in the data array: values. In the program, for each property, it is necessary to have a matching rule in the form of a structural element of the database table = property in the json data array. Since the names of properties on different resources are different, it is completely impossible to automate such a process.

Fig. 4 shows the format of the data that can be obtained through the ARI.

An example of a set of data import rules from the saveecobot.com system file is shown in Table 1.

Description of the columns of the rules table:

– the «Incoming data» column contains the name of the key in the structure of the json format, the address of the cell in the Excel file;

– the «Value in file» column contains data from the information source file and helps to visually assess the correctness of the selected key;

– the «Outgoing data» column contains a special command that determines where the incoming data will be written;

– the «Value for the command» column contains the parameter values for the command;

– the «Command» column contains the name of the command that is executed on the data;

– the «Settings» column contains the names of parameters for the command;

– the «Result» column contains the name of the variable or the name of the field in the table where the execution result will be placed.

Command options:

1. Find – searches for data in a specific database table.

2. Save – saves data to a variable.

3. Add – adds data to the database.

4. AddIfNotExist – adds data to the database if there are none.

5. Miss – skips the key and does not perform operations on it.

6. Divide – divides data into components according to a certain rule.

7. EditOrInsert – Updates data in a specific row of a database table, if any, and adds a row if none.

To import data from an xlsx file, the set of rules changes according to the specifics of the format, since the data is arranged in columns rather than blocks.

An example of information on the state of drinking water, presented in the xlsx format (Fig. 5), was taken from the website of the Municipal Enterprise «Kyivvodokanal» [16]. Apparently, unlike json, the data has a tabular structure and the parameters of each object are the headings of the corresponding columns, and the lines form tuples with a set of specific characteristics. Table 2 shows a set of rules for importing data from an xlsx file containing information about the state of drinking water.

Table 1

Automated system rules for json file

No.	Input data	Value in file	Initial data	Command value	Command	Options	Result
1	Json	Ecobot.json	–	–	–	–	–
2	id	SAVEDNOPRO_001	Database	Poi.id	Find	Source=Poi, Condition(Poi.NameOfObject=SAVEDNOPRO_001)	Id_poi
3	cityName	Dnipro	–	–	Miss	–	–
4	stationName	Heroiv Avenue, 40	–	Heroiv Avenue, 40	Save	Poi.DescOfObject	DescOfObject
5	timezone	+0200	–	–	Miss	–	–
6	latitude	48.408944	–	–	Save	Poi.Coodr_Lat	lat
7	longitude	35.072505	–	–	Save	Poi.Coord_Lan	lon
8	Id_poi	–	–	–	EditOrInsert	Poi, (DescOfObject, lat, lon)	–
9	pollutants	–	–	–	Miss	–	–
10	pol	PM 2.5	–	Elements.Id	Find	Elements, Elements.Name=«Suspended particles PM2.5»	Id_elem
11	unit	mg\/m3	–	–	Save	EmissionOnMap.measure	unit
12	time	2020-12-01 19:11:12	–	–	Divide	Date_fomat(yyyy-mm-dd)	year, month, day
13	value	52.03	–	–	Save	EmissionOnMap.measure	value
14	–	–	–	–	AddIfNotExist	EmissionOnMap, (Id_poi, Id_elem, 1, value, value, null, year, month, day,unit)	–
15	averaging	–	–	–	Miss	–	–

Fig. 5. An array of data on the state of drinking water, presented in xlsx format

Table 2

Automated system rules for xlsx file

No.	Input data	Value in file	Initial data	Command value	Command	Options	Result
1	Cells[1][0]	Drinking water	Database	Drinking water	Find	environment;name	environment;id
2	Cells[4][0]	10	Variable	Humus, very low <1.1	Save	–	value
3	Cells[5][0]	1	Variable	62.8	Save	–	value
4	Cells[6][0]	2022	Variable	–	Save	–	value
5	Cells[3][0]	Holosiivskyi Ave., 56	Database	Holosiivskyi Ave., 56	Find	poi;Name_Object	poi;Id
6	Cells[10][1]	Smell: at t 20 °C, at t 60 °C	Database	Smell: at t 20 °C, at t 60 °C	Find	elements;name	elements;code
7	Cells[10][4]	1	Variable	1	Save	–	value
8	Cells[10][2]	points	Variable	points	Save	–	value
9	Cells[11][1]	coloration	Database	coloration	Find	elements;name	elements;code
10	Cells[11][4]	12	Variable	12	Save	–	value
11	Cells[11][2]	degrees	Variable	degrees	Save	–	value
12	Cells[12][1]	Turbidity	Database	Turbidity	Find	elements;name	elements;code
13	Cells[12][4]	0.41	Variable	0.41	Save	–	value
14	Cells[12][2]	mg/dm^3	Variable	mg/dm^3	Save	–	value
15	Cells[13][1]	Hydrogen indicator	Database	Hydrogen indicator	Find	elements;name	elements;code
16	Cells[13][4]	7.7	Variable	7.7	Save	–	value
17	Cells[13][2]	pH units	Variable	pH units	Save	–	value
18	Cells[14][1]	General hardness	Database	General hardness	Find	elements;name	elements;code
19	Cells[14][4]	4.6	Variable	4.6	Save	–	value
20	Cells[14][2]	$mmol/dm^3$	Variable	$mmol/dm^3$	Save	–	value
21	Cells[15][1]	Iron total	Database	Iron total	Find	elements;name	elements;code
22	Cells[15][4]	0.1	Variable	0.1	Save	–	value
23	Cells[15][2]	mg/dm^3	Variable	mg/dm^3	Save	–	value
24	Cells[16][1]	Nitrates	Database	Nitrates	Find	elements;name	elements;code
25	Cells[16][4]	5.8	Variable	5.8	Save	–	value
26	Cells[16][2]	mg/dm^3	Variable	mg/dm^3	Save	–	value
27	Cells[17][1]	total alkalinity	Database	total alkalinity	Find	elements;name	elements;code
28	Cells[17][4]	3.6	Variable	3.6	Save	–	value
29	Cells[17][2]	$mmol/dm^3$	Variable	mmol/dm3	Save	–	value
30	–	–	Database	–	Add	emissions_on_map	idPoi,idElement,idEnvironment, 1,1,2022,1,10,'бали'
31	–	–	Database	–	Add	emissions_on_map	idPoi,idElement,idEnvironment, 12,12,2022,1,10,'градуси'
32	–	–	Database	–	Add	emissions_on_map	idPoi,idElement,idEnvironment, 0.41,0.41,2022,1,10,'мг/дм3'
33	–	–	Database	–	Add	emissions_on_map	idPoi,idElement,idEnvironment, 7.7,7.7,2022,1,10,'одиниці pH'
34	–	–	Database	–	Add	emissions_on_map	idPoi,idElement,idEnvironment, 4.6,4.6,2022,1,10,'ммоль/дм3'
35	–	–	Database	–	Add	emissions_on_map	idPoi,idElement,idEnvironment, 0.1,0.1,2022,1,10,'мг/дм3'
36	–	–	Database	–	Add	emissions_on_map	idPoi,idElement,idEnvironment, 5.8,5.8,2022,1,10,'мг/дм3'
37	–	–	Database	–	Add	emissions_on_map	idPoi,idElement,idEnvironment, 3.6,3.6,2022,1,10,'ммоль/дм3'

The rule execution algorithm is shown in Fig. 6.

Compliance with the rules does not depend on specific data values. The rule is bound to the file structure.

The proposed organization of data processing allows to automate the process of reading data, reduce the time for data processing and reduce the number of errors in manual entry.
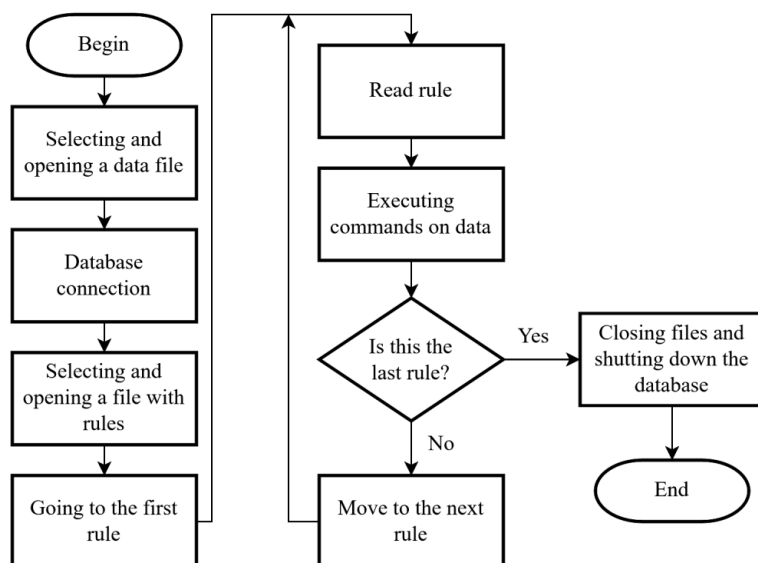

Fig. 6. The rule execution algorithm

7. SWOT analysis of research results

Strengths. The strengths of this study are that the developed automated system is flexible and easily adaptable to any monitoring system, which requires a module for collecting primary information from different sources and data formats for large monitoring systems. Also, the extension of the set of commands will allow using the developed system for import/export of data from database tables of any structure. This significantly reduces resources and time for the development of large systems and monitoring complexes.

Weaknesses. The weaknesses of the study are that there is no automated data collection from pdf and jpeg files. Therefore, the user needs to carry out additional format conversions in third-party services, but this can lead to an increase in the number of errors in the source data.

Opportunities. The developed automated system for collecting primary information can be integrated into monitoring systems of various directions, which will reduce the cost and speed up the development of such systems.

Threats. The use of an automated system for collecting primary information in economic terms may not be effective when developing small systems with a small number of information sources.

8. Conclusions

1. The analysis of sources and formats of primary information in monitoring systems is carried out. It is noted that today there is no single database of environmental, economic and social parameters, therefore, when developing systems for integrated monitoring, it is necessary to collect data from different sources and in different formats, which slows down the process of developing such systems.

2. A set of rules and commands for the system for collecting primary information is developed, which makes it possible to speed up the process of entering data from MS Excel and json files. The results of testing the developed set on real data showed that such an approach to collecting information expands the capabilities of the system for collecting primary information and makes it almost independent of the structure of the input file. The only limitation is that each file structure has its own set of rules. That is, to expand the number of sources of information, there is no need to rework the program code of the system, it is only necessary to add a new rule to the knowledge base.

3. The developed automated system for collecting primary information, implemented on the principles of an interpretive expert system, has been successfully integrated into the CEEEM. But it can be easily integrated into various monitoring systems instead of an information collection unit. To do this, it is enough to connect the automated system with the database of the monitoring system and prescribe the appropriate sets of rules. This reduces the cost of developing new software products. The module for collecting primary information CEEEM successfully uses the developed automated system to collect information necessary for the work of specialists in various fields.

References

1. *Pro systemu prohnozuvannia henetychnoho ryzyku vprovadzhennia novykh tekhnolohii ta zabrudnennia navkolyshnoho seredovyshcha* (1995). Ukaz Prezydenta Ukrainy No. 53/95. 17.01.1995 Available at: https://zakon.rada.gov.ua/laws/show/53/95#Text
2. Slipchenko, V. H., Poliahushko, L. H., Krush, O. Ye. (2021). Systema kompleksnoho eko-enerho-ekonomichnoho monitorynhu dlia optymizatsii upravlinskykh rishen (oblasti, raionu ta mista). *Visnyk Skhidnoukrainskoho Natsionalnoho Universytetu imeni Volodymyra Dalia, 4 (268),* 13–20. doi: http://doi.org/10.33216/1998-7927-2021-268-4-13-20
3. *Avtomatyzovana systema enerhomonitorynhu «Kyivshchyna enerhoefektyvna».* Available at: https://cem.drrzkg-koda.gov.ua/c=kodaenergo
4. *Riven zabrudnennia atmosfernoho povitria u misti Kyiv.* SaveEcoBot. Available at: https://www.saveecobot.com/maps/kyiv
5. *Avtomatyzovana systema enerhomonitorynhu (ASEM).* Available at: https://asem.com.ua
6. *Informatsiina systema enerhetychnoho monitorynhu (ISE).* Available at: https://www.fiatu.com.ua/it-technology/e_monitoring/
7. *A single solution to rule all. Vista Data Vision (VDV).* Available at: https://vistadatavision.com/about-us/
8. *«AirQ+: prohramnyi zasib dlia otsinky ryzyku dlia zdorovia vid zabrudnennia povitria».* Available at: https://www.euro.who.int/ru/health-topics/environment-and-health/air-quality/activities/airq-software-tool-for-health-risk-assessment-of-air-pollution
9. *FFI: A software tool for ecological monitoring.* Available at: https://www.ltrr.arizona.edu/~ellisqm/outgoing/Rowe_Mesa_2014/Monitoring/monitoring_guidelines/monitoring/FFI%20software/FFI_Overview.pdf
10. *Vsiu ekolohichnu informatsiiu v Ukraini bude zvedeno v yedynu elektronnu bazu «Vidkryte dovkillia»* (2018). Available at: https://mepr.gov.ua/news/32870.html
11. *Dostup do ekolohichnoi informatsii.* Available at: https://www.zakon-i-normativ.info/index.php/component/lica/?href=0&view=text&base=1&id=157271&menu=267259

12. *Derzhavne pidpryiemstvo «Derzhavnyi tsentr informatsiinykh resursiv Ukrainy»*. Available at: https://dir.gov.ua/
13. *Open Access Environment*. Available at: https://openaccess.org.ua
14. *Ahrokhimichna karta Ukrainy*. Available at: https://superagronom. com/karty/agrohimichna-karta-ukrainy
15. Ogidan, E. T., Dimililer, K., Ever, Y. K. (2018). Machine Learning for Expert Systems in Data Analysis. *2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. doi: http://doi.org/10.1109/ismsit.2018.8567251
16. *Rukhaiemos u maibutnie bazuiemos na dosvidi*. Available at: https://vodokanal.kiev.ua/

Volodymyr Slipchenko, Doctor of Technical Sciences, Professor, Department of Automation of Power Processes and Systems Engineering, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine, ORCID: https://orcid.org/0000-0002-3405-0781

✉*Liubov Poliahushko*, PhD, Associate Professor, Department of Automation of Power Processes and Systems Engineering, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine, e-mail: liubovpoliagushko@gmail.com, ORCID: https://orcid.org/0000-0003-3287-8523

Olha Krush, Department of Automation of Power Processes and Systems Engineering, National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute», Kyiv, Ukraine, ORCID: https://orcid.org/0000-0001-8792-2515

✉*Corresponding author*

RESEARCH OF METHODS FOR DETERMINING THE ACCURACY OF METROLOGICAL MEASUREMENTS

**Andrii Kozhevnikov,
Nataliya Bilous**

The object of research in this work are individual branches and processes of metrology, namely algorithms and methods of global positioning, analysis and research of individual processes, and their implementation in the software model to analyze their accuracy. The existing problem is that some methods or devices that produce results based on these methods are not accurate enough or their accuracy is affected by a number of factors that worsen the result.

During the work the methods and approaches of the global positioning system were analyzed, such as: precise point positioning, relative positioning of GPS, static and fast static GPS-shooting and Stop-and-go shooting. Approaches were also divided into static and kinematic. Among the methods, the method of precise point positioning was chosen, as the most common and used among ordinary GPS users, for more detailed research and analysis of accuracy. Development methods are based on the means of interaction with open GPS services, and C#, ASP.NET Core, Angular framework, and development environment Visual Studio Code and Visual Studio 2022.

In order to implement all the necessary functionality, the subject area was analyzed, the methods of determining metrological accuracy in general and within the system of global positioning as a direct object of work were considered. As a result of the work the analysis and modelling of the subject area was carried out, the methods of the global positioning system and the software system for the analysis of the accuracy of one of the GPS methods were investigated. The analysis results provided by the software help to understand the accuracy of the method and the GPS receivers that use it on different devices.

Keywords: *metrological measurements, global positioning system, precise point positioning, .NET, Angular.*

How to cite

1. Introduction

In recent years, more and more there is a trend in the automation of manual labor, and the technology that replaces manual labor is becoming more versatile and powerful. At first, computers only performed data calculations, then algorithmic tasks, and today, thanks to advances in machine learning, bioinformatics, the Internet, and increased computing power, they have become an integral part of human privacy [1]. Laptops, smartphones, tablets, smart homes are already a routine for each of us. Hundreds of tasks that were previously performed by a person with its own hands are now performed for it by machines. But not all.

The topic of this work, namely the study of methods for determining the accuracy of metrological measurements, was chosen precisely because many routine functions and