Multi-Agent Approach for the Unification of Meteorological Data

Karina Melnyk, Yaroslav Kravets, Iryna Liutenko, Svitlana Yershova, Oksana Ivashchenko, Dmytro Yershov and Olena Odyntsova

National Technical University "Kharkiv Polytechnic Institute", Kirpichova street, 2, Kharkiv, 61002, Ukraine

Abstract

This study is devoted to the problem of preparing data for further analysis and evaluation of weather providers. The review of works in forecasting weather conditions domain has been conducted. The technology for solving the task of improving the quality of managerial decisions has been developed. The formalization of this process is presented in the form of a BPMN-diagram. The stage of data preparation for decision-making is considered. It is presented as a process of obtaining unified data from mixed data. The architecture of a multi-agent system for searching, downloading and preparing data has been proposed. The Crawler Subsystem and the Unifier Subsystem were designed and implemented. A set of unified data quality indicators has proposed, as well as an approach for their evaluation.

Keywords

Weather forecast, raw data, data quality, multi-agent system, crawler, unified data, data cleaning

1. Introduction

The economic growth of many countries leads to serious environmental pollution. The military events in Ukraine in 2014-2023 significantly worsen the environment. The usage of weapons pollutes not only the atmosphere, but also the soil; it leads to a deterioration in the quality of crops and life in general as well. To reduce the negative impact on global warming, some countries are implementing green energy programs.

Green energy is renewable energy sources, which are inexhaustible in nature: sunlight, wind or water. Green energy does not produce carbon dioxide or other gases that cause global warming. The transition to green energy tops the list of measures to stop the climate crisis [1-3]. However, the process of producing green energy is associated with some features, namely, uncontrollability and dependence on weather conditions. For example, cloudiness reduces the production of electricity by solar power plants tenfold [4]; the absence of wind makes wind generation impossible [5]; the absence of precipitation leads to a decrease in hydropower [6]. Therefore, the quality of the weather forecast influences on managerial decisions regarding the management, control and use of energy.

On the other hand, the relevant weather forecast is important for the tourism industry, because it will allow making the right tourist route and avoiding irreparable consequences. The weather influences on the well-being of children under 1 year old and people over 60 years old as well [7]. The accurate weather forecast will indicate the period of using the fertilizers in agricultural activities. It will point when crops need to be sheltered from the elements. The weather also affects the meteorological support of military and civil aviation and is a base in the planning of some military ground operations [8-9].

COLINS-2023: 7th International Conference on Computational Linguistics and Intelligent Systems, April 20–21, 2023, Kharkiv, Ukraine
EMAIL: Karina.v.melnyk@gmail.com (K. Melnyk); beeengine1983@gmail.com (Y. Kravets); cherliv68@gmail.com (I. Liutenko); Svetlana.ershova.2016@gmail.com (S. Yershova); ok.v.ivashchenko@gmail.com (O. Ivashchenko); 3dimae@gmail.com (D. Yershov); elena314151@gmail.com (O. Odyntsova)
ORCID: 0000-0001-9642-5414 (K. Melnyk); 0000-0002-3893-1607 (Y.Kravets); 0000-0003-4357-1826 (I. Liutenko); 0000-0003-3893-117X (S. Yershova); 0000-0003-3636-3914 (O. Ivashchenko); 0009-0001-9284-5873 (D. Yershov); 0009-0005-8084-9526 (O. Odyntsova)

Meteorologists produce weather forecast based on synoptic maps. The results of the forecasting are presented on special websites, where the user can specify the parameters of interest: localization, weather forecast horizon, graphs of changes in atmospheric pressure, etc. [10-14]. However, often a situation arises when weather providers show different results with the same input parameters. The reason of such discrepancy is using meteorological data from climate stations located in different regions and various specification of this data. The number of input parameters for meteorologists can be up to a hundred, so weather forecasting models also vary. For any subject area, like military aviation or the agricultural sector, it is known that the more accurate and relevant the forecast from weather forecasters is, the more adequate the proposed managerial decision in this domain will be. It is impossible to change the forecast data from meteorologists. Therefore, in order to improve the quality of managerial decisions, it is necessary to choose a weather provider (a specific website) that gives the most accurate weather forecast based on a comparison of forecast and real weather data.

The format of archived meteorological data of each site differs from each other, for example: a city can be written in words in String format, or GPS coordinates can represent it. Therefore, to process and compare the results of the weather forecast, it is necessary to apply an information system (IS). Traditional ISs are a set of methods, tools and technologies for the implementation of the information process, which allow solving the tasks in accordance with the proposed algorithms. Knowledge about the domain in an intelligent IS (IIS) is stored in a knowledge base in a form of facts and rules in a declarative form. Such approach can allow changing not only individual parameters of the IIS's functioning, but also the way the system behaves as a whole. That is, IISs are systems based on knowledge, information and data. It will allow processing data of mixed nature, reducing the influence of the human factor on the data preparation process, and transforming the input parameters to a unified form in order to use definite mathematical method to solve the problem of selecting the most appropriate website with weather forecast.

Thus, the purpose of the article is to improve the quality of input information for making managerial decisions connected with the weather forecast using the IIS that will allow pre-processing of data to solve the problem of choosing the most appropriate source of weather forecast data.

2. Formal problem statement

The aforementioned information allows presenting that in order to solve the task of improving the quality of data for managerial decisions related to weather forecasting, it is necessary to solve the following set of issues:

- to make an analytical review of existing researches in weather forecast sphere;
- to conduct a domain analysis and identify separate stages of data processing;
- to develop a technology for solving the given task;
- to propose IIS architecture;
- to conduct an experiment using the developed system;
- to evaluate the quality of the proposed solution.

3. Review of related works

Many scientists were engaged in solving the problem of analyzing the quality of weather forecasts. The authors of [15] compared the results of forecasting using different methods. A weather station located in Caserta, Italy provided a data set for 4 years of observations. ARIMA, Holt-Winters and general ETS exponential smoothing techniques were used to generate weather forecasts. The result of the study was the conclusion that the exponential smoothing showed a very good result, but it happened because of the station constantly collects meteorological data with a high frequency.

The main attention in the study [16] was paid to the analysis of the prediction of solar radiation. It presented the results of a mutual comparison of selected models under clear and cloudy sky conditions. The observations for research were collected for six months from July 2013 to December 2013. The conclusions showed that the selected models overestimate short-wave radiation.

The authors of the paper [17] conducted a study of weather forecast patterns in Khyber Pakhtunkhwa, Pakistan. Rainfall and temperature data were collected for the period 1960–2020. The obtained conclusions have proven the importance of precipitation mapping, remote sensing and other processes that have a positive impact on the agricultural sector of Pakistan.

The work [18] was devoted to the analysis of weather forecast quality for two cities: Vinnytsia (Ukraine) and Fredensborg (Denmark). Weather data was taken from popular Ukrainian and foreign websites. The study showed that Ukrainian sites have better indicators of accuracy of the temperature forecast for Vinnytsia than global ones. While foreign weather providers are the best predictors of precipitation.

The result of the analysis of available scientific works showed that many researches were conducted for 1-3 settlements (geographical coordinates). In most cases, data were taken from weather stations. In addition, only 1-2 weather parameters were studied. To eliminate these shortcomings, it is proposed to create the IIS with the following requirements:

- the number of settlements may increase depending on the user's needs;
- to assess the relevance of the weather forecast, it is necessary to select all available weather parameters;
- data for analysis must be taken from open sources – weather sites.

4. Technology for improving the quality of weather forecast decisions

Let's consider the task of improving the quality of managerial decisions related with the weather forecast results in more detailed way. The technology for solving the given task consists of three main stages:

- the first stage is related to data preparation, namely, data loading, pre-processing and preparation of data from a set of websites with weather forecasts for certain areas;
- the second stage is the analysis of the received unified data;
- the last stage is formation of a set of recommendations and managerial decisions (Figure 1).

The formalization of this process is presented in the form of BPMN-diagram.

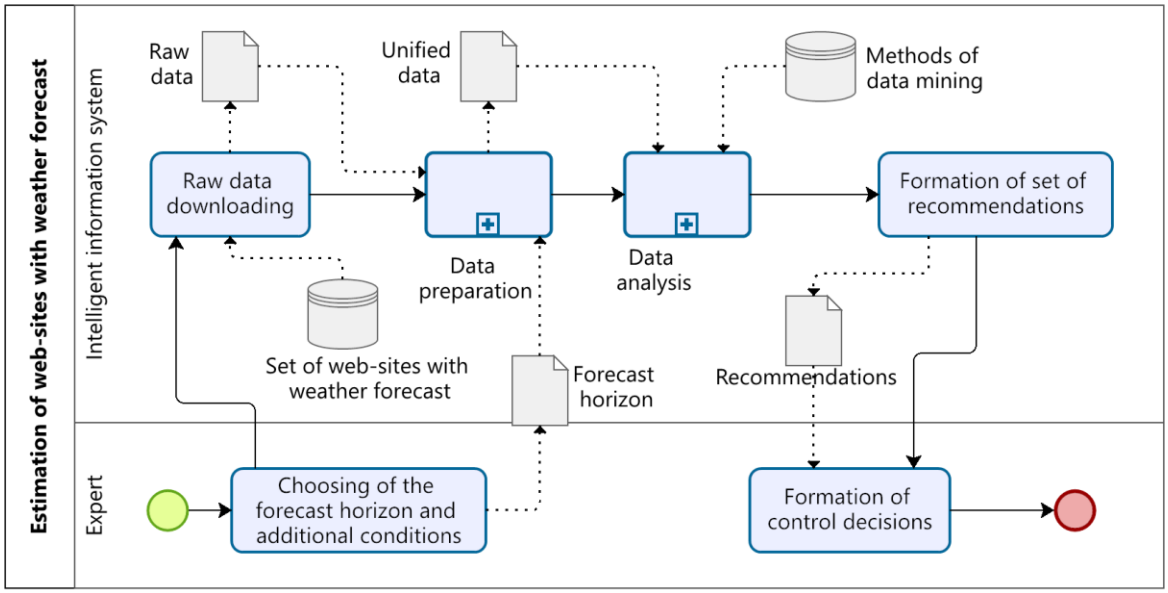

Figure 1: The technology for solving the problem of improving the quality of managerial decisions

The first stage starts from the works of an expert. The expert analyzes business processes in a certain domain and forms a list of input parameters for IIS, which must be taken into account for forming managerial decisions based on weather forecast. IIS receives archived data from a selected set of weather forecast websites. This raw data may be incomplete and contain errors. It has the following different parameters: structure, file system, forecasting period, and various forecasting models are used

to obtain it. Therefore, the main goal of the first stage of the task of improving the quality of managerial decisions is to solve the problem of data unification. It will eliminate unnecessary diversity of data by reducing the list of permissible elements. This will allow getting a high-quality result, because data pre-processing is the basis for the effective operation of data mining methods. Let's consider the details of the process of obtaining unified data:

- data cleaning – is the process of finding, removing and correcting (if it possible) irrelevant data, formatting and spelling errors, punctuation errors or non-compliant symbols, inconsistencies, duplications, structural errors and missing entries, identify outliers and smooth out noisy data from archive data;
- data integration – is the process of combining and transforming data from various sources: files, tables, etc. into a single data scheme;
- filtering or data screening – is a process of sifting, selecting and reducing the content of data according to considered conditions;
- data identification – is the process of bringing, renaming or transforming data into one accepted format for further data processing using data mining methods.

The second stage of the technology to improving the quality of weather forecasting solutions is the analysis of the received data to identify the most adequate and truthful source of the weather forecast. It includes an analytical review of data mining methods and the further selection of a mathematical model, which will allow solving the problem of choosing a certain website. There are a large number of methods that can be used for the considered task: cluster analysis, regression analysis, decision trees, associative rules, the method of analyzing hierarchies, neural networks, case-based reasoning approach, link analysis, genetic algorithms, etc. A model for solving the problem of selecting a website is built based on the chosen approach.

Next step of second stage of technology is the Cross validation process [19]. It is the process of checking the adequacy of the proposed model. The Cross validation is a set of techniques used to evaluate the performance of a mathematical model based on splitting data into training part and testing part. The easiest method is the Train Test Split. Typically, an information distribution of 80% for training and 20% for validation is chosen, but this ratio can be varied in order to find good parameters of the mathematical model. However, it is not recommended to use such a way of evaluating in the case with limited data because of the probability of high bias. Second technique is K-Folds Cross Validation method. The set of data should be divided into k different subsets. Firstly, we use (k-1) folds for the training process of the model, the last piece of data for testing the model. This process is repeated k times, while each of the pieces of data is used exactly once as the test data. Then all obtained results are summarized as average parameters, which can be used for evaluating the performance of the considered model.

After checking the model for adequacy, it can be used in real conditions. The result of using the model are two sets of sources of weather information: a subset of websites that can be trusted and another subset where the weather forecast differs from the actual weather conditions. The first set can consist only one element, and all others belong to the second set. It depends from the chosen mathematical method.

The last stage of the considered technology is the creation of recommendations about the usage of websites. Depending on the analysis model of weather forecast providers, the set of selected websites can also be ranked for confidence on a scale from threshold to maximum. Such distribution will allow experts to choose one or another website, if it is necessary to discard some input parameters over others.

5. Usage of Multi-Agent System for data preparation for weather forecast decisions

Proposed technology for resolving the selection task of definite website allow seeing that every stage is needed of own approach for resolution. There are several independent processes or stages like data preparation and data analysis. Each stage requires its own software and mathematical solutions, so there is no need to keep all data in one database or try to combine different architectural solutions or patterns into a single monolithic software product. Therefore, a multi-agent system (MAS) has been proposed as an IIS that would allow automating the analysis and comparison of weather forecasts from different

providers. The MAS consists of a set of software components – agents that can cooperate with each other and with the surrounding environment to solve certain problems. Each agent has its own capabilities and knowledge about the environment, it does not depend on the capabilities of other agent, it has the ability to make decisions and perform actions in accordance with some rule. The actions of autonomous agents can be either coordinated or competitive, depending on the problem they are trying to solve.

To find the solution of the problem of increasing the quality of managerial decisions, the following MAS architecture has been proposed. It was developed in the C# programming language for Windows, Linux, etc. operating systems, where there is a .NET 6 implementation. The system consists of two subsystems: the Crawler Subsystem and Unifier Subsystem.

The Crawler is a set of software agents with own purpose (Figure 2).

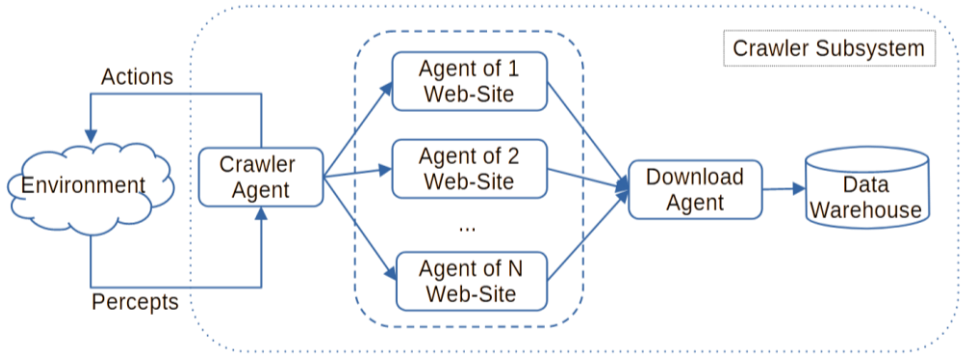

Figure 2: Architecture of Crawler subsystem

The Crawler or Search Agent finds web pages with weather forecast providers. It connects to servers using the HTTP-protocol and receives information from them. The Crawler Agent cooperates with the Agents of the selected websites, because each site has its own file structure for storing content. Next, each Site Agent generates a list of download URLs, as well as delays between downloads. This information is passed to the Download Agent, which downloads and saves the raw data as HTML-pages and additional content to the knowledge base. The operation model of Download Agent is presented in the form of an activity diagram (Figure 3).

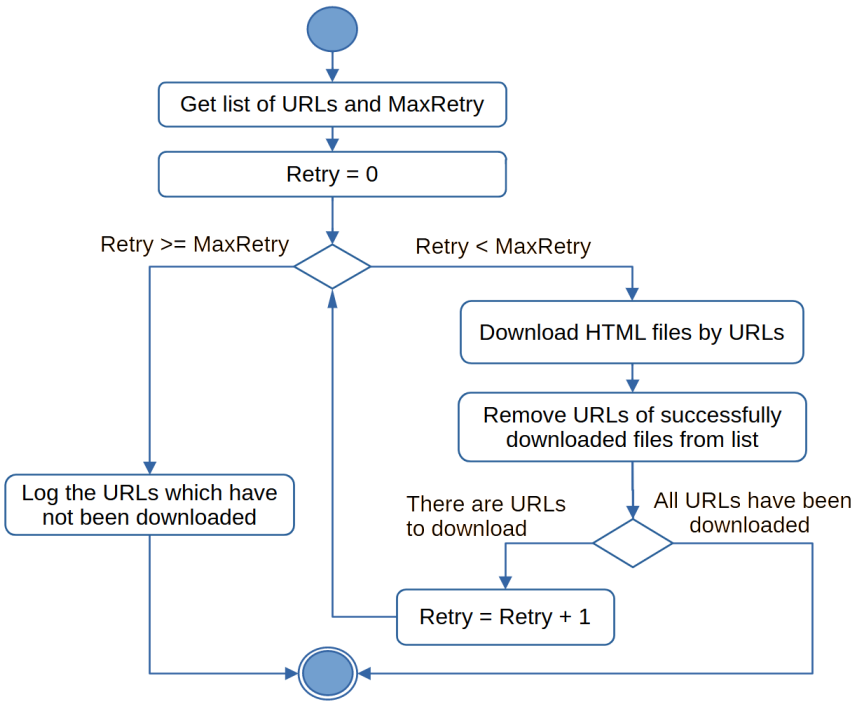

Figure 3: Download Agent model

The Download Agent also supports setting the number of simultaneous download streams. This allows not overloading the channel for a weak network. The Download Agent was developed in accordance with the principles of Polite crawling [20]. In particular, the HTTP headers represent the email and Linkedin page of one of the authors.

To analyze the performance of weather providers, the Crawler Agent has found the appropriate set of websites [10-14] that presented in Table 1. Analysis of GUI (graphical user interface) of these sites shows that each weather provider contains different time intervals for the weather forecast, so it is impossible to use non-unified data for one selected mathematical model.

Table 1
The list of weather forecast websites

Web source	Permitted time ranges for weather forecast
MeteoPost.com	Hourly weather forecast for 3 days
	Weekly forecast
Meteoprog.com	Hourly forecast for 4 days
	Forecast for 2 weeks
	Monthly forecast
Sinoptik.ua	Forecast for 10 days
Gismeteo.ua	Hourly forecast for 3 days
	Forecast for 2 weeks
	Monthly forecast
Meteo.ua	Hourly forecast for 3 days
	Forecast for 2 weeks
	Monthly forecast

After review of presented data on weather providers, it is necessary to choose list of geographical points that are subject to verification. The following list of geographical points has been created: the extreme points of Ukraine, the highest point and the lowest point, and settlements in the Western, Eastern, Central and Southern parts of Ukraine. These are the following settlements: Goverla Transcarpathian region Rakhovsky district, Gremyach Chernihiv region Novgorod-Seversky district, Donetsk, Zhytomyr, Kyiv, Kuyalnik Odessa region, Lugansk, Mariupol, Odessa, Sevastopol, Simferopol, Solomonovo Transcarpathian region Uzhgorod district, Uzhgorod, Kharkiv, Ranneye (Chervona Zirka), Ivano-Frankivsk, Lviv, Ternopil. All presented data collected by the Crawler Agent is stored in the Data Warehouse. The result of the functioning of the Crawler Subsystem is presented in Figure 4.

Figure 4: The result of the activity of the Crawler Subsystem

The Figure 4 shows the process of record the data-based actions of the MAS. This process is called Data logging [21]. It displays all actions connected with files storing, accessing or modifying on the Data Warehouse. The logging of the downloading data process from chosen websites allows monitoring

the status of the MAS in online mode. The Figure 4 shows that one of the attempts to download the page has not successful, all the others one have ended successfully. The data logging occurs both to the console and to a file on disk. There is option to add the data logging to the database and / or Elasticsearch used by Serilog logging library, which makes it easy to add logging destinations.

The stored information has to be prepared using the Unifier Subsystem for further analysis. The raw data loaded by the Crawler Subsystem is cleaned, integrated, filtered and identified by the Unifier Subsystem. The considered Subsystem, as well as the Crawler Subsystem, consists of a set of agents working for a common goal (Figure 5).

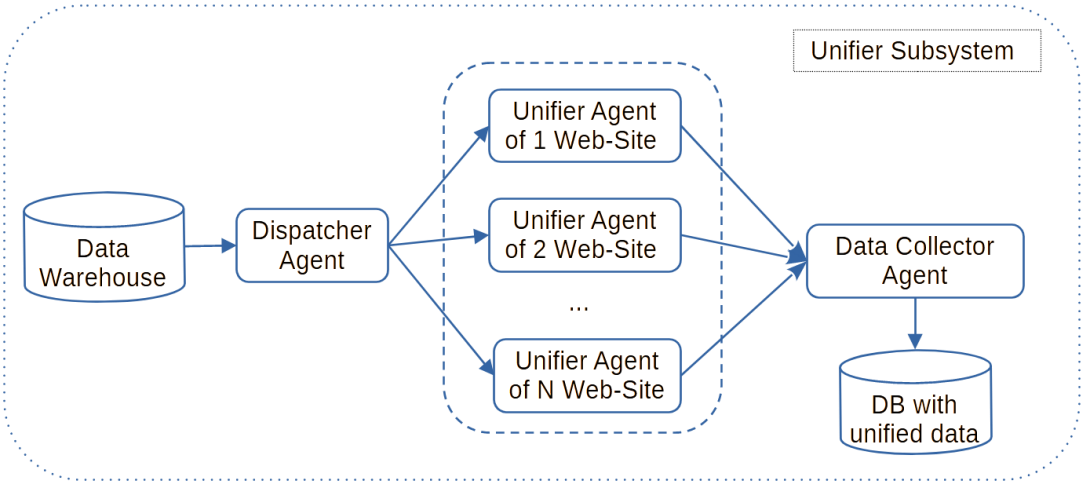

Figure 5: Architecture of Unifier Subsystem

The Dispatcher Agent sequentially reads files from the Data Warehouse and distributes it to all Site Unifier Agents. The Dispatcher Agent processes multiple files in parallel, it allows to increase productivity. Each Site Agent Unifier is designed for one specific file type, depending on the site for which it was developed. It reads from the weather forecast file the following parameters one by one: the name of the settlement, the date/time of forecast generation, temperature, humidity, wind strength and direction, precipitation. If the reading is impossible at any step (for example, the HTML element does not exist), then the file processing will be terminated and the Collector Agent will get a message that the file has not been processed. To obtain the required values from the HTML-document, the XPATH (XML Path Language) query language is used [22]. XPath is used to jump to any tag, attribute or needed text block. To get the data, it is necessary to create a query with a path that describes this data. The paths are the most useful and widely used XPath property. It consists of a set of XPath nodes relative to its starting (most often root) element.

Regular expressions are used for post-processing of the obtained data [23]. Regular expressions are patterns used to match sequences of characters in strings. An example of a regular expression for the Title of the site:

$$"\text{^}\backslash u2602\ (METEO.UA\backslash u2122|Meteo.ua)\text{:}\ Weather\ in\ .*\ during\ 10\ days,$$
$$weather\ forecast\ in\ (\text{в}in\ )\{1,2\}(?'RegExNameGroup'.*)\ 10\ days\backslash."$$

The post-processing of data is necessary, because schema DOM elements can contain raw data. For example: from the line “764 mm. rt. st ”it is necessary to get only 764, values containing percentages have to be cut similarly. The Html Agility Pack and Json.NET libraries from Newtonsof are used, which allow processing HTML and jSon files.

The Collector Agent receives the result of the work of all Unifier Agents. There are two erroneous situations, when the file is considered as problematic and should be placed in a special storage:

1. If all Agents are finished their work with errors. There are two ways for resolving this issue: the development of a new Unifier Agent for such type of file or this file can be deleted.
2. If several Unifier Agents return unified result, but not an error. The correct situation when one and only one Unifier Agent is appropriate for each particular file.

The unified data after processing of Unifier Subsystem has the following structure (Table 2).

Table 2
Data structure

Field	Description
ProviderType	Enumeration value, possible values: MeteoPost.com, Meteoprog.com, Sinoptik.ua, Gismeteo.ua, Meteo.ua, RealWeather
ProviderSubType	Data subtype for each specific type: for a month, for a week, for 2 weeks, hourly, etc., see Table 1
FilePath	The path to the file from which the data was retrieved
DownloadedAt	Date/time when the file was downloaded by the Crawler Subsystem
GeneratedAt	The date/time when the weather data was generated, DownloadedAt >= GeneratedAt
RowNumber	Number of weather records
Periodicity	Periodicity of data, the distance from the start time of one record to the start time of another. For example: the data is hourly, then it will be 1; if the data is daily, then 24; if once every 3 hours, then 3
Place	Enumeration value shows locality the data
Rows	Weather record structure

Weather information is presented as a set of meteorological characteristics. The structure of weather data for a specific locality or settlement is shows in Table 3.

Table 3
Weather record structure

Field	Description
DateTimeFrom/DateTimeTo	Start and end date for the current weather record
Temperature	Temperature, in degrees Celsius
PrecipitationProbability	Probability of precipitation (rain, thunderstorm, snow, etc.), %
PrecipitationLevel	The amount of precipitation for the period of time of the current record, mm
Pressure	Atmospheric pressure, millimeters of mercury
Humidity	Air humidity, %
WindSpeed	Wind speed, m/s
WindDirection	Direction of the wind. Magnetic Azimuth is used. Value from 0 to 360 degrees
WeatherCharacteristics	Weather characteristics. Set elements: {cloudless; mainly cloudless; light rain; rain; heavy rain; storm rain; light rain with snow; rain with snow ets}

The MAS has the possibility to save obtained data in CSV and JSON text formats. CSV is supported by all Big Data processing libraries, Data mining methods, Machine Learning approach. Nevertheless, despite the fact that this format is historically more popular in this area, the choice of CSV or JSON allows tainge into account the personal preferences of the developer. The use of these two open text formats makes the system development more flexible, since Python or R programming is mainly used for static processing. These languages contain many libraries specially tailored for data mining and the search for patterns. This is important for the considered task in this study, because this work is planned to continue in this direction.

The result of the MAS functioning is presented in Figure 6 as a snippet of unified meteorological data. The data is saved in the JSON text format, which is designed for storing and exchanging structured data. The convenience of the chosen format lies in the fact that it can be used in any programming language, and is one of the most common ways to transfer data as well.

```
{
  "Place": "Harkov",
  "ProviderType": "Meteo.Ua",
  "ProviderSubType": "hourly-data",
  "FilePath": "D:\\Projects\\Weather.Data.Test\\Weather.Meteo.Ua\\202302\\15\\14-days\\20230215 15-30-55.440-harkov.html",
  "DownloadedAt": "2023-02-15T17:30:56.5380023+02:00",
  "GeneratedAt": "2023-02-15T15:00:00+02:00",
  "RowNumber": 96,
  "Periodicity": 1,
  "Rows": [{
      "DateTimeFrom": "2023-02-15T00:00:00",
      "DateTimeTo": "2023-02-15T01:00:00",
      "Temperature": -1.47,
      "PrecipitationProbability": 0.0,
      "PrecipitationLevel": 0.0,
      "Pressure": 764.0,
      "Humidity": 94.0,
      "WindSpeed": 4.6,
      "WindDirection": 323.0,
      "WeatherCharacteristics": "Overcast"
    }, {
      "DateTimeFrom": "2023-02-15T01:00:00",
      "DateTimeTo": "2023-02-15T02:00:00",
      "Temperature": -1.68,
      "PrecipitationProbability": 0.0,
      "PrecipitationLevel": 0.0,
      "Pressure": 764.0,
      "Humidity": 90.0,
      "WindSpeed": 5.8,
      "WindDirection": 331.0,
      "WeatherCharacteristics": "Overcast"
```


Figure 6: Snippet of unified data

Thus, the experiment on the preprocessing of meteorological data from different weather forecast providers showed the effectiveness of using the developed MAS. It allowed not only to find and download data, but also to clean and identify it in order to use the presented unified data to solve the problem of assessing the quality of information on weather forecast websites.

6. Discussions

The quality of information or data is an integral characteristic that shows the degree of their suitability for decision making [24]. Many different factors can be used for evaluation of the quality of information or data. Each domain usually indicates its own specific quality criteria, but there are also more general ones. Consider the following characteristics which can help to evaluate the quality of the prepared data:

- q_1 – accuracy – data must be with a given level of detail and accuracy;
- q_2 – reliability – there should be no errors in the data;
- q_3 – time for data processing;
- q_4 – the level of complexity of data processing;
- q_5 – the influence of the human factor.

The study [25] suggests using the following integral convolution Q to assess the quality of the obtained unified data:

$$Q = \frac{\sum_{k=1}^{5} q_k \, w_k}{\sum_{k=1}^{5} w_k},$$

where q_k-th k-th criterion;
w_k – is the weight of k-th criterion.

Each criterion is proposed to be evaluated on a 10-point scale in two ways. First variant is the evaluation without using the developed MAS. Second variant is the assessment of the chosen criteria when the MAS is used to prepare data for subsequent analysis and selection task. At the same time, it should be taken into account that the criteria q_1 and q_2 make a positive contribution to the overall quality assessment, that is, the quality is improved with an increase in these parameters, and the rest criteria are negative. The results of the evaluation of the obtained data are presented in Table 4.

Table 4
The process of quality evaluation

Number of criterion	No MAS	MAS	Weight of criterion	Results of evaluation	
				No MAS	MAS
q_1	7	10	0,9	0,20	0,40
q_2	4	10	0,9	0,11	0,40
q_3	9	2	0,5	0,74	0,92
q_4	9	2	0,4	0,74	0,92
q_5	6	1	0,5	0,83	0,96
Q				0,43	0,63

Analysis of the Table 4 shows that the use of the developed MAS helps to simplify the data preparation process and improves the quality of the obtained unified data by 20%. The gained result proved the feasibility of using the developed MAS for the resolving the task of improving the quality of input information for making managerial decisions connected with the weather forecast.

7. Conclusion

The proposed approach for solving the problem of obtaining higher quality decisions by preparing data for decision making using the MAS has shown its effectiveness. The use of improvised tools or a combination of different software products for data processing can lead to errors due to the influence of the human factor. In addition, it takes quite a long time to evaluate data from several websites with weather data, since it is necessary to make a multi-criteria assessment. Therefore, the experiment and the subsequent expert evaluation of the obtained unified data showed that the MAS makes it possible to simplify the decision-making process.

The direction of further research is the development of another part of the MAS, which will analyze, evaluate and select websites based on the results from the Crawler and Unifier Subsystems.

8. References

[1] How researchers can help fight climate change in 2022 and beyond, Nature 601, 7 (2022). URL: https://www.nature.com/articles/d41586-021-03817-4. doi:10.1038/d41586-021-03817-4.

[2] These energy innovations could transform how we mitigate climate change, and save money in the process – 5 essential reads, 2022. URL: https://theconversation.com/these-energy-innovations-could-transform-how-we-mitigate-climate-change-and-save-money-in-the-process-5-essential-reads-180076.

[3] L. Milliken, T. White, M. A. Fisher, Countering Climate Change With Renewable Energy Technologies, 2021. URL: https://fas.org/blogs/sciencepolicy/countering-climate-change-with-renewable-energy-technologies/.

[4] M. Lave, J. Kleissl, Cloud speed impact on solar variability scaling – Application to the wavelet variability model, Solar Energy 91 (2013) 11-21. URL: https://www.sciencedirect.com/science/article/pii/S0038092X13000406. doi:10.1016/j.solener.2013.01.023.

[5] Using wind turbines to create clouds to combat global warming, 2022. URL: https://innovationorigins.com/en/selected/using-wind-turbines-to-create-clouds-to-combat-global-warming/.

[6] C. S. Kaunda, C. Z. Kimambo, T. K. Nielsen, Hydropower in the Context of Sustainable Energy Supply: A Review of Technologies and Challenges, ISRN Renewable Energy 2012 (2012). URL: https://downloads.hindawi.com/archive/2012/730631.pdf. doi:10.5402/2012/730631.
[7] M. Lee, S. Ohde, K. Y. Urayama, O. Takahashi, T. Fukui, Weather and Health Symptoms, International Journal of Environmental Research and Public Health 15(8):1670 (2018). URL: https://doi.org/10.3390/ijerph15081670.
[8] Weather analysis in aviation is a continuous process, 2017. URL: https://www.wing.com.ua/content/view/18789/52/ (in Russian).
[9] I. Gultepe, R. Sharman, P.D. Williams et al., A Review of High Impact Weather for Aviation Meteorology, Pure and Applied Geophysics 176 (2019) 1869-1921. URL: https://www.researchgate.net/publication/333162625_A_Review_of_High_Impact_Weather_for_Aviation_Meteorology. doi: 10.1007/s00024-019-02168-6.
[10] Weather forecast in Ukraine. URL: https://meteopost.com/ (in Ukrainian).
[11] Weather forecast. URL: https://www.meteoprog.com/ua/.
[12] Weather forecast. URL: https://sinoptik.ua/ (in Ukrainian).
[13] Weather forecast. URL: https://www.gismeteo.ua/ua/ (in Ukrainian).
[14] Weather forecast in Ukraine. URL: https://meteo.ua/ua (in Ukrainian).
[15] G. Guizzi, C. Silvestri, E. Romano, R. Revetria, Advances in Energy and Environmental Science and Engineering, Proceedings of the 4th International Conference on Energy and Environment Technologies and Equipment (EEETE '15), Volume: Advances in Energy and Environmental Science and Engineering, Michigan State University, East Lansing, MI, USA, 2015. URL: https://www.researchgate.net/publication/283449890_A_comparison_of_forecast_models_to_predict_weather_parameters.
[16] T. L. Feudo, E. Avolio, D. Gullì, S. Federico, C. R. Calidonna, A. Sempreviva, Comparison of Hourly Solar Radiation from a Ground–Based Station, Remote Sensing and Weather Forecast Models at a Coastal Site of South Italy (Lamezia Terme), Energy Procedia, Volume 76 (2015) 148-155. URL: https://www.sciencedirect.com/science/article/pii/S1876610215016604. doi: 10.1016/j.egypro.2015.07.884.
[17] F. Alam, M. Salam, N. A. Khalil et al., Rainfall trend analysis and weather forecast accuracy in selected parts of Khyber Pakhtunkhwa, Pakistan. SN Appl. Sci. 3, 575 (2021). URL: https://link.springer.com/article/10.1007/s42452-021-04457-z. doi:10.1007/s42452-021-04457-z.
[18] A. V. Kozachuk, Analysis of weather forecast quality of popular Ukrainian sites. URL: http://ir.lib.vntu.edu.ua/bitstream/handle/123456789/22559/82.Аналіз якості прогнозу погоди популярних українських сайтів. Козачук.pdf?sequence=1&isAllowed=y (in Ukrainian).
[19] M. Sanjay, Why and how to Cross Validate a Model?, 2018. URL: https://towardsdatascience.com/why-and-how-to-cross-validate-a-model-d6424b45261f.
[20] P. Han, J.-H. Sul, Polite Crawler, 2002. URL: https://courses.cs.washington.edu/courses/cse454/02au/polite-crawler.htm.
[21] Data Logging, 2016. URL: https://www.techopedia.com/definition/596/data-logging.
[22] XML Path Language (XPath) 3.0, 2014. URL: https://www.w3.org/TR/xpath-30/.
[23] Regular Expressions. The Open Group Base Specifications Issue 7, 2018 edition IEEE Std 1003.1™-2017, 2018. URL: https://pubs.opengroup.org/onlinepubs/9699919799/.
[24] Management information system – Quality of Information. URL: https://www.tutorialspoint.com/management_information_system/quality_of_information.htm.
[25] K. Melnyk, Model of the monitoring process for early diagnosis of patients' health / Karina Melnyk, Natalia Borysova, Pavel Smolin, Iryna Dehtiarova, Iuliia Gasan and Alla Voida // Proceedings of the 4th International Conference on Informatics & Data-Driven Medicine (IDDM 2021), 2021, Valencia, Spain. – P. 250-259. URL: http://ceur-ws.org/Vol-3038/paper16.pdf.