

**Побудова інформаційно-аналітичної платформи сценарного аналізу на основі великих обсягів слабкоструктурованої інформації.**

**Построение информационно-аналитической платформы сценарного анализа на основе больших объемов слабоструктурированной информации.**

**Construction of Information-Analytical Platform for Scenario Analysis Based on Large Volumes of Weakly Structured Data**

**1. Номер державної реєстрації - 0118U003779.**

**2. Науковий керівник - д.т.н., проф. Бідюк Петро Іванович, Бидюк Петр Иванович, Bidyuk Petro Ivanovych**

**3. Суть розробки, основні результати.**

(укр.) Дослідження спрямовано на побудову інформаційно-аналітичної платформи, яка забезпечить проведення сценарного аналізу соціальних, економічних, екологічних процесів у складних системах з наявністю людського фактору, використовуючи методи поглибленого аналізу великих обсягів слабкоструктурованої інформації, агрегованої з різних джерел, і експертні підходи методології передбачення. Розроблена інформаційно-аналітична платформа забезпечує збір, агрегування, обробку, класифікацію різномірної структурованої і неструктурованої інформації щодо заданої проблеми, виділення основних індикаторів і прогнозування їх поведінки, використання методів якісного аналізу для подолання неповноти, неточності, нечіткості, протирічності наявної інформації, визначення критичних факторів, передбачення плину соціальних, економічних, екологічних процесів у складних системах і побудову сценаріїв щодо них.

Досліджено сучасні підходи, методи та моделі збереження великих обсягів слабо структурованої інформації. Вивчено та структуровано потенційні джерела слабо структурованої інформації та типи документів, що можуть надходити цими джерелами. Запропоновано нові прийоми вирішення задачі вилучення оцінок, суджень та сподівань, а також емоційної забарвленості ситуацій, що описані природною мовою для супроводження процесів прийняття рішень. Реалізовано та протестовано багатокomпонентну структуру інформаційної моделі для вилучення і аналізу емоційної окраски текстів. Розроблено метод оцінювання і підвищення якості експертних оцінок парних порівнянь в методі аналітичних ієрархій. Запропоновано спосіб моделювання суджень експерта-реаліста, невизначеність якого обумовлена тільки шкалою Сааті, а також експертів-песиміста і оптиміста, в судженнях яких присутня додаткова невизначеність, обумовлена цими особистими якостями експерта. Розроблено критерії опису якості вхідних даних, що характеризують їх невизначеність, для використання модифікованого методу морфологічного аналізу. Запропоновано ряд нових прийомів і процедур модифікованого методу морфологічного аналізу, що покращують якість результатів методу, використовуючи дані, отримані шляхом обробки великих обсягів слабкоструктурованої інформації. Створено нечіткий модифікований метод морфологічного аналізу. Для розроблених в рамках роботи методів, процедур, алгоритмів створено комплексне програмне забезпечення.

(рос.) Исследование направлено на построение информационно-аналитической платформы, обеспечивающей проведение сценарного анализа социальных, экономических, экологических процессов в сложных системах с наличием человеческого фактора с использованием методов углубленного анализа больших объемов слабоструктурированной информации, агрегированной из разных источников, и экспертные подходы методологии предвидения. Разработанная информационно-аналитическая платформа обеспечивает сбор, агрегирование, обработку, классификацию разнородной структурированной и

неструктурированной информации по заданной проблеме, выделение основных индикаторов и прогнозирование их поведения, использование методов качественного анализа для преодоления неполноты, неточности, нечеткости, противоречивости имеющейся информации, определение критических факторов, предвидение протекания социальных, экономических, экологических процессов в сложных системах и построение сценариев относительно них.

Исследовано современные подходы, методы и модели хранения больших объемов слабоструктурированной информации. Изучено и структурировано потенциальные источники слабоструктурированной информации и типы документов, которые могут поступать из этих источников. Предложено новые приемы решения задачи извлечения оценок, суждений и ожиданий, а также эмоциональной окраски ситуаций, описанных естественным языком, для сопровождения процессов принятия решений. Реализовано и протестировано многокомпонентную структуру информационной модели для извлечения и анализа эмоциональной окраски текстов. Разработано метод оценивания и повышения качества экспертных оценок парных сравнений в методе аналитических иерархий. Предложено способ моделирования суждений эксперта-реалиста, неопределенность которого обусловлена только шкалой Саати, а также экспертов-пессимиста и оптимиста, в суждениях которых присутствует дополнительная неопределенность, обусловленная этими качествами эксперта. Разработано критерии описания качества исходных данных, характеризующие их неопределенность, для использования модифицированного метода морфологического анализа. Предложено ряд новых приемов и процедур модифицированного метода морфологического анализа, улучшающих качество результатов метода при использовании данных, полученных путем анализа слабоструктурированной информации. Создан нечеткий модифицированный метод морфологического анализа. Для разработанных в рамках работы методов, процедур, алгоритмов создано комплексное программное обеспечение.

**(eng.)** The research was aimed at creating an information-analytical platform that provides the conduct of scenario analysis for social, economic, ecological processes in complex systems with human factor, using deep analysis methods of large volumes of weakly structured data, aggregated from different sources, and expert approaches of foresight methodology. The developed information-analytical platform provides collection, aggregating, processing, classification of heterogeneous structured and unstructured information on a given problem, highlighting the main indicators and predicting their behavior, employing qualitative analysis methods to cope with incompleteness, lack of precision, fuzziness, contradictive nature of present data, determining critical factors, foresight of social, economic, ecological processes in complex systems and constructing scenarios regarding them.

The modern approaches, methods and models for storing large volumes of weakly structured data were studied. The potential sources of weakly structured data, as well as types of documents from these sources, are studied and classified. New techniques for solving the problem of extracting ratings, opinions and expectations, as well as emotional coloring of situations described with natural language are introduced, providing decision-making support. A multi-component information model structure for extraction and analysis of text emotional coloring was implemented and tested. A method for assessing and improving quality of expert assessment in analytical hierarchy method was developed. A technique for modeling realistic expert opinions, where the uncertainty is stipulated only by Saaty scale, pessimistic and optimistic expert opinions, that have additional uncertainty related to these qualities, was suggested. The criteria for quality of input data describing their uncertainty for modified morphological analysis method, were developed. A number of new techniques and procedures for modified morphological analysis method were suggested, improving the method results when using weakly structured information as input data/ A fuzzy modified morphological analysis method was created. All the developed within the research methods, procedures and algorithms were implemented in a complex software tool set.

#### **4. Наявність охоронних документів на об'єкти права інтелектуальної власності.**

Немає.

#### **5. Порівняння зі світовими аналогами.**

Науково-технічний рівень виконаної роботи відповідає світовому рівню, а розроблена інформаційно-аналітична платформа, а також створене математичне і методологічне забезпечення сценарного аналізу на основі великих обсягів слабкоструктурованої інформації не має аналогів у світовій практиці інформаційних технологій.

Відомі модифікації методу аналізу ієрархій ґрунтуються на припущенні, що оцінки експерта, задані в шкалах, не містять помилок і перевіряти їх якість немає необхідності. В даній роботі розроблено метод оцінювання невизначеності експертних парних порівнянь в задачі обчислення ваг альтернатив рішень. В результаті альтернативам рішень ставляться у відповідність довірчі інтервали для відносних ваг альтернатив. Ці інтервали більш достовірно, в порівнянні з точковими вагами, обчисленими, зокрема, відомим методом ЕМ, відображають відносну важливість альтернатив рішень.

В більшості сучасних робіт, що використовують метод морфологічного аналізу, цей метод розглядається тільки як допоміжний інструмент для структурованого перебору варіантів. Попередні дослідження модифікованого методу морфологічного аналізу працювали в умовах повноти і чіткості вхідних даних, тоді як в роботі створено нові підходи, прийоми, процедури, що дозволяють працювати в умовах неповноти, нечіткості, протирічності вхідних даних в модифікованому методі морфологічного аналізу.

#### **6. Економічна привабливість для просування на ринок.**

Реалізація аналітичних напрацювань у вигляді людино-машинної системи прийняття управлінських рішень із застосуванням найсучасніших комп'ютерних інформаційних технологій та програмного забезпечення дасть змогу адекватно і своєчасно розробляти і адаптувати стратегії в умовах постійного зростання об'ємів актуальної інформації і змін у навколишньому світі, при цьому враховуючи існування суттєвих невизначеностей та багатофакторних ступенів і рівнів ризику, що гарантує сталий розвиток досліджуваної системи відповідно до умов розвитку, обраних у сценаріях ключових технологій та галузей.

За допомогою ситуаційно-аналітичного центра ПСА, на базі якого реалізуються відповідні методи, створюється можливість відслідковувати дані щодо перебігу обраних соціальних, екологічних, економічних процесів, прогнозувати і передбачати їх поведінку в майбутньому.

#### **7. Потенційні користувачі.**

Застосування методології сценарного аналізу може бути впроваджено в проектах організацій, що потребують засобів підтримки прийняття рішень в умовах надходження великої кількості слабкоструктурованої інформації (наприклад, великі виробничі об'єднання та комерційні установи з глобальними мережами філій, сфери реагування на надзвичайні та різноманітні кризові ситуації, військові сфери, засоби для обробки даних космічних експериментів, тощо), відповідно, отримані результати можуть бути актуальними і використані для потреб міських державних адміністрацій, Міністерства внутрішніх справ України, Міністерства з надзвичайних ситуацій України, Служби безпеки України, Ради з національної безпеки та оборони України, інших організацій і відомств, пов'язаних з обробкою великих обсягів слабкоструктурованої інформації.

#### **8. Стан готовності розробки.**

Методологічно-методичний інструментарій у вигляді інформаційно-аналітичної системи передбачення для опрацювання неструктурованої інформації може бути застосований до впровадження. Основні інструменти імплементовані на базі тестового ситуаційно-аналітичного центру в ПСА.

## 9. Існуючі результати впровадження.

Розроблений інструментарій був використаний у ряді проектів, зокрема у таких:

«Розробка інструментарію моделювання стратегії пом'якшення соціальних лих, викликаних катастрофами і тероризмом», спільний проект з Республікою Молдова; 2018 р.

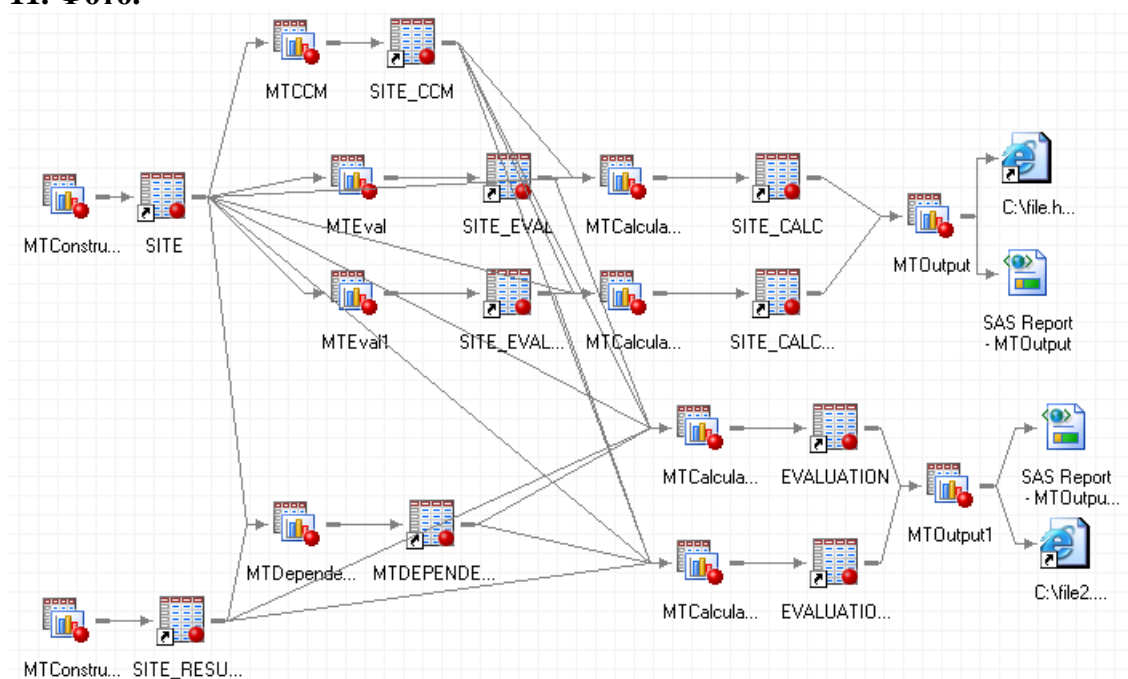
«Оцінювання стану, аналіз та передбачення поведінки складних соціально-економічних систем в контексті якості та безпеки життя людей з використанням даних аерокосмічних спостережень», спільний проект з Інститутом космічних досліджень в рамках національного сегменту проекту Горизонт-2020 ERA-PLANET, 2018 р.

Результати роботи впроваджено в навчальний процес – розроблено нові лекційні курси “Web Mining”, “Text Mining” та оновлено цикли лабораторних робіт із лекційних курсів “Системний аналіз”, “Інтелектуальні системи підтримки прийняття рішень”, “Прийняття рішень в ієрархічних системах”.

## 10. Назва підрозділу, телефон, e-mail.

Інститут прикладного системного аналізу НТУУ «КПІ ім. Ігоря Сікорського», Інститут прикладного системного аналізу, 204-84-47, natalidmp@gmail.com

## 11. Фото.



Мережа модулів модифікованого методу морфологічного аналізу в розробленому програмному забезпеченні, що імплементує модель передпроектного оцінювання ділянок підземного будівництва

## 12. Перелік публікацій за матеріалами досліджень за період виконання.

*Монографії:*

1. Панкратова Н.Д., Гайко Г.І., Савченко І.О. Розвиток підземної урбаністики як системи альтернативних проектних конфігурацій. – К.: Наукова думка, 2020. – 152 с.

*Навчальні посібники:*

1. Панкратова Н.Д. Системний аналіз. Теорія та застосування. Наук. Думка. – К. – 2018. – 348 с.

*Публікації, що входять до наукометричних баз SCOPUS, Web of Science:*

1. Pankratova N., Savchenko I., Gayko G., Kravets V. Evaluating Perspectives of Urban Underground Construction Using Modified Morphological Analysis Method// Journal of Automation and Information Sciences. – 2018. – Volume 50, Issue 10. – P. 34-46. (Scopus). DOI: 10.1615/JAutomatInfScien.v50.i10.30

2. Haiko H. I., Savchenko I.O., Matviichuk I.O. Development of a morphological model for territorial development of underground city space// *Naukovyi Visnyk NHU*, 2019, № 3. – P. 92-98. (Scopus). DOI:10.29202/nvngu/2019-3/14
3. Savchenko I.O. Using Morphological Table Networks for Modeling Social Disaster Situations //2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC), 8-12October 2018, Kyiv, Ukraine, Igor Sikorsky Kyiv Polytechnic Institute. – Pp. 105-108. DOI:10.1109/SAIC.2018.8516797
4. Pankratova N.D., Nedashkovskaya N.I. Evaluation of Ecology Projects for Black Sea Odessa Region on Basis of a Network BOCR Criteria Model // 2018 IEEE First International Conference on System Analysis & Intelligent Computing (SAIC), 8-12 October 2018, Kyiv, Ukraine, Igor Sikorsky Kyiv Polytechnic Institute. – Pp. 31-35.
5. Gorelova G.V., Pankratova N.D., Sstrategy of complex systems development based at the synthesis of methodologies foresight and cognitive modelling //Proceedings of IEEE First International Conference on System Analysis & Intelligent Computing conference (SAIC) 08-12 October, 2018 Kyiv, Ukraine. – P. 13-19.
6. Gorelova G.V., Pankratova N.D., Scientific Foresight and Cognitive Modeling of Socio-Economic Systems //18th IFAC Conference on Technology, Culture and International Stability TECIS2018 Volume 51, Issue 30, 2018, Baku, Azerbaijan, 13–15 September 2018 Pages 145-149
7. Pankratova N. D., Gorelova G.V., Pankratov V.A. Strategy for the Study of Interregional Economic and Social Exchange Based on Foresight and Cognitive Modeling Methodologies //Workshop Proceedings of the 8th International Conference on “Mathematics. Information Technologies. Education”, MoMLeT&DS-2019, Shatsk, Ukraine, June 2-4, 2019. -P136-137.
8. Pankratova N. D., Gorelova G.V., Pankratov V.A System Approach to Assessing of the Quantitative and Qualitative Characteristics of Information In Proceedings of Communicative strategies of the information society(CSIS’19). ACM, October 25-26, 2019, Saint – Petersburg, Russia. ACM, New York, NY, USA
9. I. Savchenko, New Approach for Processing Cross-Consistency Matrix in Modified Morphological Analysis Method, 2020 IEEE 2nd International Conference on System Analysis & Intelligent Computing (SAIC), Kyiv, Ukraine, 2020, pp. 167-172. doi: 10.1109/SAIC51296.2020.9239226
10. H. Haiko, I. Savchenko and I. Matviichuk, A Morphological Analysis Method-Based Model of Assessing Territories for Underground Parking Lots, 2020 IEEE 2nd International Conference on System Analysis & Intelligent Computing (SAIC), Kyiv, Ukraine, 2020, pp. 173-176. (SCOPUS) doi: 10.1109/SAIC51296.2020.9239251
11. V. Volkova, G. Gorelova and N. Pankratova, "The Development of the Cyberphysical System Concept on Base of the Interdisciplinary Theories," 2020 IEEE 2nd International Conference on System Analysis & Intelligent Computing (SAIC), Kyiv, Ukraine, 2020, pp. 1-6, doi: 10.1109/SAIC51296.2020.9239213.
12. Pankratova, N., Bidyuk, P., Golinko, I., Decision support system for microclimate control at large industrial enterprises, CEUR Workshop Proceedings, 2020, 2608, pp. 489-498.

#### *Дисертації:*

1. Зражевська, Н. Г. Методи і моделі прогнозування мір динамічних фондових ризиків : дис. канд. техн. наук : 01.05.04 – системний аналіз і теорія оптимальних рішень, 2018 р. Науковий керівник Н.Д. Панкратова.
2. Недашківська, Н. І. Методологія та інструментарій підтримки прийняття рішень на основі ієрархічних та мережевих моделей : дис. д-ра техн. наук. : 01.05.04 – системний аналіз і теорія оптимальних рішень, 2018. Науковий консультант Н.Д. Панкратова.

### **13. Надати ключові слова до розробки**

Сценарний аналіз, слабкоструктуровані дані, база знань, текстова аналітика, категоризація текстів, аналіз емоційної забарвленості, метод аналізу ієрархій, метод морфологічного аналізу.