

**НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ
«ХАРКІВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»
МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ**

Кваліфікаційна наукова
Праця на правах рукопису

ІВАЩЕНКО ОКСАНА ВІТАЛІЇВНА

УДК 004.85:004.91:519.25

ДИСЕРТАЦІЯ

**МОДЕЛІ, МЕТОДИ ТА ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ
ІНТЕЛЕКТУАЛЬНОЇ ОБРОБКИ ТЕКСТОВОГО ОПИСУ ТОВАРІВ
МАСОВОГО ВЖИТКУ**

Галузь знань 12 - Інформаційні технології

Спеціальність 122 - Комп'ютерні науки

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ О.В. Іващенко

Наукова керівниця:
Чередніченко Ольга Юріївна
доктор технічних наук, доцент

Харків 2023

АНОТАЦІЯ

Іващенко О.В. Моделі, методи та інформаційна технологія інтелектуальної обробки текстового опису товарів масового вжитку. - Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 - Комп'ютерні науки (12 - Інформаційні технології). - Національний технічний університет "Харківський політехнічний інститут" Міністерства освіти і науки України, м. Харків, 2023.

У дисертаційній роботі вирішена науково-практична задача ідентифікації комплексного сприйняття семантики товарних пропозицій на онлайн-платформах електронної комерції шляхом інтелектуальної обробки текстового опису товарів, що дозволяє систематизувати пропозиції товарів та спростити пошук.

Об'єктом дослідження є процес обробки текстової інформації в системах електронної комерції.

Предметом дослідження є моделі, методи та інформаційна технологія інтелектуальної обробки текстового опису товарів масового вжитку.

Метою дисертаційного дослідження є підвищення точності визначення схожості товарів масового вжитку на основі їх текстових описів, представлених в системі електронної комерції.

При вирішенні завдань дисертаційного дослідження використані: метод системного аналізу для розробки концепції інтелектуальної обробки текстового опису товарів масового вжитку; методи математичної статистики, методи теорії прийняття рішень; інструменти NLP для створення моделі текстового опису товарів, моделі оцінки схожості товарів, алгоритму створення тематичного ядра та моделі еталонного товару; методи теорії інтелекту для розробки моделі визначення схожості товарних пропозицій, методи експертного оцінювання для збору даних на основі краудсорсінгу; стандарт IDEF та уніфікована мова моделювання UML для проєктування інформаційної технології та компонентів

програмного забезпечення, компонентно-орієнтовані методи для розробки інформаційної технології інтелектуальної обробки текстових описів.

У вступі обґрунтовано актуальність теми дослідження, сформульовані мета, задачі та методи дослідження, відображено зв'язок дослідження з науковими програмами кафедри, наведено наукову новизну та практичне значення отриманих результатів, зазначено особистий внесок здобувача.

В першому розділі здійснено дослідження стану розвитку електронної комерції, проблем та переваг, які супроводжують даний розвиток, проаналізовано особливості реалізації рекомендаційних систем, та запропоновані можливі покращення щодо реалізації та функціонування рекомендаційних систем. Виявлено, що багато дослідників акцентувало увагу на необхідності обробки текстової інформації представлення товару для систематизації та нормалізації інформації про товари та використання отриманої структурованої інформації для визначення схожості товарів, і, як наслідок для підвищення точності рекомендацій. Проведено аналіз сучасного стану завдань текстової обробки інформації, зокрема вирішення задач розпізнавання сутностей (entity resolution - ER) та визначення схожості товарів (item matching - IM), особливостей реалізації підходів щодо вирішення цих задач, проведено дослідження сучасних методів, підходів та алгоритмів, які використовуються для обробки текстової інформації, проаналізовано характеристики текстового представлення товарів, які використовують в сучасних підходах обробки текстової інформації в системах електронної комерції.

У другому розділі обґрунтовано концепцію інтелектуальної обробки текстової інформації в системах електронної комерції, визначені основні етапи. Обґрунтовано необхідність використання алгоритмів векторизації та методів машинного навчання для обробки неструктурованих текстових даних. Розроблено модель представлення текстового опису товарів, яка є основою для подальшого аналізу та визначення схожості товарів. Запропоновано застосування етапу блокування (blocking) для зменшення складності обробки

великого обсягу текстових описів, що дозволяє за рахунок поділу всього набору текстових даних на групи схожих товарів зменшити розрахункову складність алгоритмів співставлення текстових описів. Запропоновано модель гнучкого пайплайну обробки текстових даних, яка дозволяє оптимізувати параметри блокування за рахунок вибору конфігурації на основі наявної бібліотеки алгоритмів векторизації та кластеризації.

У третьому розділі представлено методи співставлення та групування товарів в системі електронної комерції, розроблена модель визначення схожості товарів, яка базується на запропонованій у другому розділі моделі представлення текстового опису товарів. Запропоновано застосування підходу краудсорсінгу для збору додаткової інформації щодо визначення схожості товарів на основі порівняння зображень товарів та визначення колективної оцінки. Розроблено алгоритм побудови тегового ядра (множини ключових слів) для групи схожих товарів, який дозволяє визначити релевантні до кожного товару теги, які у сукупності характеризують всю групу схожих товарів та є найбільш вживаними в текстових описах товарів на онлайн платформах. Розроблена модель еталонного товару з урахуванням семантики текстового представлення та узагальнених значень атрибутів. Розглянуто пошук товарів з використанням моделі еталонного товару на основі підходу навчання з підкріпленням.

У четвертому розділі представлено розроблену інформаційну технологію інтелектуальної обробки текстових описів товарів та результати проектування компонентів програмного забезпечення. Інформаційна технологія базується на підході щодо створення гнучкого пайплайну обробки текстових даних та включає в себе ряд процесів обробки та перетворення неструктурованої текстової інформації в змістовний структурований набір ключових слів (тегове ядро), які характеризують групу схожих товарів. Представлено прототип архітектурного рішення, яке використовує принципи компонентно-орієнтованої розробки. Також наведено основні результати експериментальних досліджень та апробації розроблених моделей, підходів, алгоритмів щодо інтелектуальної

обробки текстової інформації, а також практичне впровадження підходу. Результати експериментів підтверджують працездатність запропонованих моделей.

У висновках коротко охарактеризовано вирішені завдання дисертаційної роботи, представлено теоретичну та практичну значущість отриманих результатів, наведено дані про впровадження та апробацію основних результатів дослідження.

Наукова новизна проведеного дослідження полягає в наступному:

1 Удосконалено модель текстового опису товарів масового вжитку, що на відміну від існуючих базується на ідентифікації комплексного сприйняття людиною товарних пропозицій, що дозволяє підвищити точність визначення відповідності реальному товару за рахунок врахування семантики трьох складових опису (назва, характеристика, специфікація).

2 Удосконалено модель схожості товарів, яка дозволяє порівнювати споживчі характеристики товарів електронної комерції на основі їх текстового опису та, на відміну від існуючих, враховує вподобання споживачів та сприйняття текстового опису, що дозволяє систематизувати пропозиції та спростити пошук.

3 Дістала подальшого розвитку інформаційна технологія інтелектуальної обробки текстових описів товарів за рахунок вдосконалення моделі текстового опису товару, групування схожих товарів та представлення групи схожих товарів ядром тегів та еталонним товаром, що на відміну від існуючих підходів, дозволяє підвищити ефективність пошуку та ідентифікації товарів.

Результати дисертаційної роботи впроваджено у навчальному процесі кафедри програмної інженерії та інтелектуальних технологій управління Національного технічного університету «Харківський політехнічний інститут» (акт від 14.11.2022 р.), використано у науково-дослідних роботах "Розробка інформаційно-аналітичного забезпечення управління ефективністю та якістю в складних системах за умови євроінтеграції України» (ДР № 0117U004806),

«Створення моделей та методів збору та автоматизованої переробки бізнес-інформації у веб-просторі» (ДР № 0119U002556), «Розробка методів моніторингу актуальних даних в системах організаційного управління» (ДР № 0121U108870), у яких здобувачка брала участь як виконавиця окремих етапів (акт від 01.11.2022 р.) та впроваджено для вирішення практичних завдань в діяльності ТОВ Фінпрогруп» (довідка від 06.01.2023 р.).

Ключові слова: інформаційна технологія, модель, текст, обробка природньої мови, програмна система, семантика, семантичне ядро, нейронна мережа, співставлення товарів, класифікація, електронна комерція.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

Публікації здобувача за темою дисертації, в яких опубліковані основні наукові результати дисертації:

1 Cherednichenko O., Vovk M., Ivashchenko O. Item matching based on collection and processing customer perception of images. *ICTERI : Proceedings 16th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer*. Volume II: Workshops, 2020; Kharkiv; Ukraine; 6-10 October 2020. 2020. – CEUR Workshop Proceedings, Vol-2732. pp. 329-337 337. (Scopus, Німеччина).

2 Cherednichenko O., Vovk M., Ivashchenko O., Baggia A., Stratiienko N.. Improving Item Searching On Trading Platform Based On Reinforcement Learning Approach. *Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2021)*. Volume I: Main Conference , Lviv, Ukraine, April 22-23, 2021. Lviv. 2021. pp. 1444-1455. (Scopus, Німеччина)

3 Cherednichenko O., Ivashchenko O., Lincényi M., Kováč M. Information technology for intellectual analysis of item descriptions in e-commerce. *Entrepreneurship and Sustainability Issues*. Vilnius, Lithuania. 2023. 11(1): 178-190. (Web of Science, Литва).

4 Cherednichenko O., Ivashchenko O., Cibák L. and Lincenyi M.. Item Matching Model in E-Commerce: How Users Benefit. *Economics and Culture*. Warsaw, Poland. 2023. vol.20, no.1, pp.77-90. (Scopus, Польща).

5 О. Ю. Чередніченко, Ю. М. Гонтар, О. В. Іващенко, М. А. Вовк. Аналіз компонентно-орієнтованих методів розробки програмного забезпечення для електронного бізнесу. *Вісник Вінницького політехнічного інституту*. – Вінниця, 2018. Випуск 2. с. 80-88. (Б).

6 О. Ю. Чередніченко, О. В. Іващенко, Ю. М. Гонтар, Б. М. Ворона. Інтелектуальний аналіз пропозицій товарів на основі контекстних рекомендацій. *Вісник Національного технічного університету «ХПІ»*. Серія:

Системний аналіз, управління та інформаційні технології. Харків, 2018. № 44 (1320). с.57-66. (Б).

7 О. Ю. Чередніченко, О. В. Янголенко, О. В. Іващенко, О. М. Матвєєв. Моделі формування рекомендацій у інтелектуальних системах електронної комерції. *Системи обробки інформації*. Харків, 2020. Випуск 1(160). с. 32-39. (Б).

Публікації здобувача за темою дисертації, які засвідчують апробацію матеріалів дисертації:

8 Р. І. Манєва, О. С. Ісаков, О. В. Іващенко Підготовка даних як один з найважливіших етапів інтелектуального аналізу процесів. *Інформаційні технології : наука, техніка, освіта, здоров'я: тези доповіді 26 міжнародної науково-практичної конференції (MICROCAD-2018)*, м. Харків, 16–18 трав. 2018 р. Харків. 2018. ч. 1. с. 24.

9 Yanholenko O., Sokolov D., Ivashchenko O.. Issues of Topical Search and Items Normalization on the online Marketplaces. *Тези доповідей 27 міжнародної науково-практичної конференції «Інформаційні технології : наука, техніка, освіта, здоров'я» (MICROCAD-2019)*, Харків, 15–17 травня, 2019. Харків. 2019. ч.1. с.20.

10 О. Ю. Чередніченко, О. В. Іващенко, Д. О. Мусіян. Розробка методу групування об'єктів на основі інтелектуального аналізу текстового опису. Контроль і управління в складних системах (КУСС-2020): *Тези XV міжнародної конференції, Вінниця, 8-10 жовтня 2020 року. Секція 5. Інтелектуальні технології в системах управління*. Вінниця. 2020. С.640.

11 О. В. Іващенко, Д. О. Мусіян, О. Ю. Чередніченко. Підвищення якості інформаційного пошуку на електронних ресурсах на основі методів обробки текстової інформації. Експертні оцінки елементів навчального процесу: програма та матеріали XXII міжвузівської науково-практичної конференції, Харків, 21 листопада 2020 р. Харків. 2020. С. 26.

12 М. А. Вовк, О. В. Іващенко, В. С. Павленко, Т. А. Шевченко. Пошук та порівняння товарів на майданчиках електронної торгівлі. Інформаційні технології: наука, техніка, технологія, освіта, здоров'я: тези доповідей XXIX міжнародної науково-практичної конференції MicroCAD-2021, 18-20 травня 2021 р. Харків. 2021. ч.1. с.21.

13 Ivashchenko O., Ľuboš C., Hudáková M.. Exploring customer attitude to online shopping: reasoning of research method. Proceedings of the XIII International Scientific and Practical Conference “Information activity as a component of science development”(April 04–07, 2023). Edmonton, Canada. 2023. Pp. 150-154.

14 Cherednichenko O., Ivashchenko O., Vovk M.. Towards Pipeline Construction for Product Matching Task. 41th International Conference on Organizational Science Development. 1st ed. - E-zbornik. - Maribor: University of Maribor, University Press. Maribor. 2022. pp.165-177.

15 Cherednichenko O., Ivashchenko O. Cibák L. Lincenyi M.. Item matching model in ecommerce: how users benefit. International Scientific Conference „Emerging Trends in Economics, Culture and Humanities (etECH2023)”. Riga, Latvia. 2023. pp. 71-72.

Публікації здобувача за темою дисертації, які додатково відображають наукові результати дисертації:

16 Cherednichenko O., Ivashchenko O., Kirchmayer J. Towards item matching task: systematic review/ Smerom k úlohe prirad'ovania položiek: systematická kontrola. Public Administration and Regional Development. Bratislava. Slovakia. 2023. XIX(1), pp.81-98.

ABSTRACT

Ivashchenko O.V. Models, methods and information technology of intellectual processing of textual description of consumer goods. - Qualifying scientific work on the rights of manuscript.

The dissertation for obtaining the scientific degree of Doctor of Philosophy in specialty 122 - Computer Science (12 - Information Technologies). - National Technical University «Kharkiv Polytechnic Institute» of the Ministry of Education and Science of Ukraine, Kharkiv, 2023.

In the dissertation, the scientific and practical task of identifying the complex perception of the semantics of product offers on online e-commerce platforms is solved by means of intellectual processing of the textual description of products, which allows systematizing product offers and simplifying the search.

The object of research is the proceeding of processing textual information in electronic commerce systems.

The subject of research is models, methods and information technology of intellectual processing of text descriptions of consumer goods.

The aim of the dissertation research is to increase the accuracy of determining the similarity of consumer goods based on their textual descriptions presented in the electronic commerce system.

When solving the tasks of the dissertation research, the following were used: the method of system analysis for the development of the concept of intellectual processing of the textual description of consumer goods; methods of mathematical statistics, methods of decision-making theory; NLP tools for creating a product text description model, a product similarity assessment model, a tag core creation algorithm, and a reference product model; methods of the theory of intelligence for the development of a model for determining the similarity of product offers, methods of expert evaluation for data collection based on crowdsourcing; the IDEF standard and the unified UML modeling language for designing information technology and

software components, component-oriented methods for the development of information technology for intellectual processing of text descriptions.

In the introduction, the relevance of the research topic is justified, the goal, tasks and methods of the research are formulated, the connection of the research with the scientific programs of the department are shown, the scientific novelty and practical significance of the obtained results are given, and the personal contribution of the candidate is indicated.

In the first chapter, a study of the state of development of e-commerce, problems and advantages accompanying this development was carried out, the peculiarities of the implementation of recommendation systems were analyzed, and possible improvements regarding the implementation and functioning of recommendation systems were proposed. It was revealed that many researchers focused on the need to process the textual information of product presentation to systematize and normalize product information and use the resulting structured information to determine the similarity of products, and as a result, to increase the accuracy of recommendations. An analysis of the current state of tasks of text processing of information was carried out, in particular, solving the tasks of entity resolution (ER) and determining the item matching (IM), the peculiarities of the implementation of approaches to solving these tasks, a study of modern methods, approaches and algorithms that are used for the processing of text information, the characteristics of the text presentation of goods, which are used in modern approaches to the processing of text information in e-commerce systems, are analyzed.

The second chapter substantiates the concept of intellectual processing of text information in e-commerce systems, defines the main stages. The necessity of using vectorization algorithms and machine learning methods for processing unstructured text data is substantiated. The item text description model has been developed, which is the basis for further analysis and determining the similarity of goods. The application of the blocking stage is proposed to reduce the complexity of processing a large volume of text descriptions, which allows to reduce the computational

complexity of algorithms for matching text descriptions due to the division of the entire set of text data into groups of similar goods. A model of a flexible pipeline for processing text data is proposed, which allows you to optimize blocking parameters by choosing a configuration based on the available library of vectorization and clustering algorithms.

The third chapter presents the methods of comparison and grouping of goods in the electronic commerce system, the item similarity model is developed, which is based on the item text description model proposed in the second chapter. It is proposed to use a crowdsourcing approach to collect additional information on determining the similarity of products based on comparing images of products and determining a collective assessment. The item tag core algorithm for building a tag core (a set of keywords) for a group of similar products has been developed, which allows you to determine the tags relevant to each product, which collectively characterize the entire group of similar products and are the most used in text descriptions of products on online platforms. The reference item model was developed taking into account the semantics of the text representation and the generalized values of the artifacts. Product search using the reference item model based on the reinforcement learning approach is considered.

The fourth chapter presents the developed information technology for intelligent processing of text descriptions of goods and the results of designing software components. Information technology is based on an approach to creating a flexible payplan for text data processing and includes a number of processing processes and transformation of unstructured text information into a meaningful structured set of keywords (tag core) that characterize a group of similar products. A prototype of an architectural solution that uses the principles of component-oriented development is presented. The main results of experimental research and testing of the developed models, approaches, algorithms for intelligent processing of text information are also given, as well as the practical implementation of the approach. The results of the experiments confirm the efficiency of the proposed models.

The conclusions briefly describe the solved tasks of the dissertation work, present the theoretical and practical significance of the obtained results, and provide data on the implementation and testing of the main research results.

The scientific novelty of the conducted research is as follows:

1 The item text description model has been improved, which, unlike the existing ones, is based on the identification of complex human perception of product offers, which allows to increase the accuracy of determining the conformity to the real product by taking into account the semantics of the three components of the description (name, characteristic, specification).

2 The item similarity model has been improved, which allows comparing the consumer characteristics of e-commerce products based on their text description and, unlike the existing ones, takes into account consumer preferences and perception of the text description, which allows systematizing offers and simplifying the search.

3 The information technology of intellectual processing of text descriptions of goods has received further development due to the improvement of the item text description model, grouping of similar goods and representation of a group of similar goods by the tag core and a reference item, which, unlike existing approaches, allows to increase the efficiency of the search and identification of goods.

The results of the dissertation work were implemented in the educational process of the Department of Software Engineering and Management Intellectual Technologies of National Technical University «Kharkiv Polytechnic Institute» (act dated 14.11.2022), used in research works "Development of information and analytical support for management of efficiency and quality in complex systems under the conditions of European integration of Ukraine" (RW No. 0117U004806), "Creation of models and methods of collection and automated processing of business information in the web space" (RW No. 0119U002556), "Development of methods for monitoring actual data in organizational management systems" (RW No. 0121U108870), in which the candidate participated as an executor of individual stages (act dated 01.11.2022) and implemented to solve practical tasks in the activities of LLC «Finprogroup" (certificate dated 06.01.2023).

Keywords: information technology, model, text, natural language processing, software system, semantics, semantic core, neural network, item matching, classification, e-commerce.

LIST OF APPLICANT PUBLICATIONS

Publications of the applicant on the topic of the dissertation, in which the main scientific results are published::

1 Cherednichenko O., Vovk M., Ivashchenko O. Item matching based on collection and processing customer perception of images. ICTERI : Proceedings 16th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer. Volume II: Workshops, 2020; Kharkiv; Ukraine; 6-10 October 2020. 2020. – CEUR Workshop Proceedings, Vol-2732. pp. 329-337 337. (Scopus, Germany)

2 Cherednichenko O., Vovk M., Ivashchenko O., Baggia A., Stratiienko N.. Improving Item Searching On Trading Platform Based On Reinforcement Learning Approach. Proceedings of the 5th International Conference on Computational Linguistics and Intelligent Systems (COLINS 2021). Volume I: Main Conference , Lviv, Ukraine, April 22-23, 2021. Lviv. 2021. pp. 1444-1455. (Scopus, Germany)

3 Cherednichenko O., Ivashchenko O., Lincényi M., Kováč M. Information technology for intellectual analysis of item descriptions in e-commerce. Entrepreneurship and Sustainability Issues. Vilnius, Lithuania. 2023. 11(1): 178-190. (Web of Science, Lithuania).

4 Cherednichenko O., Ivashchenko O., Cibák L. and Lincenyi M.. Item Matching Model in E-Commerce: How Users Benefit. Economics and Culture. Warsaw, Poland. 2023. vol.20, no.1, pp.77-90. (Scopus, Poland).

5 O. Yu. Cherednichenko, Yu. M. Hontar, O. V. Ivashchenko, M. A. Vovk. Analiz komponentno-orientovanykh metodiv rozrobky prohramnoho zabezpechennia dlia elektronnoho biznesu. Visnyk Vinnytskoho politekhnichnoho instytutu. – Vinnytsia, 2018. Vypusk 2. s. 80-88. (B).

6 O. Yu. Cherednichenko, O. V. Ivashchenko, Yu. M. Hontar, B. M. Vorona. Intelktualnyi analiz propozytsii tovariv na osnovi kontekstnykh rekomendatsii. Visnyk Natsionalnoho tekhnichnoho universytetu «KhPI». Serii:

Systemnyi analiz, upravlinnia ta informatsiini tekhnolohii. Kharkiv, 2018. № 44 (1320). c.57-66. (B).

7 O. Yu. Cherednichenko, O. V. Yanholenko, O. V. Ivashchenko, O. M. Matvieiev. Modeli formuvannia rekomendatsii u intelektualnykh systemakh elektronnoi komertsii. Systemy obrobky informatsii. Kharkiv, 2020. Vypusk 1(160). c. 32-39. (B).

Publications of the applicant on the topic of the dissertation, which certify the approbation of the dissertation materials:

8 R. I. Manieva, O. S. Isakov, O. V. Ivashchenko Pidhotovka danykh yak ody z naivazhlyvishykh etapiv intelektualnoho analizu protsesiv. Informatsiini tekhnolohii : nauka, tekhnika, osvita, zdorovia: tezy dopovidi 26 mizhnarodnoi naukovo-praktychnoi konferentsii (MICROCAD-2018), m. Kharkiv, 16–18 trav. 2018 r. Kharkiv. 2018. ch. 1. s. 24.

9 Yanholenko O., Sokolov D., Ivashchenko O.. Issues of Topical Search and Items Normalization on the online Marketplaces. Tezy dopovidei 27 mizhnarodnoi naukovo-praktychnoi konferentsii «Informatsiini tekhnolohii : nauka, tekhnika, osvita, zdorovia» (MICROCAD-2019), Kharkiv, 15–17 travnia, 2019. Kharkiv. 2019. ch.1. s.20.

10 O. Yu. Cherednichenko, O. V. Ivashchenko, D. O. Musiiian. Rozrobka metodu hrupuvannia obiektiv na osnovi intelektualnoho analizu tekstovoho opysu. Kontrol i upravlinnia v skladnykh systemakh (KUSS-2020) : Tezy KhV mizhnarodnoi konferentsii, Vinnytsia, 8-10 zhovtnia 2020 roku. Sektsiia 5. Intelektualni tekhnolohii v systemakh upravlinnia. Vinnytsia. 2020. S.640.

11 O. V. Ivashchenko, D. O. Musiiian, O. Yu. Cherednichenko. Pidvyshchennia yakosti informatsiinoho poshuku na elektronnykh resursakh na osnovi metodiv obrobky tekstovoi informatsii. Ekspertni otsinky elementiv navchalnoho protsesu: prohrama ta materialy XXII mizhvuzivskoi naukovo-praktychnoi konferentsii, Kharkiv, 21 lystopada 2020 r. Kharkiv. 2020. S. 26.

12 M.A. Vovk, O.V. Ivashchenko, V.S. Pavlenko, T.A. Shevchenko. Poshuk ta porivniannia tovariv na maidanchykakh elektronnoi torhivli. Informatsiini tekhnolohii: nauka, tekhnika, tekhnolohiia, osvita, zdorovia: tezy dopovidei KhXIKh mizhnarodnoi naukovo-praktychnoi konferentsii MicroCAD-2021, 18-20 travnia 2021 r. Kharkiv. 2021. ch.1. s.21.

13 Ivashchenko O., Ľuboš C., Hudáková M.. Exploring customer attitude to online shopping: reasoning of research method. Proceedings of the XIII International Scientific and Practical Conference “Information activity as a component of science development”(April 04–07, 2023). Edmonton, Canada. 2023. Pp. 150-154.

14 Cherednichenko O., Ivashchenko O., Vovk M.. Towards Pipeline Construction for Product Matching Task. 41th International Conference on Organizational Science Development. 1st ed. - E-zbornik. - Maribor: University of Maribor, University Press. Maribor. 2022. pp.165-177.

15 Cherednichenko O., Ivashchenko O. Cibák L. Lincenyi M.. Item matching model in ecommerce: how users benefit. International Scientific Conference „Emerging Trends in Economics, Culture and Humanities (etECH2023)”. Riga, Latvia. 2023. pp. 71-72.

Publications of the applicant on the topic of the dissertation, in which additionally the scientific results of the dissertation are published reflect:

16 Cherednichenko O., Ivashchenko O., Kirchmayer J. Towards item matching task: systematic review/ Smerom k úlohe prirad'ovania položiek: systematická kontrola. Public Administration and Regional Development. Bratislava. Slovakia. 2023. XIX(1), pp.81-98.

ЗМІСТ

Перелік познач та скорочень.....	5
Вступ.....	7
1 Аналіз стану питання та постановка завдань дослідження	14
1.1 Особливості реалізації рекомендаційних систем електронної комерції	14
1.2 Аналіз сучасного стану завдань обробки текстового опису товарів масового вжитку в системах електронної комерції	20
1.2.1 Основи задачі розв'язання сутностей.....	20
1.2.2 Золотий стандарт завдання ER	25
1.3 Методи та засоби інтелектуальної обробки текстової інформації в системах електронної комерції.....	29
1.4 Постановка задач дослідження	42
1.5 Висновки до розділу 1	45
2 Основні принципи створення інформаційної технології інтелектуальної обробки текстового опису товарів масового вжитку	47
2.1 Концепція інтелектуальної обробки текстового опису товарів електронної комерції.....	47
2.2 Обґрунтування вибору математичного інструментарію.....	54
2.2.1 Моделі і алгоритми у системах підтримки прийняття рішень з нечіткими результатами.....	54
2.2.2 Векторизація	55
2.3 Розробка моделі текстового опису товарів.....	65
2.4 Блокування.....	69
2.5 Висновки до розділу 2	74
3 Розробка методів зіставлення та групування товарів електронної комерції	76
3.1 Метод створення груп схожих товарів на основі текстових описів	76

3.1.1	Визначення схожості товарів на основі збору та обробки сприйняття зображень клієнтами	76
3.1.2	Модель схожості товарів	85
3.2	Алгоритм створення тегового ядра	89
3.3	Пошук пропозицій на торговій платформі на основі підходу до навчання з підкріпленням	91
3.4	Модель еталонного товару	92
3.5	Висновки до розділу 3	96
4	Проектування інформаційної технології інтелектуальної обробки текстового опису товарів масового вжитку	97
4.1	Синтез інформаційної технології інтелектуальної обробки текстового опису товарів електронної комерції	97
4.1.1	Кратка характеристика методики IDEF0	97
4.1.2	Моделювання інформаційної технології інтелектуальної обробки текстових описів	97
4.2	Проектування програмних компонентів інтелектуальної системи	106
4.2.1	Компонентно-орієнтована розробка	106
4.2.2	Проектування пайплайну	108
4.3	Експериментальні дослідження	112
4.3.1	Краудсорсинг	112
4.3.2	Експерименти з гнучким пайплайном	119
4.3.3	Експеримент по пошуку товарів на онлайн платформах за застосуванням навчання з підкріпленням	129
4.3.4	Формування рекомендацій у інтелектуальних системах електронної комерції	138
4.4	Висновки до розділу 4	143
	Висновки	144
	Список джерел інформації	146
	Додаток А Список публікацій здобувача	163

Додаток Б Документи, які підтверджують запровадження основних результатів дисертаційної роботи	167
Додаток В Глосарій.....	171