

В.О. ЛЕЩИНЬСКИЙ

ПРОЦЕСНИЙ ПІДХІД ДО ПОБУДОВИ МОДЕЛІ КОНТЕКСТУ ПОЯСНЕНЬ В ІНТЕЛЕКТУАЛЬНІЙ ІНФОРМАЦІЙНІЙ СИСТЕМІ

Розглянуто проблему побудови пояснень з урахуванням контексту прийняття рішень в інтелектуальній системі. Виконано структурування контексту пояснення з урахуванням явних та неявних знань щодо предметної області, а також об'єктних та темпоральних обмежень щодо процесу прийняття рішення. Запропоновано артефактну модель контексту пояснень, що враховує сукупність каузальних та темпоральних обмежень по роботі із об'єктами предметної області. Розроблено процесно-орієнтований метод побудови моделі контексту пояснень, який дає можливість сформувати модель з урахуванням як явних, так і неявних знань щодо предметної області.

1. Вступ

Сучасні інтелектуальні інформаційні системи використовують складні алгоритми, що базуються на методах машинного навчання і тому не завжди є зрозумілими для користувачів. Така непрозорість процесу прийняття рішень може призвести до недовіри користувачів до отриманих результатів, і, як наслідок, затримок у їх практичному використанні. Наприклад, недовіра до системи розпізнавання образів може знизити рівень безпеки підприємства, яке використовує ці системи для відеонагляду; недовіра до рекомендованих товарів у рекомендаційній системі може привести до відтоку користувачів із відповідної системи електронної комерції [1, 2].

Для вирішення проблеми непрозорості процесу отримання рішення в інтелектуальній системі використовується механізм пояснення. Пояснення дозволяють користувачам зрозуміти ключові каузальні або темпоральні залежності, які привели до відповідного рішення інтелектуальної системи як "чорного ящика" [3, 4].

Такі залежності є мультिवаріантними, оскільки вони залежать від контексту прийняття рішень. Останній визначає конкретну ситуацію, в якій безпосередньо приймається те чи інше рішення. Ситуація прийняття рішення являє собою середовище, що визначає множину умов, в яких формується та приймається рішення. Фактично контекст надає обмеження, в межах яких оцінюється та реалізується отримане рішення.

Контекст як опис ситуації прийняття рішення може містити у собі: цілі реалізації рішення; завдання, що потрібні для досягнення цілі; наявні ресурси, умови та обмеження на вирішення задач. Останні визначаються часовими рамками прийняття та реалізації рішення, потребами зацікавлених сторін, бажаними правовими та етичними характеристиками рішення.

Контекст прийняття рішень відіграє ключову роль у побудові процесу прийняття рішень, визначаючи допустимі у визначеній ситуації варіанти результуючих рішень. Тому врахування контексту є необхідним для ефективного прийняття рішень.

Аналіз контексту дає можливість визначити потенційні ризики імплементації рішення і, на цій основі, розглянути альтернативні варіанти рішення. Крім того, контекст, який містить умови та обмеження конкретного користувача, забезпечує індивідуалізацію отриманих рішень.

Тому врахування контексту при побудові пояснень має практичну значущість, а побудова моделей контексту прийняття рішень щодо пояснень в системах штучного інтелекту є актуальною задачею.

2. Аналіз літературних даних і постановка проблеми дослідження

Сучасні підходи до формування пояснень були сформовані в рамках програми самопояснювального штучного інтелекту - Explainable Artificial Intelligence (XAI) [5]. В цій програмі суттєва увага приділяється напрямку побудови пояснень з використанням причинно-наслідкових залежностей. Такі залежності між змінними визначаються з використанням байєсовського підходу [6]. Крім того, каузальні закономірності можуть бути встановлені й для структурованих об'єктів, що складаються з набору змінних [7]. Проте при поясненні процесу прийняття рішення в інтелектуальній системі розглядається і темпоральний аспект, який є відображен-

ням каузальних залежностей у динамічному середовищі. Як правило, темпоральний аспект враховується на основі виділення підмножини даних для певного проміжку часу [8].

Темпоральні залежності були виділені у формі зважених правил та проаналізовані в роботах [9-12]. В роботах [9, 10] було визначено темпоральні правила та запропоновано використати їх для формування множини альтернатив як складових процесу підтримки прийняття управлінських рішень. В роботі [11] запропоновано використати темпоральні правила та відношення між об'єктами для опису процесу прийняття рішень в інтелектуальній системі. В роботі [12] розглянуто проблему формування каузальних залежностей на основі темпоральних правил.

Таким чином, сучасні дослідження у сфері побудови пояснень орієнтовані, в першу чергу, на встановлення причинно-наслідкових залежностей між вхідними та проміжними даними, а також отриманим результатом. Частково враховується також темпоральний аспект на основі виділення часового інтервалу актуальності даних. Дослідження щодо опису процесу прийняття рішень у темпоральному аспекті свідчать про важливість використання темпоральних правил для формування процесу прийняття рішень. У сукупності результати проведених досліджень дають можливість описати процес прийняття рішень з використанням каузальних та темпоральних залежностей, що може служити основою для формування пояснень.

Однак існуючі дослідження не приділяють достатньо уваги побудові пояснень з урахуванням контексту, що має суттєвий вплив на їх актуальність та відповідність вимогам користувача. Тому розробка процесного підходу до моделювання контексту пояснень є актуальною задачею.

3. Мета і задачі дослідження

Метою даного дослідження є розробка моделі контексту пояснень інтелектуальної системи та методу її побудови з урахуванням процесного опису об'єктів, які використовуються в рамках пояснення, а також актуальності даних щодо цих об'єктів. Досягнення цієї мети створює умови для побудови пояснення за принципом чорного ящика, на основі лише інформації про підмножину подій, які виникали у процесі прийняття рішення в інтелектуальній інформаційній системі.

Для досягнення мети роботи вирішуються такі задачі:

- структуризація контексту пояснення з урахуванням явних та неявних знань щодо предметної області;
- розробка моделі контексту пояснень на основі сукупності умов та обмежень на процес прийняття рішення в інтелектуальній системі;
- розробка процесно-орієнтованого методу побудови моделі контексту пояснення з урахуванням темпорального аспекту обробки артефактів у інтелектуальній системі.

4. Структуризація контексту пояснень в системі штучного інтелекту

Контекст пояснень відноситься до конкретних обставин прийняття рішення в інтелектуальній системі і представлений інформацією та знаннями, що мають вплив на створення та використання пояснень. Фактично контекст надає додаткову інформацію, яка дає можливість користувачеві інтелектуальної системи зрозуміти конкретний варіант пояснення, причини отриманого рішення та впевнитись в релевантності останнього. Контекстна інформація також впливає безпосередньо на вибраний варіант пояснення.

Аналіз структури контексту пояснень дає можливість виділити такі його ключові складові: дані; явні знання; потреби користувача системи штучного інтелекту; неявні знання; обмеження на формування та на використання пояснень.

Дані характеризують стан предметної області та містять опис фактів, подій, а також умов і обмежень щодо їх використання, що суттєво впливають на пояснення і тому мають бути розтлумачені за додатковим запитом користувача.

Явні знання містять залежності, які обумовлюють або обмежують процес прийняття рішення в інтелектуальній системі і тому мають бути враховані в рамках пояснення.

Потреби користувача системи штучного інтелекту визначають спосіб використання отриманого результату і тому мають бути враховані у поясненні.

Неявні знання є неформалізованими та недокументованими знаннями, які мають практичне застосування в предметній області і, відповідно, впливають на процес прийняття рішення в інтелектуальній системі. Тому ці знання мають бути формалізовані для подальшого використання при побудові пояснення.

Неявні знання можна розділити на безпосередньо знання та досвід. Підходи до формалізації цих знань відрізняються. Перша категорія неявних знань (implicit-знання) може бути формалізована шляхом опитування експертів в предметній області. Досвід же не має вербального опису і тому не може бути формалізований шляхом інтерв'ювання. Досвід (tacit-знання) може бути, як правило, отриманий на основі фіксації та повторення дій користувача або аналізу логів інтелектуальної інформаційної системи.

Обмеження на формування та використання пояснень мають об'єктний та темпоральний аспекти. Темпоральний аспект задає часові рамки для використання пояснення. Об'єктний аспект охоплює ресурсні обмеження, юридичні вимоги, формалізовані вимоги до практичного застосування результуючого рішення інтелектуальної інформаційної системи.

Узагальнено ключові складові контексту пояснень представлено в таблиці 1.

Таблиця 1

Складові контексту пояснень

Елемент	Властивості
1. Дані	Факти, стани, події, умови та обмеження по них, що мають вплив на пояснення
2. Явні знання	Формалізовані залежності, які обумовлюють, визначають альтернативи або обмежують процес прийняття рішення
3. Потреби користувача	Способи використання отриманого результату, які є прийнятними для користувача інтелектуальної системи
4. Неявні знання	Implicit-знання, які можуть бути отримані шляхом опитування користувачів або експертів у предметній області, мозкових штурмів, тощо
	Tacit-знання, які зазвичай не переводяться у вербальну форму; такі знання можуть бути отримані шляхом аналізу поведінки користувачів, логів інформаційної системи, з якою працює користувач, тощо
5. Обмеження на формування та використання пояснень	Об'єктний аспект визначає обмеження на дії з об'єктами (даними про об'єкти) предметної області у процесі прийняття рішень в інтелектуальній системі
	Темпоральний аспект обмеження визначає часові рамки для використання пояснення, а також часові рамки дій в інтелектуальній системі, які потребують пояснення

Між наведеними в таблиці 1 властивостями контексту пояснень існує ряд залежностей. По-перше, дані зазвичай є елементами знань про предметну область і можуть виступати як антецеденти причинно-наслідкових залежностей.

По-друге, для побудови та персоналізації пояснень необхідно поєднати знання про предметну область: явні, неявні, знання про використання рішення інтелектуальної системи. Явні та неявні знання про процес прийняття рішення дають можливість відібрати підмножину пояснень щодо поточного результату, а знання про спосіб використання результату - персоналізувати пояснення.

По-третє, розглянута сукупність знань визначає умови та обмеження для пояснення. Умови визначають підмножину варіантів пояснення для користувача, а обмеження мають бути виконані для всіх альтернатив пояснення.

Об'єктний та темпоральний аспекти обмежень для пояснення дають можливість розглядати та тлумачити процес прийняття рішення в системі штучного інтелекту у статичній та динамічній. Об'єктний аспект дає можливість визначити підмножину об'єктів для пояснення, а також підмножину допустимих дій з цими об'єктами. Темпоральний аспект визначає час, коли можна

виконувати вказані дії. Тобто при врахуванні темпорального аспекту ми можемо надавати різні пояснення щодо отриманого в інтелектуальній системі рішення на різних інтервалах часу.

Залежності між розглянутими властивостями контексту пояснень наведено на рис. 1.

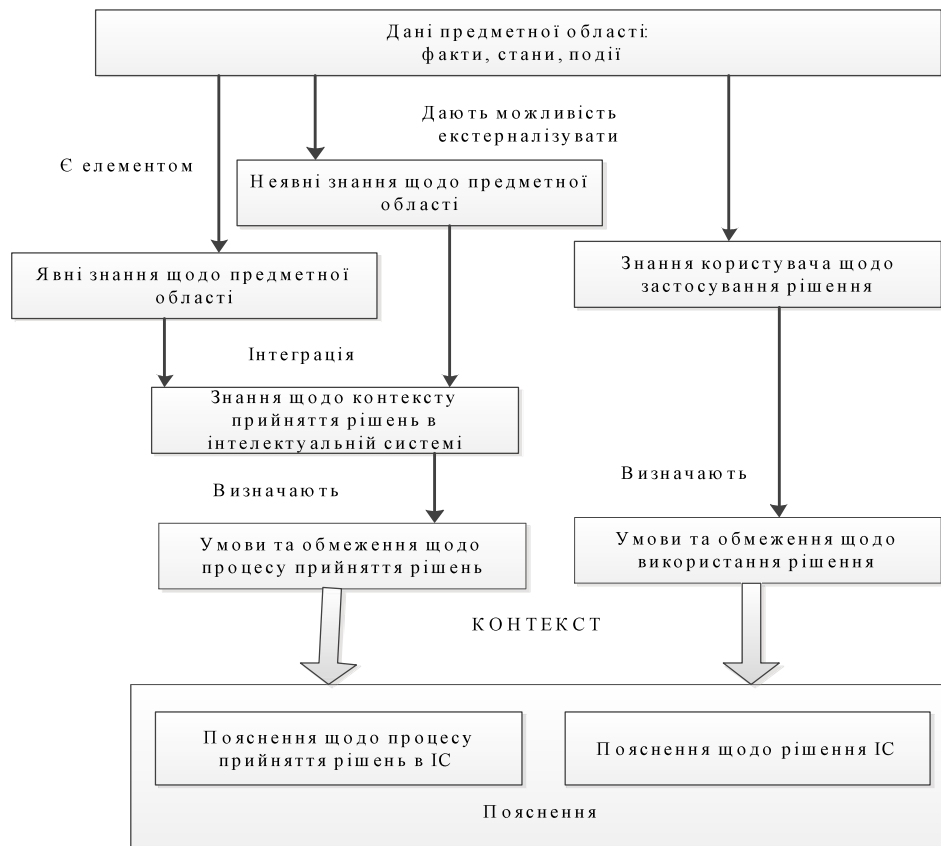


Рис. 1. Залежності між властивостями контексту пояснень

Із представленої на рис. 1 схеми можна зробити висновок про те, що суттєва для пояснення контекстна сукупність знань про процес прийняття рішення, а також про його використання може бути зведена до опису умов та обмежень щодо цього процесу у об'єктному та темпоральному аспектах.

5. Модель контексту пояснень в інтелектуальній системі

Представлена структура контексту пояснень відображає процедури оперування з об'єктами предметної області. Згідно з визначенням фірми IBM, в подальшому такі об'єкти будемо іменувати артефактами. Кожен артефакт має набір властивостей і множину допустимих дій. Крім того, для артефактів визначаються процеси обробки, які задають допустимі послідовності дій. Тобто можливий набір властивостей артефактів та дій з ними визначають контекстні умови та обмеження щодо процесу прийняття рішення у об'єктному аспекті, а допустимі послідовності дій відображають обмеження у темпоральному аспекті.

Оскільки при описі об'єктного та темпорального аспектів ми задаємо послідовність обробки артефактів у часі, такий опис контексту має процесну природу. Тобто послідовності обробки артефактів можна розглядати як підпроцеси єдиного процесу прийняття рішення в інтелектуальній системі. Відповідно, знання про контекст прийняття рішення мають процесну природу і визначають умови та обмеження у процесі отримання результату в системі штучного інтелекту. При використанні процесного опису контексту для пояснення достатньо виділити фрагменти підпроцесів у об'єктному та темпоральному аспектах і далі використати отримані знання для уточнення пояснення.

Розроблена артефактна модель контексту пояснень містить у собі множину артефактів $A = \{a_n\}$ та множину процесів обробки цих артефактів $P = \{p_l\}$:

$$M = \langle A, P : \forall p_l \in P \exists a_n \in A \rangle. \quad (1)$$

Кожен артефакт a_n задається множиною властивостей $V = \{V_n\}$ та множиною взаємопов'язаних дій з обробки артефакту:

$$D = \left\{ D_n : (\forall d_{n,s}, d_{n,k} \in D_i) \exists f_s^k \right\}, \quad (2)$$

де f_s^k - знання щодо залежності між діями $d_{n,s}, d_{n,k}$ з артефактом a_n .

Процес обробки цих артефактів складається із множини залежностей f_s^k :

$$P = \langle f_{k+1}^k, f_{k+2}^k, \dots, f_s^k \rangle. \quad (3)$$

Вказані залежності можуть бути як каузальними, так і темпоральними. Темпоральні залежності задають послідовність у часі для станів артефакту в цілому, а каузальні - зв'язки між властивостями станів.

При кожній i -реалізації процесу обробки артефакту a_n отримуємо послідовність його станів $\pi_{n,i}$. Сукупність цих послідовностей Π_n містить спільні темпоральні залежності. Якщо ці залежності виконуються для всіх послідовностей, то їх можна розглядати як обмеження. В іншому випадку - як контекстні умови реалізації процесу прийняття рішення в інтелектуальній системі.

На базі темпоральних залежностей шляхом визначення ваг можуть бути отримані каузальні залежності. Вказані залежності об'єднують дії з обробки артефактів у єдиний процес.

Таким чином, представлена артефактна модель дає можливість врахувати як об'єктний, так і темпоральний аспекти контекстних умов прийняття рішення, що і становить контекст пояснення в інтелектуальній системі.

6. Процесно-орієнтований метод побудови моделі контексту пояснень

При реалізації розглянутої моделі дані щодо предметної області відображаються через властивості артефактів, з якими оперує процес прийняття рішення. Явні та неявні контекстні знання щодо процесу прийняття рішення відображаються через опис процесу обробки кожного з цих артефактів. Знання щодо результату представляються через опис процесу використання отриманого рішення як специфічного артефакту. У об'єктному аспекті визначаються можливі та допустимі дії із обробки об'єктів, а у темпоральному - послідовність цих дій у часі.

Приклад виділення фрагментів процесного опису контексту наведено на рис. 2.

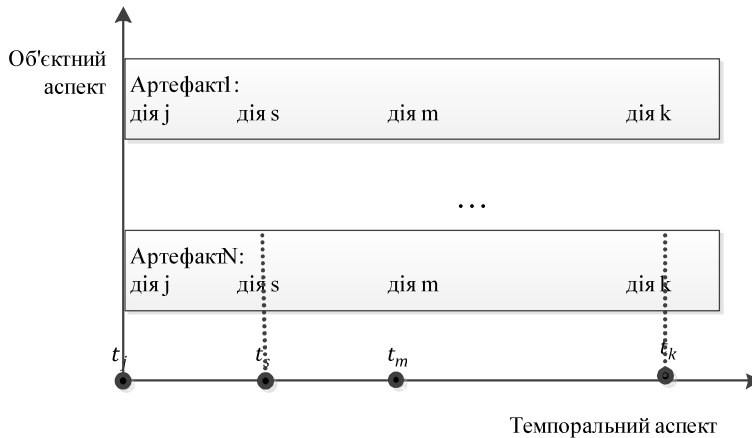


Рис. 2. Відбір обмежень і умов за об'єктним та темпоральним аспектами

Згідно з представленим прикладом, з підмножини артефактів в об'єктному аспекті виділяється a_N - N-артефакт. У темпоральному аспекті виділяються дії з артефактом $d_{n,s}, \dots, d_{n,k}$ на інтервалі $[t_s, t_k]$. Ці дії змінюють відповідні підмножини властивості артефактів $\{v_{n,q}\}$.

Для побудови пояснень суттєвою є підмножина залежностей f_s^k між діями d_s та d_k та проміжними діями.

Представлені характеристики процесного опису пояснень дають можливість визначити основні етапи побудови контекстної моделі пояснень в системі штучного інтелекту.

Процесно-орієнтований метод побудови моделі контексту пояснення містить таку послідовність етапів.

Етап 1. Формування множини артефактів $A = \{a_n\}$, що є суттєвими для побудови пояснення.

Крок 1.1. Визначення множини властивостей артефактів $V = \{V_n\}$.

Крок 1.2. Визначення множини дій D над артефактами.

Етап 2. Відбір відомих i -реалізацій процесів обробки артефактів.

На даному етапі фіксуються відомі послідовності дій із обробки кожного з артефактів.

Це дає можливість врахувати не лише явні знання щодо залежностей f_s^k , але й неявні tacit-знання, оскільки послідовність обробки артефактів відображає неформалізовані зв'язки між діями процесу обробки артефакту.

Етап 3. Побудова процесів обробки артефактів на основі узагальнення реалізацій $\pi_{n,i}$ з використанням темпоральних залежностей між діями з обробки артефактів.

Крок 3.1. Узагальнення реалізацій $f_{n,s}^{n,k}$ залежностей f_s^k у вигляді темпоральних правил r_s^k , що визначають послідовність обробки артефакту у часі, на інтервалі $[t_s, t_k]$.

Крок 3.2. Виділення підмножин темпоральних правил-обмежень та правил-умов виконання дій з обробки артефактів. Обмеження мають виконуватись на всіх реалізаціях $\pi_{n,i}$.

Крок 3.3. Побудова послідовності робіт процесу обробки кожного артефакту у часі на основі інтеграції темпоральних правил-умов при задоволенні правил-обмежень.

Етап 4. Формування каузальних залежностей, що визначають процес обробки кожного з артефактів. На відміну від темпоральних правил r_s^k , каузальні залежності задають причинно-наслідкові зв'язки між підмножинами властивостей артефактів предметної області.

Представлений метод дає можливість формувати суттєві для побудови пояснень контекстні умови і обмеження з використанням темпоральних та каузальних залежностей, представлених у процесах обробки артефактів предметної області. Це створює умови персоналізації пояснень для представленої у вигляді чорного ящика інтелектуальної системи з урахуванням підмножини актуальних вхідних та проміжних даних щодо процесу прийняття рішення.

7. Висновки та перспективи подальших досліджень

Виконано структурування контексту пояснень в інтелектуальній інформаційній системі. За результатами структурування обґрунтовано побудову опису контексту в об'єктному та темпоральному аспектах. Об'єктний аспект охоплює підмножину об'єктів предметної області, які використовуються для формування пояснення. Темпоральний аспект визначає інтервал часу, коли дії з виділеними об'єктами щодо формування рішення є актуальними. Показано, що суттєві для пояснення знання визначають умови та обмеження у об'єктному та темпоральному аспектах щодо процесу прийняття рішень в інтелектуальній системі.

Запропоновано артефактну модель контексту пояснень. Модель містить опис артефактів предметної області та процесів обробки цих артефактів. Кожен процес обробки артефактів складається із послідовності темпоральних та каузальних залежностей між станами та властивостями артефактів відповідно. Модель дає можливість персоналізувати пояснення з урахуванням їх актуальності на основі використання комбінації явних та неявних знань, що складають умови та обмеження щодо процесу прийняття рішення в інтелектуальній системі.

Розроблено процесно-орієнтований метод побудови артефактної моделі контексту пояснень, який містить етапи формування множини артефактів, процесів обробки артефактів та

залежностей, що лежать в основі цих процесів. Метод дає можливість сформувати опис контексту з використанням не лише явних, але й неявних знань шляхом узагальнення та виявлення темпоральних залежностей для упорядкованих у часі даних, що відображають процес обробки артефактів при прийнятті рішення в інтелектуальній системі.

Подальші дослідження щодо контексту пояснень пов'язані із розробкою підходу до формування каузальних контекстних залежностей на основі темпоральних правил, що описують процеси обробки артефактів. Вирішення цієї задачі орієнтовано на екстерналізацію неявних знань щодо процесу прийняття рішень в інтелектуальній інформаційній системі.

Список літератури: 1. *Miller T.* Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*. 2019. Vol. 267. P.1-38. DOI: <https://doi.org/10.1016/j.artint.2018.07.007> 2. *Camburu O.-M., Giunchiglia E., Foerster J., Lukasiewicz T., Blunsom Ph.* Can I trust the explainer? Verifying post-hoc explanatory methods. arXiv:1910.02065. 2019. 3. *Castelvecchi D.* Can we open the black box of AI? *Nature*. 2016. Vol. 538 (7623). P. 20-23. 4. *Tintarev N., Masthoff J.* A survey of explanations in recommender systems. The 3rd international workshop on web personalisation, recommender systems and intelligent user interfaces (WPRSIUI'07). 2007. P. 801-810. 5. *Gunning D., Vorm E., Wang J., Turek M.* darpa's Explainableai(xai) Program: a Retrospective. *Applied AI Letters*. 2021. Vol. 2. No. 4. <https://doi.org/10.1002/ail2.61>. 6. *Pearl J.* Causality: Models, Reasoning and Inference. 2nd ed. Cambridge University Press, USA. 2009. 7. *Maier M., Marazopoulou K., Jensen D.* Reasoning about Independence in Probabilistic Models of Relational Data. arXiv. 2014. 8. *Marazopoulou K., Maier M., Jensen D.* Learning the structure of causal models with relational and temporal dependence. Proceedings of the Thirty-First Conference on Uncertainty in Artificial Intelligence. 2015. 9. *Чала О.В.* Модель узагальненого представлення темпоральних знань для задач підтримки управлінських рішень. Вісник Національного технічного університету "ХПІ". Системний аналіз, управління та інформаційні технології. 2020. № 1(3). С. 14-18. DOI: 10.20998/2079-0023.2020.01.03. 10. *Chala O.* Models of temporal dependencies for a probabilistic knowledge base. *Econtechmod. An International Quarterly Journal*. 2008. Vol. 7. No. 3. P. 53 - 58. 11. *Чалий С.Ф.* Реляційно-темпоральна модель набору сутностей предметної області для процесу формування рішення в інтелектуальній інформаційній системі / С. Ф. Чалий, В. О. Лещинський, І. О. Лещинська. Вісник Національного технічного університету "ХПІ". Сер. : Системний аналіз, управління та інформаційні технології = Bulletin of the National Technical University "KhPI". Ser. : System analysis, control and information technology : зб. наук. пр. Харків: НТУ "ХПІ", 2022. № 1 (7). С. 84-89. 12. *Chalyi S., Leshchynskyi V.* Temporal representation of causality in the construction of explanations in intelligent systems. *Advanced Information Systems*. Kharkiv: NTU "KhPI". 2020. Vol. 4. №3. P. 113-117.

Надійшла до редколегії 10.12.2022 р.

Лещинський Володимир Олександрович, канд. техн. наук, доцент, докторант кафедри ІУС ХНУРЕ. Наукові інтереси: рекомендаційні системи, самопояснювальний штучний інтелект. Адреса: 61166, Харків, пр. Науки 14, ХНУРЕ, каф. ІУС, контактний телефон: +38(057)7021451.