

**Говорущенко Т.О.**

Хмельницький національний університет

**Павлова О.О.**

Хмельницький національний університет

**Боднар М.А.**

Хмельницький національний університет

## СУЧАСНІ ПРОБЛЕМИ СЕМАНТИЧНОГО АНАЛІЗУ СПЕЦИФІКАЦІЙ ВИМОГ ДО ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ

*У статті проводиться дослідження сучасних проблем семантичного аналізу специфікацій вимог до програмного забезпечення (ПЗ). Проведене дослідження довело необхідність розроблення інтелектуального агента на основі онтологічного підходу для автоматизації семантичного розбору (парсингу) природомовної специфікації.*

**Ключові слова:** програмне забезпечення (ПЗ), специфікація вимог до програмного забезпечення, семантичний аналіз, розбір (парсинг) специфікацій, інтелектуальний агент.

**Постановка проблеми.** Практично всі сфери людської діяльності сьогодні пов'язані з комп'ютерними системами, основою яких є програмне забезпечення (далі – ПЗ). Натепер у світі витрачається більше ніж 250 млрд. USD щорічно на розроблення приблизно 175 тис. програмних проєктів [1]. При цьому значна кількість (до 70%) програмних проєктів є неуспішними [1].

Значна кількість помилок уноситься в ПЗ на етапі формування та формулювання вимог унаслідок інформаційних утрат через неповне та різне розуміння потреб і контексту інформації. Такі помилки становлять 10–23% всіх помилок, причому чим більший обсяг ПЗ, тим більше помилок уноситься саме на цьому етапі [2, с. 637]. У роботах [3; 4] підтверджується факт, що причини майже всіх інцидентів і катастроф, пов'язаних із ПЗ, містяться саме у специфікації вимог до ПЗ. У праці [4] описано результати експерименту, проведеного для підтвердження гіпотези, що збої та помилки ПЗ, написаного різними розробниками за однією специфікацією, статистично незалежні. У результаті такого експерименту встановлено, що версії ПЗ, написані різними розробниками за однаковими вимогами, містили низку спільних помилок, пов'язаних із помилками або неточностями вимог (специфікації). Але чим раніше буде виявлено дефект, тим дешевше обійдеться його виправлення. Витрати на виправлення некоректних вимог, виявлених після випуску продукту, майже в 100 разів перевищують витрати на виправлення недоліків специфікації, виявлених на більш ранніх етапах життєвого циклу ПЗ [2, с. 27].

Отже, успішність реалізації програмного проєкту (як вчасне виконання програмного проєкту в рамках виділеного бюджету та з реалізацією всіх необхідних можливостей і функцій [1]) суттєво залежить від специфікації вимог до ПЗ, оскільки саме вона містить усі вимоги й особливості розроблюваного ПЗ. Таке експериментальне свідчення безпосередньо призводить до необхідності поглиблення аналізу специфікації вимог. Тоді *нині актуальною проблемою є аналіз специфікацій вимог до ПЗ (особливо її нефункційних вимог) з метою підвищення успішності та якості розроблюваного за специфікацією ПЗ.*

**Аналіз останніх досліджень і публікацій.** Аналіз відомих методів, призначених для роботи з вимогами та специфікаціями (Using natural language processing technique, Using CASE analysis method, QAW-method, Using global analysis method, O'Brien's approach, Method to discover missing requirement elicitation, Selection of elicitation technique, Comparison and categorization of requirements elicitation techniques, Techniques for ranking and prioritization of software requirements) [5–7], показав, що всі вони спрямовані на контроль за реалізацією вимог, тобто на верифікацію ПЗ до вимог специфікації, але жоден із них не виконує валідації вимог до потреб замовника. Тому використання таких методів для аналізу специфікацій може призвести до розроблення коректного ПЗ, яке, проте, не буде якісним унаслідок невідповідності всім потребам замовників, ураховуючи сучасне визначення якості (як здатність ПЗ

задовольнити заявлені й передбачувані потреби замовників під час його використання за певних умов [8]).

Аналіз наявних автоматизованих засобів, призначених для роботи з вимогами та специфікаціями (IBM Rational RequisitePro, IBM Rational/Telelogic DOORS, Borland Caliber RM, Sybase PowerDesigner, Open Source Requirements Management Tool, пакет засобів компанії LDRA, Sigma Software, DEVPRO, CASE.Аналитик), проведений у роботі [9], показав, що деякі з них спрямовані на контроль за реалізацією вимог у коді (тобто на верифікацію), деякі – на перевірку специфікації на достовірність, повноту й несуперечливість її інформації, але жоден із них також не перевіряє відповідність вимог специфікації потребам замовника.

У роботах [10; 11] розроблено інтелектуальний агент для аналізу нефункційних характеристик у специфікаціях вимог до ПЗ, призначений для аналізу специфікацій вимог до ПЗ на предмет достатності їх інформації щодо нефункційних характеристик-складників якості ПЗ. Але недоліком такого рішення є те, що сьогодні вибір інформації щодо нефункційних характеристик зі специфікації вимог до реального ПЗ реалізується вручну, хоча необхідним є автоматизований семантичний розбір природомовної специфікації на предмет пошуку атрибутів, необхідних для визначення нефункційних характеристик-складників якості ПЗ.

**Постановка завдання.** Проведений аналіз відомих методів і засобів, призначених для роботи з вимогами та специфікаціями, виявив, що натеper невирішеною є проблема аналізу специфікацій вимог до ПЗ на предмет відповідності вимог потребам замовника з метою підвищення якості ПЗ, розроблюваного за специфікаціями. Для вирішення такої проблеми варто розв'язати завдання автоматизації семантичного розбору (парсингу) природомовної специфікації для встановлення відповідності вимог специфікації потребам замовника. Тому метою дослідження є аналіз відомих методів і рішень семантичного аналізу специфікацій вимог до ПЗ.

**Виклад основного матеріалу дослідження.** *Аналіз відомих методів семантичного аналізу специфікацій вимог до ПЗ.* Сьогодні відомі різні підходи до семантичного аналізу специфікацій. Так, у праці [12] запропоновано методику, що забезпечує автоматичну допомогу розробникам шляхом трансформації природомовних вимог за допомогою UML-діаграм. Автори [12] зосереджуються на створенні діаграм діяльності й послідовності

на основі природомовних специфікацій вимог до ПЗ. Стандартний POS-тагер і парсер аналізують вимоги англійською мовою, надані користувачами, і «витягують» із них фрази, дії тощо. Запропонований у роботі [12] метод зменшує розрив між неформальною природною мовою та формальною мовою моделювання.

Стаття [13] присвячена вирішенню проблеми перетворення бізнес-специфікацій, написаних природною мовою, у формальні моделі, придатні для використання під час розроблення інформаційних систем. Автори пропонують метод перетворення природомовних специфікацій на основі стандарту Semantics of Business Vocabulary and Business Rules (SBVR). Цей підхід є унікальним у поєднанні Model-Driven Engineering (MDE), когнітивної лінгвістики та конфігурації на основі знань, яка дає змогу надійного семантичного оброблення специфікацій та інтеграції з наявними інструментами MDE для підвищення продуктивності й часу розроблення ПЗ за специфікацією. Метод, запропонований у праці [13], забезпечує зменшення ручної формалізації специфікацій.

Методологія, запропонована у роботі [14], складається з чотирьох процесів, а саме: розбір (парсинг) вимог (PR), відображення вимог із використанням матриці (RMM), додавання вимог до шаблону специфікації та інспекція третьою стороною. Запропонована методологія є ефективною для мінімізації неоднозначності вимог.

Автори [15] пропонують механізм автоматизації відображення вимог у формальні представлення за допомогою маркування семантичної ролі. Автори описують перший загальнодоступний набір даних для цього завдання, використовують ієрархічну структуру, що дає можливість анотувати концепцію вимог, а також обговорюють, як маркування семантичної ролі може бути адаптоване для аналізу програмних вимог. Механізм, викладений у праці [15], може бути використаний для відображення природомовних функційних вимог спочатку в специфікаціях, а потім у програмному коді, але абсолютно не підходить для нефункційних вимог.

У роботі [16] запропоновано підхід до автоматичного видобування семантичної інформації зі специфікацій вимог до ПЗ шляхом поєднання методів маркування семантичної ролі та моделювання знань предметної галузі. Автор [16] вибрав численні дієслова з доступних специфікацій у галузі електронної комерції й побудував семантичні рамки для цих дієслів. Потім обрані речення були позначені вручну, і результат використаний як навчальна вибірка для машинного навчання.

Під час процесу маркування автором приймався послідовний спосіб, під час якого попередні мічені результати використовувались для побудови динамічних характеристик для ідентифікації наступних семантичних ролей. Запропонований підхід є ефективним і надійним для моделювання мінливості та спільності функційних вимог, але не підходить для роботи з нефункційними вимогами.

Автори [17], використовуючи методи машинного навчання, розробили технологічний процес, який приймає як входи природомовні артефакти (наприклад, специфікації вимог до ПЗ). Цей процес автоматично визначає відповідні речення безпеки в артефактах і класифікує їх відповідно до цілей безпеки або явно викладених або передбачуваних речень. Авторами класифіковано 10963 речень із шести різних специфікацій і видобуто з них відповідні цілі безпеки. Аналіз показав, що 46% речень мають стосунок до безпеки, з них 28% чітко вказують на безпеку, а 72% речень – на функціональні вимоги з наслідками безпеки. Грунтуючись на такому аналізі, автори розробили контекстно-залежні шаблони, на основі яких можна створювати набір функціональних вимог безпеки, заповнюючи ключову інформацію з релевантних пропозицій безпеки.

У праці [18] розроблено метод для налаштування і створення комбінованого парсера (аналізатора) для оброблення та аналізу природомовних специфікацій, який об'єднує переваги формальних парсерів, що використовуються для оброблення описів із жорстко визначеним синтаксисом (наприклад, вихідний (сирцевий) код програми) та є дуже точними й ефективними під час оброблення формальної частини, а також аналітиків, розроблених для оброблення природних мов, які добре розуміють вільний текст. Отриманий текстовий аналізатор базується на Parsing Expression Grammars. Запропонований у праці [18] підхід дає змогу побудувати ефективний синтаксичний аналізатор для аналізу природомовних специфікацій вимог до промислового ПЗ.

Автори [19] розробили методологію та інструмент QuARS (Аналізатор якості специфікацій вимог) для систематичного й автоматичного аналізу природомовних вимог. QuARS дає змогу виконувати початковий аналіз вимог з метою автоматичного виявлення потенційних лінгвістичних дефектів, які можуть викликати проблеми інтерпретації на наступних етапах розроблення ПЗ. Цей інструмент також може частково підтримувати аналіз узгодженості та повноти шляхом кластеризації вимог відповідно до конкретних тем.

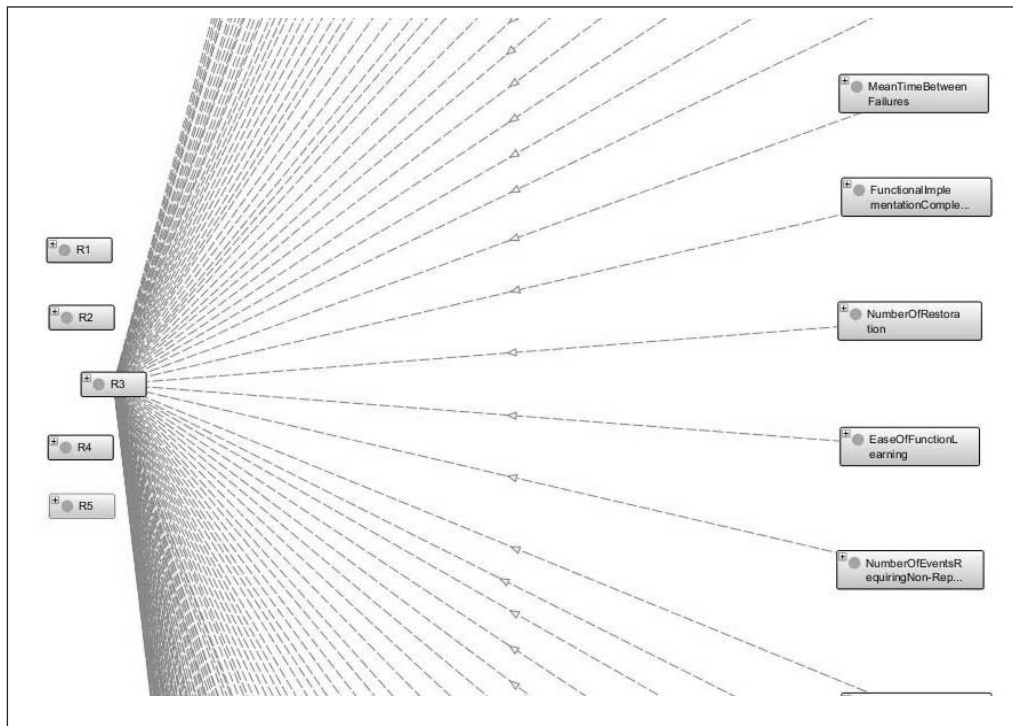
У праці [20] представлено онтологічний підхід до автоматизованої перевірки та вимірювання вимог до ПЗ, який використовується для виявлення невідповідностей, непослідовності й недоліків вимог до ПЗ. Для цього розроблено метамодель вимог – онтологію вимог, яка надає основу для валідації та інструмента підтримки, а також забезпечує концептуалізацію знань щодо вимог. Отже, онтологія вимог дає змогу документувати структуровані, багаторазові, однозначні, простежувані, повні й узгоджені вимоги відповідно до стандарту IEEE для специфікації вимог до ПЗ. Розроблена онтологія вимог може використовуватись інженером вимог як інструмент для висвітлення невідповідностей, неповноти та якісних недоліків вимог.

Автори [21] представляють прототип семантичної системи, що використовується для надання допомоги інженерам із вимог щодо видобування вимог з використанням напівформального представлення. Семантична система використовує поняття, відносини й аксіоми онтології предметної галузі, щоб надати список пропозицій, на основі яких інженер може розробити вимоги. Розроблена семантична система ефективно підтримує інженерів вимог у визначенні добре структурованих вимог.

У статті [22] запропоновано підхід до підтримки процесу інженерії вимог у процесі розробки ПЗ шляхом застосування онтології. У цій роботі подана онтологія з урахуванням особливостей методології Scrum. Під час аналізу особливостей інженерії вимог у рамках методології Scrum встановлено, що основні зусилля зосереджені на аналізі функційного складника, тому нефункційні вимоги часто не задокументовані. Беручи до уваги високу частоту оновлення вимог під час роботи відповідно до методології Scrum, можемо сказати, що застосування запропонованого підходу дає змогу швидко контролювати простежуваність і повноту вимог.

У дослідженні [23] запропоновано структуру для інтеграції різнорідних вимог за допомогою онтологій. Цей підхід ґрунтується на загальній моделі вимог користувачів, яка використовувала три наявні напівформальні мови UML, MCT і MERISE і цілеорієнтовані мови. Розроблений у праці [23] метод фокусується на мінімізації впливу неоднорідності вимог на якість ПЗ.

Отже, як показав проведений аналіз, усі розглянуті методи семантичного аналізу специфікацій вимог до ПЗ не забезпечують перевірки відповідності вимог специфікації потребам замовника (валідації), тому не спрямовані на підвищення якості розроблюваного за специфікацією ПЗ.



**Рис. 1. Фрагмент онтології предметної галузі «Інженерія програмного забезпечення» (частина «Специфікація вимог до ПЗ (атрибути якості)») – шаблон специфікації вимог до ПЗ з погляду наявності атрибутів для нефункційних характеристик-складників якості ПЗ**

*Концепція запропонованого рішення.* Для автоматизації семантичного розбору природомовної специфікації з метою перевірки відповідності її вимог потребам замовника необхідно виконати її формалізацію, наприклад, із використанням онтологій, оскільки саме онтології забезпечують системний підхід до вивчення предметної галузі, можливість цілісного подання інформації предметної галузі, виявлення дублювань і прогалів у знаннях на основі візуалізації відсутніх логічних зв'язків, можливість доступу, розуміння й аналізу інформації інтелектуальними агентами. Формалізацію специфікації вимог до ПЗ потрібно виконувати, враховуючи, на видобування яких саме вимог буде спрямовуватись семантичний аналіз (парсинг) специфікації. Якщо автоматизований семантичний розбір природомовної специфікації буде проводитись на предмет пошуку атрибутів, необхідних для визначення нефункційних характеристик-складників якості ПЗ (для розвитку рішень, запропонованих у роботах [10, 11]), то для формалізації специфікації вимог можуть бути використані структурна модель специфікації вимог до ПЗ (з погляду наявності атрибутів) і базова (універсальна) онтологічна модель предметної галузі «Інженерія програмного забезпечення» (частина «Специфікація вимог до ПЗ

(атрибути якості)»), розроблені в праці [24] на основі стандарту ISO 29148 [25].

На основі базової (універсальної) онтологічної моделі предметної галузі «Інженерія програмного забезпечення» (частина «Специфікація вимог до ПЗ (атрибути якості)») розроблено базову онтологію предметної галузі «Інженерія програмного забезпечення» (частина «Специфікація вимог до ПЗ (атрибути якості)») – рис. 1. У цій онтології атрибути, необхідні для визначення нефункційних характеристик-складників якості ПЗ, представлені з урахуванням розподілу за розділами специфікації, за рахунок чого розроблена онтологія є шаблоном специфікації вимог до ПЗ з погляду наявності атрибутів (котрий може опрацьовуватись як фахівцями, так й автоматизованими засобами для роботи зі специфікаціями або інтелектуальними агентами) та надає візуальні підказки користувачу про місце розташування тих чи інших атрибутів у специфікації вимог до ПЗ.

**Висновки.** Оскільки успішність реалізації програмного проекту суттєво залежить від специфікації вимог до ПЗ, то актуальною є проблема аналізу специфікацій вимог до ПЗ (особливо її нефункційних вимог) з метою підвищення успішності та якості розроблюваного за специфікацією ПЗ. Проведений аналіз відомих методів і засо-

бів, призначених для роботи з вимогами та специфікаціями, показав, що сьогодні невирішеним є завдання автоматизації семантичного розбору (парсингу) природомовної специфікації для встановлення відповідності вимог специфікації потребам замовника.

Проведено аналіз методів семантичного аналізу специфікацій вимог до ПЗ, який показав, що всі вони не забезпечують перевірки відповідності вимог специфікації потребам замовника (валідації), тому не спрямовані на підвищення якості розроблюваного за специфікацією ПЗ.

Тоді для автоматизації семантичного розбору природомовної специфікації з метою перевірки відповідності її вимог потребам замовника

необхідно виконати її формалізацію, наприклад, із використанням онтологій. Для такої формалізації запропоновано використати раніше розроблені структурну й онтологічну моделі специфікації вимог до ПЗ, на основі яких розроблено онтологію предметної галузі «Інженерія програмного забезпечення» (частина «Специфікація вимог до ПЗ (атрибути якості)»), що буде основою інтелектуального агента для автоматизованого семантичного розбору природомовної специфікації на предмет пошуку атрибутів, необхідних для визначення нефункційних характеристик-складників якості ПЗ, на розроблення якого будуть спрямовані подальші дослідження авторів.

#### Список літератури:

1. Hastie Shane, Wojewoda Stéphane. Standish Group 2015 Chaos Report – Q&A with Jennifer Lynch. URL: <http://www.infoq.com/articles/standish-chaos-2015> (Last accessed: January 9, 2019).
2. Макконнелл С. Совершенный код. Мастер-класс. Москва: Русская редакция, 2013. 896 с.
3. Hazard analysis of complex spacecraft using systems-theoretic process analysis / T. Ishimatsu and others. *Journal of Spacecraft and Rockets*. 2014. Vol. 51. No. 2. P. 509–522.
4. Levenson N. Software challenges in achieving space safety. *Journal of the British Interplanetary Society*. 2009. Vol. 62. P. 265–272.
5. Fatwanto A. Software requirements specification analysis using natural language processing technique. *The 13-th IEEE International Conference on Quality in Research: Proceedings* (Yogyakarta, June 25–28, 2013). Yogyakarta (Indonesia), 2013. P. 105–110.
6. Rehman T., Khan M.N.A., Riaz N. Analysis of requirement engineering processes, tools/techniques and methodologies. *International Journal of Information Technology and Computer Science*. 2013. Vol. 5. No. 3. P. 40–48.
7. A comparison of requirements specification methods from a software architecture perspective / L. Bass, J. Bergey, P. Clements, P. Merson, I. Ozkaya, R. Sangwan. URL: [http://resources.sei.cmu.edu/asset\\_files/TechnicalReport/2006\\_005\\_001\\_14786.pdf](http://resources.sei.cmu.edu/asset_files/TechnicalReport/2006_005_001_14786.pdf) (Last accessed: January 9, 2019).
8. ISO/IEC 25010:2011. Systems and software engineering. Systems and software Quality Requirements and Evaluation (SQuaRE). System and software quality models. [Introduced 01.03.2011]. Geneva (Switzerland), 2011. 34 p. (International standard).
9. Говорущенко Т.О., Павлова О.О. Сучасні проблеми оцінювання початкових етапів життєвого циклу програмного забезпечення. *Електротехнічні та комп'ютерні системи*. 2018. № 27 (103). С. 165–175.
10. Hovorushchenko T., Pavlova O. Method of Activity of Ontology-Based Intelligent Agent for Evaluating the Initial Stages of the Software Lifecycle. *Advances in Intelligent Systems and Computing*. 2019. Vol. 836. P. 169–178.
11. Павлова О.О., Говорущенко Т.О., Іванов О.В. Діяльність інтелектуального агента для оцінювання інформації у специфікаціях вимог до програмного забезпечення. *Штучний інтелект*. 2018. № 2. С. 66–75.
12. Gulia S., Choudhury T. An Efficient Automated Design to Generate UML Diagram From Natural Language Specifications. *The 6-th International Conference on Cloud System and Big Data Engineering: Proceedings* (Noida, January 14–15, 2016). Noida (India), 2016. P. 641–648.
13. Formalising natural language specifications using a cognitive linguistic/configuration based approach / M. Selway, G. Grossman, W. Mayer, M. Stumptner. *Information Systems*. 2015. Vol. 54. P. 191–208.
14. Ali S.W., Ahmed Q.A., Shafi I. Process to Enhance the Quality of Software Requirement Specification Document. *The International Conference on Engineering and Emerging Technologies: Proceedings* (Lahore, February 22–23, 2018). Lahore (Pakistan), 2018. P. 113–118.
15. Software requirements as an application domain for natural language processing / T. Diamantopoulos, M. Roth, A. Symeonidis, E. Klein. *Language Resources and Evaluation*. 2017. Vol. 51. Issue 2. P. 495–524.
16. Wang Y. Semantic Information Extraction for Software Requirements using Semantic Role Labeling. *The IEEE International Conference on Progress in Informatics and Computing: Proceedings* (Nanjing, December 18–20, 2015). Nanjing (China), 2015. P. 332–337.

17. Hidden in Plain Sight: Automatically Identifying Security Requirements from Natural Language Artifacts / M. Riaz, J. King, J. Slankas, L. Williams. *The 2-nd IEEE International Requirements Engineering Conference: Proceedings* (Karlskrona, August 25–29, 2014). Karlskrona (Sweden), 2014. P. 183–192.

18. Iwama F., Nakamura T., Takeuchi H. Constructing Parser for Industrial Software Specifications Containing Formal and Natural Language Description. *The 34-th International Conference on Software Engineering: Proceedings* (Zurich, June 02–09, 2012). Zurich (Switzerland), 2012. P. 1012–1021.

19. An automatic tool for the analysis of natural language requirements / S. Gnesi, G. Lami, G. Trentanni, F. Fabbri, M. Fusani. *Computer Systems Science and Engineering*. 2005. Vol. 20. Issue 1. Pp. 53–62.

20. Siegemund K. Contributions to Ontology-Driven Requirements Engineering: dissertation. Dresden, 2014. 249 с.

21. Ontology-Driven Guidance for Requirements Elicitation / S. Farfeleder, T. Moser, A. Krall, T. Stalhane, I. Omoroniya, H. Zojer. *Lecture Notes in Computer Science*. 2011. Vol. 6644. P. 212–226.

22. Murtazina M.Sh., Avdeenko T.V. The ontology-driven approach to support the requirements engineering process in Scrum framework. *CEUR-WS*. 2018. Vol. 2212. P. 287–295.

23. Integration of Heterogeneous Requirements using Ontologies / A. Mustafa, W.M.N. Wan-Kadir, N. Ibrahim, A. Shah, M. Younas. *International Journal of Advanced Computer Science and Applications*. 2018. Vol. 9. No. 5. P. 213–218.

24. Говорущенко Т.О. Теоретичні та прикладні засади інформаційної технології оцінювання достатності інформації щодо якості у специфікаціях вимог до програмного забезпечення: дис. ... докт. техн. наук: спец. 05.13.06. Хмельницький, 2018. 441 с.

25. ISO/IEC/IEEE 29148:2011. Systems and software engineering. Life cycle processes. Requirements engineering. [Introduced 01.12.2011]. Geneva (Switzerland), 2011. 28 p. (International standard).

#### **СОВРЕМЕННЫЕ ПРОБЛЕМЫ СЕМАНТИЧЕСКОГО АНАЛИЗА СПЕЦИФИКАЦИЙ ТРЕБОВАНИЙ К ПРОГРАММНОМУ ОБЕСПЕЧЕНИЮ**

*В статье проведено исследование современных проблем семантического анализа спецификаций требований к программному обеспечению (ПО). Проведенное исследование показало необходимость разработки интеллектуального агента на основе онтологического подхода для автоматизации семантического разбора (парсинга) спецификации, написанной естественным языком.*

**Ключевые слова:** программное обеспечение (ПО), спецификация требований к программному обеспечению, семантический анализ, разбор (парсинг) спецификаций, интеллектуальный агент.

#### **CURRENT PROBLEMS OF SEMANTIC ANALYSIS OF THE SOFTWARE REQUIREMENTS SPECIFICATIONS**

*The article deals with the current problems of semantic analysis of the software requirements specifications. The conducted research has proved the necessity of developing the ontology-based intelligent agent for automating the semantic parsing of the natural-language specifications.*

**Key words:** software, software requirements specification (SRS), semantic analysis, parsing of the SRS, intelligent agent.