



УДК 81'25

[https://doi.org/10.52058/2786-6165-2025-5\(35\)-258-265](https://doi.org/10.52058/2786-6165-2025-5(35)-258-265)

Вискушенко Світлана Андріївна кандидат філологічних наук, доцент, доцент кафедри англійської філології та перекладу, Житомирський державний університет імені Івана Франка, м. Житомир, <https://orcid.org/0000-0003-1557-1984>

СТАНДАРТИЗАЦІЯ ТЕРМІНІВ І МАШИННИЙ ПЕРЕКЛАД: ВИКЛИКИ ТА ПЕРСПЕКТИВИ

Анотація. У сучасному світі, де інформаційні потоки стають дедалі інтенсивнішими, а потреба в багатомовній комунікації зростає, машинний переклад (МП) відіграє ключову роль у забезпеченні доступу до знань та інформації. Однак, попри значні досягнення в галузі штучного інтелекту та обробки природної мови, стандартизація термінології залишається одним із головних викликів для забезпечення якості та точності перекладу.

Статтю присвячено вивченню проблем стандартизації термінів і машинного перекладу в аспекті розгляду викликів й перспектив. Підкреслюється, що однією з основних проблем є відсутність однозначних відповідників для багатьох термінів на рівні різних мов. Це особливо актуально для галузей де точність термінології є критично важливою, наприклад, медицина, право тощо. Системи МП часто не здатні адекватно передати специфіку таких термінів без наявності відповідних термінологічних баз даних. Крім того, терміни можуть не мати прямих еквівалентів в інших мовах, що ускладнює їх точний переклад. Окремо наголошується на тому, що термінологія постійно еволюціонує, особливо в галузях, що швидко розвиваються. Нові терміни з'являються швидше, ніж їх встигають включити до словників або термінологічних баз даних. Це створює додаткові труднощі для систем МП, які не мають доступу до актуальних даних. Наступною суттєвою проблемою вважається полісемантичність термінологічних одиниць. Одна лексична одиниця може мати кілька значень залежно від контексту, а системи МП повинні правильно інтерпретувати ці значення, що вимагає не лише лінгвістичних знань, але й розуміння специфіки відповідної предметної галузі. Для подолання цих викликів необхідно впроваджувати стандарти та інструменти, що сприяють уніфікації та стандартизації термінології. Більше того, розвиток технологій, таких як великі мовні моделі (LLM), відкриває нові можливості для інтеграції термінології в МП. Використання LLMs для генерації двомовних



термінологічних даних та постредагування перекладів може значно покращити точність перекладу спеціалізованих термінів.

У підсумку наголошується, що для забезпечення високоякісного машинного перекладу необхідно поєднувати технологічні інновації з лінгвістичними дослідженнями та впровадженням міжнародних стандартів. Це дозволить забезпечити точність, послідовність та ефективність перекладу в умовах зростаючої глобалізації та цифровізації.

Ключові слова: термін, стандартизація, галузевий переклад, фаховий текст, фахова лексика, полісемія, міжмовна комунікація, машинний переклад.

Vyskushenko Svitlana Andriivna Candidate of Sciences in Philology, Associate Professor, Department of English Philology and Translation, Zhytomyr Ivan Franko State University, Zhytomyr, <https://orcid.org/0000-0003-1557-1984>

STANDARDIZATION OF TERMINOLOGY AND MACHINE TRANSLATION: CHALLENGES AND PROSPECTS

Abstract. In the contemporary world, where information flows are becoming increasingly intense and the need for multilingual communication is growing, machine translation (MT) plays a key role in providing access to knowledge and information. However, despite significant advances in artificial intelligence and natural language processing, terminology standardization remains one of the main challenges in ensuring translation quality and accuracy.

The paper offers an overview of the problems of terminology standardization and machine translation in the context of challenges and prospects. It emphasizes that one of the main problems is the lack of unambiguous equivalents for many terms across different languages. This is especially relevant for fields where terminological precision is critically important, such as medicine, law, and technology. MT systems often struggle to adequately convey the specificity of such terms without access to appropriate terminological databases. Moreover, terms may not have direct equivalents in other languages, complicating their accurate translation. Research indicates that specialized domains often have unique terms and concepts that may not have direct equivalents in other languages, requiring domain-specific dictionaries and context-based models to ensure accurate translations.

Additionally, terminology is constantly evolving, particularly in rapidly developing fields. New terms emerge faster than they can be included in dictionaries or terminological databases. This creates additional difficulties for MT systems that lack access to up-to-date data. Translating cutting-edge



terminology presents challenges as these terms may not be well-established or widely recognized, necessitating continuous updates and integration of domain-specific knowledge.

Another significant problem outlined is the polysemy of terminological units. A single lexical unit can have multiple meanings depending on the context, and MT systems must correctly interpret these meanings, which requires not only linguistic knowledge but also an understanding of the specific subject area. Accurate translation of such terms requires disambiguation based on the surrounding text or domain-specific knowledge. To overcome these challenges, it is necessary to implement standards and tools that promote the unification and standardization of terminology. For example, ISO 704 establishes the basic principles and methods for preparing and compiling terminologies, describing the links between objects, concepts, and their terminological representations. Furthermore, ISO 17100 specifies requirements for all aspects of the translation process directly affecting the quality and delivery of translation services, including the management of core processes and resources.

Moreover, the development of technologies such as large language models (LLM) opens new opportunities for integrating terminology into MT. Utilizing LLMs to generate bilingual terminological data and post-edit translations can significantly improve the accuracy of translating specialized terms. Studies have demonstrated that employing LLMs to generate synthetic bilingual terminological data and post-edit translations can significantly enhance the accuracy of translating specialized terms.

In conclusion, ensuring high-quality machine translation requires combining technological innovations with linguistic research and the implementation of international standards. This approach will ensure the accuracy, consistency, and efficiency of translation in the context of increasing globalization and digitalization.

Keywords: term, standardization, specialized translation, specialized text, specialized vocabulary, polysemy, interlingual communication, machine translation.

Постановка проблеми. У сучасному глобалізованому світі, де інформаційні потоки перетинають мовні та культурні кордони, машинний переклад (МП) відіграє ключову роль у забезпеченні ефективної міжмовної комунікації. Зокрема, у галузях з високим рівнем термінологічної насиченості точність і послідовність перекладу термінів є критично важливими. Однак, незважаючи на значні досягнення в розвитку МП, навіть у рамках нейронного машинного перекладу (НМП), забезпечення термінологічної узгодженості залишається складним завданням. Стандартизація ж термінології є фундаментальним аспектом, що гарантує єдність і зрозумілість



фахової лексики. Міжнародні стандарти, такі як ISO 704:2022 «Terminology work – Principles and methods» (Термінологічна робота – Принципи та методи) та ISO 860:2007 «Terminology work – Harmonization of concepts and terms» (Термінологічна робота – Гармонізація понять і термінів), встановлюють методологічні основи для створення та гармонізації термінологічних систем фахових мов. Вищезазначені стандарти спрямовані на забезпечення однозначності термінів та їх відповідності між різними мовами та галузями знань. Однак інтеграція стандартизованої термінології в будь-які системи МП стикається з низкою викликів. Перш за все, згадані системи, навчені на великих корпусах текстів, часто не враховують специфіку термінології певної галузі знання чи професійної діяльності. Це призводить до неточного або взагалі неправильного перекладу термінів чи їх неконсистентного використання. З метою подолання цих викликів розробляються нові підходи до інтеграції термінології в НМП, наприклад, використання формату TermBase eXchange (TBX) як міжнародного стандарту для обміну термінологічними даними дозволяє забезпечити сумісність між різними програмними засобами та полегшити впровадження термінів у процес перекладу. Крім того, новітні дослідження пропонують методи навчання НМП-систем із врахуванням термінологічних обмежень, що дозволяє підвищити точність перекладу без значного збільшення обчислювальних витрат. У цьому контексті важливо також враховувати роль великих мовних моделей (LLM), які демонструють високі результати в генерації тексту та перекладі. Але відповідні моделі можуть генерувати текст без глибокого розуміння його змісту, що підкреслює необхідність ретельного контролю за термінологічною точністю та узгодженістю. Таким чином, забезпечення термінологічної узгодженості в МП є багатогранним завданням, що вимагає поєднання стандартів, технологій та лінгвістичної експертизи. У запропонованій статті буде розглянуто сучасні підходи до стандартизації термінології в контексті МП, проаналізовано основні виклики та перспективи розвитку цієї царини.

Аналіз останніх досліджень і публікацій. Різноманітними аспектами вивчення термінів, стандартизації термінології, перекладу, машинного перекладу, міжмовної комунікації цікавилися такі українські та зарубіжні науковці, як Гінзбург М. Д., Кияк Т. Р., Скороходько Е. Ф., Вискушенко С. А., Naveen P., Trojovsky P., Sager J. C. та ін.

Метою розвідки є дослідити взаємозв'язок між машинним перекладом і стандартизацією термінології, зосереджуючись на викликах, що виникають, також окреслити перспективи розвитку ефективних підходів до забезпечення термінологічної узгодженості в автоматизованому перекладі галузевих текстів.



Виклад основного матеріалу. Стандартизація термінології є фундаментальним елементом забезпечення ефективної міжмовної та міжфахової комунікації, особливо в умовах стрімкого розвитку технологій машинного перекладу. Вона спрямована на уніфікацію термінів, що дозволяє уникнути неоднозначностей та забезпечує точність передачі професійної інформації. Згідно з міжнародним стандартом ISO 704:2022 «Terminology work – Principles and methods» (Термінологічна робота – Принципи та методи), стандартизація термінології включає в себе встановлення основних принципів і методів підготовки та складання термінологій як у межах, так і поза межами стандартизації. Цей документ описує зв'язки між об'єктами, поняттями, визначеннями та позначеннями, а також встановлює загальні принципи формування термінів і власних назв та написання визначень. Отже, основні принципи, викладені у вищезазначеному стандарті, включають:

- 1) концептуальний підхід: визначення термінів повинно базуватися на чіткому розумінні відповідних понять, що забезпечує однозначність і точність у використанні термінологічних одиниць;
- 2) системність: терміни повинні бути організовані в логічні системи, що відображають взаємозв'язки між поняттями;
- 3) узгодженість: необхідно забезпечити послідовність у використанні термінів у різних контекстах і документах;
- 4) міждисциплінарність: врахування специфіки різних галузей знань при роботі над термінологією для забезпечення можливості застосування її в різних сферах [1].

Стандарт ISO 860:2007 «Terminology work – Harmonization of concepts and terms» (Термінологічна робота – Гармонізація понять і термінів) визначає методологічний підхід до гармонізації понять, систем понять, визначень і термінів. Цей стандарт застосовується до розробки гармонізованих термінологій як на національному, так і на міжнародному рівні, в одно- або багатомовному контексті. Відповідно до зазначеного стандарту гармонізація термінології включає:

- 1) гармонізацію систем понять, тобто встановлення відповідності між двома або більше тісно пов'язаними або системами понять, що перекриваються з метою покращення комунікації;
- 2) гармонізацію визначень, іншими словами, опис гармонізованого поняття за допомогою інтенціонального визначення, що відображає його положення в гармонізованій системі понять;
- 3) гармонізацію термінів, зокрема, вибір позначень для гармонізованого поняття як у різних мовах, так і в межах однієї мови. Всі ці процеси



спрямовані на покращення комунікації та забезпечення точності в передачі спеціалізованої інформації [2].

Машинний переклад (МП) пройшов значний еволюційний шлях від перших експериментів до сучасних систем, що використовують штучний інтелект. Проте, незважаючи на технологічний прогрес, передача галузевої термінології залишається складним завданням. Однією з основних проблем є відсутність прямого перекладу, тобто однозначних відповідників для багатьох термінів у різних мовах, що ускладнює їх точну передачу. Це особливо актуально для технічних, наукових та юридичних термінів, які мають специфічне значення в певному контексті. Системи машинного перекладу часто не здатні точно передати ці нюанси без належної термінологічної бази. Так, наприклад, як показують дослідження, певні вузькофахові терміни можуть не мати прямих еквівалентів в інших мовах, що ускладнює їх точний переклад. Крім того, термінологія постійно еволюціонує, особливо в галузях, що швидко розвиваються, таких як інформаційні технології, біомедицина, біоінженерія тощо. Нові термінологічні одиниці з'являються швидше, ніж їх встигають включити до словників або термінологічних баз даних, що створює додаткові труднощі для технологій автоматизованого перекладу текстів, які не мають доступу до актуальної термінології. Отже, певна кількість термінів можуть бути доступні лише на рівні однієї мови, що ускладнює міжмовну комунікацію та автоматизовану обробку даних. Наступною суттєвою проблемою є багатозначність термінів. Одна лексична одиниця може мати кілька значень залежно від контексту, що вимагає глибокого аналізу для правильного перекладу, тобто не лише лінгвістичних знань, але й розуміння специфіки предметної галузі або сфері знання, а системи МП повинні правильно інтерпретувати ці значення. Серед інших викликів особливо виділяються складні граматичні структури та обмеженість паралельних даних. Щодо граматики, то вона представляє кілька викликів у машинному перекладі, включаючи варіативність граматичних правил у різних мовах, що вимагає складних алгоритмів та лінгвістичного аналізу для точного перекладу. Інакше кажучи, певні граматичні структури на кшталт довгих речень з підрядними та пасивними конструкціями, які притаманні фаховим текстам, можуть мати різні синтаксичні правила в різних мовах, що ускладнює їх автоматичний аналіз та переклад. Якість машинного перекладу також значною мірою залежить від наявності паралельних корпусів текстів, перекладених різними мовами. Проте для багатьох мов, особливо тих, що не мають широко поширеними, таких даних недостатньо. Це обмежує можливості навчання моделей МП та знижує якість перекладу. Іншими словами, паралельні дані використовуються для навчання систем МП, але



вони доступні лише для більшості широко використовуваних мовних пар, що створює виклики для інших мов. Для подолання цих викликів необхідна ефективна стандартизація термінології. Міжнародні стандарти, які вже зазначалися раніше, надають методологічні підходи до гармонізації понять і термінів, що сприяє покращенню ефективності міжмовної комунікації. Зокрема, ISO 860 визначає методологію гармонізації понять, систем понять, визначень і термінів, що є природним продовженням ISO 704. У контексті МП важливу роль відіграє також використання термінологічних баз даних, таких як EuroTermBank, що є найбільшим багатомовним термінологічним ресурсом у Європі. Інтеграція таких ресурсів у системи МП дозволяє забезпечити точність і узгодженість перекладу фахової лексики. Таким чином, стандартизація термінів є критично важливою для забезпечення якості машинного перекладу, особливо в спеціалізованих галузях. Інтеграція міжнародних стандартів, розвиток термінологічних ресурсів та їх ефективне використання в системах МП сприятимуть покращенню точності та узгодженості перекладу, що є необхідним для ефективної міжмовної комунікації в сьогочасному світі. Одним із перспективних напрямків є також інтеграція семантичних веб-технологій у МП, що дозволяє системам краще розуміти контекст та значення термінів, сприяючи більш точному перекладу. Наприклад, використання семантичних мереж може допомогти в розрізненні значень полісемічних термінів. Іншим важливим аспектом є розвиток великих мовних моделей (LLM), які демонструють здатність до контекстуального розуміння та генерації тексту з високою точністю. Ці моделі можуть бути адаптовані до потреб МП, особливо в вузькофахових галузях, де точність термінології є критичною. Також перспективним є використання гібридних підходів, які поєднують автоматичний переклад з постредагуванням людиною. Це дозволяє забезпечити високу якість перекладу, зберігаючи при цьому ефективність та швидкість обробки текстів [1; 2; 3; 4; 5].

Висновки. У сучасному контексті глобалізації та цифровізації, машинний переклад стає невід'ємною складовою міжмовної комунікації. Однак, попри значний прогрес у розвитку технологій МП, стандартизація термінології залишається критично важливою для забезпечення точності та послідовності перекладу, особливо в спеціалізованих галузях. У підсумку, слід зазначити, що для подолання викликів у стандартизації термінології в контексті машинного перекладу необхідно поєднувати технологічні інновації з лінгвістичними дослідженнями та розробкою міжнародних стандартів. Це дозволить забезпечити точність, послідовність та ефективність перекладу в умовах зростаючої глобалізації та цифровізації.



Література:

1. ISO 704:2022 «Terminology work – Principles and methods». URL: <https://www.iso.org/standard/79077.html>
2. ISO 860:2007 «Terminology work – Harmonization of concepts and terms». URL: <https://www.iso.org/standard/40130.html>
3. Вискушенко С. А. Проблеми стандартизації термінів у міждисциплінарних дослідженнях. Вісник науки та освіти. № 3 (33). 2025. С. 177-185
4. Naveen P., Trojovsky P. Overview and challenges of machine translation for contextually appropriate translations. iScience: CellPress, 2024. P. 1-25. URL: <https://www.cell.com/action/showPdf?pii=S2589-0042%2824%2902103-5>
5. Noll R., Frischen S. L., Boeker M., Storf H., Schaaf J. Machine translation of standardised medical terminology using natural language processing: A scoping review. New Biotechnology. № 77. 2023. P. 123-129. URL: <https://www.sciencedirect.com/science/article/pii/S1871678423000432?via%3Dihub>

References:

1. ISO 704:2022 «Terminology work – Principles and methods». URL: <https://www.iso.org/standard/79077.html>
2. ISO 860:2007 «Terminology work – Harmonization of concepts and terms». URL: <https://www.iso.org/standard/40130.html>
3. Vyskushenko S. A. Problemy standartyzatsii terminiv u mizhdystsyplinarnykh doslidzhenniakh. [Problems of Terminology Standartization in Interdisciplinary Research]. Visnyk nauky ta osvity. № 3 (33). 2025. S. 177-185
4. Naveen P., Trojovsky P. Overview and challenges of machine translation for contextually appropriate translations. iScience: CellPress, 2024. P. 1-25. URL: <https://www.cell.com/action/showPdf?pii=S2589-0042%2824%2902103-5>
5. Noll R., Frischen S. L., Boeker M., Storf H., Schaaf J. Machine translation of standardised medical terminology using natural language processing: A scoping review. New Biotechnology. № 77. 2023. P. 123-129. URL: <https://www.sciencedirect.com/science/article/pii/S1871678423000432?via%3Dihub>