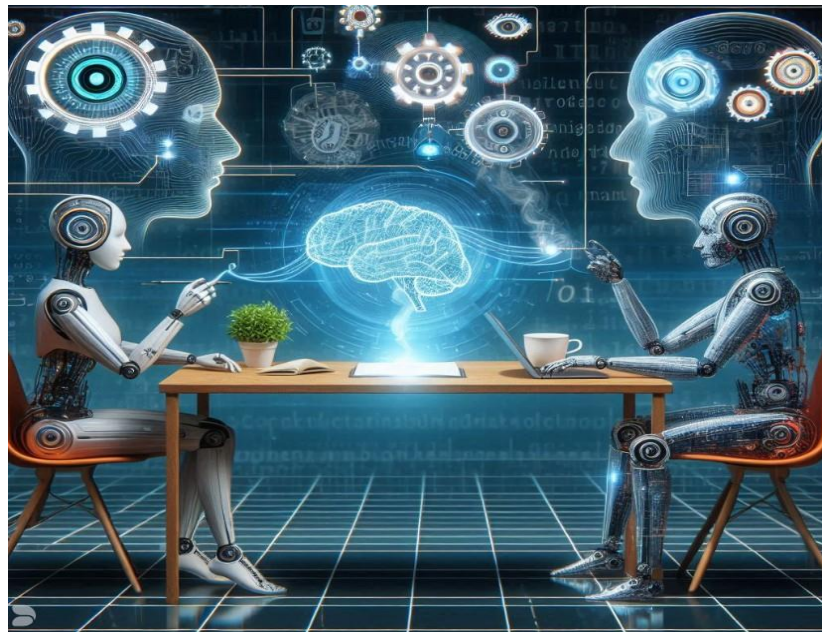


Білоцерківський національний аграрний університет
СОЦІАЛЬНО-ГУМАНІТАРНИЙ ФАКУЛЬТЕТ
Кафедра романо-германської філології та перекладу

Навчальний посібник
«Сучасні перекладацькі інформаційні
технології»

для здобувачів вищої освіти
спеціальності 035 «Філологія»



Біла Церква – 2025

УДК 811.111+811.161.2]’255

ISBN 978-966-2122-78-7

Затверджено Вченою радою Білоцерківського НАУ
(Протокол № 4 від 22 травня 2025 р.)

РЕЦЕНЗЕНТИ:

Наталія МАЙЄР, доктор педагогічних наук, професор, в.о. завідувача кафедри теорії, практики та перекладу французької мови, Київський політехнічний інститут імені Ігоря Сікорського

Катерина ГОСТРА, кандидат педагогічних наук, доцент кафедри перекладу та слов'янської філології, в.о. декана факультету іноземних мов, Криворізький державний педагогічний університет

Олена МИХАЙЛЕНКО, кандидат філологічних наук, доцент кафедри іноземних мов, Київський національний університет імені Тараса Шевченка.

Олена ДЕМ'ЯНЕНКО, кандидат педагогічних наук, доцент кафедри філології, педагогіки та методики викладання соціально-гуманітарного факультету Білоцерківського національного аграрного університету.

Ігнатенко В.Д. **Сучасні перекладацькі інформаційні технології:** навчальний посібник для здобувачів вищої освіти спеціальності «Філологія». Біла Церква, 2025. 139 с.

Навчальний посібник покликаний допомогти студентам першого (бакалаврського) рівня вищої освіти ознайомитися із сучасними перекладацькими технологіями та ефективно організувати процес вивчення обов'язкового компонента «Сучасні перекладацькі інформаційні технології».

У посібнику представлено 5 теоретичних тем, кожна з яких супроводжується інтерактивними завданнями до практичного заняття, та літературою для опрацювання. Крім завдань до теоретичного матеріалу посібника, здобувачам пропонується поглибити свої знання у процесі самостійного продовження освоєння теми за допомогою наданих матеріалів. Для розвитку творчих здібностей та самостійних наукових пошуків здобувачів пропонується виконати проєктне завдання, яке може бути виконане індивідуально, у парі або у групі.

Навчальний посібник призначений здобувачам першого (бакалаврського) рівня вищої освіти спеціальності 035 Філологія – майбутнім філологам, перекладачам. Цільовою аудиторією навчального посібника можуть бути також викладачі та магістри філології.

© Ігнатенко В.Д. 2025

ЗМІСТ

ТЕМА 1. Сучасні інформаційні інструменти перекладача.....

Цифровізація та трансформація ринку перекладацьких послуг як відповідь на виклики сьогодення

Диджитальні інструменти у перекладі: варто чи ні?

Історія машинного перекладу

Інтелектуальні Революції: застосування та порівняння Нейронного Машинного Перекладу та Великих Мовних Моделей

Нейронний машинний переклад

Штучний інтелект та перекладацька справа

Ключові ролі людини у світі машинного перекладу

Різниця між машинним перекладом та автоматизованим перекладом

ЛІТЕРАТУРА

СЕМІНАР

ТЕМА 2. CAT-інструменти: нові технології для допомоги перекладачу

Відмінність CAT від інших технологічних рішень у перекладі

Автоматизація процесу перекладу за допомогою CAT-засобів

Види CAT програм

Принцип роботи CAT-засобів

Основні елементи CAT програми

Різновиди сучасних CAT-програм для користування перекладачами

Переваги перекладів за допомогою CAT-засобів

ЛІТЕРАТУРА

СЕМІНАР

ТЕМА 3. Нові виклики у роботі перекладача: локалізація та транскреація

Поняття локалізації

Особливості локалізації

Переклад проти локалізації

Локалізація відеоігор

Специфіка роботи перекладача-локалізатора

Транскреація – творчий переклад

ЛІТЕРАТУРА

СЕМІНАР

ТЕМА 4. Аудіовізуальний переклад

Значення аудіовізуального перекладу

Історія аудіовізуального перекладу

Різновиди аудіовізуального перекладу

Дубляж

Субтитрування. Правила субтитрування

ЛІТЕРАТУРА

СЕМІНАР

ТЕМА 5. Сучасні стандарти якості перекладу та інструменти контролю

Стандарти якості перекладу. Історія.

Різновиди стандартів та критика.

Сучасні інструменти контролю якості перекладу.

Перекладач VS редактор VS коректор VS менеджер проекту.

QA – програми. Принципи роботи.

TQAUDITOR. Основні етапи роботи системи та недоліки.

ЛІТЕРАТУРА

СЕМІНАР

ЦІЛІ ВИВЧЕННЯ ДИСЦИПЛІНИ «СУЧАСНІ ПЕРЕКЛАДЦЬКІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ»

Метою вивчення дисципліни є формування сучасного рівня інформаційної та комп'ютерної культури майбутнього перекладача, зокрема: опанування системою знань, умінь і навичок, необхідних для раціонального використання засобів сучасних інформаційних технологій у галузі перекладу, професійної підготовки висококваліфікованих і конкурентно спроможних фахівців відповідно до вітчизняних та європейських стандартів.

Основним завданням вивчення дисципліни є формування у студентів сучасних компетентностей у сфері перекладу, ознайомлення з цифровими та автоматизованими інструментами, які допомагають оптимізувати перекладацьку діяльність, а також вивчення впливу штучного інтелекту на професію перекладача.

Вивчення сучасних інформаційних інструментів перекладача дозволить студентам зрозуміти, як цифровізація трансформує ринок перекладацьких послуг та які можливості й виклики виникають у зв'язку із застосуванням новітніх технологій. Особлива увага приділена історії машинного перекладу, розвитку нейронного машинного перекладу (NMT) та впровадженню великих мовних моделей (LLM) у перекладацький процес. Студенти розглянуть різницю між машинним та автоматизованим перекладом, а також ключові ролі людини у сфері машинного перекладу.

Опанування CAT-інструментів дозволить студентам зрозуміти принципи їх роботи, переваги та недоліки використання. Буде розглянуто відмінності між CAT-засобами та іншими технологічними рішеннями у перекладі, принцип роботи пам'яті перекладів, баз термінології та інших автоматизованих функцій.

Окрему увагу приділено викликам, що стоять перед перекладачами у сфері локалізації та транскреації. Студенти ознайомляться з поняттям локалізації, її відмінностями від традиційного перекладу, особливостями

локалізації відеоігор та специфікою роботи перекладача-локалізатора. Також розглянуто транскреацію як творчий підхід до перекладу контенту.

Також у фокусі аудіовізуальний переклад, який є важливою складовою сучасного перекладацького ринку, тому в межах курсу студенти вивчать основні методи аудіовізуального перекладу, такі як дубляж, субтитрування, озвучення та аудіодискрипція. Будуть розглянуті стандарти субтитрування та інші технічні аспекти роботи з аудіовізуальним контентом.

Завершальний етап курсу присвячено стандартам якості перекладу та інструментам контролю. Студенти знайомляться з міжнародними стандартами якості перекладу, такими як ISO, ASTM, а також з принципами роботи QA-програм та інших автоматизованих систем перевірки якості. Буде розглянуто функції та взаємодію перекладача, редактора, коректора та менеджера проєкту у процесі контролю якості.

Таким чином, курс «Сучасні перекладацькі інформаційні технології» допоможе студентам отримати актуальні знання та практичні навички, необхідні для успішної роботи у сфері перекладу в умовах цифрової епохи.

Під час вивчення обов'язкового компоненту «Сучасні перекладацькі інформаційні технології» майбутні філологи оволодіють загальними та фаховими компетентностями відповідно до Наказу МОН «Про затвердження стандарту вищої освіти за спеціальністю 035 «Філологія» для першого (бакалаврського) рівня вищої освіти» № 1379 від 20.06.2019 р.:

Загальні компетентності :

- Здатність учитися й оволодівати сучасними знаннями.
- Здатність застосовувати знання у практичних ситуаціях.
- Навички використання інформаційних і комунікаційних технологій.

Спеціальні компетентності:

- Здатність використовувати в професійній діяльності знання про мову як особливу знакову систему, її природу, функції, рівні.
- Здатність до збирання й аналізу, систематизації та інтерпретації мовних, літературних, фольклорних фактів, інтерпретації та перекладу тексту відповідно до спеціалізації 03541.
- Здатність вільно оперувати спеціальною термінологією (аграрною) для розв'язання професійних завдань.

ТЕМАТИЧНИЙ ПЛАН ЗАНЯТЬ

Змістовий модуль	Назва теми	Лекції	Практичні	Самостійна робота
Змістовий модуль 1	Сучасні інформаційні інструменти перекладача	4	4	12
	САТ-інструменти: нові технології для допомоги перекладачу	2	2	12
	Нові виклики у роботі перекладача: локалізація та транскреація	4	2	14
	Аудіовізуальний переклад	2	4	12
	Сучасні стандарти якості перекладу та інструменти перевірки якості перекладу.	2	2	12
	Всього 90 годин	14	14	62

ТЕМА 1. Сучасні інформаційні інструменти перекладача

1. Цифровізація та трансформація ринку перекладацьких послуг як відповідь на виклики сьогодення
2. Диджитальні інструменти у перекладі: варто чи ні?
3. Історія машинного перекладу
4. Інтелектуальні Революції: застосування та порівняння Нейронного Машинного Перекладу та Великих Мовних Моделей
5. Нейронний машинний переклад
6. Штучний інтелект та перекладацька справа
7. Ключові ролі людини у світі машинного перекладу
8. Різниця між машинним перекладом та автоматизованим перекладом

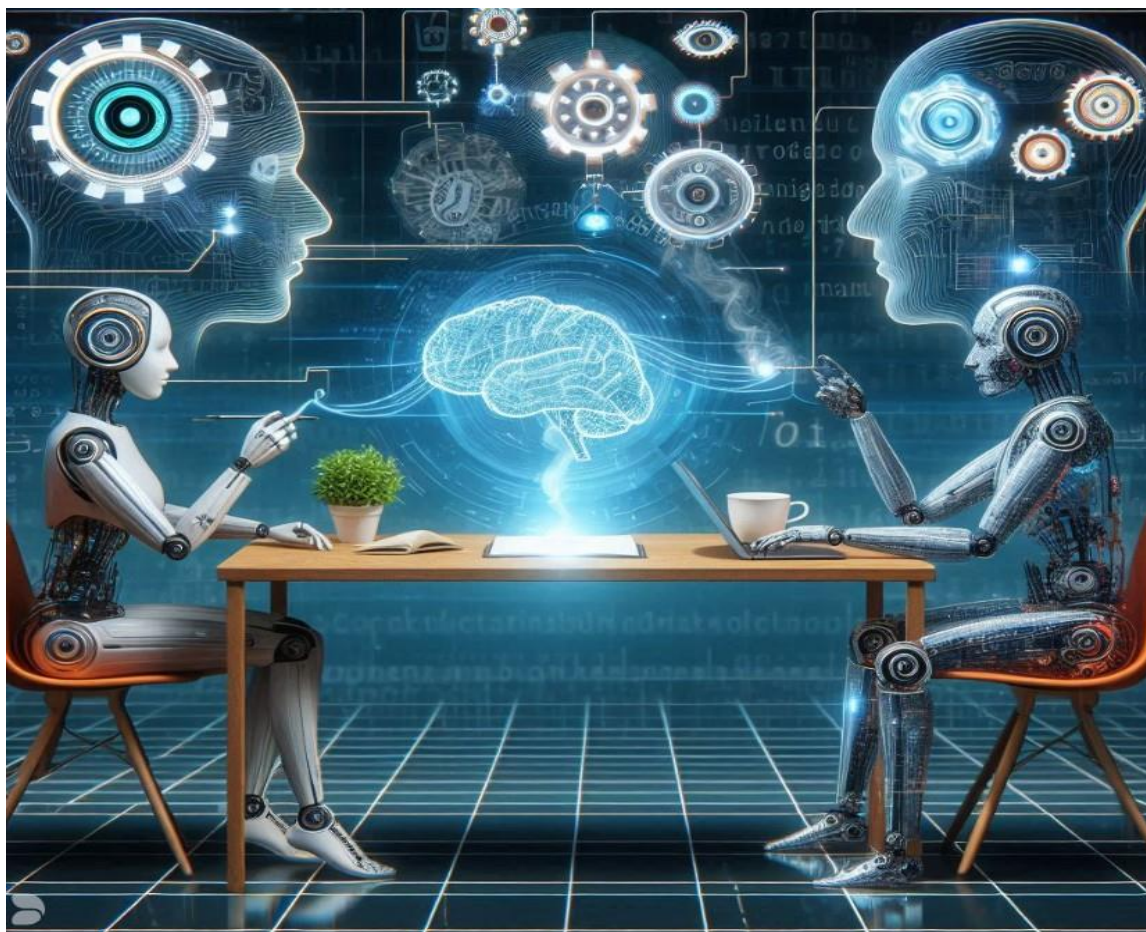


Рис. 1.1. Візуалізація концепції штучного інтелекту (згенеровано у Copilot)

ЦИФРОВІЗАЦІЯ ТА ТРАНСФОРМАЦІЯ РИНКУ ПЕРЕКЛАДАЦЬКИХ ПОСЛУГ ЯК ВІДПОВІДЬ НА ВИКЛИКИ СЬОГОДЕННЯ



Рис. 1.2. Символічне зображення зіткнення природи і технологій (згенеровано у Copilot)

В індустрії перекладу за останні роки відбулися значні зміни, які певною мірою співпадали із процесом інтернаціоналізації та глобалізації в українському суспільстві в цілому. Сьогоднішня ситуація на ринку перекладацьких послуг має певні особливості, які певним чином трансформують особистість сучасного перекладача. З економічної точки зору велика кількість гравців на ринку породжує

велику конкуренцію і разом з тим вимагає високого професіоналізму від учасників ринку. Учасникам ринку важко домовитися між собою, відсутня єдина платформа для уніфікації принципів перекладацької діяльності, що різко відрізняє український перекладацький ринок від зарубіжного ринку. Також менше монополії на перекладацькому ринку сприяє активному його розвитку, зокрема можливість працювати фрілансером.

Проте є і зворотна сторона цієї тенденції: надлишкова пропозиція породжує універсального перекладача, який має уміти перекладати тексти найрізноманітнішої тематики. На перекладацькому ринку все частіше з'являються компанії, які функціонують у відповідності до світових тенденцій, створюючи компанії з надання повноцінних перекладацьких послуг, де створюють та продають “переклад” як продукт. Мається на увазі задіявання до процесу не лише безпосередньо перекладача в єдиній ролі, а нарешті усвідомлюється цінність роботи окремо менеджера, редактора, коректора тощо. Працюючи на базі цифрової платформи у форматі онлайн офісу,

команда перекладацького бюро працює над замовленням як над цілісним проектом. Разом з тим зберігається постійна тенденція до збільшення кількості документів, що вимагають перекладу. Зокрема рішення про надання Україні статусу кандидата на членство в ЄС матиме наслідком також необхідність перекладу та оновлення різної документації, у тому числі за попередні роки на іноземну мову. Крім цього, відповідно до постанови Євросоюзу (Директива 79/112 / ЄЕС) вся технічна документація повинна бути перекладена на національні мови. Глобалізація і пов'язаний з нею зростання приведуть до колосального підйому в так званих секторах GILT (глобалізація, інтернаціоналізація, локалізація і трансація). Європейський союз асоціацій перекладацьких компаній (EUATC) вважає, що на ринку перекладів в найближчі роки очікується щорічне зростання приблизно на 5%. Все частіше перекладачі використовують інноваційні технологічні рішення у своїй роботі, щоб підвищити її швидкість і якість, доповнюючи та включаючи у свою компетентність уміння володіти машинним та автоматизованими системами перекладу (“кішки” або CAT-системи, тобто програми, які використовують технологію пам'яті перекладів (TM – translation memory)). Завдяки цим інструментам ринок перекладу виграє в усіх аспектах тому, що може швидко задовольняти потреби клієнтів, гарантуючи при цьому високу якість продукту. Штучний інтелект знаходить різноманітне застосування в перекладі та локалізації – ця тенденція ще довго буде актуальною. Потенціал штучного інтелекту величезний. Очікується, що його використання буде відрізнятися все більшою досконалістю.

Зростаюча конкуренція у світовому середовищі та можливості, котрі доступні міжнародним компаніям, що використовують можливості *локалізації*, підтверджують, що це ключова область для розвитку на довгі роки. Локалізація включає створення контенту (від реклами до веб-сайтів, відео і т.д.), який буде ідеально підходити для конкретної аудиторії. У кожній культурі є свої особливості, котрі враховує локалізація. У результаті створюється враження, що контент був створений саме для тієї аудиторії, на

яку він націлений. Клієнти можуть також звертатися до фахівців з постредагування, щоб їх переклади відповідали культурним особливостям та нормам цільової аудиторії. Також попит зростає на такі види перекладу, як **аудіовізуальний переклад, аудіодискрипція та транскреація**. Для прикладу, послуги **аудіовізуального перекладу** і, зокрема, перекладу відео, користуються величезним попитом, цей попит продовжує зростати разом з непомірним попитом на відеоконтент. **Аудіодискрипція (тифлокоментування)** полягає у створенні окремої звукової доріжки із закадровим описом персонажа, предмета, простору або дії у відеопродукції для осіб з порушеннями зору та осіб з дислексією. **Транскреація** є передбачає, що перекладений текст повинен мати схожий ефект на читача як у вихідній мові, так і мовою перекладу. Транскреація часто використовується в маркетингу та вимагає особливих творчих навичок від перекладача. У такому контексті годі і говорити про те, наскільки змінюється уявлення про перекладача у сучасному світі, яку важливу роль матиме переклад та як еволюціонує компетентність перекладача у конфронтації з великим обсягом та різноманіттям текстів для перекладу в нових умовах цифровізації. Основним гаслом для розвитку перекладацької галузі є локалізація, тобто стирання меж між різними картинами світу, максимальна адаптація до культурного коду мови перекладу і намагання зробити перекладача невидимим (“translator invisibility”) у цьому процесі.



Рис. 1.3. Символічне зображення людини і технічного прогресу як частини її буття (згенеровано у Copilot)

Диджитальні інструменти у перекладі: варто чи ні?



Рис. 1.4. Символічне зображення людини у вирі технічного прогресу (згенеровано у Copilot)

Перекладачі використовують сучасні технології та інноваційні технологічні рішення у своїй роботі, щоб підвищити її швидкість і якість. Відповідно, вони можуть перекладати краще і швидше, й за рахунок цього мати більший заробіток або ж насолоджуватися звільненим часом. Тут велику роль відіграють інструменти аналізу якості, розширення комунікаційних можливостей, а також покращена пам'ять перекладу. Завдяки цим

інструментам ринок перекладу виграє в усіх аспектах тому, що може швидко задовольняти потреби клієнтів, гарантуючи при цьому високу якість продукту.

На сьогодні використання онлайн-технологій або якщо генералізувати це поняття у терміні «автоматизація» на ринку перекладу, принаймні в Україні на сьогодні сприймається неоднозначно. Побуває думка, що це дешево, швидко, і майже завжди якість тексту перекладу нижче в порівнянні з перекладом здійсненим традиційним способом. Для багатьох перекладацьких компаній надзвичайно важливі перші два фактори, або задовольнити потреби клієнтів, перш за все, що не може не дратувати тих, хто присвятив всю свою кар'єру а подекуди і все своє професійне життя копіткому виконанню високоякісних перекладів. На жаль, в Україні поки досить фрагментований ринок перекладацьких послуг і поки кількість фрілансерів та перекладачів «з досвідом» дещо переважає кількість перекладачів, які працюють на великі компанії, які зацікавлені у розвитку.

Якщо відкинути емоції, а поглянути на факти, то щодня Гугл за офіційними даними опрацьовує близько 100 млрд. слів щодня (2018). Варто просто уявити, яку кількість перекладачів у світі треба було б залучити, щоб опрацювати таку кількість мов за один день...

Станом на травень 2022 року Google розмістив на своїй платформі Google Translate додаткові 24 нові мови, якими розмовляють понад 300 мільйонів людей. Тепер він підтримує загалом 133 мови на різних рівнях і може пропонувати переклади для 37 мов за допомогою фотографії, 32 за допомогою голосу в «режимі розмови» та 27 за допомогою живих відеозображень у режимі доповненої реальності. Після цього можна впастись у повний відчай і повністю розчаруватися у перекладі. Якщо автоматизація перекладу настільки ефективна, то чи є тут місце для людини? Чи є майбутнє у професії перекладача? Це чи не найперше питання, яке задається спікерам різних конференцій, жваво обговорюється у чатах та у всіх на вустах?

Але ринок невпинно розвивається і чомусь не відпадає потреба у перекладачах.

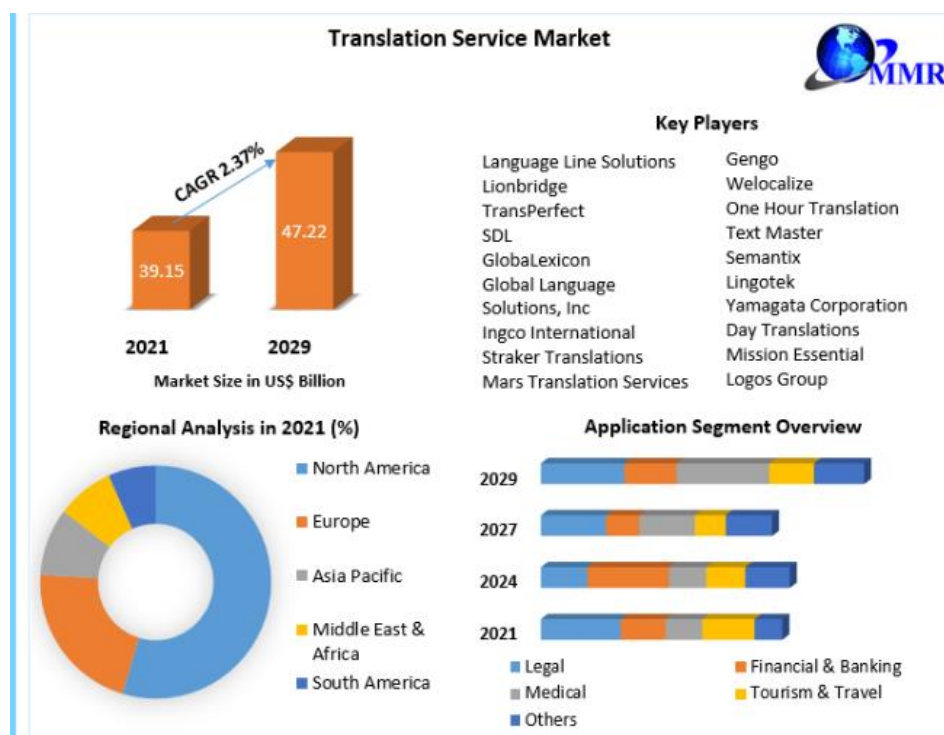


Рис. 1.5. Translation Service Market Overview [Електронний ресурс] // Maximize Market Research. – 2025. – Режим доступу: <https://www.maximizemarketresearch.com/market-report/global-translation-service-market/59434/>. – Дата звернення: 07.02.2025.

За прогнозами відомих компаній перспективи ринку перекладацьких послуг досить позитивні та багатообіцяючі. Відтак ринок перекладацьких послуг буде трансформуватися, а не зникати.

Історія машинного перекладу

Початок історії машинного перекладу датується сімнадцятим століттям, коли філософи **Лейбніц** і **Декарт** прогнозували створення кодів, що могли б зв'язати слова між мовами. Але на той час ці пропозиції залишились теоретичними і не стали основою створення фактичної машини.

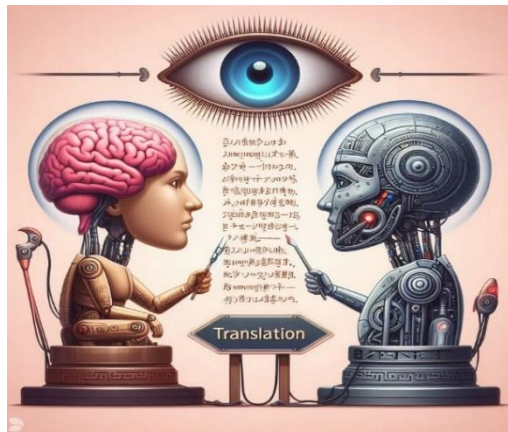


Рис. 1.6. Символічне зображення історії машинного перекладу (згенеровано у Copilot)

Перші розробки



A history of machine translation from the Cold War to deep learning

Вперше можливість машинного перекладу на практиці передбачив **Ч. Беббідж**, що у першій половині 19 століття працював над проектом цифрової аналітичної машини — механічного прототипу електронних цифрових обчислювальних машин.

Процес перекладу вчений уявляв так: *«У мене перед очима текст, написаний російською, але я збираюсь уявити, що насправді він написаний англійською, але за допомогою доволі дивних знаків. Все, що мені потрібно — це зламати код для того, щоб вилучити інформацію, що міститься у тексті».*

Перші патенти на створення перекладацьких машин було видано у середині 30-х років минулого століття. Ідея науковця з Західної Європи полягала у створенні автоматичного двомовного словника на основі перфострічки, але проект росіянина **П. Троянського** був детальнішим. Винайдений ним пристрій включав двомовний словник, здатний оперувати граматичними особливостями за принципом мови есперанто. Система поділялася на три стадії. Спочатку носій мови мав розподілити слова за їх логічними формами та синтаксичними функціями. Потім машина мала виконати переклад на потрібну мову, а носій — відредагувати висхідний матеріал.

Меморандум Вівера та подальші дослідження. Джорджтаунський експеримент

Першу пропозицію машинного перекладу за допомогою комп'ютера було висунуто **Уорреном Вівером**, дослідником з Фонду Рокфеллера у його меморандумі. Пропозиції базувалися на інформаційній теорії, успіхах у зламуванні кодів протягом другої світової війни та обговореннях універсальних та основних принципів мов. За кілька років після опублікування меморандуму розпочалися серйозні дослідження у багатьох університетах Сполучених Штатів.

Джорджтаунський експеримент у співпраці з IBM став демонстрацією машинного перекладу, і був представлений 7 січня 1954. Розроблений спільно із Джорджтаунським університетом і IBM, експеримент являв собою повністю автоматичний переклад більш ніж шістдесяти російських речень на англійську мову.

7 січня 1954 у Нью-Йорку в головному офісі IBM було вперше проведено публічну демонстрацію системи машинного перекладу (МП). Про демонстрацію повідомили в газетах, тож подія отримала широкий розголос. Попри те, що сама система мала лише 250 слів та 49 перекладених на англійську російськомовних речень (головним чином у області хімії) і була доволі примітивною, вона продемонструвала перспективи машинного перекладу, стимулювавши фінансування цього дослідження не тільки у США, а й у всьому світі. Експеримент було визнано успішним, що сповістило про початок ери вагомих капіталовкладень у дослідження машинного перекладу. Автори стверджували, що за кілька років машинний переклад буде повністю втілено в життя.

У ранніх системах використовувались великі двомовні словники та закодовані вручну правила для визначення порядку слів у висхідному продукті. У результаті цей метод було визнано занадто обмеженим, а завдяки тогочасному розвитку лінгвістики для покращення якості перекладу було запропоновано дослідження генеративної лінгвістики та трансформаційної

граматики. Але в цей час операційні системи вже застосовувались. Військово-повітряні сили США використовували систему, розроблену IBM і Вашингтонським університетом, в той час як у повітряних силах Італії працювала розробка Джорджтаунського університету. Попри те, що якість продукції була низькою, це задовольняло клієнтів, в основному з точки зору швидкості.

Наприкінці 1950-х Г. Бар, дослідник, що на замовлення США вивчав можливість створення повністю автоматичного якісного перекладу, наголосив на проблемі семантичної двозначності під час машинного перекладу. Розгляньмо наступний приклад: Little John was looking for his toy box. Finally he found it. The box was in the pen. Слово «pen» має два значення: прилад, що використовується на письмі і певний контейнер. Для людини значення є очевидним, але машина без «універсальної енциклопедії» ніколи не зможе вирішити цю проблему. Сьогодні проблема семантичної двозначності може бути вирішена шляхом написання висхідних текстів контрольованою мовою, тобто застосовуючи словник, у якому для кожного слова є тільки одне значення.

1960-ті, звіт ALPAC

Дослідження 1960-х років у СРСР та Сполучених Штатах сконцентрувались головним чином на російсько-англійській мовній парі. В основному об'єктами перекладу виступали науково-технічні документи, як от статті з наукових журналів. Недбалий переклад був достатнім для розуміння сенсу статей. Якщо стаття стосувалася інтересів безпеки, її надсилали живому перекладачу для повного перекладу, решту ж перекладали автоматично. Машинний переклад зазнав нищівного удару у 1966 році разом із публікацією звіту ALPAC (дорадчого комітету з автоматичної мовної обробки), що складався з семи вчених, скликаних американським урядом у 1964. Американський уряд був занепокоєний повільним просуванням експерименту попри значні видатки, тож було ухвалено рішення, що машинний переклад був дорожчим, менш точним та повільнішим за людський, і незважаючи на

витрати, машинний переклад навряд чи досягне якості людського найближчим часом. Однак, у звіті рекомендували продовжувати дослідження у галузі комп'ютерної лінгвістики та створити автоматичні словники, що допомагали б перекладачам.

Публікація звіту вплинула на дослідження машинного перекладу у Сполучених Штатах, меншою мірою – у Радянському Союзі та Великій Британії. Звіт, принаймні, майже повністю припинив будь-які дослідження у Сполучених Штатах майже на десятиліття. Однак, у Канаді, Франції та Німеччині дослідження продовжувались; у 1970 систему Systran використовували повітряні сили США, і, як наслідок, Комісією європейської економічної спільноти. Систему *METEO*, розроблену в Університеті Монреаля, було застосовано у 1977 року в Канаді для перекладу прогнозів погоди з англійської мови на французьку. Система перекладала близько 80 000 слів в день або 30 млн слів на рік, поки не була замінена системою конкурента 30-го вересня 2001.

1970-ті

В той час як дослідження 1960-х концентрувалися на проблемі обмеження мовних пар та системі вводу інформації, протягом 1970-х років постала потреба в недорогих системах, що змогли б перекладати діапазон технічних та комерційних документів. Ця вимога заохочувалася посиленням глобалізації та потреби перекладів в Канаді, Європі та Японії. До 1980-х років збільшилось різноманіття та число встановлених систем для машинного перекладу. Збільшилась кількість систем, що працювали на основі електронно-обчислювальних машин, як от Systran і Logos. Поширення мікрокомп'ютерів знаменувало створення дешевого ринку систем машинного перекладу, тому багато європейських, американських та японських компаній не втратили шансів цим скористатися. Схожі системи потрапили на ринки Китаю, Східної Європи, Кореї та Радянського Союзу.

Протягом 1980-х жвава діяльність у галузі машинного перекладу розгорнулася в Японії. *Завдяки комп'ютеру n'ятого покоління Японія мала*

намір перестрибнути конкурентів у галузі елементів електронних пристроїв та програмного забезпечення. Багато великих японських компаній було залучено до роботи над створенням англо-японських та японо-англійських штучних перекладачів. У дослідженнях 1980-х років переклад вбачався як певний різновид проміжного лінгвістичного відтворення, що залучає морфологічний аналіз разом з синтаксичним та семантичним. Наприкінці 1980-х відбувся вагомий стрибок у галузі створення нових методів для машинного перекладу. Система, розроблена IBM, базувалася на статистичних методах, інші групи застосовували техніку, що базується на великій кількості перекладів як зразків.

Протягом 1990-х років, після успіхів у розпізнанні мови та мовному синтезі, дослідження перейшло у стадію мовного перекладу. Відбувся значний ріст використання машинного перекладу в результаті появи дешевих та потужних комп'ютерів. На початку 1990-х машинний переклад став можливим не тільки на великих базових комп'ютерах, а й на персональних комп'ютерах та автоматизованих робочих місцях.

Можна також розглянути історію машинного перекладу через призму поколінь машинного:

1. Перше покоління систем, яке прийнято датувати до середини 1960-х рр., наближене за своєю технікою до послівного перекладу (word-for-word translation): до кожного слова або мовного звороту у вхідному тексті підбирається еквівалент у вихідному тексті, знайдений у словнику. Системи першого покоління не здатні вирішувати проблем багатозначності, не проводять жодного лінгвістичного аналізу, з огляду на що якість перекладу досить низька.

2. Наступний період розвитку систем МП (середина 1960–1970-х рр.) називають другим поколінням. Їх внутрішній устрій трохи складніший, ніж у першого покоління, такі системи МП базуються на морфологічних кореляціях між вхідною та вихідною мовами («морфологічні системи»), що є ефективними для організації перекладу в межах споріднених мов.

3. Структурно-граматичні системи МП, що ґрунтуються на синтаксичних кореляціях вхідної та вихідної мов – «синтаксичні системи». Ключовою процедурою цих систем є синтаксичний аналіз вхідної фрази із подальшою її трансформацією в структурно-синтаксичний каркас вихідної фрази.

4. Структурно-семантичні системи МП оперують глибинними структурами вхідного та вихідного контекстів. У таких системах передбачається багаторівневий аналіз та опрацювання мовного матеріалу.

Інтелектуальні революції: застосування та порівняння нейронного машинного перекладу та великих мовних моделей

Оскільки штучний інтелект (ШІ) і машинне навчання продовжують розвиватися, зростає і наша здатність обробляти та розуміти людську мову. Однією з найбільш значущих розробок у цій галузі є Large Language Model (LLM), технологія, яка має потенціал революціонізувати все, від обслуговування клієнтів до створення контенту.

Великі мовні моделі (LLM) — це тип алгоритму глибокого навчання, який обробляє та генерує текст, схожий на людину. Ці моделі є навчені на масивних наборах даних, що містять текст із різних джерел, таких як книги, статті, веб-сайти, відгуки клієнтів, публікації в соціальних мережах і огляди продуктів.

Основна мета LLM - зрозуміти та передбачити шаблони в людській мові, дозволяючи їй генерувати зв'язний та відповідний контексту текст.

Навчальний процес для LLM передбачає наступне:

- Відкриття моделі мільярдами або трильйонами речень.
- Дозволяє вивчати граматику, синтаксис і семантику.
- Дізнається фактичну інформацію.

У результаті ці моделі можуть відповідати на запитання, генерувати текст, перекладати мовами та виконувати багато інших завдань, пов'язаних із мовою, з високою точністю.

Великі мовні моделі (LLM) представляють собою значний крок вперед у нашій здатності обробляти та розуміти людську мову. Їх потенційні можливості застосування величезні: від подолання мовних бар'єрів за допомогою служб перекладу, як-от Google Translate, до генерування людського тексту та відповідей на запитання за допомогою GPT-4 OpenAI.

Основна відмінність заключається у тому, що у LLM використовується трансформер, який має свої певні особливості, як стверджує Тарас Малкович.

Нейронний машинний переклад (NMT) та великі мовні моделі (LLM) мають деякі спільні риси, але вони призначені для різних завдань та мають свої унікальні особливості.

1. **Призначення:** NMT спеціально розроблений для завдання машинного перекладу, тобто перетворення тексту з однієї мови на іншу.
2. **Архітектура:** NMT зазвичай використовує рекурентні нейронні мережі (RNN), довготривалу короткочасну пам'ять (LSTM), або трансформери. Трансформери стали найбільш популярними завдяки своїй ефективності.
3. **Обмеження:** NMT моделі спеціалізуються лише на перекладі і не призначені для широкого спектра інших завдань обробки природної мови (NLP).
4. **Навчання:** NMT моделі навчаються на паралельних корпусах текстів, де кожне речення на одній мові має відповідний переклад на іншій.

Великі мовні моделі (LLM)

1. **Призначення:** LLM, такі як GPT (Generative Pre-trained Transformer), створені для широкого спектра завдань NLP, включаючи генерацію тексту, відповідь на запитання, аналіз тональності, резюмування та навіть машинний переклад.
2. **Архітектура:** LLM зазвичай базуються на трансформерній архітектурі. GPT, наприклад, використовує багатошаровий трансформер.

3. **Універсальність:** LLM моделі є універсальними і можуть виконувати багато різних задач NLP завдяки своєму великому обсягу даних та багатозадачному підходу до навчання.
4. **Навчання:** LLM моделі навчаються на великих масивах текстових даних, що включають різні жанри, стилі та теми. Це дозволяє їм накопичувати знання та розуміння мови в широкому контексті.

Основні відмінності

- **Спеціалізація проти універсальності:** NMT моделі спеціалізовані на перекладі, тоді як LLM моделі є універсальними і можуть виконувати широкий спектр завдань NLP.
- **Навчальні дані:** NMT моделі навчаються на паралельних корпусах, а LLM моделі - на великих, різноманітних наборах текстових даних.
- **Архітектура:** Обидва підходи можуть використовувати трансформери, але NMT моделі можуть також використовувати інші типи нейронних мереж, такі як RNN або LSTM.

Хоча обидва підходи використовують сучасні нейронні мережі, NMT та LLM моделі відрізняються за своїм призначенням, обсягом завдань, які вони можуть виконувати, та типами даних, на яких вони навчаються.

Нейронний машинний переклад (NMT) та великі мовні моделі (LLM) **не є тотожними поняттями**. Вони мають різне призначення, хоча можуть використовувати схожі технології, такі як трансформери.

Ключові відмінності:

1. Призначення:

- **NMT (Neural Machine Translation):** Спрямований виключно на завдання перекладу тексту з однієї мови на іншу.
- **LLM (Large Language Models):** Створений для широкого спектра завдань обробки природної мови (NLP), включаючи генерацію тексту, відповіді на запитання, резюмування, аналіз тональності, переклад та інші.

2. Навчання:

- **NMT:** Навчається на паралельних текстах, де речення з однієї мови відповідає реченню іншої мови.
- **LLM:** Навчається на великих, різноманітних наборах текстових даних, які можуть включати тексти різних жанрів і тем.

3. Обмеження та Універсальність:

- **NMT:** Спеціалізується тільки на перекладі і, як правило, не призначений для інших завдань NLP.
- **LLM:** Універсальний, може виконувати багато різних завдань NLP, включаючи, але не обмежуючись перекладом.

4. Архітектура:

- Обидва можуть використовувати трансформери, але NMT також може використовувати інші архітектури, такі як RNN або LSTM.

Приклади:

- **NMT:** Google Translate, DeepL, які використовують спеціалізовані моделі для перекладу.
- **LLM:** GPT-3, GPT-4, BERT, які можуть виконувати багато різних завдань обробки природної мови, включаючи переклад.

Отже, хоча NMT і LLM можуть мати спільні технології та підходи, вони не є тотожними, оскільки мають різні цілі та можливості.

Нейронний машинний переклад

Нейронний машинний переклад (NMT) — різновид машинного перекладу, який передбачає використання штучної нейронної мережі для прогнозування ймовірності послідовності слів, зазвичай шляхом оперування одразу цілими реченнями в одній інтегрованій моделі.

Уперше програми глибокого навчання почали застосовувати в 1990-х роках у задачах із розпізнавання мовлення.

- Нейронний переклад використовується у сервісах перекладу багатьох компаній — як-от Google, Microsoft, Яндекс, ПРОМТ.
- Google використовує особливий різновид нейронного машинного перекладу — так званий Google Neural Machine Translation (GNMT).

- Майкрософт використовує подібну технологію для перекладу мовлення — зокрема, у Майкрософт Перекладачі та Skype Перекладачі.
- Гарвардська група з обробки природної мови випустила OpenNMT — систему нейронного машинного перекладу з відкритим вихідним кодом.

У Яндекс.Перекладачі використовується гібридна модель: свій варіант перекладу пропонує і статистична модель, і нейромережа, після чого за допомогою технології CatBoost, яка працює на основі машинного навчання, вибирається кращий з отриманих результатів.

Пропонувати системи на основі нейронних мереж почали й інші постачальники машинного перекладу, зокрема Omnisien Technologies (раніше Asia Online), KantanMT, SDL, Globalese, Systran тощо. DeepL надає загальну систему перекладу із системами штучного інтелекту глибокого навчання.

Система NMT всередині Google Перекладача використовує велику штучну нейронну мережу, придатну для глибинного навчання. Вивчаючи мільйони прикладів, GNMT покращує якість перекладу, використовуючи ширший контекст для виведення найякіснішого перекладу. Потім результат перебудовується й пристосовується для відповідності граматиці людської мови.

Проект Google Brain був створений у 2011 році в «секретній дослідницькій лабораторії Google X» Джеффом Діном, співробітником Google, Грегом Коррадо, дослідником із Google, та Ендрю Ином, професором комп'ютерних наук Стенфордського університету. Робота Ина лягла в основу одного з найбільших технологічних проривів у Google і Стенфорді.

У вересні 2016 року дослідницька група Google оголосила про розробку системи перекладу GNMT, і до листопада Google Перекладач почав використовувати нейронний машинний переклад (NMT) замість колишніх статистичних методів (SMT), які використовувалися з жовтня 2007 року всередині власної закритої SMT-системи.

Така еволюція можлива завдяки глибинному навчанню. Глибінне навчання (також відоме як глибинне структурне навчання, ієрархічне навчання, глибинне машинне навчання, англ. deep learning, deep structured learning, hierarchical learning, deep machine learning) — це галузь машинного навчання, що ґрунтується на наборі алгоритмів, які намагаються моделювати високорівневі абстракції в даних, застосовуючи глибинний граф із декількома обробними шарами, що побудовано з кількох лінійних або нелінійних перетворень.

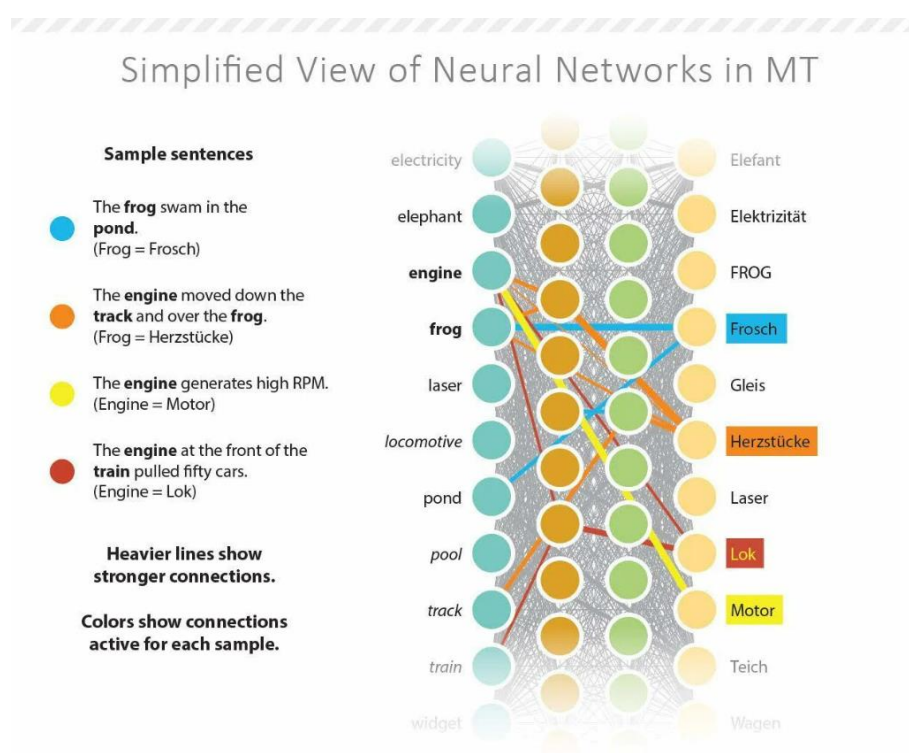


Рис. 1.7. [Електронний ресурс] // Big Language. – 2025. – Режим доступу: <https://biglanguage.com/blog/everything-to-know-about-neural-machine-translation/>– Дата звернення: 07.02.2025.

NMT функціонує, передбачаючи ймовірність певних послідовностей слів на основі шаблонів, які спостерігаються в наборах даних, що використовуються для навчання. Ключовим механізмом є векторне представлення. У NMT слова транскрибуються у вектори, кожен з яких має унікальну величину та напрямок, у процесі кодування та декодування. Движок аналізує вхідний текст, кодує його у вектори, а потім декодує в цільовий текст, передбачаючи ймовірний правильний переклад.

Хоча NMT все ще використовує імовірнісні моделі, він відрізняється від інших статистичних МП тим, що оцінює вхідні дані як єдине ціле, а не розділяє їх на підкомпоненти. Таким чином, він може виявляти зв'язки між словами, які посилюються або послаблюються залежно від того, які слова зустрічаються разом. Як результат, навчений NMT-двигун може в кінцевому підсумку робити вибір на основі цієї контекстної інформації для отримання точних і якісних результатів. На рисунку показано процес визначення зв'язків у NMT.

Штучний інтелект та перекладацька справа

Штучний інтелект (ШІ) використовується для створення комп'ютерних систем, які здатні виконувати завдання, що зазвичай потребують людського інтелекту. Основні аспекти ШІ включають машинне навчання, обробку природної мови, комп'ютерне зорове сприйняття та розпізнавання мови. Штучний інтелект використовується в різних галузях, включаючи автономні автомобілі, медицину, фінанси та інформаційні технології, з метою покращення ефективності та зручності людської діяльності.

Штучний інтелект (ШІ) стає все більш здатним у сфері перекладу текстів, і такі моделі, як ChatGPT, уже демонструють вражаючі результати. Однак питання про те, чи зможе ШІ повністю замінити людських перекладачів, залишається відкритим. У цьому розділі ми розглянемо як переваги, так і обмеження ШІ в перекладацькій справі, а також перспективи його співіснування з людськими фахівцями.

Переваги штучного інтелекту у перекладі

Швидкість. ШІ може перекладати великі обсяги тексту значно швидше, ніж людина. Це особливо корисно для термінових або великих проєктів.

Доступність. Перекладацькі інструменти на базі ШІ доступні онлайн і можуть бути використані в будь-який момент. Це забезпечує зручність для користувачів у будь-якому куточку світу.

Вартість. Використання ШІ може бути дешевшим, особливо для рутинних та великооб'ємних завдань. Це робить його привабливим варіантом для компаній і приватних осіб з обмеженим бюджетом.

Обмеження штучного інтелекту у перекладі

Якість перекладу. Хоча ШІ постійно вдосконалюється, він іноді робить помилки або неадекватно передає тон та нюанси тексту. Це може бути критично важливо для офіційних документів або художніх творів.

Культурні аспекти. Людські перекладачі краще розуміють культурні контексти та можуть адаптувати текст відповідно до них. ШІ часто не здатен врахувати всі культурні нюанси, що може призвести до невідповідностей у перекладі.

Креативність. Переклад літературних творів, поезії або текстів, де важлива креативність та стиль, все ще залишається викликом для ШІ. Люди здатні краще передати емоційний та художній зміст твору.

Майбутнє співіснування

Імовірно, в майбутньому ШІ буде використовуватися як інструмент для підвищення продуктивності перекладачів, а не як їх заміна. Людські перекладачі можуть використовувати ШІ для швидкого перекладу текстів, а потім редагувати та адаптувати їх відповідно до специфічних вимог та культурних нюансів. Це дозволить поєднати швидкість і ефективність ШІ з якістю та креативністю людської роботи.

Хоча ШІ значно покращує ефективність перекладу і може виконувати багато завдань швидко і дешево, людські перекладачі все ще залишаються незамінними для забезпечення високої якості, точності та культурної відповідності перекладів. ШІ та людські перекладачі мають потенціал для продуктивної співпраці, де кожен з них доповнює сильні сторони іншого, забезпечуючи найкращі результати в перекладацькій справі.

Таким чином, **майбутнє перекладацької справи виглядає як взаємодія між технологією та людським досвідом**, що забезпечить новий рівень ефективності та якості перекладів.

Майбутнє штучного інтелекту обіцяє багато цікавих можливостей і викликів. Прогрес у сфері штучного інтелекту передбачає розвиток більш складних систем, здатних до самонавчання і адаптації до нових умов. Очікується, що ШІ допоможе вирішувати складні глобальні проблеми, такі як зміна клімату, медична діагностика, автономний транспорт, ефективність виробництва та багато інших аспектів сучасного життя. Однак разом з тим, виникають питання щодо етики, безпеки та впливу ШІ на суспільство, які потребують уважного розгляду та регулювання.

Цікавим фактом є те, що у ChatGPT можуть з'явитися конкуренти. Anthropic розробляє власного чат-бота на ім'я Claude із штучним інтелектом, подібного до ChatGPT. Він доступний для iOS через додаток Poe від Quora і незабаром з'явиться на Android. Крім того, Anthropic представила нові моделі штучного інтелекту Claude 3: Haiku, Sonnet та Opus. Компанія також активно досліджує інтерпретованість систем машинного навчання з фокусом на архітектурі трансформера.

Ключові ролі людини у світі машинного перекладу

Pre-editing (попереднє редагування) - is the process of adjusting text before it is automatically translated in order to improve the raw quality of the machine translation out, aims to control the language and make the ST more digestible for the machine. **(перекладач редагує вхідний текст, пристосовуючи його для якіснішої обробки)**

Post-editing (постредагування) is a human translator activity which focuses on the correction of the outcome of a machine translated text. **(перекладач редагує перекладений машиною текст).** Істотне зростання обсягів машинного перекладу в останні роки привів до величезної кількості роботи для тих, хто займається його постредагуванням. Постредагування може включати виправлення граматичних помилок, невдалих формулювань, стилю і т.д., щоб перетворити переклад сумнівної якості в бездоганний. Клієнти

можуть також звертатися до фахівців з постредагування, щоб їх переклади відповідали культурним особливостям та нормам цільової аудиторії.

Постредагування має свої підвиди:

- *FAUT (Fully Automatic Useful Translation)*
- *Light post-editing*

„Light post-editing converts raw MT output into **understandable and usable**, but not linguistically or stylistically perfect, text...”

improving the output: consistent and appropriate

- *Full post-editing (GOOD ENOUGH)*

„is meant to produce human-quality output. The goal is to produce stylistically appropriate, linguistically correct output that is indistinguishable from what a good human translator can produce.” (DePalma, 2013, Online) the ‘gold standard’ human translation: accurate and fluent

Transcreation (транскреация) is a fusion of the words “translation” and “creation.” It describes copywriting content in a source text that needs to be made coherent, relevant, etc. in a new language. Термін транскреация (анг. transcreation) використовують рекламисти і їхні побратими — маркетологи. Так вони дають зрозуміти, що повідомлення потрібно перекласти з однієї мови на іншу, зберігаючи задум автора, його стиль і контекст. Транскреация застосовується, коли рекламний або маркетинговий текст був написаний, наприклад, для України, а його потрібно адаптувати для Франції. Наприклад: *булочка з маком в українському тексті може перетворитися на круасан у французькій адаптації та пончиком в англійській*. Транскреация дозволяє говорити з аудиторією зрозумілою їй мовою і домогтися від перекладу того ж впливу, що має оригінал. Творчий переклад — це робота не тільки з текстом, але і візуальними образами: зображеннями та відео. Перекласти самі лише слова для транскреации недостатньо. Слід врахувати культурні особливості цільової аудиторії, її звичаї, звички, історію, смаки, почуття гумору. Така робота під силу носію мови. Якщо відмінності між культурами значні, то іноді носію доводиться змінювати контекст, щоб спровокувати потрібні емоції в аудиторії.

Загалом процес **локалізації** не під силу машинному перекладу – це процес адаптації продукту або контенту до культурних, мовних та інших специфічних особливостей певного ринку чи регіону. Це включає переклад тексту, а також зміну формату дат, часу, валют, зображень та інших елементів, щоб забезпечити максимальну відповідність місцевим нормам і очікуванням. Локалізація виходить за межі простого перекладу, враховуючи культурні нюанси та особливості цільової аудиторії. Локалізація неможлива за допомогою лише технологій: необхідна участь перекладача. Технології машинного перекладу можуть бути корисними, але вони не враховують культурні та мовні нюанси, які є ключовими для точної і зрозумілої локалізації. Тільки професійний перекладач може забезпечити відповідність перекладу культурному контексту та зберегти його автентичність.



Рым А., lecture “AI and the future of translation. In search of evidence”

Але насправді технології не замінять повністю перекладача, про що розповідав у своєму навчальному відео Ентоні Пім. Про це також говорив Ентоні Пім на своєму YouTube каналі та присвятив цій темі цілу лекцію.



Рис. 1.8. Розвінчування міфів за Е.Пімом (власна презентація автора).

Різниця між машинним перекладом та автоматизованим перекладом

Кожен сучасний перекладач має знати, що це дві абсолютно різні речі. Це ті сучасні технології, які лежать в основі перекладацької компетентності сучасного фахівця.



Рис. 1.9. Символічне зображення протиставлення CAT програм та нейромереж (власна презентація автора)

Машинний переклад – це автоматичний переклад з вихідної мови на цільову за допомогою комп'ютерних програм або двигунів машинного перекладу без участі людини. І ось як виглядає приклад машинного перекладу.

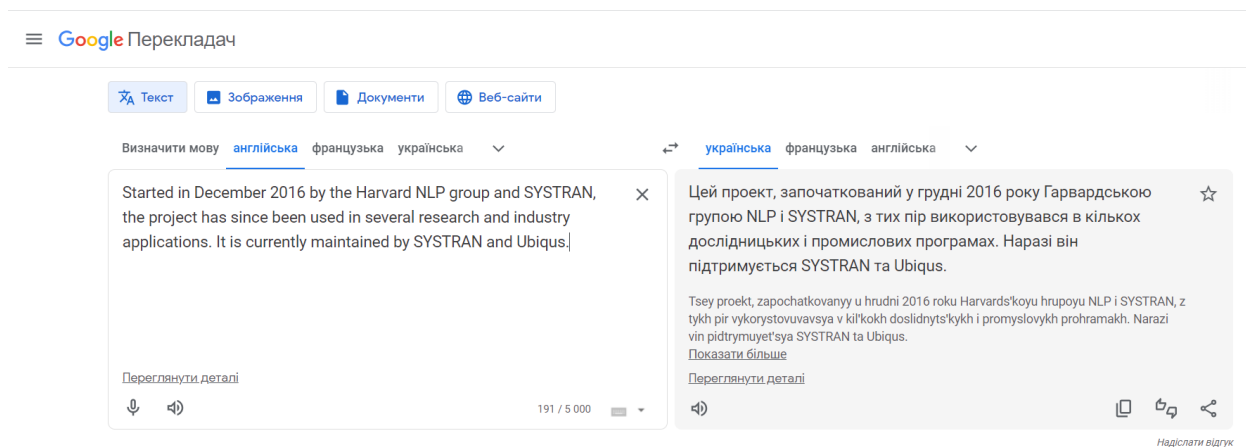


Рис. 1.10. Скріншот з Google Translate (Джерело: власний скріншот автора).

Система автоматизації перекладу (CAT-програма, CAT-tool) – це робочий інструмент професійних перекладачів і бюро перекладів. Автоматизований переклад — переклад текстів з використанням

комп'ютерних технологій (тобто використовує цю технологію саме ЛЮДИНА). Від машинного перекладу він відрізняється тим, що весь процес перекладу здійснюється людиною, комп'ютер лише допомагає їй створити готовий текст за менший час або з кращою якістю.

SDL TRADOS

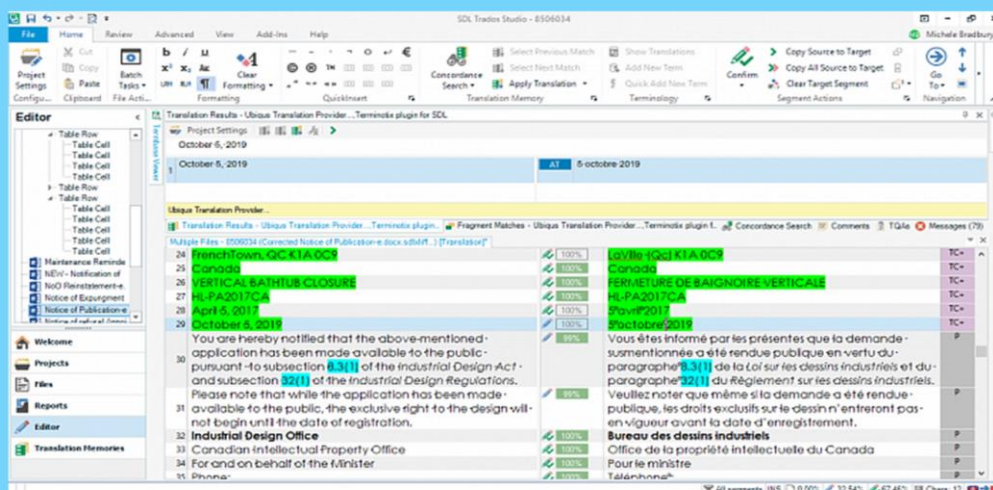


Рис. 1.11. Скріншот з SDL Trados 2019 (Джерело: власний скріншот автора).

Отже, головне – це пам'ятати і правильно використовувати сучасні технології і не забувати про відмінності та особливості кожного із видів перекладу. Отже, підсумуємо наявні сьогодні в арсеналі перекладача інформаційні технології для здійснення перекладу.

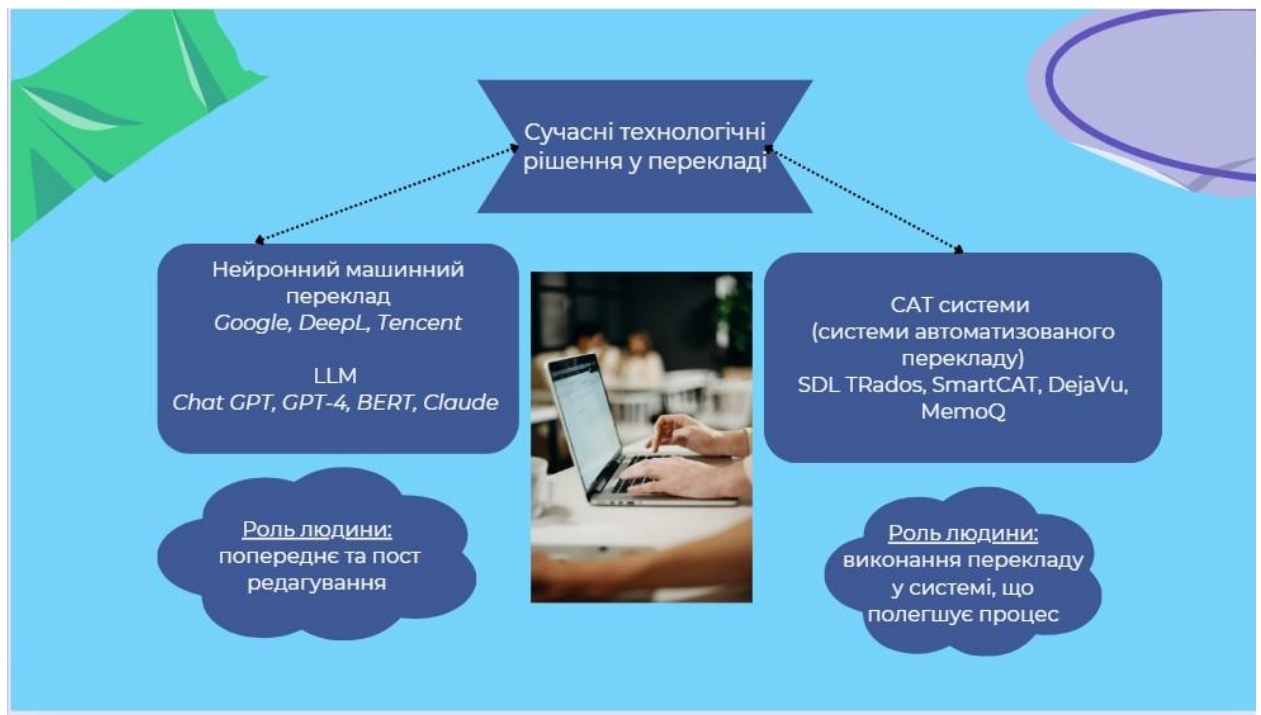


Рис. 1.12. Візуалізація класифікації сучасних технологічних рішень у перекладі (власна презентація автора).

ІНФОРМАЦІЙНІ ДЖЕРЕЛА

1. Амеліна, С. М., Тарасенко, Р. О. (2017). Підготовка перекладачів у зарубіжних університетах до застосування інформаційних технологій: Монографія. Київ: Компринт. 256 с.
2. Вергун, Л. (2016). Транскреція як засіб міжкультурної комунікації. У Інноваційні процеси економічного та соціально-культурного розвитку: вітчизняний та зарубіжний досвід (с. 170–171). THEU.
3. Притиченко, Г. (2021). Сучасні тенденції у перекладознавстві та їх вплив на роботу перекладача. Молодий вчений, 12(100), 222–225.
4. Цверкун, Ю. (2021). Тенденції перекладацької галузі у 2021 році. Взято з <https://profpereklad.ua/ru/tendencii-perevodcheskoj-otrasli-v-2021-godu/>
5. Ігнатенко В.Д. The Impact of ChatGPT on Translation Activity/ В.Д. Ігнатенко // Innovative development of science, technology and education. Proceedings of the 5th International scientific and practical conference (Vancouver, February 15-17, 2024). - Vancouver, Canada, 2024. – P. 456-458.
6. Ігнатенко В.Д. Artificial intelligence and translation: current realities and peculiarities / В.Д. Ігнатенко // Modern research in science and education : Proceedings of the 6th International scientific and practical conference (Chicago, February 8-10, 2024). - Chicago, USA, 2024. – P. 406-409.
7. Ігнатенко В.Д. Neural Network and Artificial Intelligence in Translation / В.Д. Ігнатенко // Актуальні питання інтернаціоналізації вищої освіти в Україні: лінгвістичний, правовий та психолого-педагогічний аспекти : матеріали IV Міжнар. наук.-практ. онлайн-конф. (БНАУ, 21-22 березня 2024 р.). – Біла Церква, 2024. – С. 3-5.
8. Ігнатенко В.Д. Місце машинного перекладу у діяльності сучасного перекладача: сучасні дискусії та бачення / В.Д. Ігнатенко // Modern philology: theory, history, methodology : Scientific monograph. Riga, Latvia : “Baltija Publishing”, 2024. - P. 2. - P. 523-534.

9. A history of machine translation from the Cold War to deep learning [Електронний ресурс]. - Режим доступу: <https://www.freecodecamp.org/news/a-history-of-machine-translation-from-the-cold-war-to-deep-learning-f1d335ce8b5/>.
10. Big Language Solutions. (n.d.). Everything You Need to Know about Neural Machine Translation. Взято з <https://biglanguage.com/blog/everything-to-know-about-neural-machine-translation/>
11. Milestone Localization. (n.d.). Emerging Trends in Translation Technology. Взято з <https://www.milestoneloc.com/translation-technology-trends/>
12. PYM, A. AI and the future of translation: In search of evidence [Електронний ресурс] / Anthony PYM. - 2023. - Режим доступу: <https://www.youtube.com/watch?v=46rUOOmNlkA>.
13. PYM, A. Basics of machine translation [Електронний ресурс] / Anthony PYM. - Режим доступу: https://www.youtube.com/watch?v=g_cksaRLdfY.
14. Venuti, L. (2017). Translator's Invisibility: A History of Translation. Taylor & Francis Group.