

УДК 371.2

Бородкіна Ірина Лаврентіївна,
кандидат технічних наук, доцент
кафедри комп'ютерних наук
Київського національного університету культури і мистецтв
borir@ukr.net
ORCID: <https://orcid.org/0000-0003-3667-3728>

Бородкін Георгій Олексійович,
старший викладач кафедри комп'ютерних наук
Національного університету біоресурсів і природокористування України
ORCID: <https://orcid.org/0000-0002-6488-6512>

МАШИННИЙ ПЕРЕКЛАД ЯК СКЛАДОВА ПІДГОТОВКИ МАЙБУТНІХ ДОКУМЕНТОЗНАВЦІВ

Мета роботи полягає в комплексному дослідженні систем машинного перекладу в структурі фахової діяльності документознавців. **Методологія** дослідження. Вибір матеріалів базувався на загальнонаукових методах аналізу та синтезу, порівняння та узагальнення, всебічності та об'єктивності. **Наукова новизна** дослідження обумовлюється тим, що на сьогодні практично відсутні роботи, присвячені обґрунтуванню необхідності навчання роботі з системами машинного перекладу майбутніх фахівців сфери документознавства. **Висновки.** Проведене комплексне дослідження систем машинного перекладу показало, що рушійною силою розвитку ринку електронних перекладацьких послуг є професійні потреби фахівців. На сьогодні для користувачів доступна велика кількість програмних продуктів цієї спрямованості, що зумовлює потребу формування навичок роботи з системами машинного перекладу вже на етапі фахової підготовки майбутніх документознавців та офіс-менеджерів.

Ключові слова: машинний переклад, он-лайн переклад, електронні перекладацькі послуги, документознавство, офіс-менеджер.

Бородкина Ирина Лаврентьевна,
кандидат технических наук, доцент
кафедры компьютерных наук
Киевского национального университета культуры и искусств

Бородкин Георгий Алексеевич,
старший преподаватель кафедры компьютерных наук
Национальный университет биоресурсов и природопользования Украины

МАШИННЫЙ ПЕРЕВОД КАК СОСТАВЛЯЮЩАЯ ПОДГОТОВКИ БУДУЩИХ ДОКУМЕНТОВЕДОВ

Цель работы заключается в комплексном исследовании систем машинного перевода в структуре профессиональной деятельности документоведов. **Методология исследования.** Отбор материалов базировался на общенаучных методах анализа и синтеза, сравнения и обобщения, всесторонности и объективности. **Научная новизна исследования** определяется тем, что на сегодняшний день практически отсутствуют работы, посвященные обоснованию необходимости обучения работе с системами машинного перевода будущих специалистов сферы документообращения. **Выводы** Проведенное комплексное исследование систем машинного перевода показало, что движущей силой развития рынка электронных переводческих услуг являются профессиональные потребности специалистов. На сегодня для пользователей доступно большое количество программных продуктов этой направленности, что обуславливает необходимость формирования навыков работы с системами машинного перевода уже на этапе профессиональной подготовки будущих документоведов и офис-менеджеров.

Ключевые слова: машинный перевод, онлайн перевод, электронные переводческие услуги, документообращение, офис-менеджер.

Borodkina Iryna,
PhD, Associate Professor
Department of Computer Science
Kiev National University of Culture and Arts

Borodkin Heorhii,
Senior Lecturer
Department of Computer Science
National University of Life and Environmental Sciences of Ukraine
george.borodkin@gmail.com

MACHINE TRANSLATION AS COMPONENT OF PREPARATION OF FUTURE DOCUMENT MANAGERS

The purpose of the work is a comprehensive study of machine translation systems in the structure of the professional activity of documentation specialists. **The methodology of the study** The selection of materials was based on the general scientific methods of analysis and synthesis, comparison and generalization, comprehensiveness and objectivity. **The scientific novelty** of the research is determined by the fact that today there are practically no works devoted to justifying the need for training in working with machine translation systems of future specialists in the field of documentation. **Conclusions.** A comprehensive study of machine translation systems showed that the driving force behind the development of the electronic translation services market is the professional needs of specialists. Today, a large number of software products of this type are available for users, which makes it necessary to develop skills in working with machine translation systems already at the stage of professional training of future document managers and office managers.

Keywords: machine translation, online translation, electronic translation services, document management, office manager.

Актуальність теми дослідження . У сьогоднішньому діловому спілкуванні надзвичайно швидко зростають обсяги документів, які потребують оперативного перекладу, причому кількість мов спілкування постійно зростає. Впоратися цією ситуацією можливо лише засобами сучасних перекладацьких технологій серед яких важливе місце посідають системи машинного перекладу. Такі системи здатні значною мірою спростити та прискорити роботу фахівця-документознавця без втрати інформації. З іншого боку – це дозволить краще спілкуватися з замовниками, більш чітко розуміти їх вимоги, що врешті рещт суттєво підвищить якість ділового спілкування. У той же час, на сьогодні в Україні практично відсутні науково обґрунтовані методики використання систем машинного перекладу і, відповідно, не здійснюється підготовка фахівців з цього питання. Це означає, що детального вивчення потребують питання, пов'язані з аналізом можливостей використання систем машинного перекладу в роботі фахівців-документознавців з метою розробки методики навчання майбутніх документознавців перекладу із застосуванням таких систем.

Аналіз досліджень і публікацій. Серед сервісів, які сьогодні надаються користувачам мережею інтернет, не останнє місце займає переклад з однієї мови на іншу [1]. Технології оперативного машинного перекладу дозволяють діловим людям орієнтуватися у величезному інформаційному полі, розбиратися в найновіших технологіях, легко спілкуватися з багатомовною аудиторією в багатьох країнах світу. Дуже часто виникає потреба просто зрозуміти сутність листа свого закордонного колеги або статті, знайденої в мережі Інтернет. Немає іншої реальної альтернативи як скористатися послугами електронного перекладача. Таким чином, машинний переклад виконує сьогодні дуже важливу комунікативну функцію – забезпечує можливість комунікації між людьми, які спілкуються різними мовами. Датою народження машинного перекладу як області досліджень прийнято вважати 1947 рік. Саме тоді, всього через декілька років після появи першої електронно-обчислювальної машини, через практичні потреби, пов'язані з істотно зростаючими обсягами науково-технічної інформації,

з'явилися перші програми машинного перекладу. У 80-і роки після появи перших персональних комп'ютерів доступ до яких можна було отримати будь-якої хвилини, а вартість машинного часу суттєво зменшилась, машинний переклад став економічно вигідним. Розпочався якісно новий етап розвитку цієї галузі. Вдосконалення програм машинного перекладу дозволило достатньо точно переводити багато видів текстів. Розвиток ринку персональних комп'ютерів (від настільних до кишенькових) і інформаційних технологій, поява сканерів, систем розпізнавання текстів і бурхливий розвиток мережевих технологій Internet/Intranet (які стають все більш інтернаціональними і багатомовними) у дев'яності роки двадцятого сторіччя призвели до чергового зростання інтересу до систем машинного перекладу і особливо до забезпечення можливості здійснення перекладу в інтерактивному режимі.

Машинний переклад – це процес створення тексту, представленого на одній природній мові, у еквівалентній за змістом текст іншою мовою. Як правило, цей процес виконується комп'ютером автоматично.[3]

Потреби фахівців різних сфер у послугах машинного перекладу постійно зростають, різноманітні системи машинного перекладу використовуються все більш інтенсивно. До 30 % перекладацьких компаній використовують системи машинного перекладу у професійній діяльності. Враховуючи, що якість такого перекладу постійно підвищується, можна зробити висновок про високу перспективність такого програмного забезпечення.

Відповідно до прогнозів компаній, які займаються детальним вивченням ринків програмного забезпечення [5], очікується, що ринок машинного перекладу протягом 2017 – 2023 років зросте на 20,9 %, що є доволі значним показником.

Машинний переклад, внаслідок своєї економічної ефективності сьогодні інтенсивно використовується у багатьох сферах інформаційної діяльності: для перекладу великих обсягів даних, для створення багатомовного контенту глобальних веб-сайтів, для перекладу текстів науково-технічної інформації, технічної документації (це пов'язано з постійно зростаючим експортом продукції та інформаційним обмі-

ном між науковцями різних країн). Ще однією сферою застосування систем машинного перекладу на сьогодні є великі міжнародні організації, що використовують тотожні документи, написані багатьма мовами (наприклад, ООН і офіційні структури Європейського Союзу). Документи тут повинні створюватись практично одночасно на багатьох мовах. Враховуючи, що вимоги до якості перекладу тут досить високі, документи, після перекладу з використанням систем машинного перекладу потребують додаткового редагування. В цілому можна зробити висновок, що системи машинного перекладу успішно функціонують в тих галузях, де існує потреба в перекладі великої кількості однотипних текстів, не потрібна абсолютна точність перекладу, а вхідна мова нормована і спрощена [6]. Цим умовам у великій мірі задовольняють інформаційні масиви, з якими працюють фахівці-документознавці.

Мета статті полягає в комплексному дослідженні систем машинного перекладу в структурі фахової діяльності документознавців.

Виклад основного матеріалу. 1990 року була запропонована наступна класифікація систем машинного перекладу:

- Системи повністю автоматизованого машинного перекладу (FAMT – Fully-automated machine translation)

- Системи машинного перекладу за участю людини (HAMT-Human-assisted machine translation)

- переклад, що здійснюється людиною з використанням комп'ютера (MANT – Machine-assisted human translation).

В основу іншого способу класифікації було покладено підходи до побудови алгоритмів машинного перекладу [2; 7]. Тут можна виділити машинний переклад, заснований на правилах (rule-based); машинний переклад, заснований на якості словникових баз і дозволяють визначати їх пріоритет. машинний переклад заснований на статистиці (statistical-based).

Системи машинного перекладу, що базуються на правилах (Rule-Based Machine Translation – RBMT), використовують великі збірки правил, що розробляються експертами-людьми. Людський фактор у таких системах дозволяє забезпечити досить гарні машинні переклади з передбачуваними результатами. Проте через необхідність залучення експертів

система, основана на правилах, може бути досить дорогою, а її впровадження та оновлення вимагає значних часових витрат.

Системи машинного перекладу, засновані на статистиці (Statistical Machine Translation – SMT) для створення перекладу використовують статистичні комп'ютерні алгоритми. Використання статистичних моделей та алгоритмів дозволяє мати систему низької вартості, однак дефіцит статистичних даних може знизити якість перекладу.

Останнім часом особливої популярності набуває гібридний (hybrid) підхід до розробки систем машинного перекладу. Гібридні системи машинного перекладу (Hybrid Machine Translation – HMT) розробляються на принципах, які поєднують відразу декілька алгоритмів дій.

З'являються і нові підходи до розробки систем машинного перекладу. Так, корпорація Google запропонувала використовувати нейронний машинний переклад [4]. (NMT) – алгоритм виконання машинного перекладу, що ґрунтується на штучній нейронній мережі та дає змогу подолати недоліки традиційних пофразових систем перекладу.

Можна виділити і інші ознаки класифікації існуючих на сьогодні засобів електронного перекладу.

Спосіб завантаження. За цією ознакою засоби електронного перекладу можна поділити на нерезидентні (програми, які неможна викликати з інших програм, вони автоматично перекладають наданий їм текст) та резидентні (завантажуються в пам'ять комп'ютера і підтримують інтерактивний режим перекладу для будь-якого програмного додатку).

Кількість словників, що підключається. Сучасні засоби електронного перекладу, наприклад LINGVO, працюють з великою кількістю словників.

Можливість розширення словникової бази. Сучасні електронні перекладачі мають засоби для створення власних і розширення існуючих словників.

Режим перекладу. На сьогодні користувачам надається два основних режими перекладу: автоматичний (або пакетний) і інтерактивний. У першому випадку програмі повідомляються імена файлів, призначених для перекла-

ду. Далі переклад здійснюється автоматично, без участі перекладача. Якість роботи подібних систем суттєво залежить від їх словникової бази. Інтерактивний режим перекладу передбачає завантаження програми-перекладача в оперативну пам'ять комп'ютера. Подальше використання електронного помічника може здійснюватись протягом всього часу роботи. Можливість розширення і одночасного підключення декількох словарних баз робить такі засоби дуже привабливими для користувачів.

До основних переваг електронних перекладачів слід віднести наступні.

Швидкість роботи з документом. Машинний переклад дозволяє практично миттєво перекладати вміст документа, що дозволяє швидко його зрозуміти і прийняти рішення про необхідність подальшого більш точного його перекладу[8]. Розглянемо деякі з них.

Низька вартість. Послуги професійних перекладачів не безкоштовні. Системи машинного перекладу потребують грошових витрат тільки при купівлі програмного забезпечення.

Доступність послуги. Електронний перекладач завжди доступний користувачу.

Конфіденційність. В системах машинного перекладу інформація повністю захищена.

Адаптивність. Системи машинного перекладу здатні запам'ятовувати основні терміни та фрази певної галузі. Це забезпечує єдиний мовний і термінологічний стиль при перекладі великих текстів.

Універсальність. Системи машинного перекладу універсальні, що забезпечується можливістю підключення відповідного словника.

Підтримка перегляду багатомовної інформації з мережі Інтернет. Мережа Інтернет – багатомовне інформаційне середовище. Системи машинного перекладу дозволяють переглядати і швидко перекладати сайти, створені різними мовами.

Відсутність витрат на форматування. Цей фактор є особливо важливим при перекладі електронних документів. При електронному перекладі повністю зберігається початкове форматування.

Основними недоліками систем машинного перекладу слід вважати відсутність точності (це дійсно так, як наслідок, обмеженість використання продуктів компанії InteractiveTran, NeuroTran, InterTran, PalmTran, PocketTran, MobileTran) та, як наслідок, обмеженість використання продуктів компанії InteractiveTran, NeuroTran, InterTran, PalmTran, PocketTran, MobileTran (існують певні типи документів, які неможливо перекласти електронним способом).

Окремого дослідження потребує питання оцінки якості машинного перекладу. Але слід зауважити, що кількість переваг систем машинного перекладу суттєво перевищує кількість недоліків. Крім того, розробники систем машинного перекладу постійно вдосконаджують свої програмні продукти, забезпечуючи постійне підвищення якості перекладу.

Це дозволяє зробити висновок про те, що машинний переклад є перспективним, він постійно і активно розвивається і використовується в професійній діяльності фахівців різних галузей.

У теперішній час в світі існує велика кількість компаній, які займаються розробкою комерційних систем машинного перекладу.

Systran (США) випускає наступні програмні продукти: SYSTRAN Professional Premium, SYSTRAN Professional Standard, SYSTRAN Personal, SYSTRAN Office Translator, SYSTRAN PDF Translator, SYSTRAN WebTranslator, SYSTRAN WebServer, SYSTRANet SYSTRAN-Box, які підтримують наступні напрямки перекладу: англійська – французька, англійська – німецька, англійська – іспанська, англійська – португальська, англійська – італійська, англійська – японська, китайська – англійська, англійська – російська, німецька – французька, англійська – корейська.

Transparent Language (США). Програмний продукт цієї фірми Easy Translator підтримує наступні напрямки перекладу: англійська – французька, англійська – німецька, англійська – іспанська, англійська – італійська, англійська – португальська.

Language Engineering Corporation (LEC) випускає програмні продукти: TRANSLATE Pro/Business/Personal, Power Translator, Personal/Pro/Global, Translate DotNet SDK, LEC Server, які підтримують такі напрямки перекладу: англійська – японська, англійська – іспанська, англійська – французька.

Translation Experts (США). Програмні продукти компанії InteractiveTran, NeuroTran, InterTran, PalmTran, PocketTran, MobileTran здійснюють переклади з англійської на німецьку, французьку, іспанську, угорську польську та хорватську мови.

Cross Language (Японія). Програмні продукти цієї компанії PC-Transer, PAT-Transer, MED-Transer, Kourai for Windows, Hourai V4 for Windows, Jx9 for Windows здійснюють переклад з англійської на японську мову.

Lingvistica (Канада). Електронний перекладач PARS підтримує переклад з англійської на голландську, польську, українську та російську мови.

ATA Software (Сполучене Королівство). Програмні продукти цієї компанії Al-Mutarjim Al-Arabey, golden Al-Wafi, Al-Wafi надають можливість перекладати з англійської на арабську.

ESTeam AB (Греція). ESTeam Translator розробки цієї компанії підтримує переклад всередині європейських мов.

XPLANATION (Бельгія). Сучасне інтегроване рішення цієї компанії під назвою Tstream здійснює переклад з англійської, німецької, французької, іспанської, італійської, голландської на англійську, німецьку, французьку, іспанську, італійську, голландську датську, норвезьку, фінську, шведську та португальську.

Kielikone Ltd (Фінляндія). За допомогою TranSmart Standard здійснюється переклад з англійської на фінську мову.

Крім означених вище, питаннями розробки систем машинного перекладу займаються фірми IBM, L&H (Lernout & Hauspie), SMART Communications (США), Atril, Trados, Trident Software, Lingvistica b.v., Nova Incorporated, LingoWare, IdiomaX (Швейцарія) та інші.

Онлайнні системи машинного перекладу. Онлайнні перекладачі, на відміну від локально кально, так і в онлайнному режимі. Система встановлених на персональному комп'ютері, до підтримує декілька мов і напрямів перекладу ступні з будь-якого комп'ютера, підключеного до мережі Інтернет. Сьогодні дуже поширеним став і такий зручний вид онлайн-перекладу як переклад веб-сторінок, доступний для користувачів мережевої системи машинного перекладу. Послугами автоматичних перекладачів безпосередньо в мережі Інтернет сьогодні можна скористатися за наступними адресами.

<http://pereklad.online.ua/> На сайті здійснюється переклад з української, російської, англійської, французької, німецької, латиської, польської на українську, російську, англійську, французьку, німецьку, латиську, польську мови. Система дозволяє використовувати спеціалізовані словники з наступних тематик: авто, бізнес, інколи доводиться здійснювати проміжний пе-

реклади, загальний, Інтернет, право, техніка, підтримує україномовний та російськомовний інтерфейси.

Translate.ru (www.translate.ru). Популярна система онлайнного перекладу текстів, веб-сторінок, електронної пошти та WAP, розроблена компанією ПРОМТ (www.promt.ru). Підтримує переклад з англійської, російської, німецької, французької, іспанської, італійської і португальської мов на англійську, російську, німецьку, французьку, іспанську, італійську і португальську мови (дозволяються будь-які мовні пари), за бажанням користувача може показувати варіанти перекладу. Крім того, перекладач надає можливість перекладу цілих веб-сторінок із збереженням початкового форматування, малюнків і фотографій та дозволяє перекладати електронні листи безпосередньо в браузері. Останні версії дозволяють використовувати цей перекладач на мобільних пристроях.

Перекладач компанії Google (www.google.com/translate_t) дозволяє перекладати текст і веб-сторінки. Має достатньо широкий набір мов, які можна використовувати в будь-яких комбінаціях.

Cognitive Translator (<http://cs.isa.ru:10000/ct/>) – безкоштовна онлайнна система машинного перекладу текстів компанії Cognitive Technologies. Система є інтерактивним середовищем, що забезпечує можливість перекладу, редагування і синхронного перегляду оригіналу тексту.

Перекладач Pragma (<http://www.trident.com.ua/rus/online.php>) компанії Trident

Software може здійснювати переклад як локально, так і в онлайнному режимі. Система встановлених на персональному комп'ютері, до підтримує декілька мов і напрямів перекладу ступні з будь-якого комп'ютера, підключеного до мережі Інтернет. Сьогодні дуже поширеним став і такий зручний вид онлайн-перекладу як переклад веб-сторінок, доступний для користувачів мережевої системи машинного перекладу. Послугами автоматичних перекладачів безпосередньо в мережі Інтернет сьогодні можна скористатися за наступними адресами.

Перекладацький сервіс Babel Fish Translation (<http://babelfish.yahoo.com/>) – це онлайнна версія відомої системи машинного перекладу Systran. Він пропонує переклади для великої кількості мовних пар. Дуже швидко Babel Fish Translation перекладає веб-сайти. Мовні пари тут ті ж самі.

Нескладно помітити, що не всі перекладачі підтримують можливість перекладу з української мови та на українську мову. Тому доводиться здійснювати проміжний пе-

реклад, наприклад, на російську, а потім вже з російської на українську.

Висновки. Проведене комплексне дослідження систем машинного перекладу як обов'язкових складників фахової підготовки майбутніх документознавців та офіс-менеджерів показало, що рушійною силою розвитку ринку електронних перекладацьких послуг є професійні потреби фахівців. Ринок машинного перекладу надзвичайно активно розвивається, а якість перекладу, виконуваного засобами систем машинного перекладу постійно поліпшується. В статті розглянуто способи класифікації систем машинного перекладу, визначено переваги

та недоліки їх використання. Проведене дослідження ринку програмного забезпечення систем машинного перекладу продемонструвало розмаїття таких систем та широту мовного охоплення. На сьогодні для користувачів доступна велика кількість програмних продуктів, здатних працювати як після інсталяції на комп'ютері, так і в он-лайн режимі.

Все вищезазначене дозволяє стверджувати, що наявність навичок роботи з системами машинного перекладу повинна стати обов'язковою вимогою в системі фахової підготовки майбутніх документознавців та офіс-менеджерів.

Список використаних джерел

1. Abraham-Barna C.G., Abraham-Barna T. Current trends in the translation market *Studii de știință și cultură*, 2016. Vol. XII. № 3. P. 33-43.
2. Burchardt A. Comparing Errors: Neural MT vs. Traditional Phrase-based and Rule-based MT. URL: <https://www.gala-global.org/publications/comparing-errors-neural-mt-vs-traditional-phrase-based-and-rule-based-mt>
3. Craciunescu O., Gerding-Salas C., Stringer-O'Keeffe S. Machine Translation and Computer-Assisted Translation: a New Way of Translating? *Translation Journal*. 2004. Vol. 8. № 3. URL: <http://translationjournal.net/journal/29computers.htm>
4. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. URL: <https://arxiv.org/abs/1609.08144>.
5. Machine Translation Market P&S Market Research. URL: <https://www.psmarketresearch.com/market-analysis/machine-translation-market>
6. Novak B. Machine Translation: Advantages and Disadvantages. URL: <http://dlsdc.com/blog/machine-translation-advantages-and-disadvantages>.
7. Qun L., Xiaojun Z. Machine translation: general. *The Routledge Encyclopedia of Translation Technology*. Routledge: London and New York, 2015. P. 105–119.
8. Колесник А. С. Обзор систем машинного перевода. Интеллектуальні системи та прикладна лінгвістика: матеріали II Всеукр. наук.-практ. конф., м. Харків, 28 березня 2013 р. Харків, 2013. С. 41–44.

References

1. Abraham-Barna C. G., Abraham-Barna T. (2016). Current trends in the translation market. *Studii de știință și cultură*. XII, 3, 33-43 [in English].
2. Burchardt, A. (2017). Comparing Errors: Neural MT vs. Traditional Phrase-based and Rule-based MT. Retrieved from <https://www.gala-global.org/publications/comparing-errors-neural-mt-vs-traditional-phrase-based-and-rule-based-mt> [in English].
3. Craciunescu, O. (2004). Machine Translation and Computer-Assisted Translation: a New Way of Translating? *Translation Journal*. 8(3). Retrieved from <http://translationjournal.net/journal/29computers.htm> [in English].
4. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation (2016). Retrieved from <https://arxiv.org/abs/1609.08144>. [in English].
5. Machine Translation Market P & S Market Research. (2017). Retrieved from <https://www.psmarketresearch.com/market-analysis/machine-translation-market> [in English].
6. Novak B. Machine Translation: Advantages and Disadvantages (2014). Retrieved from <http://dlsdc.com/blog/machine-translation-advantages-and-disadvantages>. [in English].
7. Qun L. (2015). Machine translation: general. *The Routledge Encyclopedia of Translation Technology*, 105–119 [in English].
8. Kolesnyk, A. S. (2013). Machine translation systems review. *Proceedings from II Ukrainian Scientific and Practical Conference "Intelligent Systems and Applied Linguistics"*. (pp. 41–44). Kharkiv. [in Russian].