

УТИЛІТА ДЛЯ АВТОМАТИЗОВАНОГО АНГЛІЙСЬКО-УКРАЇНСЬКОГО ПЕРЕКЛАДУ ІНТЕРФЕЙСУ ПРОГРАМ

Інтерфейс користувача грає велику роль під час роботи з різними обчислювальними пристроями. На сьогоднішній день можливість розробити та розповсюдити в мережі Інтернет свій програмний додаток має кожний бажуючий. Але далеко не кожний розробник досконало володіє декількома мовами світу. Мережа Інтернет надає вільний доступ до сучасних потужних систем перекладу тексту, але це все одно потребує від розробників дорогоцінний час. Це може суттєво обмежити успішне розповсюдження, кількість користувачів розробленої програми та навіть негативно вплинути на якість ключового програмного функціоналу. Тому постає питання створення якісного перекладу інтерфейсу користувача.

В роботі проведено дослідження існуючих засобів та представлено проект утиліти для розробників, бажуючих досягнути потенційних користувачів по всьому світі. У ході цього дослідження наведено приклади існуючих систем перекладу тексту в мережі Інтернет та оглянуто їх функціонал у контексті перекладу інтерфейсів користувача. Також було проведено аналіз цільової аудиторії та огляд сучасного стану ринків мобільних додатків, за допомогою якого показано актуальність проблеми мови інтерфейсів на комп'ютерах та мобільних пристроях, і визначено важливість зручного користування розробленою програмою для кінцевого споживача.

Під час дослідження було проведено огляд існуючих рішень для автоматизованого перекладу, виявлено їх ключовий функціонал та недоліки. В проекті запропонованої утиліти описано програмний інтерфейс перекладача, розроблений для збільшення зручності процесу перекладу саме інтерфейсів розроблюваного програмного забезпечення. Головний функціонал утиліти базуватиметься на використанні текстових файлів з спеціальним форматом записів для автоматичного відображення відповідних рядків в вихідному коді програми. Було проведено аналіз існуючих текстових форматів, їх використання в існуючих системах перекладу та розглянуто можливі підходи їх впровадження в проєктованій програмі. Також було запропоновано створення власного формату текстового файлу для утиліти перекладу.

Ключові слова: машинний переклад, розробка програмного забезпечення, інтерфейс користувача, мобільні додатки.

УТИЛІТА ДЛЯ АВТОМАТИЗИРОВАНОГО АНГЛІЙСЬКО-УКРАЇНСЬКОГО ПЕРЕВОДА ІНТЕРФЕЙСА ПРОГРАМ

Інтерфейс пользователя играет большую роль во время работы с разными вычислительными устройствами. На сегодняшний день возможность разработать и распространить в сети Интернет своё программное приложение имеет каждый желающий. Но далеко не каждый разработчик в совершенстве владеет несколькими языками мира. Сеть Интернет предоставляет свободный доступ к современным мощным системам перевода текста, однако это всё равно требует от разработчиков драгоценное время. Это может существенно ограничить успешное распространение, количество пользователей разработанной программы и даже

негативно повліять на якість ключового програмного функціонала. Поєтому встає питання створення якісного перекладу користувацького інтерфейсу.

В роботі проведено дослідження існуючих засобів і представлено проект утиліти для розробників, які хочуть досягти потенційних користувачів по всьому світу. В ході цього дослідження наведено приклади існуючих систем перекладу тексту в мережі Інтернет і оглянуто їх функціонал в контексті перекладу користувацьких інтерфейсів. Також було проведено аналіз цільової аудиторії і огляд сучасного стану ринку мобільних додатків, з допомогою якого визначено актуальність проблеми мовного інтерфейсу на комп'ютерах і мобільних пристроях, і показано важливість зручного користування розробленою програмою для кінцевого користувача.

В час дослідження було проведено огляд існуючих рішень автоматизованим перекладом, визначено їх ключовий функціонал і недоліки. В проекті запропонованої утиліти описано програмний інтерфейс перекладача, розроблений для підвищення зручності процесу перекладу саме інтерфейсів розроблюваного програмного забезпечення. Головний функціонал утиліти буде базуватися на використанні текстових файлів зі спеціальним форматом запису для автоматичного відображення відповідних рядків в вихідному коді програми. Було проведено аналіз існуючих текстових форматів, їх використання в існуючих системах перекладу і розглянуто можливі підходи до їх впровадження в проектувану програму. Також було запропоновано створення власного формату текстового файлу для утиліти перекладу.

Ключові слова: машинний переклад, розробка програмного забезпечення, користувацький інтерфейс, мобільні додатки.

V.D. MATUZKO, S.I. GOMENYUK
Zaporizhzhia National University

AUTOMATED ENGLISH-UKRAINIAN APPLICATION USER INTERFACE TRANSLATION TOOL

The user interface plays a big part when using various computing devices. Nowadays anyone has the ability to develop their own software application and distribute it on the Internet. However not every developer can be fluent in several world languages. The Internet provides free access to powerful modern text translation services, however that still requires precious time on behalf of the developer. This may severely limit the success and amount of users of the developed application, and even negatively impact the quality of the application's key functionality. This raises the question of creating a quality user interface translation.

This article analyses existing methods and presents a design concept of a tool for developers that wish to reach potential users across the whole world. The process of this research provides examples of existing text translation systems in the Internet and an overview of their features within the context of translating a user interface. Also conducted were analyses of the target audience and the current state of the mobile app markets, showing the relevance of the problem with user interface languages for computers and mobile devices, as well as the importance of providing a comfortable and quality application user experience for the end consumers.

The research also includes an overview of existing solutions for automated and assisted translation tools, their key functions and shortcomings. The proposed project of the translation utility describes a program interface for the translator, designed to make the translation process more comfortable and geared specifically for translating developed

software interfaces. The primary functionality of the translation tool would be based on using plain text files with a special format of records to enable automatic displaying of corresponding lines within the source code of the program. An analysis of such existing text formats was conducted, with considerations of their use and implementations in already existing translation systems and possible approaches of their implementation in the proposed translation tool. The creation of a custom text format specifically for the project was proposed as well.

Keywords: machine translation, software development, user interface, mobile applications.

Постановка проблеми

З кожним роком у світі збільшується кількість користувачів комп'ютерами та іншими пристроями. Новітні технології швидко поширюються і зараз доступні всім бажаючим. Люди мають можливість користуватися багатофункціональними персональними пристроями та підлаштовувати їх під свої унікальні потреби за допомогою програмних додатків. Велика кількість таких додатків вільно доступні у мережі Інтернет, та створені розробниками зі всього світу [1–2]. Тому виникає достатньо випадків, коли потрібна програма наявна з інтерфейсом лише на одній-двох мовах, а альтернативних рішень немає. Зазвичай це відбувається через нездатність або незручність створення пакетів локалізації розробниками додатків.

В даній роботі пропонується програмний засіб для полегшення та автоматизації створення перекладу інтерфейсу користувача. Розробники зможуть використовувати цю утиліту для поширення кола своїх користувачів не витрачаючи при цьому завеликих зусиль.

Аналіз останніх досліджень і публікацій

У сфері автоматизованого машинного перекладу уже більш двох десятиліть проводяться дослідження та розробляються все більш ефективні методи перекладу. Корпорації Google та Microsoft є одними з найбільших учасників в дослідженнях у цьому напрямку з їх багатофункціональними сервісами машинного перекладу [3–4]. Розробки цих та інших компаній вже зараз дозволяють людям зі всього світу спілкуватися між собою та користуватися іноземними веб-сайтами.

Але ще наявна проблема перекладу інтерфейсу користувача для комп'ютерних та мобільних додатків. Існує багато програм з інтерфейсом лише на рідній мові розробника, що суттєво обмежує потенційне коло користувачів. Також поширена проблема поганого використання машинного перекладу при створенні пакетів локалізації, що в багатьох випадках змушує користувачів вгадувати призначення тієї чи іншої функції в меню програми через неперевірений переклад.

У випадку додатків для мобільних пристроїв існують рішення за допомогою Google Translate [5] та Microsoft Translator [6]. Але цей підхід схожий по своїй зручності на постійне користування словником та не гарантує ідеальний результат перекладу у контексті кожного існуючого додатку. Також це вимагає додаткового часу та зусиль від кінцевого користувача.

Викладення основного матеріалу дослідження

Згідно статистиці кількість додатків, доступних в переліках Google Play та Apple App Store вимірюється мільйонами [7–8]. За їх допомогою користувачі мають можливість налаштувати свої пристрої для виконання чіткого переліку функцій. Однак велика кількість цих додатків має версії лише на одній мові, або використовує машинний переклад без перевірки коректності результату та урахування контексту в

інтерфейсі. Також якість машинного перекладу також залежить від конкретної мовної пари. Китай є одним з найбільших постачальників сучасних мобільних додатків. До того ж, машинний переклад з китайської історично відрізнявся своєю складністю у порівнянні з перекладом між європейськими мовними групами. Дослідження та розробки у цьому напрямку відбуваються постійно [9].

Щодо інтерфейсів комп'ютерних програм ситуація дуже схожа, але зазвичай розробники користуються допомогою команд перекладачів, тим самим уникаючи ситуацій з низькою якістю перекладу. Однак ці перекладачі мають виконувати переклад та знаходити методи і засоби для цього власноруч.

Існує багато програмних рішень створених для полегшення праці перекладачів, кожне з них зі своїм підходом до вирішення цього питання. Декілька з них навіть є престижними платними програмними пакетами з великим функціоналом для професіональних перекладацьких компаній [10]. Але при близькому огляді можна помітити відсутність програм націлених на переклад саме комп'ютерних інтерфейсів користувача. Цей вид перекладу відрізняється від перекладу текстів статей або книжок. Багато слів та фраз використовуються по декілька разів у різних місцях в додатках, при тому їх значення може суттєво змінюватися залежно від контексту та функціоналу програми. До того ж, зазвичай утиліти для автоматизованого перекладу мають інтерфейс більш спрямований для перекладу порівняно великих текстів, але через це не дуже ефективний при перекладі коротких рядків або окремих слів.

У цій роботі розглянуто проект утиліти, що спрямована на ефективну автоматизацію перекладу інтерфейсу користувача для невеликих команд розробників без необхідності використання спеціальних послуг перекладу. Однією з головних функцій в цьому проекті є можливість перегляду оригінального вихідного коду програми для повного розуміння контексту при перекладі інтерфейсу. Іншою важливою функцією є надання можливості використання машинного перекладу для підвищеної швидкості та ефективності роботи з подальшою можливістю провести контроль якості за допомогою зручного інтерфейсу перекладача.

Такий підхід вимагає створення спеціального інтерфейсу для перекладача, приклад якого наведений на рис. 1.

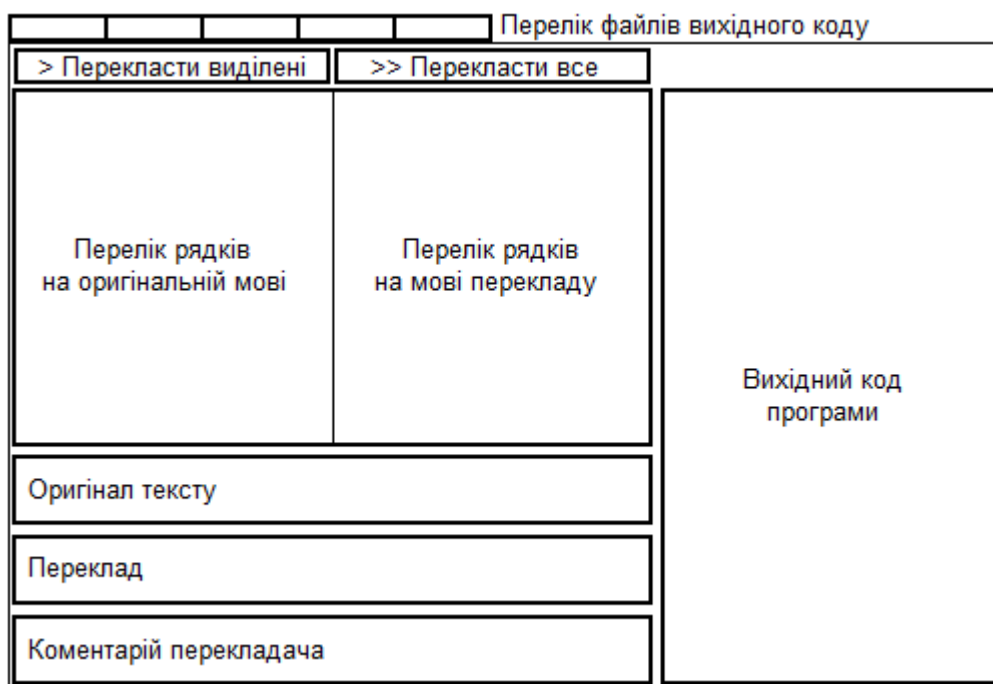


Рис. 1. Пропонована версія робочого інтерфейсу утиліти.

Цей інтерфейс має містити всю потрібну інформацію у зручному вигляді. Вікно вихідного коду програми має можливість автоматично знаходити рядок у коді, що відповідає обраному рядку оригінального тексту в переліку рядків. Таким чином перекладач миттєво бачить у якому контексті знаходяться елементи інтерфейсу в своїй програмі. Також інтерфейс надає можливість працювати з декількома файлами одночасно за допомогою вкладок. Автоматизація відбувається з використанням існуючих онлайн-сервісів машинного перекладу через програмний доступ. Як приклад можна привести Google Translate API, який дозволяє розробникам вбудовувати доступ до послуг перекладу в свої додатки [11].

Для користування машинним перекладом перекладач постачає свій ключ доступу до API. Це надає можливість автоматично локалізувати деякі або всі рядки в обраному файлі. Після цього можна з легкістю провести контроль якості проведеного перекладу та скоригувати знайдені помилки без необхідності витратити час на послідовний переклад кожного рядка.

Для використання утиліти розробник має належно підготувати інтерфейс до локалізації та експортувати текстову інформацію інтерфейсу в один з існуючих форматів даних, що використовуються для локалізації текстів. Одним з популярних таких форматів є .PO, який використовується в системі локалізації GNU gettext [12]. Ця система потребує від розробника написання коду з урахуванням його наступної локалізації. Це також забезпечує функціонування вікна перегляду вихідного коду програми в пропонованій утиліті. Після цього процесу розробник може згенерувати PO-файл, в якому буде міститись набір рядків, відповідних тексту інтерфейсу. Формат записів в файлах PO показаний на рис. 2.

```
white-space
# translator-comments
#. extracted-comments
#: reference...
#, flag...
#| msgid previous-untranslated-string
msgid untranslated-string
msgstr translated-string
```

Рис. 2. Схема запису в форматі PO.

Отриманий файл потім можна відкрити в утиліті для перекладу. Особливість такого підходу полягає у тому, що підтримка gettext реалізована тільки для конкретного переліку популярних мов програмування, тому його важко назвати універсальним рішенням для розробників [13].

Іншою альтернативою є формат XLIFF (XML Localization Interchange File Format). На відміну від попереднього методу XLIFF не має залежності від мови програмування, що дозволяє використовувати його у значно більшій кількості випадків, ніж gettext [14]. Але через це від розробника додатку потребується створення або використання існуючої бібліотеки для зчитування інформації з XLIFF-файлів. Як і в випадку з gettext інтерфейс програми також треба підготувати до локалізації, але на відміну від gettext реалізація залежить від самого розробника.

Стандарт XLIFF базується на форматі XML, тому на вигляд він більш громіздкий, ніж PO (рис. 3). Це не впливає на робочий процес при використанні утиліти.

```
<xliff xmlns="urn:oasis:names:tc:xliff:document:2.0" version="2.0"
srcLang="en-US" trgLang="ja-JP">
<file id="f1" original="Graphic Example.psd">
<skeleton href="Graphic Example.psd.skl"/>
<unit id="1">
<segment>
<source>Quetzal</source>
<target>Quetzal</target>
</segment>
</unit>
<unit id="2">
<segment>
<source>An application to manipulate and process XLIFF documents</source>
<target>XLIFF 文書を編集、または処理するアプリケーションです。</target>
</segment>
</unit>
</file>
</xliff>
```

Рис. 3 Приклад записів в форматі XLIFF.

Сама утиліта має мати можливість працювати з хоча б одним з вище зазначених форматів, а також мати алгоритм пошуку відповідних рядків тексту в вихідному коді програми. Також корисною буде можливість збереження вищезазначених ключів API в захищеному вигляді для поліпшення продуктивності перекладачів.

Висновки

В даній роботі розглянуто проблему низькоякісних інтерфейсів в різних видах програмного забезпечення, існуючі рішення цієї проблеми та недоліки цих рішень. Також було запропоновано та описано проект програмної утиліти, що спеціалізована для перекладу інтерфейсів користувача та дозволяє цей автоматизувати процес. Як подальше дослідження розробки додаткових функцій утиліти можуть бути створення і використання повністю власного формату для збереження та роботи над перекладом.

Список використаної літератури

1. How Many People Have Smartphones Worldwide (June 2020). URL: <https://www.bankmycell.com/blog/how-many-phones-are-in-the-world> (дата звернення: 04.06.2020)
2. Global Digital Overview – DataReportal – Global Digital Insights. URL: <https://datareportal.com/global-digital-overview> (дата звернення: 04.06.2020)
3. Google Translate – A Personal Interpreter on Your Phone or Computer. URL: <https://translate.google.com/intl/en/about> (дата звернення: 05.06.2020)
4. Microsoft Translator. URL: <https://www.microsoft.com/en-us/translator> (дата звернення: 05.06.2020)
5. Translate text in other apps – Android – Google Translate Help. URL: <https://support.google.com/translate/answer/6350658?co=GENIE.Platform%3DAndroid&hl=en> (дата звернення: 05.06.2020)
6. Microsoft Translator Adds Image Translation to Android – Microsoft Translator Blog. URL: <https://www.microsoft.com/en-us/translator/blog/2016/04/20/microsoft-translator-adds-image-translation-to-android/> (дата звернення: 05.06.2020)
7. Google Play: number of available apps by quarter 2020 | Statista. URL: <https://www.statista.com/statistics/289418/number-of-available-apps-in-the-google-play-store-quarter/> (дата звернення: 06.06.2020)
8. Apple App Store: number of available apps by quarter 2020 | Statista. URL: <https://www.statista.com/statistics/779768/number-of-available-apps-in-the-apple-app-store-quarter/> (дата звернення: 06.06.2020)
9. Achieving Human Parity on Automatic Chinese to English News Translation. arXiv:1803.05567 [cs.CL] URL: <https://arxiv.org/abs/1803.05567> (дата звернення: 06.06.2020)

10. Top Translation Software Tools in 2020 – Some Even Free!. URL: <https://www.smartcat.ai/blog/top-translation-software-tools-in-2019-some-even-free/> (дата звернення: 07.06.2020)
11. Try the Free Website Translator & Translate API – Google Translate. URL: <https://translate.google.com/intl/en/about/forbusiness/> (дата звернення: 07.06.2020)
12. gettext – GNU Project – Free Software Foundation (FSF). URL: <https://www.gnu.org/software/gettext/> (дата звернення: 07.06.2020)
13. GNU gettext utilities: List of Programming Languages. URL: https://www.gnu.org/software/gettext/manual/html_node/List-of-Programming-Languages.html (дата звернення: 07.06.2020)
14. XLIFF Version 2.1. URL: <https://docs.oasis-open.org/xliff/xliff-core/v2.1/xliff-core-v2.1.html> (дата звернення: 07.06.2020)

References

1. How Many People Have Smartphones Worldwide (June 2020). Retrieved from: <https://www.bankmycell.com/blog/how-many-phones-are-in-the-world> (last accessed: 04.06.2020)
2. Global Digital Overview – DataReportal – Global Digital Insights. Retrieved from: <https://datareportal.com/global-digital-overview> (last accessed: 04.06.2020)
3. Google Translate – A Personal Interpreter on Your Phone or Computer. Retrieved from: <https://translate.google.com/intl/en/about> (last accessed: 05.06.2020)
4. Microsoft Translator. Retrieved from: <https://www.microsoft.com/en-us/translator> (last accessed: 05.06.2020)
5. Translate text in other apps – Android – Google Translate Help. Retrieved from: <https://support.google.com/translate/answer/6350658?co=GENIE.Platform%3DAndroid&hl=en> (last accessed: 05.06.2020)
6. Microsoft Translator Adds Image Translation to Android – Microsoft Translator Blog. Retrieved from: <https://www.microsoft.com/en-us/translator/blog/2016/04/20/microsoft-translator-adds-image-translation-to-android/> (last accessed: 05.06.2020)
7. Google Play: number of available apps by quarter 2020 | Statista. Retrieved from: <https://www.statista.com/statistics/289418/number-of-available-apps-in-the-google-play-store-quarter/> (last accessed: 06.06.2020)
8. Apple App Store: number of available apps by quarter 2020 | Statista. Retrieved from: <https://www.statista.com/statistics/779768/number-of-available-apps-in-the-apple-app-store-quarter/> (last accessed: 06.06.2020)
9. Achieving Human Parity on Automatic Chinese to English News Translation. arXiv:1803.05567 [cs.CL]. Retrieved from: <https://arxiv.org/abs/1803.05567> (last accessed: 06.06.2020)
10. Top Translation Software Tools in 2020 – Some Even Free! Retrieved from: <https://www.smartcat.ai/blog/top-translation-software-tools-in-2019-some-even-free/> (last accessed: 07.06.2020)
11. Try the Free Website Translator & Translate API – Google Translate. Retrieved from: <https://translate.google.com/intl/en/about/forbusiness/> (last accessed: 07.06.2020)
12. gettext – GNU Project – Free Software Foundation (FSF). Retrieved from: <https://www.gnu.org/software/gettext/> (last accessed: 07.06.2020)
13. GNU gettext utilities: List of Programming Languages. Retrieved from: https://www.gnu.org/software/gettext/manual/html_node/List-of-Programming-Languages.html (last accessed: 07.06.2020)

14. XLIFF Version 2.1. Retrieved from: <https://docs.oasis-open.org/xliff/xliff-core/v2.1/xliff-core-v2.1.html> (last accessed: 07.06.2020)

Гоменюк Сергій Іванович – д.т.н., професор, завідуючий кафедри програмної інженерії Запорізького національного університету, e-mail: gserega71@gmail.com, ORCID: 0000-0001-7340-5947.

Матузко Володимир Дмитрович – аспірант кафедри програмної інженерії Запорізького національного університету, e-mail: matuzkovd@ukr.net, ORCID: 0000-0002-3005-6051.