

УДК 004.94
К.Я.Бортник,Т.С.Курило
Луцький національний технічний університет

ROBONEWS – СИСТЕМА АВТОМАТИЗОВАНОГО ПЕРЕКЛАДУ НОВИН

У роботі висвітлено питання в необхідності машинного перекладу великої кількості текстових даних. Розглянуто спосіб вирішення цієї проблеми за допомогою декількох сервісів онлайн перекладу тексту. Розкрито потенціал розробленої системи RoboNews та її переваги над іншими системами.

Мережа Інтернет є досить ефективним місцем для саморозвитку та самовдосконалення. Тому розробка різноманітних додатків, систем та модулів які використовують дані та сервіси які надає ця мережа є перспективною, та досить вигідною. Працюючи в мережі Інтернет над своїм сайтом я натрапив на необхідність машинного перекладу на іноземну мову великої кількості різноманітних текстових матеріалів. В якості іноземної мови в мене виступала Російська, але моя система автоматизованого перекладу була створена з можливістю підтримки великої кількості іноземних мов, та можливістю автоматичного визначення мови оригіналу тексту. Ручний переклад таких даних був не ефективним адже займав дуже великий об'єм часу. Тому підчас розробки системи RoboNews я вирішив використати наявні безкоштовні сервіси машинного перекладу, для того щоб полегшити, автоматизувати та пришвидшити виконання цього нелегкого та дуже необхідного завдання. Враховуючи те, що я розраховував на якісний переклад матеріалу, я розглянув наступні сервіси, та переглянув якість перекладу та всі плюси і мінуси. Серед них були "Google Translate", "Bing Translator" та "Яндекс.Перевод". Перші два сервіси мали відповідні API, тому розглядати можливість автоматичного перекладу було досить легко. Останній ж сервіс наразі немає API, тому працювати з ним в автоматичному режимі було складніше. Найбільш якісні переклади мали такі системи як "Google Translate" та "Яндекс.Перевод". При створенні системи RoboNews я надав перевагу "Google Translate", так як якість перекладу була на досить високому рівні, та було відповідне API. Саме на базі безкоштовних можливостей цього сервісу і була розроблена система, яка виконувала ряд наступних важливих можливостей, які робили її унікальною та ефективною:

- 1.Можливість автоматичного визначення мови оригінального тексту
- 2.Можливість ручного задавання мови оригінального тексту
- 3.Переклад оригінального тексту на задану адміністратором мову
- 4.Можливість побудови правил для автоматичного перекладу тексту без людського втручання

Система RoboNews була розроблена як компонент системи управління контентом "DataLife Engine". Ця система управління контентом написана на мові PHP, тому і розробка проекту RoboNews проводилась на мові PHP, яка набула досить високого вжитку при створенні персональних сайтів в мережі Інтернет.



Рисунок 1.

На Рисунку 1. можна побачити відображення системи RoboNews в адмін панелі системи управління контентом DataLife Engine.



Рисунок 2.

На Рисунку 2. зображена панель управління системою RoboNews. У цій панелі можна докладно налаштувати роботу системи, визначити правила перекладу, визначити групи з автоматичним перекладом та багато інших налаштувань, які покращують робоздатність системи. Тобто RoboNews має можливість гнучкого налаштування усіх параметрів, відповідно від поставленого адміністратором завдання. Всі ці налаштування і відрізняють систему RoboNews від «ручного» перекладу публікацій. Без них систему можна було б віднести до простого прикладу використання можливостей будь якого сервісу машинного перекладу, але це не так. Система RoboNews складається з модулю машинного перекладу і системи правил, за допомогою яких і визначається напрямок перекладу тексту, необхідність у перекладі публікації, унікалізація тексту, встановлення правил для використання можливостей системи окремими групами користувачів сайту, встановлення правил для перекладу окремих частин матеріалу (заголовка, короткого і повного описів новини) та ряду інших налаштувань та додатків. Саме поєднання цих компонентів і відрізняють RoboNews від інших систем перекладу, і наразі їй немає аналогів. Новизна системи полягає в тому, що було поєднано та використано дві системи (машинний переклад і гнучкий ряд можливих для встановлення правил), що і дозволило досягнути поставлених цілей та значно полегшити і пришвидшити роботу у перекладі великої кількості текстових матеріалів. На даний момент система RoboNews працює уже більше року та доказала свою ефективність та надійність.

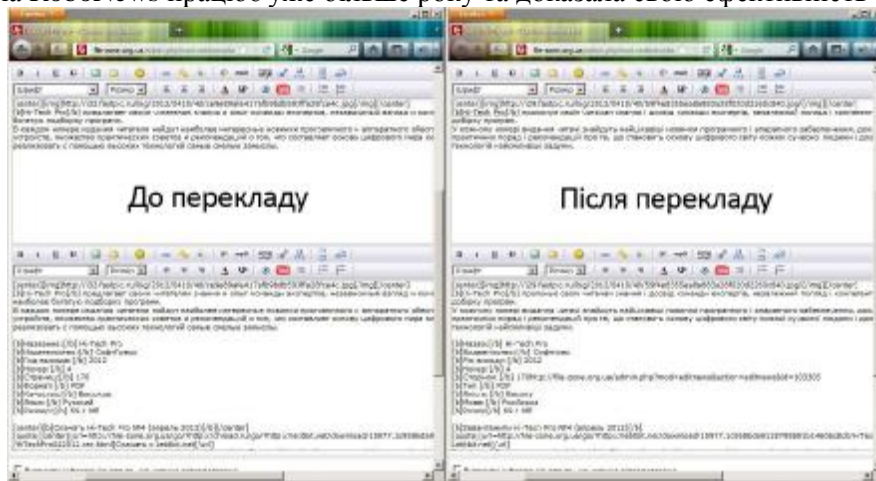


Рисунок 3.

На Рисунку 3. можна побачити приклад роботи системи, а саме оригінал матеріалу, та матеріал після обробки системою RoboNews. Як можна побачити на рисунку форматування матеріалу було збережено, а сам текст публікації був перекладений на іншу мову (з Російської на Українську). Ці дії були виконані автоматично, на основі вже встановлених правил у панелі управління системою RoboNews.

Отже, я можу зробити висновок, що розробка додатків та систем що використовують безкоштовні можливості мережі Інтернет є досить ефективною, та приносить очікувані результати полегшуючи та покращуючи роботу багатьох систем. Прикладом цього є система RoboNews, яка є унікальною у своєму роді і немає аналогів.

1. <http://dle-news.ru/>
2. <http://www.microsoft.com/web/post/using-the-free-bing-translation-apis>
3. <https://developers.google.com/translate/>