

ПОДДЕРЖКА И УПРАВЛЕНИЕ ПРОЦЕССАМИ ЛОКАЛИЗАЦИИ

Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь

Поддубицкий А.Ф.

Ворвиль А.А. - магистр техн. наук

Одной из задач, с которыми часто сталкиваются организации, деятельность которых не ограничена территорией одного культурно-лингвистического сообщества, является локализация документов. Объем средств, затрачиваемых на решение проблемы перевода зависит от целевых и исходных языков, размера текстов, подлежащих переводу и многих других факторов. Одной из проблем, эффективное решение которых могло бы в значительной мере сократить расходы на локализацию является эффективная интеграция ручного и машинного перевода в единый процесс локализации.

На сегодняшний день в мире существует ряд решений, призванных облегчить работу переводчика:

- системы типа «память перевода» [1], доступные для установки на ПК [2];
- сервисы, доступные по протоколу HTTP, доступные как через веб-клиенты, предоставляемые производителем, так и через программный интерфейс (прим. – «Google Translate»);
- словари, обеспечивающие перевод отдельных слов и словосочетаний в рамках одной или нескольких языковых пар [2];
- встраиваемые в популярные офисные приложения модули для ускорения перевода, использующие ресурсы вышеописанных систем, и др. виды инструментов.

Все эти средства помогают частично или полностью автоматизировать процесс перевода, но качество перевода часто оказывается на неприемлемом уровне. Высокое качество полностью автоматического перевода доступно лишь для формализованных текстов, написанных в техническом или официально-деловом стиле. Ручной перевод даёт необходимый уровень качества, но является значительно более медленным и дорогим решением, требующим участия хорошо подготовленных специалистов.

На данный момент существуют системы, позволяющие организовать процесс локализации с участием человека и средств машинного перевода [1, 2]. Переводчик-человек выполняет вспомогательную роль, участвуя в процессе с целью повышения качества результата. С точки зрения способа интеграции ручного и машинного перевода существующие системы делятся на следующие типы:

- с постредактированием: на первом этапе текст на исходном языке переводится специальной программой, далее для улучшения результата текст на целевом языке, полученный машиной, просматривает и правит переводчик-человек.
- с предредактированием: зная особенности работы программы-переводчика, человек готовит текст к программной обработке. Этот процесс может включать разметку, упрощение текста и перевод заведомо сложных для программы-переводчика кусков. Далее следует перевод с помощью программы.
- с интерредактированием: человек вмешивается в работу системы перевода при необходимости с целью разрешения трудных задач.
- смешанные системы: например, с пред- и постредактированием. Могут включать и все три первых способа.

Существующие системы, однако, не предусматривают эффективного мониторинга процесса, поддержки локализации для нескольких языковых пар, большого числа документов, не предоставляют средств работы в команде при активном использовании средства машинного перевода.

Целью работы является создание системы управления процессами локализации, предоставляющей следующие возможности:

- построение сложных, гибких и повторно используемых процессов локализации с применением машинных и ручных операций;
- выполнение и контроль созданных шаблонов процессов локализации;
- эффективная работа команды переводчиков и средств машинного перевода в рамках одного процесса;
- интеграция со сторонними системами машинного перевода [3];
- возможность оперирования результатами отдельных шагов процесса.

Система имеет архитектуру клиент-сервер. Сервисы разработаны с помощью Windows Communication Foundation 4. Основными технологиями разработки серверной части выбраны .NET, Microsoft Workflow Foundation 4 [4], IIS. Также система имеет программный интерфейс, построенный по принципам REST [3], предоставляющий возможность интеграции с любым внешним ПО.

Список использованных источников:

1. XLIFF version 1.2. OASIS Standard / OASIS Xml Localization Interchange File Format (XLIFF) TC – OASIS, Xliff Technical Committee – 2007 – 71 p.
2. Machine Translation. An Introductory Guide / D. Arnold – USA, Cambridge, Blackwell Publishers, – 1994 – 206 p.
3. Architectural Styles and the Design of Network-based Software Architectures / R.T. Fielding – University of California, Irvine, – 2000 – 180 p.
4. Pro WF 4.5 / B. White – Microsoft, Apress, – 2013 – 638 p.