

Министерство образования Республики Беларусь  
Учреждение образования  
«Белорусский государственный университет  
информатики и радиоэлектроники»

УДК 004.453

БУЛИЦКАЯ  
Ольга Евгеньевна

**АВТОМАТИЗИРОВАННЫЙ ПРОГРАММНЫЙ КОМПЛЕКС  
ПЕРЕВОДА ДОКУМЕНТОВ КОМПАНИИ**

АВТОРЕФЕРАТ  
магистерской диссертации на соискание степени  
магистра технических наук

по специальности 1-40 80 02 «Системный анализ, управление и обработка  
информации (по отраслям)»

Научный руководитель  
канд. техн. наук, доцент  
Севернёв А.М.

Минск 2017

Работа выполнена на кафедре информационных технологий автоматизированных систем учреждения образования «Белорусский государственный университет информатики и радиоэлектроники»

Научный руководитель: **Севернёв Александр Михайлович**,  
кандидат технических наук, доцент  
доцент кафедры ИТАС УО «БГУИР»

Рецензент: **Шмаков Михаил Сергеевич**,  
кандидат технических наук, доцент  
зав. кафедрой полиграфического  
оборудования и систем обработки  
информации УО «БГТУ»

Защита диссертации состоится «25» января 2016 г. года в часов на заседании Государственной комиссии по защите магистерских диссертаций в учреждении образования «Белорусский государственный университет информатики и радиоэлектроники» по адресу: 220013, г.Минск, ул. П.Бровки, 6, 4-й уч. корп., ауд.423, тел.: 293-88-23, e-mail: [kafitas@bsuir.by](mailto:kafitas@bsuir.by).

С диссертацией можно ознакомиться в библиотеке учреждения образования «Белорусский государственный университет информатики и радиоэлектроники».

## **ВВЕДЕНИЕ**

Одной из основополагающих задач организаций, деятельность которых не ограничена территорией одного государства, является перевод различных документов с одного языка на другие. В зависимости от объёма переводимых текстов, языка оригинала и целевого языка перевода (языка, на который осуществляется перевод) на перевод могут потребоваться время и значительные средства.

Существует ряд возможностей для того, чтобы перевод был достаточно качественным, быстрым и недорогим. Традиционный способ перевода текста предполагает обращение к сторонним организациям, либо наём работников соответствующей профессии. Любой из данных вариантов предполагает перевод одним человеком или группой людей. Такой способ решения задачи имеет свои недостатки и преимущества.

Главным преимуществом ручного перевода является его качество. На сегодняшний день доступные способы автоматизированного перевода текста могут обеспечить относительно корректную передачу содержания текста на целевом языке.

К автоматизированным способам перевода можно отнести программы со встроенными словарями для определённых языков. Такие программы обычно рассчитаны на одного пользователя и являются настольными приложениями.

Ручной способ перевода также обладает существенными недостатками. Ручной перевод гораздо дороже машинного, особенно, если организация пользуется услугами переводчиков регулярно или же содержит собственный штат. В случаях, когда переводится текст большого объёма, перевод осуществляется группой переводчиков, и при оценке трудоёмкости, а значит и затрат на решение поставленной задачи, следует учитывать также коммуникации между членами команды переводчиков. Если команда достаточно большая, то требуется руководящий персонал. Следовательно, появляются проблемы разделения задач, проверки результатов по различным критериям, ответственности за разные этапы работ или элементы текста.

## **ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ**

### **Актуальность темы исследования**

Актуальность темы исследования магистерской диссертации обусловлена тем, что на многих предприятиях необходимо осуществлять перевод документации с языка оригинала на различные целевые языки. При этом предприятия хотят получить качественный перевод документов, сэкономив на затратах.

Для решения данной проблемы необходимо объединить ручной и машинный переводы в целях сохранения качества и увеличения производительности. Кроме того, нужно постоянно работать над отказоустойчивостью

системы и повышать её производительность.

### **Степень разработанности проблемы**

На данный момент существует большое количество систем для автоматизации перевода. Но большинство из них имеют различные недостатки.

Зачастую эти недостатки связаны со структурой конкретного языка. Кроме того, все языки имеют слова, значения которых меняется в зависимости от контекста. В каждом языке существуют устойчивые выражения, не распознаваемые существующими системами.

В связи с этим необходимо объединять машинный перевод документа с его ручным переводом. Но сделать это надо так, чтобы машина переводила, а человек выступал только в качестве проверяющего.

Так как разрабатываемая в диссертации система будет основана на такой технологии как Windows WF, необходимо изучить, как именно система запоминает момент остановки рабочего процесса перевода. И в случае, если предоставленный механизм нас не устраивает, его необходимо улучшить.

### **Цель и задачи исследования**

Изучение существующих механизмов технологии Windows WF компании Microsoft, разработанной для создания и выполнения потоков работ, с целью улучшения отказоустойчивости разрабатываемой системы, а также повышения скорости работы рабочих процессов (в данном случае – процессов перевода).

**Объект исследования диссертации** – материальные потоки, локализация.

**Предмет исследования** – процесс сохранения момента остановки рабочих процессов технологии Windows WF.

**Цель данной работы:** исследование процесса сохранения момента остановки рабочих процессов с помощью технологии Windows WF и дальнейшее улучшение этого процесса.

Для достижения цели необходимо решить следующие **задачи**:

- а) обобщить теоретические основы работы SQL;
- б) проанализировать временные затраты на выполнение SQL запросов, используя стандартный метод работы Windows WF с моментом остановки процесса;
- в) изучить современные способы сжатия данных;
- г) изучить способы преобразования данных в XML и их последующее сжатие;
- д) применить теоретические разработки на практике.

В диссертации используются такие **методы исследования** как анализ, классификация, сравнение, аналогия, моделирование.

### **Область исследования**

Содержание диссертационной работы соответствует образовательному стандарту высшего образования второй ступени (магистратуры) специальности 1-53 80 01 – «Автоматизация и управление технологическими процессами и производствами (по отраслям)».

### **Теоретическая и методологическая основа исследования**

В основу диссертации легли материалы по работе с Windows WF (на языке-оригинале), а также официальный сайт компании Microsoft – MSDN.

Для получения результатов, связанных с производительностью SQL-запросов, использовался Microsoft SQL Server, а также теоретические сведения, взятые с официальных источников.

Информация по работе с библиотекой для сжатия данных, а также информация об используемых алгоритмах, была получены из Википедии.

**Информационная база** исследования для анализа сформированы на основе данных полученных из базы данных, используемой на производстве.

**Научная новизна** диссертационной работы заключается в разработке модели, помогающей ускорить процесс перевода, увеличить отказоустойчивость, а также объединить машинный и ручной способы перевода.

#### **Основные положения, выносимые на защиту**

Разработанный в рамках магистерской работы способ сохранения рабочего процесса позволяет ускорить процесс перевода и повысить показатель отказоустойчивости.

**Теоретическая значимость** диссертации заключается в том, что в ней описан процесс улучшения работы стандартной схемы сохранения момента остановки рабочего процесса.

**Практическая значимость** диссертации состоит в том, что на основе предложенной модели была увеличена производительность работы рабочих процессов, а также увеличена отказоустойчивость.

#### **Апробация и внедрение результатов исследования**

Результаты исследования были представлены на Международная научно-практическая интернет-конференция молодых ученых и студентов «Актуальные проблемы автоматизации и управления».

Результаты магистерской диссертации внедрены в рабочий процесс компании, занимающейся переводами.

#### **Публикации**

Основные положения работы и результаты диссертации изложены в опубликованной работе общим объемом 5,0 п.л. (авторский объем – 5,0 п.л.).

**Структура и объем работы.** Структура диссертационной работы обусловлена целью, задачами и логикой исследования. Работа состоит из введения, шести глав, заключения и двух приложений. Общий объем диссертации – 59 страниц. Работа содержит 34 рисунка.

## ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** рассмотрены: текущее состояние проблемы переводов документов компаний; существующие способы переводов; основные достоинства и недостатки машинного и ручного переводов.

В **первой главе** рассматриваются общие теоретические сведения о машинном переводе. Проведен обзор основных существующих средств автоматизации перевода, выявлены их преимущества и недостатки. На его основе сформулированы основные требования к разрабатываемому программному комплексу автоматизированного перевода.

Во **второй главе** приводятся основные сведения о Windows Workflow Foundation. Также рассматриваются различия между последовательными процессами и процессами, основанными на состояниях.

В **третьей главе** приводятся примеры того, в каких системах необходимо использовать возможность сохранения момента остановки и каковы причины использования этой возможности.

В **четвёртой главе** рассматривается стандартный Persistence Store, разработанный компанией Microsoft, приводятся основные недостатки работы с ним.

В **пятой главе** описана разработка собственного способа хранения момента остановки рабочего. Приведены примеры алгоритмов, которые используются библиотекой сжатия данных Gzip. Показано, как результаты внедрения собственного процесса улучшили работу приложения.

В **шестой главе** приведена методика использования разработанного программного комплекса на практике.

В **заключении** подведены итоги работы.

В **приложениях** приведены примеры сериализованного рабочего процесса, а также исходный текст программного модуля.

## ЗАКЛЮЧЕНИЕ

1. Проведен анализ существующих систем автоматизации процесса перевода. Рассмотрены плюсы и минусы машинного и ручного способов перевода.

2. Рассмотрен и проанализирован стандартный механизм работы технологии Windows WF для сохранения момента остановки рабочего процесса в базе данных.

3. Разработан собственный механизм сохранения момента остановки рабочего процесса в базе данных.

4. Внедрён собственный способ сохранения момента остановки рабочего процесса.

#### **Список опубликованных работ**

1. Булицкая, О.Е. / Автоматизированный программный комплекс перевода документов компании / О.Е. Булицкая //Международная научно-практическая интернет-конференция молодых ученых и студентов «Актуальные проблемы автоматизации и управления», Украина, 25 октября 2016 г. – С. 188.

Библиотека БГУИР