

**ИСПОЛЬЗОВАНИЕ МЕТОДОВ МАШИННОГО ПЕРЕВОДА ПО ПРАВИЛАМ  
ДЛЯ СОЗДАНИЯ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ  
ПРЕОБРАЗОВАНИЯ УЗКОСПЕЦИАЛИЗИРОВАННОГО МЕДИЦИНСКОГО  
КОНТЕНТА В ПАЦИЕНТО-ОРИЕНТИРОВАННЫХ ТЕЛЕМЕДИЦИНСКИХ  
РЕШЕНИЯХ**

**Е.А. Машина (Университет ИТМО)  
Научный руководитель – к.т.н., П.В. Балакшин  
(Университет ИТМО)**

В работе описываются подходы к созданию системы перевода узкоспециализированного медицинского контента, предназначенного для предоставления информации пользователям, не обладающим необходимой медицинской квалификацией, для дальнейшего использования в пациенто-ориентированных телемедицинских системах, создаваемых для информационной поддержки непосредственно пациенту и его близким, основанные на технологии машинного перевода Example-based Machine Translation.

**Цель работы** – обоснование выбора набора оптимальных механизмов приведения узкоспециализированных медицинских текстов к виду, понятному пациентам, не обладающим врачебной квалификацией.

Одним из наиболее эффективных областей применения систем искусственного интеллекта на сегодняшний день являются телемедицинские технологии, постоянно повышающие процентное соотношение специализированных действий, осуществляющихся без непосредственного участия медицинского персонала. Очередным этапом развития телемедицины является создание пациенто-ориентированных систем, оказывающих справочно-информационную поддержку непосредственно пациенту и его близким [1].

При создании пациенто-ориентированных телемедицинских систем приходится решать большое количество существенных проблем, одной из которых является адаптация узкоспециализированного медицинского контента для неквалифицированного пользователя.

В связи с этим важнейшим направлением работ является реализация технологии перевода узкоспециализированного медицинского контента к виду, доступному для понимания пользователем, не обладающим медицинской квалификацией [2].

При выборе механизмов для построения указанной двунаправленной системы перевода следует обратить внимание на то, что перевод со «специализированного медицинского» на «повседневный» будет проводиться внутри одного языка.

Учитывая высокую степень риска при неправильном трактовании узкоспециализированных медицинских терминов и понятий на язык естественной речи, основными требованиями при создании рассматриваемого переводчика являются точность и однозначность формулировок.

Выбирая общую технологию автоматической адаптации профессионального медицинского контента для нужд непрофессионального пользователя, следует обратить внимание на то, что такие узкоспециализированные тексты хорошо структурированы. А используемые в них термины и описания действий большей частью стандартизованы в виде специализированных протоколов [2].

В связи с этим за основу технологии построения машинного перевода (с медицинского на каждыйдневный) взять Машинный Перевод Основанный на Примерах — Example-based Machine Translation (ЕВМТ) [3], расширенный возможностями технологии является технологии Translation Memory (ТМ), повышающей скорость перевода за счет уменьшения объема проводимой работы при переводе повторяющихся фрагментов текстов.

Результативность технологии ЕВМТ для адаптации узкоспециализированных текстов обуславливается тем, что она с одной стороны не требует обширных баз данных «зеркальных

текстов», необходимых для обучения, с другой стороны, не требует делать и глубокий лингвистический анализ, необходимый для создания систем перевода, основанных на правилах [4].

При использовании ЕВМТ предложения переводимого (адаптированного) текста разбиваются на определённые фразы, затем происходит перевод этих простейших фраз с учетом уже накопленного ранее опыта перевода. А после этого из этих фраз-фрагментов строится окончательный перевод предложения [5]. При этом технология ЕВМТ позволяет не проводить предварительное выстраивание всеобъемлющей системы правил и исключений, необходимых для построения фраз. Вместо этого в ЕВМТ при осуществлении перевода происходит использование фрагментов текстов, имеющих аналоги на языке, на который требуется перевод. Это позволяет говорить о ее универсальности для различных отраслей знания [6].

Так как в основе узкоспециальной медицинской терминологии лежит хорошо структурированный и грамматически изученный латинский язык, задача построения двунаправленного переводчика, создаваемого для пациенто-ориентированных телемедицинских систем, существенно облегчается. Поэтому применение описанных в работе механизмов для реализации двунаправленного переводчика в составе пациенто-ориентированного телемедицинского решения позволит неквалифицированному пользователю быстро и однозначно понимать смысл предоставляемых ему специализированных медицинских данных и самостоятельно генерировать запросы на предоставление специализированной информации без использования узкоспециализированных медицинских терминов.

#### **Литература:**

1. Машина Е.А. Особенности создания пациенто-ориентированных телемедицинских систем // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19 апреля 2019г.) - 2019. - Т. 3. - С. 256-259.
2. Машина Е.А., Нелепко Л.Н. Структура и источники специализированных данных системы информационной поддержки муниципальных волонтеров, оказывающих доврачебную медицинскую помощь // Сборник трудов VIII конгресса молодых ученых (Санкт-Петербург, 15-19 апреля 2019г.) - 2019. - Т. 3. - С. 267-270.
3. Somers, H. Example-based Machine Translation // Machine Translation, 1999. № 14, p. 113—157.
4. Нелюбин, Л.Л. Введение в технику перевода (когнитивный теоретико-прагматичный аспект) //Л.Л. Нелюбин – 5-е изд., стер. – М.: ФЛИНТА, 2016. – 216 с.
5. Kit C., Pan H., Webster J. J. Example-based Machine Translation: A New Paradigm // Translation and Information Technology / S.W. Chan (ed.) -Hong Kong: Chinese University of Hong Kong, 2002. p. 57-78.
6. Computers and Translation: A Translator's Guide. In: Somers, H. (ed.) John Benjamins Publ. Company, Amsterdam, 2003, p. 351.