

ОСНОВНЫЕ НАПРАВЛЕНИЯ РЕАЛИЗАЦИИ СИСТЕМЫ ПЕРЕВОДА УЗКОСПЕЦИАЛИЗИРОВАННОГО МЕДИЦИНСКОГО КОНТЕНТА В ПАЦИЕНТО- ОРИЕНТИРОВАННЫХ ТЕЛЕМЕДИЦИНСКИХ СИСТЕМАХ

Е.А. Машина (Университет ИТМО)

Научный руководитель – к.т.н., доцент П.В. Балакшин
(Университет ИТМО)

В работе описываются основные направления реализации 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, а так же описываются необходимые для этого этапы предварительной обработки специализированных медицинских данных.

Цель работы – обоснование выбора оптимальных механизмов приведения узкоспециализированных медицинских текстов к виду, понятному пациентам, не обладающим врачебной квалификацией для дальнейшего использования в качестве информационной поддержки муниципальных медицинских волонтеров.

Основной особенностью реализуемой совместной мультидисциплинарной рабочей группой студентов Университета ИТМО, МГУ и СПбГУ (PMH.Team) пациенто-центричной системы PersonalMedHelper является то, что она создается для оказания информационной поддержки путем предоставления непрофессиональному пользователю массивов данных, построенных на основе автоматизированной обработки узкоспециализированного медицинского контента [1]. В связи с этим важнейшим направлением работ, проводимых на данном этапе, является реализация технологии перевода узкоспециализированного медицинского контента к виду, доступному для понимания по понятийному ряду пользователям, не обладающим медицинской квалификацией [2].

Создание такого механизма позволит решить следующие задачи:

- данные, предназначенные для неквалифицированных пользователей, будут генерироваться на уже существующей базе узкоспециализированной медицинской информации, разрешая при этом проблемы, связанные с достоверностью исходной информации,

- преобразование автоматически обрабатываемой специализированной медицинской информации в вид, доступный неквалифицированному пользователю, позволит исключить процедуру дополнительной подготовки пациенто-ориентированных массивов данных, что приведет к ускорению подготовки систем к использованию и снижению затрат на предварительную подготовку контента,

- наличие создаваемого решения для перевода позволит проводить постоянную актуализацию информации, доступной неквалифицированному пациенту, одновременно с обновлением и расширением подобной узкоспециализированной медицинской информации.

При выборе механизмов для построения указанной двунаправленной системы перевода следует обратить внимание на то, что перевод со «специализированного медицинского» на «каждодневный» 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.

В результате проведенного на ранних этапах работ сравнительного анализа наиболее обоснованной методикой построения системы перевода узкоспециализированного медицинского контента признана технология перевода текста, основанная на правилах: Rule-based Machine Translation (RBMT) [3].

Наиболее кратким обоснованием этого является тот факт, что RBMT-системы имеют более высокое качество перевода, чем системы, основанные на анализе статистики (Statistical Machine Translation, SMT-системы) [4].

В качестве основного принципа реализации RBMT-системы является построение связи структур исходного и конечного текстов. При этом наиболее значимыми преимуществами RBMT-систем являются грамматическая и синтаксическая точность, и высокая стабильность получаемого результата. Важно отметить, что метод дополнительно предоставляет широкие возможности настройки на рассматриваемую предметную область.

В связи с этим, именно RBMT-решение является наиболее предпочтительным решением в случаях, когда однозначность и качество перевода имеют первостепенное значение.

В целом, выбранное в качестве технологической основы решение представляет собой метод построения перевода, основанный на словарной информации и анализе грамматических правил. Данная словарная информация, в свою очередь, базируется на лингвистическом описании структур узкоспециализированных профессиональных текстов [5]. При этом качество перевода будет находиться в прямой зависимости от подробности описания понятий и объемов предварительно обработанных лингвистических баз данных [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.

В дальнейшем применение описанных в работе механизмов для реализации двунаправленного переводчика в составе решения PersonalMedHelper 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, 15-19апреля 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, 15-19апреля 2019г.) - 2019. - Т. 3. - С. 267-270.
3. Нелюбин, Л.Л. Введение в технику перевода. Флинта, М. (2009)
4. Bowker, L. Computer-Aided Translation Technology: A Practical Introduction: Univ.of Ottawa Press, Ottawa (2002).
5. Сдобников В.В., Петрова О.В. Теория перевода. М.: АСТ Восток-Запад (2006).
6. Computers and Translation: A Translator's Guide. In: Somers, H. (ed.) John Benjamins Publ. Company, Amsterdam (2003).