

DOI: <https://doi.org/10.17816/DD321420>

Технология распознавания речи в лучевой диагностике

Н.Д. Кудрявцев¹, К.А. Бардасова², А.Н. Хоружая¹¹ Научно-практический клинический центр диагностики и телемедицинских технологий, Москва, Российская Федерация² Уральский государственный медицинский университет, Екатеринбург, Российская Федерация

АННОТАЦИЯ

Устройства, способные распознавать речь, являются перспективным инструментом для системы здравоохранения. Технология распознавания речи имеет довольно длинную историю применения в западных системах здравоохранения (с 1970-х годов), однако широкое распространение она получила лишь в начале XXI века, заменив медицинских транскрипционистов. Для отечественного здравоохранения данная технология относительно новая. Её активная разработка началась лишь в начале 2010-х годов, а повсеместное внедрение в здравоохранение — в конце 2010-х годов. Такая задержка связана с особенностями русского языка и ограничением вычислительных мощностей, присутствующих в начале XXI века.

В настоящее время комплексы устройств и программного обеспечения для распознавания речи используются в голосовом заполнении медицинской документации и позволяют сократить время подготовки протоколов рентгенологических исследований при сравнении с традиционным (клавиатурным) вводом текста.

В литературном обзоре отражена краткая история развития и применения технологии распознавания речи в лучевой диагностике. Отражены ключевые научные исследования, подтверждающие эффективность её использования в западных системах здравоохранения. Продемонстрирован отечественный опыт применения технологии распознавания речи и оценена её эффективность. Описаны перспективы дальнейшего развития данной технологии в российском здравоохранении.

Ключевые слова: научный обзор; технология распознавания речи; голосовой ввод; лучевая диагностика; рентгенология; заполнение медицинской документации.

Как цитировать

Кудрявцев Н.Д., Бардасова К.А., Хоружая А.Н. Технология распознавания речи в лучевой диагностике // *Digital Diagnostics*. 2023. Т. 4, № 2. С. 185–196.DOI: <https://doi.org/10.17816/DD321420>

DOI: <https://doi.org/10.17816/DD321420>

Speech recognition technology in radiology

Nikita D. Kudryavtsev¹, Kristina A. Bardasova², Anna N. Khoruzhaya¹

¹ Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

² Ural State Medical University, Ekaterinburg, Russian Federation

ABSTRACT

Speech recognition devices are promising tools for the healthcare system. Speech recognition technology has had a relatively long history of use in Western healthcare systems since the 1970s. However, it became widely used at the beginning of the 21st century, replacing medical transcriptionists. This technology is relatively new in home healthcare. Its active development began only in the early 2010s, and its implementation in healthcare started in late 2010. This delay is due to the idiosyncrasies of the Russian language and the limited computational power present at the beginning of the 21st century.

Currently, complexes of devices and software for speech recognition are used in the voice filling of medical records and can reduce the time for preparing reports for radiological examinations compared with traditional (keyboard) text input.

The literature review provides a brief history of speech recognition technology development and application in radiography. Key scientific studies showing its efficacy in Western healthcare systems are reflected. Voice recognition technology in the home is demonstrated, and its effectiveness is evaluated. The prospects for further development of this technology in Russian healthcare are described.

Keywords: medical records; radiation diagnostics; radiology; speech recognition software; voice input.

To cite this article

Kudryavtsev ND, Bardasova KA, Khoruzhaya AN. Speech recognition technology in radiology. *Digital Diagnostics*. 2023;4(2):185–196.

DOI: <https://doi.org/10.17816/DD321420>

Received: 17.03.2023

Accepted: 19.04.2023

Published: 25.05.2023

DOI: <https://doi.org/10.17816/DD321420>

语音识别技术在放射诊断中的应用

Nikita D. Kudryavtsev¹, Kristina A. Bardasova², Anna N. Khoruzhaya¹

¹ Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

² Ural State Medical University, Ekaterinburg, Russian Federation

简评

能够进行语音识别的设备是保健系统的一个有前途的工具。语音识别技术在西方医疗系统中有相当长的使用历史（自20世纪70年代以来），但它在21世纪初才得到了广泛推广，取代了医疗抄写员。对于国内的医疗保健来说，该技术是相对较新的。它的积极开发是在2010年代初才开始，并2010年代末才在保健事业广泛采用的。这种延迟是由于俄语的特点和21世纪初计算能力的限制而导致的。

语音识别的设备和软件包现在被用于通过语言输入填写病历，此外，与传统（用键盘）文本输入相比，减少了准备X射线学协议所需的时间。

本文综述简要介绍了语音识别技术在放射诊断中的发展和应用的历史。介绍了证实其在西方医疗系统中使用的有效性的主要科学研究。展示了国内使用语音识别技术的经验，并对其有效性进行了评估。描述了该技术在俄罗斯保健事业进一步发展的前景。

关键词：科学综述；语音识别技术；语音输入；放射诊断；X射线学；填写病历。

To cite this article

Kudryavtsev ND, Bardasova KA, Khoruzhaya AN. 语音识别技术在放射诊断中的应用. *Digital Diagnostics*. 2023;4(2):185–196.

DOI: <https://doi.org/10.17816/DD321420>

收到: 17.03.2023

接受: 19.04.2023

发布日期: 25.05.2023

ВВЕДЕНИЕ

В настоящее время голосовое управление превратилось в стандартную функцию для многих бытовых «умных» устройств. Это стало возможным благодаря развитию технологии распознавания речи. Системы, построенные на её основе, анализируют речь пользователя и трансформируют её в цифровые данные. Помимо управления «умными» устройствами, технология распознавания речи получила широкое распространение в телефонной связи. Сейчас при звонках во многие государственные и коммерческие организации пользователя встречает автоответчик, который распознаёт голосовой запрос звонящего и маршрутизирует его на выбранного специалиста. В Москве в 2019 году был запущен проект по уведомлению граждан о записи к врачу и напоминании о диспансерном наблюдении с помощью голосового помощника. Во время подобного звонка гражданин мог записаться к медицинскому специалисту, отменить или перенести свою консультацию, также система опрашивала пациента о наличии жалоб¹.

В области здравоохранения системы распознавания речи получили широкое распространение при голосовом заполнении медицинской документации. Это связано с тем, что большую часть своего рабочего времени врачи тратят на подготовку медицинской документации [1–4]. Данный фактор отрицательно сказывается на качестве оказания медицинской помощи, особенно в условиях ограниченного времени приёма пациента. Перспектива применения технологии, например, в лучевой диагностике связана с сокращением длительности заполнения протоколов диагностических исследований и возможностью уделять больше времени на изучение диагностических изображений, сопроводительной медицинской документации и общение с пациентами. Именно поэтому наибольшую популярность системы голосового ввода получили в рентгенологических отделениях, поскольку организация рабочего процесса в них наиболее удобна для внедрения подобной технологии. Современные систематические обзоры [5–7] продемонстрировали, что применение систем распознавания речи в данных условиях эффективно, а хорошая восприимчивость объясняется большими объёмами текстовой информации, которую врачи-рентгенологи обязаны заносить в протоколы.

ИСТОРИЯ ПРИМЕНЕНИЯ ТЕХНОЛОГИИ РАСПОЗНАВАНИЯ РЕЧИ В РЕНТГЕНОЛОГИИ

Ранние годы применения

Первые попытки медицинского применения технологии распознавания речи были предприняты в 70–80-х

годах прошлого столетия. В 1975 году был описан прототип системы распознавания речи, которая могла проанализировать крайне ограниченный объём медицинской лексики и терминологии [8]. В отделении лучевой диагностики технология была впервые апробирована в 1981 году [9]. В израильской больнице Beth Israel Hospital её начали использовать для подготовки протоколов рентгенологических исследований. Изначально врачи этого учреждения применяли систему CLIP (Coded Language Information Processing) — иерархический стандартизированный язык медицинских терминов, разработанный М. Simon и В.В. Leeming [10]. Структура этого языка содержала медицинские термины, закодированные определённым образом. Например, значению А соответствовала *anatomic area* (анатомическая область), А6 — *the lower limb* (нижняя конечность), А61 — *the thigh* (бедро), в свою очередь значениями В, В6, В61, В611 и В612 были закодированы *bone* (кость), *the lower limb bones* (кости нижней конечности), *the femur* (бедренная кость), *the lesser trochanter* (малый вертел) и *the greater trochanter* (большой вертел) соответственно. Набирая различные комбинации кодов с помощью клавиатуры, врач-рентгенолог мог составить протокол любой сложности и объёма [11].

Системы распознавания речи тех времён имели главный недостаток — ограниченный словарь (около 200 уникальных медицинских терминов), что было связано с малым объёмом оперативной памяти. Этот фактор не позволял в полном объёме описывать диагностические исследования. Одним из решений, позволяющих обойти ограничение объёма рабочей памяти, стало использование технологии распознавания речи совместно с системой CLIP.

В исследовании 1981 года было проведено сравнение клавиатурного и голосового вводов [9]. Оценивались скорость и качество подготовки 60 протоколов радиологических исследований. В результате оказалось, что объём протокола не зависел от способа заполнения, что косвенно подтверждало и схожую сложность описываемых исследований. Длительность же заполнения протокола с помощью голосового ввода в 4 раза превышала длительность клавиатурного ввода. Протоколы, заполненные голосом, содержали в среднем 12 ошибок распознавания, в то время как протоколы, подготовленные с помощью клавиатуры, не имели таковых. В ходе этого исследования были описаны также дополнительные ограничения технологии. Первые системы не могли полностью подавить внешние шумы, поэтому качество распознавания врачебной речи было низким, что способствовало появлению ошибок. Увеличение длительности заполнения медицинского документа было связано с тем, что система могла распознать только отдельно произнесённые слова. Врачу приходилось делать паузу между словами,

¹ Официальный сайт Мэра Москвы [интернет]. Голосовой помощник соберёт жалобы на самочувствие пациента по телефону перед приемом в поликлинике. Режим доступа: <https://www.mos.ru/news/item/89302073/>.

а в случае с системой CLIP — между кодами. Такой метод голосового заполнения был неудобным и неестественным для человека. Стоит также отметить, что система голосового ввода требовала от 3 до 6 часов предварительной настройки и адаптации под конкретную речь пользователя (врача).

Все эти ограничения не позволили системам распознавания речи тех времён получить распространение в здравоохранении. По этой причине привлечение медицинских транскрипционистов на тот период времени было оправданным. Тем не менее появился тот базис, который позволил учесть «слабые места» технологии на последующих этапах развития.

Возрастающая доступность рентгенологических исследований, появление компьютерных и магнитно-резонансных томографов, переход от аналоговых носителей к цифровым привели к повышению рабочей нагрузки на врачей-рентгенологов и увеличению длительности подготовки заключений. Для решения этой проблемы и оптимизации кадровых ресурсов в середине 80-х годов в западных странах были организованы диктофонные центры. Врачи-рентгенологи диктовали на звукозаписывающие устройства описания рентгенологических находок, выявленных во время интерпретации диагностических изображений. Аудиозаписи передавались в диктофонный центр, где медицинские транскрипционисты расшифровывали аудиозаписи и оформляли протоколы исследований в текстовом виде. Подготовленные протоколы проверял и визирует врач. В некоторых случаях они возвращались транскрипционистам для исправления ошибок [12]. Диктофонные центры были популярными в зарубежных медицинских учреждениях вплоть до 2010-х годов [5, 13].

Ряд авторов провели сравнение эффективности подготовки протоколов рентгенологических исследований с помощью медицинских транскрипционистов и технологии распознавания речи. Были сделаны выводы, что основное преимущество медицинского транскрипциониста заключается в возможности заметить грамматические ошибки и учесть контекстную информацию о пациенте. Эти преимущества позволяют правильно понять и согласовать аудиозапись врача, даже если она плохого качества [14]. Однако в связи с дефицитом медицинских транскрипционистов зачастую возникали ситуации, когда протокол в печатном виде возвращался врачу только через 16 часов после диктовки [15]. В зависимости от мощности диктофонного центра для расшифровки аудиозаписей требовалось от 6 до 24 часов.

Дальнейшее развитие систем распознавания речи со временем вытеснило медицинских транскрипционистов [14]. Несмотря на то что технологии распознавания речи уступают в точности распознавания медицинским транскрипционистам, они позволяют сократить

ежемесячные экономические расходы отделения лучевой диагностики на 81% и значительно сократить время подготовки протоколов исследований [5, 13, 16].

В отечественной системе здравоохранения также предпринимались попытки внедрения диктофонных центров², однако в открытом доступе отсутствуют научные исследования, оценивающие эффективность их применения. В связи с отсутствием подобных центров в современной отечественной системе здравоохранения можно предположить также, что их применение оказалось нецелесообразным.

1990-е: новый виток развития инновации

В конце XX века в системах распознавания речи был увеличен объём памяти и словарей (до 19 000 рентгенологических терминов), сокращено время предварительной настройки до нескольких минут и повышена точность распознавания. В 1995 году в США были представлены первые программы расшифровки естественного языка, позволяющие определять слитную английскую речь. Теперь врач мог диктовать в удобном разговорном темпе, не делая пауз между словами. На данном этапе эти программы уступали в точности распознавания системам с отдельным вводом [17]. Дальнейшее развитие технологий и повышение точности распознавания речи позволило системам слитного голосового ввода заменить системы с отдельным вводом.

XXI век

Широкомасштабное внедрение и применение автоматического распознавания речи в отделениях лучевой диагностики началось в западных странах в начале 2000-х годов. В зарубежных исследованиях сравнили скорость голосового заполнения англоязычной медицинской документации с клавиатурным вводом, объёмы медицинского документа и удовлетворённость врачей. Полученные данные показали, что использование технологии ведёт к повышению скорости заполнения документации на 26% и увеличению объёма протоколов. Голосовой ввод позволил также оптимизировать рабочий процесс, сократив время, необходимое для подготовки медицинской документации, улучшив качество протоколов по содержанию. У врачей повысилась удовлетворённость от работы с документацией [16, 18]. Частота ошибок также сократилась, преимущественно встречались пунктуационные ошибки [7]. Вышеперечисленные факторы привели к сокращению общего времени подготовки протоколов исследований с 16 до 5 часов [15]. Результаты исследований, оценивающие затрату времени на подготовку протокола, среднее количество символов в минуту, количество и частоту ошибок, отражают позитивную динамику в развитии

² Официальный сайт ГБУЗ «ИКБ № 1 ДЗМ» [интернет]. История больницы. Режим доступа: <https://ikb1.ru/about/>.

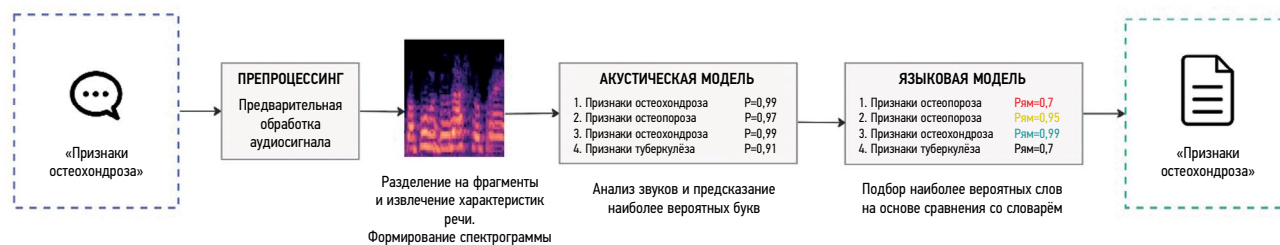


Рис. 1. Упрощённая схема работы классической системы распознавания речи. Представлен алгоритм распознавания фразы «Признаки остеохондроза».

технологии и внедрении её в отделениях лучевой диагностики. Процент готовых протоколов в течение одного часа увеличился с 26 до 58%, протоколы стали более структурированы по содержанию [19].

Экономические затраты с течением времени тоже сократились. Так, в 5 из 7 систематических обзоров [5], оценивающих затраты, сообщалось об уменьшении, в 2 — об увеличении экономических затрат. Похожие результаты были продемонстрированы в эндокринологии и психиатрии, где использование систем распознавания речи позволило повысить продуктивность и эффективность работы врачей [20]. Применение этих систем в хирургии помогло сократить время подготовки протоколов операций с 4 до 3 дней, а количество подготовленных протоколов в течение одного дня повысилось с 22 до 37% [21]. Авторы зарубежных исследований 2019–2020-х годов делают вывод, что технология распознавания речи экономит врачам время, повышает их эффективность и позволяет документировать более важные детали при заполнении медицинских бумаг [22–24], однако основным препятствием для внедрения систем голосового ввода среди врачей можно считать человеческий фактор — их сопротивление переменам и страх перед новыми технологиями [25].

В 2016 году в исследовании Microsoft research было продемонстрировано, что точность систем распознавания речи достигла уровня человеческих возможностей и составила 94% [26]. Сейчас эта технология прочно вошла в медицинскую практику в англоязычных странах, а уровень внедрения голосового ввода в отделениях лучевой диагностики достиг 85%³. На сегодняшний день доля рынка, занимаемая подобными программами для нужд здравоохранения, составляет около 25% от общемирового⁴. Лидерами в разработке программного обеспечения

для распознавания речи являются компании Nuance Communications, IBM и Philips⁵.

Распознавание речи приобрело всеобщую значимость в медицине в англоязычных странах приблизительно за 45 лет. Оно охватило все уровни здравоохранения — от первичного звена и неотложной помощи до высоко-специализированных отделений. Современные медицинские системы распознавания речи для английского языка обладают точностью до 99%, способны адаптироваться под разные акценты и не требуют предварительного обучения голосовому профилю врача⁶.

СОВРЕМЕННАЯ СИСТЕМА РАСПОЗНАВАНИЯ РЕЧИ

Как было сказано ранее, технология распознавания речи строится на процессе перевода устной речи человека в текст с использованием компьютера. В современных системах голосового ввода применяются алгоритмы искусственного интеллекта, которые позволяют значительно повысить качество и скорость распознавания речи пользователя [27]. Процесс распознавания состоит из нескольких этапов, каждый из которых имеет свои особенности (рис. 1).

1. *Получение аудиосигнала.* Первый этап распознавания речи — это запись звуковой волны. Она может быть выполнена с помощью микрофона, встроенного в любое устройство записи аудио, например в смартфон. Затем звуковая волна преобразуется в цифровой формат, чтобы его мог обработать компьютер.
2. *Предварительная обработка аудиосигнала.* На этом этапе полученные звуковые данные проходят предварительную обработку, позволяющую устранить посторонние фоновые шумы и выделить речь пользователя. Это повышает качество распознавания речи.

³ Reaction Data [интернет]. Speech Rec in Radiology-State of the Market. 2019 [cited 2019 Dec 23]. Режим доступа: <https://www.reactiondata.com/report/speech-recognition-in-radiology-state-of-the-market/>.

⁴ Grand View Research [интернет]. Voice And Speech Recognition Market Size Report, 2030. Режим доступа: <https://www.grandviewresearch.com/industry-analysis/voice-recognition-market>.

⁵ Nuance Communications. Healthcare Clinical Documentation AI Solutions & Services for the NHS (<https://www.nuance.com/en-gb/healthcare.html>); Philips. Healthcare--Philips (<https://www.dictation.philips.com/gb/industries/industry/healthcare-professionals/>); IBM. Watson Speech to Text (<https://www.ibm.com/cloud/watson-speech-to-text>).

⁶ Nuance Communications [интернет]. Dragon Medical One--#1 Clinical Documentation Companion. Режим доступа: <https://www.nuance.com/healthcare/provider-solutions/speech-recognition/dragon-medical-one.html>.

3. *Разделение на фрагменты.* Затем аудиофайл разбивается на небольшие фрагменты длиной около 10–25 миллисекунд. Это необходимо для оптимизации анализа звуковых данных. Каждый фрагмент анализируется системой распознавания речи отдельно.
4. *Извлечение характеристик речи.* На этом этапе вычисляются характеристики речи, такие как частота, продолжительность и амплитуда звуков. Они используются для определения фонем, которые составляют произнесённые слова.
5. *Сопоставление с моделями распознавания.* Сопоставление фонем происходит с помощью акустических моделей, которые обучаются на большом количестве образцов речи. Эти модели могут использовать различные методы, включая скрытые модели Маркова, нейронные сети и другие алгоритмы машинного обучения [28–30].
6. *Составление слов и текста.* На этом этапе языковая модель объединяет фонемы в слова и фразы, а затем в полный текст. Этот текст может потребовать дополнительной обработки для исправления орфографических ошибок и других неточностей.
7. *Вставка сформированного текста.* Это финальный этап. Сформированный текст вставляется в медицинский документ. Учитывая то, что системы распознавания речи ещё не достигли 100% точности, в некоторых случаях врачу приходится вручную вносить коррективы в распознанный текст.

Чтобы система распознала звуки вне зависимости от пола, возраста и интонации диктующего, преобразовала их в буквы с большей точностью, акустические и языковые модели используют модули искусственного интеллекта. Разработчики обучают нейронные сети на наборе данных. При этом набор данных включает в себя большое количество разнообразных аудиозаписей и примеров текста. Когда поступает голосовой сигнал, нейронная сеть «ищет» соответствие в базе данных. В процессе использования нейронная сеть продолжает своё обучение и создаёт новые комбинации пар «звук-буква», что позволяет с большей вероятностью воспроизводить задуманный текст без орфографических ошибок. В процессе обучения компьютер распознаёт наиболее важные признаки произношения фонем и записывает полученные данные в виде профиля пользователя⁷ [31].

В последнее время получили широкое распространение так называемые end-to-end-подходы к построению систем распознавания речи. End-to-end-подходы в технологии распознавания речи представляют собой методы, в которых вся обработка речевого сигнала выполняется автоматически, без необходимости выполнения отдельных этапов обработки, таких как извлечение признаков и создание модели распознавания речи [32].

Одним из наиболее популярных end-to-end-подходов является глубокое обучение (deep learning). В этом случае нейронная сеть обучается напрямую на сырых аудиоданных, без необходимости выполнения предварительной обработки [32]. Другим примером end-to-end-подхода является использование рекуррентных нейронных сетей (recurrent neural network, RNN) или свёрточных нейронных сетей (convolutional neural networks, CNN) для распознавания речи. В этом случае нейронная сеть обучается на входном сигнале и выходном тексте, используя технику обучения с учителем [32].

Преимущество end-to-end-подходов заключается в том, что они могут обеспечить более высокую точность распознавания, поскольку вся информация о речи используется при обучении нейронной сети. Однако эти подходы могут быть более сложными в реализации и требовать большего количества данных для обучения [33].

ОТЕЧЕСТВЕННЫЙ ОПЫТ ПРИМЕНЕНИЯ ТЕХНОЛОГИИ РАСПОЗНАВАНИЯ РЕЧИ В ЛУЧЕВОЙ ДИАГНОСТИКЕ

Первые системы распознавания речи для русского языка появились в середине 2000-х годов [34, 35], но применение общеразговорного словаря не позволяло использовать их в медицинской практике. Потребовалось ещё несколько лет для разработки систем голосового ввода, позволяющих расшифровывать русскую речь с медицинскими терминами [36]. Активное развитие технологии распознавания речи пришлось лишь на вторую половину 2010-х годов. Такая задержка была связана со сложностями распознавания русской речи.

Русский язык имеет более сложную структуру словообразования по сравнению с английским, так как является синтетическим языком с наличием большого количества словоформ. Чтобы распознавать слова, необходимо использовать словари большего объёма. Это замедляет работоспособность системы [37]. Например, современные системы распознавания речи для английского языка используют словарь, содержащий до 300 тыс. слов и терминов, а для русского языка словарь может содержать более 5 млн слов, словоформ и словосочетаний [33, 38, 39]. Кроме того, большинство словоформ одного и того же слова отличаются только окончаниями, которые часто произносятся пользователями нечётко. Это приводит к ошибке распознавания всей фразы и необходимости корректировки финального документа. В том числе в русском языке имеется больше вариантов расположения слов в предложении в отличие от английского языка, где используются строгие грамматические конструкции. Это затрудняет создание языковых моделей системы распознавания речи и понижает точность их работы.

⁷ Cloud.mts.ru [интернет]. Технология распознавания речи и ее значение для бизнеса. Режим доступа: <https://cloud.mts.ru/cloud-thinking/blog/tehnologiya-raspoznavaniya-rechi/>.

На сегодняшний день в России ведущей компанией по разработке систем голосового ввода для здравоохранения является группа компаний «ЦРТ»⁸. Первое исследование эффективности применения технологии распознавания речи в отделениях лучевой диагностики было проведено в 2020 году на базе семи городских поликлиник Департамента здравоохранения города Москвы с использованием системы голосового ввода «Voice2Med» (группа компаний «ЦРТ») ранней версии с точностью распознавания 93%. В исследовании проводилось сравнение скорости заполнения медицинской документации с помощью клавиатурного ввода и системы распознавания речи. Врачи-рентгенологи заполняли протоколы компьютерных и магнитно-резонансных томографических исследований. Хронометражное исследование продемонстрировало, что средняя длительность, необходимая для описания одного исследования с помощью клавиатурного ввода, составила 10 минут 15 секунд, а при использовании системы распознавания речи — 8 минут 2 секунды. На современном этапе развития точность распознавания медицинской терминологии на русском языке достигла 98%. Это стало возможным благодаря формированию словаря медицинских терминов, разработанного на основе 2,5 млн протоколов рентгенологических исследований, и анализу обратной связи от врачей-рентгенологов⁹ (рис. 2).

Опрос врачей-рентгенологов, проведённый в 2022 году, продемонстрировал, что 62,8% респондентов отмечают повышение эффективности своей работы при использовании системы распознавания речи. Большинство врачей, использующих голосовой ввод в повседневной работе, оценивают качество распознавания

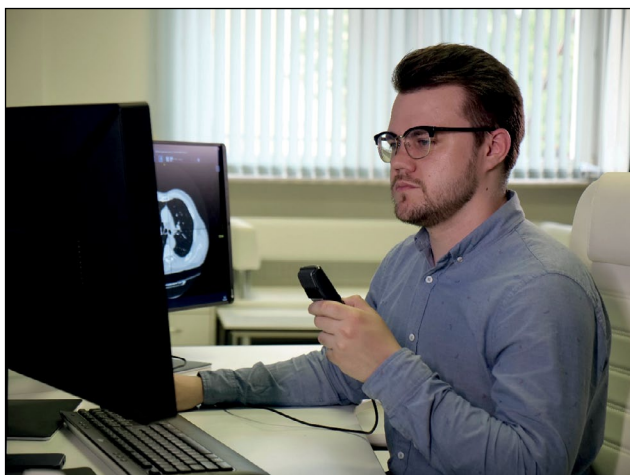


Рис. 2. Рабочее место врача-рентгенолога в Московском референс-центре лучевой диагностики, оснащённое системой распознавания речи. Процесс заполнения медицинской документации.

рентгенологических терминов как хорошее и отличное. Респонденты отмечают, что возникали ситуации, когда распознавалась чужая речь, случалось некорректное распознавание окончаний слов. Выявлено также, что на качество распознавания может негативно влиять посторонний фоновый шум (работа диагностического оборудования, общение медицинского персонала с пациентом или с коллегами) и некачественные устройства для записи звука. Немаловажным фактором, влияющим на приверженность использованию новой технологии, стал возраст врачей и их заинтересованность инновацией. Молодые специалисты более открыты к технологии, также более склонными к применению систем голосового ввода в работе оказались специалисты в возрасте 30–40 лет. Результаты проведённых опросов показали положительную динамику в отношении врачей-рентгенологов к технологии распознавания речи в течение 2 лет с момента начала её применения [40].

ПЕРСПЕКТИВЫ РАЗВИТИЯ ТЕХНОЛОГИИ РАСПОЗНАВАНИЯ РЕЧИ

Дальнейшее улучшение точности распознавания позволит ещё больше сократить время подготовки электронных медицинских документов. Одной из основных задач, стоящих перед разработчиками систем распознавания речи, является обеспечение высокой точности анализа речи в сложных акустических условиях, когда на записи присутствуют множество шумов или голоса посторонних людей. Кроме того, как уже упоминалось, из-за особенностей русского языка одна из сложностей, с которой сталкиваются пользователи, — это распознавание окончаний слов. Поэтому для систем распознавания речи русского языка важнейшей задачей является разработка языковой модели, способной с высокой точностью предсказывать и согласовывать слова в предложениях.

Интеграция программ голосового ввода с медицинскими информационными системами позволит осуществлять дистанционное заполнение структурированных электронных медицинских документов. Развиваясь, система способна не просто распознавать фразы врача, но и понимать, в каком разделе медицинского документа распознанный текст должен быть размещён. Реализация такого функционала позволит врачам ультразвуковой диагностики, патоморфологам, эндоскопистам, хирургам заполнять медицинские документы непосредственно во время выполнения медицинской манипуляции, а не постфактум, что значительным образом скажется на качестве документов и скорости их подготовки.

⁸ Группа компаний ЦРТ [интернет]. Синтез и распознавание речи, запись и анализ, идентификация лица и голоса. Режим доступа: <http://www.speechpro.ru/>.

⁹ Группа компаний ЦРТ [интернет]. Voice2Med: Программа для голосового заполнения медицинской документации. Режим доступа: <https://www.speechpro.ru/product/programmy-dlya-raspoznaniya-rechi-v-tekst/voice2med>.

Технология распознавания речи имеет также большой потенциал в стандартизации и унификации лексики, используемой при подготовке медицинских документов, в том числе рентгенологических протоколов. На сегодняшний день не существует единого универсального списка терминов для описания одного и того же патологического состояния в рентгенологии [41]. Даже два разных врача-рентгенолога, работающих в одном отделении, могут использовать при подготовке протоколов исследования различные синонимы, описывающие одну и ту же патологическую находку. В ряде работ было отмечено, что использование структурированных и стандартизированных протоколов с унифицированной терминологией упрощает восприятие и получение нужной информации как другим врачам-рентгенологам, так и врачам других специальностей [42–44].

На сегодняшний день предпринята попытка разработать тезаурус, который позволил бы стандартизировать описание патологических изменений, выявленных при компьютерных томографических исследованиях. Тезаурус содержит 120 русскоязычных рентгенологических терминов и примеров их описания [45]. Однако разработка тезауруса — это сложная задача, требующая согласования терминологии большим количеством специалистов и рентгенологическим сообществом.

ЗАКЛЮЧЕНИЕ

В литературном обзоре представлена краткая историческая справка о развитии технологии распознавания речи в отделениях лучевой диагностики, подробно описана её эволюция, а также оценены достоинства и недостатки инновации на основе литературных данных. Особое внимание уделено применению технологии распознавания речи в российских отделениях лучевой диагностики. В некоторых работах продемонстрировано значительное улучшение точности распознавания медицинской терминологии на русском языке. В перспективе её применение способно сокращать длительность подготовки медицинской документации и уделять больше времени общению с пациентом и изучению его истории болезни, что открывает новые возможности для развития персонализированной медицины. Тем не менее всё ещё сохраняются ошибки распознавания окончаний и согласования слов в предложении, что требует от врачей дополнительного времени на их исправление. В дальнейшем эти проблемы будут решены за счёт применения новых алгоритмов искусственного интеллекта.

Результаты некоторых исследований демонстрируют позитивный настрой врачей-рентгенологов по отношению к системам распознавания речи, что проявляется в более частом их использовании в работе. Технология, однозначно, должна продолжать своё развитие в здравоохранении России, так как пилотные отечественные и устоявшиеся зарубежные примеры её внедрения свидетельствуют

о позитивной динамике применения. Дальнейшее повышение точности распознавания медицинских терминов позволит найти ещё больше сторонников технологии распознавания речи среди медицинских специалистов.

ДОПОЛНИТЕЛЬНО

Источник финансирования. Данная статья подготовлена авторским коллективом в рамках НИР «Научно-методические основы цифровой трансформации службы лучевой диагностики» (№ ЕГИСУ: № 123031400118-0) в соответствии с Приказом от 21.12.2022 г. № 1196 «Об утверждении государственных заданий, финансовое обеспечение которых осуществляется за счет средств бюджета города Москвы государственным бюджетным (автономным) учреждениям, подведомственным Департаменту здравоохранения города Москвы, на 2023 год и плановый период 2024 и 2025 годов» Департамента здравоохранения города Москвы.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Вклад авторов. Все авторы подтверждают соответствие своего авторства международным критериям ICMJE (все авторы внесли существенный вклад в разработку концепции, проведение поисково-аналитической работы и подготовку статьи, прочли и одобрили финальную версию перед публикацией). Наибольший вклад распределён следующим образом: Н.Д. Кудрявцев — концепция научной публикации, поиск и анализ литературных данных, написание текста рукописи; К.А. Бардасова — поиск и анализ научных данных, сбор и подготовка материалов; А.Н. Хоружая — написание и редактирование текста рукописи.

ADDITIONAL INFORMATION

Funding source. This article was prepared by a group of authors as a part of the research and development effort titled “Theoretical and methodological framework for digital transformation in radiology” (USIS No. 123031400118-0) in accordance with the Order No. 1196 dated December 21, 2022 “On approval of state assignments funded by means of allocations from the budget of the city of Moscow to the state budgetary (autonomous) institutions subordinate to the Moscow Health Care Department, for 2023 and the planned period of 2024 and 2025” issued by the Moscow Health Care Department.

Competing interests. The authors declare that they have no competing interests.

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the version to be published and agree to be accountable for all aspects of the work. N. Kudryavtsev — concept and design of the paper, text writing, approval of the final version of the paper; K. Bardasova — data collection and analysis; A. Khoruzhaya — data interpretation, text editing.

СПИСОК ЛИТЕРАТУРЫ

1. Вечорко В.И. Распределение рабочего времени на амбулаторном приеме врача-терапевта участкового с медицинской сестрой в поликлинике города Москвы (фотохронометражное наблюдение) // Социальные аспекты здоровья населения. 2016. № 6. С. 4.
2. Каплиева О.В., Марега Л.А., Воробьева Л.П. Хронометраж рабочего времени врачей детского консультативно-диагностического отделения // Дальневосточный медицинский журнал. 2018. № 4. С. 72–76.
3. Ryabchikov I.V., Zagafarov R.R., Mukhina V.V., et al. Distribution of the traumatologist-orthopaedician's working time with outpatients // *Moscow Sur J.* 2018. N 6. P. 38–43. doi: 10.17238/issn2072-3180.2018.6.38-43
4. Kudryavtsev N.D., Sergunova K.A., Ivanova G.V., et al. Evaluation of the effectiveness of the implementation of speech recognition technology for the preparation of radiological protocols // *VIT.* 2020. Vol. 6, N S1. P. 58–64. doi: 10.37690/1811-0193-2020-S1-58-64
5. Blackley S.V., Huynh J., Wang L., et al. Speech recognition for clinical documentation from 1990 to 2018: A systematic review // *J Am Med Inf Association.* 2019. Vol. 26, N 4. P. 324–338. doi: 10.1093/jamia/ocy179
6. Motyer R.E., Liddy S., Torreggiani W.C., Buckley O. Frequency and analysis of non-clinical errors made in radiology reports using the National Integrated Medical Imaging System voice recognition dictation software // *Ir J Med Sci.* 2016. Vol. 185, N 4. P. 921–927. doi: 10.1007/s11845-016-1507-6
7. Hodgson T., Coiera E. Risks and benefits of speech recognition for clinical documentation: A systematic review // *J Am Med Inf Association.* 2016. Vol. 23, N e1. P. e169–e179. doi: 10.1093/jamia/ocv152
8. Itakura F. Minimum prediction residual principle applied to speech recognition // *IEEE Trans Acoust Speech Signal Process.* 1975. Vol. 23, N 1. P. 67–72. doi: 10.1109/TASSP.1975.1162641
9. Leeming W., Porter D., Jackson J.D., et al. Computerized radiologic reporting with voice data-entry // *Radiology.* 1981. Vol. 138, N 3. P. 585–588. doi: 10.1148/radiology.138.3.7465833
10. Simon M., Leeming B.W., Bleich H.L., et al. Computerized radiology reporting using coded language // *Radiology.* 1974. Vol. 113, N 2. P. 343–349. doi: 10.1148/113.2.343
11. Vogel M., Kaisers W., Wassmuth R., Mayatepek E. Analysis of documentation speed using web-based medical speech recognition technology: Randomized controlled trial // *J Med Internet Res.* 2015. Vol. 17, N 11. P. e247. doi: 10.2196/jmir.5072
12. Ramaswamy M.R., Chaljub G., Esch O., et al. Continuous speech recognition in MR imaging reporting // *Am J Roentgenol.* 2000. Vol. 174, N 3. P. 617–622. doi: 10.2214/ajr.174.3.1740617
13. Poder T.G., Fisette J.F., Déry V. Speech recognition for medical dictation: Overview in quebec and systematic review // *J Med Systems.* 2018. Vol. 42, N 5. P. 89. doi: 10.1007/s10916-018-0947-0
14. Sankaranarayanan B., David G., Vishwanath K.R., et al. Would technology obliterate medical transcription? // *Proceedings of the 2017 ACM SIGMIS Conference on Computers and People Research.* New York, NY, USA: ACM, 2017. P. 97–104. doi: 10.1145/3084381.3084414
15. Houston J.D., Rupp F.W. Experience with implementation of a radiology speech recognition system // *J Digital Imaging.* 2000. Vol. 13, N 3. P. 124–128. doi: 10.1007/BF03168385
16. Saxena K., Diamond R., Conant R.F., et al. Provider adoption of speech recognition and its impact on satisfaction, documentation quality, efficiency, and cost in an inpatient EHR // *AMIA Jt Summits Transl Sci Proc.* 2018. Vol. 2017. P. 186–195.
17. Schwartz L.H., Kijewski P., Hertogen H., et al. Voice recognition in radiology reporting // *Am J Roentgenol.* 1997. Vol. 169, N 1. P. 27–29. doi: 10.2214/ajr.169.1.9207496
18. Vogel M., Kaisers W., Wassmuth R., Mayatepek E. Analysis of documentation speed using web-based medical speech recognition technology: Randomized controlled trial // *J Med Int Research.* 2015. Vol. 17, N 11. P. e247. doi: 10.2196/jmir.5072
19. Hammana I., Lepanto L., Poder T., et al. Speech recognition in the radiology department: A systematic review // *Health Inf Manag.* 2015. Vol. 44, N 2. P. 4–10. doi: 10.1177/183335831504400201
20. Mohr D.N., Turner D.W., Pond G.R., et al. Speech recognition as a transcription aid: A randomized comparison with standard transcription // *J Am Med Inf Association.* 2003. Vol. 10, N 1. P. 85–93. doi: 10.1197/jamia.m1130
21. Singh M., Pal T.R. Voice recognition technology implementation in surgical pathology: Advantages and limitations // *Arch Pathol Laboratory Med.* 2011. Vol. 135, N 11. P. 1476–1481. doi: 10.5858/arpa.2010-0714-OA
22. Goss F.R., Blackley S.V., Ortega C.A., et al. A clinician survey of using speech recognition for clinical documentation in the electronic health record // *Int J Med Inf.* 2019. N 130. P. 103938. doi: 10.1016/j.ijmedinf.2019.07.017
23. Blackley S.V., Schubert V.D., Goss F.R., et al. Physician use of speech recognition versus typing in clinical documentation: A controlled observational study // *Int J Med Inform.* 2020. N 141. P. 104178. doi: 10.1016/j.ijmedinf.2020.104178
24. Yang L., Ene I.C., Belaghi R.A., et al. Stakeholders' perspectives on the future of artificial intelligence in radiology: A scoping review // *Eur Radiol.* 2022. Vol. 32, N 3. P. 1477–1495. doi: 10.1007/s00330-021-08214-z
25. European Society of Radiology (ESR). Impact of artificial intelligence on radiology: A EuroAIM survey among members of the European Society of Radiology // *Insights Imaging.* 2019. Vol. 10, N 1. P. 105. doi: 10.1186/s13244-019-0798-3
26. Szymański P., Żelasko P., Morzyet M., et al. WER we are and WER we think we are // *arXiv.* arXiv:2010.03432.2020. doi: 10.48550/arXiv.2010.03432
27. Li J. Recent advances in end-to-end automatic speech recognition // *arXiv.* arXiv:2111.01690.2022. doi: 10.48550/arXiv.2111.01690
28. Juang B.H., Rabiner L.R. Hidden markov models for speech recognition // *Technometrics.* 1991. Vol. 33, N 3. P. 251–272.
29. Graves A., Mohamed A., Hinton G. Speech recognition with deep recurrent neural networks // *2013 IEEE International Conference on Acoustics, Speech and Signal Processing.* Vancouver, BC, Canada: IEEE, 2013. P. 6645–6649. doi: 10.48550/arXiv.1303.5778
30. Deng L., Li X. Machine learning paradigms for speech recognition: An overview // *IEEE Trans Audio Speech Lang Process.* 2013. Vol. 21, N 5. P. 1060–1089.
31. Казачкин А.Е. Методы распознавания речи, современные речевые технологии // *Молодой ученый.* 2019. № 39. С. 6–8.
32. Kamath U., Liu J., Whitaker J. Deep learning for NLP and speech recognition. Cham: Springer International Publishing, 2019. 621 p.

33. Wang D., Wang X., Lv S. An overview of end-to-end automatic speech recognition // *Symmetry*. 2019. Vol. 11, N 8. P. 1018. doi: 10.3390/sym11081018
34. Zhzhikashvili V.A., Farkhadov M.P., Petukhova N.V., Zhzhikashvili A.V. The first voice recognition applications in Russian language for use in The Interactive Information Systems // *Speech and Computer*. Saint-Petersburg, SPECOM, 2004. P. 304–307.
35. Карпов А.А., Ронжин А.А., Ли И.В. SIRIUS система дикторне-зависимого распознавания слитной русской речи // *Известия Южного федерального университета. Технические науки*. 2005. Т. 54, № 10. С. 44–54.
36. Ирзаев М.Г. Использование голосового ввода информации в медицинских учреждениях для заполнения электронных карт и историй болезней пациентов // *Новые технологии и техника в медицине, биологии и экологии: сборник научных трудов*. 2013. № 3. С. 149–154.
37. Vazhenina D., Markov K., Karpov A., et al. State-of-the-art speech recognition technologies for Russian language // *Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments*. Aizu-Wakamatsu Japan: ACM, 2012. P. 59–63. doi: 10.1145/2160749.2160763
38. Kamvar M., Chelba C. Optimal size, freshness and time-frame for voice search vocabulary // *arXiv*. arXiv:1210.8436.2012.
39. Kipyatkova I.S., Karpov A.A. An analytical survey of large vocabulary Russian speech recognition

- systems // *SPIRAS Proceedings*. 2014. Vol. 1, N 12. P. 7. doi: 10.15622/sp.12.1
40. Kudryavtsev N.D., Semenov D.S., Kozhikhina D.D., Vladzimirskyy A.V. Speech recognition technology: Results of a survey of radiologists at the Moscow reference center for diagnostic radiology // *Healthcare Management*. 2022. Vol. 8, N 3. P. 95–104. doi: 10.33029/2411-8621-2022-8-3-95-104
41. Синицын В.Е., Комарова М.А., Мершина Е.А. Протокол рентгенологического описания: прошлое, настоящее, будущее // *Вестник рентгенологии и радиологии*. 2014. № 3. С. 35–40.
42. Sobez L.M., Kim S.H., Angstwurm M., et al. Creating high-quality radiology reports in foreign languages through multilingual structured reporting // *Eur Radiol*. 2019. Vol. 29, N 11. P. 6038–6048. doi: 10.1007/s00330-019-06206-8
43. Ganeshan D., Duong P.A., Probyn L., et al. Structured reporting in radiology // *Academic Radiology*. 2018. Vol. 25, N 1. P. 66–73. doi: 10.1016/j.acra.2017.08.005
44. Dos Santos P.D., Hempel J.M., Mildenerger P., et al. Structured reporting in clinical routine // *Rofo*. 2019. Vol. 191, N 01. P. 33–39. doi: 10.1055/a-0636-3851
45. Андрианова М.Г., Кудрявцев Н.Д., Петряйкин А.В. Разработка тезауруса рентгенологических терминов для голосового заполнения протоколов диагностических исследований // *Digital Diagnostics*. 2022. Т. 3, № S1. С. 21–22. doi: 10.17816/DD105703

REFERENCES

1. Vechorko V.I. Distribution of working time at an outpatient appointment of a district therapist with a nurse in a polyclinic in Moscow (photochromometric observation). *Social Aspects Public Health*. 2016;(6):4. (In Russ).
2. Kaplieva O.V., Marega L.A., Vorobyeva L.P. Timekeeping of working hours of doctors of the children's consultative and diagnostic department. *Far Eastern Med J*. 2018;(4):72–76. (In Russ).
3. Ryabchikov I.V., Zagafarov R.R., Mukhina V.V., et al. Distribution of the traumatologist-orthopaedician's working time with outpatients. *Moscow Sur J*. 2018;(6):38–43. (In Russ). doi: 10.17238/issn2072-3180.2018.6.38-43
4. Kudryavtsev N.D., Sergunova K.A., Ivanova G.V., et al. Evaluation of the effectiveness of the implementation of speech recognition technology for the preparation of radiological protocols. *VIT*. 2020;6(S1):58–64. (In Russ). doi: 10.37690/1811-0193-2020-S1-58-64
5. Blackley S.V., Huynh J., Wang L., et al. Speech recognition for clinical documentation from 1990 to 2018: A systematic review. *J Am Med Inf Association*. 2019;26(4):324–338. doi: 10.1093/jamia/ocy179
6. Motyer R.E., Liddy S., Torreggiani W.C., Buckley O. Frequency and analysis of non-clinical errors made in radiology reports using the National Integrated Medical Imaging System voice recognition dictation software. *Ir J Med Sci*. 2016;185(4):921–927. doi: 10.1007/s11845-016-1507-6
7. Hodgson T., Coiera E. Risks and benefits of speech recognition for clinical documentation: A systematic review. *J Am Med Inf Association*. 2016;23(e1):e169–e179. doi: 10.1093/jamia/ocv152
8. Itakura F. Minimum prediction residual principle applied to speech recognition. *IEEE Trans Acoust Speech Signal Process*. 1975;23(1):67–72. doi: 10.1109/TASSP.1975.1162641
9. Leeming W., Porter D., Jackson J.D., et al. Computerized radiologic reporting with voice data-entry. *Radiology*. 1981;138(3):585–588. doi: 10.1148/radiology.138.3.7465833
10. Simon M., Leeming B.W., Bleich H.L., et al. Computerized radiology reporting using coded language. *Radiology*. 1974;113(2):343–349. doi: 10.1148/113.2.343
11. Vogel M., Kaisers W., Wassmuth R., Mayatepek E. Analysis of documentation speed using web-based medical speech recognition technology: Randomized controlled trial. *J Med Internet Res*. 2015;17(11):e247. doi: 10.2196/jmir.5072
12. Ramaswamy M.R., Chaljub G., Esch O., et al. Continuous speech recognition in MR imaging reporting. *Am J Roentgenol*. 2000;174(3):617–622. doi: 10.2214/ajr.174.3.1740617
13. Poder T.G., Fiset J.F., Déry V. Speech recognition for medical dictation: Overview in quebec and systematic review. *J Med Systems*. 2018;42(5):89. doi: 10.1007/s10916-018-0947-0
14. Sankaranarayanan B., David G., Vishwanath K.R., et al. Would technology obliterate medical transcription? In: *Proceedings of the 2017 ACM SIGMIS Conference on Computers and People Research*. New York, NY, USA: ACM; 2017. P. 97–104. doi: 10.1145/3084381.3084414
15. Houston J.D., Rupp F.W. Experience with implementation of a radiology speech recognition system. *J Digital Imaging*. 2000;13(3):124–128. doi: 10.1007/BF03168385
16. Saxena K., Diamond R., Conant R.F., et al. Provider adoption of speech recognition and its impact on satisfaction, documentation quality, efficiency, and cost in an inpatient EHR. *AMIA Jt Summits Transl Sci Proc*. 2018;2017:186–195.
17. Schwartz L.H., Kijewski P., Hertogen H., et al. Voice recognition in radiology reporting. *Am J Roentgenol*. 1997;169(1):27–29. doi: 10.2214/ajr.169.1.9207496
18. Vogel M., Kaisers W., Wassmuth R., Mayatepek E. Analysis of documentation speed using web-based medical speech recognition technology: Randomized controlled trial. *J Med Int Research*. 2015;17(11):e247. doi: 10.2196/jmir.5072

19. Hammana I, Lepanto L, Poder T, et al. Speech recognition in the radiology department: A systematic review. *Health Inf Manag.* 2015;44(2):4–10. doi: 10.1177/183335831504400201
20. Mohr DN, Turner DW, Pond GR, et al. speech recognition as a transcription aid: A randomized comparison with standard transcription. *J Am Med Inf Association.* 2003;10(1):85–93. doi: 10.1197/jamia.m1130
21. Singh M, Pal TR. Voice recognition technology implementation in surgical pathology: Advantages and limitations. *Arch Pathol Laboratory Med.* 2011;135(11):1476–1481. doi: 10.5858/arpa.2010-0714-OA
22. Goss FR, Blackley SV, Ortega CA, et al. A clinician survey of using speech recognition for clinical documentation in the electronic health record. *Int J Med Inf.* 2019;(130):103938. doi: 10.1016/j.ijmedinf.2019.07.017
23. Blackley SV, Schubert VD, Goss FR, et al. Physician use of speech recognition versus typing in clinical documentation: A controlled observational study. *Int J Med Inform.* 2020;(141):104178. doi: 10.1016/j.ijmedinf.2020.104178
24. Yang L, Ene IC, Belaghi RA, et al. Stakeholders' perspectives on the future of artificial intelligence in radiology: A scoping review. *Eur Radiol.* 2022;32(3):1477–1495. doi: 10.1007/s00330-021-08214-z
25. European Society of Radiology (ESR). Impact of artificial intelligence on radiology: A EuroAIM survey among members of the European Society of Radiology. *Insights Imaging.* 2019;10(1):105. doi: 10.1186/s13244-019-0798-3
26. Szymański P, Żelasko P, Morzyet M, et al. WER we are and WER we think we are. *arXiv.* arXiv:2010.03432.2020. doi: 10.48550/arXiv.2010.03432
27. Li J. Recent advances in end-to-end automatic speech recognition. *arXiv.* arXiv:2111.01690.2022. doi: 10.48550/arXiv.2111.01690
28. Juang BH, Rabiner LR. Hidden markov models for speech recognition. *Technometrics.* 1991;33(3):251–272.
29. Graves A, Mohamed A, Hinton G. Speech recognition with deep recurrent neural networks In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, BC, Canada: IEEE; 2013. P. 6645–6649. doi: 10.48550/arXiv.1303.5778
30. Deng L, Li X. Machine learning paradigms for speech recognition: An overview. *IEEE Trans Audio Speech Lang Process.* 2013;21(5):1060–1089.
31. Kazachkin AE. Speech recognition methods, modern speech technologies. *Young Scientist.* 2019;(39):6–8. (In Russ).
32. Kamath U, Liu J, Whitaker J. Deep learning for NLP and speech recognition. Cham: Springer International Publishing; 2019. 621 p.
33. Wang D, Wang X, Lv S. An overview of end-to-end automatic speech recognition. *Symmetry.* 2019;11(8):1018. doi: 10.3390/sym11081018
34. Zhzhikashvili VA, Farkhadov MP, Petukhova NV, Zhzhikashvili AV. The first voice recognition applications in Russian language for use in The Interactive Information Systems. In: Speech and Computer. Saint-Petersburg, SPECOM; 2004. P. 304–307. (In Russ).
35. Karpov AA, Ronzhin AA, Li IV. SIRIUS system of dictoron-independent recognition of the merged Russian speech. *Izvestia Southern Federal University. Technical Sci.* 2005;54(10):44–54. (In Russ).
36. Irzaev MG. The use of voice input of information in medical institutions to fill in electronic charts and patient medical histories. *New technologies and techniques in medicine, biology and ecology: Collection of scientific papers.* 2013;(3):149–154. (In Russ).
37. Vazhenina D, Markov K, Karpov A, et al. State-of-the-art speech recognition technologies for Russian language. In: Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments. Aizu-Wakamatsu Japan: ACM; 2012. P. 59–63. doi: 10.1145/2160749.2160763
38. Kamvar M, Chelba C. Optimal size, freshness and time-frame for voice search vocabulary. *arXiv.* arXiv:1210.8436.2012.
39. Kipyatkova IS, Karpov AA. An analytical survey of large vocabulary Russian speech recognition systems. *SPIIRAS Proceedings.* 2014;1(12):7. (In Russ). doi: 10.15622/sp.12.1
40. Kudryavtsev ND, Semenov DS, Kozhikhina DD, Vladzmyrskyy AV. Speech recognition technology: Results of a survey of radiologists at the Moscow reference center for diagnostic radiology. *Healthcare Management.* 2022;8(3):95–104. (In Russ). doi: 10.33029/2411-8621-2022-8-3-95-104
41. Sinityn VE, Komarova MA, Mershina EA. Protocol of radiological description: Past, present, future. *Bulletin Radiology Radiology.* 2014;(3):35–40. (In Russ).
42. Sobez LM, Kim SH, Angstwurm M, et al. Creating high-quality radiology reports in foreign languages through multilingual structured reporting. *Eur Radiol.* 2019;29(11):6038–6048. doi: 10.1007/s00330-019-06206-8
43. Ganeshan D, Duong PA, Probyn L, et al. Structured reporting in radiology. *Academic Radiology.* 2018;25(1):66–73. doi: 10.1016/j.acra.2017.08.005
44. Dos Santos PD, Hempel JM, Mildenerger P, et al. Structured reporting in clinical routine. *Rofo.* 2019. Vol. 191, N 01. P. 33–39. doi: 10.1055/a-0636-3851
45. Andrianova MG, Kudryavtsev ND, Petryaykin AV. Development of a thesaurus of radiological terms for voice filling of diagnostic research protocols. *Digital Diagnostics.* 2022;3(S1):21–22. (In Russ). doi: 10.17816/DD105703

ОБ АВТОРАХ

* Кудрявцев Никита Дмитриевич;

адрес: Россия, 127051, Москва, ул. Петровка, д. 24, стр. 1;
ORCID: <http://orcid.org/0000-0003-4203-0630>;
eLibrary SPIN: 1125-8637; e-mail: KudryavtsevND@zdrav.mos.ru

Бардасова Кристина Алексеевна;

ORCID: <http://orcid.org/0009-0002-4310-1357>;
eLibrary SPIN: 1156-7627; e-mail: bardasovakris@mail.ru

Хоружая Анна Николаевна;

ORCID: <http://orcid.org/0000-0003-4857-5404>;
eLibrary SPIN: 7948-6427; e-mail: KhoruzhayaAN@zdrav.mos.ru

AUTHORS' INFO

* Nikita D. Kudryavtsev;

address: 24/1 Petrovka street, 127051 Moscow, Russia;
ORCID: <http://orcid.org/0000-0003-4203-0630>;
eLibrary SPIN: 1125-8637; e-mail: KudryavtsevND@zdrav.mos.ru

Kristina A. Bardasova;

ORCID: <http://orcid.org/0009-0002-4310-1357>;
eLibrary SPIN: 1156-7627; e-mail: bardasovakris@mail.ru

Anna N. Khoruzhaya;

ORCID: <http://orcid.org/0000-0003-4857-5404>;
eLibrary SPIN: 7948-6427; e-mail: KhoruzhayaAN@zdrav.mos.ru

* Автор, ответственный за переписку / Corresponding author