**МИНИСТЕРСТВО ЦИФРОВОГО РАЗВИТИЯ СВЯЗИ И МАССОВЫХ КОММУНИКАЦИЙ**

**Ордена Трудового Красного Знамени**

**Федеральное государственное бюджетное образовательное учреждение высшего образования**

**«Московский технический университет связи и информатики»**

Кафедра «Математическая Кибернетика и Информационные технологии»

Проектный практикум

Использование ИИ в медицине

Выполнил: Студент группы

БВТ2201

Стасюк Вячеслав

Москва

2023

ОГЛАВЛЕНИЕ

[ВВЕДЕНИЕ 3](#__RefHeading___1)

[АКТУАЛЬНОСТЬ 3](#__RefHeading___2)

[Глава 1. Анализ предметной области: 4](#__RefHeading___3)

[1.2 Разбор использования методов ИИ на конкретных примерах. 6](#__RefHeading___4)

[Автоматизация списка медицинских проблем с использованием обработки естественного языка 6](#__RefHeading___5)

[COVID-Net: адаптированная глубокая конволюционная нейронная сеть для выявления случаев COVID-19 по рентгеновским изображениям грудной клетки 14](#__RefHeading___6)

[Обнаружение речевых аномалий с помощью классификатора последовательностей, основанного на восприятии и использующего универсальную модель речи 28](#__RefHeading___7)

[СПИСОК ИСТОЧНИКОВ 38](#__RefHeading___8)

# ВВЕДЕНИЕ

Медицине очень важно использовать актуальные технологии. Каждая новая разработка в любой области человека может принести пользу в спасении или улучшении жизни человека. Изучение катодных лучей Вильгельмом Рентгеном привело к созданию рентген-аппарата, очень важного изобретения для диагностики. Открытие ядерного магнитного резонанса привело к созданию МРТ. Прямо сейчас происходит активное развитие искусственного интеллекта (ИИ). ИИ можно использовать в анализе изображений, больших данных, работы с языком и многим другим. Это невероятно перспективный и бурно развивающийся инструмент, который просто не может обойти стороной медицина. В данной работе мы рассмотрим современное применение технологий ИИ в медицине.

# АКТУАЛЬНОСТЬ

Внедрение искусственного интеллекта (ИИ) одна из самых значимых современных тенденций, меняющих подходы к диагностике, лечению и управлению пациентами. Согласно данным аналитического отчёта MarketsandMarkets[1], глобальный рынок ИИ в здравоохранении в 2023 году оценивался в 14 миллиардов долларов, и прогнозируется его рост до 102 миллиардов долларов к 2028 году со среднегодовым темпом прироста (CAGR) более 47%. Это обусловлено стремлением повысить точность диагностики, оптимизировать управление медицинскими данными и улучшить результаты лечения.

Диагностические ошибки являются одной из ключевых проблем в медицине, приводя к осложнениям у 10-15% пациентов. Как показывают исследования, опубликованные в *The Lancet Digital Health*[2], алгоритмы глубокого обучения, применяемые для анализа изображений, таких как КТ и МРТ, позволяют сократить количество ошибок до 40%, предоставляя врачам дополнительную поддержку в выявлении патологий, включая онкологические и сердечно-сосудистые заболевания.

В то же время Всемирная организация здравоохранения (ВОЗ)[3] прогнозирует, что к 2030 году миру будет не хватать 18 миллионов квалифицированных медицинских специалистов. В этих условиях ИИ становится важным инструментом для автоматизации рутинных задач, таких как ведение медицинской документации, обработка результатов анализов и управление логистикой пациентов. Это позволяет высвобождать время для медицинского персонала и повышать эффективность систем здравоохранения.

Использование ИИ не только помогает решать текущие вызовы медицины, но и открывает новые возможности для персонализированного лечения и профилактики заболеваний.

**Цель задания:**

Изучить современные подходы, технологии и примеры использования искусственного интеллекта (ИИ) в медицине, а также сформировать аналитический отчет, содержащий ключевую информацию о текущем состоянии и перспективах применения ИИ в этой области.

Задачи:

1. Изучить существующие направления ИИ
2. Изучить какие из этих направлений сейчас применяются в медицине
3. Подробно разобрать отдельные исследования по применению отдельной технологии искусственного интеллекта в медицине.

# Глава 1. Анализ предметной области:

Искусственный интеллект это не просто универсальный алгоритм, который можно применить к разным задачам. Искусственный интеллект это целая область компьютерных наук, которая подразделяется на множество направлений. Можно выделить следующие направления ИИ, которые либо уже сейчас используются в медицине, либо у которых есть потенциал для использования в будущем.

* Обработка естественного языка (*Natural Language Processing, NLP*) - это форма машинного обучения, позволяющая обрабатывать и анализировать различные тексты. При использовании в медицинских записях она может помочь в прогнозировании исхода заболевания, дополнить системы сортировки в больницах и создать диагностические модели для выявления хронических заболеваний на ранних стадиях. Эти приложения могут быть особенно полезны в реанимации, где требуется анализировать больше данных о пациентах и прогнозировать их смертность. Помимо способности понимать естественный язык (NLU), NLP также может создавать естественный язык (NLG), предоставляя пациентам возможность задавать вопросы и получать доступ к соответствующей информации в виде чат-ботов[4].
* Компьютерное зрение (Computer Vision, CV) основывается на алгоритмах глубокого обучения (deep learning), в частности сверточных нейронных сетях (CNN). Эти сети используют большие наборы данных изображений для обучения задач, таких как классификация объектов, их локализация, сегментация и реконструкция изображений.[5] В медицине применяются специализированные версии этих алгоритмов, адаптированные к особенностям медицинских данных (например, большие размеры изображений или трёхмерные структуры).
* Автоматическое распознавание речи (ASR, Automatic Speech Recognition) — это технология, преобразующая устную речь в текстовый формат. Она основана на алгоритмах искусственного интеллекта и машинного обучения, которые анализируют аудиосигналы, выделяют фонемы и составляют из них слова и предложения. ASR демонстрирует значительный потенциал в выявлении речевых нарушений, связанных с неврологическими состояниями.[6]

## 1.2 Разбор использования методов ИИ на конкретных примерах.

### **Автоматизация списка медицинских проблем с использованием обработки естественного языка**

**Проблемный список** (The medical problem list) является важной частью медицинской документации в США, а также центральным элементом ориентированной на проблемы электронной медицинской записи, разрабатываемой в учреждении Intermountain Health Care, штата Юта. Для того чтобы выполнять свои функции, проблемный список должен быть максимально точным и актуальным. В большинстве стационарных отделений этого учреждения переходят от бумажного инструмента к электронной версии списка. Однако текущий проблемный список, записываемый вручную в бумажной карте, обычно является неполным, неточным и зачастую совсем не используется. Авторы надеются, что преимущества доступного для всех списка проблем изменят эту ситуацию. В качестве дополнительного стимула они создают среду, где ведение списка проблем становится легким и эффективным процессом.

**Медицинский проблемный список**

Более трех десятилетий назад Ларри Уид предложил **ориентированную на проблемы медицинскую запись (Problem-Oriented Medical Record)** как средство решения проблемы сложности медицинских знаний, клинических данных и недостатков в документации медицинской помощи [8,9]. Он отметил отсутствие последовательной структуры и содержания в прогресс-нотах (записях о состоянии пациента), которые составляют значительную часть медицинской документации.

Уид предложил стандартный подход, акцентируя внимание на списке медицинских проблем пациента, который должен тщательно поддерживаться специалистами, оказывающими уход. Этот проблемный список служит двум основным целям:

1. Предоставляет краткое формальное резюме заболеваний пациента.
2. Является инструментом для организации рутинной документации врачебных решений, а также планов и результатов лечения.

**Кодирование записей в проблемном списке**

Каждая медицинская проблема должна иметь соответствующий код в системе контролируемой терминологии.

Ниже примеры различных медицинских словарей, используемых для кодирования проблемных списков:

* **ICD-9-CM** (Международная классификация болезней) [10],
* **SNOMED** (Systematized Nomenclature of Medicine) [11],
* **UMLS®** (Unified Medical Language System) [12–14],
* **Локально разработанные словари** [15].

В статье подчеркивается, что **SNOMED-CT** позволяет охватить 98,5% терминов, используемых в проблемных списках [11]. Кодирование медицинских проблем может быть выполнено вручную, когда код назначается при добавлении проблемы, либо автоматически с использованием **техник обработки естественного языка (NLP)**. Авторы акцентируют внимание на том, что использование NLP позволяет вводить данные на естественном языке, который является наиболее удобным и выразительным способом записи информации.

**Обработка естественного языка в медицине**

Авторы подчеркивают, что медицинская запись пациента содержит значительное количество информации, помимо записей в проблемном списке. Однако большая часть клинической информации представлена в неструктурированном виде, в так называемом **свободном тексте**. Эти документы включают историю болезни пациента, отчеты о терапевтических вмешательствах и описание клинического прогресса, составляя значительную часть медицинской документации.

С ростом использования кодированных данных в системах поддержки принятия решений и увеличением требований к стандартным медицинским наборам данных возникает потребность в преобразовании свободного текста в структурированные и кодированные данные. NLP становится ключевым инструментом для этой задачи, позволяя преобразовывать текст в кодированную информацию, что расширяет функциональность **ориентированной на проблемы компьютерной записи пациента (CPR)** [16].

**Исследования и применение NLP в медицине**

Авторы описывают работу различных групп, которые разрабатывали методы автоматического кодирования текстовых медицинских документов:

1. **Linguistic String Project** [17,18]: Разработка инструментов для анализа медицинского текста.
2. **Системы для анализа рентгеновских отчетов**:
   * Zingmond [19] применил инструмент семантического кодирования для анализа рентгеновских отчетов, выявляя аномалии, требующие дальнейшего наблюдения.
   * Friedman и коллеги [20–24] изучили методы кодирования интерпретаций, содержащихся в этих отчетах, и применяли NLP к отчетам по маммографии [25], нейрорадиологии [26], выпискам пациентов [27] и патологическим отчетам [28].
3. **MedLEE (Medical Language Extraction and Encoding system)** [29]:
   * Приложение MedLEE было разработано для извлечения и кодирования медицинского текста, достигая **83% полноты (recall)** и **89% точности (precision)** [30].
   * Оно также стало первым биомедицинским NLP-инструментом, который был применен к данным из другого учреждения, где после корректировок достигло производительности, аналогичной исходной [24].

**Разработки группы медицинской информатики Университета Юты**

Группа медицинской информатики LDS Hospital и Университета Юты активно занимается исследованиями NLP на протяжении многих лет:

1. **SPRUS (Special Purpose Radiology Understanding System)** [31,32]: Первое NLP-приложение, ориентированное только на семантический анализ.
2. **SymText** [33–35]: Объединил синтаксический и вероятностный семантический анализ.
3. **MPLUS** [36]: Последняя версия, включающая синтаксический анализ на основе контекстно-свободной грамматики с использованием **bottom-up chart parser**, совмещенного с объектно-ориентированным семантическим анализом.

SymText и MPLUS активно используют **семантические сети**, реализованные как **байесовские сети** (belief networks) [37]. Эти сети обучаются для определения вероятностных связей между извлеченными терминами и их значениями. Такой подход позволяет:

* Быть устойчивым к шумным данным.
* Улучшать производительность через обучение.

**Тестирование методов NLP в радиологии**

Авторы подчеркивают, что ключевой областью тестирования подходов NLP стала радиология:

* Основное внимание было уделено анализу отчетов о **рентгенографии грудной клетки** [32–34], особенно данных, поддерживающих диагноз **пневмонии** [38,39].
* Для проверки методов NLP использовались также диагнозы при поступлении [40] и отчеты о сканировании перфузии легких [41].

В рамках описанного в статье проекта система **MPLUS** была адаптирована и обучена для извлечения медицинских проблем из электронных текстовых документов [68].

Авторы делают вывод, что использование NLP позволяет эффективно преобразовывать свободный текст в структурированную информацию, облегчая анализ данных и их последующее использование в медицинской практике.

**Целевые медицинские проблемы**

Первым шагом было определение набора медицинских проблем, которые система будет обрабатывать:

* В рамках пилотного проекта выбрано **80 медицинских проблем**, связанных с кардиоваскулярными заболеваниями.
* Источники отбора проблем:
  + **Лог центральной клинической базы данных (Clinical Data Repository, CDR)** – концепции, закодированные как медицинские проблемы, отсортированные по частоте использования за 2002 год.
  + **Список топ-25 диагнозов кардиоваскулярных пациентов** в LDS Hospital.
  + Итоговый список был проверен двумя опытными врачами: специалистом по внутренней медицине и специалистом по интенсивной терапии.

Итоговый список включал такие проблемы, как анемия, инфаркт миокарда, гипертензия и пневмония, и охватывал значительную долю частых диагнозов.

**Система автоматического списка медицинских проблем (APL) [7]**

APL состоит из двух основных компонентов:

1. **Фоновое приложение:**
   * Выполняет обработку свободного текста медицинских документов.
   * Извлекает упоминания медицинских проблем, которые сохраняются в CDR.
2. **Приложение управления списком проблем:**
   * Позволяет пользователям просматривать, подтверждать и редактировать извлеченные проблемы.
   * Обеспечивает доступ к источникам знаний через кнопку "инфокнопка" для каждой закодированной проблемы.

**Модель данных**

Для обмена данными между компонентами использовалась модель данных:

* Документы представлены в формате **HL7 Clinical Document Architecture (CDA)**.
* Медицинские проблемы кодируются как **наблюдения (Observations)**.
* Проблемы сохраняются в CDR в формате **ASN.1 Clinical Event**.

XML-формат позволяет связать извлеченную проблему с исходным предложением, чтобы пользователи могли просматривать источник проблемы.

**Обработка текста**

1. **Определение секций:**
   * Секции документа (например, "История текущего заболевания") идентифицируются с помощью регулярных выражений.
   * 539 различных заголовков секций были сопоставлены с 20 стандартными заголовками для унификации.
2. **Определение предложений:**
   * После идентификации секций текст разбивается на предложения с использованием регулярных выражений.
   * Регулярные выражения учитывали знаки препинания и другие маркеры для разделения текста на предложения.

**Модуль NLP**

1. **Анализ предложений:**
   * NLP-модуль анализирует каждое предложение с учетом контекста (тип документа и секция).
   * Проблемы классифицируются как "присутствующие" или "отсутствующие".
   * Если одна и та же проблема упоминается как "присутствующая" и "отсутствующая", приоритет отдается статусу "отсутствует" на уровне документа.
2. **Тренировка:**
   * Создано 4436 обучающих случаев с использованием полуавтоматизированного подхода:
     + Сначала выбраны предложения, содержащие целевые проблемы, с использованием регулярных выражений и фраз из UMLS Metathesaurus.
     + Затем вручную определено состояние проблемы (присутствует/отсутствует).
3. **Алгоритмы:**
   * Первая версия модуля базировалась на SymText с использованием **байесовских сетей** для анализа.
   * Вторая версия использовала **MetaMap Transfer (MMTx)** для обнаружения медицинских проблем и **NegEx** для определения отрицаний.

**Формирование документов CDA**

После анализа документов:

* Создавались версии документов в формате CDA с закодированными проблемами.
* Пользователь мог просматривать документы с подсвеченными предложениями, содержащими медицинские проблемы.

**Сохранение данных**

* Извлеченные проблемы, которые уже были сохранены в CDR, не добавляются повторно.
* Новые проблемы сохраняются со статусом "предложенные" для последующего подтверждения пользователем.

**Приложение управления списком проблем**

* Пользователи взаимодействуют с системой через интерфейс "Clinical Desktop".
* Основные функции:
  + Просмотр и редактирование проблем.
  + Подтверждение ("активные" или "неактивные") или отклонение ("ошибка") предложенных проблем.
  + Кнопка "источник" позволяет быстро перейти к документу, откуда была извлечена проблема, с подсвеченными предложениями.

Приложение значительно упрощает процесс добавления проблем, обеспечивая высокую скорость анализа (менее одной секунды на извлечение и отображение источника).

Для оценки доли кодированных диагнозов, охватываемых набором из 80 целевых проблем, были получены все уникальные коды МКБ-9-CM, присвоенные сердечно-сосудистым взрослым стационарам в 2003 г. в учреждении (LDS Hospital, Солт-Лейк-Сити, штат Юта). Всего было использовано 1531 различных кодов. Таким образом, набор из 80 целевых проблем охватывал только 5,2% всех возможных кодов. В наборе данных было найдено в общей сложности 24160 кодов по МКБ-9-CM, и 15449 из них соответствовали одной из 80 целевых задач. Таким образом, набор проблем охватывает 63,9 % всех случаев кода.

Небольшая пробная оценка алгоритмов обнаружения разделов и предложений показала хорошую производительность. Алгоритм обнаружения разделов достиг чувствительности 1,00 и положительной предсказательной ценности 1,00. Он обнаружил все 146 разделов, присутствующих в 20 случайно выбранных документах из нашего набора данных, и обнаружил только эти 146 разделов. Чувствительность алгоритма обнаружения предложений составила 0,889 (95% доверительный интервал 0,78-0,998), а положительное предсказательное значение - 0,946 (0,907-0,985). Было обнаружено 687 из 731 предложения, присутствующих в тестовом наборе из 20 документов. Из этих 687 предложений 662 были обнаружены правильно (т. е. истинно положительные), а 25 - ложно положительные.

### **COVID-Net: адаптированная глубокая конволюционная нейронная сеть для выявления случаев COVID-19 по рентгеновским изображениям грудной клетки**

Пандемия коронавирусной инфекции 2019 года (COVID-19) оказала разрушительное воздействие на здоровье и благополучие население планеты. Важнейшим шагом против борьбы с COVID-19 является эффективный метод диагностики, который включает в себя рентгенологическое обследование с использованием рентгенографии грудной клетки. На рентгенограммах грудной клетки пациенты демонстрируют отклонения от нормы, характерные для инфицированных COVID-19.[43, 44] В данном исследовании представляют COVID-Net[42] - глубокую конволюционную нейронную сеть, предназначенную для выявления случаев COVID-19 по рентгеновским снимкам грудной клетки (CXR), с открытым исходным кодом и доступную широкой общественности.

**Проблемы существующих методов диагностики**:

* **RT-PCR (реакция полимеразной цепи с обратной транскрипцией)[45]**:
  + Является золотым стандартом диагностики COVID-19.
  + Проблемы:
    - Процесс медленный, трудоемкий, требует высококвалифицированного персонала и ресурсов.
    - Высокая вариативность чувствительности (зависит от методики сбора образцов, времени с момента появления симптомов).
  + Нехватка тестов RT-PCR делает процесс массового тестирования сложным, особенно в условиях высокого спроса.
* **Рентгенография грудной клетки (CXR)[46]**:
  + Выступает альтернативным методом диагностики.
  + На ранних стадиях исследования было показано, что у пациентов с COVID-19 наблюдаются характерные аномалии на CXR.
  + Преимущества CXR:
    - Быстрота получения результата, что позволяет использовать его для триажа.
    - Доступность оборудования в большинстве медицинских учреждений, в том числе в развивающихся странах.
    - Наличие портативных систем, позволяющих минимизировать риск передачи вируса при транспортировке пациентов.

**Проблема нехватки специалистов**:

* Для интерпретации рентгеновских снимков необходимы высококвалифицированные радиологи.
* Компьютерные системы диагностики могут помочь врачам быстрее и точнее интерпретировать CXR-снимки.

**COVID-Net**:

* Создана специализированная архитектура сверточной нейронной сети для диагностики COVID-19 на основе CXR.
* Особенности архитектуры:
  + Применение стратегии "коллаборативного проектирования человек-машина".
  + Легковесная структура с высокой вычислительной эффективностью.
* COVID-Net стала одной из первых открытых архитектур для детекции COVID-19, доступной для исследовательского сообщества.

**COVIDx Dataset**:

* Создан крупнейший открытый набор данных COVIDx, состоящий из 13,975 рентгеновских изображений 13,870 пациентов.
* Набор данных объединяет изображения из пяти публичных репозиториев:
  + COVID-19 Image Data Collection.
  + Fig. 1 COVID-19 Chest X-ray Dataset Initiative.
  + ActualMed COVID-19 Chest X-ray Dataset Initiative.
  + RSNA Pneumonia Detection Challenge dataset.
  + COVID-19 Radiography Database.

**Методы объяснимости**:

* Впервые применены методы объяснимости для анализа решений модели.
* Цель:
  + Выявить ключевые факторы, на которые модель опирается при диагностике.
  + Убедиться, что модель принимает решения на основе релевантных данных.
  + Сделать систему прозрачной и надежной для клинического использования.

**Развитие COVID-Net и COVIDx:**

С момента публикации набора данных COVIDx и модели COVID-Net было проведено множество исследований в области диагностики COVID-19 на основе рентгенографии. Некоторые из этих исследований использовали модифицированные версии COVIDx или COVID-Net. Основное внимание уделялось изучению глубокого обучения, особенно сверточных нейронных сетей, благодаря их значительным успехам и достижениям в различных задачах компьютерного зрения.

В данном исследовании для создания COVID-Net была применена стратегия коллаборативного проектирования "человек-машина", где прототипирование сети, основанное на принципах, разработанных человеком, было объединено с машинным исследованием дизайна для разработки архитектуры сети, оптимизированной для диагностики COVID-19 на основе рентгенограмм грудной клетки (CXR). Также был создан открытый эталонный набор данных COVIDx для обучения и оценки COVID-Net.

**COVIDx dataset (Набор данных COVIDx)**

1. **Общие характеристики**:
   * COVIDx включает 13,975 CXR-изображений из 13,870 случаев пациентов.
   * На момент написания статьи это крупнейший открытый эталонный набор данных по количеству случаев пациентов, инфицированных COVID-19.
2. **Процесс создания**:
   * Набор данных COVIDx был создан путем объединения и модификации пяти различных публичных репозиториев данных:
     + **COVID-19 Image Data Collection**.
     + **COVID-19 Chest X-ray Dataset Initiative** (создана в сотрудничестве с Fig. 1).
     + **ActualMed COVID-19 Chest X-ray Dataset Initiative**.
     + **RSNA Pneumonia Detection Challenge dataset** (использует открытые данные CXR из других источников).
     + **COVID-19 Radiography Database**.
3. **Примеры изображений**:
   * В статье приведены примеры CXR-изображений из набора данных COVIDx, демонстрирующие разнообразие случаев пациентов.
4. **Классы данных**:
   * В набор данных включены следующие классы:
     + Случаи без пневмонии (нормальные изображения).
     + Случаи пневмонии, не связанные с COVID-19.
     + Случаи пневмонии, вызванной COVID-19.
5. **Распределение данных**:
   * **COVID-19**: Ограниченное количество случаев, что отражает недостаток данных о COVID-19 в публичном доступе.
   * **Другие классы**:
     + 8,066 случаев без пневмонии.
     + 5,538 случаев пневмонии, не связанной с COVID-19.
   * Набор данных включает 358 CXR-изображений от 266 пациентов с COVID-19.
6. **Доступность**:
   * Скрипты для генерации набора данных COVIDx опубликованы в открытом доступе на GitHub (ссылка предоставлена в статье).

Прототипирование сети на основе принципиального проектирования

Первая стадия стратегии коллаборативного проектирования "человек-машина", использованная для создания COVID-Net, заключается в **прототипировании сети на основе принципиального проектирования**. На этой стадии прототип сети разрабатывается с использованием принципов проектирования и лучших практик, разработанных человеком.

**Основные аспекты:**

1. **Выбор архитектуры**:
   * В данном исследовании использовались принципы проектирования **остаточных архитектур** (residual architectures), которые многократно доказали свою эффективность:
     + Позволяют разрабатывать надежные архитектуры нейронных сетей, которые проще обучать для достижения высокой производительности.
     + Обеспечивают возможность построения более глубоких архитектур.
2. **Цели проектирования прототипа**:
   * Конструкция сети была разработана таким образом, чтобы обеспечивать один из следующих трех возможных выходов (предсказаний):
     + **Нормальный**: отсутствие инфекции.
     + **Не связанная с COVID-19 инфекция**: например, бактериальная или вирусная пневмония, не вызванная SARS-CoV-2.
     + **COVID-19**: вирусная инфекция, вызванная SARS-CoV-2.
3. **Мотивация выбора классов предсказания**:
   * Такой подход помогает клиницистам:
     + Определить, кто должен быть в приоритете для проведения RT-PCR тестирования для подтверждения диагноза COVID-19.
     + Выбрать стратегию лечения в зависимости от типа инфекции, так как COVID-19 и другие виды инфекций требуют различных подходов.
4. **Примеры изображений**:
   * В статье приведены примеры рентгенограмм из набора данных COVIDx, показывающие различия между снимками при инфекциях, не связанных с COVID-19, и снимками пациентов с COVID-19.

Вторая стадия стратегии коллаборативного проектирования "человек-машина", использованная для создания COVID-Net, заключается в **машинном исследовании дизайна**. На этой стадии используется предварительно созданный человеком прототип сети, данные и специфические требования для автоматического изучения и поиска оптимальных макро- и микроархитектур сети.

**Основные аспекты:**

1. **Цель стадии**:
   * Оптимизация архитектуры для конкретной задачи детекции COVID-19 на основе рентгенограмм.
   * Достижение более высокой степени детализации и гибкости, чем это возможно при ручном проектировании.
2. **Методология**:
   * Используется стратегия **генеративного синтеза** (generative synthesis), в основе которой лежит взаимодействие двух компонентов:
     + **Генератор (Generator, G)**: отвечает за создание вариантов архитектуры сети на основе заданных параметров.
     + **Инквизитор (Inquisitor, I)**: анализирует созданные варианты и дает обратную связь для оптимизации генератора.
3. **Процесс генеративного синтеза**:
   * Генератор создает нейронную сеть NsN\_sNs​ с заданными параметрами (например, топологией сети).
   * Инквизитор оценивает производительность созданной сети, учитывая заданные человеком требования.
   * На основе анализа инквизитора параметры генератора обновляются, чтобы генерировать улучшенные варианты сети.
   * Этот процесс повторяется до тех пор, пока сеть не будет соответствовать заданным критериям.
4. **Требования к дизайну**:
   * В исследовании использовались следующие требования, чтобы управлять процессом генерации:
     + **Чувствительность к COVID-19** ≥ 80%.
     + **Положительная прогностическая ценность (PPV)** ≥ 80%.
   * Эти критерии важны для минимизации пропущенных случаев COVID-19 и снижения числа ложноположительных предсказаний, которые могут перегрузить клинические ресурсы.
5. **Результаты автоматизированного проектирования**:
   * Сгенерированная архитектура COVID-Net учитывает:
     + Баланс между высокой емкостью представления данных (representational capacity) и вычислительной эффективностью.
     + Чувствительность к COVID-19, что критически важно для медицинских приложений.

**Основные характеристики архитектуры COVID-Net:**

1. **Доступность** - архитектура COVID-Net доступна в открытом доступе на GitHub (ссылка предоставлена в статье), что способствует воспроизводимости и дальнейшему развитию.
2. **Интересные особенности архитектуры** - созданная с помощью стратегии коллаборативного проектирования "человек-машина", архитектура COVID-Net включает ряд уникальных характеристик, которые способствуют эффективности при детекции COVID-19 на основе рентгенограмм грудной клетки.

**3. Легковесный дизайн:**

COVID-Net активно использует легковесный паттерн **Projection-Expansion-Projection-Extension (PEPX)**, который включает следующие стадии:

* **Первая проекция**:
  + 1×1 свертки для проекции входных признаков в пространство с меньшей размерностью.
* **Расширение**:
  + 1×1 свертки для увеличения размерности признаков до более высокого уровня.
* **Глубинные сверточные слои**:
  + Эффективные 3×3 глубинные сверточные операции для изучения пространственных характеристик с минимизацией вычислительных затрат.
* **Вторая проекция**:
  + 1×1 свертки для возврата признаков в пространство с меньшей размерностью.
* **Расширение**:
  + Финальная 1×1 свертка для увеличения размерности каналов до конечного уровня.

Этот дизайн, открытый в процессе машинного исследования, увеличивает представляемость данных (representational capacity) при одновременном снижении вычислительной сложности.

**4. Избирательная долгосрочная связь:**

* Архитектура включает избирательные долгосрочные соединения (long-range connectivity), которые повышают представляемость данных и упрощают обучение.
* Особенности:
  + Четыре плотно соединенных слоя с 1×1 свертками служат центральными "хабами", соединяющими ранние слои с более поздними.
  + Этот подход позволяет сохранить вычислительную и память-эффективность, избегая чрезмерного увеличения сложности, характерного для плотно соединенных сетей.

**5. Архитектурное разнообразие:**

* COVID-Net обладает значительным разнообразием архитектуры, включая:
  + Сверточные слои с различными размерами ядер (от 7×7 до 1×1).
  + Различные конфигурации группировки слоев (от неразделенных до глубинных).
* Это разнообразие позволяет модели быть тонко настроенной для задачи детекции COVID-19 и обеспечивает высокую представляемость данных.

**6. Детали реализации:**

* COVID-Net была предварительно обучена на наборе данных ImageNet, а затем дообучена на COVIDx.
* Гиперпараметры:
  + learning rate: 2×10−4.
  + Количество эпох: 22.
  + Размер батча: 64.
  + Фактор уменьшения скорости обучения: 0.7.
  + "Терпение" (patience): 5 (снижение скорости обучения при стагнации).
* Применена стратегия балансировки батчей для равномерного представления каждого типа инфекции.

**COVID Net auditing via explainability (Аудит COVID-Net с использованием методов объяснимости)**

1. **Необходимость объяснимости**:
   * Важность объяснимости в клинических приложениях, таких как детекция COVID-19:
     + Обеспечение уверенности, что модель делает правильные выводы на основе релевантной информации.
     + Исключение случаев, когда решения основаны на неправильных данных, таких как артефакты изображения или ненужные визуальные метки.
2. **Методика объяснимости**:
   * Применен подход **GSInquire**, который является ключевым компонентом стратегии генеративного синтеза.
   * Основные принципы GSInquire:
     + Генератор GGG создает архитектуру сети.
     + Инквизитор III анализирует сеть и выявляет критические факторы, которые используются для принятия решений.
   * Для объяснения решений модели используется интерпретация z(x;Nref), которая визуализирует ключевые области, влияющие на предсказания.
3. **Цели аудита**:
   * Проверить, что COVID-Net делает предсказания, основываясь на корректной информации (например, областях легких), а не на нерелевантных признаках.
   * Повысить прозрачность модели и доверие клиницистов к ее результатам.

**Ethics approval (Этическое одобрение)**

* Исследование получило одобрение комитета по этике Университета Ватерлоо (номер протокола: 42235).
* Все действия, связанные с данными, соответствуют требованиям этических стандартов.

**Experimental setup (Настройка эксперимента)**

1. **Подготовка данных**:
   * Использовался набор данных **COVIDx**.
   * Предобработка данных включала:
     + Обрезку верхних 8% изображений для удаления текстовой информации.
     + Аугментацию данных (сдвиги, вращение, горизонтальное отражение, масштабирование, изменение интенсивности).
2. **Метрики оценки**:
   * **Точность (Accuracy)**.
   * **Чувствительность (Sensitivity)**: показывает, насколько хорошо модель выявляет положительные случаи.
   * **Положительная прогностическая ценность (PPV)**: указывает на долю правильных положительных предсказаний.
3. **Гиперпараметры и обучение**:
   * Модель обучалась с использованием оптимизатора **Adam**.
   * Снижение скорости обучения происходило по мере стагнации улучшений (основано на терпении в 5 эпох).
   * COVID-Net обучалась в течение 22 эпох с батчем размером 64.

**Количественный анализ:**

* **Точность**:
  + COVID-Net достигла точности **93.3%** на тестовом наборе данных COVIDx.
  + Модель превосходит альтернативы: VGG-19 (83.0%) и ResNet-50 (90.6%).
* **Чувствительность (Sensitivity)**:
  + COVID-Net показала чувствительность к COVID-19 **91%**, что выше, чем у ResNet-50 (83%) и значительно выше, чем у VGG-19 (58.7%).
* **PPV (Положительная прогностическая ценность)**:
  + COVID-Net достигла PPV **98.9%** для COVID-19, минимизируя ложноположительные случаи.
* **Вычислительная сложность**:
  + COVID-Net имеет значительно меньшую сложность (7.50 GMACs), чем VGG-19 (89.63 GMACs) и ResNet-50 (17.75 GMACs).

**Сравнение архитектур:**

* Уникальные особенности COVID-Net, такие как легковесные паттерны PEPX и избирательные долгосрочные связи, обеспечили баланс между производительностью и эффективностью.
* COVID-Net требует меньше вычислительных ресурсов, при этом показывая лучшую точность и чувствительность.

**Качественный анализ (Объяснимость):**

* С помощью метода GSInquire было подтверждено, что COVID-Net делает предсказания на основе релевантных данных, таких как области легких на рентгенограммах.
* Модель корректно идентифицировала ключевые визуальные индикаторы, такие как двусторонние аномалии и стекловидная опалесценция, используемые врачами для диагностики.

**Основные выводы:**

* COVID-Net демонстрирует высокую точность и надежность в детекции COVID-19 на основе CXR.
* Анализ объяснимости повысил доверие к модели, исключив сценарии "правильное решение по неправильным причинам".

**COVID-Net**:

* Представлена новая архитектура сверточной нейронной сети для детекции COVID-19 на основе рентгенограмм грудной клетки.
* Модель доступна в открытом доступе, что способствует развитию исследовательских инициатив.

**COVIDx Dataset**:

* Создан крупнейший открытый эталонный набор данных для диагностики COVID-19, объединяющий данные из пяти источников.

**Объяснимость**:

* Использование методов объяснимости показало, что модель принимает решения на основе релевантных данных, повышая доверие клиницистов.

**Результаты**:

* COVID-Net достигла высокой точности (93.3%), чувствительности к COVID-19 (91%) и положительной прогностической ценности (98.9%).

### **Обнаружение речевых аномалий с помощью классификатора последовательностей, основанного на восприятии и использующего универсальную модель речи**

Авторы разработали модель классификации речевых аномалий[47], основанную на архитектуре Perceiver, в сочетании с Универсальной речевой моделью (Universal Speech Model, USM). Исследование направлено на решение задач автоматического выявления речевых отклонений, связанных с неврологическими заболеваниями, таких как нарушение координации дыхания, голосовых связок и артикуляции.

Неврологические повреждения или заболевания являются одной из ведущих причин смертности и заболеваемости в мире. Одной из ключевых причин этого является недостаток доступа к специализированным экспертам, что приводит к задержкам в диагностике, ошибочным диагнозам и отсутствию соответствующего лечения.

Многие неврологические заболевания проявляются через нарушения речи. Например, медленная или хриплая речь характерна для моторных речевых расстройств (например, дизартрии), которые часто связаны с такими заболеваниями, как боковой амиотрофический склероз (БАС). Для выявления аномалий речи специалисты по речи и языку (SLP) проводят клинические оценки, которые помогают в дифференциальной диагностике. Однако доступ к такой экспертизе ограничен.

Авторы утверждают, что автоматизация выявления аномалий речи может значительно повысить точность диагностики и качество лечения неврологических заболеваний. В статье представлена разработка моделей для выявления аномалий речи в рамках трех задач, часто используемых в клинической практике:

1. Продление гласного звука (Vowel Prolongation, VP).
2. Альтернативная частота движений (Alternate Motion Rate, AMR).
3. Последовательная частота движений (Sequential Motion Rate, SMR).

Эти задачи выполнялись на основе набора данных, который является значительным для клинической речи, но относительно небольшим для использования методов глубокого обучения. Авторы отмечают, что получение размеченных данных в этой области представляет сложности, связанные с:

* Ограничениями, связанными с конфиденциальностью медицинских данных.
* Трудоемкостью разметки и редкостью необходимой экспертизы.

Для преодоления ограничения в данных авторы использовали фундаментальные модели, обученные на больших объемах информации. Несмотря на долгую историю разработки "универсальных" моделей в области речи, их определение со временем изменялось: от устойчивости к различным акустическим средам и независимости от домена до мультиязычности. Современные фундаментальные модели основаны на значительно больших объемах данных и количестве параметров. Такие модели способны генерировать мощные представления речи, моделируя долгосрочные временные зависимости (вплоть до 120 секунд в используемом корпусе данных). Однако выделение релевантных сигналов для классификации последовательностей остается сложной задачей.

Для решения этой проблемы была предложена архитектура, основанная на Perceiver, для классификации последовательностей. Данная модель преобразует последовательности переменной длины в небольшой фиксированный набор латентных представлений, специфичных для классов, что делает её подходящей для задач классификации последовательностей.

Существуют многочисленные исследования, в которых применяются методы машинного обучения для выявления и количественной оценки речевых или неврологических нарушений на основе коротких речевых записей. Например, традиционные методы извлечения признаков в сочетании со статистическими или классическими моделями машинного обучения были применены для диагностики таких заболеваний, как боковой амиотрофический склероз (БАС)[48, 49], болезнь Альцгеймера (AD)[50], болезнь Паркинсона (PD) [51, 52], депрессия [53] и аутизм [54]

Хотя методы глубокого обучения на необработанных или неструктурированных аудиоданных показывают лучшие результаты, чем традиционные подходы, такие как распознавание эмоций [55], их использование для анализа речевых или неврологических нарушений встречается реже. Например, в недавнем исследовании использовались мел-спектрограммы и сверточные нейронные сети для различения пациентов с болезнью Паркинсона и здоровых людей. Точность составила 81.6% для анализа отдельных гласных звуков и 92% для комбинации различных звуков [56]. В другом исследовании автоэнкодеры использовались для извлечения признаков из спектрограмм, после чего на этих признаках обучался классификатор для различения здоровых людей и пациентов с когнитивными нарушениями, достигнув точности 90.57%[57]. Использование предварительно обученных моделей в этой области встречается ещё реже. Например, недавнее исследование применило модель Data2vec[58] для извлечения голосовых признаков, которые затем использовались в классификаторе для диагностики болезни Альцгеймера[59].

На первый взгляд точность, сообщаемая в этих исследованиях, выглядит многообещающе. Однако методологические ограничения снижают применимость таких решений в реальных условиях. Например:

* Задача бинарной классификации (здоровый/больной) не отражает сложность реальной клинической диагностики, где медицинские работники могут распознать аномалии речи, но не всегда могут определить конкретное заболевание.
* Многие исследования не используют настоящие тестовые наборы данных и работают с небольшими выборками. Например, наиболее популярные корпусы UA-Speech и TORGO, используемые для анализа дизартрии, содержат записи всего от 15 и 7 человек с речевыми нарушениями соответственно[60, 61].
* Малый объем данных вызывает сомнения в достоверности таких исследований. Более того, недавние работы показывают, что модели, обученные на данных, не включающих голосовые сегменты, могут превосходить модели, использующие голосовые части записи [62].
* Корпусы, используемые в этих исследованиях, часто сильно упрощены, и производительность моделей на таких наборах данных не переносится на более сложные и реалистичные данные, связанные с неврологическими заболеваниями [63].

С точки зрения моделирования задача классификации последовательностей решается различными подходами. В текстовых задачах классификаторы часто используют специальный "токен классификации", обобщающий информацию о последовательности [64]. Хотя этот подход можно адаптировать к аудиоданным, он будет менее оптимальным, так как разные классы требуют анализа различных областей сигнала. Альтернативно, для аудиопоследовательностей используются I-векторы [65], которые конденсируют сигнал в представление фиксированной длины, например, для задач верификации дикторов. Также применяются LSTM-сети, предоставляющие "нейронное" представление состояния последовательности [66]. Однако оба метода имеют ограничения, поскольку не позволяют различным классам анализировать различные регионы сигнала.

Модели внимания с латентными единицами, такие как Set Transformer[67], были предложены для снижения квадратичной временной сложности функций внимания. В модели Perceiver [68] перекрестное внимание с набором латентных единиц используется итеративно для сжатия данных высокой размерности из разных модальностей с последующей их интеграцией с блоками трансформеров. Однако в Perceiver для классификации используется усреднение выходных данных блока самовнимания, что отличается от предложенного подхода, где усреднение не выполняется, а применяется факторизованная проекция (подробности в разделе 4.1).

**Описание задачи**

Целью исследования является помощь клиницистам в выявлении аномальных характеристик речи, что может быть полезно для диагностики неврологических заболеваний. Как указано в разделе **Related Work**, существующие исследования имеют ограничения в экспериментальных настройках. Авторы стремятся преодолеть эти ограничения, используя более крупный набор данных, собранный в клинических условиях, с целью выявления различных аномалий речи в ряде задач. В данной работе рассматриваются три основные задачи:

1. **Продление гласных (Vowel Prolongation)**:
   * Пациенты просят сделать глубокий вдох и произнести звук "аааа" как можно дольше и равномернее, пока у них не закончится дыхание.
   * Эта задача направлена на изоляцию характеристик, связанных с дыханием и фонацией.
2. **Альтернативная частота движений (Alternating Motion Rate, AMR)**:
   * Пациентов просят сделать глубокий вдох и как можно быстрее и равномернее повторять звуки "пух".
   * Эта задача используется для оценки скорости и регулярности быстрых повторяющихся движений, а также для выявления нарушений артикуляции.
3. **Последовательная частота движений (Sequential Motion Rate, SMR)**:
   * Пациентов просят сделать глубокий вдох и как можно быстрее и равномернее произносить звуки "пух-тух-кух".
   * Эта задача используется для оценки способности быстро и точно воспроизводить последовательности звуков, что помогает выявлять артикуляционные и просодические нарушения.

**3.1. Моделируемые атрибуты (Метки)**

Для каждой задачи требуется предсказать ряд характеристик речи (всего 14 меток):

1. Ненормальная изменчивость громкости (abnormal loudness variability): может указывать на плохую координацию или контроль дыхательного потока.
2. Ненормальная изменчивость высоты звука (abnormal pitch variability): может указывать на плохую координацию движений гортани.
3. Хриплость (breathy): может указывать на слабость голосовых связок или неполное смыкание голосовой щели.
4. Искажения (distortions): могут быть связаны с дефицитом программирования или слабостью.
5. Дрожь (flutter): может указывать на слабость голосовых связок или быстрый тремор.
6. Хриплый/грубый голос (hoarse/harsh): может указывать на чрезмерное смыкание голосовых связок.
7. Нерегулярные нарушения артикуляции (irregular articulatory breakdowns): могут быть следствием плохой координации.
8. Уменьшение громкости (loudness decay): может указывать на снижение функций дыхательной или гортанной систем.
9. Быстрый темп (rapid rate): может указывать на дефицит контроля базальных ганглиев.
10. Медленный темп (slow rate): может быть следствием дефицита программирования, контроля или выполнения.
11. Напряженность (strained): может быть следствием гипераддукции голосовых связок.
12. Сегментация слогов (syllable segmentation): может указывать на плохое программирование или координацию.
13. Тремор (tremor): может быть связан с дефицитом в системе контроля базальных ганглиев.
14. Нестабильность (unsteady): может быть следствием плохой координации или контроля дыхательной и/или гортанной систем.

Каждая запись аннотировалась бинарными метками для всех 14 характеристик отдельно для каждой задачи. Так как эти задачи содержат чувствительные данные пациентов, набор данных недоступен публично. Однако для примера авторы ссылаются на небольшой публичный набор данных для задачи продления гласных у пациентов с болезнью Паркинсона.

**3.2. Метрики оценки**

Цель состоит в том, чтобы предсказать наличие каждой характеристики (метки) для каждой аудиозаписи. Одна запись может содержать несколько меток. Задача по сути представляет собой бинарную классификацию для каждой из 14 меток. Основной метрикой является средняя точность предсказания.

**3.3. Данные и разделение на тренировочные/тестовые выборки**

* Набор данных для всех трех задач включает **2594 аудиофайла** со средней длиной записи 37.5 секунд.
* Всего в выборке **782 уникальных спикера**, с средней частотой 3.3 записи на спикера.
* Данные сбалансированы по полу.
* Данные разделены на тренировочную (80%) и тестовую (20%) выборки таким образом, чтобы в них не пересекались спикеры. Разделение выполнено случайным образом, но в детерминированной манере.
* Эксперименты подтвердили, что результаты моделей стабильны при различных разделениях данных.

Для всех экспериментов данные из всех задач объединены, и модель выполняет классификацию совместно для всех задач.

Авторы представляют результаты экспериментов, направленных на оценку производительности предложенной архитектуры на основе Perceiver для классификации речевых аномалий.

**4.1. Архитектура Perceiver**

* Общая архитектура модели состоит из двух частей:
  + **Universal Speech Model (USM)**: предварительно обученная модель для генерации речевых признаков.
  + **Perceiver**: использует перекрестное внимание для сжатия последовательностей в фиксированный набор классовых латентных представлений.
* Преимущества предложенной архитектуры:
  + Латентные переменные специфичны для классов, что позволяет анализировать разные регионы сигнала для разных меток.
  + Факторизованная проекция захватывает зависимости во времени и признаках, улучшая классификацию.
* Сравнение с базовыми моделями:
  + Точность модели: **83.1%**, превосходя трансформеры (80.9%) и стандартный Perceiver (81.8%).

**4.2. Кодировки USM и цель предобучения**

* Использование USM обеспечивает значительное улучшение точности по сравнению с лог-мел-признаками:
  + Без предобучения: **76.4%**.
  + Предобучение на основе BEST-RQ: **82.0%**.
  + Предобучение для распознавания речи (ASR): **83.1%**.
* Неожиданно оказалось, что ASR-оптимизированные кодировки дают лучшие результаты для медицинских задач, несмотря на их различие с задачами распознавания речи.

**4.3. Кодировки разных слоев**

* Анализ показал, что слои USM различаются по эффективности:
  + Средние слои обеспечивают лучшую точность (до **83.1%**) за счет баланса акустической и фонетической информации.
  + Низкие слои лучше для акустических характеристик, а высокие — для фонетических.

**4.4. Объединение задач**

* Объединение данных из разных задач (например, продление гласных и AMR) показало улучшение точности с **83.6%** до **84.1%**.
* Совместная модель для всех задач упрощает развертывание и использование в клинической практике.

**Основные выводы**: Предложенная архитектура на основе Perceiver, в сочетании с USM, превосходит базовые модели. Использование предобучения и латентных переменных, специфичных для классов, значительно улучшает точность классификации. Дальнейшее развитие модели может включать адаптацию кодировок для разных меток и задач.

Авторы представили модель, которая способна предсказывать аномалии речи, важные для диагностики неврологических заболеваний. Основные достижения работы:

1. **Новая архитектура**:
   * Предложена факторизованная архитектура на основе Perceiver, которая преобразует длинные аудиопоследовательности в фиксированный набор классовых латентных представлений. Это позволяет учитывать разные регионы сигнала для разных меток.
2. **Высокая точность**:
   * Модель показала лучшую производительность (**83.1% точности**) по сравнению с трансформерами (80.9%) и стандартным Perceiver (81.8%).
3. **Эффективность USM**:
   * Использование Universal Speech Model, обученной для задач распознавания речи (ASR), значительно улучшило точность классификации. Это показывает, что предобученные модели могут быть эффективны даже для задач, отличающихся от их основной цели.
4. **Инсайты о слоях USM**:
   * Средние слои USM обеспечивают лучший баланс акустической и фонетической информации, что дает максимальную точность предсказаний.
5. **Практическая значимость**:
   * Модель способна работать с ограниченными данными, что делает её подходящей для использования в клинической практике.

**Направления для дальнейших исследований:**

1. Изучение представлений разных слоев USM для дальнейшей оптимизации классификации.
2. Разработка методов для адаптации модели к более широкому спектру аномалий и задач.

**Заключение**: Разработанная модель открывает новые возможности для автоматизации диагностики неврологических заболеваний, предоставляя эффективный инструмент для выявления речевых аномалий в клинических условиях.

# СПИСОК ИСТОЧНИКОВ

1. MarketsandMarkets. Artificial Intelligence in Healthcare Market - Global Forecast to 2028.<https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-healthcare-market-54679303.html>
2. The Lancet Digital Health. The role of AI in improving diagnostic accuracy: Challenges and opportunities. https://arxiv.org/abs/2306.07999
3. World Health Organization. [Global strategy on human resources for health: Workforce 2030](https://www.who.int/publications/i/item/9789241511131)
4. Natural Language Processing Applications in Healthcare: Opportunities and Challenges // AccScience : электр. версия. 2023. URL: https://accscience.com/journal/AIH/1/1/10.36922/aih.2147 (дата обращения: 23.12.2024). — Дата публикации: 15.02.2023.
5. Applications of Computer Vision in Medical Image Analysis // OAE Publish : науч.-техн. журн. : электр. версия. 2021. Т. 3, № 15. URL: https://www.oaepublish.com/articles/ais.2021.15 (дата обращения: 23.12.2024). — Дата публикации: 01.08.2021.
6. The Role of Automatic Speech Recognition in Detecting Neurological Disorders // MDPI : науч.-техн. журн. : электр. версия. 2022. Т. 13, № 10. URL: https://www.mdpi.com/2078-2489/13/10/499 (дата обращения: 23.12.2024). — Дата публикации: 10.10.2022.
7. Automation of a problem list using natural language processing // PubMed Central : электр. версия. URL:<https://pmc.ncbi.nlm.nih.gov/articles/PMC1208893/> (дата обращения: 20.12.2024). — Дата публикации: 05.12.2005.
8. Weed LL. Medical records that guide and teach // New England Journal of Medicine : науч.-техн. журн. : электр. версия. 1968. Т. 278, № 593–600. DOI: <https://doi.org/10.1056/NEJM196803142781105> (дата обращения: 23.12.2024). — Дата публикации: 14.03.1968.
9. Weed LL. Medical records that guide and teach // New England Journal of Medicine : науч.-техн. журн. : электр. версия. 1968. Т. 278, № 652–657. DOI: <https://doi.org/10.1056/NEJM196803212781204> (дата обращения: 23.12.2024). — Дата публикации: 21.03.1968.
10. Scherpbier HJ, Abrams RS, Roth DH, Hail JJ. A simple approach to physician entry of patient problem list // Proceedings of the Annual Symposium on Computer Applications in Medical Care : науч.-техн. журн. : электр. версия. 1994. С. 206–210. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2247883/> (дата обращения: 23.12.2024). — Дата публикации: 1994.
11. Wasserman H, Wang J. An Applied Evaluation of SNOMED CT as a Clinical Vocabulary for the Computerized Diagnosis and Problem List // Proceedings of the AMIA Symposium : науч.-техн. журн. : электр. версия. 2003. С. 699–703. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1479891/> (дата обращения: 23.12.2024). — Дата публикации: 2003.
12. Unified Medical Language System (UMLS) // Национальная медицинская библиотека США : электр. ресурс. URL:<https://www.nlm.nih.gov/research/umls/index.html> (дата обращения: 23.12.2024). — Дата публикации: 1993.
13. Payne T, Martin DR. How useful is the UMLS metathesaurus in developing a controlled vocabulary for an automated problem list? // Proceedings of the Annual Symposium on Computer Applications in Medical Care : науч.-техн. журн. : электр. версия. 1993. С. 705–709. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2247681/> (дата обращения: 23.12.2024). — Дата публикации: 1993.
14. Goldberg H, Hsu C, Law V, Safran C. Validation of clinical problems using a UMLS-based semantic parser // Proceedings of the AMIA Symposium : науч.-техн. журн. : электр. версия. 1998. С. 805–809. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2232212/> (дата обращения: 23.12.2024). — Дата публикации: 1998.
15. Zelingher J, Rind DM, Caraballo E, Tuttle M, Olson N, Safran C. Categorization of free-text problem lists: an effective method of capturing clinical data // Proceedings of the Annual Symposium on Computer Applications in Medical Care : науч.-техн. журн. : электр. версия. 1995. С. 416–420. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2579060/> (дата обращения: 23.12.2024). — Дата публикации: 1995.
16. Spyns P. Natural language processing in medicine: an overview // Methods of Information in Medicine : науч.-техн. журн. : электр. версия. 1996. Т. 35. С. 285–301. URL:<https://pubmed.ncbi.nlm.nih.gov/8960545/> (дата обращения: 23.12.2024). — Дата публикации: 1996.
17. Chi E, Lyman M, Sager N, Friedman C. Database of computer-structured narrative: methods of computing complex relations // IEEE, редактор. SCAMC 85 : науч.-техн. конф. : электр. версия. 1985. С. 221–226. URL: доступ к статье недоступен в открытом виде (дата обращения: 23.12.2024). — Дата публикации: 1985.
18. Sager N, Friedman C, Chi E. The analysis and processing of clinical narrative // Medinfo 86 : науч.-техн. конф. : электр. версия. Амстердам (Голландия). 1986. С. 1101–1105. URL: доступ к статье недоступен в открытом виде (дата обращения: 23.12.2024). — Дата публикации: 1986.
19. Zingmond D, Lenert LA. Monitoring free-text data using medical language processing // Computers and Biomedical Research : науч.-техн. журн. : электр. версия. 1993. Т. 26. С. 467–481. DOI: https://doi.org/10.1006/cbmr.1993.1033 (дата обращения: 23.12.2024). — Дата публикации: 1993.
20. Friedman C, Alderson PO, Austin JH, Cimino JJ, Johnson SB. A general natural-language text processor for clinical radiology // Journal of the American Medical Informatics Association : науч.-техн. журн. : электр. версия. 1994. Т. 1. С. 161–174. DOI: https://doi.org/10.1136/jamia.1994.95236146 (дата обращения: 23.12.2024). — Дата публикации: 1994.
21. Friedman C, Hripcsak G, Shagina L, Liu H. Representing information in patient reports using natural language processing and the extensible markup language // Journal of the American Medical Informatics Association : науч.-техн. журн. : электр. версия. 1999. Т. 6. С. 76–87. DOI: https://doi.org/10.1136/jamia.1999.0060076 (дата обращения: 23.12.2024). — Дата публикации: 1999.
22. Hripcsak G, Friedman C, Alderson PO, DuMouchel W, Johnson SB, Clayton PD. Unlocking clinical data from narrative reports: a study of natural language processing // Annals of Internal Medicine : науч.-техн. журн. : электр. версия. 1995. Т. 122. С. 681–688. DOI: https://doi.org/10.7326/0003-4819-122-9-199505010-00007 (дата обращения: 23.12.2024). — Дата публикации: 1995.
23. Hripcsak G, Kuperman GJ, Friedman C. Extracting findings from narrative reports: software transferability and sources of physician disagreement // Methods of Information in Medicine : науч.-техн. журн. : электр. версия. 1998. Т. 37. С. 1–7. URL: доступ к статье недоступен в открытом виде (дата обращения: 23.12.2024). — Дата публикации: 1998.
24. Hripcsak G, Kuperman GJ, Friedman C. Extracting findings from narrative reports: software transferability and sources of physician disagreement // Methods of Information in Medicine : науч.-техн. журн. : электр. версия. 1998. Т. 37. С. 1–7. URL: доступ к статье недоступен в открытом виде (дата обращения: 23.12.2024). — Дата публикации: 1998.
25. Jain NL, Friedman C. Identification of findings suspicious for breast cancer based on natural language processing of mammogram reports // Proceedings of the AMIA Annual Fall Symposium : науч.-техн. журн. : электр. версия. 1997. С. 829–833. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2233454/> (дата обращения: 23.12.2024). — Дата публикации: 1997.
26. Elkins JS, Friedman C, Boden-Albala B, Sacco RL, Hripcsak G. Coding neuroradiology reports for the Northern Manhattan Stroke Study: a comparison of natural language processing and manual review // Computers and Biomedical Research : науч.-техн. журн. : электр. версия. 2000. Т. 33. С. 1–10. DOI: https://doi.org/10.1006/cbmr.1999.1535 (дата обращения: 23.12.2024). — Дата публикации: 2000.
27. Friedman C, Knirsch C, Shagina L, Hripcsak G. Automating a severity score guideline for community-acquired pneumonia employing medical language processing of discharge summaries // Proceedings of the AMIA Symposium : науч.-техн. журн. : электр. версия. 1999. С. 256–260. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2232517/> (дата обращения: 23.12.2024). — Дата публикации: 1999.
28. Xu H, Friedman C. Facilitating Research in Pathology using Natural Language Processing // Proceedings of the AMIA Symposium : науч.-техн. журн. : электр. версия. 2003. С. 1057. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1480165/> (дата обращения: 23.12.2024). — Дата публикации: 2003.
29. Friedman C, Johnson SB, Forman B, Starren J. Architectural requirements for a multipurpose natural language processor in the clinical environment // Proceedings of the Annual Symposium on Computer Applications in Medical Care : науч.-техн. журн. : электр. версия. 1995. С. 347–351. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2578982/> (дата обращения: 23.12.2024). — Дата публикации: 1995.
30. Friedman C, Shagina L, Lussier Y, Hripcsak G. Automated Encoding of Clinical Documents Based on Natural Language Processing // Journal of the American Medical Informatics Association : науч.-техн. журн. : электр. версия. 2004. DOI: <https://doi.org/10.1197/jamia.M1552> (дата обращения: 23.12.2024). — Дата публикации: 2004.
31. Ranum DL. Knowledge-based understanding of radiology texts // Computer Methods and Programs in Biomedicine : науч.-техн. журн. : электр. версия. 1989. Т. 30. С. 209–215. DOI: <https://doi.org/10.1016/0169-2607(89)90073-4> (дата обращения: 23.12.2024). — Дата публикации: 1989.
32. Haug PJ, Ranum DL, Frederick PR. Computerized extraction of coded findings from free-text radiologic reports. Work in progress // Radiology : науч.-техн. журн. : электр. версия. 1990. Т. 174. С. 543–548. DOI: <https://doi.org/10.1148/radiology.174.2.2404321> (дата обращения: 23.12.2024). — Дата публикации: 1990.
33. Haug P, Koehler S, Lau LM, Wang P, Rocha R, Huff S. A natural language understanding system combining syntactic and semantic techniques // Proceedings of the Annual Symposium on Computer Applications in Medical Care : науч.-техн. журн. : электр. версия. 1994. С. 247–251. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2247867/> (дата обращения: 23.12.2024). — Дата публикации: 1994.
34. Haug PJ, Koehler S, Lau LM, Wang P, Rocha R, Huff SM. Experience with a mixed semantic/syntactic parser // Proceedings of the Annual Symposium on Computer Applications in Medical Care : науч.-техн. журн. : электр. версия. 1995. С. 284–288. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2578971/> (дата обращения: 23.12.2024). — Дата публикации: 1995.
35. Koehler SB. SymText: a natural language understanding system for encoding free text medical data // Диссертация : электр. версия. 1998. URL: доступ к полной версии отсутствует в открытом доступе (дата обращения: 23.12.2024). — Дата публикации: 1998.
36. Christensen L, Haug P, Fiszman M. MPLUS: a probabilistic medical language understanding system // Proceedings of the Workshop on Natural Language Processing in the Biomedical Domain : науч.-техн. конф. : электр. версия. 2002. С. 29–36. URL: доступ к статье отсутствует в открытом доступе (дата обращения: 23.12.2024). — Дата публикации: 2002.
37. Nivre J. On Statistical Methods in Natural Language Processing // In: Bubenko JjW, Benkt, editor. Promote IT: Second Conference for the Promotion of Research in IT at New Universities and University Colleges in Sweden : науч.-техн. конф. : электр. версия. Скёвде (Швеция): University of Skövde. 2002. С. 684–694. URL: доступ к полной версии отсутствует в открытом доступе (дата обращения: 23.12.2024). — Дата публикации: 2002.
38. Fiszman M, Chapman WW, Aronsky D, Evans RS, Haug PJ. Automatic detection of acute bacterial pneumonia from chest X-ray reports // Journal of the American Medical Informatics Association : науч.-техн. журн. : электр. версия. 2000. Т. 7. С. 593–604. DOI: <https://doi.org/10.1136/jamia.2000.0070593> (дата обращения: 23.12.2024). — Дата публикации: 2000.
39. Fiszman M, Haug PJ. Using medical language processing to support real-time evaluation of pneumonia guidelines // Proceedings of the AMIA Symposium : науч.-техн. журн. : электр. версия. 2000. С. 235–239. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2243975/> (дата обращения: 23.12.2024). — Дата публикации: 2000.
40. Haug PJ, Christensen L, Gundersen M, Clemons B, Koehler S, Bauer K. A natural language parsing system for encoding admitting diagnoses // Proceedings of the AMIA Annual Fall Symposium : науч.-техн. журн. : электр. версия. 1997. С. 814–818. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2233449/> (дата обращения: 23.12.2024). — Дата публикации: 1997.
41. Fiszman M, Haug PJ, Frederick PR. Automatic extraction of PIOPED interpretations from ventilation/perfusion lung scan reports // Proceedings of the AMIA Symposium : науч.-техн. журн. : электр. версия. 1998. С. 860–864. URL:<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2232298/> (дата обращения: 23.12.2024). — Дата публикации: 1998.
42. COVID‑Net: a tailored deep convolutional neural network design for detection of COVID‑19 cases from chest X‑ray images // *Scientific Reports* : науч.-техн. журн. : электр. версия. 2020. Т. 10, № 19549. DOI:<https://doi.org/10.1038/s41598-020-76550-z> (дата обращения: 23.12.2024). — Дата публикации: 2020.
43. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China // The Lancet. 2020;395(10223):497–506. DOI: [https://doi.org/10.1016/S0140-6736(20)30183-5.](https://doi.org/10.1016/S0140-6736(20)30183-5)
44. Guan W, Ni Z, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China // New England Journal of Medicine. 2020;382(18):1708–1720. DOI: [https://doi.org/10.1056/NEJMoa2002032.](https://doi.org/10.1056/NEJMoa2002032)
45. Wang, W. et al. Detection of SARS-CoV-2 in different types of clinical specimens // JAMA : науч.-техн. журн. : электр. версия. 2020. Т. 323, № 18. С. 1843–1844. DOI: <https://doi.org/10.1001/jama.2020.3786> (дата обращения: 23.12.2024). — Дата публикации: 2020.
46. Ng, M.-Y. et al. Imaging profile of the COVID-19 infection: radiologic findings and literature review // Radiology: Cardiothoracic Imaging : науч.-техн. журн. : электр. версия. 2020. Т. 2, № 1. Статья e200034. DOI: <https://doi.org/10.1148/ryct.2020200034> (дата обращения: 23.12.2024). — Дата публикации: 2020.
47. Assessing the Effectiveness of Automatic Speech Recognition Technology in Emergency Medicine Settings: A Comparative Study of Four AI-powered Engines // Research Square : электр. версия. 2023. URL:<https://www.researchsquare.com/article/rs-4727659/v1> (дата обращения: 20.12.2024). — Дата публикации: 18.09.2023.
48. Park SY, Lee SY, Chen Y. The effects of EMR deployment on doctors’ work practices: A qualitative study in the emergency department of a teaching hospital // International Journal of Medical Informatics : науч.-техн. журн. : электр. версия. 2012. Т. 81, № 3. С. 204–217. DOI: https://doi.org/10.1016/j.ijmedinf.2011.12.001 (дата обращения: 23.12.2024). — Дата публикации: 2012.
49. van Buchem MM, Boosman H, Bauer MP, Kant IM, Cammel SA, Steyerberg EW. The digital scribe in clinical practice: a scoping review and research agenda // NPJ Digital Medicine : науч.-техн. журн. : электр. версия. 2021. Т. 4. С. 1. DOI: https://doi.org/10.1038/s41746-021-00454-3 (дата обращения: 23.12.2024). — Дата публикации: 2021.
50. Shafran I, Du N, Tran L, Perry A, Keyes L, Knichel M, Domin A, Huang L, Chen Y-h, Li G. The Medical Scribe: Corpus Development and Model Performance Analyses // In Proceedings of the Twelfth Language Resources and Evaluation Conference : науч.-техн. конф. : электр. версия. 2020. С. 2036–2044. URL: https://aclanthology.org/2020.lrec-1.251/ (дата обращения: 23.12.2024). — Дата публикации: 2020.
51. Nuance Announces the General Availability of Ambient Clinical Intelligence // Nuance : электр. ресурс. 2020. URL:<https://news.nuance.com/2020-02-24-Nuance-Announces-the-General-Availability-of-Ambient-Clinical-Intelligence> (дата обращения: 23.12.2024). — Дата публикации: 24.02.2020.
52. Amazon Comprehend Medical // Amazon Web Services : электр. ресурс. URL:<https://aws.amazon.com/comprehend/medical/> (дата обращения: 23.12.2024). — Дата публикации: информация отсутствует.
53. Tran BD, Latif K, Reynolds TL, Park J, Elston Lafata J, Tai-Seale M, Zheng K. Mm-hm, Uh-uh: are non-lexical conversational sounds deal breakers for the ambient clinical documentation technology? // Journal of the American Medical Informatics Association : науч.-техн. журн. : электр. версия. 2023. Т. 30. С. 703–711. DOI: <https://doi.org/10.1093/jamia/ocac270> (дата обращения: 23.12.2024). — Дата публикации: 2023.
54. Zhang Z, Joy K, Harris R, Park SY. Characteristics and challenges of clinical documentation in self-organized fast-paced medical work // Proceedings of the ACM on Human-Computer Interaction : науч.-техн. журн. : электр. версия. 2022. Т. 6, № CSCW2. С. 1–21. DOI: <https://doi.org/10.1145/3555541> (дата обращения: 23.12.2024). — Дата публикации: 2022.
55. Pilerot O, Maurin Söderholm H. A conceptual framework for investigating documentary practices in prehospital emergency care // *Proceedings of the Association for Information Science and Technology* : науч.-техн. журн. : электр. версия. 2019. Т. 56, № 1. С. 587–589. DOI:<https://doi.org/10.1002/pra2.91> (дата обращения: 23.12.2024). — Дата публикации: 2019.Google Speech to Text // Google Cloud : электр. ресурс. URL:<https://cloud.google.com/speech-to-text?hl=en> (дата обращения: 23.12.2024). — Дата публикации: информация отсутствует.
56. Introducing Whisper - OpenAI // OpenAI : электр. ресурс. URL:<https://openai.com/index/whisper> (дата обращения: 23.12.2024). — Дата публикации: информация отсутствует.
57. Azure Speech to Text // Microsoft Azure : электр. ресурс. URL:<https://azure.microsoft.com/en-us/products/ai-services/speech-to-text> (дата обращения: 23.12.2024). — Дата публикации: информация отсутствует.
58. Zhang Z, Joy K, Upadhyayula P, Ozkaynak M, Harris R, Adelgais K. Data work and decision making in emergency medical services: a distributed cognition perspective // Proceedings of the ACM on Human-Computer Interaction : науч.-техн. журн. : электр. версия. 2021. Т. 5, № CSCW2. С. 1–32. DOI: https://doi.org/10.1145/3479564 (дата обращения: 23.12.2024). — Дата публикации: 2021.
59. Ozkaynak M, Dolen C, Dollin Y, Rappaport K, Adelgais K. Simulating teamwork for better decision making in pediatric emergency medical services // AMIA Annual Symposium Proceedings : науч.-техн. конф. : электр. версия. American Medical Informatics Association. 2020. С. 993. URL: <https://knowledge.amia.org/69810-1> (дата обращения: 23.12.2024). — Дата публикации: 2020.
60. NEMSIS Data Dictionary EMS Data Standard // NEMSIS : электр. ресурс. URL:<https://nemsis.org/media/nemsis_v3/release-3.5.0/DataDictionary/PDFHTML/EMSDEMSTATE/index.html> (дата обращения: 23.12.2024). — Дата публикации: информация отсутствует.
61. Yujian L, Bo L. A normalized Levenshtein distance metric // IEEE Transactions on Pattern Analysis and Machine Intelligence : науч.-техн. журн. : электр. версия. 2007. Т. 29. С. 1091–1095. DOI: <https://doi.org/10.1109/TPAMI.2007.1078> (дата обращения: 23.12.2024). — Дата публикации: 2007.
62. Şen Z. Fuzzy string matching procedure // Open Bioinformatics Journal : науч.-техн. журн. : электр. версия. 2020. Т. 13. С. 1. DOI: <https://doi.org/10.2174/1875036202013010001> (дата обращения: 23.12.2024). — Дата публикации: 2020.
63. Bosker HR. Using fuzzy string matching for automated assessment of listener transcripts in speech intelligibility studies // Behavior Research Methods : науч.-техн. журн. : электр. версия. 2021. Т. 53, № 5. С. 1945–1953. DOI: <https://doi.org/10.3758/s13428-020-01502-4> (дата обращения: 23.12.2024). — Дата публикации: 2021.
64. Reimers N, Gurevych I. Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks // Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) : науч.-техн. конф. : электр. версия. 2019. С. 3982–3992. DOI: <https://doi.org/10.18653/v1/D19-1410> (дата обращения: 23.12.2024). — Дата публикации: 2019.
65. Wu Y, Jiang M, Xu J, Zhi D, Xu H. Clinical named entity recognition using deep learning models // AMIA Annual Symposium Proceedings : науч.-техн. конф. : электр. версия. American Medical Informatics Association. 2017. С. 1812. URL: <https://knowledge.amia.org/69812-1> (дата обращения: 23.12.2024). — Дата публикации: 2017.
66. Luo X, Gandhi P, Storey S, Huang K. A Deep Language Model for Symptom Extraction from Clinical Text and Its Application to Extract COVID-19 symptoms from Social Media // IEEE Journal of Biomedical and Health Informatics : науч.-техн. журн. : электр. версия. 2021. Т. 25, № 4. С. 1246–1255. DOI: <https://doi.org/10.1109/JBHI.2021.3059294> (дата обращения: 23.12.2024). — Дата публикации: 2021.
67. Navarro DF, Ijaz K, Rezazadegan D, Rahimi-Ardabili H, Dras M, Coiera E, Berkovsky S. Clinical named entity recognition and relation extraction using natural language processing of medical free text: A systematic review // International Journal of Medical Informatics : науч.-техн. журн. : электр. версия. 2023. Статья 105122. DOI: <https://doi.org/10.1016/j.ijmedinf.2023.105122> (дата обращения: 23.12.2024). — Дата публикации: 2023.
68. Park SY, Chen Y. Adaptation as design: learning from an EMR deployment study // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems : науч.-техн. конф. : электр. версия. 2012. С. 2097–2106. DOI: <https://doi.org/10.1145/2207676.2208351> (дата обращения: 23.12.2024). — Дата публикации: 2012.