

**В.В. ГРИБОВА,**

д.т.н., Институт автоматки и процессов управления ДВО РАН, г. Владивосток, Россия,
e-mail: gribova@dvo.ru

Л.А. ФЕДОРИЩЕВ,

к.т.н., Институт автоматки и процессов управления ДВО РАН, г. Владивосток, Россия,
e-mail: fleo1987@mail.ru

ОБЛАЧНЫЙ СЕРВИС ДЛЯ ФОРМИРОВАНИЯ ФОРМАЛИЗОВАННЫХ ИСТОРИЙ БОЛЕЗНИ

УДК: 004.514.6

DOI: 10.37690/1811-0193-2020-S1-51-57

Грибова В.В., Федорищев Л.А. *Облачный сервис для формирования формализованных историй болезни* (Институт автоматки и процессов управления ДВО РАН, г. Владивосток, Россия)

Аннотация. В работе представлен облачный сервис для формирования историй болезней, особенностью которого является формальное представление всех введенных данных о пациенте. Формальное представление истории болезни необходимо для дальнейшей обработки данных пациента интеллектуальными сервисами поддержки принятия решений, а также получения новых знаний о диагностике и лечении пациентов на основе методов машинного обучения. Сервис реализован на медицинском портале облачной платформы IACPaaS. Представлены основные принципы разработки, архитектура и интерфейс программного средства. Описана технология использования сервиса.

Ключевые слова: медицинский облачный сервис, онтология, редактор, знания, база знаний, диагностика заболеваний.

UDC: 004.514.6

Gribova V.V., Fedorishev L.A. *Cloud service for the formation of formalized case records* (Institute of Automation and Control Processes Far Eastern Branch of RAS, Vladivostok, Russia)

Abstract. The paper presents a cloud service for the formation of case records, the feature of which is a formal presentation of all entered patient data. A formal presentation of the case record is necessary for further processing of patient data by intelligent decision support services, as well as for obtaining new knowledge about the diagnosis and treatment of patients based on machine learning methods. The service is implemented on the medical portal of the IACPaaS cloud platform. The basic principles of development, architecture and software interface are presented. The technology of using the service is described.

Keywords: medical cloud service, ontology, editor, knowledge, knowledge base, diagnosis of diseases.

ВВЕДЕНИЕ

Развитие информационных технологий в медицине идет высокими темпами, и в настоящее время большинство медицинских учреждений имеет программное обеспечение, связанное со специализированным медицинским оборудованием, системы поддержки принятия решений, медицинские информационные системы, которые помогают в управлении медицинским учреждением, создании отчетов, ведении историй болезни [1–5]. Вместе с тем необходимо констатировать, что внедрение в повседневную деятельность врачей интеллектуальных систем поддержки принятия решений значительно тормозится вследствие того, что данные о пациенте в формируемых историях болезни либо не формализованы, либо формализованы частично (как правило, это паспортная информация о пациенте, метрики пациента и ряд других признаков, имеющих простую структуру). Вместе с тем наиболее важная и информативная часть, связанная с жалобами пациента, его объективным осмотром, анамнезом заболевания, результатами исследований представлена в неформализованном текстовом виде (для формирования этой информации разработчики предлагают вводить данные в текстовые поля), что затрудняет их дальнейшее использование интеллектуальными сервисами поддержки принятия решений. Эта ситу-



ация также значительно тормозит как разработку самих интеллектуальных сервисов, так и анализ данных методами машинного обучения для получения новых научных знаний о диагностике, лечении, прогнозе заболеваний, и др. Основная причина такой ситуации – огромные объемы медицинских данных (десятки и сотни тысяч) – жалобы, данные объективного исследования, лабораторных и инструментальных исследований, история жизни, настоящего заболевания и др., которые необходимы в истории болезни пациента для правильной диагностики и назначении лечения. Наличие такого объема данных, во-первых, значительно затрудняет разработку пользовательского интерфейса таких систем, особенно с учетом требований юзабилити¹, во-вторых, даже при частичной формализации данных (как правило, для ввода узко специализированных историй болезни), приводит к неэффективной работе врачей, вследствие необходимости навигации по большой и неудобной структуре данных. В результате текстовый ввод данных является единственно приемлемым для таких систем.

Учитывая важность и актуальность формализованных историй болезни для дальнейшего развития медицины в целом, актуальной проблемой является разработка методов создания сервисов их формализованного ввода.

Целью работы является описание подхода, методов реализации и сервисов на облачной платформе IASaaS [6] для формализованного ввода историй болезни.

МЕТОДЫ

Основные проблемы, с которыми сталкиваются разработчики электронных историй болезни или электронных медицинских карт, – это объем медицинских данных, постоянно пополняемый новыми данными (медицина – это область знаний, которая активно развивается: появляются новые инструментальные, лабораторные исследования, действующие вещества и многое другое). Соответственно, реализовать пользовательский интерфейс истории болезни или электронной медицинской карты (ЭМК), «покрывающий» все разнообразие возможных данных очень сложно:

¹ **Юзабилити** – способность продукта быть понимаемым, используемым и привлекательным для пользователя в заданных условиях (ISO/IEC25010)[2]; свойство системы, продукта или услуги, при наличии которого конкретный пользователь может эксплуатировать систему в определенных условиях для достижения установленных целей с необходимой результативностью, эффективностью и удовлетворенностью (ISO 9241–210).

каждый элемент данных необходимо описать, выбрать интерфейсный элемент(ы) для его реализации, расположить на экране монитора. Учитывая, что таких данных десятки тысяч, количество которых постоянно растет, задача формализации истории болезни или ЭМК (далее для удобства будем использовать термин ЭМК) становится практически невыполнимой. Также немаловажной проблемой является время ввода информации (известно, что врачи очень ограничены временем, выделяемым им на прием пациента), поэтому ввод данных не должен занимать много времени, при этом используемая терминология должна быть согласована с различными системами обработки данных. Указанные проблемы определяют основные требования к системам такого класса:

- обеспечение удобства и простоты ввода, его интуитивная понятность;
- использование единой, разделяемой медицинским сообществом терминологии (допускающей использование синонимов);
- универсализм реализации (поскольку реализация таких систем является сложной и трудоемкой проблемой, занимающей многие человеко-месяцы) с одной стороны, с другой – легкая настройка на конкретные требования к представлению ЭМК (известно, что несмотря на введение общего стандарта, ЭМК часто ориентированы на конкретные медицинские учреждения, отражающие их специфику).

Для реализации указанных требований предлагается использовать двухуровневый онтологический подход для формирования всех компонентов системы [7–8]. Его особенность заключается в том, что на первом этапе формируется модель (или онтология) информационного ресурса, затем по нему формируется информационный ресурс. При этом онтология явно отделена от информационного ресурса, но формируемые по ней информационные ресурсы связаны с терминами онтологии. Для формализованного ввода ЭМК используются следующие онтологии: онтология медицинской терминологии и наблюдений и онтология ЭМК.

Ключевым компонентом, обеспечивающим формализацию ЭМК, является *онтология медицинской терминологии и наблюдений* [9], в соответствии с которой формируется *база медицинских терминов и наблюдений*. В настоящее время общепризнано, что необходим общий, разделяемый медицинским сообществом терминологический словарь, который должен использоваться при формировании

медицинской документации различных типов и назначения. В России и за рубежом активно ведутся работы по формированию таких медицинских справочников, например, Национальная База Медицинских Знаний (НБМЗ), Хеликс, United Medical Knowledge Base (УМКВ), Snomed и другие. Аналогичная работа проводится и в коллективе, в котором работают авторы. Принципиальное отличие – терминологический справочник создается на основе онтологии, которая задает типы и структуру медицинской терминологии. Так, онтология включает следующие базовые понятия: признаки (симптомы), факторы, вспомогательные термины и нечеткие шкалы. Все базовые понятия объединяются в логически связанные группы–множества терминов, наблюдений.

В разделе **Термины** описаны основные разделы медицины: морфология, физиология, патология, фармакология, патогенные факторы, методы лечения, единицы измерения и др., которые являются *группами терминов*. Каждая группа терминов включает подгруппы терминов, например, группа терминов *морфология* включает: *анатомию человека* и *цитологию*, группа терминов *анатомия человека* включает группы терминов: *части тела*, *анатомические плоскости*, *описание внутренних органов* и т.д.

События обозначают понятия, соответствующие событиям, которые могут происходить с пациентом до заболевания и являются причиной заболевания. Они описаны в терминах *группа событий* и *событие*.

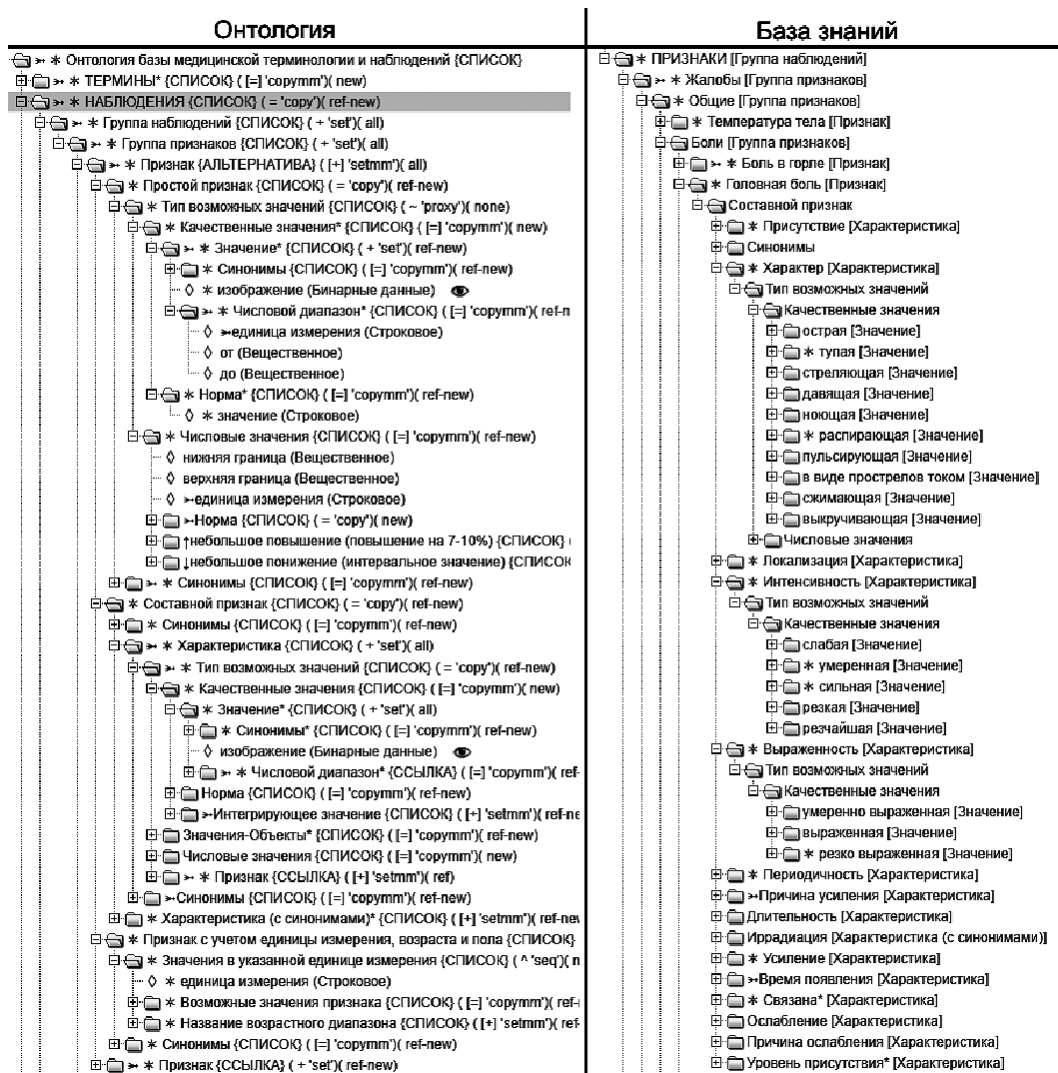


Рис. 1. Скриншот фрагмента онтологии (слева) и базы медицинской терминологии и наблюдений (справа)

Факторы – это класс понятий, соответствующий внешним условиям и анатомо-физиологическим особенностям пациента, которые могут повлиять на течение и исход заболевания. Они описаны в терминах *группа факторов, фактор*.

Признаками заболеваний являются: жалобы, данные объективного исследования, лабораторные и инструментальные исследования, они описаны в онтологических терминах: *группа признаков, признак, характеристика, значение*.

Признаки могут быть *простыми* – с качественными, числовыми или интервальными значениями, *составными*, описываемыми характеристиками (с аналогичными вариантами типов значений) и *признаками с учетом единицы измерения, возраста, пола, беременности, спорта и других особенностей*, которые введены для возможности описывать референсные значения лабораторных и инструментальных исследований. Для единообразного представления и однозначного толкования участниками разработки и пользователями формализуемых знаний требуется свод не только всех используемых на практике терминов, названий наблюдений, всех возможных их значений, но и их распространенных синонимов, употребляемых при заполнении историй болезней, поскольку медицинской терминологии свойственна «подвижность лексического состава» и его постоянное развитие. Термин *синоним* введен для групп признаков, признаков и качественных значений признаков

На основе описанной выше онтологии на платформе IASaaS разработана *База знаний*

медицинской терминологии и наблюдений, включающая несколько десятков тысяч различных элементов. Эта база является общей универсальной базой терминов и наблюдений, которая может использоваться в медицинских сервисах различного назначения. На *рис. 1* представлен фрагмент онтологии, описывающей «Составной признак», и построенный по онтологии признак «Головная боль» в терминологии наблюдений.

Онтология ЭМК, необходимая для представления ее структуры, по которой будут формироваться ЭМК пациентов, содержит следующие основные структурные элементы: паспортные данные пациента, история заболевания, жалобы при поступлении, история настоящего заболевания, история жизни, объективное состояние больного, лабораторные и инструментальные исследования и др. Каждый основной структурный элемент ЭМК определяется датой и множествами признаков, событий или факторов, структура которых совпадает со структурой соответствующих элементов онтологии медицинской терминологии и наблюдений и ссылается на соответствующие фрагменты базы медицинской терминологии и наблюдений (*см. рис. 2*).

Онтология ЭМК разработана на платформе IASaaS. На *рис. 2* (слева) представлен верхний уровень онтологии ЭМК, справа – описание одного из основных структурных элементов.

Онтологический подход к созданию ЭМК и базы медицинской терминологии и наблюдений позволил разработать универсальный сервис (онтолого-ориентированный) заполнения и ведения

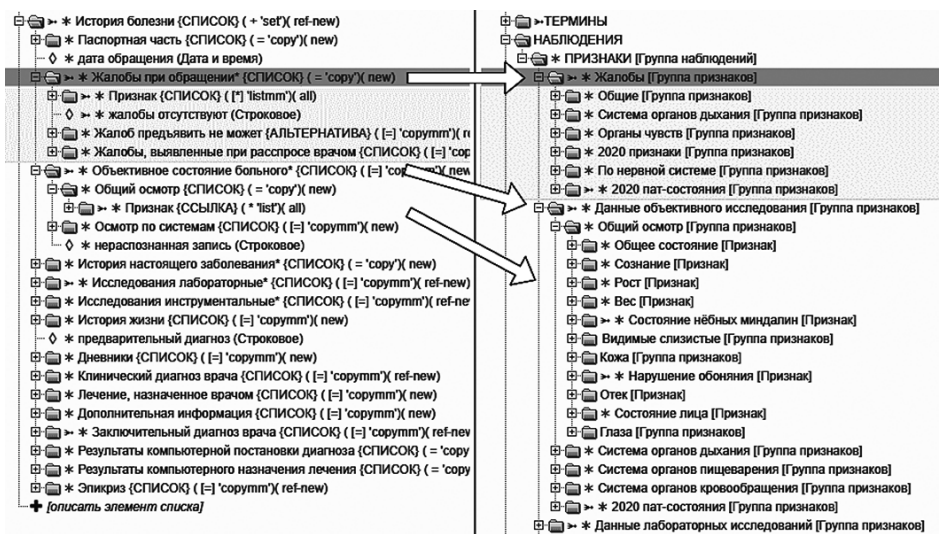


Рис. 2. Связь онтологии ЭМК и БЗ терминологии и наблюдений



ЭМК пациента независимо от набора используемых признаков (см. рис. 3). Основные принципы работы сервиса заключаются в следующем:

1. Область экрана разбивается на три фрейма: фрейм структуры (онтологии) ЭМК, фрейм поиска и отображения признаков, фрейм описания признаков. В первом фрейме в виде древовидной структуры отображается онтология ЭМК, по которой доступна удобная навигация. С каждым структурным элементом ЭМК связаны фрагменты базы терминологии и наблюдений. При выборе пользователем структурного элемента ЭМК автоматически «подгружается» часть базы терминов и наблюдений, соответствующая этому структурному элементу. Например, при выборе элемента *Жалобы ЭМК* в фрейм поиска и отображения (в алфавитном порядке) выводится возможный набор жалоб из базы терминологии и наблюдений, из которого пользователь может выбрать необходимую.

2. Фрейм поиска и отображения признаков. *База медицинской терминологии и наблюдений* включает десятки тысяч элементов. Поэтому для быстрого поиска используется метод так называемой косметической адаптации – автоматический ввод ответов: интерфейс предоставляет поле ввода ключевого слова для поиска нужного элемента. В соответствии с введенным ключевым словом (достаточно части слова) поисковая система сервиса находит все релевантные элементы в базе и отображает их пользователю в виде прокручиваемого списка. Поисковая система считает элемент релевантным, если его название или часть названия, или

его синоним содержит введенное ключевое слово. Также интерфейс поиска может быть настроен так, чтобы релевантные элементы находились и по значениям внутренних характеристик элементов.

3. Фрейм описания признаков. В соответствии с онтологией каждый признак может быть *простым* или *составным*. Значения признаков и характеристики могут иметь разные типы значений: *качественные, числовые*, для некоторых значений признаков также важны единицы измерения (например, единица измерения признаков *Ретикулоциты, Моноциты в крови* может быть % или 10⁹/л). Значения также могут быть совместными и несовместными (например, *Головная боль* может быть *умеренно выраженной, выраженной* или *резко выраженной* – это несовместные значения характеристики *выраженность*; также этот признак может иметь *характер* боли: *острая, давящая, пульсирующая* и другие – эти значения являются совместными, так как не исключают друг друга). Каждый признак (термин) формируется на основе онтологии. В соответствии с ее структурой, а также с учетом требований юзабилити генерируется интерфейс визуального представления каждого признака. Так, если число значений признака или характеристики меньше 5, то используется группа радио-кнопок (для несовместных значений) или чекбоксов (для совместных значений). Если число значений больше 5, то используется выпадающий список (например, *интенсивность* Головной боли может иметь 5 вариантов значений: *слабая, умеренная, сильная, резкая, резчайшая*, которые отображаются

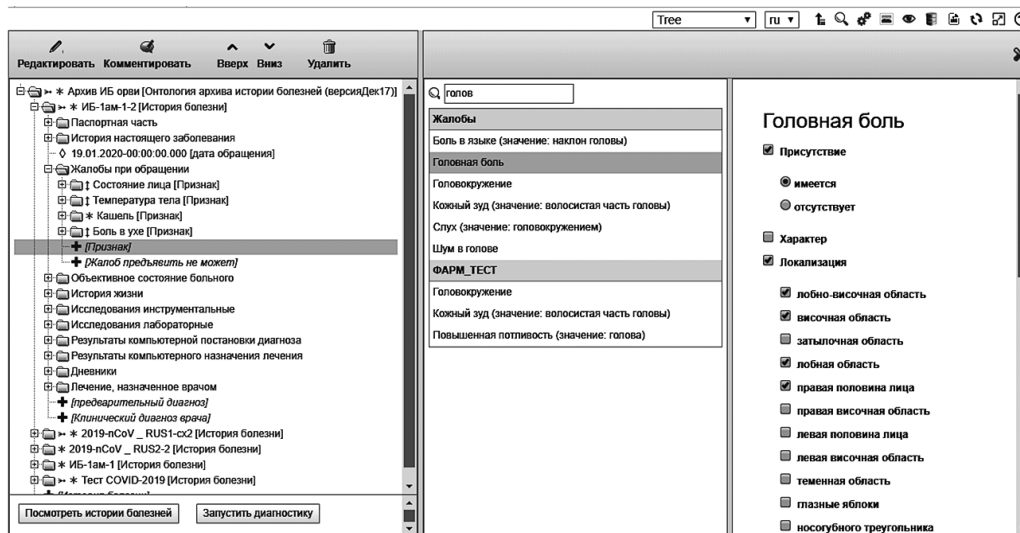


Рис. 3. Сервис ЭМК



в интерфейсе в форме выпадающего списка; а 2 варианта *присутствия* (имеется или отсутствует) отображаются в форме радиокнопок.

4. Пользователь имеет возможность просмотреть все введенные значения, добавить новые или изменить введенные.

Реализация алгоритма формирования истории болезни выполнена с использованием традиционного веб-ориентированного подхода на основе HTML и Javascript. Реализация сервиса выполнена на облачной платформе IASaaS. Возможна интеграция данного сервиса с произвольными медицинскими информационными системами.

РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

В настоящее время сервис доступен на платформе IASaaS и интегрирован с интеллектуальными сервисами по диагностике и лечению ряда заболеваний на платформе IASaaS. Перед его практическим использованием проводилось сравнение времени текстового ввода и предложенного в данной работе. Экспериментальное сравнение традиционного описания истории болезни (через текстовые поля) и предложенного в данной работе показало примерно одинаковые результаты по затраченному времени при вводе данных в историю болезни (очевидно, что ценность полностью формализованной истории болезни во много раз превышает аналогичное неформализованное представление). Однако, тут стоит учесть, что в экспериментах пользователям предлагалось ввести фиксированные фразы в текстовое поле и с использованием предлагаемого сервиса. Очевидно, что в реальной ситуации врачу для ввода неформализованного текста

необходимо дополнительное время на составление правильных фраз.

В качестве примера рассмотрим фрагмент *Жалобы* одной из историй болезни с инфарктом миокарда в текстовом виде:

Жалобы на момент поступления: интенсивная боль за грудиной давящего характера, иррадиирующая в левую лопатку, левое плечо, левую руку, возникающая при физической нагрузке, боль не купируется нитроглицерином, продолжительная. Одышка, общая слабость и потоотделение.

Описание этих же жалоб с помощью сервиса ЭМК представлено на рис. 4.

ВЫВОДЫ

Актуальной задачей для интеллектуализации медицины является формальное представление всей информации в ЭМК. Данные, представленные в формальном виде, обеспечивают возможность их интеллектуальной обработки, во-первых, системами поддержки принятия решений (для помощи врачу в диагностике заболеваний, назначении лечения, его мониторинга и прогноза); во-вторых, для обнаружения новых знаний из больших массивов таких данных. На сегодняшний день подавляющее большинство ЭМК (за исключением некоторых опросников) не поддерживает полное формальное представление всех данных (например, жалобы, история настоящего заболевания и др. описываются в текстовых полях, которые требуют отдельных методов анализа для их преобразования в формальный вид). Качество таких преобразований (из текстового представления в формальный) является неудовлетворительным [10, 11]. Основная проблема, с которой сталкиваются разработчики, – огромный объем данных,

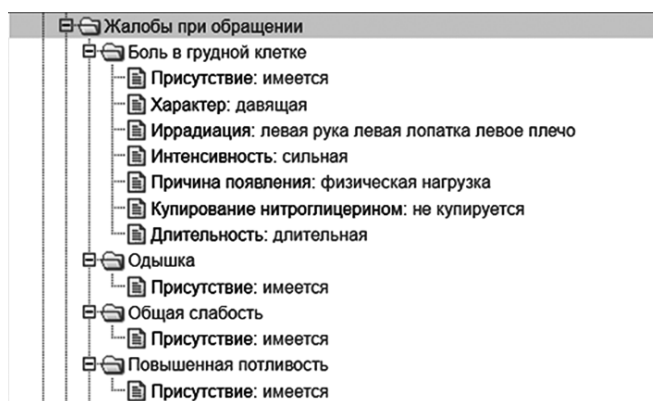


Рис. 4. Описание жалоб истории болезни



которые потенциально могут быть внесены в ЭМК, и, соответственно, сложность для программистов в реализации пользовательского интерфейса, необходимость его постоянной модификации при появлении новых признаков. Предложенный авторами онтологический подход обеспечивает автоматическую генерацию пользовательского интерфейса ЭМК. При этом изменение/дополнение терминологической базы (с помощью отдельно реализованного редактора базы медицинской терминологии и наблюдений) не требует модификации алгоритма сервиса.

Немаловажным требованием к сервису является простота и удобство его использования для врачей. Оно обеспечивается следующими решениями: системой интеллектуального поиска (с учетом синонимии) необходимых для ввода признаков и структурным представлением их описаний в соответствии с требованиями юзабилити. Известный любому программисту принцип создания интерфейса «не заставляйте пользователя думать» основан на предоставлении ему выбора элементов

вместо ввода, поскольку таким образом пользователь (врач) не должен ни вспоминать названия характеристик, их значений, ни обдумывать, как правильно и корректно составить фразу. Ему предоставляется полное описание признака, всех его характеристик и возможных значений, из которых необходимо выбрать требуемые.

Ожидается, что реализованный сервис, во-первых, заметно упростит работу врача по вводу данных в ЭМК по сравнению с используемыми в настоящее время различными аналогичными сервисами, во-вторых, позволит в его практической деятельности использовать интеллектуальные системы поддержки принятия решений, в третьих, формализованные датасеты (ЭМК) помогут обнаружить (с использованием методов машинного обучения) новые знания в медицине в целом (в практической медицине и организации здравоохранения).

Работа выполнена при частичной финансовой поддержке фонда РФФИ № 19-07-00244 и № 20-07-00670.

ЛИТЕРАТУРА



1. Ваганова Е.В. Медицинские информационные системы как объект оценки: факторы и тенденции развития // Вестн. Том. гос. ун-та. Экономика. – 2017. – № 37.
2. Берестнева О.Г., Шаропин К.А., Старикова А.В., Кабанова Л.И. Технология формирования баз знаний в медицинских информационных системах // Известия ЮФУ. Технические науки. – 2010. – № 8.
3. Бледжанц Г.А., Саркисян М.А., Исакова Ю.А., Туманов Н.А., Попов А.Н., Бегмуродова Н.Ш. Ключевые технологии формирования искусственного интеллекта в медицине // Ремедиум. – 2015. – № 12.
4. Бурцева А.Л., Берестнева Е.В., Степаненко Н.П. Создание базы знаний для медицинской экспертной системы // Современные наукоемкие технологии. – 2016. – № 3–1. – С. 14–17.
5. Кобринский Б.А. Компьютеризированные и дистанционные обучающие системы (на примере медицинской диагностики) // Открытое образование. – 2018. – № 2. URL: <https://cyberleninka.ru/article/n/kompyuterizirovannye-i-distantsionnye-obuchayuschie-sistemy-na-primere-meditsinskoy-diaagnostiki> (Дата обращения: 23.04.2020).
6. Грибова В.В., Клещев А.С., Москаленко Ф.М., Тимченко В.А., Федорищев Л.А., Шалфеева Е.А. Облачная платформа IASaaS для разработки оболочек интеллектуальных сервисов: состояние и перспективы развития // Программные продукты и системы. – 2018. – Т. 31. – № 3. – С. 527–536.
7. Грибова В.В., Клещев А.С., Москаленко Ф.М., Тимченко В.А. Модель редактора сложноструктурированных информационных единиц, управляемого метаинформацией, и его реализация // Научно-техническая информация. Сер. 2. – 2016. – № 2. – С. 1–13.
8. Грибова В.В., Клещев А.С., Москаленко Ф.М., Тимченко В.А. Модель порождения орграфов информации по орграфу метаинформации для двухуровневой модели сложноструктурированных информационных единиц // Научно-техническая информация. Сер. 2. – 2015. – № 12. – С. 26–38.
9. Грибова В.В., Петряева М.В., Окунь Д.Б., Шалфеева Е.А. Онтология медицинской диагностики для интеллектуальных систем поддержки принятия решений // Онтология проектирования. – 2018. – № 1 (27).
10. Ненашева Е.О., Пальчунов Д.Е. Разработка автоматизированных методов преобразования предложений естественного языка в бескванторные формулы логики предикатов // Вестник НГУ. Серия: Информационные технологии. – 2017. – № 3.
11. Klyshinsky E., Gribova V.V., Shakhgeldyan C., Shalfееva E.A., Okun D.B., Geltser B.I., Gorbach T.A., Karpik O.D. Formalization of Medical Records Using an Ontology: Patient Complaints // Communications in Computer and Information Science. – 2020. – Vol. 1086. – Pp. 141–152. ISBN 978-3-030-39574-2. DOI https://doi.org/10.1007/978-3-030-39575-9_14.