

УНИФИКАЦИЯ РАЗРАБОТКИ СИСТЕМ ПОДДЕРЖКИ ПРИНЯТИЯ РЕШЕНИЙ НА ОСНОВЕ НАЦИОНАЛЬНОЙ МЕДИЦИНСКОЙ НОМЕНКЛАТУРЫ

Т.В. Зарубина, С.Е. Раузина, Е.Н. Николаиди, П.А. Астанин, Л.В. Ронжин, А.А. Борисов, П.А. Ядгарова, Ю.И. Королева, А.В. Усова, И.И. Потапова, А.П. Милютина, А.В. Будыкина

VIII Всероссийская научно-практическая конференция по Искусственному интеллекту в здравоохранении
и системам поддержки принятия врачебных решений ИТМ ИИ (ITM-AI.ru)

Москва, 08-09 февраля 2024 г.

«КЛАССИЧЕСКИЙ» ВАРИАНТ РАЗРАБОТКИ СППР И ПРОБЛЕМЫ МАСШТАБИРОВАНИЯ



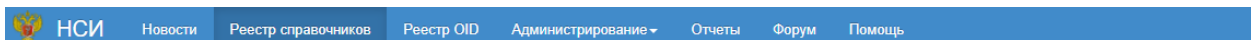


Регламентная служба ФРНСИ Минздрава России
на базе ФГБУ ЦНИИОИЗ Минздрава России создана в
2015 г.

Есть разные группы справочников
Каждый справочник имеет паспорт



nsi.rosminzdrav.ru



Реестр справочников / Просмотр справочника

Действующие вещества лекарственных препаратов для медицинского применения, в том числе необходимых для льготного обеспечения граждан лекарственными средствами

Код	Наименование на русском	Наименование на латинском	Версия	Наименование на английском	Код АТХ
4	L-аспарагин		3.10, 15.09.2020 18:47		
6	Абакавир	Abacavirum	3.9, 09.07.2020 16:00	Abacavir	J05AF06
9	Абатацепт	Abataceptum	3.8, 27.05.2020 17:40	Abatacept	L04AA24
10	Абциксимаб	Abciximabum	3.7, 13.04.2020 17:07	Abciximab	B01AC13
13	Агомелатин	Agomelatinum	3.6, 06.04.2020 13:50	Agomelatine	N06AX22
14	Адалимумаб	Adalimumabum	3.5, 18.03.2020 14:16	Adalimumab	L04AB04
15	Адамантилбромфениламин		3.4, 12.12.2019 17:01	Adamantylbromphenylamine	A13A
16	Адапален	Adapalenum	3.3, 27.06.2019 19:29	Adapalene	D10AD03
18	Адеметионин	Ademetioninum	3.2, 09.01.2019 13:33	Ademetionine	A16AA02
22	Аденозин		3.1, 02.07.2018 00:00	Adenosine	C01EB10
24	Аденозина фосфат	Adenosini phosphas	2.3, 18.01.2018 00:00	Adenosine phosphate	C01EB10
25	Адипиодон	Adipiodonum	2.2, 22.11.2017 00:00	Adipiodone	V08AC04
26	Адифенин	Adipheninum	2.1, 11.10.2017 00:00	Adiphenine	
27	Адонизид		1.0, 06.12.2016 00:00	Adoniside	
31	Азаклорзин	Azaclorzinum		Azaclorzine	

Поддерживается версияемость справочников

Единый электронный адрес
Регламентной службы ФРНСИ:
nsi@nsiminzdrav.ru

В составе Регламентной службы ФРНСИ
работает отдел разработки
структурированных электронных
медицинских документов (СЭМД)

1. Лаборатория семантического анализа медицинской информации
2. Лаборатория разработки систем поддержки принятия клинических и управленческих решений
3. Лаборатория цифрового развития медицинского образования

2022 г.

Цель:

Создание научно-образовательной экосистемы для обеспечения здравоохранения принципиально новыми инструментами поддержки деятельности медицинских работников и исследователей, созданными и развиваемыми на основе инновационных цифровых технологий, искусственного интеллекта и Big Data и подготовки высококвалифицированных кадров с высокой степенью исследовательской и цифровой готовности

Задачи:

- Информационно-поисковая система для предоставления врачу семантически связанной информации
- Сервисы поддержки принятия клинических решений
- Учебная МИС МО для обучения студентов «врачебных» факультетов и последипломного образования
- Семантический тренажер для подготовки врачей-кибернетиков

-
- Информационно-поисковый сервис врача-исследователя
 - ...



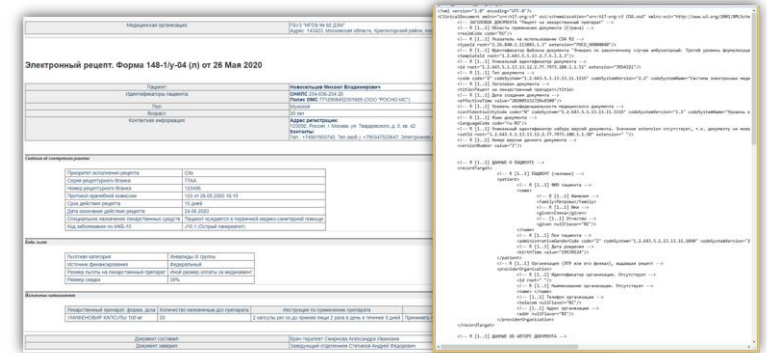
Разработка унифицированной медицинской терминологии



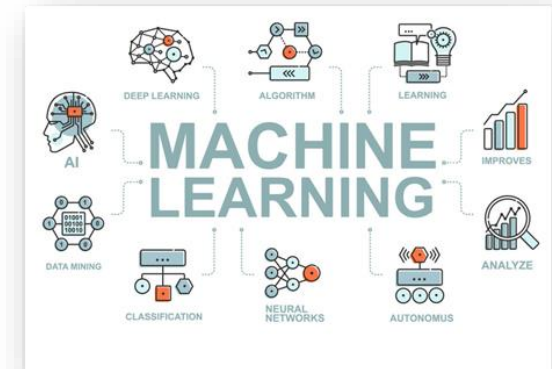
Инженеры по знаниям



1. Разработка структурированных электронных медицинских документов (СЭМД)



2. Развитие алгоритмов обработки неструктурированных медицинских текстов



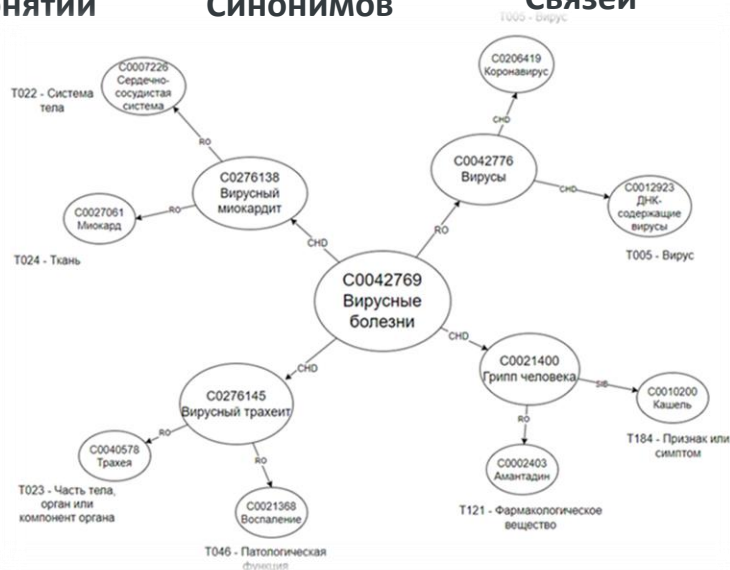
Систематизированный словарь терминов/концептов

Их связанное представление Все области медицины

Зарубежный опыт

Система унифицированного медицинского языка,
Unified Medical Language System **UMLS**

4,6млн Понятий **17,2млн** Синонимов **98млн** Связей **76** Справочника



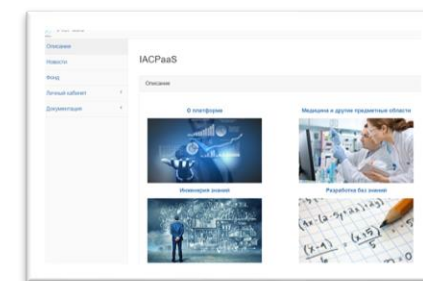
ICD-10-CM
RxNorm

Отечественный опыт

УМКВ - Объединенная база медицинских знаний

Терминология, созданная на платформе **IACPaaS** (Институт автоматизации и процессов управления ДВО РАН)

~ 2,5 млн. терминов (без учета синонимов)



№	Наименование КР	ID	Год принят.
1	Аллергический ринит	КР261	2020
2	Острый синусит	КР313	2021
3	Острые респираторные вирусные инфекции у взрослых	КР724	2021
4	Хронический тонзиллит	КР683	2021
5	Острый тонзиллит и фарингит	КР306	2021
6	Острый обструктивный ларингит и эпиглоттит	КР352	2021
7	Паратонзиллярный абсцесс	КР664	2021
8	Грипп у взрослых	Проект	2021
9	Бронхиальная астма	КР359	2021
10	Внебольничная пневмония у взрослых	КР654	2021
11	Хроническая обструктивная болезнь легких	КР603	2021
12	Хронический бронхит	КР655	2021
13	Эмфизема легких	КР656	2021
14	Идиопатический легочный фиброз	КР677	2021
15	Новая коронавирусная инфекция (Covid-19)	Врем. МР	2022
16	Стабильная ишемическая болезнь сердца	КР155	2020
17	Ишемический инсульт и транзиторная ишемическая атака у взрослых	КР171	2021
18	Гастрит и дуоденит	КР708	2021
19	Язвенная болезнь	КР277	2020
20	Хронический панкреатит	КР273	2020
21	Рак желудка	КР574	2020
22	Рак молочной железы	КР379	2021

Первичная валидация:

оценка степени покрытия концептами UMLS образов заболеваний – на основе **22 клинических рекомендаций** в части описания клинической картины, данных анамнеза и факторов риска

- Аналогичное или полностью схожее наименование – **74%**
- Частичное совпадение – **16%** терминов
- **Не найдено - 10% терминов**
Причины: сложное или иное звучание в англоязычной среде

Проблемы:

1. Низкая доля русских переводов < 3% от объема терминов.
2. Большая неравномерность семантической сети.
3. Недостаток прямых связей между концептами-заболеваниями и концептами-симптомами.



СОЗДАНИЕ УНИФИЦИРОВАННОЙ НАЦИОНАЛЬНОЙ МЕДИЦИНСКОЙ НОМЕНКЛАТУРЫ (УНМН) **1 версия**



➤ Переводы

Нейросетевые переводы, а также переводы Google Translate, Yandex Translate, PROMT терминов из семантических групп *Симптомы и признаки, Клинические находки, Заболевания и синдромы* с последующим экспертным уточнением → переведена **170** тыс. понятий

➤ Сопоставление с национальными справочниками (ФР НСИ МЗ РФ)

Добавлено переводов \cong **117** тыс.

- | | | | | |
|-----------------|-------------------------------|-----------------------------------|--------------|--------------|
| ✓ МКБ-10 (100%) | ✓ Выявленные патологии (100%) | ✓ Анатомические локализации (90%) | ✓ ФСДИ (75%) | ✓ ФСЛИ (32%) |
|-----------------|-------------------------------|-----------------------------------|--------------|--------------|

Исходно было представлено на русском языке в UMLS - **304** тыс. понятий
Добавлено экспертно верифицированных переводов - > **287** тыс.

- медицинская литература (публикации):
27 млн абстрактов **PubMed**

- данные реальной клинической практики:
dataset **MIMIC-IV**: (**330** тыс. ЭМК)

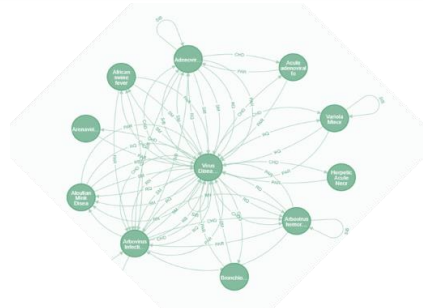


Добавлено \approx 16 млн прямых связей между симптомами и заболеваниями

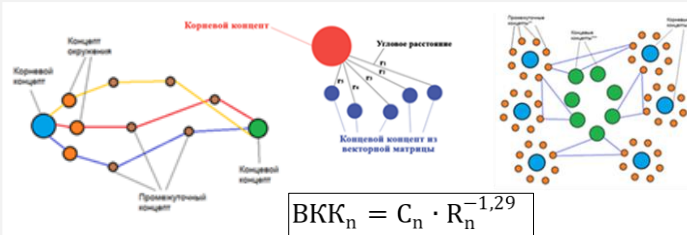
✓ **Закономерности, исходно представленные в UMLS**

✓ **Существующие методы анализа семантических сетей**
 ✓ **Разработанные ИЦТМ метрики**

✓ **Ранги, полученные при анализе связей из dataset MIMIC-IV и PubMed**



паттерны для выделения симптомов, диагностических методов, лечения и т.п.



- кратчайшие пути
- графовое расстояние
- геометрические фигуры
- метрики клинической значимости, связности, специфичности



Лаборатория вычислительной физиологии Массачусетского технологического института

от 70% покрытия основных задач по 22 группам нозологий до около 90%

Реляционная форма СУБД:



- Хранение терминов и связей, результатов работы аналитических метрик, весов;
- Точечный поиск информации;
- Создание графовой модели УНМН;
- Обеспечение взаимодействия пользователя с интерфейсными решениями.

Графовая форма СУБД:



- Оптимизация процессов обработки данных;
- Представление информации в виде графа;
- Выполнение сложных аналитических запросов с большим количеством операций объединения данных;
- Реализация параллельных вычислений.

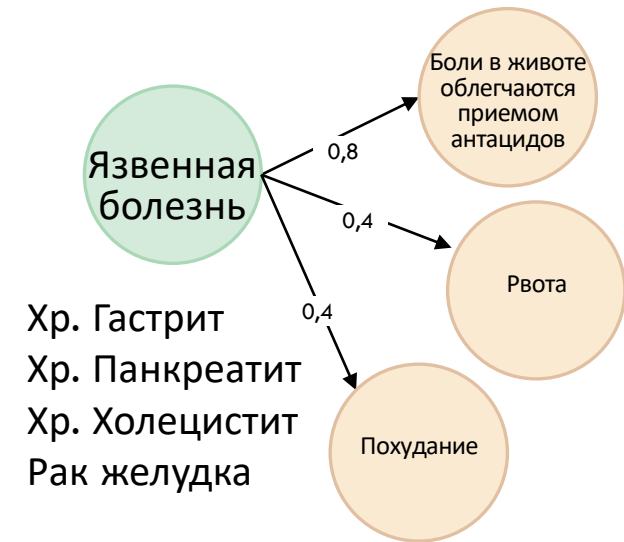
Оценка значимости терминов и связей

➤ Метрики автоматизированного построения значимости

- ✓ Частота встречаемости концептов в графовой модели, медицинских текстах, ЭМК
- ✓ Разметка концептов по дополнительным семантическим признакам (например, концептов-симптомов к нозологическим группам) для ранжирования выдаваемых диагностических заключений
- ✓ Клиническая значимость концептов

➤ Получение весовых коэффициентов экспертным путем: изучение медицинских публикаций, привлечение экспертов

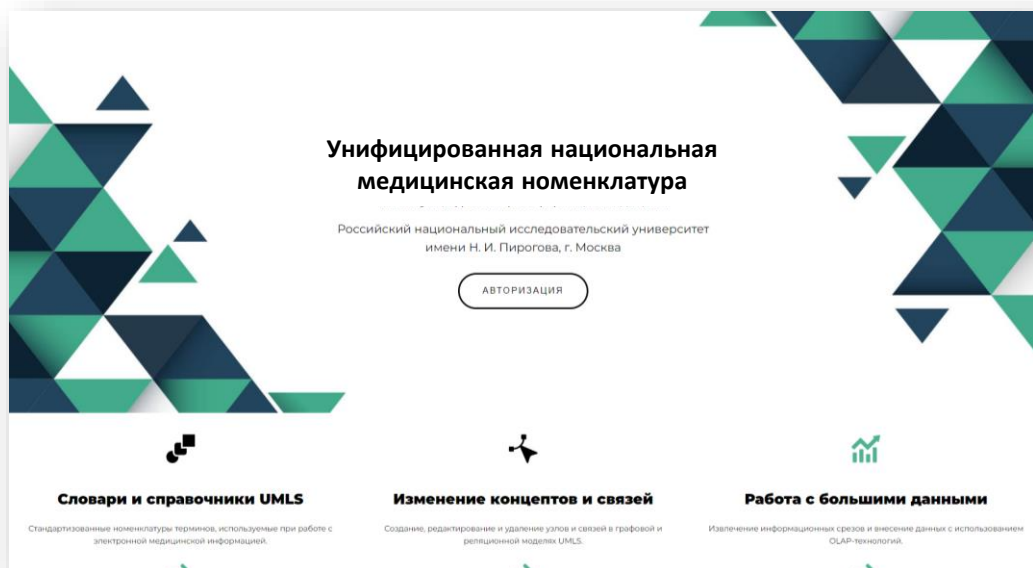
- ... Заболевания органов дыхания неопухолевого характера
- ... Заболевания дифференциального ряда аксиального спондилоартрита
- ... Заболевания дифференциального ряда язвенной болезни



Рубрикатор
клинических рекомендаций



Аналитическая система для проведения исследовательских работ и обучения



- ✓ Нормативно-справочная информация – международные и национальные словари
- ✓ Редактор базы знаний – инструмент для наполнения УНМН экспертными знаниями
- ✓ Интеллектуальное ядро – комплекс аналитических решений, предназначенных для извлечения OLAP-срезов из УНМН
- ✓ + АРМ аналитика, АРМ эксперта

Прототип информационно-поисковой системы (ИПС) для автоматического определения наиболее вероятного класса заболеваний при составлении дифференциально-диагностического ряда произвольного клинического образа и выбора методов диагностики для уточнения



Искомый диагноз встречается среди трёх первых заболеваний в ранжированном перечне

свободно
вводимый
текст

понятия,
выделенные из
свободно
вводимого текста

Укажите жалобы или симптомы

боль в правом подреберье
отрыжка
горьким диспепсия
тяжесть в животе

Поиск 🔍

Очистить

Возможно подходят:

C0000726 ЖИВОТ
C0030193 боль

С0549272 Тяжесть в желудке ✖
C0014724 отрыжка ✖
C0235299 боль в правом верхнем квадранте живота ✖
C0013395 диспепсия ✖

Подобрать диагноз

Заболевания | Диагностические методы

Диспепсия
1

Холецистит
0.5

Язвенная болезнь желудка и/или двенадцатиперстной кишки
0.32

Гастродуоденит
0.285

Аппендицит
0.282

понятия,
наиболее тесно
связанные с
выделенными

вывод ранжированного списка заболеваний на основе введенных жалоб, симптомов, клинических находок

Создание полноценных ИПС
для пользователя-врача,
исследователя, пациента;
разработка СППКР с
использованием созданных
инструментов

Развитие собственно
УНМН



Разработка новых национальных
словарей, например,
справочника аббревиатур,
справочника симптомов и
клинических находок

Применение УНМН для
формализации и унификации
ЭМК, а также при создании
СЭМД (единое кодирование
информации в ЭМК)

ЭТАПЫ РАЗРАБОТКИ

СППКР

Базы знаний:

Эксперты/ Медицинская литература/ Методы машинного обучения

Онтологическая модель УНМН:

Закономерности UMLS/ Графовые метрики/
Ранги DataSet

УНМН:

Концепты, связи, семантическая организация UMLS/
Переводы/ справочники ФНСИ
Прямые связи из PubMed и DataSet MIMIC-IV



БЛАГОДАРЮ ЗА ВНИМАНИЕ

zarubina@rsmu.ru