



Институт цифровой медицины

Кафедра информационных и
интернет-технологий

Москва
2024

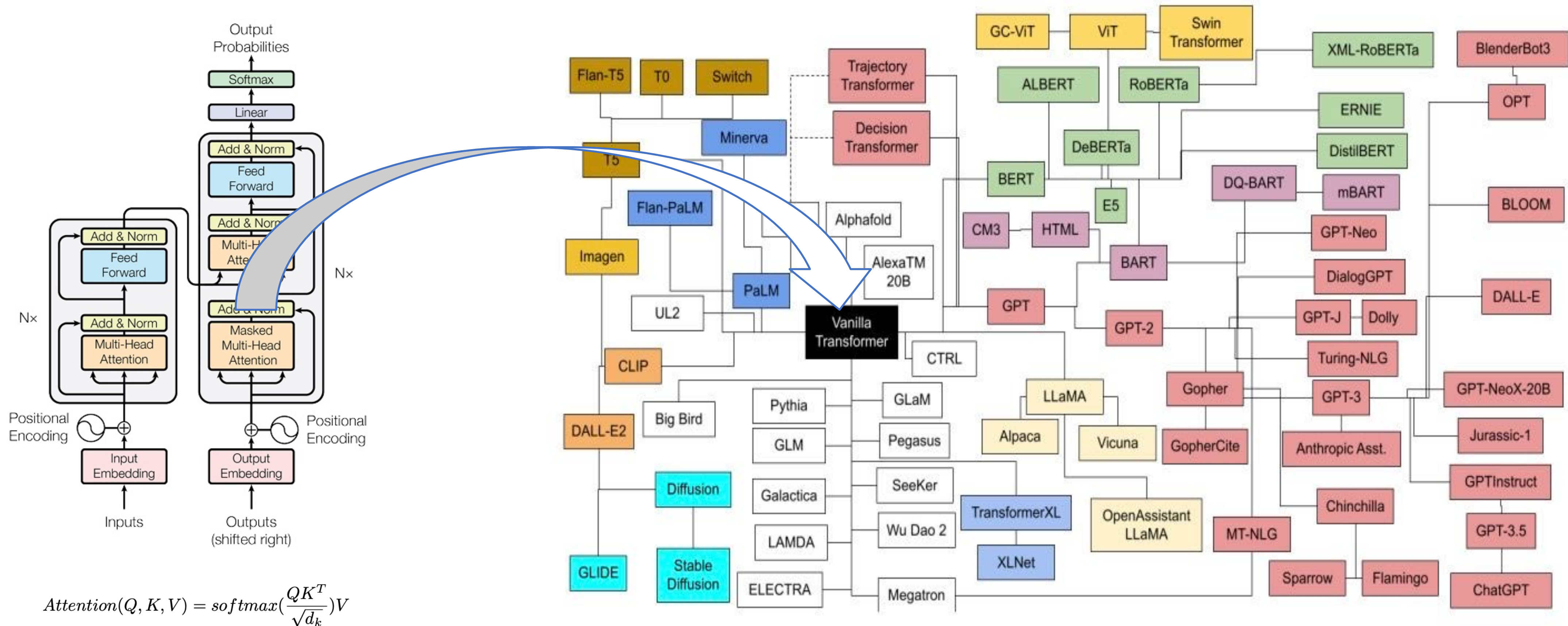
Большие языковые модели в медицине

Фартушный Эдуард Николаевич



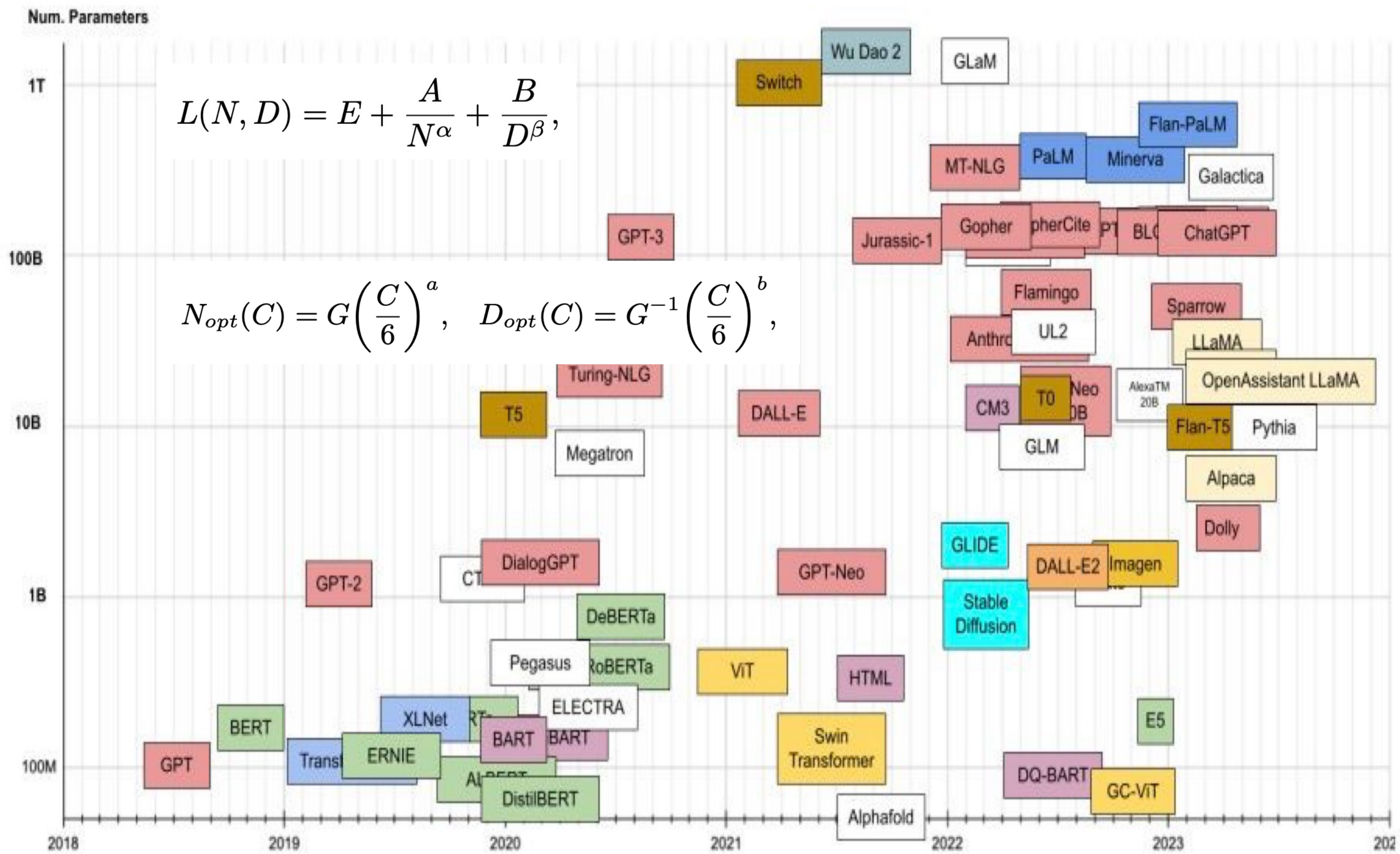
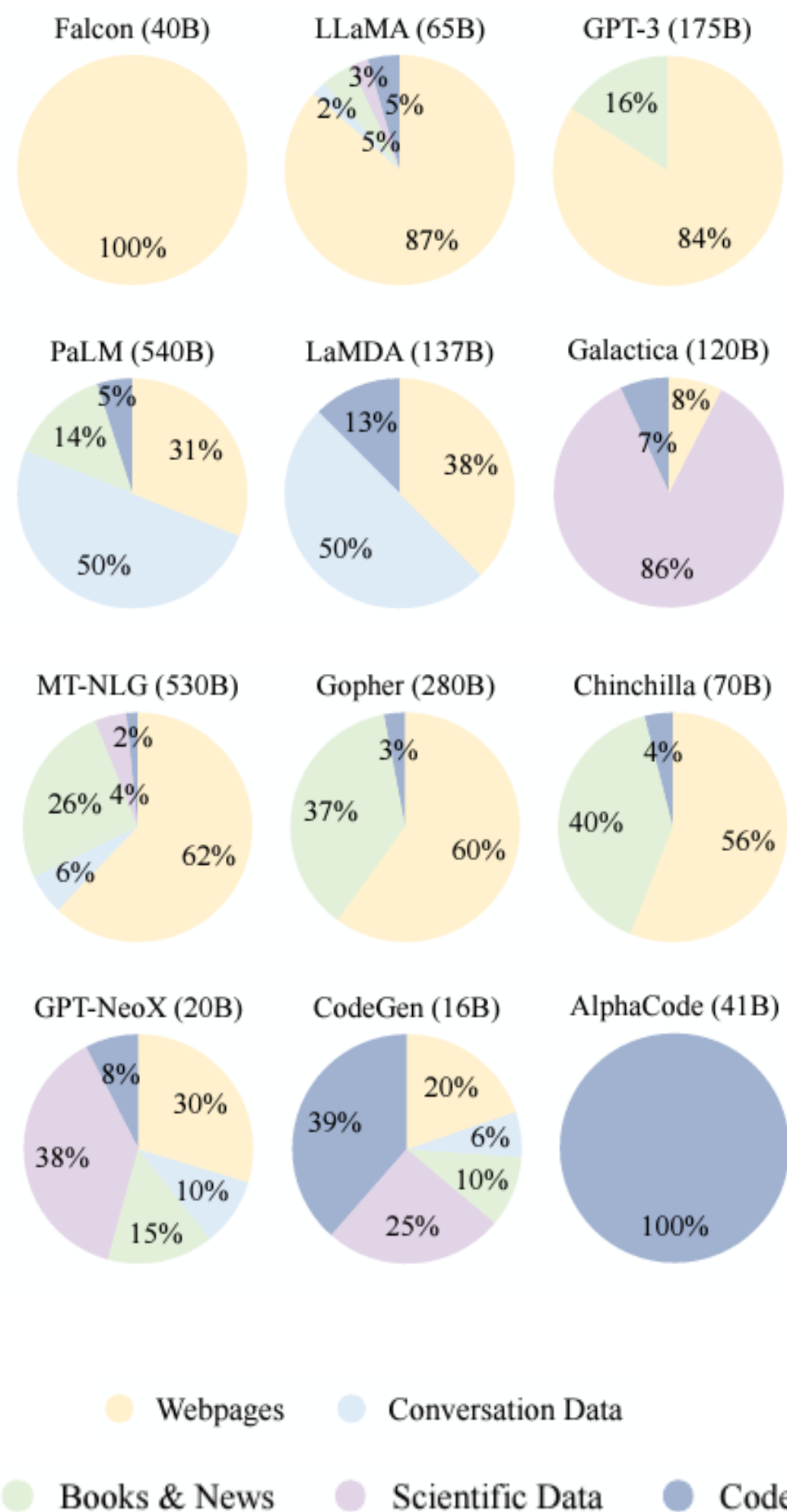


1. Что такое LLM? Эволюция от "маленьких" к большим языковым моделям.
2. Архитектура трансформеров для задач NLP. Качественный переход при увеличении размера модели и обучающего датасета.
3. Применение LLM в современной медицине, краткий обзор решений.
4. Работы в Сеченовском Университете по медицинскому применению LLM.
5. Вызовы и проблемные вопросы в разработке и применении LLM. Направления работ и перспективы.
6. Предварительные итоги и выводы

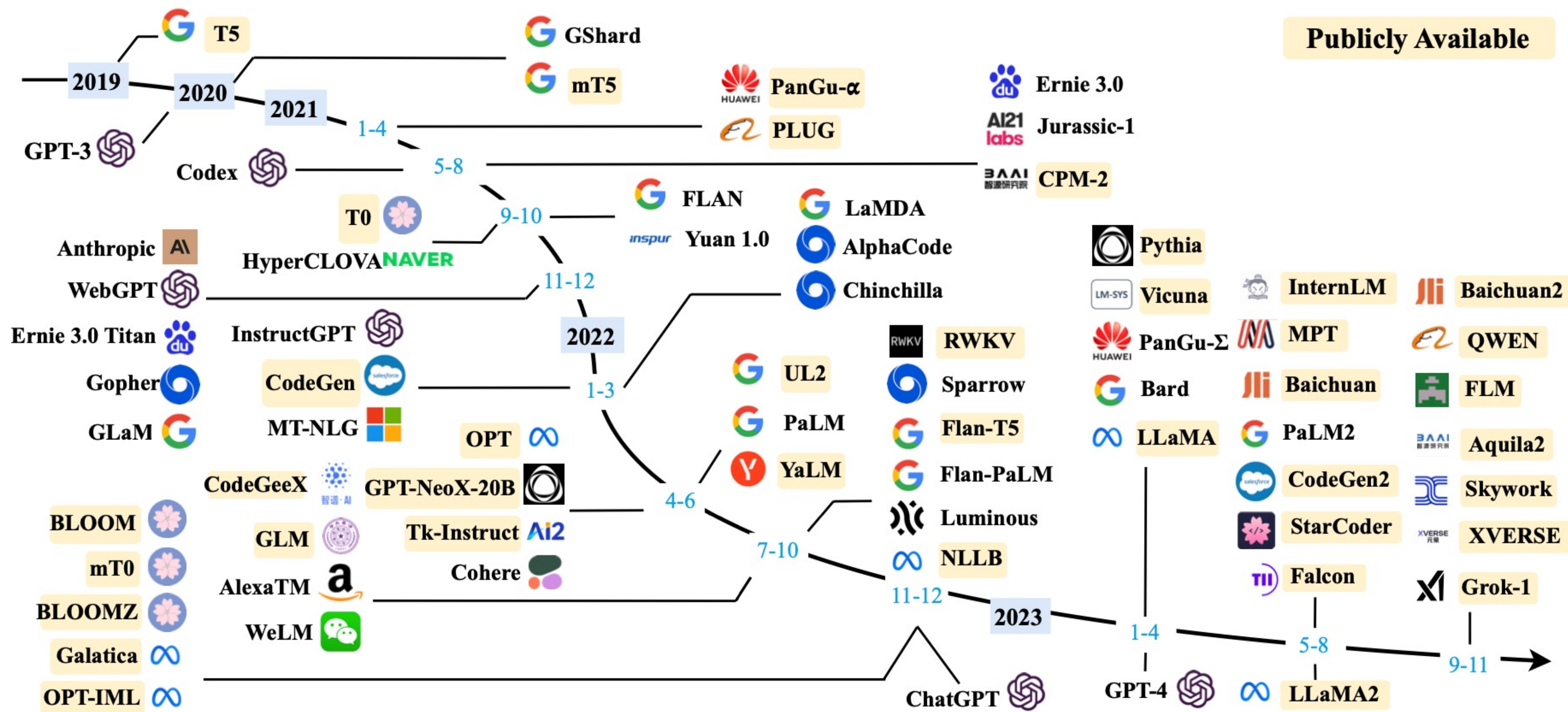


$$Attention(Q, K, V) = softmax(\frac{QK^T}{\sqrt{d_k}})V$$

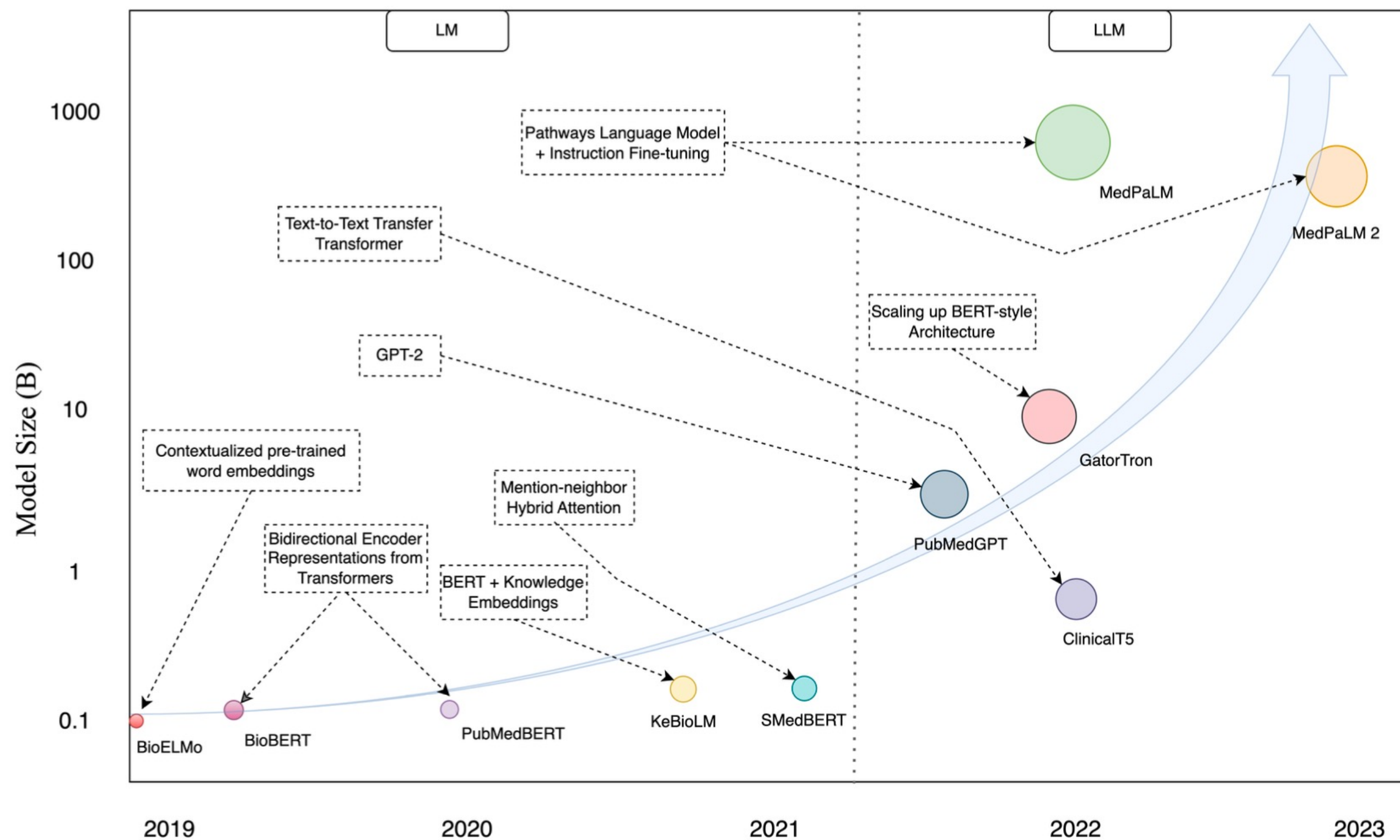
AttentionIsAllYouNeed, 31stConferenceonNeuralInformationProcessingSystems(NIPS2017).



Формула “шинишлы” от команды Google DeepMind 2022 год: Hoffmann, S. Borgeaud, ... “Training compute-optimal large language models,” vol. abs/2203.15556, 2022.



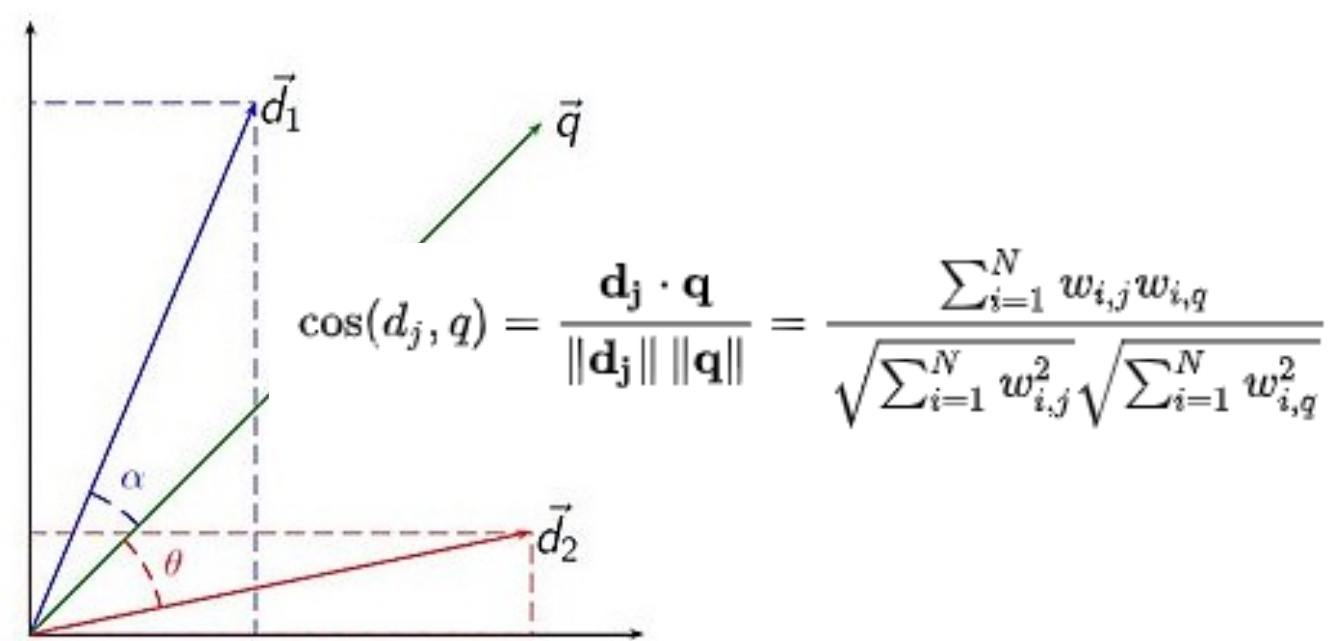
Эволюция больших языковых моделей (размером более 10Gb.) за последние годы.



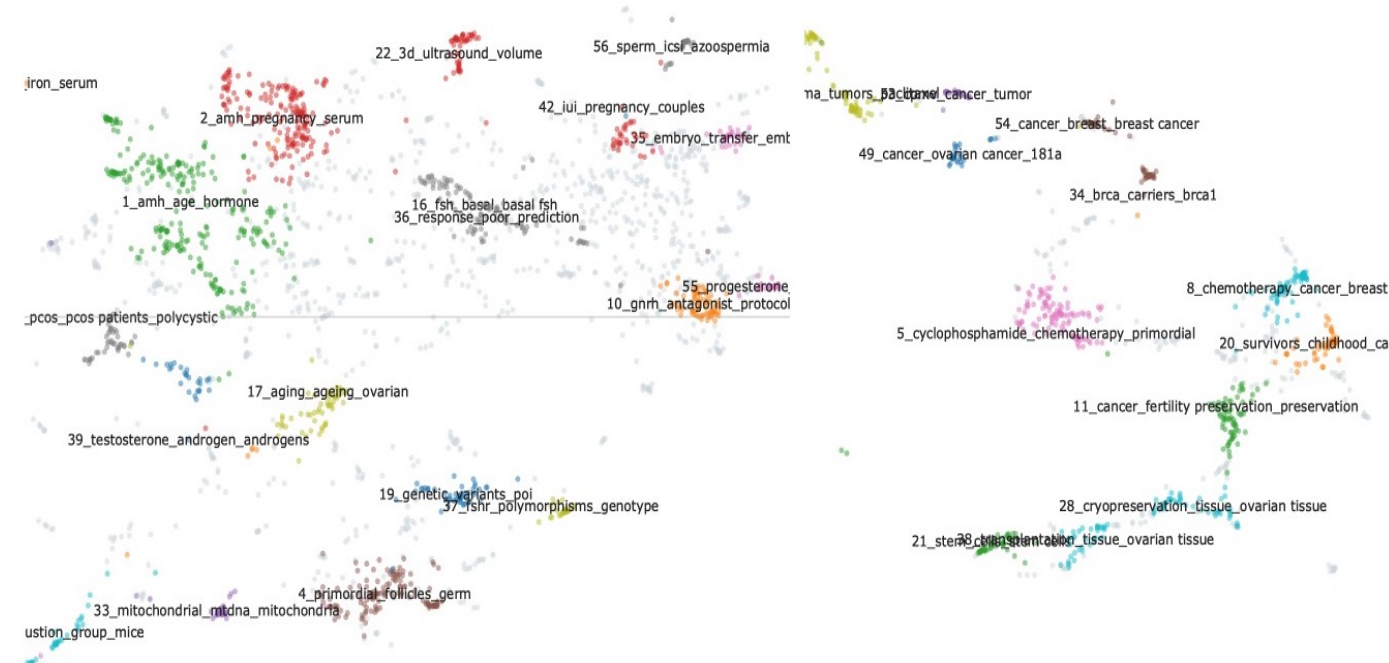
Модель	Год	Задачи	Автор
MedPALM-2	2023	MQA	Google
Radiology-Llama2	2023	Радиология	Университет Джоржии
Med-HALT	2023	Тесты на галлюцинирование	Saama AI
BioGPT	2023	Классификация, MQA,	MS
GatorTron	2023	Семантическое сходство, MQA	Университет Флориды
ClinicalT5	2022	Классификация, NER	Университет Орегон
BioMedLM	2022	BMQA	Стэнфорд
BioBART	2022	Саммаризация, диалоги, NER	Университет Цинхуа



1. Задачи медицинской диагностики, анализ симптомов, записей в ЭМК ... и дальнейшая классификация
2. Уход за пациентами и применение LLM в соответствующих сервисах
3. Системы поддержки принятия врачебных решений СППВР (AI'Doctor - НТИ, Сколково др.)
4. Обработка биомедицинских текстов (SechenovDatamed и др.)
5. Разработка лекарственных препаратов
6. Виртуальные медицинские помощники и боты
7. Радиология – мультимодальные LLM, тексты + медицинские изображения
8. Генерация медицинского текста при описании результатов исследований (Центр телемедицины и радиологии г.Москвы)



Ищем похожие на q-документ из коллекции D, документ q представлен вектором нормализованных термов, в т.ч. N-граммами.



При построении векторных эмбедингов используем в т.ч. LLM-модели

Search interface for SechenovDataMed.AI. Search query: "Factors influencing the effectiveness of ruxolitinib therapy for myelofibrosis?".

Filters:

- Дата публикации: От: 01.01.2000, До: dd.mm.gggg
- Тип статьи:
 - Клиническое исследование
 - Метаанализ
 - Рандомизированное контролируемое исследование
 - Обзор
 - Регулярный обзор
 - Журнальная статья
- Возраст пациента:
 - Ребенок: рождение-18 лет
 - Новорожденный: рождение-1 месяц
 - Младенец: рождение-23 мес.
 - Младенец: 1-23 месяца
 - Дошкольник: 2-5 лет
 - Ребенок: 6-12 лет
 - Подросток: 13-18 лет
 - Взрослый: 19+ лет
 - Молодёжь: 19-24 года.
 - Взрослый: 19-44 лет
 - Средний возраст: + Возраст 45+ лет
 - Средний возраст: 45-65 лет
 - Возраст: 65+ лет
 - 80 и старше: 80+ лет

Results: 5 sources found.

PMID	Заголовок	Дата выхода
26056473	Критическая оценка роли руксолитиниба при миелофибро...	2015 год
32932806	Белки теплового шока и PD-1/PD-L1 как потенциальные тер...	2020 11 сен
32201152	Реактивация туберкулеза после терапии руксолитинибом ...	2021 сентябрь
30010187	Руксолитиниб у пожилых пациентов с миелофиброзом: вл...	2018 октябрь
32334833	Ингибирование JAK2 в лейкозных клетках, несущих JAK2(...	2020 25 июня

Buttons: "Найти", "Результаты поиска", "Обработать", "Разметить".

Annotation for the first result:

Терапия рака с помощью Т-клеток, созданных с помощью TCR: текущие стратегии, проблемы и перспективы.

Авторы: Shafer, Paul; Kelly, Lauren M; Hoyos, Valentina

Аннотация:

Рак 1,00 Терапия с использованием TCR-сконструированных Т-клеток 0,89 TCR 0,99 - Сконструированных Т-клеток: текущие стратегии, проблемы и перенаправление Т-клеток на опухолевые клетки 0,50 Опухолевые 1,00 клетки, Т-клетки 0,78 могут быть сконструированы ex vivo для экспрессии рака 0,71 -антигенспецифические Т-клеточные рецепторы (TCR 1,00), генерирующие продукты, известные как TCR-сконструированные Т-клетки 0,97 TCR 0,94 -сконструированные Т-клетки (TCR T 0,94 TCR 0,99 T). В отличие от химерных антигенных рецепторов 0,99 (CAR 0,97), TCR 1,00 распознают пептиды, представленные HLA 0,54 полученные из белков всех клеточных компартов. Использование Т-клеток TCR 0,91 TCR 0,92 Т-клеток для адаптивной клеточной терапии (ACT) привлекло повышенное внимание, особенно в связи с активизацией усилий по лечению солидного рака 1,00 с помощью ACT. В этом обзоре мы описываем различные механизмы распознавания антигена 0,65 Т-клетками и передачи сигнала, опосредованные CAR 0,78 и TCR 0,99 . Мы описываем классы раковых антигенов 0,97 , распознаваемые современными методами лечения TCR 0,97 T, и обсуждаем как классические, так и новые доклинические стратегии открытия, улучшения и проверки антигенспецифического TCR 0,99 . Наконец, мы рассматриваем текущую картину клинических испытаний терапии TCR 0,99 T и обсуждаем, что эти текущие результаты указывают на разработку будущих инженерных TCR 0,78 .



ebm.datamed.pro/tematic_review

EBM Sechenov DataMed.AI Тематический анализ Поговорим Факты для EBM

adoptive T cells therapy

Дата публикации: 01.01.2000 - 00.00.0000

Тип статьи: Clinical Trial, Meta Analysis, Randomized Controlled Trial, Review, Systematic Review, Journal Article

Возраст пациента: Child: birth-18 years, Newborn: birth-1 months, Infant: 1-23 months, Preschool Child: 2-5 years, Child: 6-12 years, Adolescent: 13-18 years, Adult: 19+ years, Young Adult: 19-24 years

Суммаризация прошла успешно.

2_activation_tcr_cell_cells_signalling_cd28_kinase_signaling_receptor_protein

Огрисовать граф

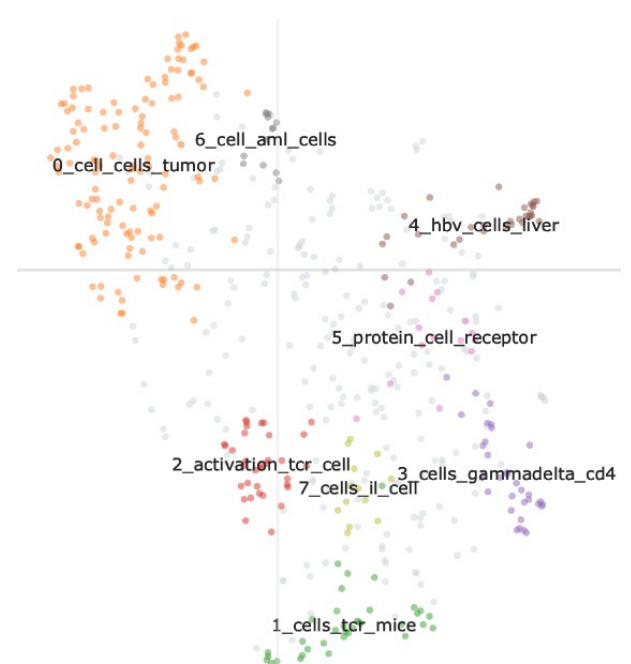
Заголовок	Дата выхода	Av
CD25 and CD69 induction by alpha4beta1 outside-in signallin...	2013 Aug 15	Ci
The adaptor protein EBP50 is important for localization of the ...	2009 Dec 23	St
The membrane adaptor LAT is proteolytically cleaved followin...	2013 Mar 15	Ga
Binding of Staphylococcal Enterotoxin B (SEB) to B7 Receptor...	2021	Ku
CD8(-) T cell transfectants that express a high affinity T cell re...	2001 Oct 15	Hc
Control of amino-acid transport by antigen receptors coordina...	2013 May	Sir
T-cell receptor early signalling complex activation in response ...	2010 May 27	St
A role for the thermal environment in defining co-stimulation r...	2015	Zy
BCL11B enhances TCR/CD28-triggered NF-kappaB activation t...	2009 Jan 15	Ci
The low affinity Fc receptor for IgG functions as an effective c...	2005 Feb 1	Dh
New insights into the catalytic mechanism of histidine phosph...	2013 Jul 1	Sa

Подробное описание

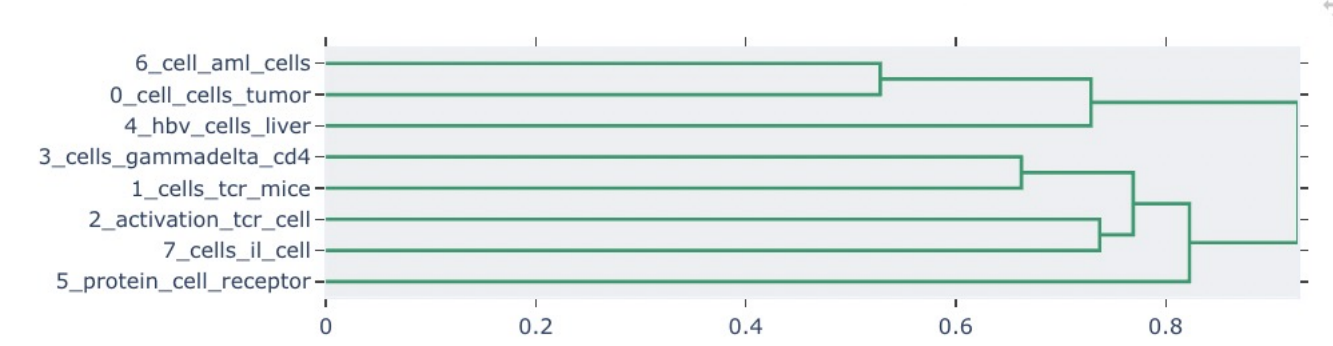
PSMA-targeting TGFbeta-insensitive armored CAR T cells in metastatic castration-resistant prostate cancer: a phase 1 trial.

Авторы: Narayan, Vivek; Barber-Rotenberg, Julie S; Jung, In-Young; Lacey, Simon F; Rech, Andrew J; Davis, Megan M; Hwang, Wei-Ting; Lal, Priti; Carpenter, Erica L; Maude, Shannon L; Plesa, Gabriela; Vapiwala, Neha; Chew, Anne; Moniak, Michael; Sebro, Ronnie A; Farwell, Michael D; Marshall, Amy; Gilmore, Joan; Lledo, Lester; Dengel, Karen; Church, Sarah E; Hether, Tyler D; Xu, Jun; Gohil, Mercy; Buckingham, Thomas H; Yee, Stephanie S; Gonzalez, Vanessa E; Kulkovskaya, Irina; Chen, Fang; Tian, Lifeng; Tien, Kyle; Gladney, Whitney; Nobles, Christopher L; Raymond, Hayley E; Hexner, Elizabeth O; Siegel, Donald L; Bushman, Frederic D; June, Carl H; Fraietta, Joseph A; Haas, Naomi B

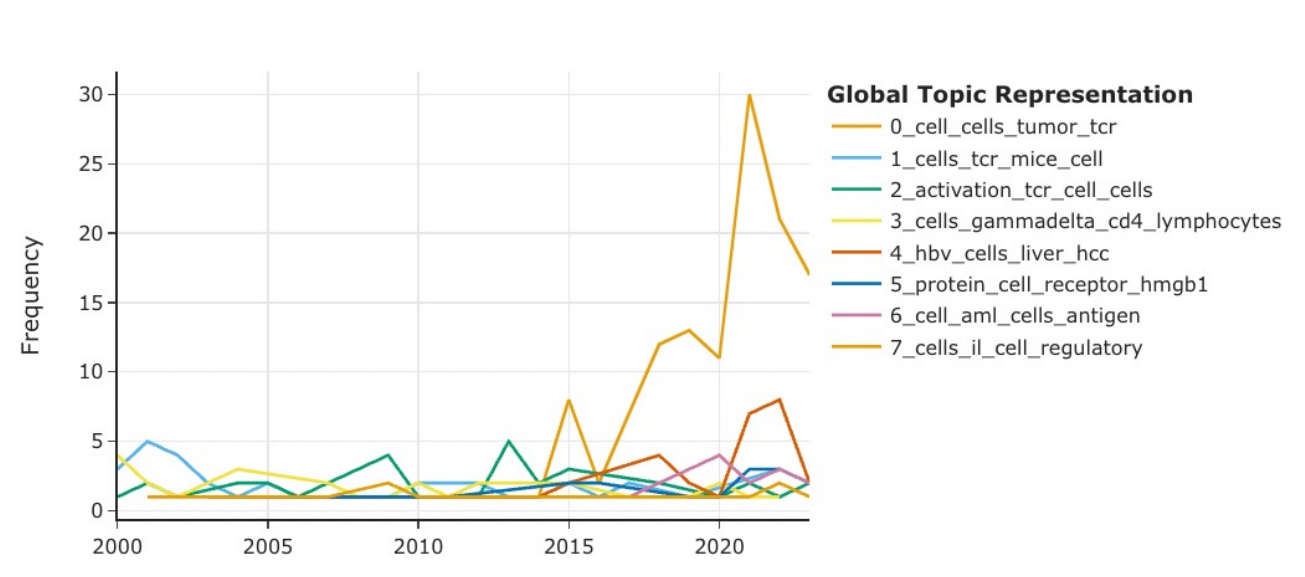
Аннотация: PSMA_{0.99}-targeting TGF-insensitive armored_{0.80}TGF_{0.92}-insensitive armored CAR Tcells_{0.77} in metastatic Chimeric antigen receptor (CAR) Tcells_{0.44}Chimeric antigen receptor (CAR) Tcells have demonstrated promising efficacy, particularly in hematologic malignancies_{0.04}. One challenge regarding CAR Tcells_{0.78}CAR_{0.99} Tcells in solid tumors_{0.01} is the immunosuppressive tumor_{0.98} microenvironment (TME), characterized by high levels of multiple inhibitory factors, including transforming growth factor_{0.97} (TGF_{0.95}). We report results from an in-human_{0.98} phase 1 trial of castration-resistant, prostate cancer-directed CAR Tcells_{0.98}prostate cancer_{0.87}-directed CAR Tcells armored with a dominant-negative TGF-receptor_{1.00} (NCT03089203). Primary endpoints were safety and feasibility, while secondary objectives included assessment of CAR Tcell_{0.90}CAR_{0.70} Tcell distribution, bioactivity and disease response. All prespecified endpoints were met. Eighteen patients_{0.98} enrolled, and 13 subjects received therapy across four dose levels.



Hierarchical Clustering



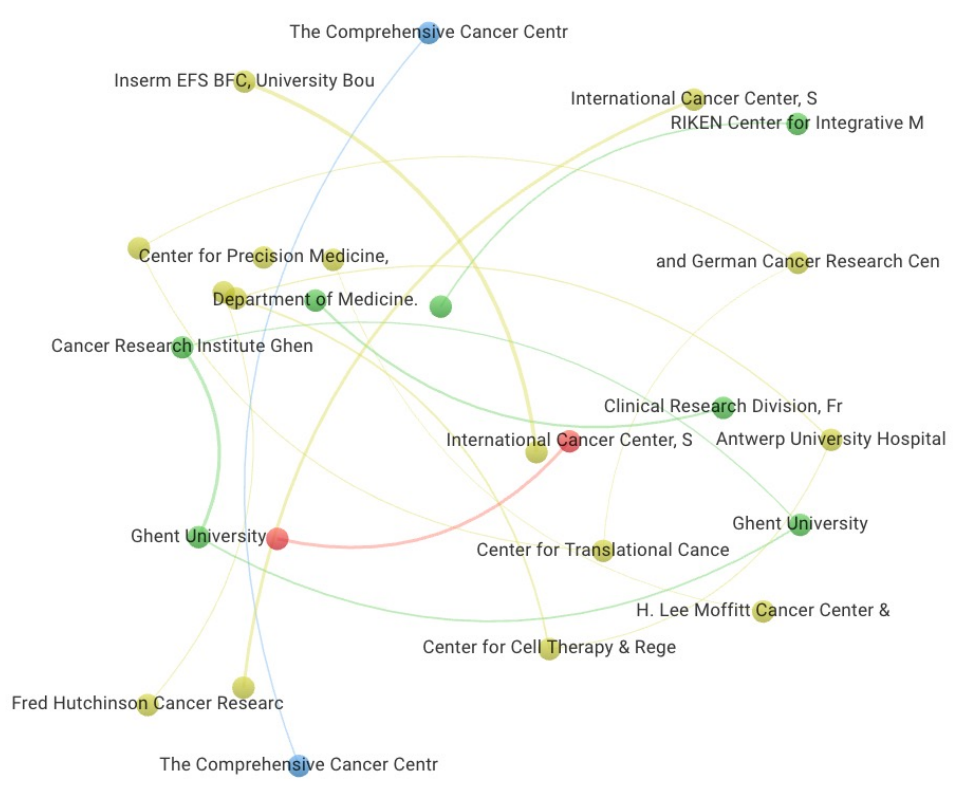
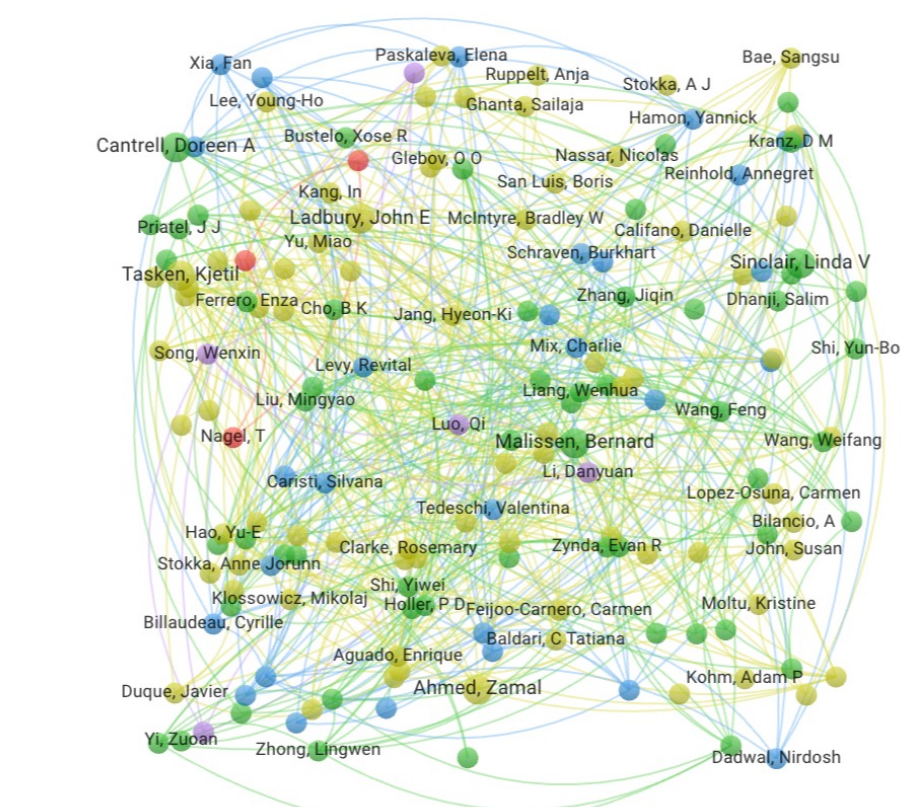
Topics over Time



Суммаризация

Подведем итог
Вовлечение TCR инициирует серию молекулярных событий, приводящих к секреции цитокинов, пролиферации и дифференцировке Т-клеток. Таким образом, наша работа предоставила новое понимание взаимодействия между ранними сигнальными путями TCR и CD28, имеющими решающее значение для запуска активации наивных Т-клеток. Задействование TCR (рецептор Т-клеток) индуцирует фосфорилирование тирозина адаптера LAT (линкер активации Т-клеток), и тем самым он рекрутирует несколько цитозольных медиаторов для нижестоящих сигнальных путей. Эти сигналы приводят к фосфорилированию мотивов активации иммунорецепторов на основе тирозина, присутствующих в цитозольных хвостах TCR-ассоциированных субъединиц CD3, которые после фосфорилирования служат каркасом для сборки большого супрамолекулярного сигнального комплекса, ответственного за активацию Т-клеток. Сопутствующая стимуляция TCR и CD28 вызывает выраженный кальциевый ответ в наивных Т-клетках. Активация Т-клеток инициируется при взаимодействии лигандов рецептора Т-клеток (TCR) и костимулирующих рецепторов. Фактически, активацию можно было обнаружить при значительно более низких концентрациях пептидов, чем у Т-клеток, экспрессирующих TCR дикого типа. Задействование рецептора Т-клеток (TCR) специфическим лигандом пептид-МНС инициирует трансмембранную передачу сигналов, вызывающую активацию Т-клеток, что является ключевым этапом в большинстве адаптивных иммунных ответов. Эти подходы производят сигнал TCR недостаточной силы, чтобы вызвать активацию CD4(+) Т-клеток, и вместо этого вызывают функциональную анергию или делецию Т-клеток, избегая при этом генерализованной долгосрочной иммуносупрессии.

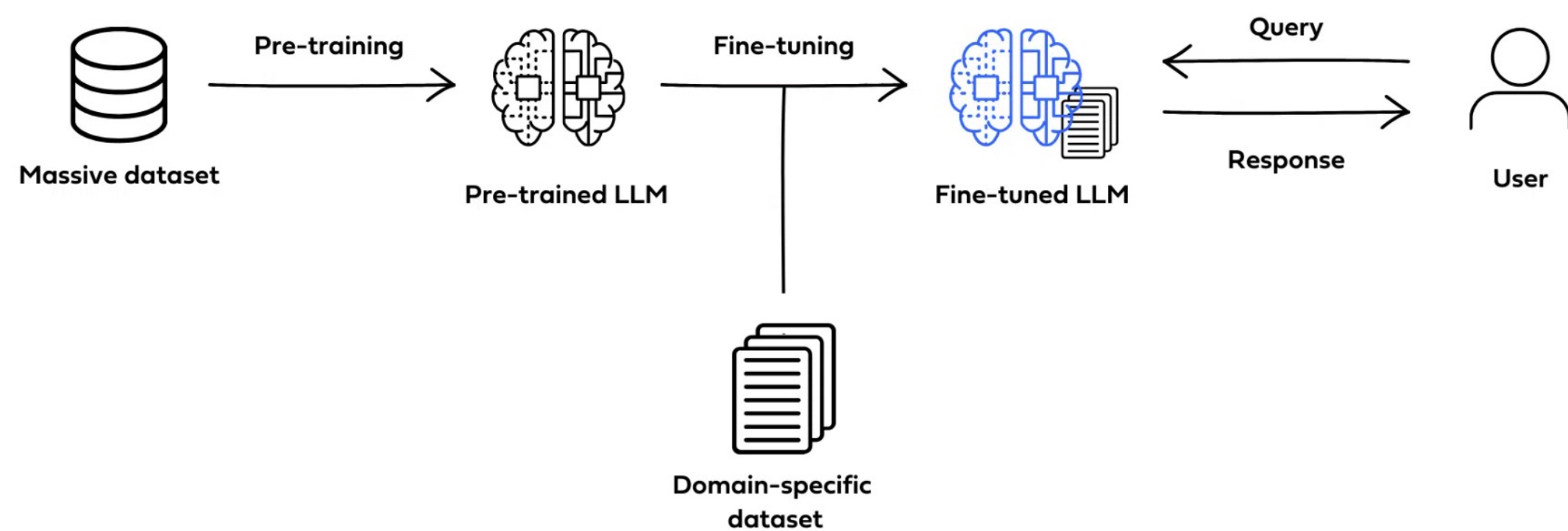
Проектор



Впервые применен RAG для генерации поисковых ответов на домене биомедицинского текста.

Тестировалась модель YandexGPT.

Full Fine Tuning

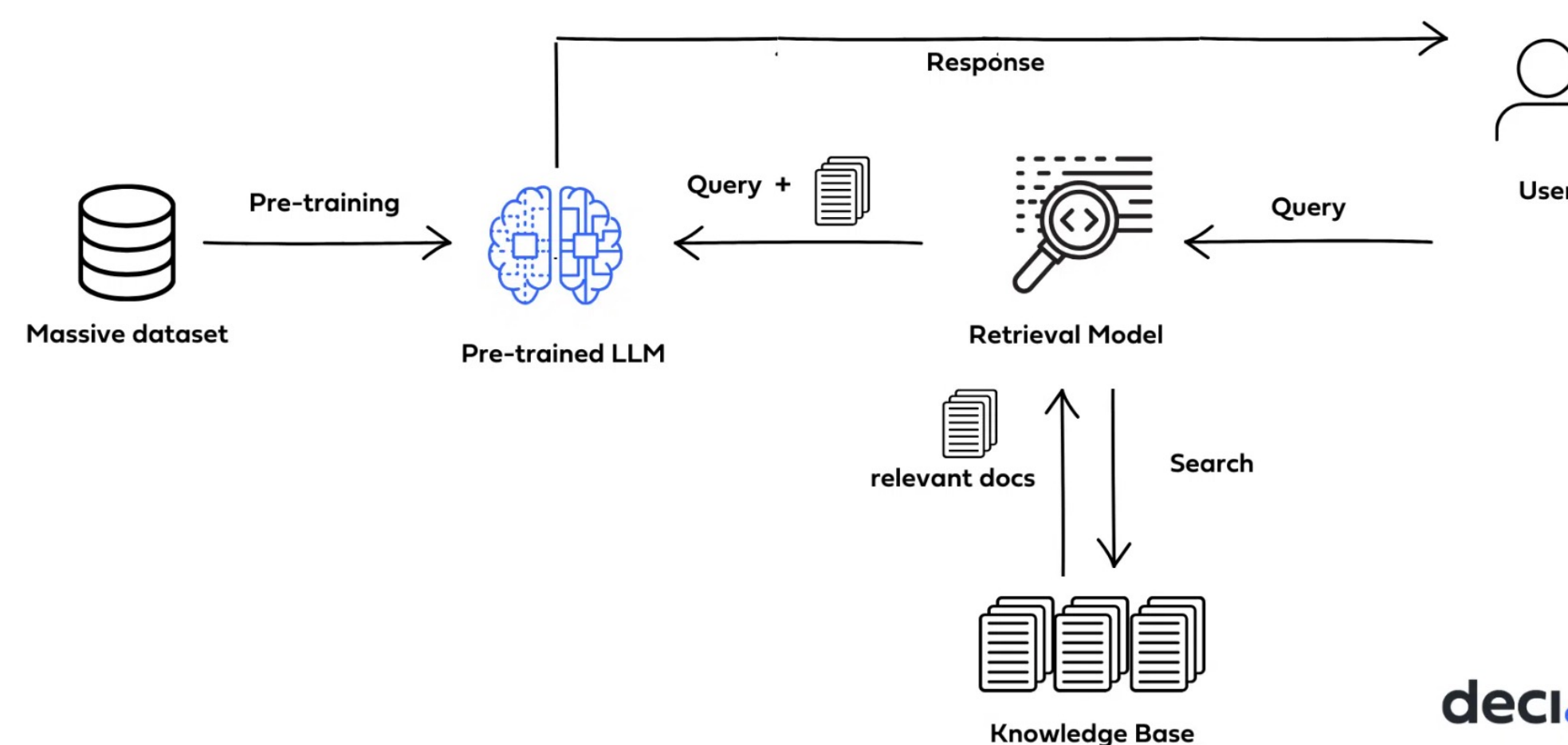


Относительное сравнение стоимости реализации и поддержки



1. Промпт-инжиниринг
2. RAG – генерация дополненная поисковыми данными
3. Точечный тюнинг модели PEFT LoRA «Низкоранговая адаптация больших языковых моделей», был представлен в 2023 года.
4. Full fine-tuning обучение базовой модели на своем домене знаний

Retrieval Augmented Generation





Опыты и оценка качества различных моделей

Медицинские диалоги из предпрома

Медицинский чат бот Sechenov-DatMed.LLM [ver=0.0.2]

Параметры моделей
 Temperature: 0.40
 Context_length: 1024
 Max_new_tokens: 512
 Top_K: 30

Выбор модели:
 llama-2 (13B)

21:05
 System: I hope this information helps! Do you have any other questions or concerns?

21:26
 User: What is **Sensitivity of EBUS-TBNA** in the diagnosis of lung cancer?

21:28
 Llama-2 (13B): Hello! I'm here to help you with your question about the sensitivity of EBUS-TBNA in the diagnosis of lung cancer. EBUS-TBNA (Endobronchial Ultrasound-Transbronchial Needle Aspiration) is a minimally invasive diagnostic procedure that uses a flexible bronchoscope to obtain tissue samples from the lungs and surrounding areas. The sensitivity of EBUS-TBNA in diagnosing lung cancer refers to its ability to correctly identify cancerous cells or tissues. Studies have shown that EBUS-TBNA has a high sensitivity for detecting lung cancer, particularly for adenocarcinoma and squamous cell carcinoma. One study published in the Journal of Thoracic Oncology found that EBUS-TBNA had a sensitivity of **83.9%** for diagnosing lung cancer, with a specificity of **95.6%**. Another study published in the European Respiratory Journal reported a sensitivity of **87.2%** and a specificity of **94.1%**. However, it's important to note that the sensitivity of EBUS-TBNA can vary depending on several factors, such as the experience of the operator, the quality of the equipment, and the patient's individual characteristics. Additionally, EBUS-TBNA may not be as effective in detecting certain types of lung cancer, such as small cell lung cancer or metastatic disease. In summary, EBUS-TBNA is a highly sensitive diagnostic tool for lung cancer, with studies showing an accuracy rate of around **85-90%**. However, the sensitivity can vary depending on several factors, and it may not be as effective in certain cases. If you have any further questions or concerns, please don't hesitate to ask!

21:31
 User: What is **Sensitivity of PET** for lung cancer diagnosis?

21:33
 Llama-2 (13B): Hello! I'm here to help you with your question about the sensitivity of PET for lung cancer diagnosis. PET (Positron Emission Tomography) is a powerful imaging tool that can help doctors diagnose and monitor various types of cancer, including lung cancer. The sensitivity of PET for lung cancer diagnosis refers to its ability to accurately detect and identify lung cancer tumors. Studies have shown that PET has high sensitivity for detecting lung cancer, particularly for adenocarcinoma and squamous cell carcinoma, which are the most common types of lung cancer. In fact, a systematic review and meta-analysis published in the journal Radiology found that PET had an overall sensitivity of **87%** for detecting lung cancer. However, it's important to note that the sensitivity of PET can vary depending on several factors, such as the type of tracer used, the dose of tracer administered, and the quality of the imaging equipment. Additionally, PET may not be as sensitive for detecting small tumors or early-stage lung cancer, and it may not be able to distinguish between benign and malignant lesions in some cases. It's also important to note that while PET is highly sensitive for detecting lung cancer, it is not perfect and can produce false negative results in some cases. Therefore, it should be used in combination with other diagnostic tests, such as CT scans and biopsies, to ensure accurate diagnosis and treatment. In summary, the sensitivity of PET for lung cancer diagnosis is high, but it may not be perfect and should be used in combination with other diagnostic tests to ensure accurate diagnosis and treatment. If you have any further questions or concerns, please don't hesitate to ask!

Send a message [Send]

Система: Задавайте вопрос модели DataMed.LLM (сборка v.0.1.12)! 08:23

Пользователь: Сравнить эффективность гифтитиниба, эрлотиниба и афатиниба при лечении мелкоклеточного рака легких? 11:25

Модель: Как чат-бот, обученный работе с большими объемами медицинской информации, я могу предоставить вам информацию об эффективности гилотиниба, эрлотиниба и афатиниба в лечении мелкоклеточного рака легких (МРЛ). Согласно систематическому обзору и метаанализу, опубликованным в журнале Thorax, все три препарата доказали свою эффективность в терапии первой линии МРЛ. Однако существуют некоторые различия в их эффективности, которые могут повлиять на решения о лечении. Гилотиниб (Тепметко) — ингибитор тирозинкиназы, одобренный FDA в 2018 году для лечения метастатического МРЛ. В клинических исследованиях было показано, что препарат улучшает общую выживаемость, выживаемость без прогрессирования и частоту ответа по сравнению с химиотерапией. Метаанализ, опубликованный в журнале Cancer, показал, что гилотиниб связан со значительным улучшением общей выживаемости по сравнению с эрлотинибом (Тарцева) и афатинибом (Гилотриф). Эрлотиниб — ингибитор тирозинкиназы EGFR, который уже несколько лет используется для лечения МРЛ. В клинических исследованиях было показано, что препарат улучшает общую выживаемость, выживаемость без прогрессирования и частоту ответа по сравнению с химиотерапией. Метаанализ, опубликованный в журнале Oncotarget, показал, что эрлотиниб связан со значительным улучшением общей выживаемости по сравнению с афатинибом. Афатиниб является ингибитором тирозинкиназы EGFR, который также используется в течение нескольких лет для лечения МРЛ. В клинических исследованиях было показано, что препарат улучшает общую выживаемость, выживаемость без прогрессирования и частоту ответа по сравнению с химиотерапией. Метаанализ, опубликованный в журнале Lung Cancer, показал, что афатиниб связан со значительным улучшением общей выживаемости по сравнению с эрлотинибом.

Чувствительность спирометрического теста для выявления хронической обструктивной болезни легких (ХОБЛ) **высокая — 74%**. Однако специфичность этого теста составляет **всего около 65%**, а это означает, что отрицательный результат не исключает ХОБЛ. В результате положительная прогностическая ценность может быть низкой в некоторых группах населения. ХОБЛ, или хроническая обструктивная болезнь легких, — это состояние, при котором дыхательные пути сужаются и затрудняется дыхание. Спирометрический тест измеряет, сколько воздуха вы можете выдохнуть после глубокого вдоха. Высокая чувствительность означает, что тест позволяет выявить ХОБЛ при ее наличии. Однако, поскольку не все люди с ХОБЛ имеют аномальные результаты спирометрии, специфичность теста ниже. Это означает, что у некоторых людей, не страдающих ХОБЛ, этот тест все же может дать положительный результат. Таким образом, хотя чувствительность спирометрического теста для выявления ХОБЛ высока, его специфичность лишь умеренная. В результате положительная прогностическая ценность может быть низкой в некоторых группах населения. Важно помнить, что отрицательный результат не исключает ХОБЛ и может потребоваться дальнейшее тестирование или осмотр у врача.



1. Галлюционирование LLM – генерация не существующих знаний
2. Безопасность – намеренное внедрение токсичных инструкций или предвзятых текстов в домен
3. Обеспечение конфиденциальности данных – особенно при обучении на данных пациентов, проблемы очистки и деперсонализации данных
4. Доступность ресурсов для обучения или тюнинга LLM
5. Разработка системы оценки качества LLM медицинского назначения
6. Подготовка датасетов с медицинскими знаниями для обучения LLM
7. Интерпретируемость и объяснимость генерируемых результатов LLM



Перплексия - вероятностная метрикой, которая количественно определяет степень неопределенности или информационного содержания в прогнозах, сделанных языковой моделью, при этом более низкие значения отражают более высокую точность и последовательность прогнозов.

BLEU - Оценка Bilingual Evaluation Understudy (BLEU) оценивает качество результата машинного перевода путем сравнения его с одним или несколькими эталонными переводами.

GLEU - GLEU вычисляет средний балл различных n-грамм, предлагая более полную оценку сгенерированного текста.

ROUGE - Оценка, Recall-Oriented Understudy for GistEvaluation оценивает системы автоматического реферирования и машинного перевода, измеряя перекрытие между сгенерированным резюме и справочными резюме.

Distinct-1/2 — это дополнительная мера, используемая для оценки разнообразия сгенерированных ответов путем количественной оценки уникальности n-грамм.

Оценка F1 — это показатель машинного обучения, который сочетает в себе точность и полноту для измерения точности модели двоичной классификации, балансируя между правильной идентификацией положительных примеров и минимизацией ложных результатов.

BERTScore — это метрика, которая вычисляет оценку сходства между токенами в предложении-кандидате и токенами в эталонном предложении, используя контекстные встраивания для сходства токенов, а не полагаясь на точные совпадения.

Человеческая оценка, как показатель качества для оценки большой языковой модели в академическом контексте, предполагает участие экспертов-оценщиков, которые оценивают качество и адекватность контента, созданного с помощью модели, предоставляя ценную качественную информацию о его производительности и согласованности с человеческим опытом. - ожидания и стандарты.



Институт цифровой
медицины

Большие языковые модели

Москва, 2024

СПАСИБО ЗА ВНИМАНИЕ!

Фартушный Эдуард Николаевич

fartushny_e_n@staff.sechenov.ru