

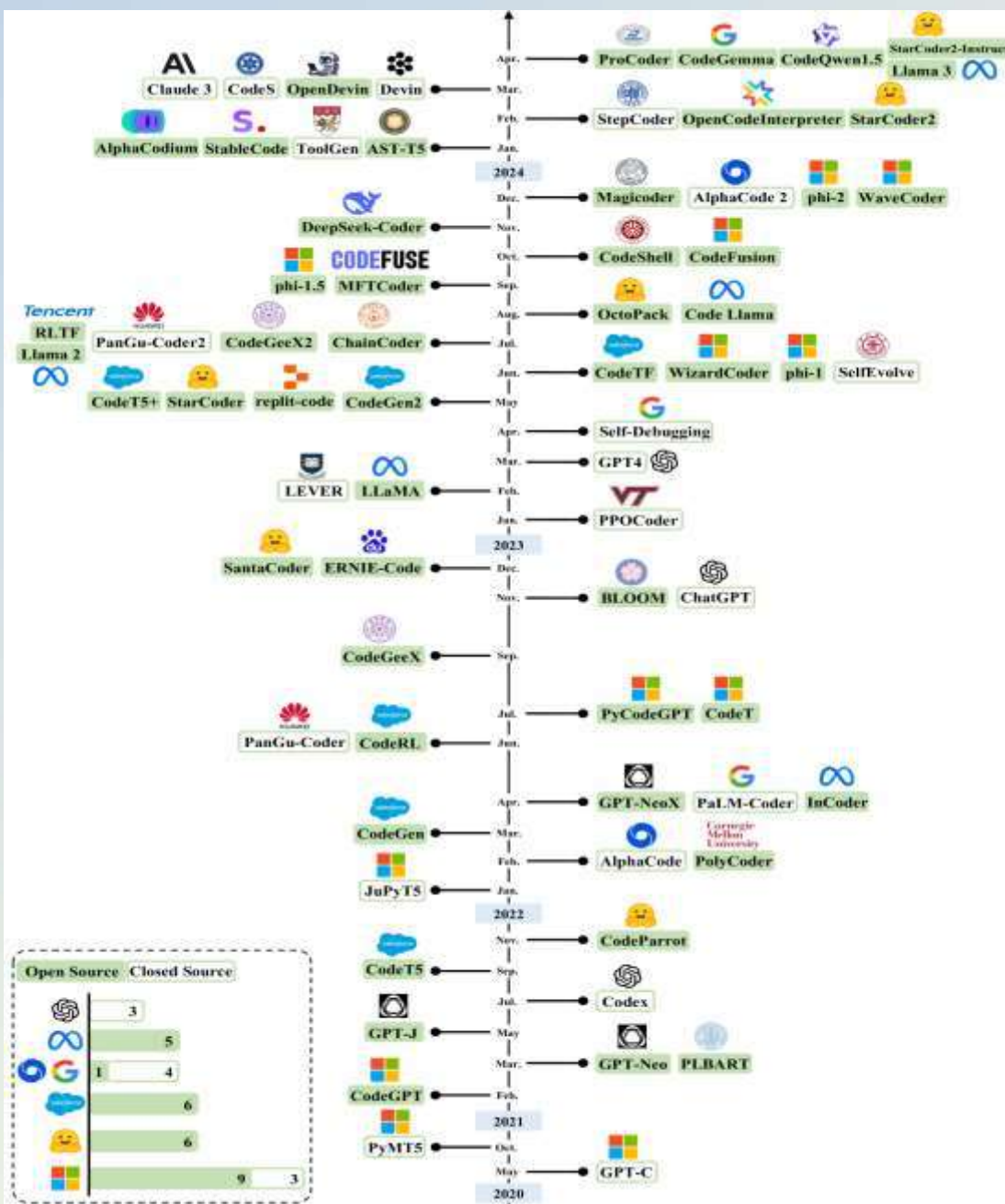
Научный семинар по исследованиям
цифровой экономики
*«Экономика искусственного интеллекта:
китайский феномен и русский след»*

**Генеративный искусственный
интеллект в задачах
экономического прогнозирования**

Бахтизин А.Р.

Москва, 05.03.2025

Эволюция больших LLM



<https://arxiv.org/html/2406.00515v1>

Boston Consulting Group:

Следующий шаг в развитии ИИ – **автономные агенты** станут мейнстримом в ближайшие годы и кардинально усилят сегодняшние LLM. По оценкам они станут массовыми в течение **3-5 лет**.

- 1. Автоматизация рабочих процессов:** выявление узких мест, снижение вмешательства человека
- 2. Создание виртуальных сообществ агентов** (различных персонажей), в том числе для проведения социологических, экономических, психологических и др. исследований

<https://www.bcg.com/publications/2023/gpt-was-only-the-beginning-autonomous-agents-are-coming>



Разработка компании IBM

Large language models revolutionized AI. LLM agents are what's next (2024)

Модернизация устаревших программных приложений для крупных предприятий и организаций. Для программистов-людей рефакторинг кода (просмотр и переработка миллионов строк старой кодовой базы) и интеграция с другими системами – дорогостоящий и трудоемкий процесс

Агенты на основе LLM значительно упрощают этот процесс – агент-эксперт по устаревшему программному обеспечению анализирует старый код и перерабатывает отдельные сегменты. Агент-эксперт по качеству обеспечивает тестирование и проверяет соответствие корпоративным стандартам.

<https://research.ibm.com/blog/what-are-ai-agents-llm>

<https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/why-agents-are-the-next-frontier-of-generative-ai>

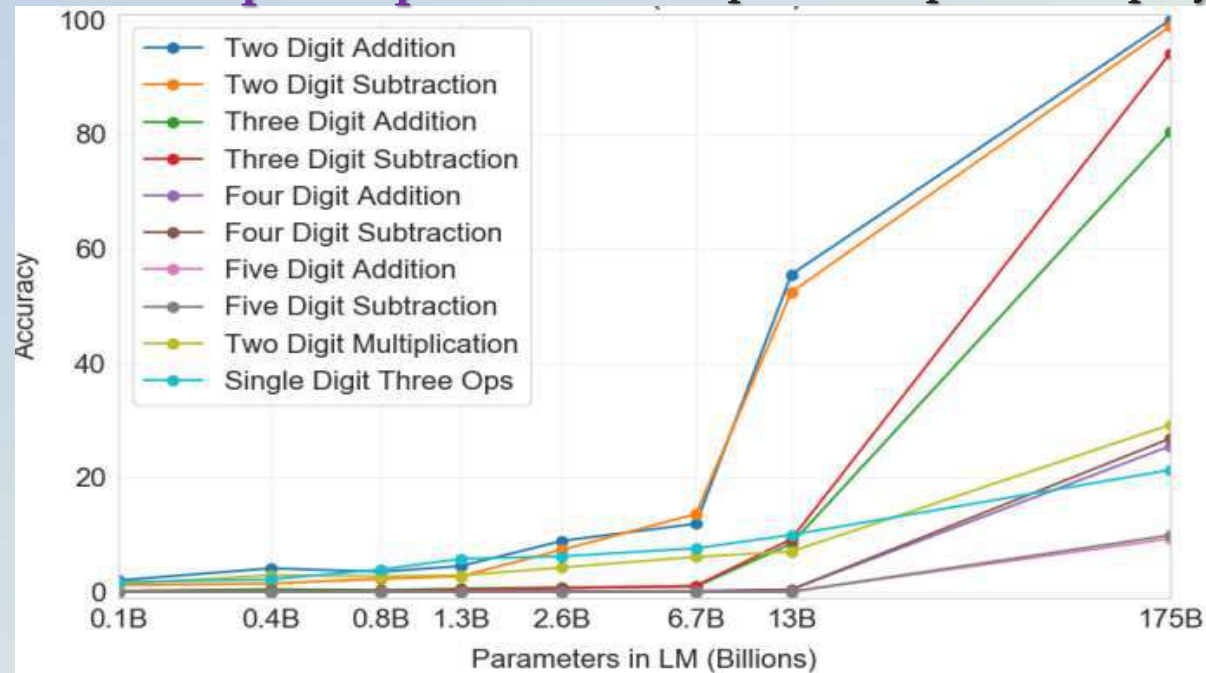


Агенты ИИ могут расширить возможности человека в сфере здравоохранения, образования, программирования, прогнозирования и др.

<https://www.weforum.org/stories/2024/12/ai-agents-empower-human-potential-while-mitigating-risks/>

	Традиционное эконометрическое прогнозирование	Экономическое прогнозирование на основе больших языковых моделей
Источники данных	Структурированные данные (официальная статистика, опросы и т.д.), ограниченные по объему	Большие неструктурированные данные (новости, статьи, финансовые отчеты, социальные сети и т.д.). Возможно обновление в режиме реального времени.
Формализация	Как правило линейные зависимости, предопределенные на основе теоретических гипотез. При этом сложные закономерности могут быть упущены	Большие языковые модели могут самостоятельно выявлять сложные нелинейные взаимосвязи и скрытые закономерности.
Реалистичность	Модели редко обновляются и могут давать серьезные погрешности для нестабильных периодов	Обучение может происходить в режиме реального времени, модели обновляются с использованием новых данных
Интерпретируемость	Прозрачность получаемых результатов - понятные уравнения, рассчитанные с помощью унифицированных методов	Черный ящик – сотни млрд необъяснимых параметров
Точность результатов	Хорошо работает в стабильных условиях	Могут давать результаты высокой точности, что особенно важно в нестабильные периоды, но возможна галлюцинация; ангажированность
Исследовательский вклад	Определение формы модели, отбор переменных, формулировка гипотез	Минимизация вмешательства человека

Точность решения 10 арифметических задач моделями различных размеров: точность резко возрастает при увеличении параметров



Tom B. Brown et al. Language Models are Few-Shot Learners
<https://doi.org/10.48550/arXiv.2005.14165>

OpenAI: доступны веса и технические детали нейронной сети **GPT-2**, технические подробности **GPT-3**, для **GPT-4** пока нет информации.

Generative Pre-trained Transformers (GPT)

Генеративные предварительно обученные трансформеры

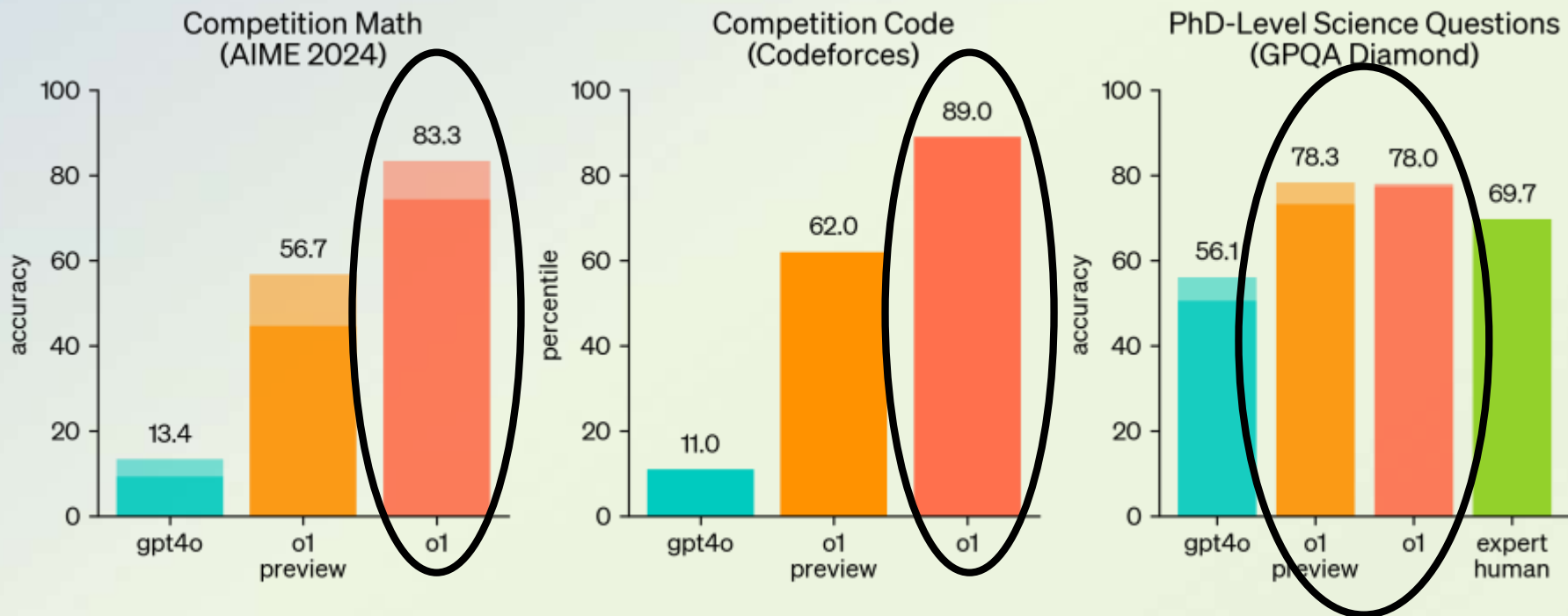
Средний путь преобразования данных через ChatGPT ≈ **400 слоев**

Языковые модели	Количество параметров	Тренировочные данные	Дата выпуска
GPT-1	0,117 миллиарда	4,5 ГБ текста из 7000 книг разных жанров	11 июня 2018 г.
GPT-2	1,5 миллиарда	40 ГБ текста, 8 миллионов документов	14 февраля 2019 г.
GPT-3	175 миллиардов	570 ГБ , в основном CommonCrawl, WebText, английская Википедия	11 июня 2020 г.
ChatGPT (на базе GPT-3.5)	175 миллиардов	Дополнительное обучение	30 ноября 2022 г.
GPT-4	1,7 трлн	Неизвестно	14 марта 2023 г.

GPT-4o1:

- ❑ выстраивает длинную внутреннюю цепочку мыслей, разбивает запрос на части, рассматривает несколько вариантов
- ❑ 89-й перцентиль - соревновательное программирование (Codeforces)
- ❑ входит в число 500 лучших студентов США математической олимпиады США (AIME)
- ❑ превосходит уровень PhD в решении задач по физике, биологии и химии (GPQA).

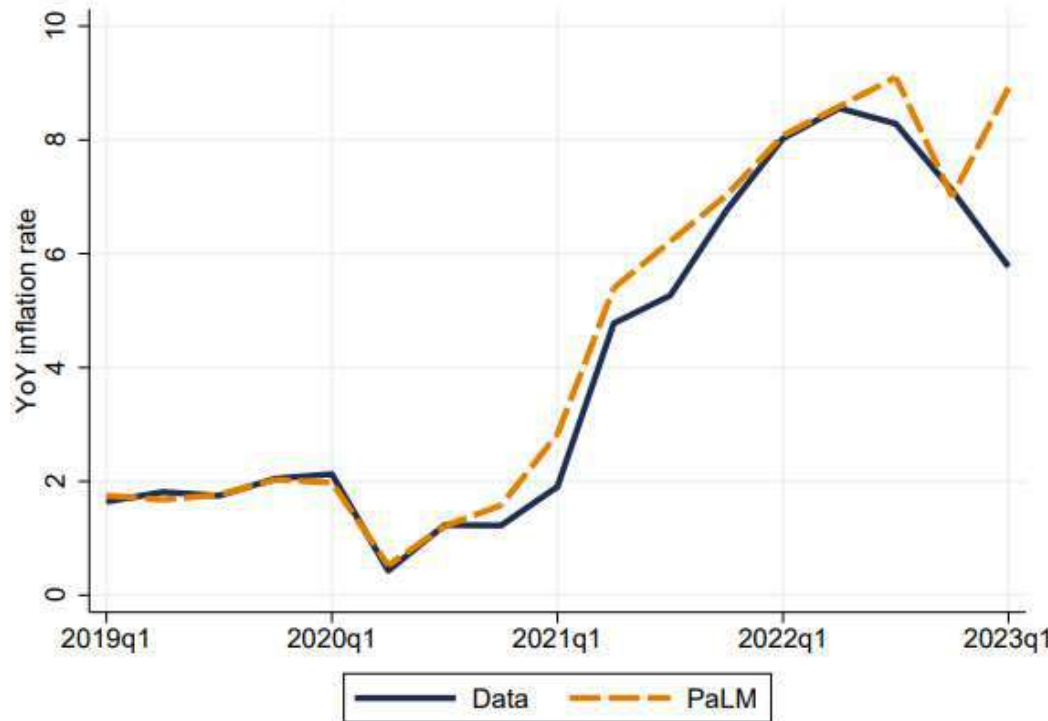
<https://openai.com/index/learning-to-reason-with-llms/>



Использование языковых моделей

Pathways Language Model Google AI

Исследование Федерального резервного банка Сент-Луиса от июля 2023 г. по ретроспективным прогнозам инфляции за период 2019–2023 гг. с использованием Google PaLM. В среднем прогнозы LLM показывают меньшие среднеквадратические ошибки, чем альтернативные прогнозы. Использование LLM рекомендуется в дополнение к другим прогнозам.



Faria e Castro, M., Leibovici, F., 2023; Artificial Intelligence and Inflation Forecasts, Federal Reserve Bank of St. Louis Working Paper 2023-015. URL <https://doi.org/10.20955/wp.2023.015>

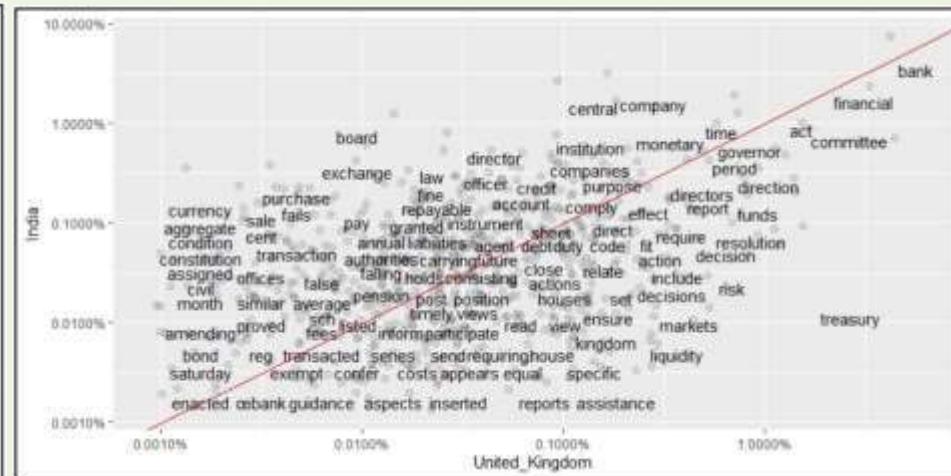
Исследование Национального банка Словакии

- Применение ChatGPT для реализации DSGE моделей и выявления бизнес-циклов
- Настройка эконометрических моделей

МВФ с использованием технологий ИИ провел анализ базы документов «*Central Bank Legislation Database*» по 175 центральным банкам по 273 специальным категориям.

Анализ законодательных документов ЦБ, включающих биграммы о «независимости» и «автономии»

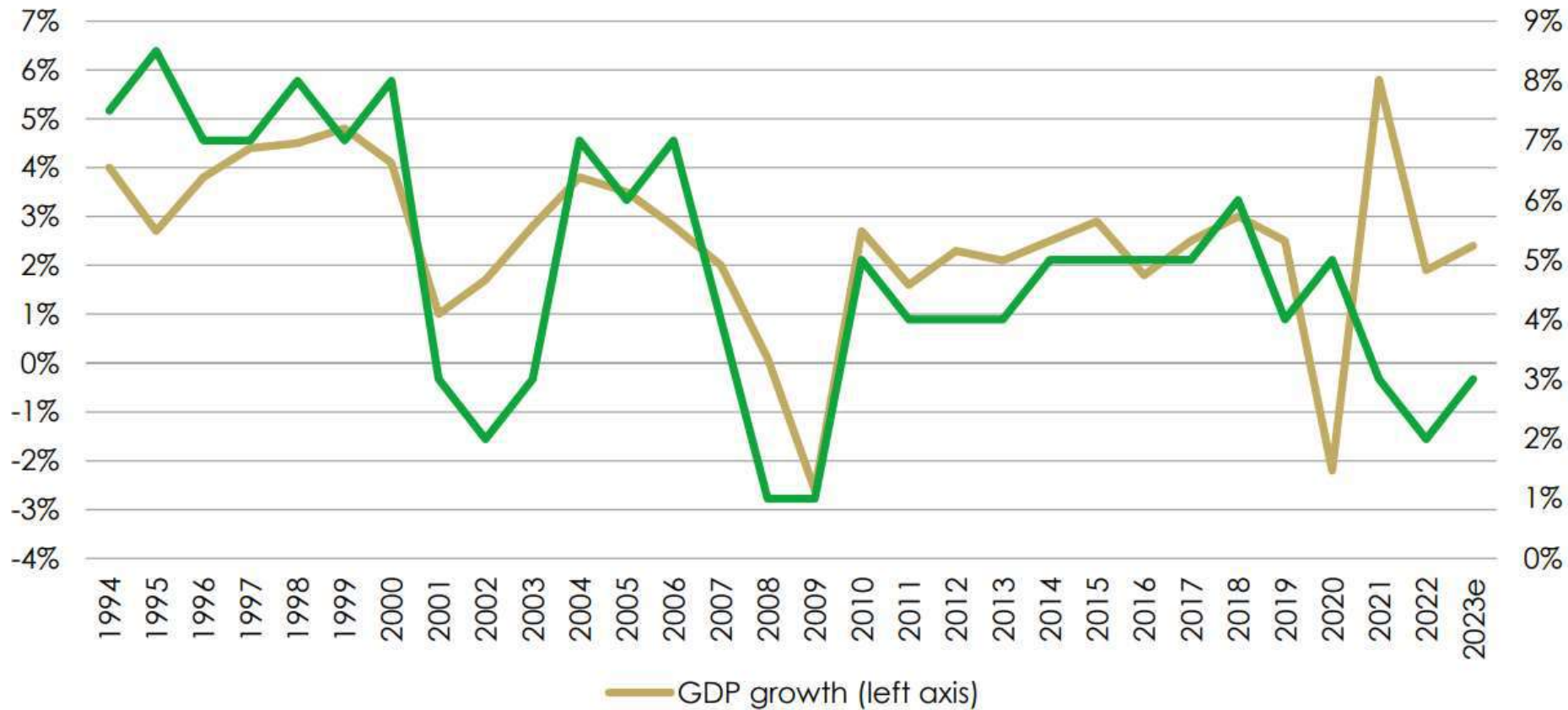
Количество биграмм



Al Ajmi, K., J. Deodoro, A. Khan, K. Moriya, 2023, **Predicting the Law: Artificial Intelligence Findings from the IMF's Central Bank Legislation Database**. IMF Working Paper 23/241. Washington, D.C.: International Monetary Fund

Прогноз ВВП США с использованием LLM GPT-4 (инвестиционная компания Liontrust)

Исходные данные: протоколы заседаний ФРС США, содержащие информацию о сложившейся ситуации в экономической системе



Source: Liontrust research and Bloomberg, December 2023

<https://www.liontrust.co.uk/-/media/liontrust/files/insights/blogs/2023/12/2312-james-dowey-using-large-language-models-to-forecast-us-growth.pdf>

Wisdom of the Silicon Crowd: LLM Ensemble Prediction Capabilities Rival Human Crowd Accuracy (17/06/2024)

Прогнозы **12 LLM VS** прогнозы от **925 прогнозистов** с высокой репутацией из системы Metaculus (американской системы массового онлайн-прогнозирования)

Примеры вопросов:

- ☐ Будет ли повышена ставка по федеральным фондам США до 14 декабря 2023 года? (**НЕТ**)
- ☐ Достигнет ли Биткойн \$40 000 до 1 января 2024 года? (**ДА**)

И примеры вопросов связанных с политикой:

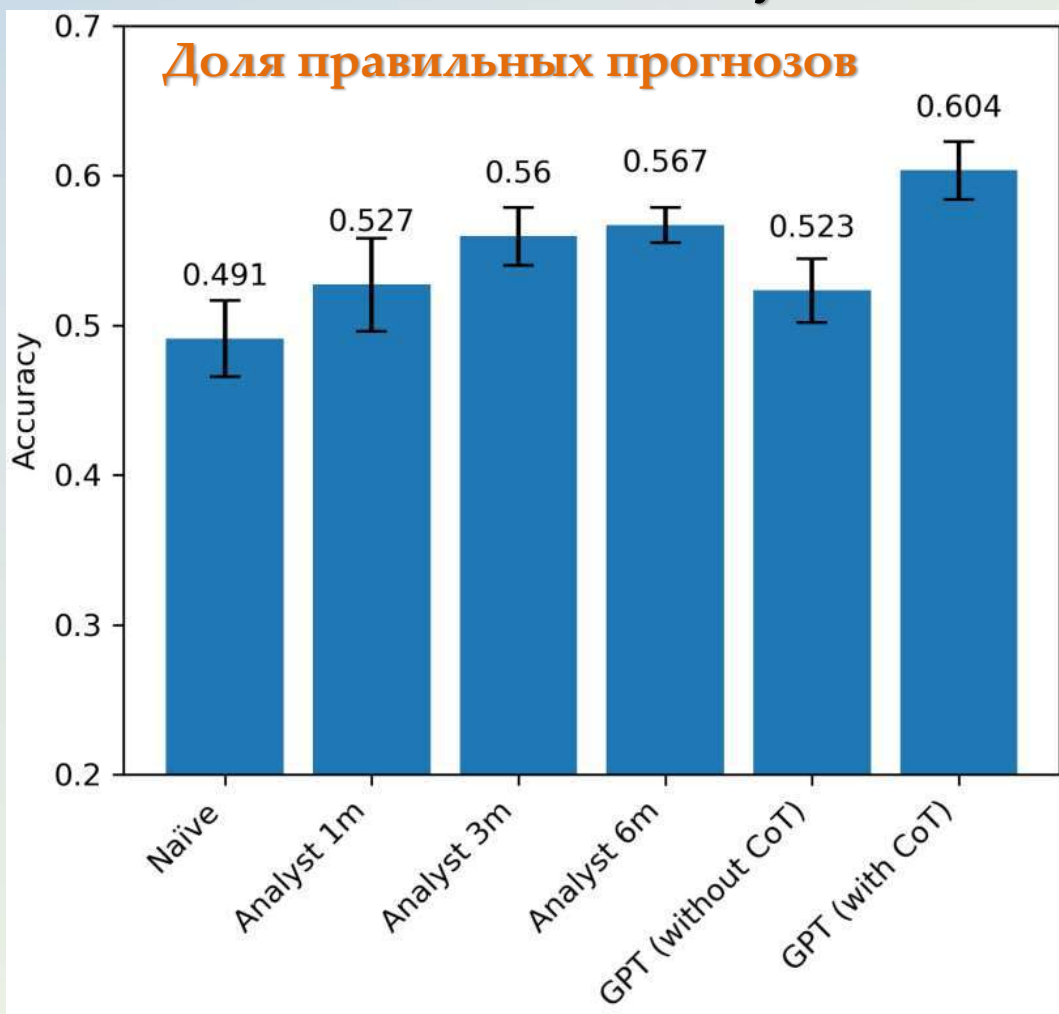
- ☐ Потеряет ли ХАМАС контроль над сектором Газа до 2024 года? (**НЕТ**)
- ☐ Посетит ли Владимир Зеленский Израиль до 2024 года? (**НЕТ**)
- ☐ Проведет ли Дональд Трамп хотя бы один час в тюремной камере до 1 января 2024 года? (**НЕТ**)

Результаты показали, что группа LLM превосходит прогнозистов-людей в % правильных ответов

Financial Statement Analysis with Large Language Models (University of Chicago, 20/05/2024)

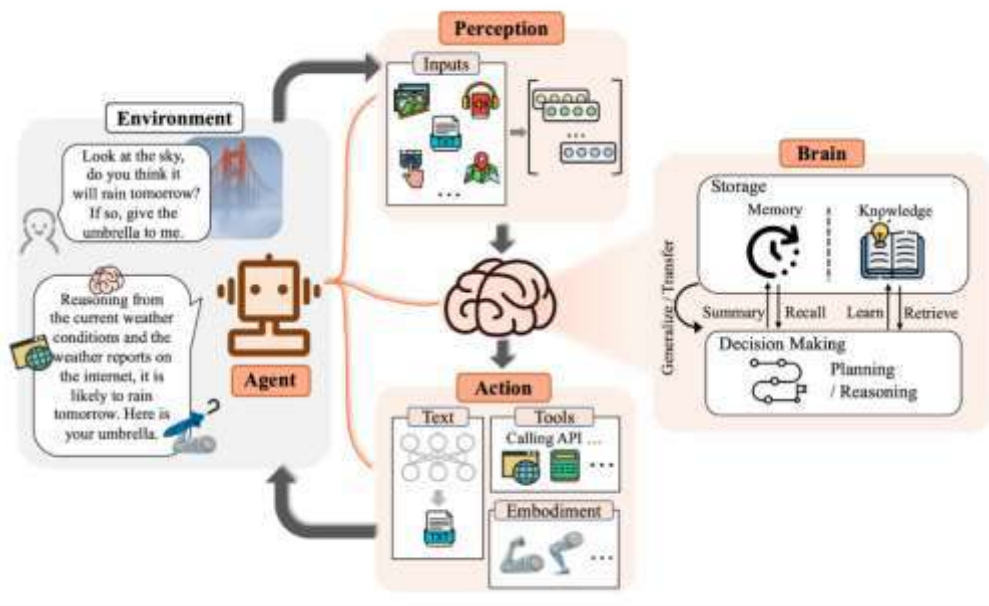
В ряде случаев LLM превосходят финансовых аналитиков в части прогноза изменения доходов.

GPT vs. Human Analysts



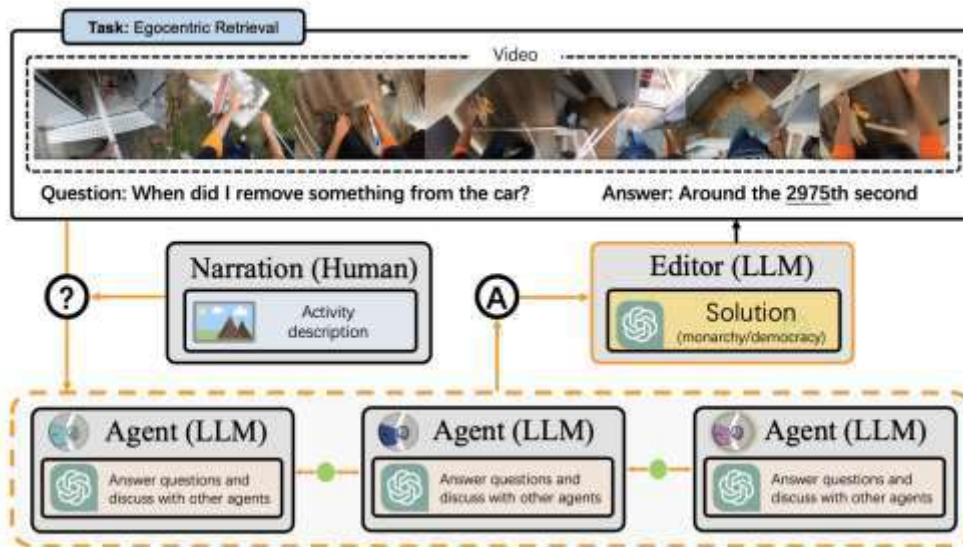
Внедрение **цепочек мыслей** (chain-of-thought), эффективно «обучающих» модель позволяет имитировать финансового аналитика – выполнять набор инструкций:
вычислять финансовые показатели, выявлять тенденции

Kim, Alex G. and Muhn, Maximilian and Nikolaev, Valeri V., Financial Statement Analysis with Large Language Models (May 20, 2024). Chicago Booth Research Paper, Fama-Miller Working Paper, Available at SSRN: <https://ssrn.com/abstract=4835311> or <http://dx.doi.org/10.2139/ssrn.4835311>



- ## Автономные агенты на базе LLM
- LLM является **мозгом** агента, а наш мир – **средой** функционирования
 - **Планирование**
 - **Получение** (текст, камеры, сенсоры) и **обработка** информации
 - Выбор **подходящего действия** для текущей ситуации
 - Оценивание **эффективности** на основе обратной связи
 - Запись в **краткосрочную** и **долгосрочную** память

Обратная связь для улучшения навыков и улучшения исследования среды



- Обратная связь от **человека** (недостаток – задержка на написание текста)
- Обратная связь от **модели**
- Интеграция **долгосрочного планирования** и **системы обработки обратной связи** позволяет серьезно развивать агентов
- **Модуль памяти:** агент сам пишет подпрограммы на JavaScript для будущего исполнения
- **Планирование через дебаты** между агентами (устойчивый результат через 3-4 итерации)

Wang et al. "Voyager: An Open-Ended Embodied Agent with Large Language Models" arXiv preprint [arXiv:2305.16291](https://arxiv.org/abs/2305.16291) (2023).

- ❑ Yue Guo and Yi Yang. 2024. EconNLI: **Evaluating Large Language Models on Economics Reasoning**. In Findings of the Association for Computational Linguistics ACL 2024, pages 982–994, Bangkok, Thailand and virtual meeting. Association for Computational Linguistics.
- ❑ Ash, E, S Hansen and Y Muvdi (2024), ‘DP19479 **Large Language Models in Economics**’, CEPR Discussion Paper No. 19479. CEPR Press, Paris & London. <https://cepr.org/publications/dp19479>
- ❑ Korinek, Anton. **"Generative AI for Economic Research: Use Cases and Implications for Economists."** Journal of Economic Literature, 61 (4): 1281–1317.
- ❑ Andrea Carriero, Davide Pettenuzzo, Shubhranshu Shekhar Macroeconomic **Forecasting with Large Language Models** <https://doi.org/10.48550/arXiv.2407.0089>
- ❑ Jesus Fernandez-Villaverde **Large Language Models in Economics** March 5, 2024, https://www.sas.upenn.edu/~jesusfv/Course_Oxford_July_2023_Block_13.pdf
- ❑ Kim, Alex G. and Muhn, Maximilian and Nikolaev, Valeri V., **Financial Statement Analysis with Large Language Models** (May 20, 2024). Chicago Booth Research Paper, Fama-Miller Working Paper, Available at SSRN: <https://ssrn.com/abstract=4835311> or <http://dx.doi.org/10.2139/ssrn.4835311>
- ❑ Andrea Carriero & Davide Pettenuzzo & Shubhranshu Shekhar, 2024. **Macroeconomic Forecasting with Large Language Models Papers** 2407.00890, arXiv.org, revised Jan 2025.
- ❑ Ming Jin, Shiyu Wang, Lintao Ma, Zhixuan Chu, James Y. Zhang, Xiaoming Shi, Pin-Yu Chen, Yuxuan Liang, Yuan-Fang Li, Shirui Pan, Qingsong Wen Time-LLM: **Time Series Forecasting by Reprogramming Large Language Models** <https://doi.org/10.48550/arXiv.2310.01728>
- ❑ Jens Ludwig, Sendhil Mullainathan, Ashesh Rambachan **Large Language Models: An Applied Econometric Framework** <https://doi.org/10.48550/arXiv.2412.07031>
- ❑ Gao, C., Lan, X., Li, N. et al. **Large language models empowered agent-based modeling and simulation: a survey and perspectives** <https://www.nature.com/articles/s41599-024-03611-3>

Развертывание больших языковых моделей в ситуационном центре ЦЭМИ РАН



Open Web UI – open-source вебинтерфейс с открытым исходным кодом, позволяющий работать с несколькими генеративными моделями как через API, так и локально.

На текущий момент в системе подключены модели от **OpenAI**, **Anthropic** и **DeepSeek**, а также реализован RAG (Retrieval-Augmented Generation), дополняющий запросы отобранными материалами (специализированные статьи). Этот метод позволяет улучшать качество ответов модели за счет поиска релевантной информации в базе документов перед генерацией ответа.

Для этого используется embedding-модель **text-embedding-3-large** от OpenAI. Embedding — это представление текста в виде числовых векторов, позволяющее модели эффективно искать релевантную информацию.

Информационные войны современности и моделирование распространения новостей

Согласно отчету Global Digital Report (начало 2025 г.) Интернетом пользуются примерно **5,56 млрд человек**, что составляет **67,9%** от населения планеты. Число аккаунтов в социальных сетях достигло **5,24 млрд**, но при этом отмечается, что эта цифра не отражает количества уникальных пользователей, поскольку содержит в себе дублирующие учетные записи и фейковые аккаунты.

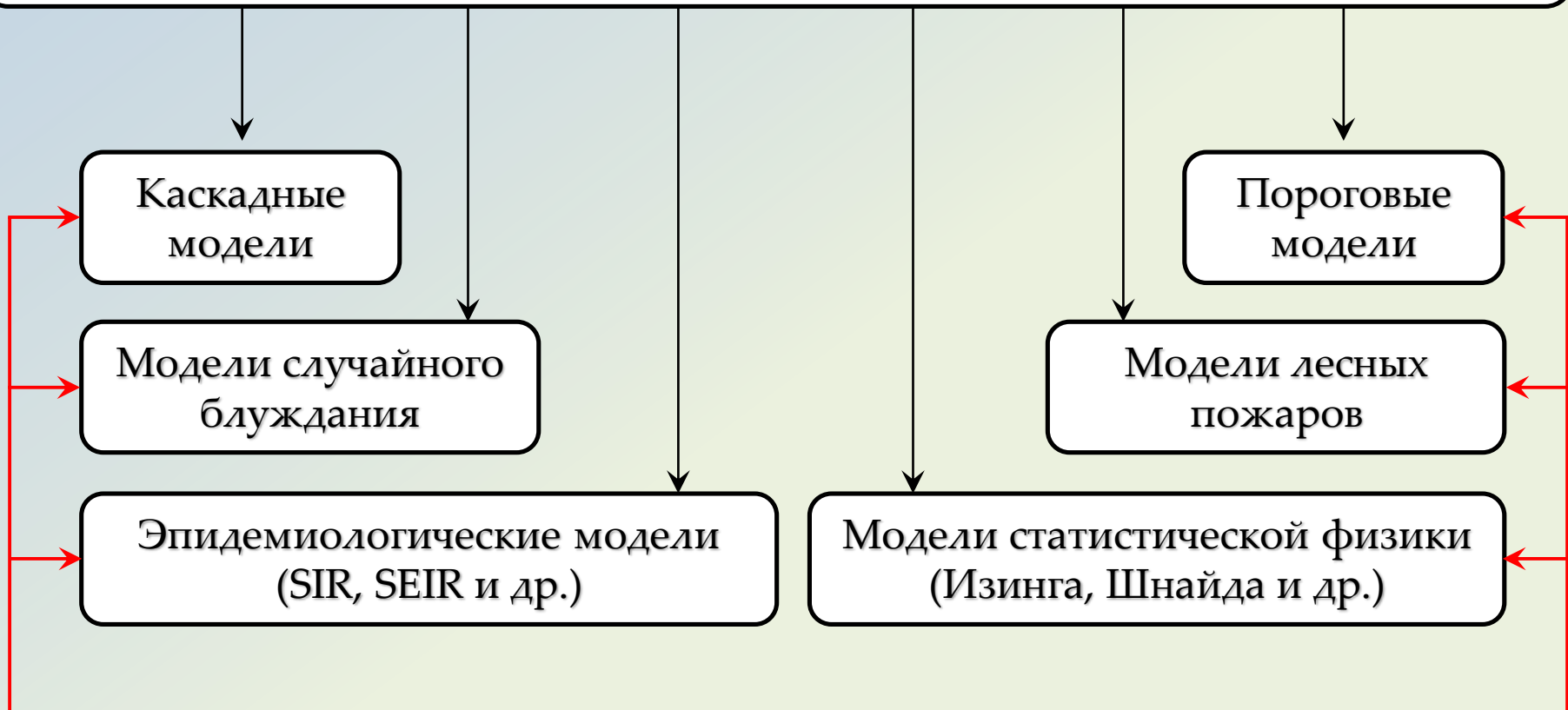
В среднем пользователь Интернета за счет подключенных к сети приложений, установленных на компьютер и/или смартфон, проводит в сети **6 часов 38 минут**. Для сравнения, среднее время просмотра телевизора составляет **3 часа 12 минут**.

Forbes: СВО России определена как «Первая война в социальных сетях» (First social media war)

Forbes: Конфликт во Вьетнаме заработал репутацию «Первой телевизионной войны» (First television war)

Типологизация моделей распространения информации

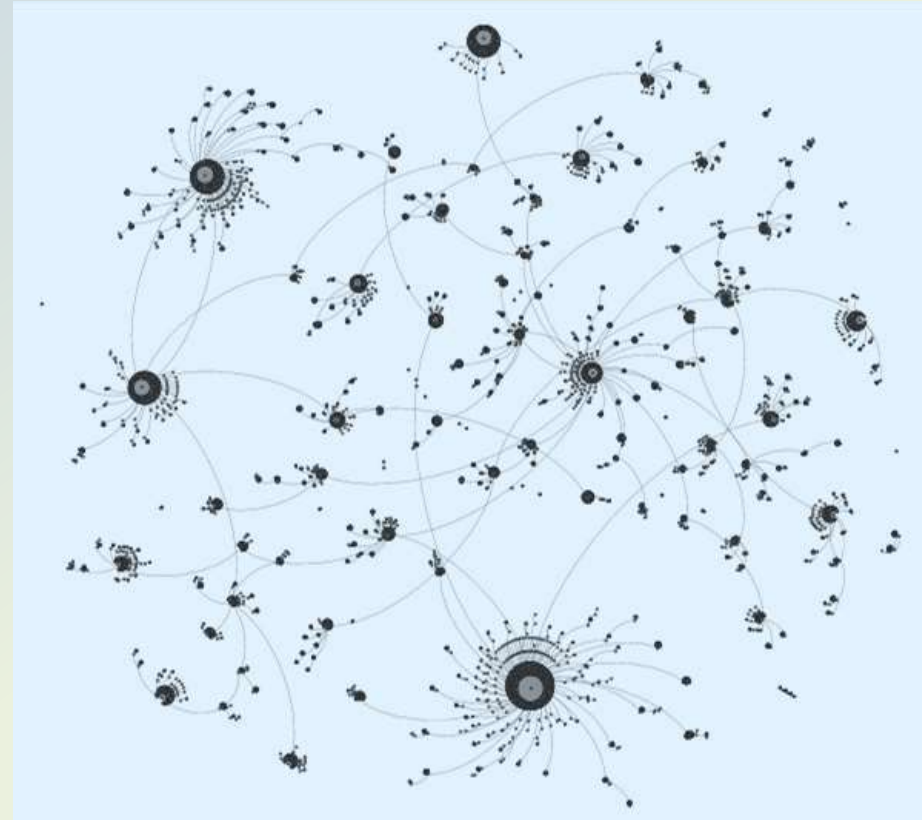
Модели распространения информации в онлайн социальных сетях



Способы численного разрешения: агент-ориентированные, дискретно-событийные модели, дифференциальные уравнения, клеточные автоматы, нейронные сети и др.

Моделирование информационных войн (агент-ориентированных модели)

- ❖ Скорость распространения информации (топология социальной сети, плотность сети, количество кластеров, влияющих на мощность информационных потоков, механизм ввода информации)
- ❖ Способы распространения - линейные пороговые, случайного блуждания (диффузионное влияние)
- ❖ Расчет количества внедренных лидеров мнений, необходимого для дискредитации противоположных точек зрения и поляризации общественного мнения
- ❖ Расчет точного времени для повышения эффективности от оглашения информации



- ❑ **Razaque A., Rizvi S., Khan M.J., Almiani M., Rahayfeh A.A. (2022): State-of-art review of information diffusion models and their impact on social network vulnerabilities**, Journal of King Saud University - Computer and Information Sciences, Volume 34, Issue 1, 2022, Pages 1275-1294, ISSN 1319-1578, <https://doi.org/10.1016/j.jksuci.2019.08.008>
- ❑ **Arruda H.F., Cardoso F.M., Arruda G.F., Hernandez A.R., Costa L.D., Moreno Y. (2022): Modelling how social network algorithms can influence opinion polarization**, Information Sciences, Volume 588, 2022, Pages 265-278, ISSN 0020-0255, <https://doi.org/10.1016/j.ins.2021.12.069>.
- ❑ **Pond T., Magsarjav S., South T., Mitchell L., and Bagrow J.P. (2020): Complex contagion features without social reinforcement in a model of social information flow**. Entropy, 22(3):265, 2020.
- ❑ **Röchert D., Cargnino M., Neubaum G. (2022): Two sides of the same leader: an agent-based model to analyze the effect of ambivalent opinion leaders in social networks**. J Comput Soc Sc 5, 2022, 1159–1205. <https://doi.org/10.1007/s42001-022-00161-z>

Инициализация модели и установка начального состояния

1 Считывание информации из базы данных

2 Вычисление структуры цифрового двойника общества

3 Создание агентов с соответствующими значениями свойств и связями между ними

Изменение управляющих параметров

8 Вывод результатов на экран и ожидание действий от пользователя

Расчет шага(ов) модельной симуляции

4 Для каждого агента определение количества контактов, обладающих новой информацией

5 Для каждого агента рассчитывается значение логистической функции с учетом его характеристик и количества контактов

6 Для каждого агента реализуется процедура, определяющая преодолел ли поступающий к ним импульс определенный порог

7 Сбор и обработка статистики по агентам и пересчет всех программных коллекций



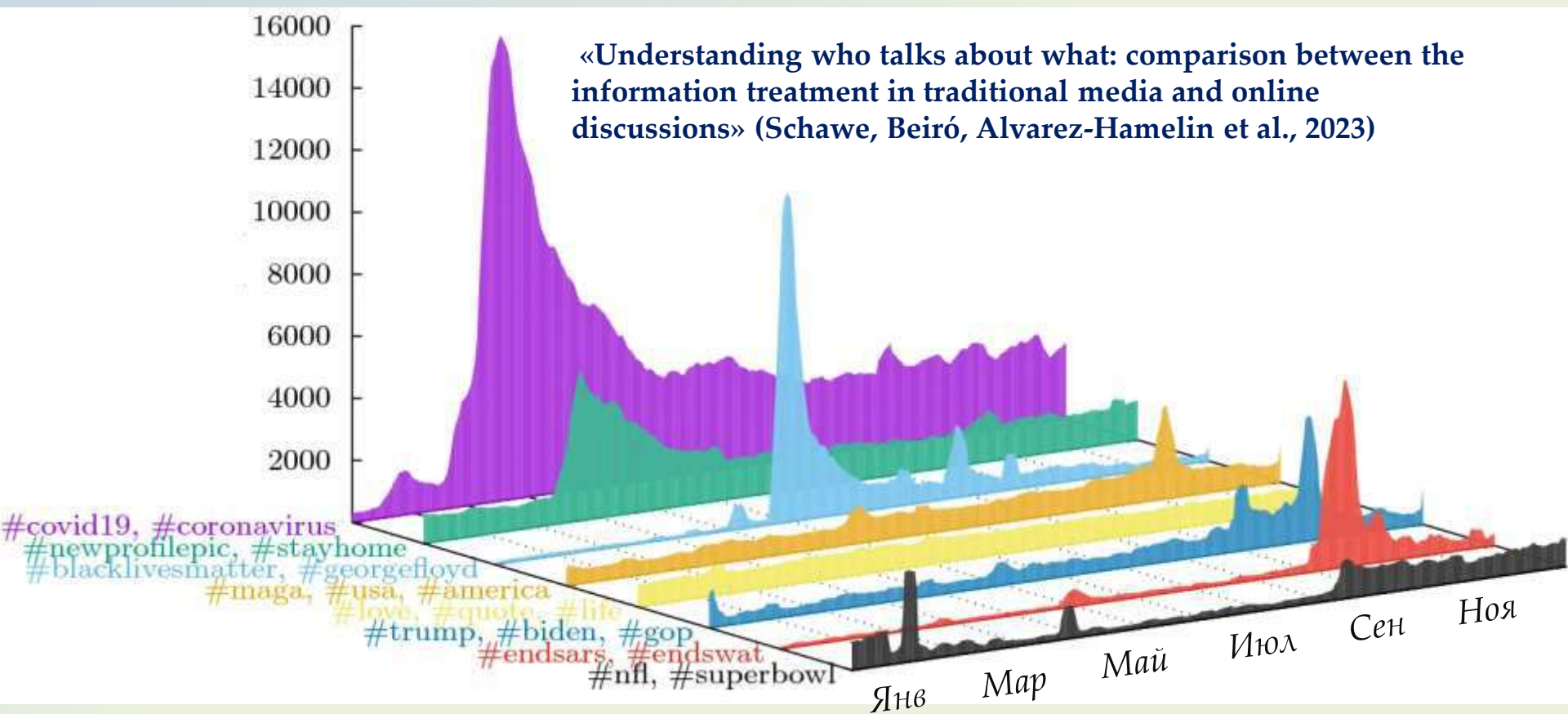


Темп прироста информированных агентов: базовый вариант и 3 сценария (ось абсцисс – такты модельного времени, ось ординат – темп прироста, в %)

С точки зрения скорости распространения информации самым действенным сценарием является первоначальный охват целевой аудитории. Вовлечение лидеров мнений является более эффективным механизмом, чем повышение уровня репутации агентов.

Результаты исследования сети Twitter в части анализа динамики обсуждения наиболее значимых общественных процессов **2020–2021 гг.** (вертикальная ось – количество уникальных участников сети, использующих соответствующие хэштеги, отложенные по одной из горизонтальных осей)

Был обработан репрезентативный участок Twitter, включающий около **8 млн** пользователей, опубликовавших около **полумиллиарда твитов**.



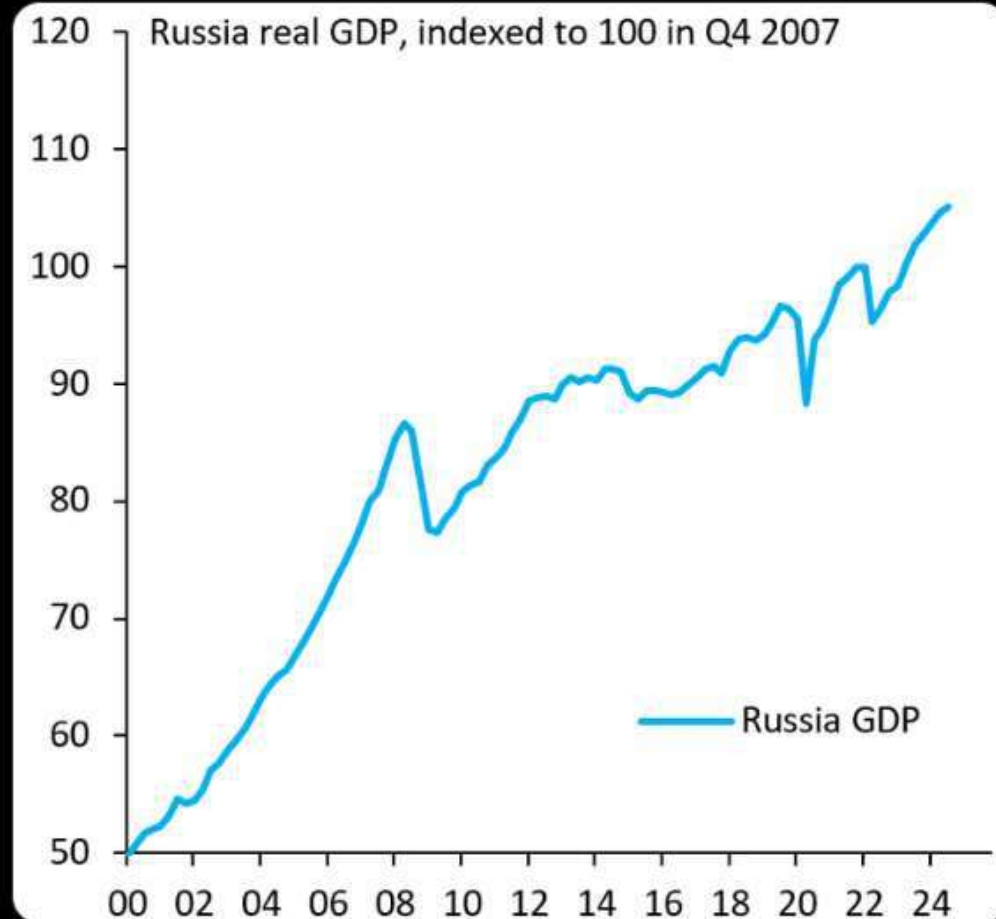


Робин Брукс
@robin_j_brooks

x1 ...

Необходимо запретить любые научные работы, которые интерпретируют устойчивый ВВП России как признак того, что санкции не работают. Устойчивость России является симптомом нерешительности Запада в отношении санкций. Если бы Запад взялся за дело, Россия бы сейчас рухнула...

[Перевести пост](#)



8:28 PM · 31 дек. 2024 г. · 691,3 тыс. просмотров

Robin Brooks is a senior fellow in the Global Economy and Development program at the Brookings Institution.

Chief Economist and Managing Director, **Institute of International Finance**

Chief FX Strategist and Managing Director, **Goldman Sachs**, Global Investment Research

Economist, Brevan Howard Asset Management

Vice President, Goldman Sachs Asset Management

Senior Economist, Asia and Pacific and Research, **International Monetary Fund**

Спасибо за внимание!