

GEO-оптимизация - основные механики

1. Как ChatGPT выбирает контент

- В бесплатной версии модель опирается на тренировочные данные (до середины 2024) — всё общедоступное в сети.
- В платной версии есть Web Browsing — модель обращается к интернету в реальном времени, берёт свежие данные.
- При запросах часто происходит переформулирование и уточнение, чтобы дать более точный ответ.

2. Факторы, влияющие на попадание сайта в выдачу GPT

- Индексация в **Bing** — крайне важно, т. к. Web Browsing часто использует Bing.
- Охват семантики — работа с семантическим ядром — основа успеха не только в SEO, но и в GEO, LLM-модели оценивают не только полноту ответа на конкретный запрос, но и полноту тематического охвата.
- Корректный HTML-markup, логическая структура, семантические теги, структурированные данные (FAQ/Q&A schema и др.), особенно важна Оптимизация под вопросы (Q&A-контент) - Форматирование под LLM-чтение: Чёткая структура: H2, H3, списки, таблицы. Короткие абзацы с одной мыслью → GPT проще «резать» текст. Заголовки и подзаголовки — помогают ИИ «достать» именно нужную информацию.
- Работа с «сущностями» — бренды, темы, эксперты, форматы. Чтобы ИИ распознавал, о чём сайт и кто его делает.
- Экспертность, факты, статистика, исследования — сайты с качественным экспертным контентом выше оценены. Формат «данные + выводы»
- Избегать «воды» — LLM плохо цитируют растянутый текст.
- Обновление контента — особенно в темах с быстрыми изменениями (например, финансовые, YMYL, технологии).
- Фокус на узкой тематике — сайты с тематическим фокусом воспринимаются как более релевантные.
- Авторство, прозрачность, доверие (опыт автора, команды, ссылки с авторитетных источников)
- Упоминания на авторитетных площадках и форумах, присутствие в рейтингах, новостях.
- llms.txt — это экспериментальный файл-протокол (похож на robots.txt), где владелец сайта указывает правила для LLM, что индексировать, а что нет.

Пример рекомендаций по GEO-оптимизации

В данном блоке рекомендаций приводятся требования для попадания сайта в ИИ ответы чат-бота ChatGPT и других ИИ ботов.

Контентная составляющая

Уникальный контент

На сайте должен размещаться уникальный контент, который несет какую-то ценность для пользователя. ИИ сгенерированные тексты желательно не публиковать, тк подобные тексты, как правило, поверхностные, содержат много “воды” и мало конкретики, такие тексты, даже безупречные с точки зрения фактчекинга, все равно нельзя признать качественными.

Контент должен иметь “добавленную ценность”, например, если мы делаем обзор на web3-шлюз, это должен быть не просто обзор работы криптовалютного шлюза, которых в Сети может быть несколько десятков. А уникальный взгляд на вещи, или пространный комментарий приглашенного эксперта. Т.е. каждый материал должен чем-то выделяться на фоне других статей в интернете.

Длинный “хвост” запросов и разговорные фразы

Рекомендуется в текстах использовать длинные вхождения фраз, например *“which cryptocurrency gateway is best for businesses in the field of ecommerce”*. Это так называемые long-tail запросы, т.е. длинные разговорные фразы/запросы (conversational queries), которые ИИ боты часто берут для своих ответов на соответствующие вопросы пользователей.

Подробности <https://searchengineland.com/adapt-seo-strategy-stronger-ai-visibility-453641>.

LSI копирайтинг

Как для обычного поиска, так и для ИИ чат-ботов в материалах следует использовать LSI фразы. Это фразы, имеющие отношение к теме статьи, основному продвигаемому термину. С помощью таких фраз ботами устанавливаются ассоциации между терминами, увеличивается релевантность статьи какой-либо теме.

Например, для темы “Криптовалютный шлюз” такими LSI фразами будут являться:

- криптовалюта
- процессинг
- платежный шлюз
- модуль оплаты
- крипто эквайринг

и проч.

Использование связанных терминов в статьях и коммерческих текстах дает преимущество подобным материалам.

Оформление текстов

При размещении текстов и в целом контента страниц следует придерживаться стандартов оформления. Это соблюдение вложенности заголовков (H1-H6), использование маркированных списков, таблиц и других приемов оформления текстов.

Подобное грамотное оформление делает контент легко читаемым для ИИ.

Кроме того, чтобы выделяться в ИИ выдаче Google Overview требуется оформлять контент с использованием: статистики, цитат экспертов.

Внешние размещения

Площадки

Следует публиковать контент на тех площадках, которые ИИ часто цитирует (и на контенте которых обучался), это: Reddit, Quora и проч.

Представитель проекта может выступить топикстартером по какой-либо теме. Либо отвечать от лица фирмы на вопросы пользователей, где это возможно оставлять ссылку на продвигаемый сайт в качестве пруфа.

Помимо указанных выше сайтов, рекомендуется публиковаться на тематических ресурсах с высоким трастом. А также просто общеизвестных сайтах авторитетных сайтах, которые участвуют в обучающих выборках ИИ, такие как en.wikipedia.org.

Важно: Необходимо публиковать оригинальный, интересный контент, который вносит что-то новое в уже существующие в сети Интернет статьи по теме. Например, если это PR статья по криптовалютам регулярным платежам, нужно привести примеры подписок, которые можно оплачивать с помощью этого способа. Или дать прямую речь эксперта по какой-либо теме.

Здесь отметим, что частое упоминание бренда в позитивном ключе на рассмотренных сайтах повышает вероятность его появления в ответах ИИ.

Техническая составляющая

Микроразметка

На сайте требуется внедрить микроразметку по схемам:

- [BreadcrumbList](#)
- [Article](#)
- [WebPage](#)
- [FAQ](#)
- [OpenGraph](#)
- [Organization](#)

Наличие размеченных данных позволяет попадать сайту как в ИИ ответы Google, так и в ответы ChatGPT. Размеченные данные сразу передают ботам необходимую информацию о содержании страницы в структурированном виде.

HTML карта

HTML карта упрощает для ботов понимание структуры сайта, а также предоставляет быстрый доступ к ключевым разделам ресурса для краулеров.

LLMs.txt

сентябрь 2025 - на данный момент не существует общепринятого формата файла.

llms.txt - это аналог robots.txt, но для LLM (большие языковые модели), помогающий AI-моделям лучше понимать контент сайта. llms.txt улучшает видимость в AI-поиске, предоставляя чистую, структурированную версию контента сайта, которую большие языковые модели могут легко обрабатывать и понимать. По сути это список важного контента в простом формате, чтобы модели могли легко его парсить, цитировать и использовать в генерации ответов.

GPTBot, PerplexityBot и ClaudeBot ежедневно активно парсят эти файлы, используя их для лучшей интерпретации и индексации контента сайта.

Файл llms.txt требует специального markdown-форматирования (остается только чистый текст из HTML-страниц: без стилей и скриптов в корневой директории сайта). Отметим, что формат Markdown llms.txt очень близок к простому тексту, с минимальной разметкой или форматированием, но все же предоставляет возможность отображать важную структуру документа. Основные LLM, такие как GPT-4o от OpenAI, изначально «говорят» на Markdown и часто включают Markdown в свои ответы. Это свидетельствует о том, что они были обучены на огромных объемах текста, отформатированного в Markdown, и хорошо его понимают. Дополнительным преимуществом является то, что конвенции Markdown также очень эффективны с точки зрения траты токенов (т.е. ИИ бот сможет забрать больше информации).

Директория llms.txt (<https://directory.llmstxt.cloud/>) показывает более 70 крупных AI-компаний, уже использующих этот стандарт, включая Cursor AI, Anthropic и Perplexity.

Правильный llms.txt содержит основные элементы:

- название сайта/проекта,
 - краткое описание сайта (Обзор сайта, ключевые факты, миссия или основная информация),
- Список важных страниц, документов или разделов в формате Markdown-ссылок (см. ниже).

Синтаксис:

Заголовок и подзаголовки выделяются знаком решетки: # и ## соответственно, описание сайта (дисклеймер) знак >, название страницы в квадратных скобках []. Пример:

Название сайта

> Описание сайта

Название раздела

[Link title](https://example.com/red-toys-for-kids) - Optional link details

Пример

Toys for kids

> Online store selling handmade toys for children aged 5 to 16. Bright toys from Russian craftsmen.

Catalog

[Red toys](https://example.com/red-toys-for-kids) - Red toys for kids with big hat

По мере необходимости файл может обновляться.

Пример эффективной имплементации llms.txt можно увидеть здесь (<https://wordlift.io/llms.txt>). Еще пример <https://www.fastht.ml/docs/llms.txt>.

Пример файла [www.coinbase.com](https://docs.cdp.coinbase.com/llms.txt) - <https://docs.cdp.coinbase.com/llms.txt>.

Сайт организации разработчика стандарта файла <https://llmstxt.org/>.

Важно

Инструменты для разработчика приводятся в этой статье <https://wordlift.io/blog/en/mastering-llms-txt-for-genai-optimized-indexing/>. Среди генераторов файла Markdown llms.txt выделяются:

- <https://sitespeak.ai/tools/llms-txt-generator>
- <https://github.com/microsoft/markitdown> - решение от Microsoft
- <https://wordlift.io/generate-llms-txt/>
- <https://www.firecrawl.dev/playground>

Можно протестировать разные, но зачастую проще составить файл вручную.

Запрет к индексации

Чтобы файл robots.txt не участвовал в индексировании классическими поисковыми системами (т.е. не появился в выдаче), страница должна отдавать следующий ответ сервера:

HTTP Header – X-Robots-Tag: noindex

Это предотвратит индексацию, но позволит ботам читать файл.

Важно: Не добавляйте Disallow: /llms.txt в robots.txt, так как это заблокирует краулинг для всех ботов, включая полезные ИИ.

Техническая часть

Для имплементации файл должен быть помечен как noindex для традиционных поисковых систем, оставаясь при этом доступным для AI-краулеров через добавление директивы robots.txt:

User-agent: Googlebot

Disallow: /llms.txt

Важно

Инструменты для разработчика приводятся в этой статье <https://wordlift.io/blog/en/mastering-llms-txt-for-genai-optimized-indexing/>. Среди генераторов файла Markdown llms.txt выделяются:

- <https://sitespeak.ai/tools/llms-txt-generator>
- <https://github.com/microsoft/markitdown> - решение от Microsoft
- <https://wordlift.io/generate-llms-txt/>
- <https://www.firecrawl.dev/playground>

Можно протестировать разные.

Подробнее про синтаксис и заполнение файла - <https://searchengineland.com/llms-txt-isnt-robots-txt-its-a-treasure-map-for-ai-456586>.

Robots.txt

В файл robots.txt необходимо добавить новые директивы.

Disallow: /llms.txt

User-agent: OAI-SearchBot

Allow: /llms.txt

User-agent: GPTBot

Allow: /

Allow: /llms.txt

Эти директивы напрямую разрешат ботам ChatGPT (и другим) парсить сайт, показывать его контент, а также обучаться на материалах сайта. А для поискового бота Google файл будет закрыт (чтобы избежать дублирования с основным контентом).

Справочная информация по ботам ChatGPT <https://platform.openai.com/docs/bots>.

Оптимизация под мобильные устройства

Большинство запросов к чат-ботам и к поиску Google происходит с мобильных устройств. В ответах ИИ показывает ссылки на страницы источники информации, поэтому очень важно, чтобы ресурс был адаптирован под смартфоны и имел высокую скорость загрузки на них (этот момент может быть оценен ИИ).

