

Научно-исследовательский препринт

Тема: Голосовой поиск как социальная архитектура обнаружения медиаконтента

Аннотация

Статья посвящена анализу голосового поиска как механизм социальной архитектуры цифровых медиасред. Рассматривается, как голосовой поиск трансформирует процессы обнаружения контента, пользовательскую навигацию, структуру выбора и распределение внимания внутри цифровых платформ.

В исследовании использовались музыкальные платформы (Spotify и YouTube Music), видеосервисы (Netflix и YouTube), smart TV-системы и Android TV, а также голосовые ассистенты (Siri и Алиса). Анализ осуществлялся при помощи пользовательских сценариев, интерфейсных маршрутов и структуры выдачи результатов при голосовом поиске.

Исследование показывает, что голосовой поиск формирует новую модель платформенного посредничества, в которой пользователь взаимодействует не с каталогом контента, а с алгоритмически организованным механизмом интерпретации намерений.

Введение

Цифровые медиасреды переживают переход от визуально-интерфейсной логики взаимодействия к голосовой и контекстной модели доступа к контенту. Если ранние цифровые платформы строились вокруг экранной навигации, поисковых строк и каталогов, то современные экосистемы больше

ориентируются на голосовые интерфейсы, работающие через естественный язык.

Голосовой поиск перестает быть вспомогательной функцией и становится инфраструктурой обнаружения контента. Пользователь больше не исследует платформу вручную, а формулирует намерение, которое интерпретируется алгоритмической системой.

Подобные изменения заметны в музыкальных сервисах, видеоплатформах и экосистемах smart TV. Например, при использовании Spotify или YouTube Music пользователь может не выбирать конкретную композицию, а формулировать настроение или контекст: «включи музыку для тренировки» или «поставь что-нибудь спокойное». Аналогичная ситуация наблюдается и в видеосервисах, где запросы вроде «что посмотреть вечером» постепенно вытесняют классический поиск по жанрам или названиям фильмов. В результате платформа получает значительно больше возможностей для алгоритмической интерпретации намерений пользователя.

Распространение голосового поиска связано сразу с несколькими процессами: развитием систем распознавания речи, ростом популярности smart TV и голосовых ассистентов, увеличением объема мобильного и фоновоего потребления контента, а также стремлением платформ сократить путь пользователя к медиаматериалу. Если ранние цифровые сервисы строились вокруг визуального исследования каталога, то современные платформы все чаще ориентируются на модель мгновенного алгоритмического ответа.

В отличие от текстового поиска, где пользователь способен самостоятельно исследовать каталог и сравнивать альтернативы, голосовой поиск чаще предлагает ограниченное количество результатов. Алгоритм становится посредником между пользовательским запросом и медиасредой, определяя, какой контент окажется видимым. Именно поэтому голосовой поиск

представляет интерес не только как технология интерфейса, но и как социальная архитектура распределения внимания.

Теоретическая рамка исследования

Исследование выполнено в рамках подхода социальных архитектур Н. Н. Калмыкова, рассматривающего цифровую среду как проектируемую систему организации пользовательского поведения. В данной логике интерфейс, алгоритм рекомендаций и поисковая система анализируются не как нейтральные технические инструменты, а как механизмы управления пользовательскими маршрутами и видимостью контента. Для анализа голосового поиска особое значение имеет понятие discoverability. Под discoverability понимается совокупность механизмов, определяющих, какой контент становится доступным и заметным для пользователя внутри платформы. В условиях алгоритмической медиасреды discoverability зависит не только от пользовательского запроса, но и от системы персонализации, ранжирования и рекомендаций.

Голосовой поиск усиливает роль подобных механизмов. Если текстовый поиск позволяет пользователю самостоятельно рассмотреть результаты, то голосовой поиск уменьшает время взаимодействия и чаще предлагает готовое решение. В результате пользователь получает не столько список вариантов, сколько интерпретацию собственного запроса, выполненную платформой.

В рамках исследования голосовой поиск рассматривается как форма алгоритмического посредничества, влияющая на:

- структуру пользовательской навигации
- вариативность выбора

- прозрачность поиска
- распределение внимания
- зависимость медиапотребления от платформенных рекомендаций

Объект и предмет исследования

Объектом исследования являются голосовые интерфейсы цифровых медиасервисов.

Предмет исследования - механизмы алгоритмического посредничества и discoverability в голосовом поиске.

Цель исследования заключается в выявлении того, каким образом голосовой поиск трансформирует архитектуру обнаружения контента и влияет на пользовательское взаимодействие с цифровыми медиасервисами.

В рамках исследования предполагается решить несколько задач: проанализировать особенности голосового поиска в различных типах медиасервисов, исследовать влияние голосового поиска на навигацию и выбор контента, сравнить текстовую и голосовую модели поиска, а также выявить архитектурные дисбалансы, возникающие в условиях алгоритмического посредничества.

Гипотеза исследования заключается в том, что голосовой поиск усиливает централизованный характер обнаружения контента, сокращает видимость альтернатив и повышает зависимость пользователя от платформенной логики рекомендаций.

Методология исследования

Исследование выполнено в рамках сравнительно-явленческого маршрута, рекомендованного для анализа цифровых механизмов и эффектов. Основу работы составляет сочетание сценарного, интерфейсного и сравнительного анализа.

Сценарный анализ позволяет исследовать реальные пользовательские маршруты взаимодействия с медиасервисами. В центре внимания находятся типовые ситуации использования голосового поиска: поиск музыки во время тренировки, запуск фильма через smart TV, взаимодействие с голосовым ассистентом в свободном режиме. Анализируются количество действий, скорость доступа к контенту и степень зависимости пользователя от алгоритмической выдачи.

Интерфейсный анализ используется для исследования структуры голосового взаимодействия. Рассматриваются механизмы активации голосового поиска, способы подтверждения запросов, структура выдачи результатов, а также особенности исправления ошибок распознавания речи. Отдельное внимание уделяется тому, как интерфейс скрывает или, напротив, демонстрирует логику рекомендаций.

Эмпирическая база и протокол исследования

Эмпирическая часть исследования была построена на сравнительном анализе голосового поиска в различных типах цифровых медиасервисов. В качестве объектов наблюдения были выбраны сервисы из таблицы:

Типы исследуемых сервисов

Тип сервиса	Основная функция голосового поиска	Особенность архитектуры
Музыкальные	Быстрый запуск музыки	Эмоционально-

платформы (Spotify, YouTube Music)		контекстный поиск
Видеосервисы (YouTube, Netflix)	Поиск фильмов и сериалов	Алгоритмическая рекомендация
Smart TV	Навигация по устройству	Минимизация ручного ввода
Голосовые ассистенты (Алиса, Siri)	Универсальное посредничество	Централизация управления

Выбор данных платформ обусловлен их различной логикой организации discoverability и разным уровнем алгоритмического посредничества.

Наблюдение проводилось в период с марта по май 2026 года. Дополнительно фиксировались различия между авторизованными аккаунтами со сформированной историей медиапотребления и новыми аккаунтами без выраженной пользовательской персонализации.

Анализ проводился на мобильных устройствах iPhone 14 (iOS 18), Samsung Galaxy S24 (Android 15), smart TV LG webOS и Android TV. Все сценарии тестировались в регионе Российской Федерации с использованием русскоязычных голосовых запросов. Каждый голосовой запрос повторялся не менее трех раз для проверки устойчивости результатов. Для проверки использовались как авторизованные аккаунты с историей, так и новые аккаунты без пользовательской истории.

В ходе исследования фиксировалось:

- тип голосового запроса
- число результатов в выдаче
- наличие визуально доступных альтернатив

- возможность ручной корректировки выбора
- характер рекомендательной логики платформы
- наличие автоматического запуска контента без подтверждения пользователя

Для анализа использовались типовые запросы пользователей:

- включи музыку для сна
- что посмотреть вечером
- включи популярную музыку
- найди спортивные трансляции
- включи спокойный джаз
- продолжи подкаст

Эмпирические наблюдения голосового поиска

Платформа	Тестовый запрос	Устройство	Характер голосовой выдачи	Особенности
Spotify	Включи музыку для сна	Samsung Galaxy S24	Ограниченная рекомендательная подборка	Автоматический запуск
YouTube Music	Спокойный джаз	iPhone 14	Расширенная смешанная	Отображение нескольких

			выдача	альтернативных подборок
Netflix	Что посмотреть вечером	Android TV	Краткая персонализированная подборка	Ограниченная видимость альтернатив
YouTube	Найди спортивные трансляции»	LG webOS	Расширенная поисковая выдача	Высокая вариативность результатов
Siri	Включи популярную музыку	iPhone 14	Один приоритетный результат	Автоматическое определение сервиса воспроизведения
Алиса	Продокжи подкаст	Samsung Galaxy S24	Автоматизированный сценарий воспроизведения	Минимальная ручная корректировка

В ходе наблюдений было установлено, что результаты голосового поиска в большинстве сервисов демонстрировали частичную повторяемость, однако вариативность зависела от типа платформы и персонализации аккаунта. Наиболее стабильные результаты были зафиксированы в Siri и Алисе, где пользователь чаще всего получал один приоритетный результат. В YouTube и YouTube Music наблюдалась более вариативная выдача с более большим набором альтернативных результатов.

Схема работы голосового поиска

Пользовательский запрос → Распознавание речи → интерпретация намерения → алгоритмическое ранжирование → ограниченная выдача → Автоматическое воспроизведение контента

Сценарный анализ голосового поиска

Spotify: контекстный поиск и автоматическая рекомендация

В музыкальных сервисах голосовой поиск преимущественно используется для быстрого запуска контента в фоновом режиме. Характерным примером является запрос «включи музыку для сна». В Spotify подобный запрос не приводит пользователя к расширенной поисковой выдаче. Вместо этого сервис автоматически запускает рекомендательный плейлист, сформированный на основе пользовательского профиля, популярности треков и внутренних алгоритмов персонализации.

В текстовом поиске пользователь получает значительно более широкую выдачу: плейлисты, альбомы, исполнителей и подборки. Голосовой интерфейс, напротив, сокращает пространство выбора и ориентирован на максимально быстрое воспроизведение контента.

Таким образом, голосовой поиск в Spotify выполняет не только поисковую, но и рекомендательную функцию. Платформа фактически берет на себя решение о том, какой контент должен быть предложен пользователю в конкретной ситуации.

Netflix: ограниченная discoverability

В Netflix голосовой поиск особенно заметно влияет на структуру discoverability. Запрос «что посмотреть вечером» не предполагает поиск конкретного фильма или жанра. Алгоритм интерпретирует пользовательское намерение и формирует ограниченную подборку персонализированных рекомендаций.

При использовании текстового поиска пользователь способен просматривать расширенный каталог, переходить между категориями и сравнивать результаты. Голосовой поиск существенно сокращает этот маршрут. Пользователь получает несколько наиболее релевантных вариантов, тогда как значительная часть каталога остается невидимой.

Подобная модель делает discoverability более централизованной. Видимость контента определяется не исследовательской активностью пользователя, а системой алгоритмического ранжирования.

Smart TV: сокращение навигационного маршрута

В smart TV-системах голосовой поиск прежде всего компенсирует неудобство ручного ввода через пульт управления. Голосовые команды позволяют запускать приложения, искать фильмы и управлять воспроизведением без необходимости проходить через несколько интерфейсных уровней.

Например, команда «открой YouTube» в LG webOS приводит к мгновенному запуску приложения. Аналогичным образом запрос «найди спортивные трансляции» запускает готовую подборку контента без полноценного просмотра категорий и разделов.

С одной стороны, подобная модель действительно упрощает взаимодействие с платформой. С другой стороны, сокращение пользовательского маршрута одновременно уменьшает исследовательскую

навигацию. Пользователь все реже взаимодействует с медиасредой напрямую и все чаще получает заранее структурированный результат.

Алиса и Siri: максимальное посредничество

Наиболее высокий уровень алгоритмического посредничества наблюдается в голосовых ассистентах. Siri и Алиса функционируют не как отдельные медиасервисы, а как централизованные посредники между пользователем и цифровой экосистемой.

Запрос «включи популярную музыку» приводит не к отображению списка сервисов или вариантов, а к автоматическому запуску платформы, которую ассистент считает наиболее релевантной. Пользователь практически не участвует в выборе маршрута.

В подобных системах discoverability становится максимально централизованной. Алгоритм одновременно определяет:

- какой сервис будет использован
- какой контент окажется видимым
- какой результат будет воспроизведен первым

Сравнительная матрица голосового поиска

Сервис	Тип поиска	Число результатов	Видимость альтернатив	Возможность ручной коррекции
Spotify	Контекстный	3 - 5	Частичная	Высокая
YouTube Music	Контекстный	5 - 10	Средняя	Высокая
Smart TV	Навигационный	1 - 3	Низкая	Средняя
Алиса	Ассистивный	1	минимальная	Низкая

Siri	Ассистивный	1	минимальная	Низкая
YouTube	Гибридный	10+	Высокая	Высокая
Netflix	Рекомендательный	1 - 3	Низкая	Ограниченная

Матрица показывает, что уровень алгоритмического посредничества возрастает по мере перехода от платформ с расширенной выдачей к голосовым ассистентам. Наиболее ограниченная discoverability характерна именно для ассистивных систем, где пользователь получает один готовый результат вместо полноценного набора альтернатив.

Четырехконтурная диагностика голосового поиска

1. Институционально-регулятивный контур
2. Пространственно-сценарный контур
3. Семиотико-интерфейсный контур
4. Цифрово-данный контур

1. Институционально-регулятивный контур

На институциональном уровне голосовой поиск усиливает зависимость пользователя от платформенной экосистемы. Голосовые ассистенты и smart TV-системы способны приоритетно продвигать собственные сервисы или партнерский контент. В результате платформа получает дополнительный контроль над распределением внимания.

Особенно заметно это в ассистивных системах, где пользователь зачастую не видит альтернативных маршрутов. Голосовой интерфейс скрывает сам механизм выбора платформы и делает посредничество практически незаметным.

2. Пространственно-сценарный контур

Голосовой поиск радикально сокращает пользовательский маршрут. В традиционной модели поиска пользователь последовательно взаимодействует с несколькими интерфейсными уровнями: открывает приложение, вводит запрос, просматривает выдачу и выбирает результат.

Голосовой поиск минимизирует этот процесс до одной команды. Подобное сокращение действительно ускоряет доступ к контенту, однако одновременно уменьшает исследовательское взаимодействие с медиасредой.

3. Семиотико-интерфейсный контур

Голосовые интерфейсы меняют сам характер коммуникации между пользователем и платформой. Запросы становятся разговорными, эмоциональными и контекстными. Пользователь обращается к системе не как к поисковому механизму, а как к посреднику, способному «понять» намерение. Подобная модель усиливает антропоморфизацию платформы и создает эффект естественного диалога. При этом сама логика алгоритмической фильтрации остается скрытой.

4. Цифрово-данный контур

На цифрово-данном уровне голосовой поиск усиливает значение алгоритмов распознавания речи, анализа намерений и персонализации. При этом пользователь практически не имеет доступа к механизмам ранжирования.

Главной особенностью голосового поиска становится ограниченная выдача. Во многих случаях пользователь получает один результат или небольшую подборку рекомендаций. Это делает discoverability более централизованной и усиливает роль алгоритмического посредничества.

Архитектурные эффекты голосового поиска

Проведенное тестирование позволяет точнее интерпретировать архитектурные эффекты голосового поиска. В ходе наблюдений было

установлено, что voice search действительно сокращает пространство пользовательского выбора, однако степень этого сокращения различается в зависимости от типа платформы.

Наиболее высокая вариативность выдачи сохраняется в YouTube и YouTube Music, где голосовой поиск все еще интегрирован в визуальную интерфейсную среду. Пользователь может быстро перейти к ручной корректировке результатов и самостоятельно исследовать альтернативный контент.

В Netflix, Siri и Алисе ситуация иная. Здесь голосовой поиск преимущественно ориентирован на быстрое предоставление одного или нескольких наиболее релевантных результатов. Подобная логика существенно снижает исследовательскую навигацию и усиливает зависимость пользователя от алгоритмов платформы.

Тестирование также показало, что голосовой поиск способствует росту пассивного медиапотребления. Во многих сценариях контент запускался автоматически сразу после голосового запроса, без необходимости дополнительного подтверждения со стороны пользователя. Особенно заметно это проявлялось в музыкальных сервисах и голосовых ассистентах.

Таким образом, речь идет не просто об изменении интерфейсного механизма поиска, а о трансформации самой архитектуры обнаружения контента, при которой алгоритм начинает играть роль центрального посредника между пользователем и медиасредой.

Рекомендации и сценарные выводы

Проведенное исследование показывает необходимость повышения прозрачности голосового поиска в цифровых медиасервисах. В современных

платформах голосовой поиск всё больше работает не только как интерфейсный механизм, но и как система алгоритмического распределения внимания.

Одним из возможных направлений развития является расширение видимости альтернативных результатов внутри voice-интерфейсов. Важным также представляется предоставление пользователю возможности быстрого перехода от автоматизированной voice-выдачи к расширенной ручной навигации.

Для медиаплатформ перспективными направлениями могут стать:

- маркировка рекомендательного характера voice-выдачи
- отображение нескольких альтернативных сценариев выбора
- повышение прозрачности персонализации
- снижение степени скрытого алгоритмического посредничества

Для исследовательской практики перспективным направлением выступает дальнейшее изучение voice interfaces как инфраструктуры платформенного управления discoverability и пользовательским вниманием.

Заключение

Голосовой поиск постепенно становится важной частью пользовательской навигации и обнаружения медиаконтента внутри цифровых платформ. В отличие от текстового поиска, голосовые-интерфейсы ориентированы на сокращение пользовательского маршрута и автоматическую интерпретацию запросов.

В наблюдаемых сценариях это приводило к уменьшению количества визуально доступных альтернатив и усилению алгоритмического посредничества. Наиболее высокий уровень подобного посредничества был зафиксирован в голосовых ассистентах Siri и Алисе, где пользователь чаще всего получал заранее определенный сценарий воспроизведения. В музыкальных и видеосервисах степень алгоритмического контроля была менее выраженной, однако голосовой поиск и здесь постепенно совмещает функции поиска, рекомендации и навигации.

Полученные результаты позволяют рассматривать голосовой поиск как значимый элемент современной платформенной медиасреды, влияющий на пользовательский выбор, видимость контента и способы взаимодействия с цифровыми сервисами.

Список литературы

1. Калмыков Н. Н. Социальные архитектуры в социологии управления: четырехслойная модель и диагностический протокол анализа управляемости // Социология. — 2026. — №3. — С. 121–126.
2. Калмыков Н. Н. Социальные архитектуры как проектируемые многомерные среды: теоретическое ядро и управленческая операционализация // Социология. — 2026. — №2.
3. Калмыков Н. Н. Парадокс цифровой прозрачности в кадровых социальных архитектурах: семантическая автономия, контрконтроль и суверенитет цифрового слоя // Russian Economic Bulletin. — 2026. — №1.
4. Gillespie T. Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media. — Yale University Press, 2018.
5. Porcheron M., Fischer J., Reeves S. Voice Interfaces in Everyday Life // Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. — 2018.
6. Luger E., Sellen A. “Like Having a Really Bad PA”: The Gulf between User Expectation and Experience of Conversational Agents // Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. — 2016.
7. Napoli P. Social Media and the Public Interest: Media Regulation in the Disinformation Age. — New York: Columbia University Press, 2019.
8. Sundar S. S. Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction // Journal of Computer-Mediated Communication. — 2020. — Vol. 25. — № 1. — P. 74–88.

9. Spotify Support. Voice Search and Voice Control. URL: <https://support.spotify.com/> (дата обращения: 06.06.2026).
10. YouTube Help Center. Voice Search Functions. URL: <https://support.google.com/youtube/> (дата обращения: 06.06.2026).
11. Apple Support. Siri User Guide. URL: <https://support.apple.com/siri> (дата обращения: 06.06.2026).
12. Яндекс Справка. Алиса и голосовое управление сервисами. URL: <https://yandex.ru/support/alice/> (дата обращения: 06.06.2026).
13. Netflix TechBlog. Recommendations, Personalization, and Search. URL: <https://netflixtechblog.com> (дата обращения: 06.06.2026).
14. Android Developers. Voice Interaction API for Android TV. URL: <https://developer.android.com/reference/android/service/voice/VoiceInteractionSession> search (дата обращения: 06.06.2026).