

Хабр



КАК СТАТЬ АВТОРОМ



Зарплаты IT-специалистов



Войти



empenoso

29 июл в 05:26

Экологичное расставание с Алисой: строим полностью локальный и приватный голосовой ассистент

Простой

8 мин

62K

Умный дом, Гаджеты, Настройка Linux*, Open source*

Мнение



Космотекст

Идея отказаться от использования Яндекс Алисы в системе умного дома возникла у меня после новости о принятии Госдумой законопроекта, [касающегося штрафов за поиск и доступ к экстремистским материалам в интернете](#). Казалось бы, при чём тут голосовой помощник? Однако Яндекс входит в реестр организаторов распространения информации, что означает определённые юридические и технические обязательства по хранению и передаче данных.

Хотя я не ищу ничего, выходящего за рамки интересов автоматизации, желание иметь полностью автономный, локально работающий умный дом — без зависимости от интернета и облачных сервисов — стало для меня ещё актуальнее.

Тем более что сейчас единственным слабым звеном в моём умном доме остается Яндекс Алиса — которая требует постоянного интернет-соединения даже для выполнения простейших команд управления локальными устройствами.

В этой статье я расскажу, как и на что планирую заменить Алису, чтобы сохранить привычный голосовой контроль, но без сторонних подключений и рисков для приватности.

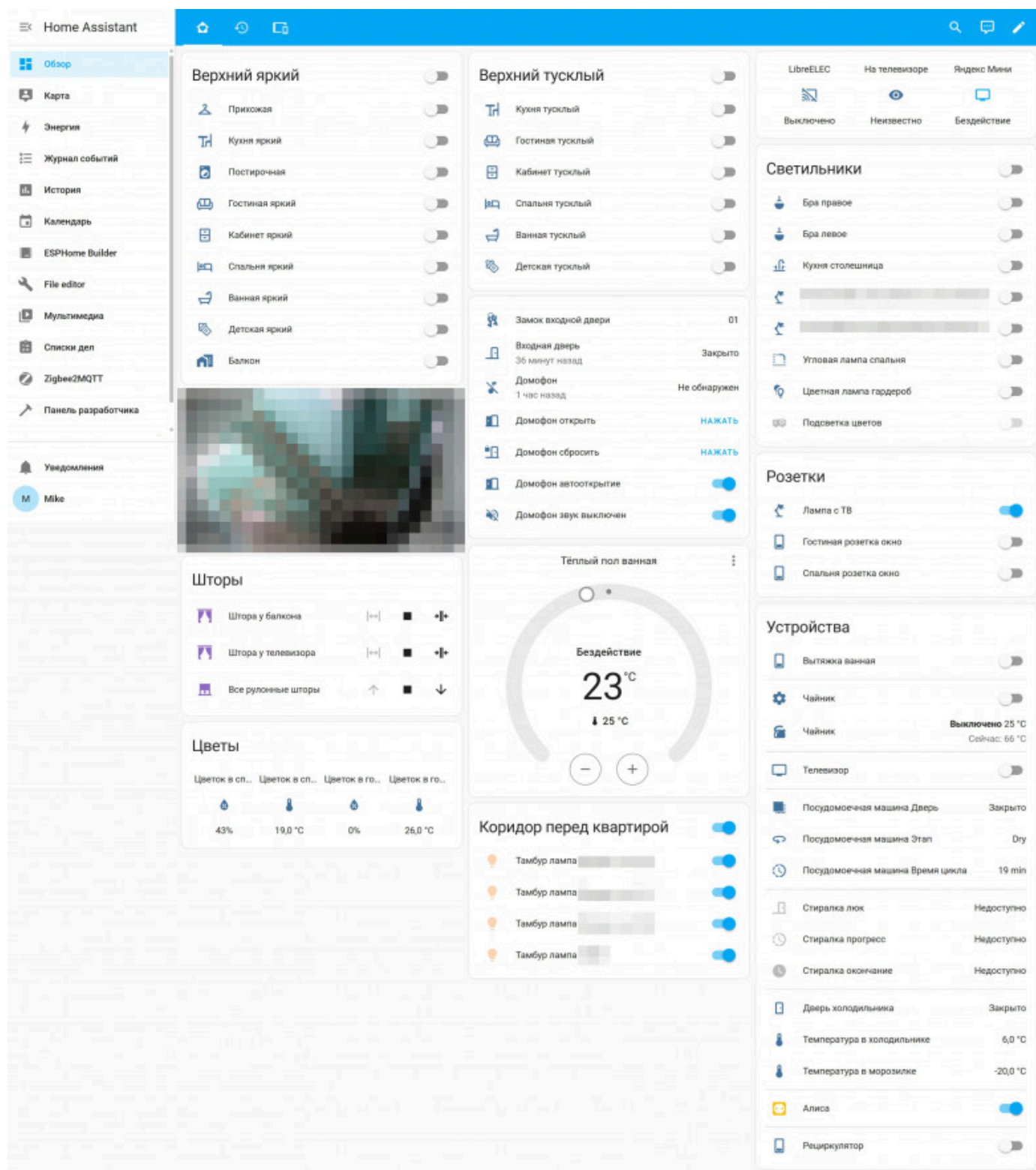
Конфигурация моего умного дома: чем будем управлять

РЕКЛАМА



IT-пинбол

Играй и выигрывай AirPods Max!



Мой Home Assistant в «человеко читаемом» виде

Мой умный дом строился с прицелом на автономность, надежность и открытые стандарты — так, чтобы управление работало даже при полном отсутствии интернета. На данный момент архитектура системы выглядит следующим образом.

**IT-пинбол**

Играй и выигрывай AirPods Max!

Мозг системы: центральный контроллер — это Raspberry Pi 4 Model B с 2 ГБ оперативной памяти, установлен в 2022 году. На него установлена Home Assistant OS — полноценная операционная система, заточенная под локальное управление умным домом — подробнее описывал [в другой статье](#). Вся логика автоматизаций, интерфейс управления и интеграции работают исключительно локально, без необходимости в сторонних облаках.

Извиняюсь за скриншот, но с прокруткой только PicPick под Windows умеет делать - и вот результат :(



IT-пинбол

Играй и выигрывай AirPods Max!

В светильниках.

Ключевую нагрузку по управлению берет на себя Zigbee-сеть: 42 устройства, объединённые с помощью USB-донгла Sonoff Zigbee 3.0 Plus и интеграции Zigbee2MQTT. Это датчики, реле освещения и другие элементы.

Что управляется:

- Освещение: в каждой комнате — два контура: тусклый (вечерний) и яркий, плюс светодиодная лента в спальне, [освещение общего коридора с двумя режимами](#).
- Климат: кондиционеры, обогрев ванной комнаты через реле теплого пола.
- Электропитание и бытовая техника: управляемая розетка для ТВ, [стиралка](#), холодильник, посудомойка, чайник.
- Датчики: движения, открытия, температуры и влажности.
- Шторы: [моторизованные рулонные](#) и классические.
- Мультимедиа: [управление Kodi на медиаплеере](#) и доступ к медиатеке NAS Synology, [панель управления умным домом](#).
- Безопасность: [камера видеонаблюдения из подъездного домофона](#), IP-камера у лифтов, управление домофоном в многоквартирном доме — автовахтер по моим правилам.

Все эти устройства уже управляются локально, без облачных зависимостей — кроме стиралки Bosch, купленной ещё в 2022 году.

Теоретический минимум: из чего состоит локальный голосовой помощник

Однако чтобы убрать колонку Яндексa и заменить Алису на полностью автономного голосового помощника, нужно понять, из каких компонентов он состоит. Это не “одна программа”, а целая цепочка взаимодействующих модулей, каждый из которых выполняет свою задачу:



IT-пинбол

Играй и выигрывай AirPods Max!

ESP32-S3-BOX-3. Фото из интернета

Микрофон и динамик («Уши и рот» системы) — это устройства, которые слышат пользователя. Не должно быть колхоза из датчиков. Устройство должно выглядеть современно и не портить интерьер.

В моем случае я присматриваюсь к двум: компактный M5Stack ATOM Echo для комнат и более продвинутый ESP32-S3-BOX для гостиной.



IT-пинбол

Играй и выигрывай AirPods Max!

Официальный комплект для разработки умных динамиков ATOM Echo M5Stack

Они захватывают звук и отправляют его на сервер для дальнейшей обработки.

100% новый ESP32-S3-BOX-3 ESP32-S3-BOX-3B модуль комплекта разработки приложений AIOT 2,4 ГГц Wi-Fi + Bluetooth 5

Wake Word движок: нужен, чтобы система слушала нас постоянно, но реагировала только по ключевой фразе (например, «Привет, пирожок!»). Используем OpenWakeWord — полностью локальный и настраиваемый.

Speech-to-Text (STT): этот модуль превращает речь в текст. Здесь смотрю на Whisper от OpenAI — пишут что это один из самых точных и устойчивых к шуму движков, работающий прямо на локальном сервере. Про его выбор чуть ниже.



IT-пинбол

Играй и выигрывай AirPods Max!

Символ команды. Эта задача ложится на встроенный в Home Assistant механизм Assist,

который сопоставляет текст с действиями и сущностями в системе.

Text-to-Speech (TTS): чтобы система могла отвечать голосом, нужен синтез речи. Я планирую использовать Piper — современный, быстрый, качественный, легко интегрируется как Add-on в HA. Как вариант RHVoice — тоже отличный вариант, но Piper сейчас является де-факто стандартом в сообществе HA за простоту и качество.

Wyoming Protocol: связующее звено. Простой, но мощный протокол, через который все эти модули общаются между собой и с Home Assistant.

Речь в текст: почему именно такой стек?

Давайте будем честны: моя Raspberry Pi 4 с 2 ГБ памяти — отличный мозг для автоматизации, но для тяжелых вычислений, таких как распознавание речи в реальном времени, её мощности не хватит.

Поэтому, помимо «ушей» в виде ESP32-S3-BOX и M5Stack ATOM Echo, в систему придется докупить отдельный мини-ПК. Это может быть недорогой китайский NUC-подобный компьютер, который возьмет на себя самую ресурсоемкую задачу — преобразование речи в текст (Speech-to-Text (STT)).

Или может быть [Raspberry Pi 5 с 16 ГБ оперативной памяти](#) — цены сопоставимы.



IT-пинбол

Играй и выигрывай AirPods Max!

Speech-to-Phrase (от Open Home Foundation): это самый легковесный вариант. Он не распознает речь, а просто ищет точное совпадение с заранее заданными фразами. К тому же это не конкретный движок, а концепция pipeline в НА. По умолчанию он использует тот же Whisper, но его самую легкую модель, чтобы хоть как-то работать на слабых устройствах вроде RPi. Плюс: минимальные требования к железу. Минус: абсолютная негибкость. Система поймет «включи свет на кухне», но проигнорирует «сделай на кухне посветлее». Это не интеллект, а поиск по словарю.

Rhasspy: ветеран мира локальных ассистентов. Мощный, но сложный в настройке комбайн. Главный аргумент против него сегодня: проект развивается медленнее, чем экосистема Home Assistant. Пока Rhasspy остается монолитной системой, связка Assist + Wyoming-протокол ушла далеко вперед в плане гибкости и интеграции.

Whisper от OpenAI — современный стандарт транскрипции. Понимает естественную речь в свободной форме, работает с русским языком. Различные модели (tiny, base, small, medium) позволяют балансировать между скоростью и качеством. Активно развивается, поддерживается сообществом НА, появляются оптимизированные версии вроде distil-whisper. Это выбор на перспективу.

Как избавиться от голосового помощника Алисы

Поскольку я нахожусь в активном поиске оптимального решения и уже покупаю компоненты, то буду признателен за ваши комментарии, критику и предложения.

Вариант 1: простой и дешевый

Лично для себя я не рассматриваю этот вариант, однако этот путь подойдет тем, кто хочет попробовать локальное голосовое управление с минимальными затратами времени и денег. Как раз, чтобы «пощупать» концепцию и понять, насколько она жизнеспособна.



IT-пинбол

Играй и выигрывай AirPods Max!

M5Stack ATOM Echo. Микроразмер. Фото из интернета

Или если вы только планируете сделать умный дом — можно изначально заложить более мощное железо — чтобы всё было на одном севере.

Все компоненты — Home Assistant, распознавание речи (STT) и синтез голоса (TTS) — работают прямо на Raspberry Pi. Один микрофон, одна точка входа, минимум зависимостей.

То есть:

```
[M5Stack ATOM Echo] ← Wi-Fi → [Raspberry Pi 4 (HA + STT + TTS)]
```

Если брать мой случай:

- Уже есть: Raspberry Pi 4 (2 ГБ) с установленной Home Assistant OS.
- **Нужно купить: M5Stack ATOM Echo (примерно 1 400 рублей).** Это крошечное



IT-пинбол

Играй и выигрывай AirPods Max!

Настройка:

1. **Прошивка ATOM Echo:** через ESPHome. Готовый YAML-конфиг для голосового ассистента легко [найти в официальных примерах](#).

2. Pipeline в HA:

- **STT:** Используем [Assist pipeline от Open Home Foundation](#) с движком faster-whisper и моделью tiny. Запустится скорее всего даже на Pi 4.
- **TTS:** Устанавливаем [Add-on Piper](#) — быстрый и качественный синтезатор, особенно с голосами на русском.

Плюсы этого решения:

- Минимальные вложения — только 1 400 рублей и немного времени.
- Простота — всё работает на одном устройстве.
- Быстрый старт — можно реализовать за один вечер.

Минусы:

- Скорее всего заметная задержка из-за слабого железа.
- Нагрузка на Home Assistant — может тормозить работу системы во время STT.
- Плохо масштабируется: один микрофон — ещё приемлемо, но два и больше будут проблемой.

Вариант 2: «правильная» архитектура с заделом на будущее

Это мой приоритетный путь — вынести ресурсоёмкие задачи обработки речи на отдельный сервер, а Raspberry Pi остаётся заниматься только управлением умным домом. Подход масштабируемый, стабильный и в моём случае надеюсь что будет в разы быстрее.



IT-пинбол

Играй и выигрывай AirPods Max!

ESP32-S3-BOX. Фото из интернета

Схема сложнее:

```
[Пользователь]
  ↓ говорит
[ESP32-S3-BOX / M5Stack ATOM Echo] ← микрофон + wake word ("Привет, пирожок!")
  ↓ захватывает аудио
  (по Wi-Fi)
  ↓
[Мини-ПК: Whisper STT-сервер]
  ↓ распознаёт речь в текст (Whisper STT)
  ↓
[Home Assistant на Raspberry Pi 4]
  ↓ определяет намерение (Assist)
```



IT-пинбол

Играй и выигрывай AirPods Max!

```

    ↓ синтезирует голосовой ответ
    (по Wi-Fi)
    ↓
    [ESP32-S3-BOX / M5Stack ATOM Echo] ← динамик
    ↓ озвучивает ответ
    [Пользователь]

```

Железо:

- Уже есть Raspberry Pi 4 (2 ГБ) — Home Assistant, Zigbee, автоматизации.
- Примерно 14 т.р.: Mini PC (Intel N100 или N95) — сервер обработки голоса.
- Примерно 6 т.р. ESP32-S3-BOX — «умный» ассистент для гостиной.
- Примерно 1,4 т.р. M5Stack ATOM Echo — недорогие ассистенты для других комнат.

Сервер обработки голоса (Mini PC):

Устанавливаем легкий Linux (Debian/Ubuntu Server), затем — Docker и Docker Compose. В docker-compose.yml разворачиваем сразу три контейнера:

- Whisper — для распознавания речи (STT).
- Piper — синтез речи (TTS).
- OpenWakeWord — «ключевая фраза» для активации.

С мощностями N100 можно использовать модель Whisper уровня small или даже medium, получая более точное и быстрое распознавание речи, чем на Pi.

Настройка Home Assistant: на Raspberry Pi в этом случае не используется голосовых add-on'ов — только интеграция через Wyoming:

- Заходим в *Настройки* → *Устройства и службы* → *Добавить интеграцию*.
- Добавляем *Wyoming Protocol* трижды — для каждого из сервисов (Whisper, Piper, WakeWord), указав IP и порты Mini PC.



IT-пинбол

Играй и выигрывай AirPods Max!

Спутники (ESP32-S3-BOX и ATOM Echo): прошиваются через ESPHome. У ESP32-S3-BOX можно задействовать экран: отображать статус («Слушаю», «Думаю», «Выполняю»), добавляя интерактивности.

Плюсы:

- Ожидаемая быстрая реакция.
- Ожидание распознавания сложных фраз.
- Не грузит Home Assistant.
- Масштабируемость: добавляем спутники - и всё.

Минусы:

- Дороже (нужен Mini PC).
- Потребуется базовые навыки Linux и Docker.

Вариант 3: дорого и сложно

Можно полностью избавиться от Raspberry Pi 4 с 2 ГБ памяти и абсолютно всё перевести на новый мощный сервер. RAM видимо выбрать 16-32 ГБ чтобы с запасом на все. Может быть даже купить NVIDIA VRAM 6 ГБ, но это тогда сильно увеличит стоимость и можно будет забыть о безвентиляторности.



IT-пинбол

Играй и выигрывай AirPods Max!

Сборка в mini-ITX. Фото из интернета

Можно тоже будет использовать Home Assistant OS или Linux (Ubuntu/Debian) + Docker.

Правда это большая работа — много устройств. Пока склоняюсь к второму варианту.

Заключение: свобода выбора

Переход на локального голосового ассистента — это не просто технический эксперимент, а осознанный шаг к созданию по-настоящему приватного и независимого умного дома.

Первый вариант — это отличная, почти бесплатная возможность «пощупать» технологию и понять ее ограничения. Второй — полноценное решение, которое по скорости и качеству скорее всего не уступит Алисе, при этом полностью оставаясь под контролем. Третий вариант — если есть бюджет.

Все пути ведут к одной цели — избавлению от «облачного рабства». До сентября ещё есть время. А расставание с Алисой может быть не только экологичным, но и очень увлекательным!

А каким голосовым помощником пользуетесь вы?

Автор: Михаил Шардин

 [Моя онлайн-визитка](#)

 [Telegram «Умный Дом Инвестора»](#)

29 июля 2025 года

Теги: [космотекст](#), [алиса](#), [яндекс](#), [whisper](#), [rhasspy](#)

Хабы: [Умный дом](#), [Гаджеты](#), [Настройка Linux](#), [Open source](#)

Редакторский дайджест

Присылаем лучшие статьи раз в месяц



IT-пинбол

Играй и выигрывай AirPods Max!

**211**

Карма

191.3

Рейтинг

Михаил Шардин @empenoso

Автоматизация / Данные / Финансы / Умные дома

[Подписаться](#)[Сайт](#) [Сайт](#) [GitHub](#) **Комментарии 113****IT-пинбол**

Играй и выигрывай AirPods Max!

ЛУЧШИЕ ЗА СУТКИ

ПОХОЖИЕ



alexeybashuk

16 часов назад

Они «запатентовали» на себя обычное слово, а потом потребовали с нас пять миллионов рублей. Вот что было дальше



Простой



6 мин



7.4K

Кейс

+70

30

29



beget_com

20 часов назад

«Инженеры, делайте хуже!» — как в 1924 году родилось запланированное устаревание



5 мин



6.4K

Ретроспектива

+46

32

137



RationalAnswer

22 часа назад

Дешифровка переписок в мессенджерах в ЕС, а также публикация частных диалогов с Гроком



5 мин



12K

Дайджест

+35

14

75



alizar

18 часов назад

Lottie — новый стандарт векторной анимации



IT-пинбол

Играй и выигрывай AirPods Max!

 +32 28 9**IlinValery**

20 часов назад

Как мы сделали робота-комплектующего: от идеи до первого прототипа

**Простой**

9 мин



2K

 +32 22 7**omyhosts**

13 часов назад

Что на самом деле делала кнопка «Турбо»



5 мин



6.5K

Ретроспектива

 +31 18 26**alan_dani**

22 часа назад

Я перепробовал все приложения для списков дел и в итоге остался с .txt файлом

**Простой**

4 мин



4.6K

Перевод

 +31 35 35**interpres**

14 часов назад

Самодельная монтировка телескопа на ESP32

**Средний**

9 мин



2.4K

Обзор

Перевод

**IT-пинбол**

Играй и выигрывай AirPods Max!

**romanova_ase**

18 часов назад

Геоданные в PostgreSQL: зачем нужен PostGIS и как он работает



11 мин



2.5K

**+25**

39



3

**Albert_Wesker**

17 часов назад

Создаём собственный системный вызов в Linux 6.8

**Средний**

4 мин



1.9K

Тutorial

Перевод

**+20**

51



3

OSMP, Kubernetes и прод-трафик: как я пилотировала российский NGFW

Турбо

[Показать еще](#)

ИСТОРИИ

**IT-пинбол**

Играй и выигрывай AirPods Max!



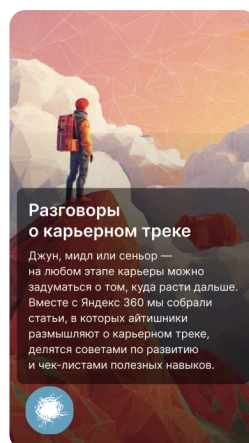
В поисках пятого элемента ИТ



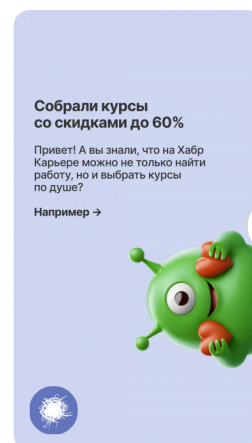
Твой вид из окна в стиле киберпанк



Топ-7 годных статей из блогов компаний



Как расти в ИТ: советы, гайды и опыт сеньоров



Курсы со скидками до 60%

ВОПРОСЫ И ОТВЕТЫ

Как отменить инвертирование текста в yandex почте в dark mode?

HTML · Простой · 0 ответов

Почему Яндекс в один день может массово удалить из индекса страницы нашего сайта, которые он считает низкоценными и мало востребованными?

Яндекс · Простой · 2 ответа

Почему Postfix не может отправить письмо в yandex?

Google · Простой · 2 ответа

Яндекс Браузер на MacOS: как убрать или подвинуть иконку автозаполнения?

CSS · Средний · 0 ответов

Yandex не индексирует изображения на сайте, почему?

Поисковая оптимизация · Средний · 2 ответа

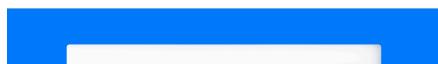
[Больше вопросов на Хабр Q&A](#)

МИНУТОЧКУ ВНИМАНИЯ



ИТ-пинбол

Играй и выигрывай AirPods Max!



Турбо

Лего-мир / аниме / киберпанк
за окном? Почему бы и нет?



Да начнётся битва: выбираем
лучший IT-бренд работодателя

IT-комикс

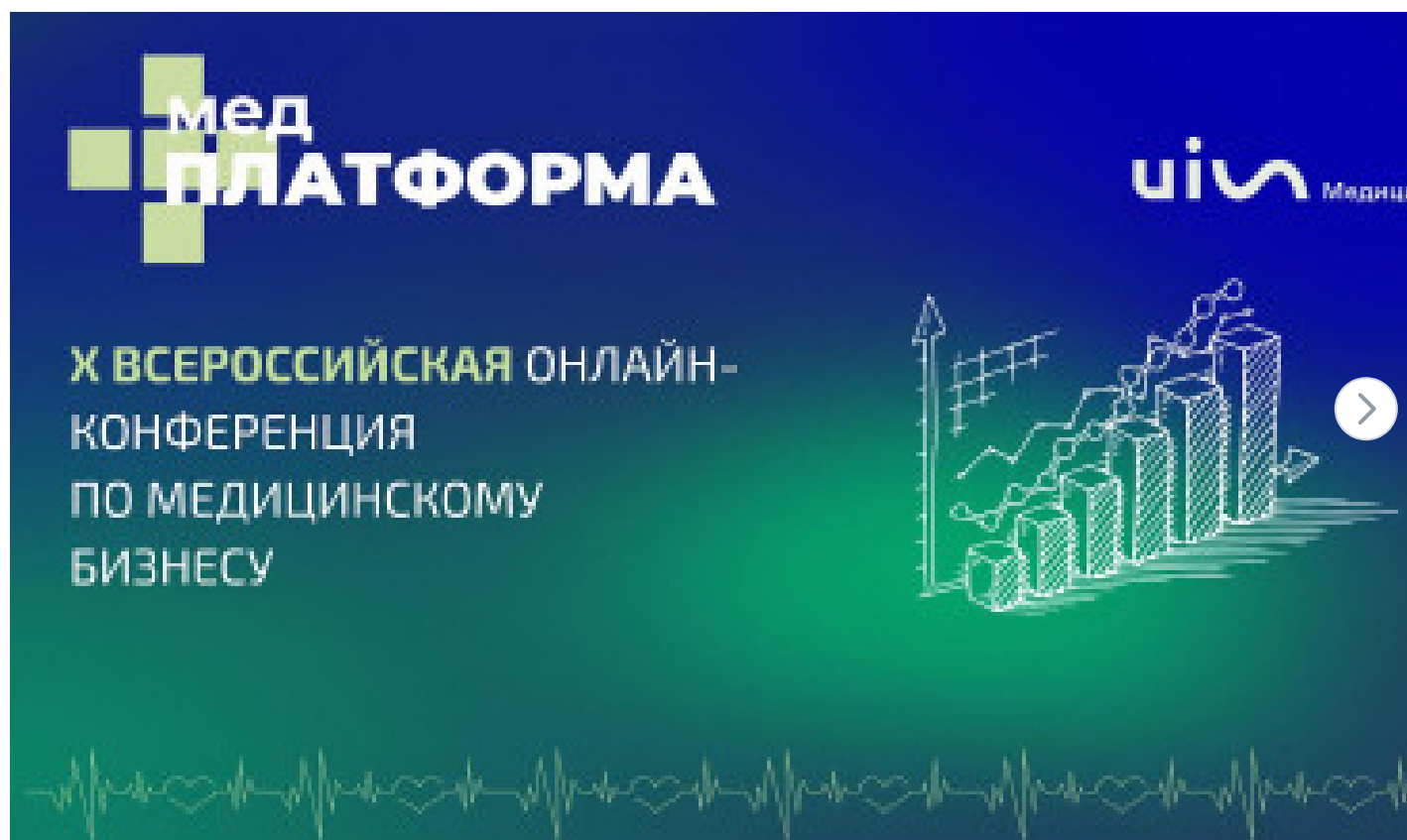
Смогу ли я? Комикс для будущих
айтишников

БЛИЖАЙШИЕ СОБЫТИЯ



IT-пинбол

Играй и выигрывай AirPods Max!



28 – 29 августа

Х Медплатформа: Всероссийская онлайн-конференция по медицинскому бизнесу

Онлайн

Другое

[Больше событий в календаре](#)

Хабр



IT-пинбол

Играй и выигрывай AirPods Max!

Техническая поддержка

© 2006–2025, Habr



IT-пинбол

Играй и выигрывай AirPods Max!