

Искин по имени Алиса

"Умные колонки" уже пару лет как очень популярны. Но не в России. Пока. По итогам прошлого (уже!) 2018 года, мировой рынок smart speaker - больше 20 млн. проданных устройств. Проблема популярности в России - язык. И он был преодолен усилиями Яндекс. За что мы им всячески уважаем.

Тот факт, что "умных колонок" в мире продается довольно много, подтверждается вот такими цифрами:

Worldwide smart speaker market Q3 2018				
Rank	Vendor	Q3 2017	Q3 2018	Growth
#1	amazon	74.7%	31.9%	1.3%
#2	Google	24.6%	29.8%	187%
#3	Alibaba 天猫 TMALL.COM	0.1%	11.1%	>200 times
#4	MI Xiaomi	0.1%	9.7%	>300 times
Others		0.5%	17.4%	>70 times
Overall market		8.3 million	19.7 million	137.0%

Source: Canalys estimates (sell-in volume), Smart Speaker Analysis, November 2018

canalys

[Источник](#) - данные за 3 квартал прошлого (непривычно еще писать про 2018-ый - прошлый) года. Порядка 20 миллионов устройств продано по всему миру. Это не так много, но и немало. Если учесть, что средняя цена "умной колонки" 100 долларов, то весь рынок можно оценить в два миллиарда. Это приличный рынок, который еще и растет на 40% в год.

Видно, что лидеры "умных колонок" США и Китай. США, причем, в разы. И львиную долю занимает [Amazon со своим хитом Alexa](#), про которую стоило бы написать отдельно много букв (хотя на "наге" статья про этот девайс [уже была](#)), но до меня прибор пока не доехал (я бы хотел [EchoSpot](#) заюзать - это уже даже не "умная колонка", а "умный экран". Но со звуком!).

Зато, доехал российский продукт под названием [Яндекс.Станция](#).

[\[Искин по имени Алиса\]](#)

И это, нужно отметить, революционный продукт на российском рынке. Конкуренты появятся нескоро. И даже, скорее всего, не появятся. Потому что уровень хайтека в такой простой на вид железке, зашкаливает.

Не без косяков, конечно. Но обо всем по порядку.

Распаковку "Станции" я показывать не стану - это довольно заурядная процедура и на Ютубе сотни роликов про то, что и как там в коробочку положили. Я пойду дальше и разберу прибор на винтики (забегая вперед, придется писать две статьи, потому что получилось слишком много букв).

Но для начала нужно немного описать "как это работает" и для чего это вообще. Хотя, я и уверен, что читатели техносайта прекрасно понимают. Но описать пользовательский опыт - это довольно интересно. Хотя бы и для меня.

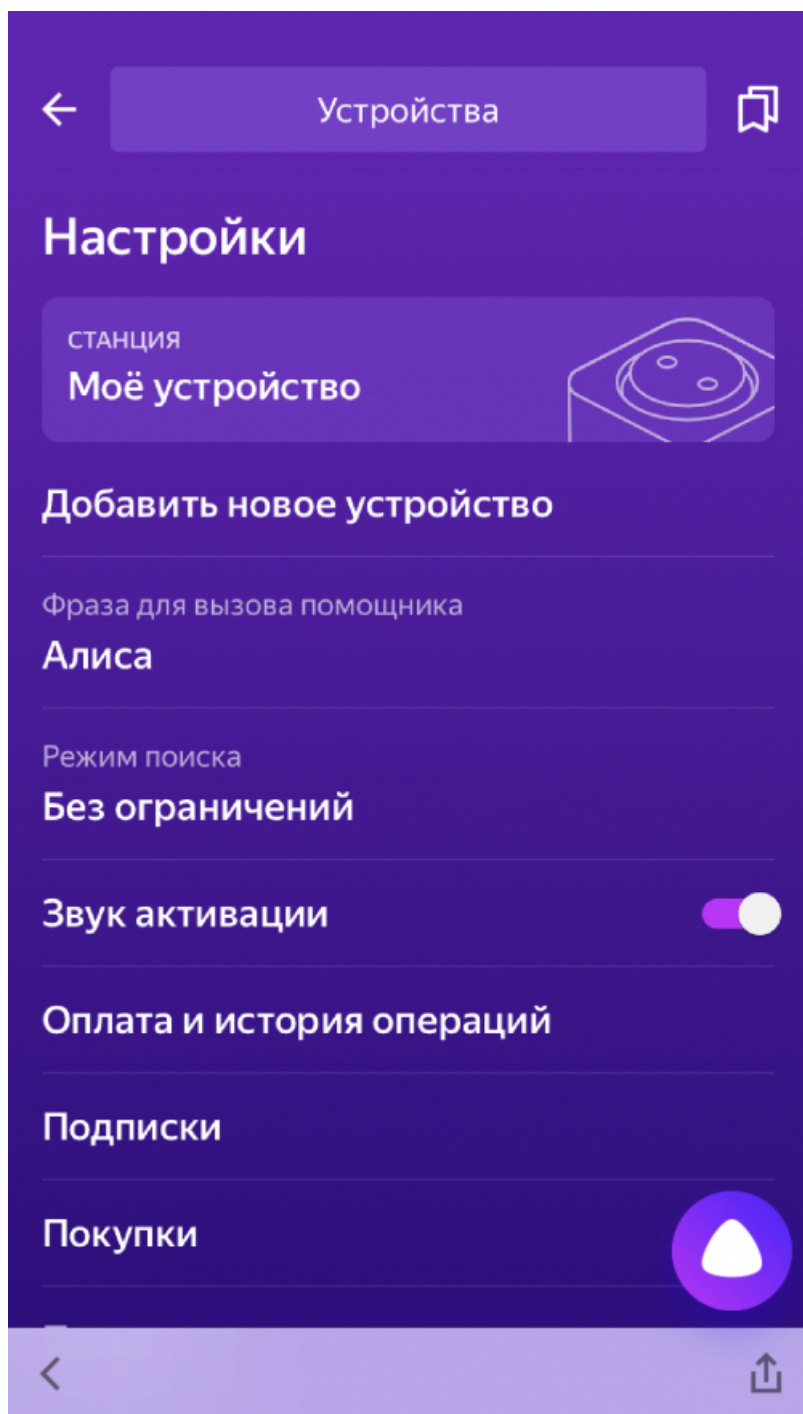
Вот такая машинка:

Главное отличие Яндекс.Станции от Алексы или [Google Home](#) - в Станции есть не только звук. Но и видео. Впрочем, у Amazon есть [fire TV Cube](#) с весьма похожим экспириенсом. Но первичным свойством Станции все же является звук. Телевизор можно не подключать.

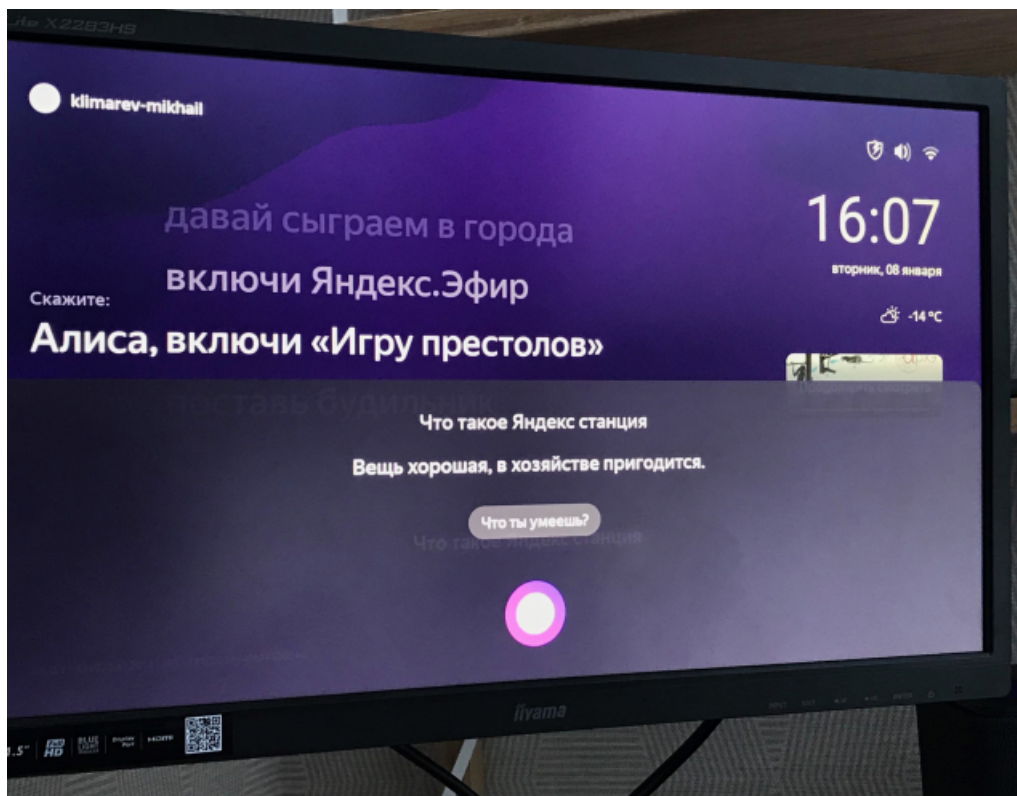
Вот так происходит активация и первый запуск Станции (все равно я б лучше не снял такой ролик - потому от производителя):

Все довольно просто и понятно. Далее вы простым русским языком общаетесь с Алисой. Диалог вопрос-ответ. Для активации достаточно сказать "Алиса," + ваш вопрос. Ну, например, "Алиса, какая погода на завтра". И Алиса ответит.

Вообще, общаться голосом с "роботом" в текущей реализации - опыт непривычный. Вам нужно сначала мысленно сформулировать вопрос и только потом произнести его вслух, не забыв сказать "Алиса" для активации. Кстати, если у вас в доме уже есть Алиса, то в настройках можно изменить активацию на "Яндекс":



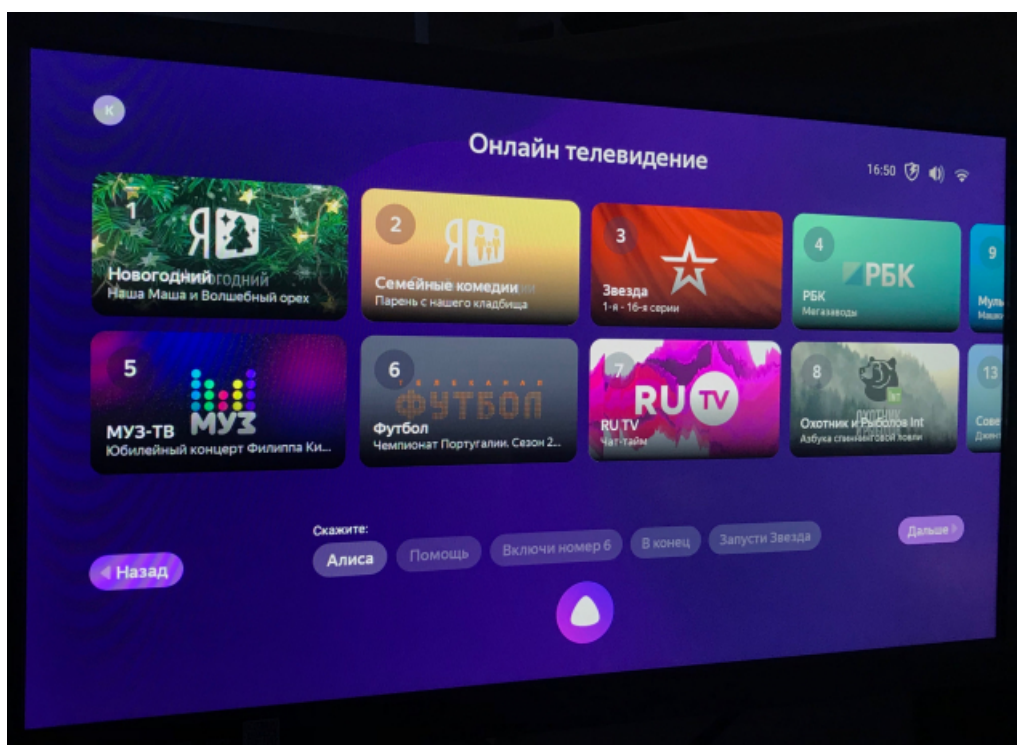
При подключенном экране (у меня просто монитор с HDMI-входом) процесс "осмысления диалога" отображается на экране.



Еще "на главном" элементарно-бытовые цифры - дата, время, погода за бортом в местонахождении пользователя. Еще две "печеньки" справа-снизу - "продолжить смотреть" для видео и "продолжить слушать" для аудио.

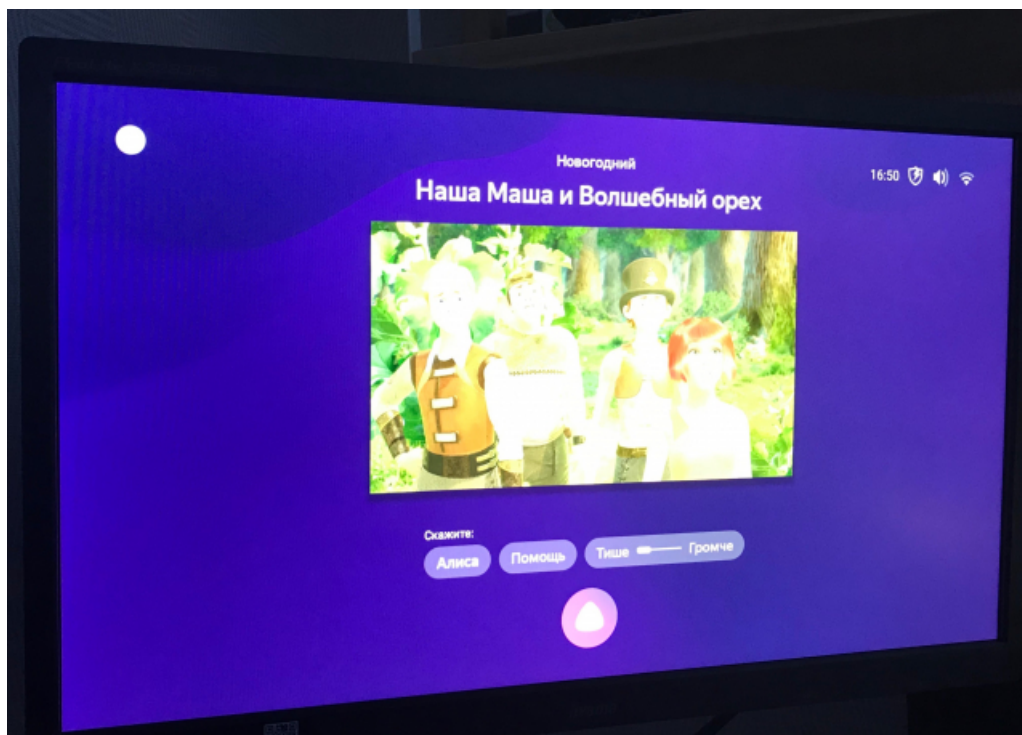
Довольно спартанский интерфейс. Но нужно учитывать, что никаким handled-устройством ни в один элемент интерфейса вы ткнуть не сможете. У Станции нет пульта ДУ, как у других телевизионных приставок. Нет мыши или тач-интерфейса. Только голос.

Недавно появившаяся функция Станции - Яндекс.Эфир:



Да, это онлайн-телевидение - линейное, как железнодорожная рельса. Голосом (Алиса, дальше) можно перемещаться по меню. Голосом же можно называть канал, который вы хотите запустить. Или цифрой, или названием, если помните. По названию, кстати, выходят казусы: например, канал "Юрган" мне не удалось запустить по названию. Алиса интерпретирует "Юрган", как "Курган" и жаловалась на отсутствие карты на устройстве.

Вот так выглядит процесс буферизации канала перед запуском:



А так - собственно, воспроизводимый контент:



... обратите внимание, что SD-картинка не заполнила экран, оставив черные пространства вокруг. Ну, может, это связано с моим монитором конкретно. Но осадочек-таки есть.

Всего в Яндекс.Эфире на Станции мне показало 98 каналов и ориентироваться только голосовыми командами в них непросто. Откровенно говоря, в таком количестве и в браузере не очень просто найти что-то. [Можете попробовать](#). Но на Станции каналы нельзя группировать или фильтровать совсем. Чтоб понять, что смотреть вам придется листать страницы унылым "Дальше" и "Назад". Можно быстро переместиться "В конец" и ждать несколько секунд, чтоб все это великолепие пролисталось.

Ну, и запомнить синонимы: включи/покажи/поставь/запусти. Воспроизведи - не работает.

Ещё, Алиса умеет искать видео "по всему интернету" ([так написано в руководстве](#)) для чего нужно еще указать источник. Поскольку большинство видео в Интернете находится на Youtube, то это почти правда. Но можно вызвать и другие сервисы. Например, Vimeo - работает прекрасно. Но попытка вызвать RuTube провалилась - Алиса распознает команду, как Youtube. Ну, или у меня плохая

дикция. Или "что-то еще". Возможно это ["что-еще"](#) не позволило также включить в список каналов Эфира "Первый канал", "Россия", "НТВ" и прочие причастные к копирастической травле Яндекса генераторы контента. Я думаю, что это даже к лучшему.

Как итог - видео на Станции смотреть можно. Но найти что-то - навык для сильных духом, с хорошей дикцией и высшим лингвистическим образованием. Ну, или по меньшей мере, вы должны точно знать, что хотите смотреть, ибо рекомендательный навык у Алисы, скажем так, не очень.

Но тут, скорее, проблема во мне. Вкусы и предпочтения у меня странные и, очевидно, отличаются от "среднего россиянина". Плюс, отношение правообладателей к своим "сокровищам". Кстати, о "Плюсах". В комплекте с прибором идет годовая подписка на ["Яндекс.Плюс"](#) (1690 рублей), два месяца IVI.ru (798 рублей) и три месяца Амедиатеки (1599 рублей). Но я решительно не понял, почему бы я захотел там за что-то платить. Но, повторюсь, это мои личные проблемы и вкусовщина. Надеюсь, кому-то нравятся все эти "елки" и "ходячие мертвецы".

А вот то, для чего собственно предназначена "умная колонка" - для музыки - Станция очень хороша. Звук прекрасный. Выбор колоссальный. Рекомендации работают безупречно. Но это я связываю с тем, что пользуюсь сервисом давно и даже есть собственная "публичная радиостанция". [Можно пользоваться.](#)

Еще пара слов о лингвистике и мифологии.

Во-первых, у голоса Алисы есть лицо. Актриса, разумеется. По задумке разработчиков, было решено, что голосовым ассистентом станет молодая ироничная девушка, готовая помочь владельцу смартфона. Голосом "Алисы" стала актриса дубляжа [Татьяна Шитова](#), озвучившая большинство героинь Скарлетт Йоханссон и голос операционной системы "Саманта" в российском дубляже фильма Спайка Джонза "Она". Вот эта девушка:



Во-вторых, "голосовой ассистент" может помогать не только в "делах контента", но и вообще помогать. Алисе можно задавать вопросы (поиск Яндекса), попросить что-то напомнить или поставить таймер, погода, курсы основных валют. Можно даже поиграть в несложные голосовые игры. Вот тут полный [список доступных навыков Алисы](#). И список постоянно обновляется, что радует.

Но, как всегда, есть ложка дёгтя, и не одна. Первая - это русский язык. Который, в отличие от английского, является менее [аналитическим](#) и больше [синтетическим](#). Если не углубляться в дебри типологий языков, то в аналитических языках передача смысла идет (в большей мере) через синтаксис, где важен порядок слов и используются служебные слова (артикли и/или модальные глаголы, как пример). В синтетических языках смысл изменяется через изменение самих слов, путем добавления всевозможных приставок, суффиксов, окончаний. Синтетические языки (это не только русский, но еще, например, в некоторой степени немецкий и большинство семитских языков) труднее использовать для автоматической обработки - вариантов построения фразы и смысловых оттенков в предложении сильно больше.

Для примера: в английском языке вы всегда отличите вопрос от утверждения. По порядку слов: в вопросе всегда сначала сказуемое, а в утверждении - подлежащее. В русском языке вопрос вообще передается интонациями. Причем, не всегда.

И потому, "говорящий искин" для русского значительно сложнее, чем для английского. Есть мнение, что для китайского языка (который вообще в большей степени [изолирующий](#)) создать "говорящий на естественном языке автомат" еще сложнее. Но я лингвист не настоящий вообще, потому утверждать наверняка не могу.

Весь вышеописанный пассаж заключается в том, что да - пока "Алиса" не совсем все понимает и к ней нужно привыкать.

Особенно это заметно, когда "навыки" для Алисы начинают писать люди, которые не очень понимают, что делают. И да - "навыки" можно создавать. [Есть даже целый маркет-плейс навыков](#). И, разумеется, [документация](#).

Языковая проблема тут заключается в том, что для сторонних навыков вам нужно выстраивать сложные конструкции, которые довольно громоздко выглядят, если это действительно "естественный язык":

Активация навыка с командой

Некоторые навыки (например, справочные) может быть удобно запускать сразу с входными данными. Для этого нужно использовать специальные активационные слова:

- «скажи [активационное имя] [команда]»;
- «узнай у [активационное имя] [команда]»;
- «попроси [активационное имя] [команда]»;
- «спроси у [активационное имя] [команда]».

Например, «спроси Сбербанк где ближайшее отделение».

Причем, "активационное имя" - история, когда вам нужно с одной стороны обеспечить уникальность "навыка". А с другой стороны - "Алиса" плохо понимает всяческие неологизмы или аббревиатуры. Напомню случай с каналом "Юрган", который постоянно превращается в "Курган". Или "Рутюб", который распознается как "Ютюб".

Но направление с интеграцией "Алисы" и сторонних сервисов - очень хорошее.

Впрочем, это не единственная "ложка дегтя". Их много, на самом деле, но вторая из главных - наличие отсутствия инструментов аутентификации. Технологически сложная задача - отличать одного пользователя от другого.

Ну, то есть, вы можете как-то отличить конкретное устройство в миллионах аккаунтов. Вы можете даже это устройство привязать к платежной системе с помощью нативного приложения от Яндекс (что, собственно, и сделано при приобретении Яндекс.Плюс и Амедиатеки). Но действительно полезное приложение можно создать только если вы одного "собеседника" с "умной колонкой" сможете отличить от другого.

Для тех, кто разбирается в современных web-технологиях - для "Алисы" нужен аналог cookie-файла. Или привязку к приложению на устройстве. Для смартфонов это работает. А для стационарной колонки - нет. Тут нет никаких приложений. И нет дополнительного устройства ввода, например, пин-кода. И это проблема.

Как, Впрочем, проблемой является и потеря данных в диалогах, разнесенных во времени.

Зато, на Станции есть скринсейверы!



Там прекрасные красивые фото из космоса. Но картинок мало, а новых добавить нельзя - я не нашел как. Как и не нашел активационной фразы, чтоб скринсейвер запустился принудительно. Иногда хочется.

Мораль статьи - банальна. Яндекс.Станция это действительно прорывной продукт, который заставляет переосмыслить способы взаимодействия с информационными системами. И продукт впервые на русском языке, что важно. Ибо на английском "голосовые помощники" существуют уже пару-тройку лет и именно в факторе "умной колонки".

И очевидно, что голосовые интерфейсы довольно скоро займут существенную долю рынка "бытовых интеллектуальных устройств". Ну, потому что "голосовой помощник" очевидным образом может быть встроен в любой прибор, который может быть соединен с интернетом и имеет достаточно мощный процессор. И систему аудио ввода/вывода.

Первым на ум из таких приборов приходит телевизор. Собственно, именно потому в Яндексе и прикрутили видеоплату к первому серийному прибору со [SpeechKit](#) - "просто голос" слишком радикально для входа на российский рынок.

Впрочем, очень быстро (счет на месяцы) появились и более дешевые устройства:

[\[Искин по имени Алиса\]](#)

Это [Irbis A](#) и [DEXP Smartbox](#), которые умеют только голос.

Но вообще, "Алису" можно встроить вообще во что угодно - платформер Yandex IO к вашим услугам.

Так что, ждем новых клевых применений "Алисы" во всевозможной бытовой технике. Разработку, уверен, уже начали с десятка команд.

Насколько мне известно, работа ведется над несколькими моделями телевизоров (самое логичное) и телеприставок (самое разумное), "умным дверным замком", "универсальным домашним медиацентром", Wi-Fi-маршрутизатором и микроволновкой.

Очень надеюсь, что приборы появятся в продаже уже в этом году.

А пока, предлагаю посмотреть небольшой художественный (!) фильм, где "Алису" встроили в автомобиль и к чему это привело.

Ну, и хочется анонсировать вторую часть, в которой мы будем искать в Станции "товарища майора". Скоро на Hage!