

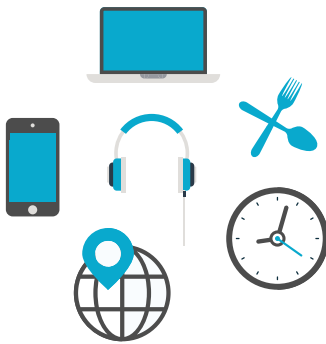
Мобильная Реклама: отдельное внимание

Интеллектуальное Определение Аудитории

Вы не могли обсуждать мобильную рекламу и не слышать термин «Большие данные». Издатели и разработчики приложений получают все больше и больше данных о своих пользователях, и их способность эффективно собирать, хранить и получать доступ к данным для таргетированной рекламы становится одним из наиболее важных элементов в индустрии мобильной рекламы.

Все начинается со сбора исходных данных из трех столпов информации: контекста, активности и демографии.

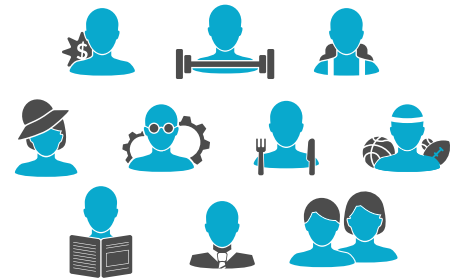
Контекст “Прямо сейчас”



Активность Контекст во времени



Демография Персональные характеристики



Но сбор данных - это только одна сторона медали; превращение этих данных в полезную информацию об аудитории - это настоящий вызов. В общем, использование в той или иной форме контекстной информации для таргетирования кампаний является нормой в течение многих лет.

Тем не менее, использование методик создания определенных аудиторий набирает обороты, в некоторых регионах более 85% кампаний ориентированы на аудитории, которые создаются специально для этих рекламных кампаний.

Так как создать аудиторию, учитывая все сложности? Начиная с данных, собранных из наших трех столпов, нужно интегрировать эти данные с четырьмя основными направлениями деятельности:

- 1. Раскройте контекстную информацию с сегментацией**
- 2. Уточните аудиторию по паттернам поведения**
- 3. Ускорьте с машинным обучением и стратегией измерения**
- 4. Будьте честными и прозрачными с конфиденциальностью**

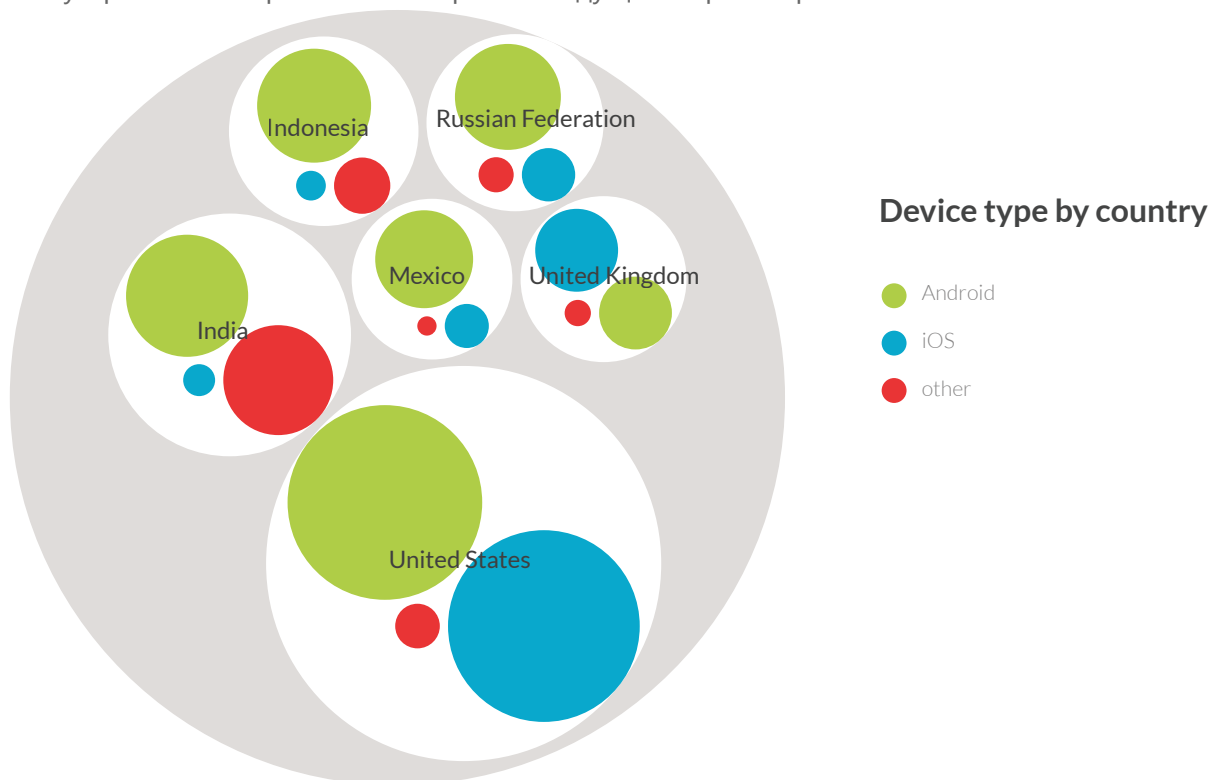
С помощью этих четырех направлений деятельности, рекламодатели могут определить аудиторию, состоящую из очень определенной группы потребителей, которые будут наиболее восприимчивы к рекламному сообщению кампании. Далее представлен наш взгляд на углубленное использование каждого из этих видов деятельности.

1. Раскройте контекстную информацию с сегментацией

Что ваша аудитория делает прямо сейчас? Это ответ, который можно получить из контекстной информации. Получаемая в реальном времени, она была основой таргетированной пользовательской рекламы в Интернете в течение нескольких лет. Когда вы видите объявления спортивной обуви на спортивных сайтах и объявления кино на развлекательных сайтах это значит, что маркетологи используют контекстный таргетинг. Рекламодатель предполагает, что, посетив этот сайт, вы демонстрируете заинтересованность в спортивных и развлекательных мероприятиях, и, следовательно, будете заинтересованы в продукции в этой же категории.

В мобильной экосистеме, примерами контекстной информации, используемыми для таргетирования, являются тип устройства, длина сессии, время суток, день недели, установленные приложения (или используемые), оператор сети и тип сети, а также местоположение пользователя.

Собирая и объединяя контекстные наблюдения по сегментам аудитории, мы можем разработать более полное представление о намерениях пользователей и вероятности их взаимодействия с конкретным типом объявления. Ниже представлено несколько сегментов аудитории, созданных путем объединения данных о типах устройства и стран на некоторых из ведущих мировых рынков.



Тем не менее, размещение объявлений является гораздо более сложной задачей, чем оптимизация по контекстной информации конкретного сайта, местонахождению пользователя или устройству. Могут быть сотни компаний, размещающих объявления с таргетингом на спортивные сайты, iPhone или пользователей социальных приложений в Лондоне. Какие объявления скорее всего получат отклик аудитории?

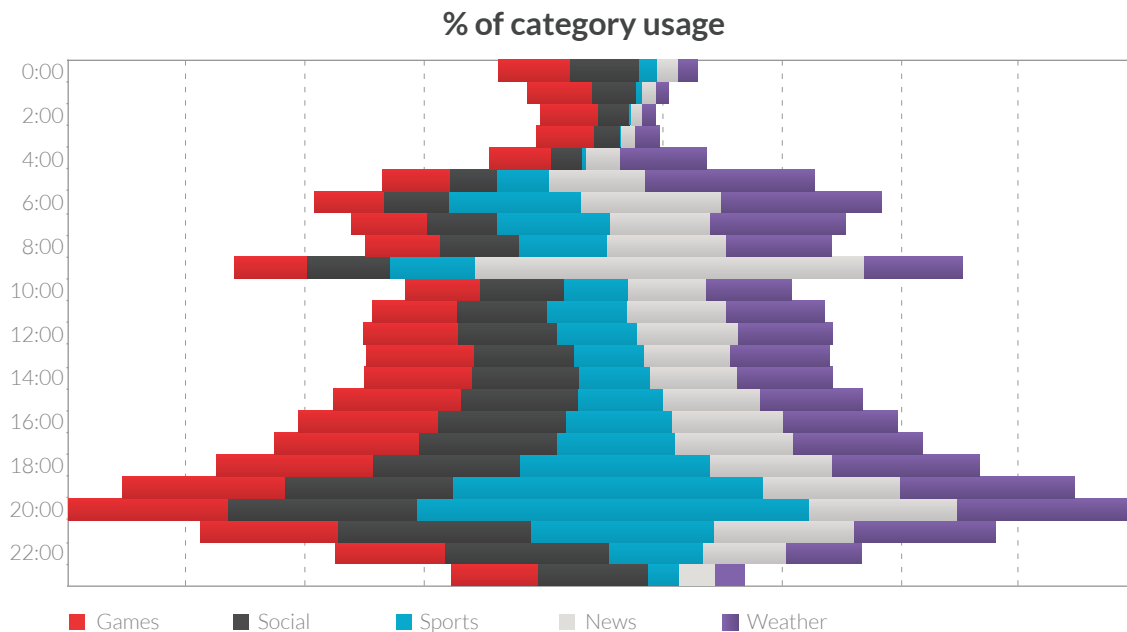
Чтобы определить истинную релевантность объявления для пользователя, требуется более глубокий анализ, который будет принимать во внимание больше переменных, таких как предыдущее поведение пользователя и демографическая информация.

2. Уточните аудиторию по паттернам поведения

Как мы можем точно наблюдать и понимать поведенческие паттерны? Вот где имеет первостепенное значение сбор данных на протяжении времени, состоящий из серии контекстных событий. К примеру, мы можем сказать, что мы знаем, что пользователь установил приложение в прошлом в результате контакта с объявлением. Это означает, что люди, скорее всего, сделают это в будущем, так что рекламодатель скорее покажет им подобный тип объявления, ожидая ответа. Пользователю, который не установил приложение в прошлом, с другой стороны, может быть показано низкооплачиваемое собственное объявление, так как ожидания от этого пользователя ниже.

Когда: Время суток, день недели

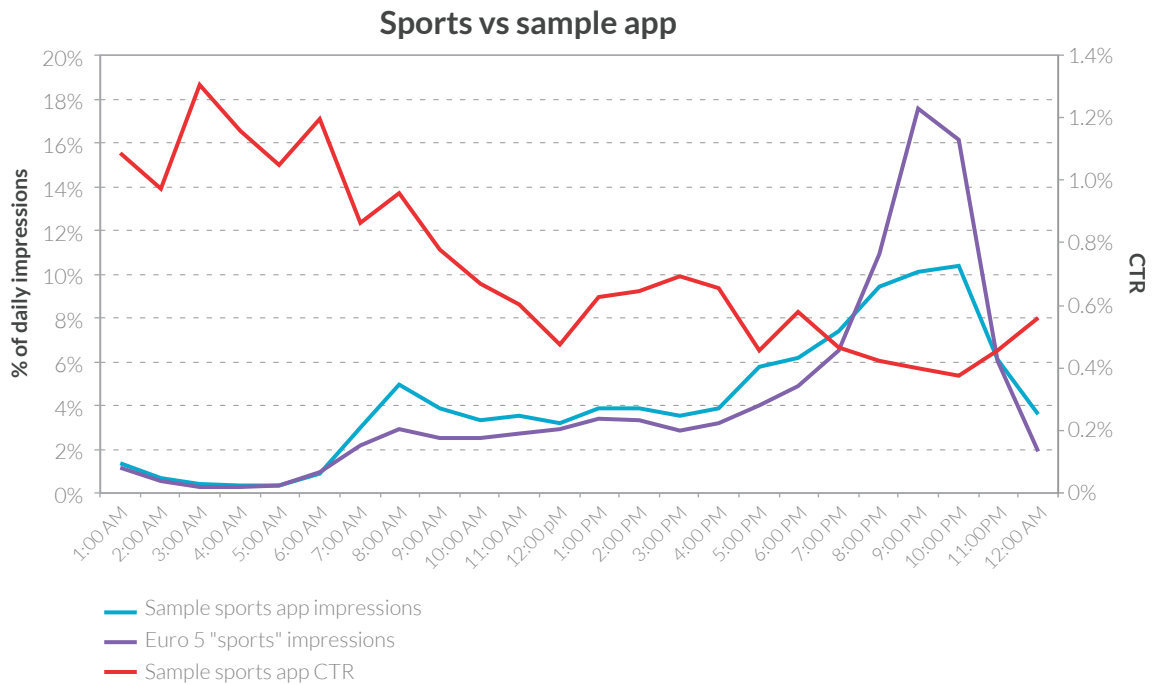
Наши данные показывают, что взаимодействие пользователей с рекламой существенно меняется в течение дня, а также зависит от дня недели. Например, в Топ-5 рынков Западной Европы, вовлечение пользователей по выбранным категориям в течение 24 часов показывает два основных периода: утром (с 6 до 10 часов), что объясняется высоким интересом к новостям, и позже в тот же день (с 6 до 9 вечера), где мы видим повышенный интерес к спорту, погоде и социальным сетям.



«Нужно» против «хочу»

Где разница между необходимостью и желанием? Когда потребителю "нужно", он требует, чтобы мобильный сайт или приложение немедленно исполняло его желание, например, осуществило поиск ближайшей кофейни или правильного инструмента для своего проекта. В состоянии "хочу", однако, потребитель наиболее восприимчив к рекламному сообщению бренда.

На графике мы видим сравнение объема показов и уровня кликабельности (CTR) на примере одного из самых популярных спортивных приложений региона. Когда объем показов высокий, CTR находится на самом низком уровне, это указывает на то, что пользователям "нужна" эта спортивная информация и, таким образом, они менее восприимчивы к рекламным объявлениям. Когда пользователь переходит в режим "хочу", CTR увеличивается, так как они вовлечены в общение с брендами посредством этого приложения.



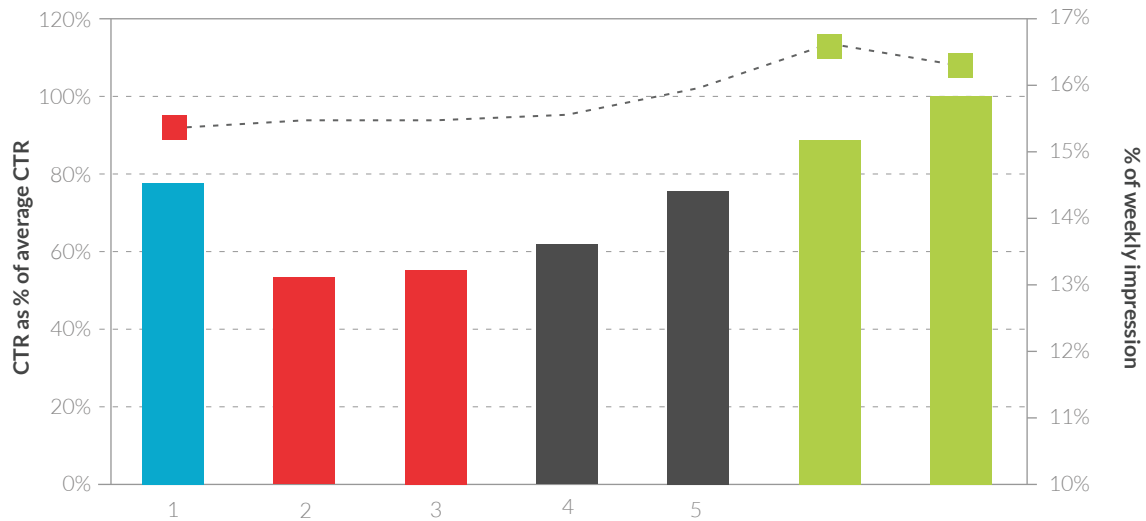
² <http://www.operamediaworks.com>, <http://www.flurry.com>

³ <http://www.flurry.com/bid/109749/Apps-Solidify-Leadership-Six-Years-into-the-Mobile-Revolution#.VRXrsTqRQkl>

На нашем суточном графике также можно увидеть, что категория Игры имеет очень низкую дисперсию на протяжении 24 часов. Более глубокий анализ этой категории выявляет влияние дня недели на поведение аудитории. График показывает объем показов и уровень кликабельности (CTR) для видео-рекламы в популярных игровых приложениях. Зеленый цвет показывает CTR, который значительно выше среднего показателя. Красный цвет отражает дни, когда этот уровень значительно ниже среднего. Теперь нам нужно определить свойства пользователей в различное время суток и день недели.

Video impressions served and CTR for a top game app by DOW

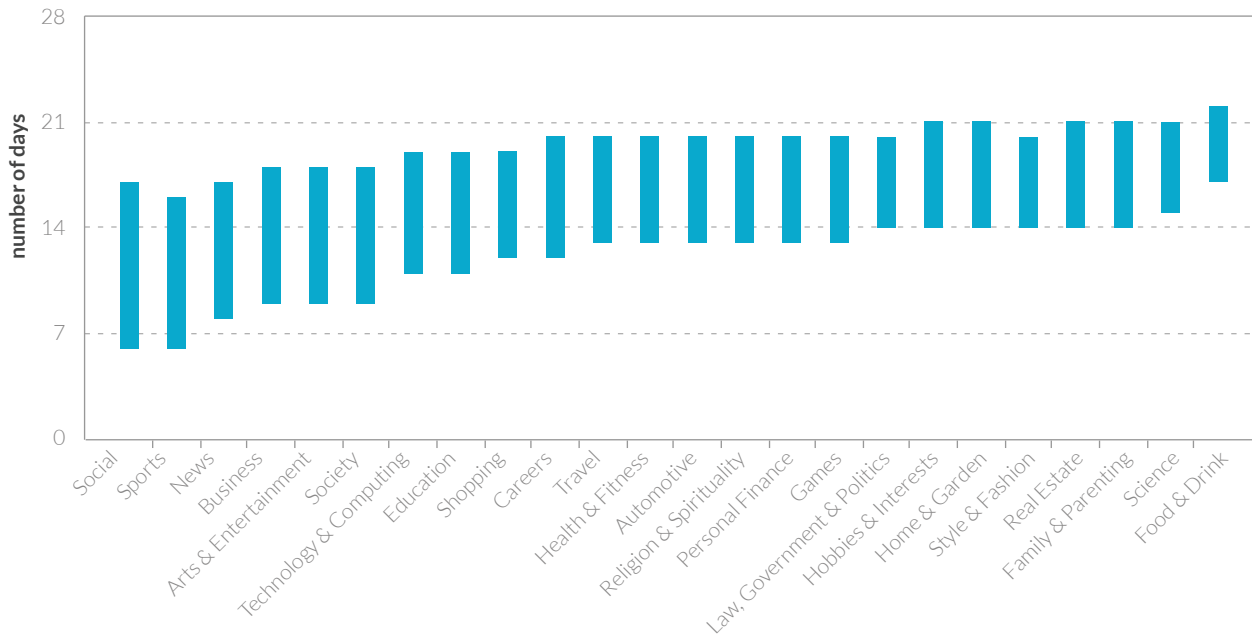
Red and green bars denote significant difference from mean



Как часто: частота использования

Частота использования сайта или приложения, часто измеряемая «сессиями» за определенный период времени, является еще одним важным показателем. Однако, понимание периода времени, в который ядро пользователей наиболее часто посещает приложение, также является важным. Как показано на графике, некоторые категории сайтов раньше встречают своих пользователей во временном цикле, в то время как другие имеют более постепенный цикл спроса. Этот цикл приобретения формируется из частоты сессии. В нашем примере, спортивные сайты получают более 50% своих ежемесячных пользователей в течение первой недели месячного цикла анализа, а дополнительные 20% появятся в ближайшие 10 дней. Напротив, сайты в категории Еда и Напитки ждут гораздо дольше, прежде чем приобретают соответствующие доли ежемесячных пользователей.

Days required to acquire 50% and 70% of monthly unique users



Как долго: проведенное время

В то время как частота запросов указывает на интерес и помогает нам решить, какой период времени нужно использовать для сбора данных, проведенное время - ключевой показатель для понимания того, насколько действительно глубок интерес пользователя. В общем, проведенное время формируется рядом факторов, таких как тариф, доступность и качество высокоскоростных сетей, способ взаимодействия (веб или мобильное приложение).

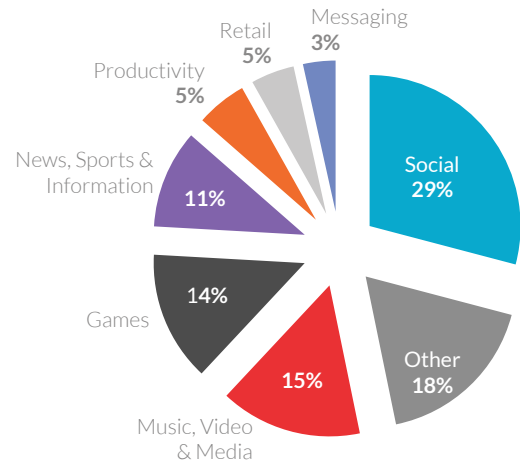
В Соединенных Штатах, например, мобильные приложения популярны и сети передачи данных повсеместны и цены на них разумны. Мы можем увидеть, как американские пользователи при таких условиях, проводят свое мобильное время. (Другой набор мобильных "норм", например, в других странах, может по-другому повлиять на распределение времени.)

Где: анализ интересов

До этого момента, акцент был сделан на информации, которая может быть собрана и проанализирована для определения аудитории на одном сайте или мобильном приложении. Деятельность на нескольких сайтах также может быть использована для определения сегментов аудитории, и это обеспечивает более глубокое понимание пользователей. В то же время реклама, основанная на поведении на нескольких сайтах, порождает многочисленные проблемы, связанные с частной жизнью, сегодняшние платформы управления данными (data management platforms, DMP) сосредотачивают усилия в этой области, чтобы предоставить рекламодателю полезную информацию, одновременно защищая неприкосновенность частной жизни потребителя.

Собирая и объединяя контекстные наблюдения, информацию о деятельности и демографии сегментов аудитории, мы можем разработать более полное представление о намерениях пользователей и вероятности их взаимодействия с конкретным типом объявления.

Time spent by site/app category
in the United States

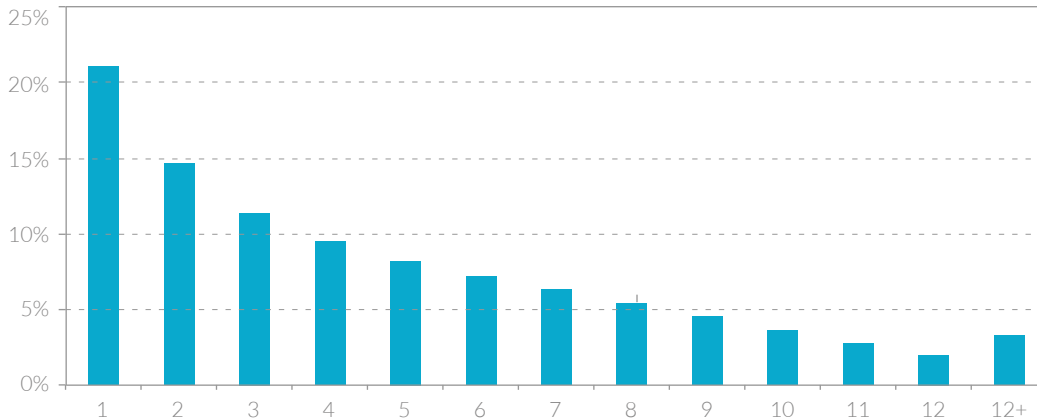


Reference: Opera Mediaworks and ComScore's "U.S. Mobile App Report, 2014"

Посредством нашего анализа мы выявили три основных типа пользователей:

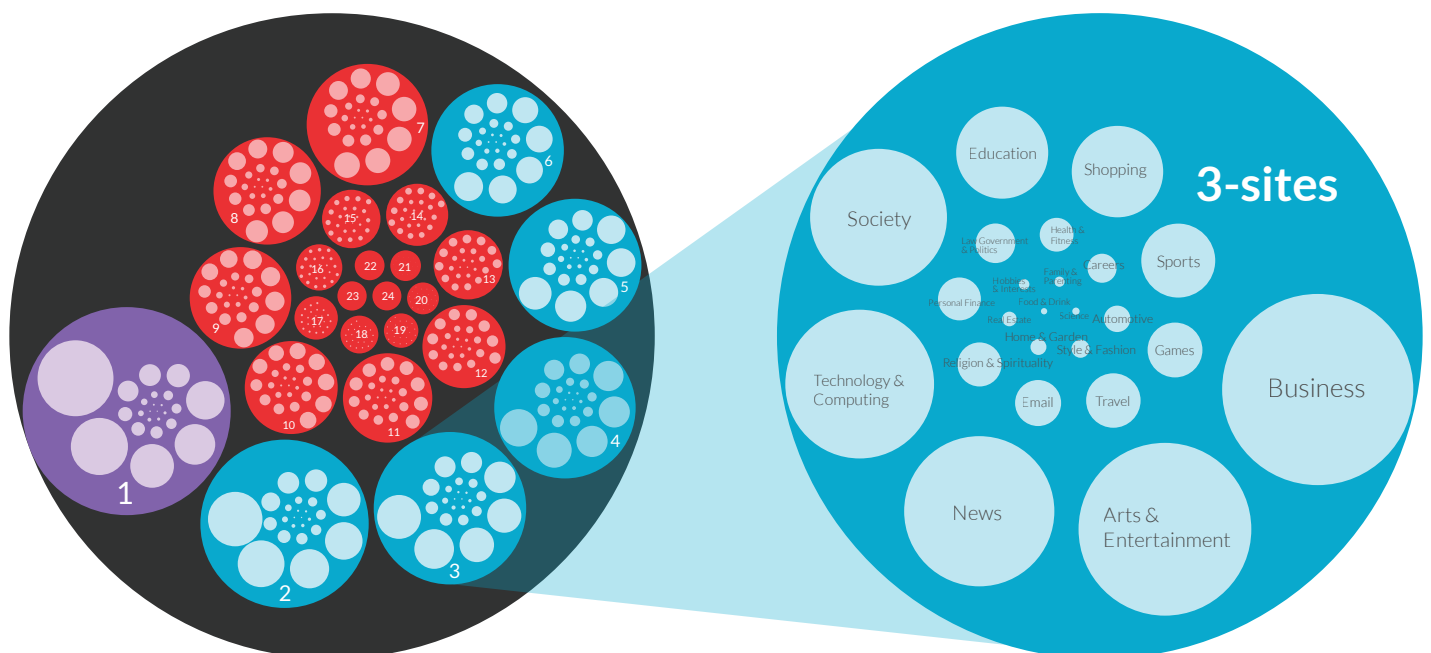
- **Мобильный однолюб.** Эта группа, которая представляет собой примерно 21% всех пользователей, фокусирует внимание на одном типе сайтов или приложений (как правило, в категории социальной или бизнес направленности), о ней мало можно узнать из поведения на остальных сайтах.

% of uniques by # of site categories visited in month

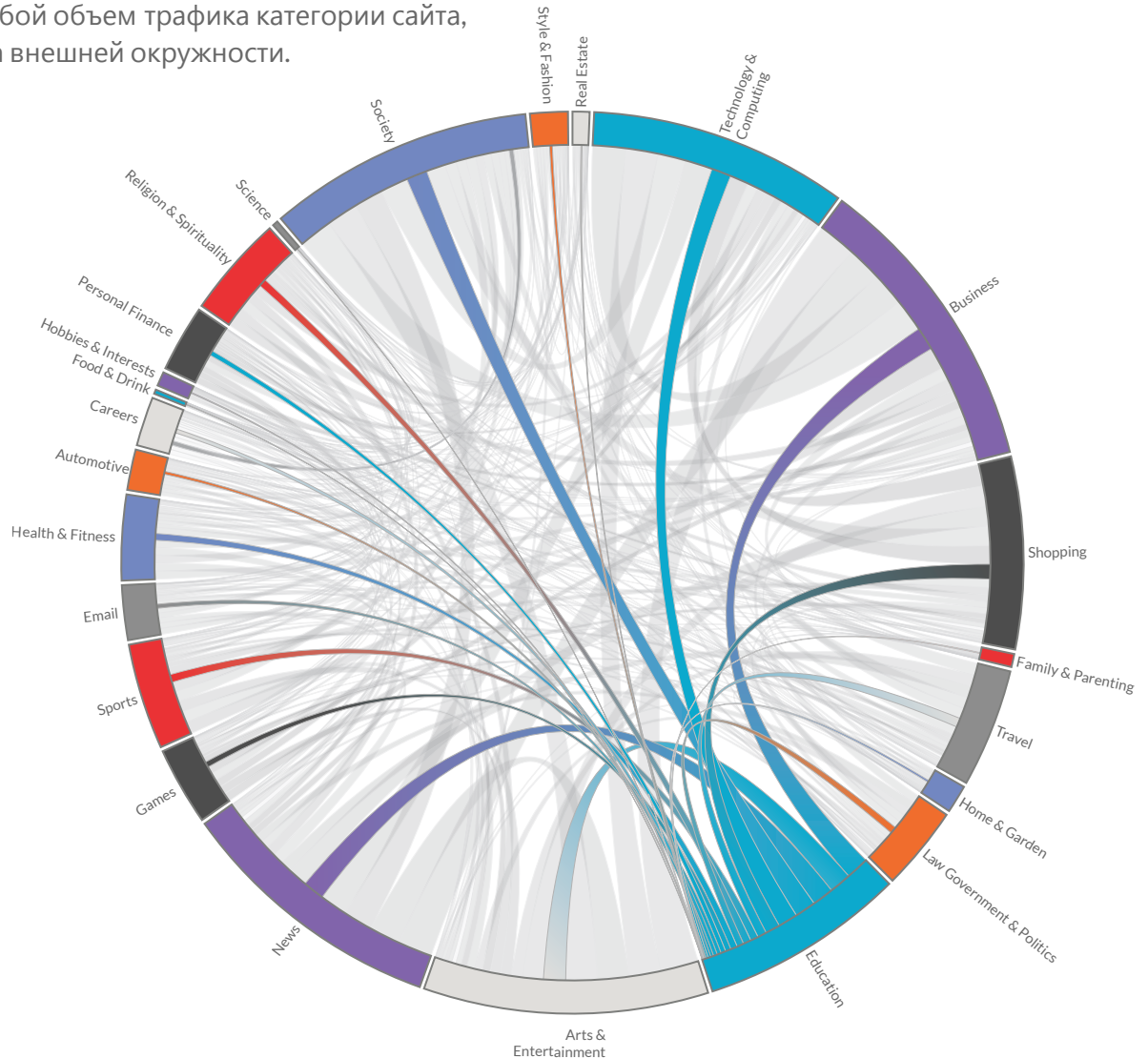


- **Любители разнообразия.** Эта группа посещает от 2 до 6 категорий и состоит примерно из 51% всех пользователей. График слева изображает три группы наших пользователей. Справа вы увидите разбивку сайтов, посещаемых пользователями трех различных категорий сайтов. Размеры пузырьков пропорциональны числу уникальных пользователей в каждой группе.

User clusters by number of sites visited



• **Контекстное разнообразие.** Эта группа (28% всех пользователей) имеет очень широкий круг интересов, посещает семь или более категорий сайтов. Диаграмма ниже показывает, как распределены визиты по другим категориям сайта одного из этих сегментов (пользователи, которые проводят большую часть своего времени на образовательных порталах). Размер соединительной линии представляет собой объем трафика категории сайта, приведенных на внешней окружности.



Так как вы лучше сегментируете аудиторию по ряду свойств, то можно ожидать, что результаты станут лучше. По факту, стремление к более высоким eCPM и успешным показам формируют постоянно растущий спрос на данные и их анализ, являются также движущей силой большей части инноваций в мобильной рекламе. Однако, вместе с улучшением вашей способности собирать и анализировать данные для определения аудитории, вы столкнетесь с двумя дилеммами:

1. Чем больше сегментированна ваша аудитория, тем меньше в результате будет эта аудитория. Чтобы быть успешным, вы должны сбалансировать эффективность таргетинга с масштабом.

2. По мере увеличения количества и разнообразия данных, которые вы собираете, с определением сегментов аудитории, выявленных из данных, показ объявлений будет все более затруднителен. Для человека (или даже группы людей) невозможен запуск такого вида таргетированных оптимизированных объявлений "в реальном времени". Вот где важна вычислительная мощность.

3. Ускорение с машинным обучением и измерением стратегий

Машинное обучение - это технология, которая использует исторические данные, чтобы сделать прогноз по новым. В подходах с машинным обучением, аналитические модели создаются с использованием алгоритмов для выявления закономерностей в исторических данных, которые называются обучающей выборкой. Затем мы можем использовать эти аналитические модели, чтобы сформировать прогноз для новых данных, как только они будут доступны.

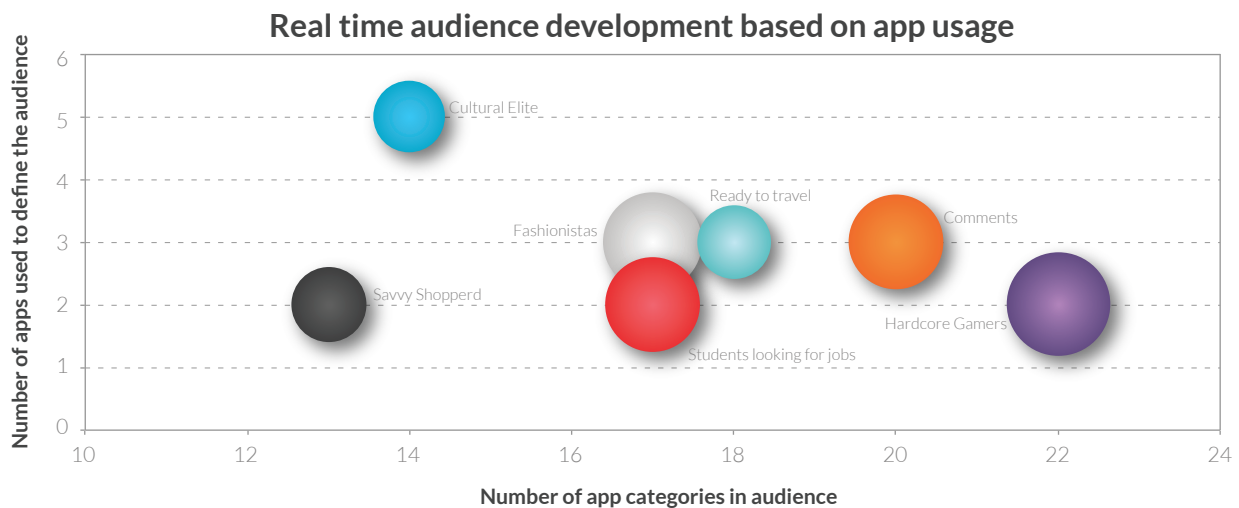
Алгоритмы машинного обучения гораздо эффективнее, чем человек при одновременном оценивании нескольких контекстуальных, поведенческих или демографических свойств, которые влияют на производительность рекламной кампании. Компьютерные алгоритмы анализируют несколько свойств данных одновременно, в то время как люди серьезно ограничены в своих возможностях по нахождению закономерностей в наборах данных рекламного трафика. Машинное обучение используется для создания широкого спектра прогнозов. В мобильной рекламе, машинное обучение часто используется для решения задач:

- Кластеризации аудитории
- Увеличение данных
- Оптимизация в реальном времени цены объявления
- Динамической адаптации созданного объявления

Кластеризация аудитории

При большой аудитории, как выделить сегмент для таргетинга из такого большого количества данных? Алгоритмы машинного обучения способны определить из широкого набора свойств те, которые наиболее актуальны для достижения цели кампании, таким образом, создается кластер аудитории. Мы можем создать правила, которые будут автоматически относить нового пользователя к кластеру таким образом, чтобы пользователь с наибольшей вероятностью соответствовал цели кампании и ему показалась определенная реклама.

Для иллюстрации подхода к кластеризации, мы взяли группу пользователей, вовлеченных в различные рекламные кампании и создали кластеры в зависимости от их поведения с помощью специальных приложений и набора контекстных переменных. Зная эти кластеры аудитории, система с машинным обучением может ассоциировать новых пользователей с кластером в режиме реального времени, одновременно информируя систему показа объявлений для отправки нужного объявления для этого пользователя на основе выбранного кластера.

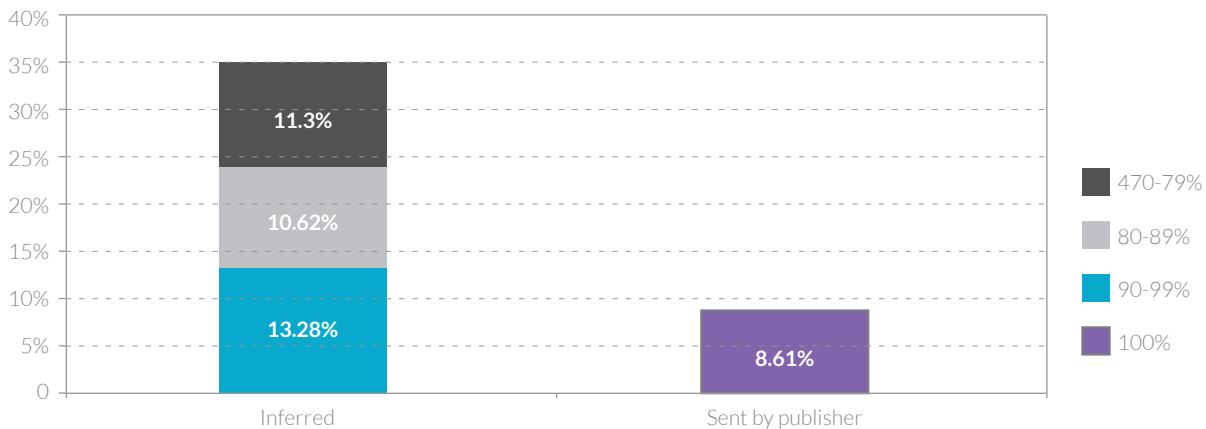


Увеличение данных

Публишеры предоставляют информацию покупателям о своих пользователях, и довольно часто этого достаточно для решения о показе объявлений и участия в аукционе в режиме реального времени (RTB), но часто отсутствуют некоторые ключевые элементы для покупателя. Например, сравнительно мало публишеров могут предоставить демографические данные о своих пользователях, но эти данные считаются критическими для покупателей.

Вот почему DMP набирают обороты на рынке; они могут дополнить предлагаемые данные для того, чтобы рекламодатели имели лучшее представление о том, кому они продают. Это дополнение может возникнуть из внешнего набора данных, или может быть получено с помощью статистических моделей. Мы можем также предположить неизвестные атрибуты, такие как пол. Например, недавний анализ RTB площадок показал, что только 9% запросов выдают пол. Для сравнения, алгоритмы машинного обучения сделали вывод о поле 34% пользователей. В то время как доверительный интервал сформированных данных был установлен на уровне 70%, а, следовательно, не является точным, присутствующие неточности при таком подходе могут быть смягчены за счет дополнительных расчетов и кластеризации аудитории.

Identification of gender for publisher and machine learning methods
by confidence level



Оптимизация аукционов в реальном времени

RTB - растущая часть экосистемы мобильной рекламы. С RTB, платформы спроса (DSP) должны определить оптимальное предложение для показа каждого объявления в режиме реального времени. Работа этой модели требует анализа на основе данных каждого показа для прогнозирования показателей производительности, таких как количество кликов (CTR), уровень конверсии и установок (CR/IR), и даже таких метрик после установки, как жизненная ценность (LTV).

На площадке RTB, DSP должен определять оптимальную цену за каждый конкретный показ. У большинства RTB-площадок максимальное время ожидания отклика 100 миллисекунд, что означает, что оценка показа по данным должна быть сформирована в течение очень короткого промежутка времени.

Для того, чтобы определить предлагаемую цену, алгоритм должен оценить вероятности результата показа на основе эффективных показателей, таких как CTR, CR/IR и тех метрик, появляющихся после установки, которые позволят нам узнать примерный LTV.

Модели машинного обучения в этой предметной области могут определять метрики вовлечения и эффективности, основанные на некоторых или всех типах данных об аудитории, обсуждаемых здесь. Довольно часто, самая трудная часть создания этих моделей заключается в определении нужных переменных для использования в анализе. В таблице ниже показан пример регрессионного модели, используемой для идентификации переменных, которые важны для более высокой вовлеченности.

Ad type	Variables		Page Context			
	Day of week	screen size	Entertainment	News	Social	sales
Media entertainment	Insignificant	Negative	Strong (+)	Insignificant	Positive	Insignificant
Product sales	Insignificant	Insignificant	Insignificant	Positive	Insignificant	Strong (+)

В конечном счете, ценность машинного обучения для RTB в том, что оно предоставляет рекламодателям инструменты, с помощью которых можно рассмотреть каждого пользователя как полноценную аудиторию.

Динамическая адаптация креатива

Поскольку мы приближаемся к персонализации, мы должны не только показать нужное рекламное сообщение пользователю, мы должны также динамически изменить объявление, учитывая контекстную составляющую. Это означает изменение таких элементов, как цвет, изображение, продукт и сообщение в режиме реального времени на основе того, что будет наиболее актуально для аудитории.

Как показано ниже, объявление можно динамически изменяться на основе таких факторов, как местоположение, язык, погода и время суток. Мы также можем адаптировать контент для просмотра в приложении или на мобильном сайте, или мы можем использовать результаты прошлых показов других объявлений с различными элементами для принятия решения о том, что показать. Кластеризация аудитории и оптимизационные модели, описанные выше, включились в игру в этой среде машинного обучения.



Погодные условия



Время дня



Место

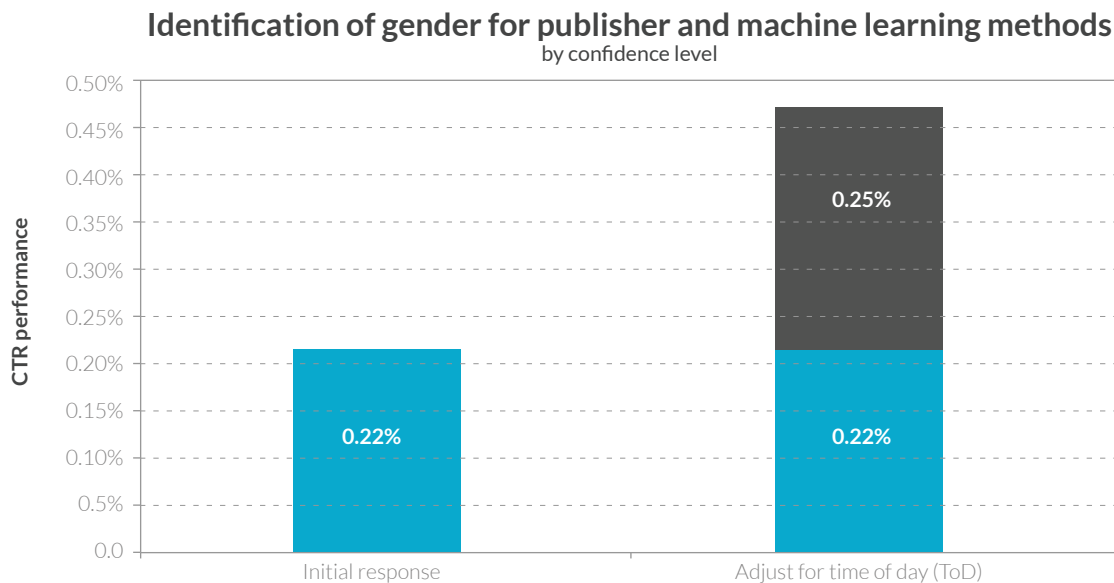
Языковые
настройки

Измерение всего

Наряду с задачами машинного обучения, сегодняшнее управление данными и системы анализа играют важную роль в измерении производительности. Способность быстро адаптировать кампании для оптимизации производительности, часто в режиме реального времени, имеет решающее значение для успеха маркетинга.

Но для того, чтобы определить "правильное" объявление для каждого целевого сегмента рынка, мы должны детально проверить объявления. Маркетологи должны быть знакомы с A/B тестированием. При определении аудитории, A/B тестирование означает разбивку аудитории на две части, отличающихся одним параметром (А и В метрики), которые определяют успех (например, конверсию). Чтобы определить, какая модель аудитории лучше, вы проводите эксперимент на обеих группах одновременно. Наиболее эффективный вариант может быть выбран для реального использования в кампании, или выставлен против другой модели для дальнейшего тестирования.

Как показано на графике, даже простые изменения могут принести большие результаты. В этом примере показаны результаты испытаний А, где время суток (ToD) не учитывается. Тест В подчеркивает разницу, достигнутую, когда переменная о времени суток скорректировала то, что аудитория может взаимодействовать в большей степени с рекламой в определенные периоды времени.



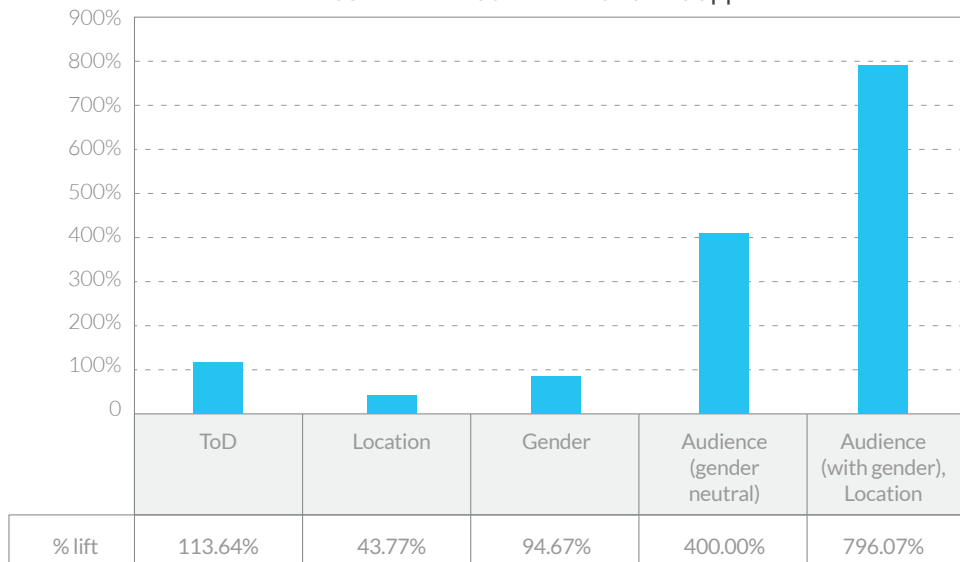
Сегодня A/B тестирование уступает многовариантному (многопараметрическому) тестированию. Многовариантное тестирование использует ту же основную предпосылку A/B тестирования, которая сравнивает производительность различных версий объявления. Однако, как следует из ее названия, у многовариантного тестирования несколько переменных изменяются одновременно. Объявление состоит из комбинации элементов, таких как фоновое изображения, положение заголовка и сообщения, и цель многовариантного тестирования в том, чтобы определить, какая комбинация этих вариантов является лучшей из всех возможных комбинаций.

Корректировка на лету и тестирование

Модели машинного обучения могут потерять эффективность с течением времени, так как поведение потребителя и контекстные переменные изменяются. Регулярный сбор пробных записей прогнозов для анализа и аудита целей (истинные ответы) помогает уточнять модели машинного обучения. Если качество прогноза начинается снижаться, потому что исходные данные больше не соответствуют реальному положению вещей, вы должны придумать новую модель или составить новые данные для обучения. Даже после начала кампании рекламодатели могут продолжить тонкую настройку моделирования аудитории, чтобы максимизировать эффективность. Этот вид "тонкой настройки" дает лучшие результаты, чем статический подход. Заметим, однако, как отмечалось ранее, различные свойства аудитория будут иметь различное влияние на производительность кампаний, в зависимости от характеристик этой аудитории или кампании. Итак, нужно измерять все - все время - и корректировать соответствующим образом.

Percentage of campaign performance lift

Baseline = no audience variables applied



Будьте честными и прозрачными с конфиденциальностью

Определение аудитории 100% опирается на собранные данные - но медиа издатели, рекламодатели и сторонние сервисы, такие как Opera Mediaworks, не могут вводить потребителей в заблуждение относительно необъявленного или скрытого сбора данных, их использования или политики безопасности.

В Opera Mediaworks мы считаем сохранение конфиденциальность ключевым предметом ответственности перед нашей отраслью, нашими клиентами и, в конечном счете, нашими друзьями, родными и нашими работниками. Ведь мы все также являемся частью аудитории. Чтобы обеспечить соблюдение наших обязанностей по отношению к потребителю, мы отстаиваем право каждого потребителя на:

- Простую в понимании политику конфиденциальности, доступную везде: в Интернете или в мобильном Интернете
- Выбор в отношении сбора, использования и распространения персональной информации
- Обоснованный доступ к реальным сотрудникам для обращения по вопросам или проблемам, связанным с политикой конфиденциальности и практике исполнения обязательств, а также доступ к любой личной информации обратившегося с возможностью исправления, изменения или удаления информации, показавшейся неточной или неполной
- Уверенность в том, что данные защищены от потери, неправильного использования и несанкционированного доступа, раскрытия, изменения и уничтожения