

предисловие



Эта книга — небольшой путеводитель по грамотной работе с данными, источник подходов и советов по созданию визуальных форм ваших данных. Мы хотим, чтобы вы научились создавать красоту из хитросплетений абстрактных показателей, умели задавать аналитические вопросы, овладели навыками целеориентированной разработки интерактивных дашбордов и визуализации.

Многие начинающие специалисты в этой области ощущают нехватку лаконичных материалов, которые бы охватывали весь спектр темы грамотной работы с данными. Поэтому мы сделали путеводитель. Мы постарались объяснить наиболее важные, ключевые темы в деталях и сразу же дать необходимые ссылки на более развернутые материалы для дальнейшего самостоятельного изучения. Это поможет вам ориентироваться в пространстве данных.

В основе путеводителя — результаты системной работы целого сообщества специалистов в рамках образовательных проектов DataYoga. Наше сообщество сформировалось и многое сделало за последние несколько лет благодаря экспериментам с новыми форматами обучения в онлайн и опыту почти 10 000 участников марафонов по работе с визуализацией данных.

В путеводитель включены элементы последних шести тематических марафонов по работе с программными продуктами для визуализации данных. Здесь вы найдете наработки из корпоративных тренингов ДатаЙоги (DataYoga), проводимых для крупнейших компаний в разных отраслях экономики.

Мы хотим поделиться опытом и начать системное исследование лучших практик, применяемых нашими клиентами и партнерами. Мы приглашаем к сотрудничеству участников профессиональных комьюнити, работающих с дизайном, программным обеспечением для бизнес-анализа, специалистов из бизнеса.

Как практикующие преподаватели с многолетним стажем, курирующие курсы в вузах и корпоративных университетах, мы видим, насколько быстро меняются подходы, технологии и требования в сфере анализа и визуализации данных. Классические руководства по дизайну данных так и остаются не переведенными на русский язык, и мы стараемся восполнить потребность в этих материалах, собирая по крупицам полезные практики. Мы делимся ими в рамках обмена опытом в наших онлайн-сообществах, которые ведем последние годы.

Современные комьюнити для новичка в этой сфере — основа для быстрого обучения, практики, обмена опытом. Мы благодарны всем нашим коллегам, которые внесли свой вклад в систематизацию материалов, поделились знаниями и продолжают просветительскую деятельность в этом направлении.

Недостаток необходимых русскоязычных материалов вынуждает многих специалистов самостоятельно обучаться работе с программными продуктами для визуализации данных. Многим это дается тяжело: непросто соединять технические навыки с визуальным анализом и критическим мышлением. Форма может теряться по мере того, как на первый план выходят срочные бизнес-задачи. Но именно визуальная форма и понимание того, как современный человек воспринимает информацию, во многих проектах становятся краеугольным камнем и влияют на успешность реализуемых задач.

Визуализация данных сильно недооценена в современном бизнесе. Но, как мы видим в последнее время, начинают набирать силу долгосрочные корпоративные образовательные проекты, нацеленные на грамотную работу с данными и визуальной аналитикой. Такие программы

ставят перед собой цель научить сотрудников видеть и понимать данные, чтобы принять обоснованные бизнес-решения.

Эта книга предназначена для тех, кто только пришел в мир данных и ищет хороший путеводитель для долгосрочной работы и накопления опыта. Мы пытались выстроить системный взгляд на область разработки информационных продуктов: интерактивные отчеты, презентации и дашборды. Мы объединили материалы наших отдельных марафонов, построили между ними связи и показали направления для дальнейшего изучения и развития навыков в этой сфере.

Многие примеры путеводителя представлены в скетчах, для того чтобы быстро передать некую идею и проиллюстрировать подход, но более подробные материалы разбора, множественные практические примеры вы, конечно же, найдете на страницах наших марафонов, которые находятся в бесплатном открытом доступе на сайте datayoga.ru

Учитесь с удовольствием. Виз-виз-ура-а-а!

зачем вам Data Yoga и чему вы научитесь

Для легкого старта: отметьте для себя те темы книги, которые кажутся вам наиболее интересными.

- Разобраться в управленческих технологиях и видеть ключевые показатели эффективности.
- Уметь правильно выбирать и настраивать виды визуализации под вашу задачу.
- Понимать пользователя и его поведение.
- Узнать, как различные шрифты управляют вниманием, что такое изящность и какой шрифт подобрать для цифр.
- Как переносить стили художников и их цветовые палитры в дизайн своих информационных продуктов.
- Уметь различать 50 оттенков серого, знать их названия и понимать природу цвета, уверенно подбирать свои целевые палитры для данных.
- Различать зрительные иллюзии и понимать людей с особенностями восприятия.
- Уметь выстраивать аргументы за использование таблиц для работы с цифрами и настаивать на применении визуализации для аналитических целей.
- Противодействовать когнитивным искажениям, знать их природу и не допускать ошибок в визуализациях и презентациях.
- Освоить составление информативных заголовков, аналитических выводов и аннотаций, правильно размещать метаданные.
- Найти эффективный алгоритм для разбора чужих дашбордов в деталях, уметь их переделывать и видеть нюансы визуализации.
- Научиться выявлять и собирать требования к интерактивному продукту.
- Соприкоснуться с историей первых визуализаций и появления дашбордов в разных областях человеческой деятельности.
- Поразмышлять над особенностью дизайна для мобильных устройств.
- Обосновать перечень компетенций и того, что нужно знать и уметь, чтобы быть грамотным в работе с данными.

- Сравнить требования и детали современных стандартов по визуализации данных и разработать свой гид по стилю дашбордов.
- Хорошо ориентироваться в графических стилях и уметь подбирать их для визуализаций.
- Научиться скетчить, графически передавать свои идеи быстрыми набросками, и собрать свой визуальный словарь.
- Научиться проектировать дата-истории с эмоциональным вовлечением зрителя.
- Проектировать слайды презентаций и не тратить на это много времени.
- Освоить приемы интерактивности для удобной аналитики.
- Разбираться в различных типах данных и понимать их связь со шкалами и визуальным представлением.
- Практиковать принципы DataYoga — осознанность и концентрацию на текущем моменте.
- Уметь научить даже ребенка понимать, любить и проектировать визуализации.



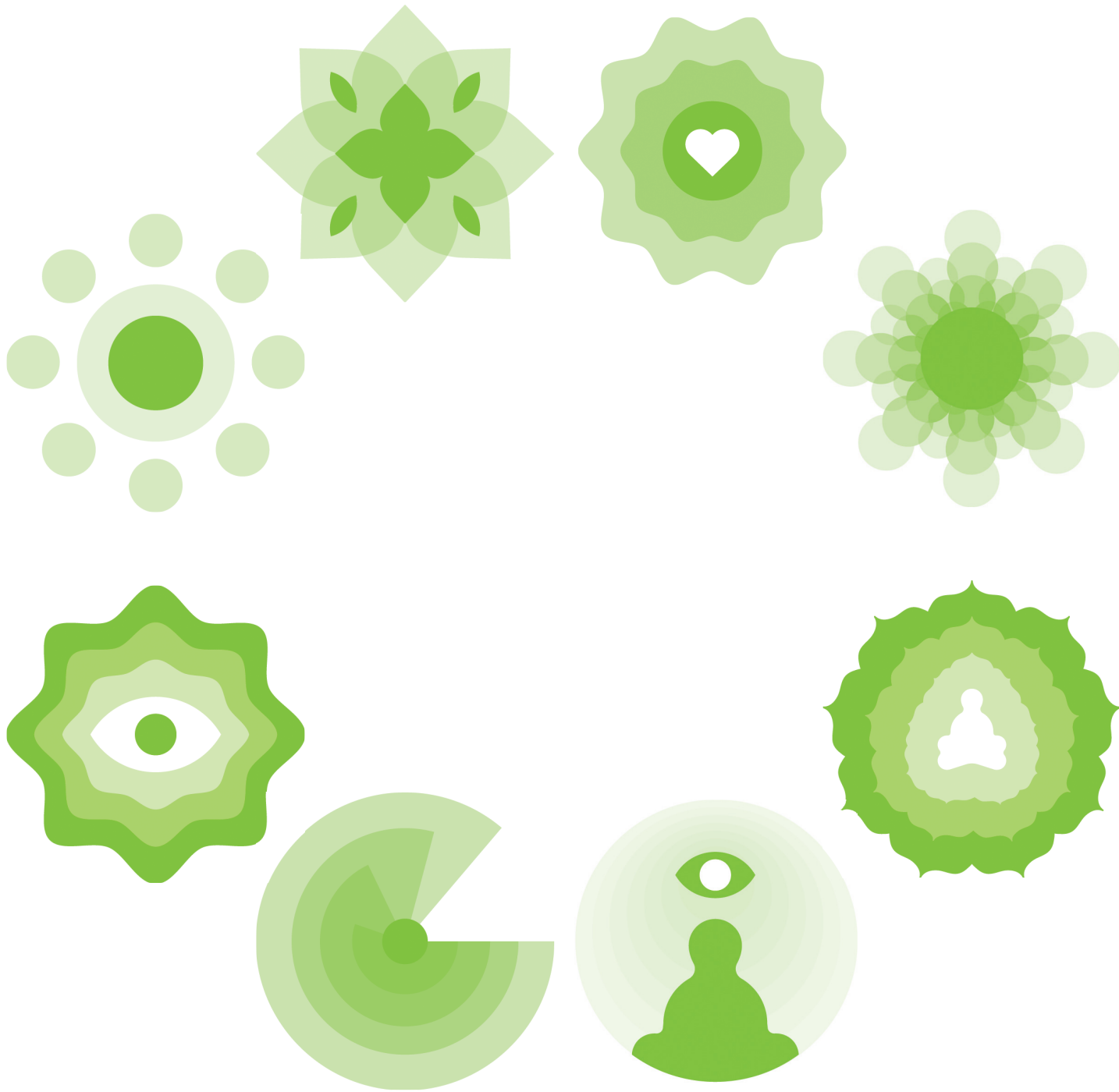
раздел



o DataYoga

В основе традиционной йоги, из которой развились принципы работы с данными DataYoga, лежат практики и способы видения, помогающие человеку лучше узнать самого себя, сформировать свой особый способ видения.

За время, прошедшее со времен систематизации великими мистиками древнейших текстов и опыта индийской культуры, появилось множество интерпретаций слова «йога». Мы хотели бы обозначить, в каких контекстах мы будем трактовать термин «Датаяога», возникший на стыке современных технологий обработки и визуализации данных и потребности людей привнести в свою жизнь желаемые перемены, опираясь на растущий объем данных нашего быстро меняющегося мира.



DataYoga — это



Идея достижения того, что прежде было недостижимо, лежало за пределами наших возможностей, — это идея изменений, которые происходят, когда мы предпринимаем действия, позволяющие нам выйти за свои границы и в ходе дискуссии лучше понять себя или других.



Это образ действия, при котором наше внимание целиком и полностью сконцентрировано на том, чем мы занимаемся в настоящий момент, на тех задачах, которые мы решаем с помощью имеющихся у нас данных. Мы можем быть погружены в процесс,

но на самом деле уделять ему очень мало внимания; действовать, но не присутствовать в своем действии.

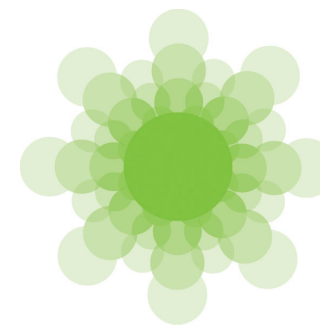
DataYoga стремится к достижению такого состояния, когда вы присутствуете в каждом действии, которое вы осуществляете со своими данными. Концентрация внимания позволит вам выполнять работу лучше, чем прежде, и при этом осознавать все, что вы делаете. Ваша внимательность снижает вероятность ошибок, а осознанность во всех ваших действиях избавит от старых ненужных привычек. Вам уже не нужно будет делать что-то

сегодня только потому, что вы делали это вчера. Так вы сможете избежать бессмысленных повторений в своих поступках.

В истинной практике DataYoga каждый человек идет своей дорогой. И для движения вперед нужно действовать, следя за направлением своего движения. Это поможет всегда знать, где находится цель и что нужно, чтобы ее достичь. Такое внимательное наблюдение позволит вам постоянно обнаруживать что-то новое.

принципы DataYoga

- Не использовать визуализацию данных во вред обществу и самому себе.
- Избегать манипуляций с данными и отказываться от лжи, профанаций и самообмана.
- Отказаться от бесцельного накопления данных, не имеющих смысла.
- Постоянно, с помощью размышлений, общения и обучения, приближаться к самому себе и изучать свои способности эффективно понимать и обрабатывать данные.
- Предпринимать максимально эффективные действия по работе с данными независимо от нашего эмоционального состояния.
- Стараться совершать дела на благо всего общества, чтобы сделать мир вокруг себя менее хаотичным и более структурированным.
- Содержать данные в чистоте и при возможности обогащать их, чтобы добавлять глубину видения при проведении аналитики.



8 шагов DataYoga

Наше внимание целиком и полностью сконцентрировано на том, чем мы занимаемся в настоящий момент.

Обретаем здоровый разум и баланс при работе с данными внешнего мира путем соблюдения определенных правил: использования визуализации во благо, отказ от обмана, манипуляций и профанаций, отказ от накопления бесполезных данных и желания обладать ими бесцельно. Жизнь в ладу с внешним миром.

- Применение визуализации в разных доменах.

Когда же мы обладаем ясным пониманием, мы чувствуем глубоко внутри себя тишину и покой.

Налаживаем гармоничное взаимодействие со своим внутренним миром и своими данными. Учимся держать данные в чистоте и преобразовывать их для визуализации. Понимаем свои компетенции и выстраиваем путь самообразования. Самодисциплина приведет нас к осознанной удовлетворенности и возможности делать полезные визуализации на благо всем окружающим.

- Саморазвитие в области работы с данными.
- Форматы обучения.

Методологические практики для понимания чувств аудитории.

Учимся концентрироваться на конечной цели и держать в фокусе задачи нашей аудитории: максимально погружаемся во внутренний мир пользователей наших визуализаций и вовлекаем их в циклы прототипирования дата-продуктов.

- Понимание аудитории.
- Проведение интервью.
- Важность KPI.
- Подходы к ведению дата-проектов.

Асаны присоединения к данным помогают объединять источники.

Налаживаем навыки подключения к данным и их объединения. Понимаем происходящее за пределами зоны видимости для достижения максимального внимания и эффективности. Учимся практически.

- Вопросы к данным.
- Как мы измеряем и понимаем данные.
- Виды агрегации.
- Способы объединения.
- Корпоративное управление данными.



Визуальное дыхание дает почувствовать данные.

Визуальная пранаяма — «аяма» означает расширение — описывает воздействие пранаямы. «Прана» — это то, что есть повсюду, то есть данные. Учимся удерживать внимание на дате и расширяем видение находящейся за данными реальности.

- История визуализации.
- Принципы визуального дыхания.
- Как работают диаграммы.
- Как читать визуализации.
- Каталог визуализаций.
- Прототипирование и скетчинг.
- Стандартизация.
- Важность цвета и шрифта.



Управление текущим моментом, концентрация на дашборде.

Собираем все свои визуальные навыки и, концентрируя внимание, используем их в разработке идеального дашборда и его адаптации под мобильные устройства для получения идеального дата-продукта. Правило «быть здесь и нигде больше» является ключом к полной концентрации. Есть только текущий момент. Нет ничего больше.

- Что такое дашборд.
- Основные виды дашбордов.
- Точки внимания и сетки.
- Мобильные макеты.
- ТВ-дашборды.
- Ошибки на дашбордах.



Медитация и интерактивность проявляют связи в данных.

Отбрасываем все и ощущаем только предмет нашей медитации, настраивая интерактивные элементы на дашборде. Это позволит создать визуальный ответ на наши целевые вопросы. Учимся проектировать пользовательский опыт с помощью интерактивных элементов. Когда человек постоянно концентрируется на одном объекте, он обретает все более и более глубокое понимание этого объекта.

- Виды интерактивности.
- Процесс аналитики.
- Пять видов анализа данных.



Сверхсознание — это уже ваша история.

Убираем временные и пространственные рамки: практикуемся подавать наши инсайты с помощью различных подходов выстраивания дата-историй. Медитируем регулярно. На восьмом шаге ДатаЙоги деятельность ума характеризуется устойчивой фиксацией внимания.

- Истории против анализа данных.
- Приемы сторителлинга.
- Идеальная история на языке данных.



раздел



о Data Literacy, или что такое грамотность в работе с данными

DataYoga вместе с глобальной инициативой Data Literacy Project стремится сформировать сообщество увлеченных людей, которые осознают важность навыков грамотной работы с данными.

На практике термин Data Literacy означает набор навыков, который помогает работать с данными без ошибок, решая свои задачи. При этом это не только технические навыки работы в программном обеспечении — компетенции включают в себя понимание концепций управления данными, знание базовой статистики, а также умение выдвигать гипотезы и принимать обоснованные решения на основе данных.

Грамотность работы с данными — это умение понимать данные, работать с ними и анализировать их, а также правильно интерпретировать и использовать. Необходимо также уметь обосновывать свое мнение и выводы, опираясь на данные.

Последние несколько лет компании по всему миру включают программу развития навыков Data Literacy в свои внутренние стандарты обучения сотрудников. Менеджмент верит, что эти навыки нужны не только аналитикам, но и вообще всем. При этом грамотная работа с данными востребована не только в бизнесе.

Сайт российского проекта Data Literacy, курируемого DataYoga совместно с партнерами Qlik и Accenture

dataliteracy.ru

Манифест и миссия Data Literacy. Ключевые принципы развития навыков Data Literacy

dataliteracy.ru/manifest

Все ссылки книги собраны на странице

datayoga.ru/book-links

почему вам стоит обратить внимание на работу с данными?

.01

Данных становится больше

Человечество производит все больше и больше данных год от года. Новые инструменты и технологии, развитие способов хранения и обработки данных, увеличивающееся число транзакций и изменение пользовательских привычек. NFT- аукционы, глобальные базы данных по заболеваемости, хранилища рекламных предпочтений в интернете, поставленные лайки и реакции, отслеживание здоровья и заполненности холодильника.

.02

Данные становятся более ценными

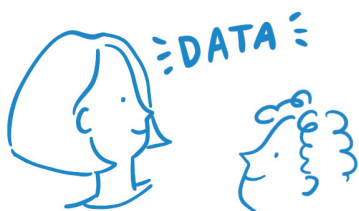
Компания делает свои данные доступными для сотрудников, чтобы они сами могли строить отчеты или работать с данными. К счастью, уже многие компании знакомы с self-service-аналитикой и не боятся отдать данные рядовым сотрудникам для самостоятельного анализа. В обществе формируются порталы и организации по работе с открытыми данными, создаются площадки для краудсорсинга данных: каждый может добавить свой кусочек к мировому пазлу данных.

.03

Данные становятся доступнее

Поскольку объем данных в последние годы растет экспоненциально, аналитика коренным образом изменила методы работы компаний. Компании пытаются перевести все процессы в цифровой вид: оцифровать и тегировать. Для организаций очень важно использовать эти данные для принятия правильных решений — и они готовы давать людям доступ к данным. Без обученного персонала продвинуться дальше этого шага практически невозможно.

культура грамотной работы с данными



Как выглядит культура грамотной работы с данными в организациях?

Развитие культуры работы с данными в компании требует от ее сотрудников как технических навыков, так и soft-skills — того, как вы подходите к данным и обращаетесь с ними. Совокупность этих компетенций образует особый свод характеристик культуры Data Literacy:

Аналитические навыки.

Это не статистические навыки, а способность анализировать, интерпретировать и понимать данные. Задавать правильные вопросы — ключ к успеху.

Статистические методологии.

Хотя не каждый должен иметь степень по высшей математике, каждая организация нуждается в статистических навыках в какой-то степени, будь то несколько специалистов

по данным или целая команда по data science.

Визуализация данных. Она упрощает их понимание и должна быть неотъемлемой частью культуры информационной грамотности, помогая широкому кругу людей общаться друг с другом на языке данных.

Обучение. Необходимо иметь внутренний центр обучения и повышения квалификации. Сотрудники должны иметь

возможность постоянно изучать новые концепции данных без отрыва от рабочего процесса. Будет здорово, если развивать компетенции разрешается в рабочее время.

Менторство. Необходимы наставники по грамотности данных — те, кто желает помочь в обучении и формировании растущих навыков своих коллег.

ГОВОРИТЬ НА ЯЗЫКЕ ДАННЫХ — ЭТО КАК?

Навык читать и понимать данные снимает «языковой» барьер в понимании всего процесса работы с данными. Что такое гистограмма? Из чего состоят базы данных? Когда нужно использовать медиану? Мы учимся разбираться в терминологии и без проблем считывать данные.

Навык работы с данными — это способность взаимодействовать с полученными и (уже успешно) прочитанными данными. Понимание инструментов для их обработки, использование методов визуализации, подготовка данных для презентации.

Навыки анализа данных дают возможность приходиться к выводам. Вы можете владеть статистическими методами и алгоритмами

для проведения самостоятельного анализа и поиска инсайтов (озарений) в данных. Или же читать уже подготовленные кем-то аналитические отчеты и делать выводы на основе полученной информации. Статистические навыки не равны аналитическим, а скорее являются их частью.

Навык вести аргументацию и коммуникацию с помощью данных — это умение отстаивать свою точку зрения (гипотезу, предположение, действие) с помощью данных. Это новая альтернатива традиционной интуиции, которой часто руководствовались при принятии решений.

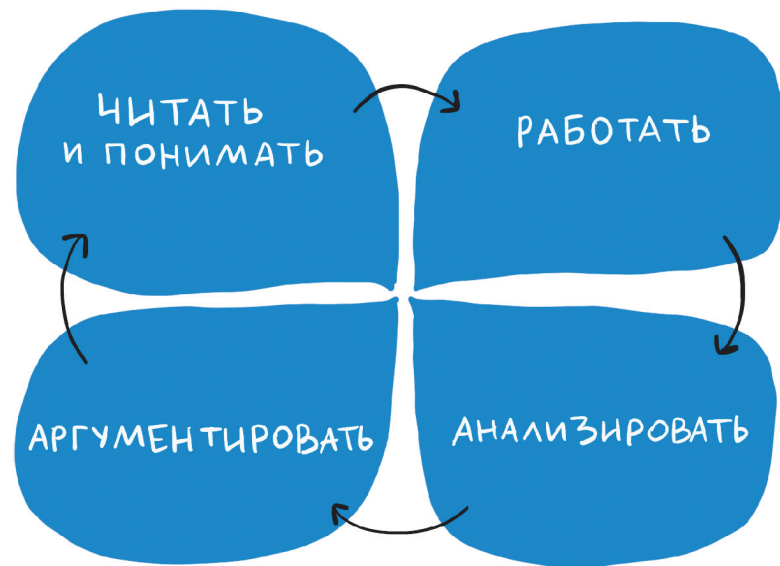
Компаниям с высоким уровнем грамотности работы с данными при принятии решений не нужно опираться только на нескольких дата-сайентистов, которые и без этого обычно сильно загружены. У большего числа сотрудников тоже будут навыки анализа данных, и в организациях смогут приходиться к эффективным решениям быстрее.

Сотрудники научатся задавать правильные вопросы и выдвигать гипотезы, а также чувствовать себя более комфортно в общении с руководством при помощи данных. Компании будут конкурентоспособны, потому что смогут быстро разобраться в своих данных, чтобы адаптироваться к происходящему и опережать своих конкурентов. В конце концов, весь персонал будет чувствовать свою причастность к общему процессу анализа данных и процессу принятия решений.

В компании с развитой культурой данных сами данные рассматриваются как второй язык, на котором говорят все. Все, кто когда-либо учил иностранный язык, понимают, что изучение любого языка — комплексное освоение навыков чтения, понимания, разговора и восприятия.

На специальной доске Miro вы сможете узнать больше про навыки и компетенции по работе с данными и модель оценки уровня зрелости организаций в рамках подходов Data Literacy

miro.com/app/board/o9J_kpnxz0k=



Свободное владение данными включает в себя четыре ключевые компетенции.

Эти компетенции не конечны и не являются строго обязательными. Степень развития каждой группы навыков зависит от вида, частоты, уровня, профиля деятельности каждого отдельного человека. Но такой список может служить хорошим ориентиром для возможных стремлений.

Истории успеха

Изучите примеры развития культуры данных в разных компаниях на сайте проекта Data Literacy

dataliteracy.ru/cases

К языку данных применяются те же самые подходы, которые мы используем в нашем «разговорном» мире. Мы учимся читать и изучаем основные блоки языка. Мы учимся слушать других и обрабатывать то, что воспроизводится другими. Мы начинаем формулировать собственные сообщения. Мы делимся ими с другими людьми и наблюдаем за реакцией окружающих на наши мысли и предложения. И конечно же, мы стремимся пообщаться с теми, кто свободно говорит на новом для нас языке.

- Знание языка данных помогает произвести хорошее впечатление и получить достойную оплату своих умений. Сегодня в компаниях нужно не только знать предметную область, но и понимать принципы обработки и анализа данных для принятия решений.
- Знание языка развивает такие качества, как уверенность в себе и самоконтроль, повышает эффективность работы и коммуникацию с коллегами. Изучение языка — непростое, но благодарное занятие, повышающее самооценку и уверенность в себе.
- Знание языка расширяет ваши горизонты. Язык — самый надежный ключ к другой культуре. Современный глобальный мир, который стремительно становится все меньше, требует умения понимать людей, принадлежащих к другой культуре или занимающихся в организации смежными направлениями. Это становится жизненной необходимостью.

Людам свойственно бояться того, чего они не знают. Чем лучше вы узнаете ценности и особенности жизни других, тем более очевидным для вас станет тот факт, что различия между людьми не так велики, как может показаться на первый взгляд.



Развитие компетенций Data Literacy неразрывно связано с развитием культуры работы с данными в компании, креативности среди ее сотрудников и открытой коммуникации, где могут появляться смелые идеи и новые подходы.

СКОЛЬКО ЛЮДЕЙ ОБЛАДАЕТ НАВЫКАМИ ГРАМОТНОЙ РАБОТЫ С ДАННЫМИ



Исследование Censuwide от имени Qlik* показало, что только 24% лиц, принимающих решения в сфере бизнеса, от младших менеджеров до руководителей высшего звена, полностью уверены в своей способности читать, работать с данными, анализировать и спорить на основе данных — то есть имеют фундаментальные навыки, которые определяют грамотность человека в области данных.

В России компания Qlik тоже проводила сертификацию: было задано 50 вопросов по темам понимания данных, базовой статистики, методов анализа данных и интерпретации бизнес-требований. Сертификация была частью подготовки образовательного марафона от DataYoga и Qlik «Заговори на языке данных».

** В исследовании участвовали 7377 лиц, принимающих решения в бизнесе. Исследование проводилось в период с августа 2017 года по февраль 2018 года. Респонденты были отобраны из Европы, Азии и США.*

Из 1500 участников только 11% набрали проходной балл и получили сертификат.

Почти 50% участников не смогли сопоставить приведенный график распределения с указанной величиной моды.

Возможно, те же 15% ответили, что бимодальное распределение, вероятнее всего, имеет распределение веса новорожденных малышей.

Для 11% участников утверждение «значение показателя увеличилось на 200%» можно описать в цифрах как «было 100, стало 200».

Тем не менее сертификация не позиционируется как полная и объективная проверка знаний. Это скорее возможность для каждого участника узнать больше о многообразии факторов, которые стоит учитывать при работе с данными. Одновременно это возможность узнать свои слабые стороны и наметить для себя новые горизонты.

Марафон «Заговори на языке данных» — это бесплатная онлайн-программа получения навыков понимания и анализа данных.

dataliteracy.ru/speakdata



ПРОТЕСТИРУЙТЕ СВОИ НАВЫКИ РАБОТЫ С ДАННЫМИ

Вы тоже можете пройти тестирование, чтобы узнать свой уровень знаний: от простого и развлекательного теста до более серьезной сертификации по работе с данными, включая вопросы на знание BI-инструментов.

Для начинающих подходит тест на определение вашей роли в работе с данными. От скептика до эксперта — насколько вы доверяете данным и верите в их силу для принятия решений?

Онлайн-тест из 10 коротких вопросов для тех, кто хочет немного повеселиться

dataliteracy.ru/novice

Онлайн-тестирование из 50 вопросов по навыкам работы с данными

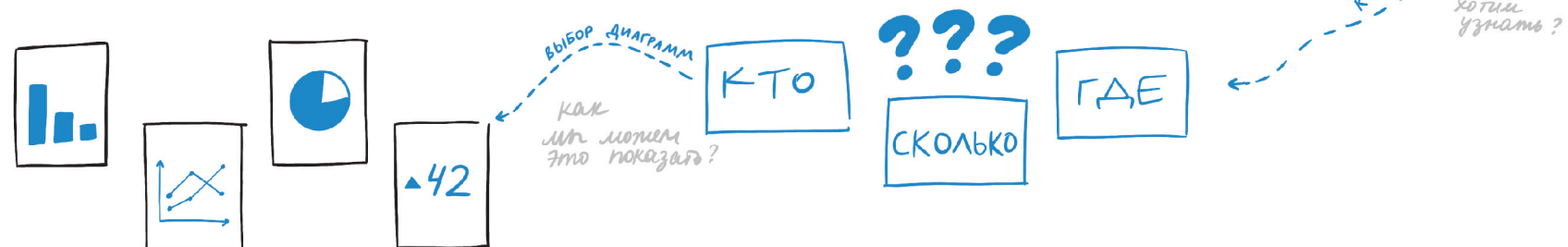
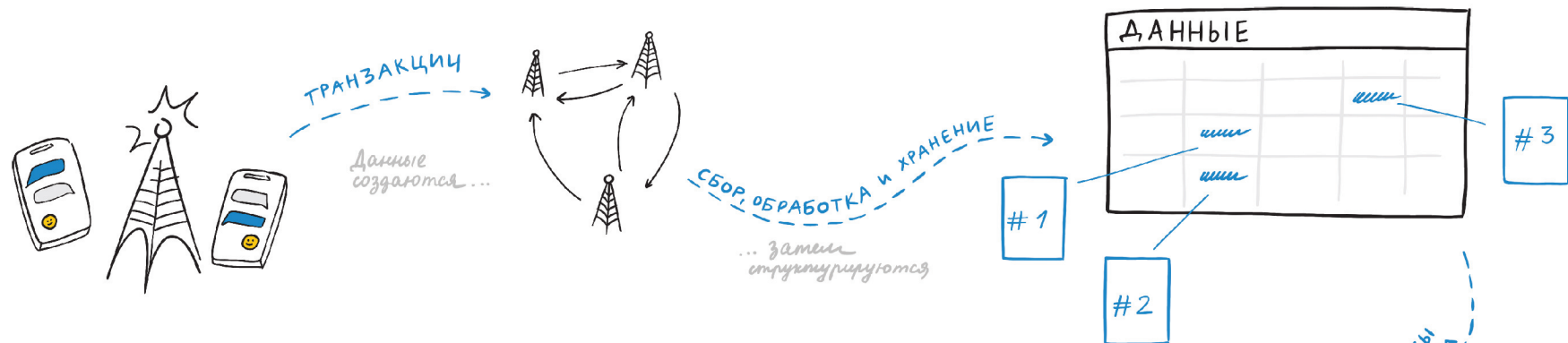
dataliteracy.ru/datatest



Георгий Нанейшвили,
*директор по развитию партнерской сети Qlik СНГ,
автор книги «Оптимизируй ЭТО немедленно!»*

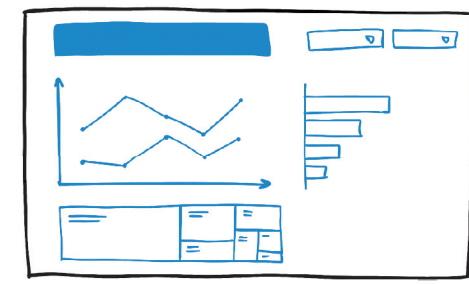
Дата-грамотность среди сотрудников — это последняя миля в работе с аналитикой. Представьте, что вы вложили серьезные ресурсы в проект по созданию аналитического решения, добились полноценной и достоверной информации, грамотно визуализировали показатели, нужные для принятия решений, но три четверти сотрудников компании просто не готовы

работать с данными и аналитикой. Это — как минимум стагнация для любого бизнеса, особенно в современной бизнес-среде. Поэтому для предпринимателей формирование корпоративной аналитической культуры должно стать целью № 1, которая будет поддерживать рост бизнеса.



ВИЗУАЛИЗАЦИЯ

А теперь всё на одной дашборде!



Путь данных

Посмотрим на то, как собираемые в огромном количестве данные из нашей повседневной жизни обретают ценность.

Вот наглядный пример, знакомый каждому: данные вашего мобильного телефона. Каждый день мы пользуемся смартфонами и планшетами, совершая звонки, отправляя и получая СМС (все еще!) и используя мобильный интернет.

Каждое наше действие фиксируется в специальных системах хранения данных: кто, кому, когда, из какой точки и в какую точку на карте звонил, сколько времени проговорил и сколько мегабайтов интернета потратил.

Данные собираются в том виде, в котором они генерируются устройствами передачи связи, в техническом формате, совсем неудобном для восприятия человека. Это огромные таблицы данных с большим количеством метрик.

Если компания (в нашем случае — сотовый оператор связи) решает использовать эти данные для проведения аналитики и принятия решений любого уровня — от операционных до стратегических, то ей нужно продолжить цепочку данных дальше. Начать обрабатывать «сырые» таблицы с помощью специальных процедур и правил,

формируя курируемые источники данных.

На этом этапе данные фильтруются — компания оставляет только те поля, которые ей нужны. Некоторые поля могут пройти через правила трансформации. «Сырые» данные преобразовывают в более агрегированный источник, которым удобно управлять с точки зрения хранения данных. Агрегированные источники уже используются разработчиками ПО и аналитиками — к ним удобно обращаться.

На подготовленном источнике данных заинтересованные сотрудники уже могут задавать

вопросы к данным. Ведь данные сами по себе не имеют никакого смысла. Они должны помогать нам в оценке обстановки, принятии решений, разрешении проблем или постановке целей.



МНЕНИЕ

Как развивать компетенции внутри компании?

Татьяна Хорошева,

*руководитель направления развития компетенций по анализу данных,
«Газпромнефть — Региональные продажи»*

Мы в компании используем единую систему, включающую в себя перечень общих и профессиональных компетенций, их индикаторов, позволяющих оценить уровень развития каждой компетенции.

Использование модели компетенций дает возможность спрогнозировать программу обучения, также применяется она при оценке эффективности проведенного обучения, корректировке планов развития сотрудников, оценке соответствия кандидата планируемому повышению, что, в свою очередь, позволяет сформировать прозрачную систему мотивации персонала.

Узнайте ваш уровень владения инструментом Qlik Sense, чтобы найти свои слабые и сильные стороны

dataliteracy.ru/qlik-check

Модель аналитических компетенций состоит из общих и профессиональных компетенций.

Data Literacy — все сотрудники.

- Жизненный цикл аналитики данных.
- Качество данных.
- Основы анализа данных.
- Основы машинного обучения.

— Визуализация и интерпретация данных.

Data Science — профессиональные компетенции продвинутого анализа данных.

- Подготовка и организация данных.
- Основы разработки программного обеспечения.
- Математические основы анализа данных.
- Методы оптимизации и имитационного моделирования.
- Машинное обучение и анализ данных.

какую выгоду приносят данные

За последние несколько лет и коммерческие, и общественные организации сделали огромные инвестиции в инфраструктуру работы с данными, настраивая и закупая серверы, оборудование, информационные системы и программы. И все это в надежде извлечь максимальную пользу из имеющихся в компании данных.

Данные обещают нам понимание паттернов и тенденций, которые окружают нас в нашей личной и профессиональной жизни. Мы можем использовать эти знания для принятия решений и для достижения целей.

Но любые выгоды не будут иметь смысла, если человек не станет пользоваться своими навыками и опытом в работе с данными.

Несмотря на фокус на работе с данными, немногие люди в организациях могут сказать, что они делают все правильно и уверенно. Данные могут стать основой для роста, привести нас к лучшим результатам в жизни или... утаить весь прогресс и скрыть возможности. Данные — это инструмент, и нам придется выбирать, насколько осознанно мы будем его использовать.

ОТ ДАННЫХ К МУДРОСТИ



Концепция пирамиды DIKW. Чтобы прийти к мудрости, нужно пройти путь извлечения выгоды из данных через получение информации и формирование знаний.

От сырых данных к мудрости, которая помогает нам делать выводы и принимать решения. Навыки грамотной работы с данными нужны и для того, чтобы понимать, какие данные могут быть полезны вообще, и уметь преобразовывать их от простого набора цифр в ценность.

Нам зачастую не нужны все данные, которые собираются и становятся доступны нам в различных системах хранения. Не все из них можно трактовать с пользой. Некоторые просто создают шум, из которого нельзя получить ничего полезного.

- **Исходные данные**

Это пока еще только потенциально полезная информация. Не все данные имеют ценность. Но если вы сможете предварительно «просканировать» исходные данные и понять их потенциальную пользу, то данные можно собрать и проверить на качество — узнать, пригодны ли они.

- **Информация**

Данные, которые имеют цель и значение. Изучаем взаимосвязи между элементами исходных данных, например с помощью визуализации.

- **Знание**

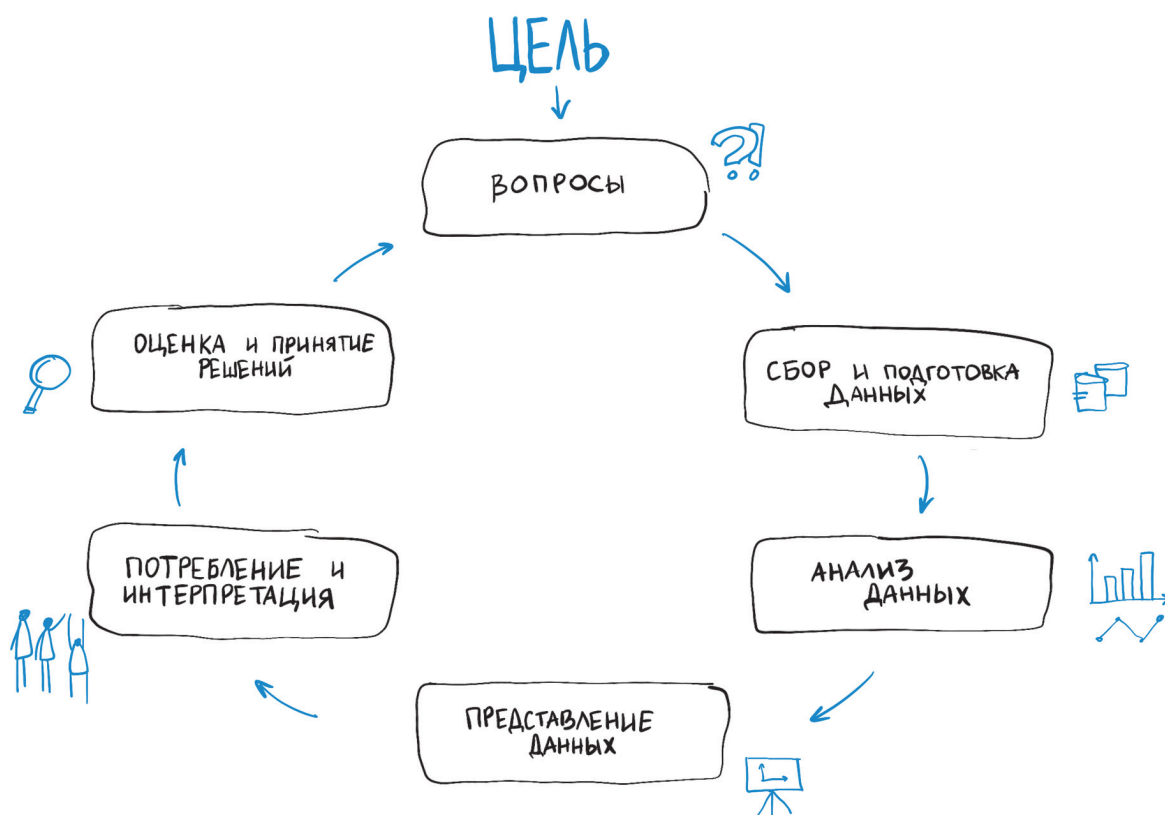
Информация, которая была собрана и обработана. С одной стороны, это синтез всей имеющейся информации с целью получения результата. С другой — знание индивидуально. Знание можно передать кому-то только в виде информации.

- **Мудрость**

Знание, трансформирующееся в принятие обоснованного решения. Здесь задаются вопросы «Что происходит?» и «Почему это происходит?».

Давайте на примере. Представьте, что вы пилот самолета. Вдруг на приборной панели загорается надпись: «10 000». О чем вы подумаете? Сейчас для вас эта цифра — просто данные. Вы пока не знаете, как их применить к своему опыту управления самолетом. Затем появляется надпись: «над уровнем моря». Окей, теперь у этих данных есть значение и цель — они показывают высоту полета. Это стало для вас информацией. Из своего прошлого опыта вы знаете, что впереди на вашем пути есть большая-большая вершина — гора Олимп высотой в 12 000 метров. И, как успешный мудрый пилот, вы понимаете, что самолету нужно подняться еще выше, чтобы долететь до конечного пункта маршрута в целости.

процесс принятия решений



До того как превратиться в мудрость, на физическом уровне данные проходят шесть больших этапов: от непосредственного создания данных и источников, в которых они хранятся, до конечного потребления и принятия решений.

Путь циклический — начинается с цели и вопросов, которые у вас есть, а заканчивается принятием решений и постановкой уже новых вопросов.

Цель и вопросы

Умения и личные компетенции человека, сформированные на основе имеющегося личного и профессионального опыта.

Сбор и подготовка данных

От Excel и Google Sheets до корпоративных хранилищ данных. На физических серверах в дата-офисе компании до аренды места в облачных хранилищах. В базах данных, платных и бесплатных. Инструменты для очистки и преобразования данных. ETL-инструменты. Самостоятельные скрипты и алгоритмы. Макросы в Excel. Ручная обработка.

Анализ данных

Снова Excel и формулы. Языки программирования. Статистические методы. Программы для статисти-

ческого анализа. BI-платформы для проведения визуальной аналитики.

Представление данных

Инструменты визуализации данных и BI-платформы. Онлайн-сервисы для создания диаграмм и графиков.

Потребление и интерпретация

Интерактивные дашборды и приложения в инструментах визуализации, отчеты, презентации, встречи. Порталы для распространения информации. Презентации. Инфографика.

Оценка и принятие решений

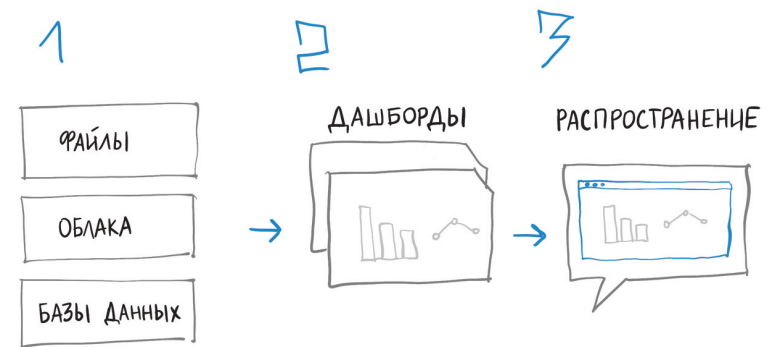
Только личные качества и накопленный опыт. Здесь необходимо понимание предметной области и контекста решаемой задачи.

про self-service в организациях

Скажем очевидную вещь: с ростом количества собираемых данных компаниям хочется видеть и рост их использования, анализа и принимаемых на их основе решений.

При традиционном подходе к аналитике пользователи достаточно ограничены в своих возможностях, так как подготовку отчетов на основе нужных данных выполняет IT-служба. В то же время IT-специалисты оторваны от бизнеса и не знают всех тонкостей работы с показателями конкретного бизнес-подразделения. А владельцы бизнес-данных, прекрасно понимая свои внутренние потребности, не могут разобраться в сложных механизмах работы с инструментами анализа данных.

Поэтому, вместо того чтобы возлагать функции анализа данных на небольшое количество очень дорогих и занятых экспертов IT, компании сегодня стремятся предоставлять возможности самостоятельного анализа бизнес-сотрудникам, используя практику self-service.



При self-service-подходе бизнес-специалисты получают в свои руки понятные инструменты прямого общения с данными. Самое главное, что эти люди являются экспертами в своих профессиональных областях и точно знают, какие вопросы им нужно задать. Решения будут приниматься оперативнее. Более того, такие сотрудники будут быстро получать ответы на свои вопросы и делиться ими с коллегами, популяризируя культуру работы с данными в организации в целом.

корпоративные марафоны

Марафоны DataYoga — это формат обучения визуализации данных через коммуникацию с широкой аудиторией. Обычно в марафонах участвует сразу большое количество человек, поэтому это хороший шанс для повышения популярности тем анализа и визуализации данных в сообществе, в том числе корпоративном.

Основу корпоративных марафонов составляют внутренние истории успеха компании: кто и каких результатов достиг благодаря визуальному анализу данных. Во время марафона его участники общаются в Telegram-чате, формируя внутреннее сообщество и первых экспертов. Более того,

так как все необходимые материалы, ссылки и ресурсы собираются в одном месте, то марафон одновременно строит и фундамент внутренней базы знаний.

DataYoga успешно провела несколько открытых марафонов по работе с различными BI-инструментами, а также несколько корпоративных марафонов в таких компаниях, как «Вымпелком», «Ростелеком» и «Почта России».

Марафоны DataYoga

<https://datayoga.ru/marathons>