

# Платформа подготовки данных Trifacta

# Платформа подготовки данных Trifacta

Рейтинг №1 пользователями и аналитиками



Самый высокий рейтинг поставщика подготовки данных среди пользователей в G2



Поставщик услуг по подготовке данных №1 от Dresner Advisory Services 6 лет подряд



Оценка 4,6 из 5 по версии Gartner Peer Insights для инструментов подготовки данных

# Одна из самых продвинутых платформ для специалистов по обработке данных для оценки качества, преобразования и автоматизации конвейеров данных

## Обеспечение качества данных

Низкое качество данных может погубить любой аналитический проект, нанеся ущерб авторитету и доверию.

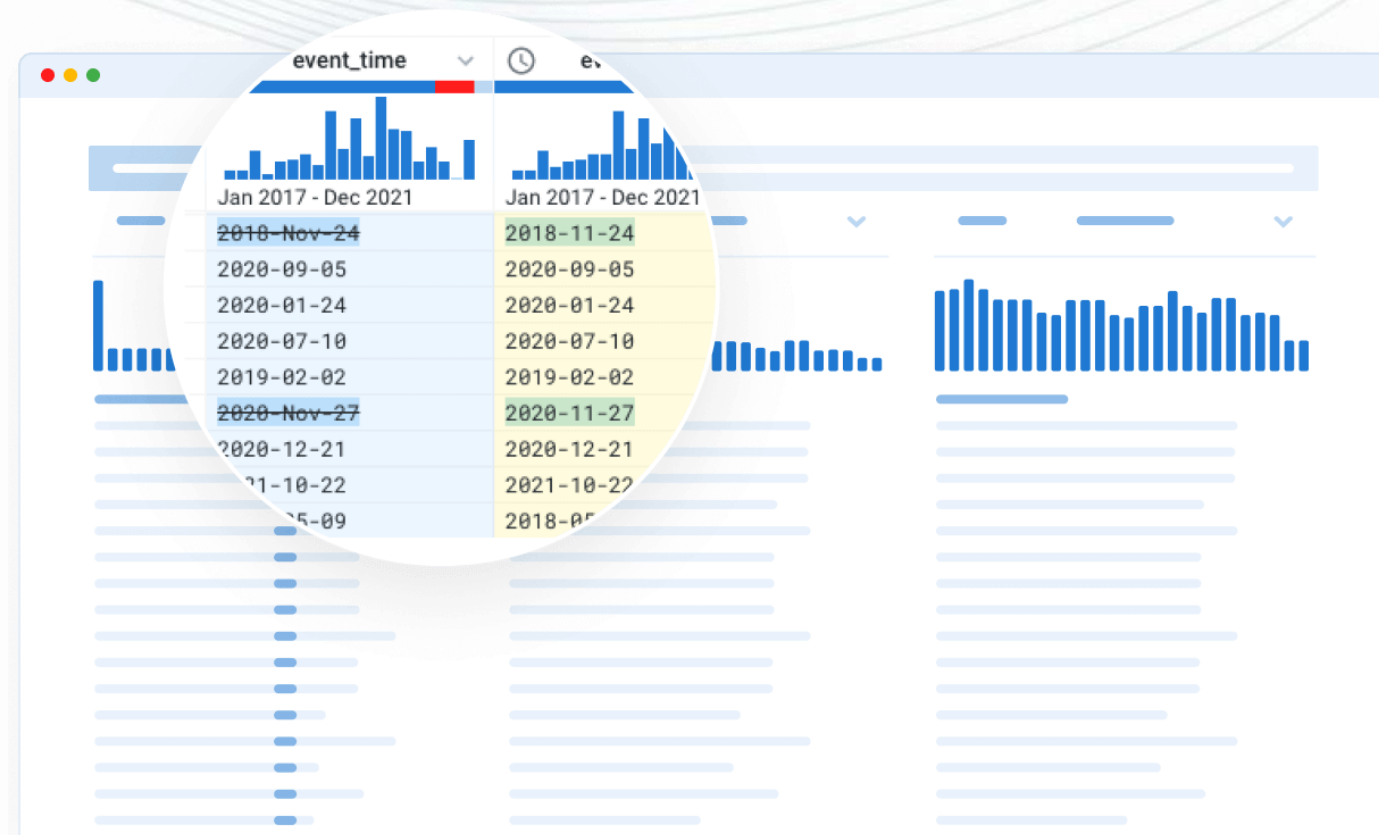
Только Trifacta сочетает визуальное взаимодействие с подсказками искусственного интеллекта для обеспечения беспрецедентного качества данных. Легко исправляйте выбросы и ошибки, а также автоматизируйте мониторинг качества, чтобы на ваши аналитические платформы не попадали неверные данные.



# Визуализировать и ускорить преобразование данных

Преобразования ручного кодирования могут работать, но они медленные, подвержены ошибкам и не масштабируются.

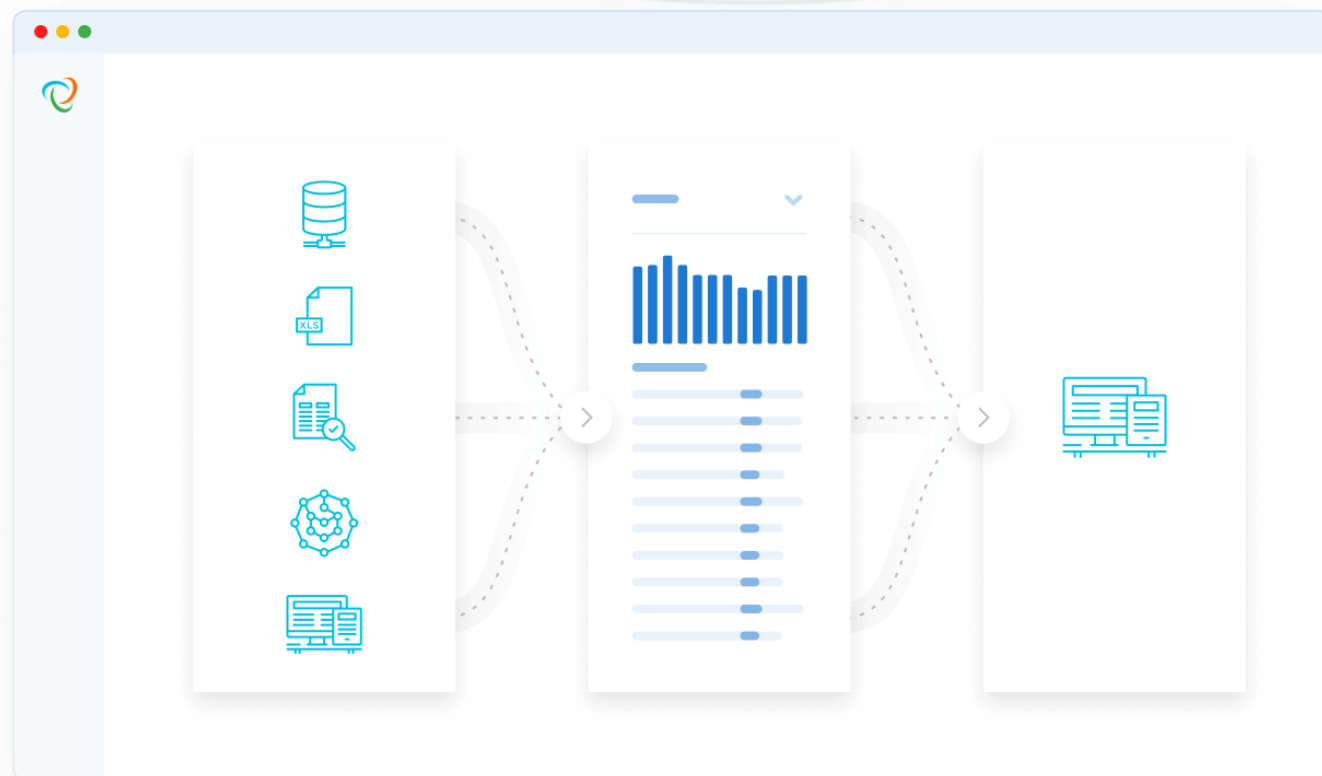
Trifacta предоставляет интеллектуальные визуальные инструкции, позволяющие прогнозировать каждое преобразование в режиме реального времени. Он без проблем поддерживает код, когда он вам нужен, и без кода, когда он вам не нужен.



# Автоматизируйте конвейеры данных

Ручные, утомительные процессы подготовки данных не масштабируются.

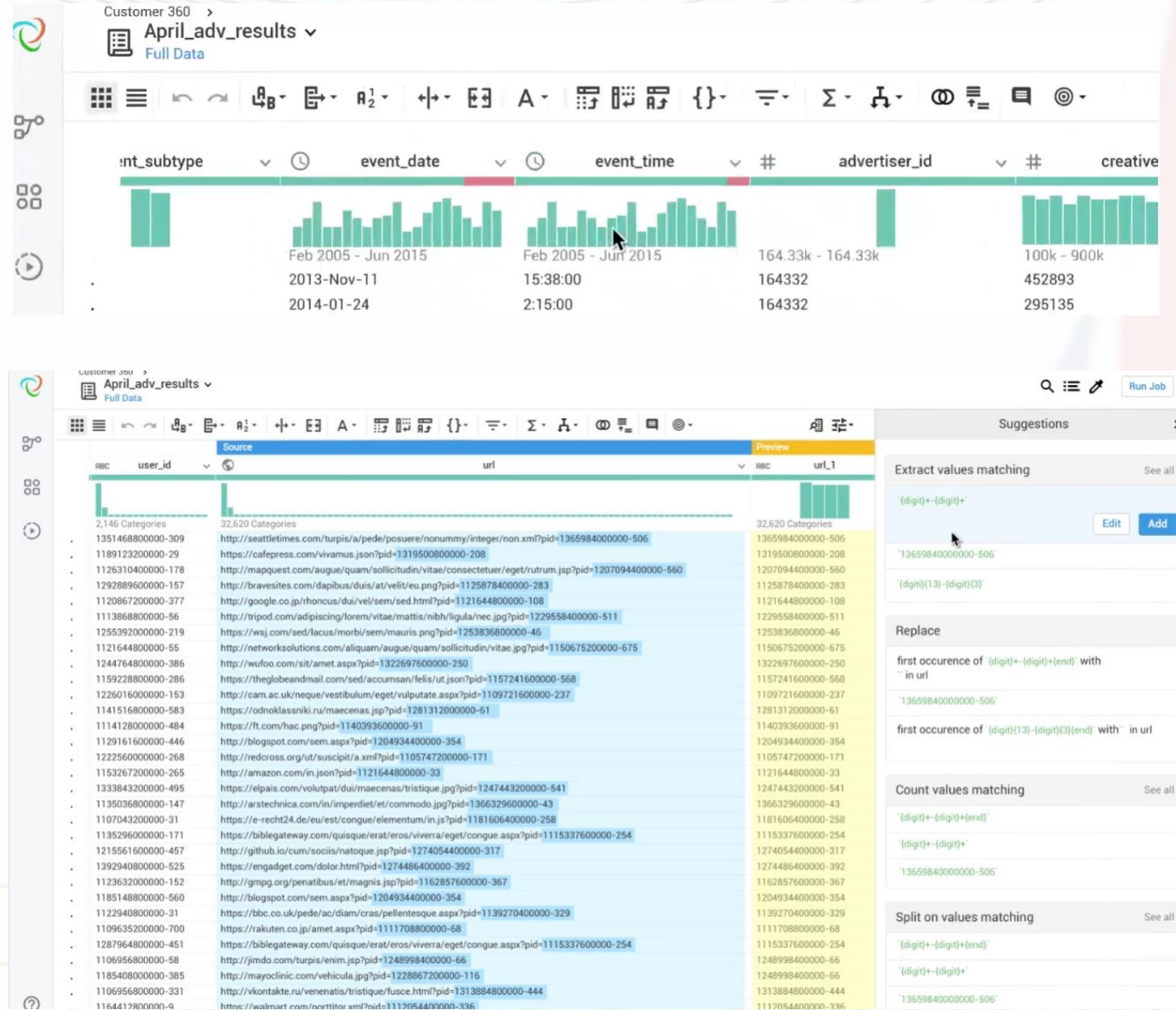
Trifacta помогает создавать, развертывать автоматизированные конвейеры данных и управлять ими за считанные минуты, а не месяцы.





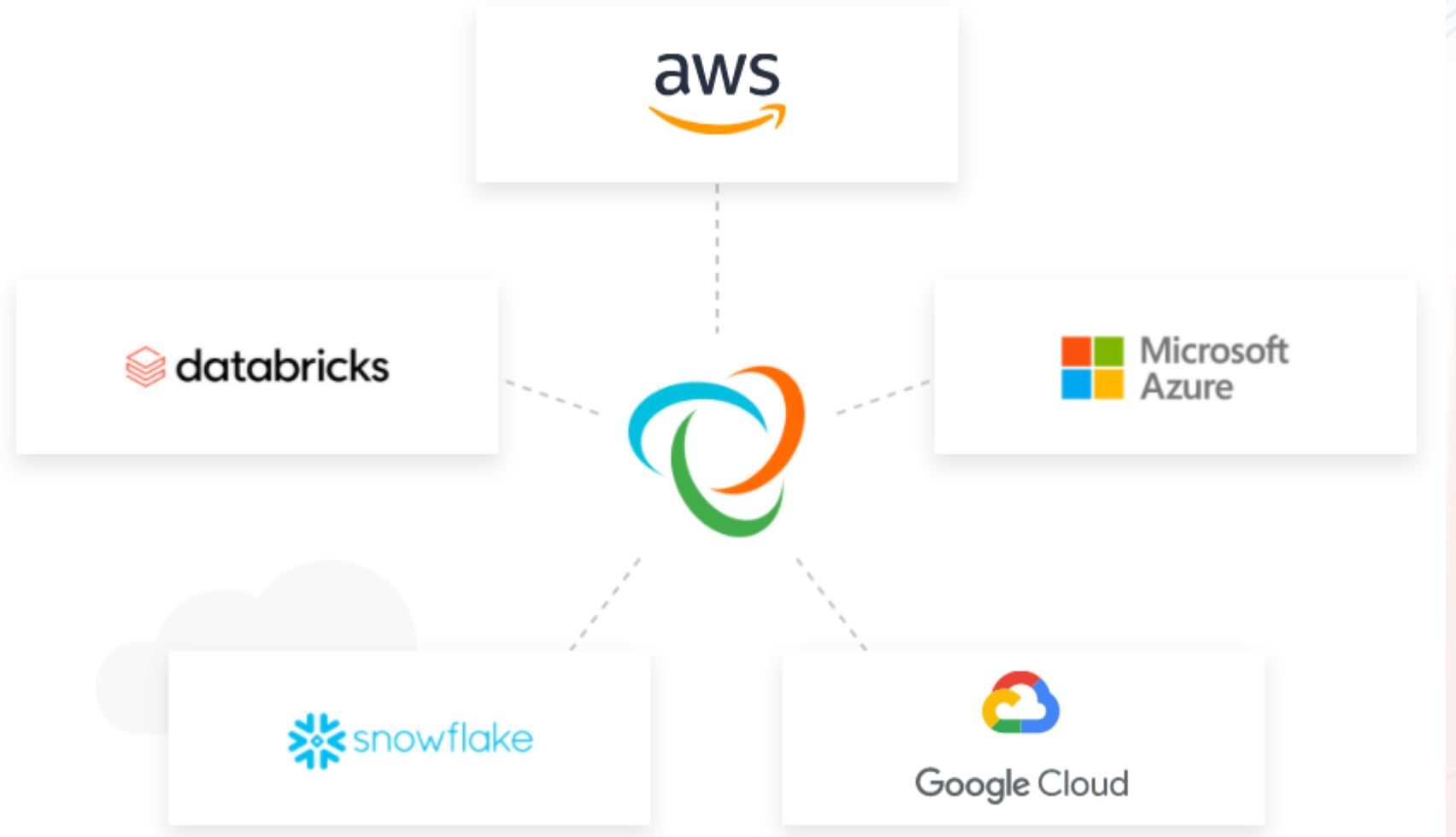
# Обеспечение качества данных

Успешные инженеры по обработке данных полагаются на точные, хорошо структурированные данные, отформатированные для решения конкретной задачи. Тем не менее, сегодняшние данные более беспорядочные и сложные, чем когда-либо прежде. Преобразовать его в формат для анализа - трудоемкий и технически сложный процесс. Инжиниринг данных - это процесс, который вы должны пройти для преобразования исходных входных данных из различных источников в подготовленные выходные данные, которые будут использоваться в анализе, отчетности и науке о данных.



## Безупречная работа в любом облаке

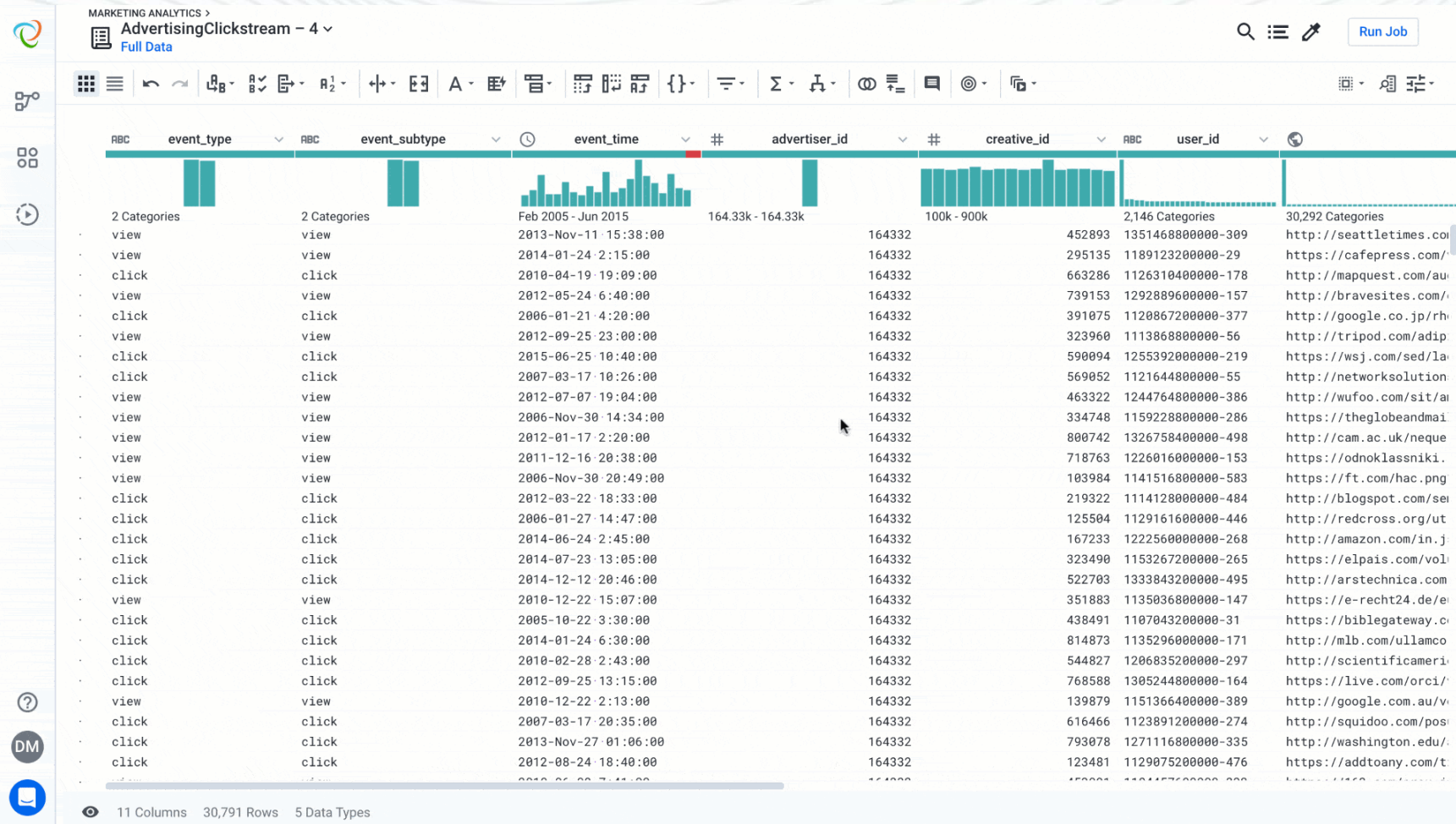
Безупречная инженерия данных в любой облачной, гибридной или мультиоблачной среде.



Trifacta делает любые данные понятными и полезными для анализа, независимо от того, где они хранятся и как потребляются. Партнерские отношения на высшем уровне с ведущими поставщиками облачных услуг, включая Amazon Web Services, Microsoft Azure, Google Cloud, Snowflake и Databricks, гарантируют, что мы можем поддерживать рабочие нагрузки по подготовке данных в любой среде

# Обеспечение качества данных

Trifacta представляет автоматизированные визуальные представления данных на основе их содержания в наиболее привлекательном визуальном профиле. Кроме того, каждый профиль полностью интерактивен, что позволяет пользователю просто выбрать определенные элементы профиля, чтобы запросить предложения по преобразованию.

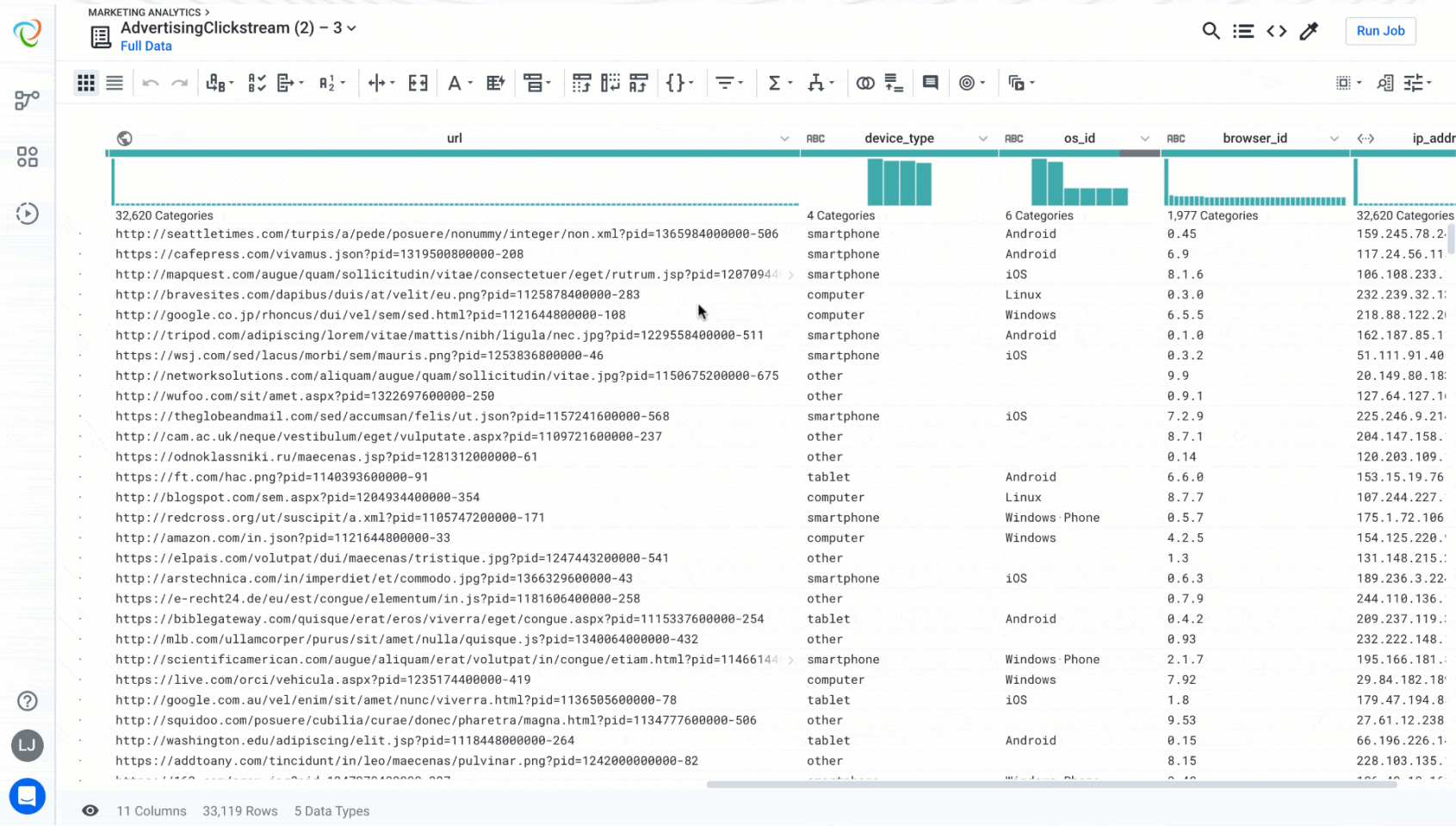


Подробные визуальные представления позволяют более глубоко исследовать данные, обеспечивая автоматическое понимание данных на самом детальном уровне. Выбросы и недопустимые данные легко выявлять и исправлять, что предотвращает влияние ошибочных данных на последующую аналитику, а постоянный мониторинг и управление качеством данных означает, что на ваши аналитические платформы не попадут плохие данные.



# Визуализировать и ускорить преобразование данных

Каждый клик, перетаскивание или выбор в Trifacta приводит к предсказанию - система интеллектуально оценивает имеющиеся данные, чтобы рекомендовать ранжированный список предлагаемых преобразований с предварительным просмотром каждого преобразования в реальном времени для оценки или редактирования пользователями.

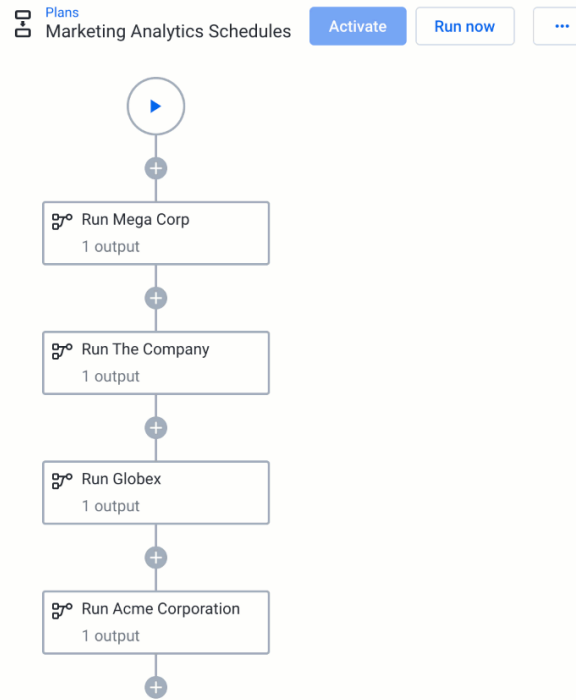


В отличие от любого другого продукта для инженерии данных, Trifacta дает пользователям возможность делать больше со своими данными, постоянно направляя их через процесс, используя интеллектуальные предложения на основе машинного обучения.

## Автоматизировать конвейеры данных

Ручные повторяющиеся процессы инженерии данных не масштабируются. В то же время большинство инструментов автоматизации предназначены для ИТ-аудитории с высокотехнологичным образованием.

Trifacta дает вам возможность создавать конвейеры проектирования данных в масштабе с помощью щелчков мыши, а не кода, и помогает развертывать конвейеры данных самообслуживания и управлять ими за считанные минуты, а не месяцы.



# Отчетность и аналитика

Сократите время на отправку данных для аналитики до 80%

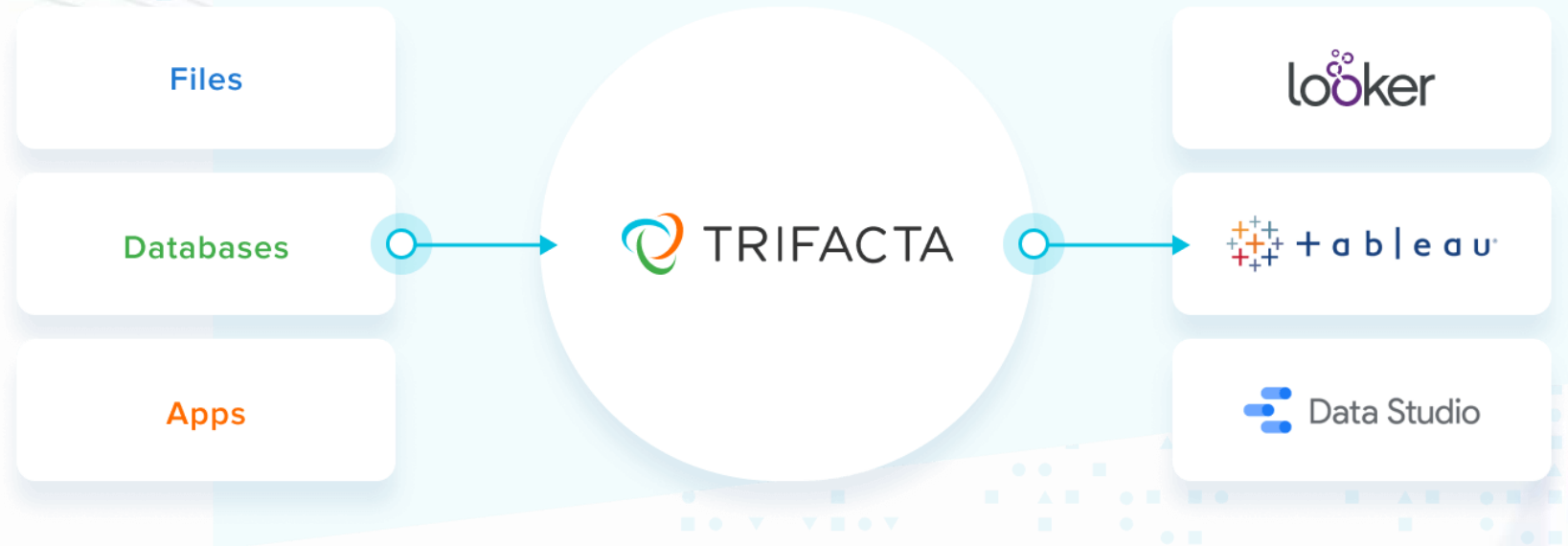


Поддерживаемые услуги



# Отчетность и аналитика

Сократите время на отправку данных для аналитики до 80%



Более 80% аналитических проектов тратится на подготовку данных для аналитики. Современная аналитика требует емкости, масштабирования и автоматизации, а также предоставляет интерфейс для профессионалов в области данных - тех, кто лучше всего разбирается в данных - для изучения, взаимодействия и преобразования данных для анализа. Сочетая интерактивную визуализацию с пользовательским интерфейсом без кода, Trifacta делает обнаружение, смешивание, фильтрацию, очистку и агрегирование данных быстрым и интуитивно понятным. После завершения подготовительной работы автоматизируйте потоки подготовки данных, используя эластичное масштабирование облака.



## Создавайте отчеты быстрее

Самое большое препятствие на пути к успеху аналитики - это время. Процессы подготовки данных вручную могут помешать специалистам по обработке данных использовать свой опыт для поиска новых идей. Trifacta сокращает время подготовки с часов до минут и может автоматизировать конвейеры аналитики для получения повторяемой ценности.

## Обеспечьте высокое качество данных

Низкое качество данных может привести к тому, что организации будут принимать решения на основе ложной или неточной информации. Trifacta позволяет быстрее обнаруживать и устранять проблемы с качеством данных. Trifacta дает вам уверенность в точности ваших данных и информации, которую они используют.

## Используйте новые источники данных

Trifacta подключается к различным файловым системам, хранилищам данных, приложениям и рынкам данных, чтобы предоставить пользователям доступ к данным в необработанном виде. Возможность подключаться к данным из множества различных источников открывает возможность для поиска новых идей и принятия более эффективных решений.

## Избавьтесь от зависимости от ИТ

Слишком часто выполнение аналитических проектов замедляется из-за ручной работы, которую можно автоматизировать с помощью правильных инструментов. Но часто эти инструменты автоматизации предназначены для ИТ-специалистов или высокотехнологичных пользователей. Trifacta предлагает конвейеры данных самообслуживания для автоматизированных аналитических процессов для получения повторяемой ценности.

# Машинное обучение и наука о данных

Создавайте высококачественные модели машинного обучения для своих проектов в области науки о данных



Поддерживаемые услуги



# Машинное обучение и наука о данных

Создавайте высококачественные модели машинного обучения для своих проектов в области науки о данных



Trifacta устраняет утомительные и трудоемкие задачи, связанные с задачами подготовки данных для науки о данных, такие как структурирование неструктурированного текста, быстрое кодирование, масштабирование, стандартизация и нормализация данных. С Trifacta у специалистов по обработке данных есть инструменты для создания согласованных, высококачественных, дифференцированных данных для моделей искусственного интеллекта и машинного обучения. Визуальные подсказки и встроенное машинное обучение Trifacta создают интерфейс, ориентированный на простоту использования, мгновенную проверку и эффективные преобразования. Эти функции позволяют специалистам по обработке данных выполнять необходимые преобразования и проверки качества данных без необходимости отладки, ускоряя развертывание моделей и повышая точность поддерживаемых платформ обработки данных.

## Извлечение и создание функций с легкостью

Легко инжинируйте функции в ваших данных путем структурирования текста, извлечения важных значений, быстрого кодирования, масштабирования, стандартизации и нормализации данных. Trifacta упрощает понимание данных, с которыми вы работаете, и построение на их основе функций.

## Мониторинг и оценка качества данных

Следите за качеством данных и изменениями с течением времени, чтобы не автоматизировать принятие решений по неверным данным. Настройте предупреждения на основе пороговых значений качества данных, чтобы постоянно устранять любые аномалии, и создавайте правила качества данных для управления запуском заданий и создаваемыми выходными данными.

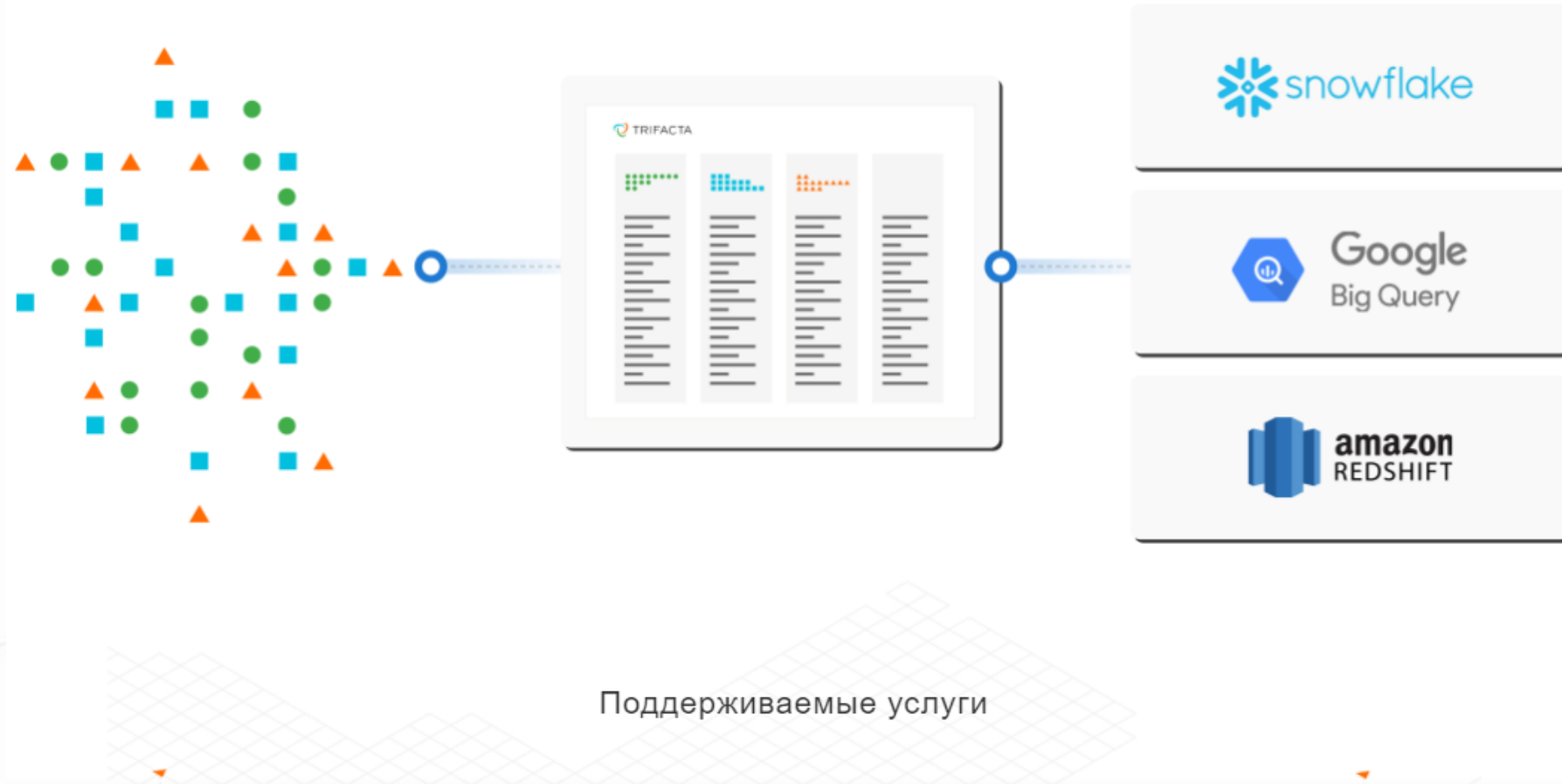
## Быстрое создание и развертывание конвейеров

Автоматизируйте рабочие процессы машинного обучения, чтобы ваши модели постоянно пополнялись новыми данными. Больше не нужно отправлять свою работу в ИТ-команду для развертывания в производственной среде, Trifacta позволяет вам управлять своими собственными конвейерами и настраивать их по мере необходимости.



# Подключение к Cloud Data Warehouse

Используйте облако самообслуживания для проектирования данных Trifacta, чтобы переместить свои данные в облачное хранилище данных



## Подключение к Cloud Data Warehouse

Используйте облако самообслуживания для проектирования данных Trifacta, чтобы переместить свои данные в облачное хранилище данных



Центр тяжести аналитики перемещается в облачные хранилища данных, такие как Snowflake, BigQuery и RedShift. Это означает, что все больше и больше данных необходимо очищать, смешивать и интегрировать в эти среды для более широкого аналитического использования. Trifacta сокращает время, необходимое для переноса разнообразных данных в ваше облачное хранилище данных, позволяя аналитикам, специалистам по обработке данных и инженерам данных выполнять эту работу самостоятельно с помощью платформы самообслуживания, созданной для облака.

## Разнообразные данные на борту быстрее

Обнаруживайте закономерности, аномалии и выбросы с помощью визуального руководства. Очистите, структурируйте и объедините различные источники данных, чтобы их можно было опубликовать в вашем хранилище данных.

## Автоматизация и управление конвейерами данных

Один раз поспоришь, потребляешь непрерывно. Настройте конвейеры данных самообслуживания для каждого из ваших проектов или клиентов, чтобы легко автоматизировать и управлять всеми вашими потребностями в очистке данных.

## Мониторинг и оценка качества данных

Следите за качеством данных и изменениями с течением времени, чтобы не автоматизировать принятие решений по неверным данным. Настройте оповещения на основе пороговых значений качества данных, чтобы постоянно устранять любые аномалии.

# Dataprep от Trifacta в GCP

Интеллектуальная облачная служба данных для визуального изучения, очистки и подготовки данных для анализа и машинного обучения.

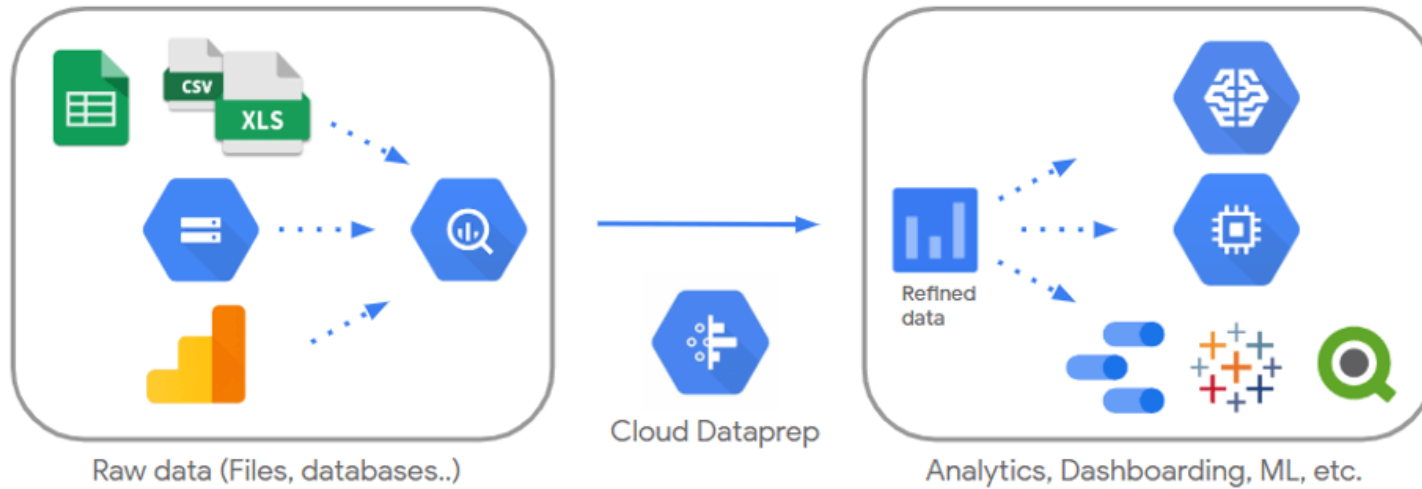
## Интеллектуальная подготовка данных

Dataprep от Trifacta - это интеллектуальная служба данных для визуального изучения, очистки и подготовки структурированных и неструктурированных данных для анализа, составления отчетов и машинного обучения. Поскольку Dataprep является бессерверным и работает в любом масштабе, нет инфраструктуры для развертывания или управления. Следующее идеальное преобразование данных предлагается и прогнозируется с каждым вводом пользовательского интерфейса, поэтому вам не нужно писать код.





# Автоматизируйте свое хранилище данных Google Cloud в BigQuery с помощью Cloud Dataprep

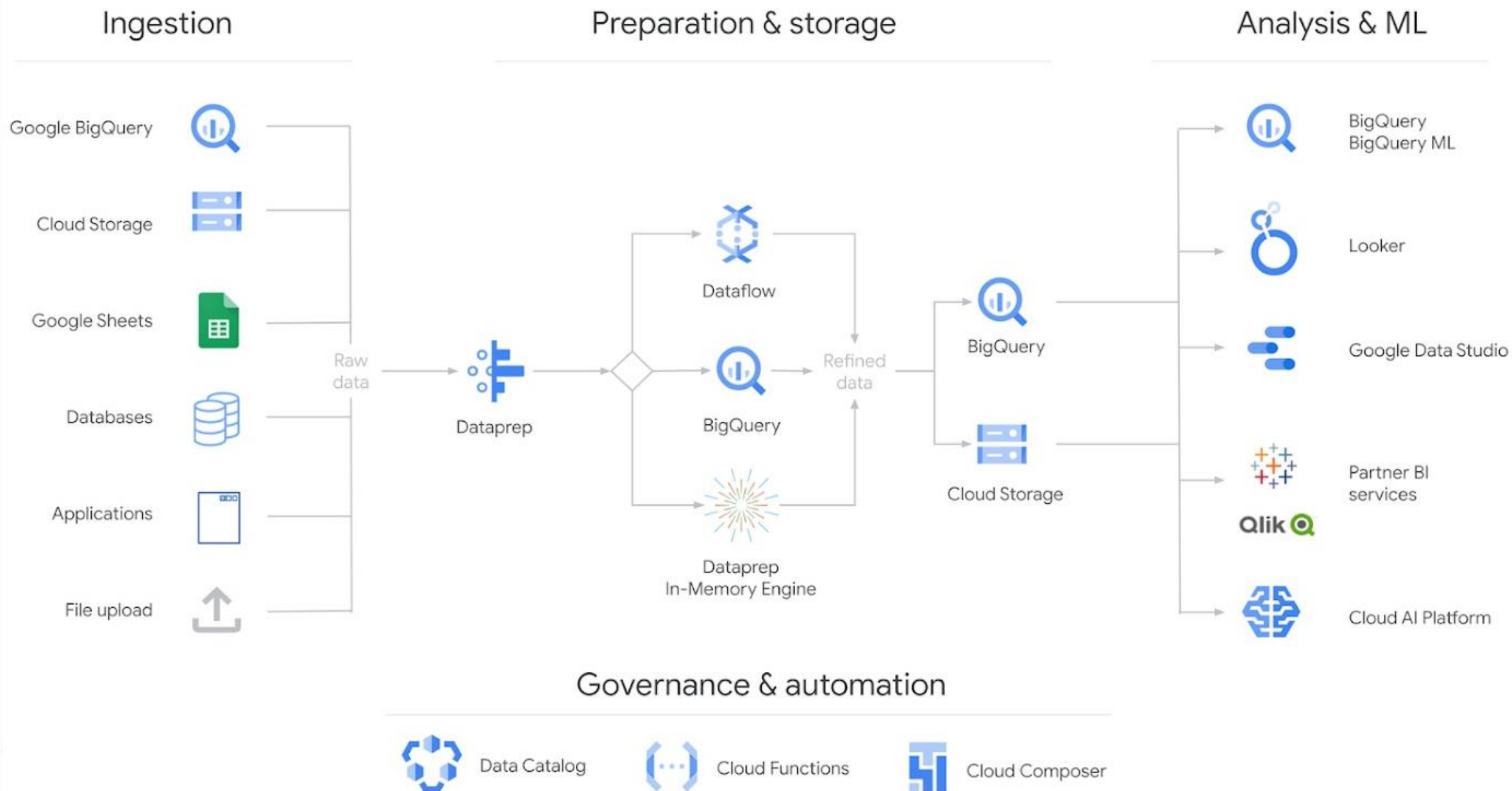


Необработанные данные относятся к вашим входным данным до того, как они прошли подготовку (структурирование, очистка, обогащение, проверка, дедупликация, объединение и т.д.). Входные данные могут поступать из различных источников, таких как файлы (таблицы Excel или Google, CSV и др.), облачные службы, такие как BigQuery (облачное хранилище данных Google), облачное хранилище Google, а также дополнительные базы данных и приложения.

Под уточненными данными понимаются подготовленные данные, которые часто находятся в вашем хранилище данных. После того, как данные были уточнены, они готовы к использованию различными способами. Вот некоторые примеры:

- для прямого анализа
- реализованы в качестве дашбордов с такими инструментами, как Data Studio, Tableau Software, Qlik и др.
- используются для создания моделей машинного обучения

# Архитектура конвейера Dataprep ELT в GCP



# Google Cloud Dataprep в 20 раз ускоряет задачи инженерии данных за счет работы внутри BigQuery

С помощью раскрывающегося меню BigQuery Google Cloud Dataprep поддерживает рабочие процессы извлечения нагрузки и преобразования (ELT), которые работают непосредственно с данными в хранилище. Используя возможности SQL, данные можно преобразовывать на месте, что делает их эффективными для таких манипуляций с данными, как фильтры, объединения, объединения и агрегирование. Dataprep автоматически определяет, когда конвейер данных может быть частично или полностью преобразован в операторы BigQuery SQL, позволяя предприятиям запускать рабочие нагрузки для анализа данных в любом масштабе с помощью Google Dataprep от Trifacta. Выполнение BigQuery обеспечивает невероятно высокую производительность и полную гибкость при оптимизации затрат.

## Расширение маркетинговой аналитики

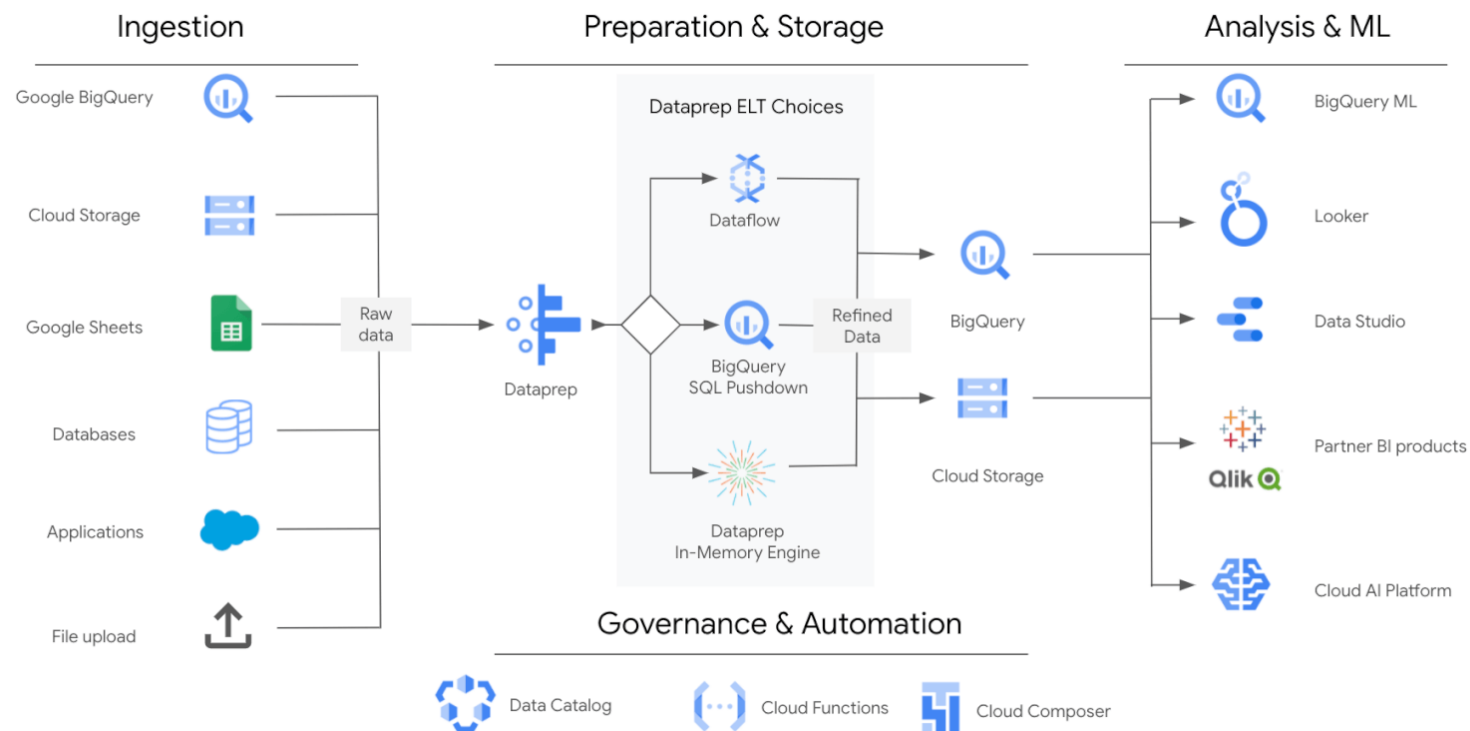
Объедините различные источники маркетинговых данных из различных сервисов Google, таких как Google Analytics и Double Click, чтобы получить более **целостное представление о ваших маркетинговых кампаниях**.

## Расширенное внедрение облачных хранилищ данных

Повысьте качество и ценность данных, доступных в BigQuery, с помощью функций автоматической очистки и преобразования данных Dataprep.

## Усовершенствуйте свое облачное озеро данных

Предоставьте аналитикам, специалистам по обработке данных и инженерам **доступ к необработанным данным и их уточнение** в облачном хранилище Google на платформе с централизованным управлением.



## Supported Services

- Cloud Dataflow
- BigQuery ML
- Looker
- Cloud ML Engine
- Google Cloud Storage
- Google BigQuery
- Google Data Studio
- Google AutoML



## Ускорение внедрения озера данных

Trifacta предотвращает превращение озера данных Amazon S3 в болото данных с помощью самостоятельной подготовки данных на основе машинного обучения, интегрированной с AWS, чтобы ускорить внедрение озера данных и улучшить услуги обработки



AWS Cloud

## Ускоренная бизнес-аналитика

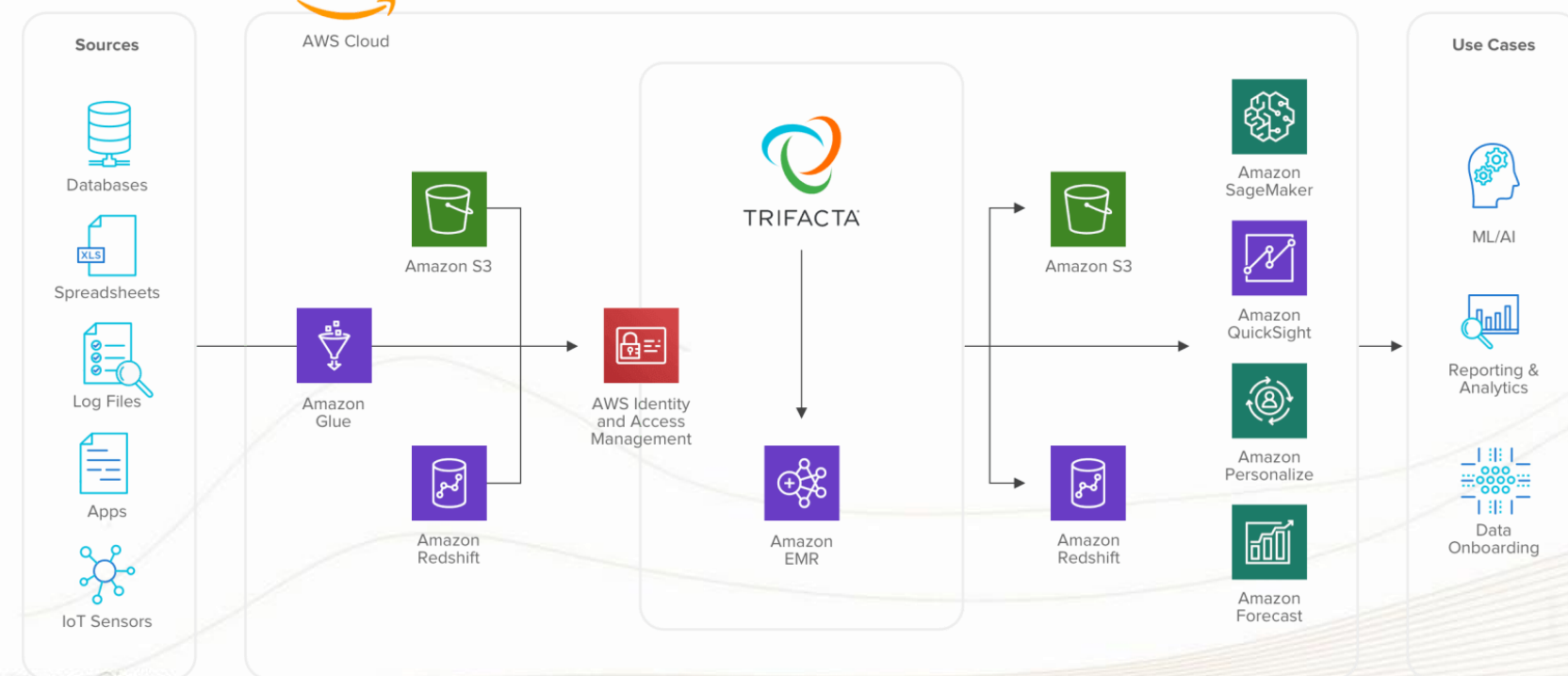
Trifacta быстро преобразует беспорядочные данные в вашем хранилище данных Amazon, например Amazon Redshift, в чистые и надежные данные, позволяя вашим аналитикам сразу же приступить к созданию отчетов бизнес-аналитики и другим проектам расширенной аналитики.

## Стимулируйте успех AI / ML

Trifacta автоматизирует подготовку данных для искусственного интеллекта и машинного обучения на AWS, освобождая время для специалистов по обработке данных, чтобы они могли сосредоточиться на разработке и обучении моделей, а не на обработке данных вручную.

## Supported Services

- Amazon Redshift
- Amazon Glue
- Amazon Aurora
- Amazon SageMaker
- Amazon IAM
- Amazon S3
- Amazon EMR
- Amazon Athena
- Amazon RDS
- Amazon QuickSight



## Усовершенствуйте свое озеро данных Azure

Тесно интегрированная с экосистемой Azure, Trifacta автоматизирует профилирование, смешивание и совместную работу данных в вашем озере данных Microsoft Azure, чтобы обеспечить успешное внедрение и аналитику.

## Ускорение анализа

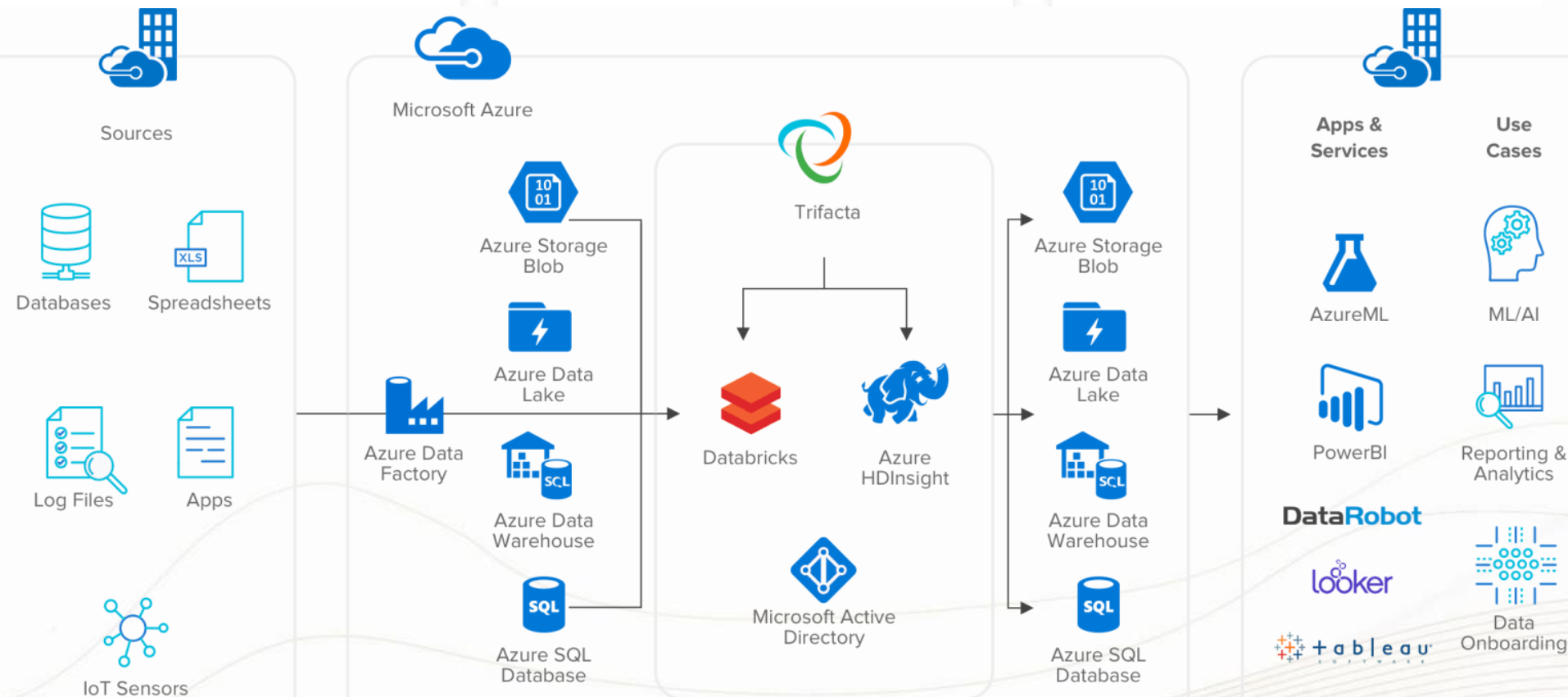
Trifacta доставляет чистые, хорошо подготовленные данные в хранилище SQL Azure, предоставляя интерактивное решение для подготовки данных, которое может использовать любой аналитик для ускорения создания отчетов бизнес-аналитики.

## Улучшение результатов AI / ML

Сократите время обработки данных для ваших проектов AI / ML с помощью простого в использовании решения для подготовки данных в Azure, которое обеспечивает более быстрые и точные прогнозы.

## Supported Services

- Azure Data Lake
- Azure SQL Data Warehouse
- Azure HDInsights
- Azure Data Factory
- Azure Storage Blob
- Azure Databricks
- Azure Active Directory



## Разнообразные данные на борту быстрее

Обнаруживайте аномалии, очищайте беспорядочные данные и объединяйте разрозненные источники данных перед публикацией в хранилище данных Snowflake для аналитики.

## Модернизируйте отчетность и аналитику

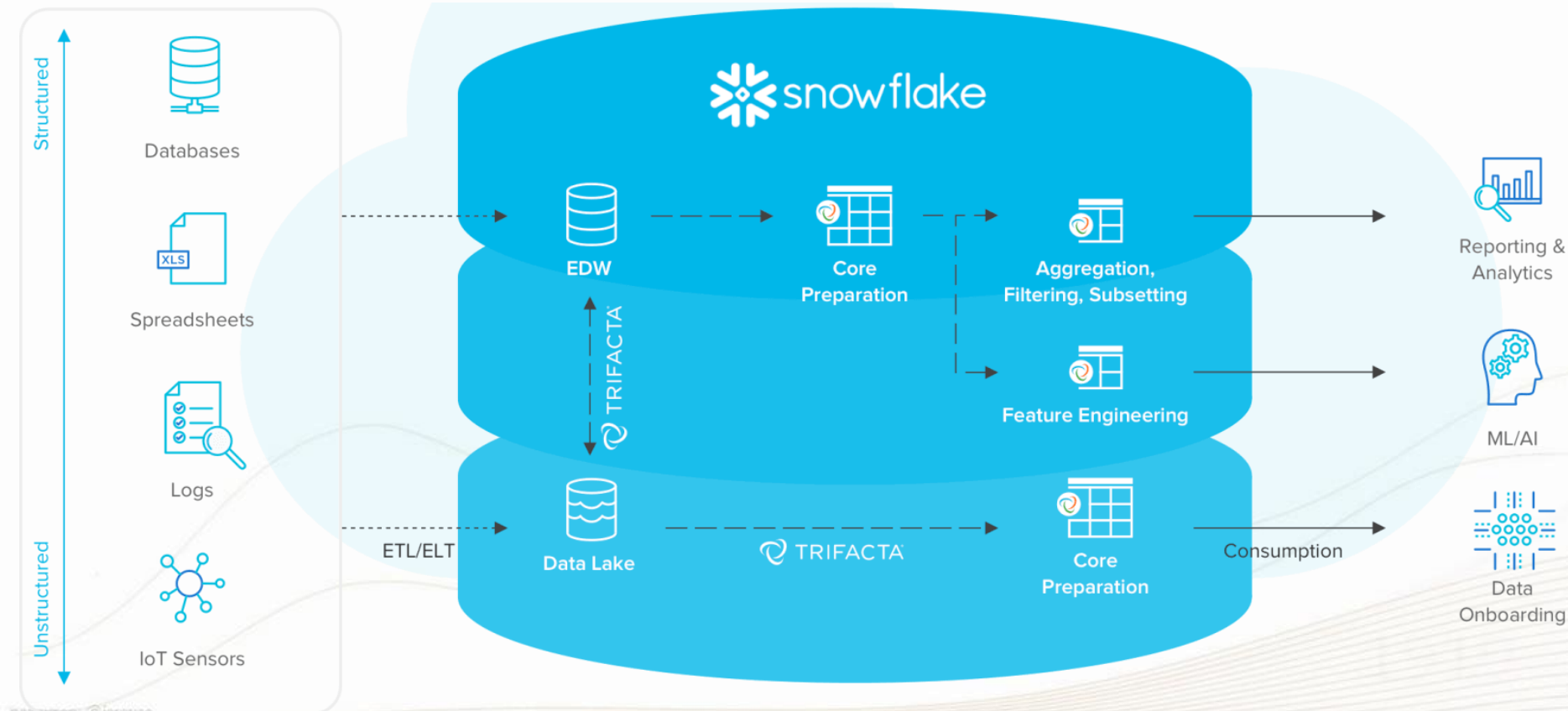
Используйте визуальное руководство и руководство на основе машинного обучения для расширения возможностей различных пользователей, а также гибкое масштабирование и автоматизацию современных облачных платформ.

## Ускорение инициатив ML / AI

Повысьте скорость и качество общих задач подготовки данных для машинного обучения, таких как проектирование функций, стандартизация атрибутов и быстрое кодирование.

## Поддерживаемые услуги

Trifacta обеспечивает поддержку ряда встроенных функций Snowflake, таких как безопасность, контроль доступа и шифрование, а также сервисов облачной платформы, необходимых для развертывания Snowflake на AWS, Azure и Google Cloud.



# Trifacta анонсирует облако для обработки данных

**Открытость:** среда разработки данных должна быть гибкой, чтобы беспрепятственно работать в рамках сквозных аналитических рабочих процессов и полностью интегрироваться в современные цепочки инструментов. По замыслу, среда должна быть достаточно расширяемой, чтобы пользователи могли использовать свою магию, чтобы заполнить пробелы и освободить их от тирании монолитных стеков и привязки к поставщикам. Открытая среда разработки данных поддерживает независимость и слабую взаимосвязь, в то же время способствуя компоновке сервисов и взаимодействию, что позволяет пользователям использовать лучшие в своем классе инструменты, которые могут помочь быстрее решать сложные проблемы с данными. Чтобы поддержать это, Trifacta теперь предлагает:

- ❑ **Поддержка нескольких облаков:** облако инженерии данных Trifacta предлагает собственные решения, оптимизированные для каждой платформы: Google, AWS, Azure. Поддержка нескольких облаков означает свободу выбора и свободу от кода. Пользователи могут менять свое мнение о том, какую платформу они предпочитают, или запускать разные рабочие нагрузки в разных средах, не переписывая какой-либо код.
- ❑ **Гибкое выполнение:** пользователи могут выбирать между ETL или ELT или оптимальной комбинацией двух в зависимости от стоимости. Гибкое выполнение также означает, что пользователи могут свободно генерировать SQL, Spark, Dataflow/Beam или Python.
- ❑ **Универсальное подключение:** пользователи могут подключать любое приложение к данным из более чем 180 корпоративных источников данных - как локальных, так и в облаке - и публиковать уточненные данные в электронных таблицах, средствах бизнес-аналитики и отчетности, а также в записных книжках для анализа данных. Возможности подключения к облаку обработки данных Trifacta расширены за счет фреймворков Trifacta REST, XML и JDBC.
- ❑ **На основе API:** облако Trifacta интегрируется со всеми без исключения цепочками инструментов. С помощью SDK и стандартов OpenAPI, доступных на множестве языков, пользователи могут интегрировать Trifacta в существующие рабочие процессы или использовать Trifacta для координации работы сторонних приложений, от инструментов управления версиями, приема и репликации до каталогов и бизнес-гlossариев.



**Интеллектуальность:** современные инструменты инженерии данных должны быть интеллектуальными и учиться на самих данных и на взаимодействии пользователей, чтобы автоматизировать наиболее сложные и трудоемкие части очистки и преобразования данных, улучшая взаимодействие с пользователем и ускоряя инновации, основанные на данных. Этот интеллект также должен применяться для улучшения текущих операций с данными, обеспечения самоконтроля и, в некоторых случаях, автоматического устранения проблем, которые обычно нарушают конвейеры данных. Чтобы поддержать это, Trifacta теперь предлагает:

- ❑ Прогнозируемое преобразование. Облако для проектирования данных Trifacta имеет визуальный интерфейс «направляй и решай», который использует машинное обучение, чтобы сделать понимание и решение проблем преобразования данных интуитивно понятным для пользователей любого уровня подготовки, независимо от их технической проницательности. Возможности прогнозного преобразования включают: автоматическое обнаружение и применение формата к неструктурированным и полуструктурированным наборам данных; использование примеров для вывода логики преобразования; синтез моделей данных из исходных данных; и автоматическое сопоставление данных с заранее определенной целью.
- ❑ Адаптивное качество данных: активное профилирование данных Trifacta выходит за рамки традиционных правил качества данных. Теперь пользователям стало проще обнаруживать и проверять проблемы с качеством данных. Профили статистических данных используются для выявления сложных шаблонов, автоматически предлагающих возможные правила качества, такие как ограничения целостности, шаблоны форматирования и зависимости столбцов. Пользователям предлагаются на рассмотрение преобразования, основанные на классификаторах для вероятностных правил качества данных, и они могут более легко стандартизировать данные с помощью поддержки сложной кластеризации.
- ❑ Интеллектуальные конвейеры данных: Trifacta позволяет пользователям моделировать потоки данных, управляя взаимосвязями между наборами данных и рецептами. Они могут вводить в действие и автоматизировать потоки данных с помощью планов, которые обеспечивают параллельное и условное выполнение, а также предварительную и постобработку. Кроме того, мониторинг тенденций качества данных на определенный момент времени и исторических данных обеспечивает контекст для упреждающего оповещения (через электронную почту, Slack, PagerDuty и другие платформы) об изменениях схемы и распределения данных, которые могут повлиять на пригодность данных для последующего использования.

**Самообслуживание:** Платформа инженерии данных Trifacta увеличивает количество и разнообразие людей, которые могут участвовать в процессе разработки данных. Он предоставляет современным специалистам по обработке данных и новаторам в области аналитики инструменты, необходимые для совместной работы и обмена знаниями. Он обслуживает менее технических экспертов в предметной области, а также более технических разработчиков, которые хотят работать на разных уровнях абстракции от необработанных данных. Чтобы поддержать это, Trifacta предлагает:

- ❑ Разработка с минимальным количеством кода / без кода: независимо от того, создаете ли логику инженерии данных визуально или из командной строки с помощью SQL, Python или Javascript. Trifacta поддерживает сочетание обоих подходов в зависимости от навыков и потребностей пользователей. Пользователи также могут создавать популярные фреймворки, такие как dbt и Google Dataform. Благодаря обратной связи и непрерывной проверке пользователи всегда следят за данными в процессе работы, чтобы сразу увидеть результаты своей работы.
- ❑ Макросы, Шаблоны, Обмен : Инженерные данные облака Trifacta такие как макросы, потоки разделяемые данных, рецепты и шаблоны уменьшают объем повторяющийся задач, увеличивают последовательность реализации и позволяют создавать новые проекты. Эти повторно используемые активы расширяют обмен знаниями внутри и между организациями, стандартизируя подход к общим проблемам инженерии данных.
- ❑ Обмен знаниями: лучшие практики, практические советы, обсуждения, сертификаты - все это вместе в новом сообществе Trifacta. Сообщество расширяет приверженность Trifacta к самообслуживанию, давая пользователям возможность использовать инновации друг друга. Этот форум - это не просто место для обучения, это место, где пользователи могут легко поделиться своими лучшими идеями и активами, а также клонировать и настраивать то, что сделали другие пользователи, в широком диапазоне сценариев ETL, качества данных и автоматизации. Вместо того, чтобы всегда начинать с нуля, пользователи могут подключиться к сообществу Trifacta, чтобы получать выгоду от тяжелой работы друг друга, творческого мышления и решения проблем.
- ❑ Ценообразование, основанное на использовании: технологии недостаточно для поощрения демократизации; его цена и упаковка также должны обеспечивать широкое распространение. Trifacta позволяет пользователям начать работу с облаком инженерии данных бесплатно и платить только тогда, когда они видят ценность. Удобные онлайн-покупки по кредитным картам и опции «оплата по факту» предоставляют возможности подготовки данных, ETL и качества данных для отдельных лиц, групп и организаций любого размера.

**Самообслуживание:** Платформа инженерии данных Trifacta увеличивает количество и разнообразие людей, которые могут участвовать в процессе разработки данных. Он предоставляет современным специалистам по обработке данных и новаторам в области аналитики инструменты, необходимые для совместной работы и обмена знаниями. Он обслуживает менее технических экспертов в предметной области, а также более технических разработчиков, которые хотят работать на разных уровнях абстракции от необработанных данных. Чтобы поддержать это, Trifacta предлагает:

- ❑ Разработка с минимальным количеством кода / без кода: независимо от того, создаете ли логику инженерии данных визуально или из командной строки с помощью SQL, Python или Javascript. Trifacta поддерживает сочетание обоих подходов в зависимости от навыков и потребностей пользователей. Пользователи также могут создавать популярные фреймворки, такие как dbt и Google Dataform. Благодаря обратной связи и непрерывной проверке пользователи всегда следят за данными в процессе работы, чтобы сразу увидеть результаты своей работы.
- ❑ Макросы, Шаблоны, Обмен : Инженерные данные облака Trifacta такие как макросы, потоки разделяемые данных, рецепты и шаблоны уменьшают объем повторяющийся задач, увеличивают последовательность реализации и позволяют создавать новые проекты. Эти повторно используемые активы расширяют обмен знаниями внутри и между организациями, стандартизируя подход к общим проблемам инженерии данных.
- ❑ Обмен знаниями: лучшие практики, практические советы, обсуждения, сертификаты - все это вместе в новом сообществе Trifacta. Сообщество расширяет приверженность Trifacta к самообслуживанию, давая пользователям возможность использовать инновации друг друга. Этот форум - это не просто место для обучения, это место, где пользователи могут легко поделиться своими лучшими идеями и активами, а также клонировать и настраивать то, что сделали другие пользователи, в широком диапазоне сценариев ETL, качества данных и автоматизации. Вместо того, чтобы всегда начинать с нуля, пользователи могут подключиться к сообществу Trifacta, чтобы получать выгоду от тяжелой работы друг друга, творческого мышления и решения проблем.
- ❑ Ценообразование, основанное на использовании: технологии недостаточно для поощрения демократизации; его цена и упаковка также должны обеспечивать широкое распространение. Trifacta позволяет пользователям начать работу с облаком инженерии данных бесплатно и платить только тогда, когда они видят ценность. Удобные онлайн-покупки по кредитным картам и опции «оплата по факту» предоставляют возможности подготовки данных, ETL и качества данных для отдельных лиц, групп и организаций любого размера.



# Trifacta сотрудничает с Databricks, чтобы повысить окупаемость инвестиций в хранилища данных

**Trifacta теперь интегрирована с платформой Databricks Lakehouse.** Платформа Databricks Lakehouse, являющаяся типичным хранилищем данных, оптимизирована, открыта и способна поддерживать широкий спектр аналитических рабочих нагрузок и типов данных. Для тех, кто не знаком с этим термином, Lakehouse («домик у озера») сочетает в себе элементы как озера данных, так и традиционного хранилища данных и может упростить настройку нескольких систем, включая озеро данных, несколько хранилищ данных и другие специализированные системы. Благодаря интеграции Trifacta пользователи Databricks Lakehouse Platform теперь могут ускорить процесс разработки и организации конвейеров данных в своих хранилищах данных.

Исторически сложилось так, что небольшая группа инженеров по обработке данных полностью управляла процессом настройки и обслуживания конвейеров данных, используя чрезвычайно трудоемкие инструменты и процессы. Теперь инженеры по обработке данных ищут более гибкие и гибкие варианты, которые позволяют быстрее подготавливать и передавать данные в руки бизнес-пользователей, например Trifacta.

Интеграция **Trifacta с Databricks** также включает поддержку Delta Lake, что позволяет работникам, работающим с данными, ускорить процесс уточнения данных для аналитических рабочих нагрузок. Любое преобразование, выполняемое в Trifacta, транслируется в исполняемый код Spark, который выполняется через Databricks, в то время как Databricks автоматически масштабирует выполнение обработки на основе параметров задания преобразования. Результат - более быстрые и надежные конвейеры данных, которые работают так же усердно, как и аналитики.



# Trifacta позволяет пользователям сосредоточиться на анализе данных, а не на их подготовке

В чем ключ к разработке точных, меняющих правила игры прогнозов на основе данных, которые могут улучшить ваш бизнес и открыть новые возможности? Хорошие, чистые, структурированные данные.

**Результаты машинного обучения (ML) хороши ровно настолько, насколько хороши данные, на которых они построены.**

Любой, кто занимался обработкой данных - процессом обнаружения, структурирования, очистки, обогащения, проверки и публикации набора данных - знает, насколько это может быть сложно, особенно по мере того, как объем точек данных увеличивается, а форматы и источники данных становятся более разнообразными. Опасности, связанные с подготовкой данных, могут помешать даже самым опытным экспертам в предметной области проанализировать свои данные.

Миссия компании Trifacta- ускорить процесс подготовки данных, чтобы помочь клиентам сосредоточиться на самом важном: создании ценности для бизнеса с помощью точной аналитической информации, основанной на данных.

Trifacta, партнер по передовым технологиям партнерской сети AWS и компетентный партнер AWS ML, использует результаты десятилетий инновационных исследований, масштабируемого управления данными и машинного обучения, чтобы сделать процесс подготовки данных более быстрым и интуитивно понятным. Trifacta Wrangler автоматически представляет убедительные визуальные представления ваших данных. Когда вы закрашиваете или нажимаете на определенные элементы, Trifacta предлагает логические преобразования на основе машинного обучения, которые вы можете выбирать, редактировать или создавать с нуля с обратной связью в реальном времени. Затем Trifacta позволяет публиковать подготовленные данные в ряде приложений и инструментов бизнес-аналитики.