

Обогащение данных (Data enrichment)

Разделы: [Бизнес-задачи](#)

Обогащение представляет собой процесс насыщения данных новой информацией, которая позволяет сделать их более ценными и значимыми с точки зрения решения той или иной задачи [анализа](#).

Существует два основных метода обогащения данных — **внешнее** и **внутреннее**.

Внешнее обогащение предполагает привлечение дополнительной информации из источников, которые находятся вне информационной системы предприятия. Таковыми могут быть другие предприятия и организации, финансово-кредитные учреждения, банки, страховые компании, государственные налоговые и статистические службы, органы государственного и муниципального управления, различные службы социальной сферы — миграционная служба, органы труда и занятости, система здравоохранения, пенсионные фонды и т.д.

Источником информации для обогащения данных могут быть любые организации, которые в процессе своей деятельности собирают, структурируют и хранят сведения, связанные с их деятельностью. Внешнее обогащение особенно эффективно при проведении стратегического анализа, когда необходимо учесть все [факторы](#) и тенденции, действующие в [бизнес-окружении](#) предприятия.

Например, при [прогнозировании](#) продаж российских товаров внешним обогащающим признаком, который даст дополнительную информацию для построения прогностической модели, может стать курс доллара. Чем выше курс доллара, тем выше цены на импортные аналоги, что повышает спрос на отечественные образцы.

В настоящее время многие компании используют для внешнего обогащения данных информацию из соцсетей. В интернете работают специализированные сервисы, реализующие системы сбора, хранения, унификации и управления собственными, приобретенными и собранными из открытых источников данными о целевой аудитории и ее поведении для использования в задачах клиентской аналитики. Услугами этих сервисов могут пользоваться компании для внешнего обогащения своих данных.

Внутренне обогащение не предполагает привлечения какой-либо внешней информации. Оно обычно связано с получением и включением в набор данных полезной информации, которая отсутствует в явном виде, но может быть тем или иным способом получена с помощью манипуляций с имеющимися данными. Затем эта информация встраивается в

виде новых полей или даже таблиц в хранилище данных и может быть использована для дальнейшего анализа. Для обогащения данных могут использоваться и результаты их анализа.

Например, при анализе клиентской базы оператора мобильной связи могут использоваться признаки, содержащие продолжительность разговоров в ночное, утреннее, дневное и вечернее время. Возможно, будет полезным обогатить данные полем, в котором будет среднесуточное потребление трафика клиентом. Поскольку среднее значение вычисляется только с использованием собственных данных, такое обогащение является внутренним.

Решение Loginom Data Quality включает множество видов обогащения клиентских данных: дополнение полезными данными, заполнение пропусков, актуализацию. Одним из способов обогащения из популярных сервисов и систем в Loginom является возможность использования внешних данных, а именно интеграция со внешними источниками.