

Золотая запись (Golden record)

Синонимы: Эталонная запись, Эталонные данные, Single version of truth, Единая версия истины

Разделы: [Бизнес-задачи](#)

Золотая запись — понятие в области управления мастер-данными, которое означает создание наиболее достоверного, проверенного, непротиворечивого и полного представления о каждом объекте данных компании (клиенте, товаре, контрагенте и т.д.).

Золотая запись — это единственная, четко определенная версия для каждого объекта данных в информационной системе компании. В этом смысле золотую запись иногда называют «единой версией правды», к которой могут обращаться пользователи, если хотят убедиться, что находящиеся у них данные являются корректными.

Одной из самых сложных задач при реализации MDM-решений является организация рабочего процесса по созданию золотой записи. Необходимо рассмотреть все источники данных и оценить, какие поля в них являются наиболее достоверными. Другими словами, если имеется две системы, регистрирующие имя клиента, то какая из них наиболее корректно фиксирует это имя.

Например, имя клиента может быть представлено: полностью фамилией, именем и отчеством; фамилией и инициалами; просто именем и фамилией. Какую из полученных записей можно рассматривать в качестве золотой? Существует ли другая система регистрации, которая более надежна?

Основными методами при создании и поддержке золотых записей является сопоставление и слияние записей. Если есть две записи, которые очень похожи, каков процесс их включения в золотую запись? Например, рассмотрим две записи:

Last Name	First Name	Member #	Phone Number	Street Address	City, State
Wayland	Jennifer	201215	7065842	123 Maine Street	Camden, Maine
Wayland	Jenn	201211	2078675309	123 Main Street	Camden, Maine

Значение поля Last Name совпадают со значениями поля City, State, а остальные поля разные — для создания золотой записи они не позволяют реализовать автоматическое сопоставление или слияние.

Для примера допустим, что первая запись поступила из источника, который обладает большей надежностью для имен и адресов, а вторая запись — из источника, в котором наиболее точными являются номера клиентов и их телефоны.

При разработке MDM-решения должен быть создан набор инструментов, который позволяет максимально автоматизировать сопоставление и слияние записей. Так, в приведенном примере можно настроить рабочий процесс для получения записей из двух источников, установив правило для слияния/сопоставления: «Взять имя из первой записи, номер клиента и номер телефона из второй записи и адрес из первой записи». Тогда золотая запись может быть сформирована следующим образом:

Last Name	First Name	Member #	Phone Number	Street Address	City, State
Wayland	Jennifer	201211	7065842	123 Maine Street	Camden, Maine

Сопоставление и слияние представляют наибольший интерес, когда исходные поля не являются очевидными «победителями». Возможны ситуации, требующие «ручного» вмешательства. В этом случае решение, какое поле из какой записи должно иметь приоритет, может принять специалист по управлению данными, основываясь на своем опыте и знаниях.

Из-за сложности внедрения MDM-решений для формирования золотой записи полезно ответить на следующие вопросы.

1. Какая информация должна содержаться в золотой записи?
2. Какие источники данных используются для получения записи?
3. Все ли источники данных в настоящее время интегрированы?
4. Какой источник для каждого из полей является наиболее надежным?
5. Какова процедура утверждения золотой записи?
6. Каковы дата и время последних изменений записи (нет ли более новых версий)?

Кроме этого, для определения золотой записи могут использоваться следующие методы:

1. Метод наиболее новой записи (выбирается запись с самыми поздними датой и временем, потенциально несущая наиболее актуальную информацию).
2. Метод наиболее часто встречающейся записи.
3. Метод наибольшей полноты (записи, содержащие больше значений, внесенных в каждое доступное поле, считаются наиболее подходящими).
4. Метод расчетных показателей (по длине записи, по степени доверия к источнику данных и т.д.).

Решение [Loginom Data Quality](#) автоматически производит очистку и исправляет ошибки в данных, приводит их к стандартному виду, восстанавливает и обогащает. Конечная цель очистки клиентских данных — создание золотой записи: единого, точного, актуального и полного варианта записи клиента.

