

Техническое задание

Модуль конвертации многомерных кубов Microsoft SQL Server Analysis Services в описание куба XLTable

Заказчик: ТОО «BR Systems» (участник Astana Hub), разработчик OLAP-сервера XLTable. Версия ТЗ: 1.0 от 19.10.2026.

1. Назначение

Многие компании годами строили отчётность на многомерных кубах Microsoft SQL Server Analysis Services (SSAS Multidimensional), а теперь переносят хранилища на ClickHouse, DuckDB и другие аналитические СУБД. XLTable — OLAP-сервер, к которому Excel подключается так же, как к SSAS (протокол XMLA, язык MDX). При этом данные остаются в СУБД, а XLTable переводит запросы Excel в SQL.

Сейчас каждый куб SSAS при переходе на XLTable описывают заново вручную, и для крупной модели это недели работы. Нужен модуль, который строит описание куба XLTable автоматически из описания куба SSAS и прямо сообщает, что перенести не удалось.

2. Что такое куб XLTable

Куб XLTable — это текстовый файл `.sql`: набор обычных запросов `SELECT` с разметкой в комментариях `-- olap_*`. В нём описаны группы мер, измерения, иерархии, связи, вычисляемые поля, роли и ограничения доступа. Формат открыт:

- описание формата: <https://xltable-olap.readthedocs.io/en/latest/cubes.html>
- примеры для ClickHouse и DuckDB: https://xltable-olap.readthedocs.io/en/latest/clickhouse_sample.html, https://xltable-olap.readthedocs.io/en/latest/duckdb_sample.html
- вся документация одним файлом: <https://xltable-olap.readthedocs.io/en/stable/lms-full.txt>

Исходный код XLTable для работы не нужен. Результат проверяется на бесплатной версии XLTable Desktop (<https://xltable.com/download>), куб кладётся в её папку `cubes`.

3. Входные данные

Многомерная модель SSAS (Multidimensional, SQL Server 2012 и новее) в одном из двух видов:

1. XMLA-скрипт базы, выгруженный из SQL Server Management Studio (Script Database as → CREATE To);
2. папка проекта SQL Server Data Tools / Visual Studio: файлы `.ds`, `.dsv`, `.dim`, `.cube`, `.role`, `.database`.

Оба вида описывают модель на языке ASSL (XML). Подключение к работающему серверу SSAS для конвертации не требуется, модуль работает по файлам.

Дополнительно на вход подаётся необязательный файл соответствия имён (YAML или JSON): как называются таблицы и столбцы источника в новой СУБД (схема, имя таблицы, имена столбцов, регистр). Без него имена переносятся как есть.

Табличные модели SSAS (Tabular, Model.bim) в эту задачу не входят.

4. Выходные данные

1. **Файл куба XLTable** `.sql` — по одному на каждый куб базы SSAS — для выбранной целевой СУБД: ClickHouse или DuckDB. Файл соблюдает все правила формата XLTable. Он загружается в XLTable с включённой проверкой описания (`--definition_check_on`) без ошибок и без ручных правок.
2. **Отчёт о конвертации** в двух видах: для человека (HTML или Markdown) и машиночитаемый (JSON). По каждому объекту модели в отчёте указаны тип, имя, статус («перенесён», «перенесён с изменениями», «не

перенесён»), причина и исходный фрагмент (XML или MDX). В начале отчёта — итог: сколько мер, атрибутов, иерархий, вычислений и ролей перенесено автоматически, в штуках и процентах.

Объект, который не удалось перенести, в файл куба не попадает (файл всегда остаётся рабочим) и обязательно попадает в отчёт.

5. Правила соответствия SSAS → XTable

Объект SSAS	Во что превращается в XTable
Таблица представления источника данных (DSV)	FROM <схема.таблица> <псевдоним> в блоке источника
Именованный запрос DSV	общее табличное выражение (CTE) перед --olap_cube или подзапрос; T-SQL переводится в диалект целевой СУБД
Именованное вычисление DSV (вычисляемый столбец)	выражение в списке полей; T-SQL переводится в диалект целевой СУБД
Группа мер	блок --olap_source <имя> с --olap_measures
Мера с агрегацией Sum, Count, Min, Max	sum, count, min, max
Мера Distinct Count	count(distinct ...) (DuckDB), uniqExact(...) (ClickHouse)
Свойства меры FormatString, Visible = False, DisplayFolder, Description	--format, --hide, --folder, --description
Измерение базы и его атрибуты (KeyColumns, NameColumn)	блок --olap_source с --olap_dimensions; поле строится по NameColumn
Измерение из нескольких таблиц («снежинка»)	соединения с --relationship='part-source' внутри блока измерения
Пользовательская иерархия	уровни с одинаковым --hierarchy, сверху вниз, в одном источнике
Измерение времени (Type = Time, типы атрибутов Years, Quarters, Months, Days)	метки --date на уровнях календаря
Связь измерения с группой мер Regular	LEFT JOIN <таблица> <псевдоним> ON ... в блоке группы мер
Связь Referenced (через другое измерение)	цепочка соединений, которую XTable разбирает сам; либо part-source
Связь Fact (вырожденное измерение)	--relationship='one-table'
Связь Many-to-Many через промежуточную группу мер	--relationship='many-to-many', если связь выражается соединением через таблицу-мост; иначе — в отчёт
Измерение, использованное несколько раз (role-playing: дата заказа, дата отгрузки)	отдельные источники с разными псевдонимами и разными отображаемыми именами
Вычисляемая мера CREATE MEMBER [Measures].[X] AS ... с арифметикой мер и IIF	поле --olap_calculated_fields; деление защищается nullIf
Вычисляемая мера «тот же показатель год назад» (ParallelPeriod по году, кварталу, месяцу)	мера с --offset
Переводы (Translations) на выбранный язык, Caption	--translation
Действие Drillthrough	--olap_drillthrough со списком полей

Объект SSAS	Во что превращается в XTable
Роль: члены роли (группы Windows)	<code>--olap_user_groups</code>
Роль: доступ к данным измерения (AllowedSet / DeniedSet) с явным списком членов	<code>--olap_access_filters</code> вида <code>поле in (...) / not in (...)</code>
Роль: видимость мер и измерений	<code>--olap_measures_visible, --olap_dimensions_visible</code>

Модуль соблюдает ограничения формата XTable, описанные в документации. Самые важные из них:

- имена полей и отображаемые имена уникальны в пределах куба (при совпадении модуль делает их уникальными и пишет об этом в отчёт);
- строковые литералы записываются в обратных апострофах;
- в выражениях полей измерений нет слова `FROM`, комментариев `--` и `LEFT JOIN`;
- имена полей строятся по соглашению XTable (`<псевдоним>_<агрегат>_<столбец>`).

Таблица описывает обязательный минимум. Исполнитель может перенести больше, если результат проходит проверку описания и даёт те же цифры, что SSAS.

6. Что не переносится и должно попасть в отчёт

- полуаддитивные агрегации (LastNonEmpty, FirstChild, AverageOfChildren, ByAccount, None) и MeasureExpression;
- MDX-скрипт сверх простых вычисляемых мер: SCOPE, присваивания ячеек, сложная работа со временем;
- именованные наборы, KPI, перспективы, связанные группы мер и связанные измерения;
- динамическая безопасность (MDX-выражения в AllowedSet / DeniedSet), права на ячейки;
- секции, агрегации и настройки обработки: в XTable расчёт выполняет сама СУБД, это не ошибка, а пояснение в отчёте;
- выражения T-SQL, которые не удалось перевести в диалект целевой СУБД.

По каждому такому объекту в отчёте указывается причина и, где можно, подсказка, как сделать вручную.

7. Требования к модулю

- Консольная утилита на Python 3.10 и новее; работает в Windows и Linux.
- Пример вызова: `ssas2xtable <вход> --target clickhouse|duckdb --lang ru --mapping names.yaml --out <папка>`; параметр `--cube` — конвертировать только один куб.
- Перевод T-SQL в диалекты ClickHouse и DuckDB — средствами открытой библиотеки (например, sqlglot) с доработками.
- Сторонние библиотеки — только со свободными лицензиями (MIT, BSD, Apache 2.0); без GPL и AGPL.
- Модель уровня Adventure Works конвертируется меньше чем за минуту.
- Код покрыт автотестами (pytest): разбор ASSL, каждое правило из раздела 5, формирование отчёта.
- Модуль спроектирован так, чтобы позже его можно было встроить в XTable: сначала из XML строится промежуточная модель, затем из неё — файл куба. Добавить новую целевую СУБД должно быть можно, не переписывая разбор.

8. Тестовые модели и данные

1. **Adventure Works DW, многомерная модель** — открытый учебный пример Microsoft (проект SSAS и база AdventureWorksDW).
2. **Две синтетические модели**, которые составляет исполнитель. Вместе они должны покрыть всё, чего нет в Adventure Works: Many-to-Many, Referenced, Fact, role-playing-измерения, роли с ограничениями доступа, переводы, Drillthrough.
3. Заказчик может дополнительно передать обезличенную модель реального клиента.

Исполнитель загружает данные AdventureWorksDW в ClickHouse и DuckDB. Скрипт загрузки входит в результат работы.

9. Приёмка

1. Все кубы тестовых моделей конвертируются и загружаются в XLTable с `--definition_check_on` без ошибок — и для ClickHouse, и для DuckDB.
2. Не меньше 90 % мер, атрибутов и иерархий тестовых моделей переносятся автоматически.
3. **Совпадение цифр.** 20 контрольных запросов (сочетания мер, измерений и иерархий, включая вычисляемые меры, фильтры и итоги) дают в XLTable те же значения, что в SSAS, с точностью до округления. Сравнение автоматизировано: скрипт отправляет одни и те же MDX-запросы в SSAS и XLTable по XMLA и сравнивает ответы. Скрипт входит в результат.
4. Проверка в Excel: сводная таблица по сконвертированному кубу открывается, раскрывает иерархии и делает Drillthrough.
5. В отчёте о конвертации перечислены все непереведённые объекты тестовых моделей — ни один не потерян молча.
6. Роли: пользователь из группы роли видит в XLTable только разрешённые данные (проверяется на синтетической модели).

10. Этапы

Этап	Содержание	Результат
1. Прототип	разбор ASSL (оба вида входа), промежуточная модель; DSV, измерения, иерархии, группы мер с Sum / Count / Distinct Count; цель DuckDB	базовый куб Adventure Works загружается в XLTable на DuckDB
2. Полный перенос	все правила раздела 5, перевод T-SQL, ClickHouse, отчёт о конвертации, файл соответствия имён	все тестовые модели конвертируются, отчёт полный
3. Сверка и передача	скрипт сравнения цифр, синтетические модели, документация, автотесты	приёмка по разделу 9

Срок и стоимость каждого этапа исполнитель указывает в заявке. Каждый этап заказчик проверяет и принимает отдельно.

11. Что передаёт исполнитель

- исходный код модуля в git-репозитории заказчика, с историей изменений;
- инструкцию пользователя (установка, параметры, файл соответствия имён, чтение отчёта) на русском или английском;
- описание внутреннего устройства: промежуточная модель и как добавить новую целевую СУБД;
- автотесты, синтетические модели, скрипты загрузки данных и сравнения цифр.

Исключительные права на результат, включая исходный код, переходят заказчику. Исполнитель не публикует код и материалы без согласия заказчика.

12. Что предоставляет заказчик

- консультации по формату куба XLTable (созвон раз в неделю и чат);
- бесплатную версию XLTable Desktop, а для проверки ролей — временную лицензию серверной версии;
- проверку промежуточных результатов в течение 5 рабочих дней после сдачи этапа.