SPSync

Column type mapping (PDF)

The SQL Server and PostgreSQL column SPSync writes for each SharePoint column type, with child tables for multi-value columns and the types being verified.

Anyone writing a migration script or a Power BI model against SharePoint data ends up building this table by hand. A multi-select choice can't sit in one cell, and a currency column needs a precision. This is the mapping SPSync uses for SQL Server and PostgreSQL, from our schema reference, with a note wherever a type is still being checked against a live tenant.

How SharePoint columns become SQL columns

SharePoint type | SQL Server | PostgreSQL | Columns written
Single line of text | NVARCHAR(MAX) | character varying | {col}
Multiple lines of text | NVARCHAR(MAX) | character varying | {col}
Number | FLOAT | double precision | {col}
Currency | DECIMAL(38,9) | numeric(38,9) | {col}
Yes/No | BIT | boolean | {col}
Date only | DATE | date | {col}
Date and time | DATETIMEOFFSET(7) | timestamptz | {col}
Choice | NVARCHAR(MAX) | character varying | {col}
Choice, multiple selections | child table | child table | {table}__{col} with value and ordinal
Lookup | BIGINT and NVARCHAR(MAX) | bigint and character varying | {col}_id, {col}_value
Lookup, multiple values | child table | child table | {table}__{col} with lookup_id, lookup_value, ordinal
Person or group | UNIQUEIDENTIFIER | uuid | {col}_user_id, joins to sp_users.id
Person or group, multiple | child table | child table | {table}__{col} with user_id, ordinal
Hyperlink or picture | NVARCHAR(MAX) twice | character varying twice | {col}_url, {col}_description
Location | NVARCHAR(MAX), FLOAT, FLOAT | jsonb, double precision twice | {col}_json, {col}_lat, {col}_lng
Calculated | the result type's mapping | the result type's mapping | {col}
Managed metadata | {col}_term_id, {col}_label | child table: term_id, label, ordinal | none

Every column that comes from a SharePoint column is nullable in the database, whatever the Required setting says. SharePoint doesn't apply Required to items that existed before the setting changed, and a sync has to accept those rows.

Lookups and people

A single lookup gives you two columns. {col}_id holds the SharePoint item id of the target, so it joins to that list's table on sp_item_id when you replicate both lists. {col}_value holds the display text, so a report can show the name without a join. SPSync writes no foreign key constraint, which means load order can never fail a sync. Add the constraint yourself if you want one.

People work differently. A person column stores only a user id, and the name and email live once in the shared sp_users table. The id is a UUID derived from the person's login, so the same person gets the same id in every list. Reporting views, if you turn them on, do the sp_users join for you.

Multi-value columns become child tables

A multi-select choice, a multi-value lookup and a multi-person column each get their own table, named after the list table and the column with a double underscore. A Projects list with a multi-select Tags column produces ks_projects and ks_projects__tags, with one row per tag per item. Each child row carries sp_item_id and sp_site_id to join back to the parent, plus an ordinal so you can rebuild the order a user picked.

A flat export gets this wrong. A semicolon-joined string looks fine in Excel but can't be counted in SQL without string splitting. With a child table, "how many projects carry the Safety tag" is a join and a count(*).

Types to check before you rely on them

Text, number, choice and yes/no columns, date columns, single lookups (the id and the display text) and the system columns have been confirmed against a live tenant. A few mappings above are written to the schema but are still being checked against real Microsoft Graph responses.

Type | Status today

Currency, hyperlink | Mapping in place, being verified on a live tenant

Multi-value lookup and person | Not confirmed on a live tenant

Date only, on a site east of UTC | Set site_time_zone in the replication, or the date reads one day early. Being verified on a live tenant

Calculated | Graph sends every calculated value as text. SPSync converts it to the formula's result type and writes NULL if it doesn't parse. Only the number result has been recorded live

Managed metadata | The term's GUID and label. Graph sends no type for these columns, so SPSync recognises one from its values; a column with no values yet is left out with a warning in the run log until it has one. Being verified end to end on a live tenant

Library file name, URL and size | Not confirmed on a live tenant

Metadata columns on every table

Each list table also carries system columns. sp_item_id is the primary key, and sp_created, sp_modified, sp_author_id and sp_editor_id replace SharePoint's own Created, Modified and author fields. sp_deleted_at gets a timestamp when an item is deleted in SharePoint, so the row stays unless you set hard deletes in a command-line replication.

Names, renames and type changes

Column names come from SharePoint's internal name, never the display name, so renaming a column in SharePoint never renames it in the database. SPSync decodes _x0020_ back to a space

and converts to lower case with underscores, so Due_x0020_Date becomes due_date. A name that clashes with a SQL reserved word gets _col, which turns Order into order_col.

A new SharePoint column appears on the next run, NULL for existing rows, and a removed column stays in the database. If someone changes a column's type, SPSync creates {col}_v2 with the new type on SQL Server and PostgreSQL and leaves the old column alone, with a warning in the run log naming both.

Full, up to date version: https://spsync.io/resources/sharepoint-column-to-sql-map/