

SpreadsheetMapper

Der **Mapper** ist der wichtigste Step in Synesty Studio. Mit diesem Step kann man ein Spreadsheet bearbeiten und in ein anderes Spreadsheet umwandeln.

Der Step bietet einen umfangreichen Mapping-Editor, mit dem die Spalten des Spreadsheets bearbeitet werden können. Eine Anleitung zu den einzelnen Funktionen dieses Steps [finden Sie hier](#).

- 1 [Erweiterte Optionen](#)
 - 1.1 [mappingMode](#)
 - 1.2 [Warnungen Deaktivieren](#)
 - 1.3 [CacheMode Aktivieren](#)
- 2 [Anwendungsbeispiel](#)
- 3 [Ziel-Schema für Validierung und Filterung](#)
- 4 [Warnungen und Fehlermeldungen](#)
 - 4.1 [Warnung bei nicht \(mehr\) existierenden Spalten](#)
 - 4.2 [Warnung bei ungemappten Werten](#)
- 5 [Wie kann ich den Dateinamen der eingelesenen Datei in einer Spalte ausgeben?](#)

The screenshot shows the SpreadsheetMapper interface with three mapping columns and a function list. The top bar contains buttons for 'Vorschau anzeigen', 'Spalten', 'Einstellungen', 'Finish Configuration', and 'Abbrechen'. The three mapping columns are:

- Column 1: Titel 'name', Quelle 'name', Wert field.
- Column 2: Titel 'brandname', Quelle 'brandname', Wert field.
- Column 3: Titel 'description', Quelle 'description', Wert field.

Below the columns is a 'Funktionen' section with a search bar and a list of functions: Suchen & Ersetzen, Rechenoperation, Runden, Datumsfunktionen, Querverweis, Mappingset, Skript, and Hilfe zu Funktionen. The bottom part of the screenshot shows a preview of the mapped data:

Super Shoe XI	K-Star	Description of Super Shoe XI
Vintage	3-Stripes	Description of Vintage. This is a very large product description which spans multiple lines and usually is challenging to display for common spreadsheet processing tools. Lorem ipsum dolor sit amet, c(...)
Infinity	Runnerschoice	Description of Infinity

Erweiterte Optionen

mappingMode

Mit dem mappingMode kann man einstellen, wie das eingestellte Mapping angewendet wird. Es gibt 2 Optionen:

1. Default - output only mapped columns

Hier erscheinen nur gemappte Spalten im Output-Spreadsheet.

2. Pass-through unmapped source columns

In diesem Modus beinhaltet das Output-Spreadsheet alle im Mapping definierten Spalten und **alle** Spalten des Input-Spreadsheets. Pass-through bedeutet, dass alle Spalten des Input-Spreadsheets durchgeschleift werden.

Spalten im Mapping überschreiben Spalten mit gleichem Namen im Input-Spreadsheet. Alle zusätzlichen Spalten im Mapping (die nicht im Input-Spreadsheet vorkommen), werden am Ende angefügt.

Beispiel für MappingMode

Input-Spreadsheet

a	b
1	2

Mapping

a=a+1, neueSpalte=Test

Ergebnis mappingMode = Default

a	neueSpalte
2	Test

Ergebnis mappingMode = Pass-through unmapped source columns

a	b	neueSpalte
2	2	Test

Hier wird Spalte a überschrieben (Funktion wird angewendet), Spalte b wird durchgeschleift (pass-through) und die neue Spalte wird hinten angefügt.


Wann brauchen ich den Pass-Through Modus?

Der Modus *Pass-through unmapped source columns* ist sehr praktisch, wenn man z.B. ein großes Input-Spreadsheet mit 40 Spalten hat, aber man nur zwei Spalten per Funktion bearbeiten will, aber trotzdem alle Spalten des Input-Spreadsheets auch wieder (unangetastet) im Output-Spreadsheet sehen möchten. In diesem Fall reicht es, wenn man dem Mapping nur die zwei zu bearbeitenden Spalten hinzufügt, statt alle 40 Spalten im Mapping zu definieren, wovon 38 einfach nur durchschleifen. Ein weiterer Effekt davon ist, dass auf diese Weise auch neue Spalten im Input-Spreadsheet automatisch mit durchgeschleift werden. Das ist praktisch z.B. ein Lieferant neue Felder in einer CSV-Datei schickt, die man automatisch unangetastet importieren will, aber auf wenige bekannte Felder noch Funktionen anwenden will. D.h. man kümmert sich nur um die wenigen Spalten, die man bearbeiten möchte, alles andere bleibt unangetastet.

Warnungen Deaktivieren

disableWarnings

Disable unmapped Values Warning
Disable processing errors (e.g. syntax error)





Die Option *disableWarnings* kann genutzt werden, um Warnungen zu deaktivieren. Siehe [Warnungen und Fehlermeldungen](#).

CacheMode Aktivieren

cacheMode

Ja



Ja, um Caching zu aktivieren. Das kann langsame Verarbeitung bei Verwendung des SpreadsheetMapper beschleunigen, wenn das Ergebnis-Spreadsheet von min. zwei weiteren Steps verarbeitet wird.

Die Option cacheMode kann genutzt werden, um das Ergebnis des Mappers zwischen zu speichern.

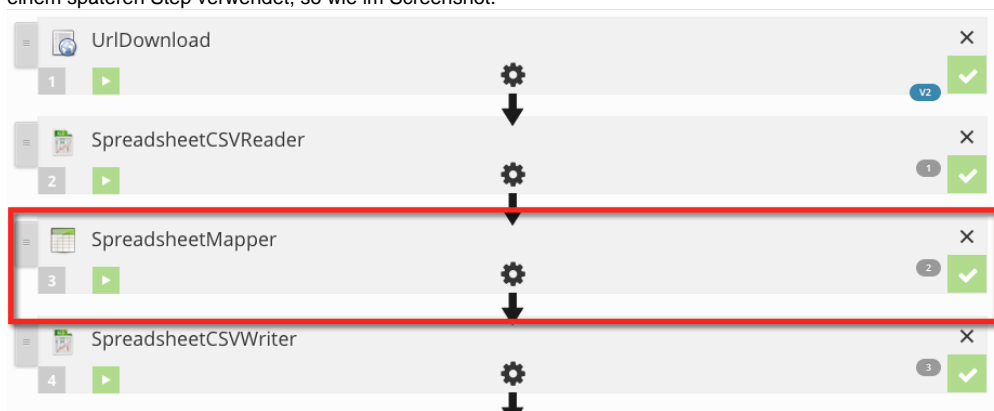
Wann sollte das aktiviert werden? Sobald im Flow zwei oder mehr Steps als Input ein und denselben Mappers haben. Dann sollte bei diesen Mapper der CacheMode aktiviert werden und die Laufzeit zu verringern.



Hinweis: Bei aktiviertem CacheMode gibt es eine Einschränkung. Der Zugriff von Varianten-Zeilen auf Parent-Zeilen funktioniert, aber der Zugriff von Child-Zeilen auf Master-Zeilen funktioniert nicht. D.h. man kann `#{parent['columnname']}` benutzen, jedoch nicht `#{master['columnname']}`. (Siehe auch [Auf Relationen zugreifen](#)). Aus diesem Grund ist der CacheMode standardmäßig deaktiviert.

Anwendungsbeispiel

Den Mapper findet man meistens nach Steps, welche ein SPREADSHEET als Output haben. Dieses Spreadsheet wird dann damit aufbereitet und in einem späteren Step verwendet, so wie im Screenshot.



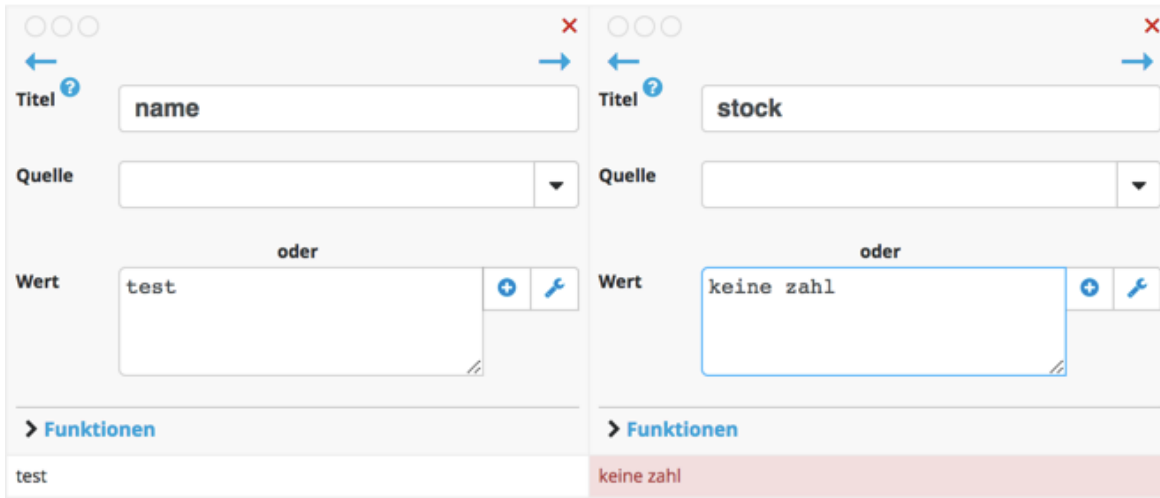
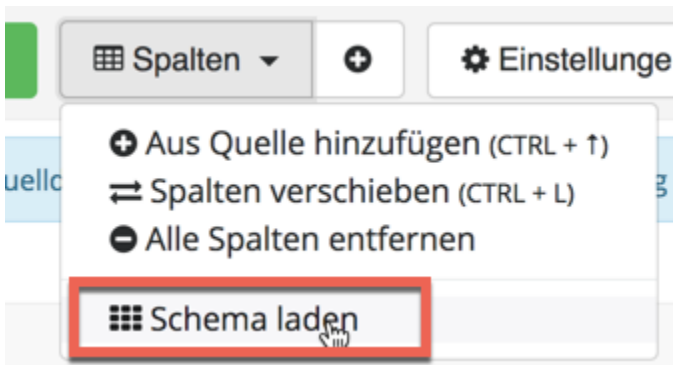
Ziel-Schema für Validierung und Filterung

Man kann im Mapper auch ein Schema hinterlegen (siehe [Datastores#Schema](#)). Damit hat man die Möglichkeit die Zielwerte die durch das Mapping entstehen auch zu validieren, d.h. zu prüfen, ob die Werte gemäß der Schema-Definition korrekt sind.

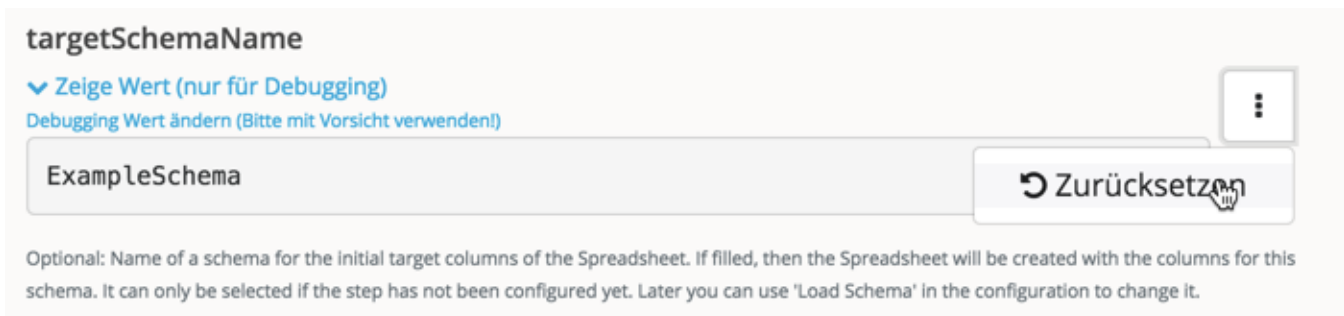
Das ist praktisch, da Fehler visuell hervorgehoben werden und man invalide Daten auch direkt herausfiltern kann.

Um ein Schema auf ein Spreadsheet anzuwenden, klicken Sie auf *Spalten / Schema laden* und wählen das Schema.

! Achtung: Durch das Laden eines Schemas wird die aktuelle Spaltenkonfiguration überschrieben. Wir empfehlen ein Schema gleich zu Beginn der Bearbeitung des Spreadsheets zu laden.



Das geladene Schema wird am Mapper Step gespeichert unter *Erweiterte Optionen* / *targetSchemaName* und kann dort auch zurückgesetzt werden.



Warnungen und Fehlermeldungen

Warnung bei nicht (mehr) existierenden Spalten

Manchmal kommt es vor, dass sich Spaltenbezeichnungen von CSV- oder XLS-Dateien ändern. Nicht immer hat man darauf Einfluss. Damit man besser mitbekommt, wenn z.B. ein Lieferant Spaltenbezeichnungen der CSV-Datei ändert, wird jetzt eine Warnung im Flow-Eventlog ausgegeben. Diese Warnungen kommen vom Mapper Step, **wenn eine Quellspalte benutzt wird, die es in der Quelle nicht mehr gibt**.

Beispiel:



Beispiel

```
Row: 1  
{article_id=Error: SourceColumn does not exist: ean}
```

Die Beispiel-Fehlermeldung sagt, dass in der Zeile 1 und Spalte *article_id* eine Quellspalte *ean* **erwartet wird**, die aber im Quell-Spreadsheet nicht mehr existiert. Das könnte z.B. passieren, wenn diese Spalte in der Quelle z.B. von *"ean"* in *"ean13"* umbenannt, oder gar entfernt wurde.

Sie sollten darauf hin den entsprechenden Mapper kontrollieren und in der Spalte *article_id* die Quellspalte anpassen und durch die neue korrekte Quellspalte ersetzen.

i Hinweis

Falls Sie sich wundern, dass ihre Flows plötzlich mit einer Warnung versehen werden, dann könnte genau das die Ursache sein. Wenn Sie im Eventlog "*SourceColumn name does not exist*" finden, dann prüfen Sie alle Mapper, ob sie Quellspalten nutzen, die es im Quell-Spreadsheet nicht mehr gibt, oder die anders geschrieben werden.

Eine andere Ursache kann sein, dass sie ein Spreadsheet eines Steps verwenden, der je nach konfiguriertem "Mode" unterschiedliche Spalten in im Ausgabe-Spreadsheet hat. D.h. wenn sie den Mode in der Quelle anpassen und die Ausgabe dadurch plötzlich weniger Spalten hat, dann werden Sie wahrscheinlich diese Warnung im folgenden Mapper sehen, falls dieser eine der entfernten Quell-Spalten erwartet.

Der Flow wird auch trotz Warnung bis zum Ende ausgeführt. Aber diese Warnung kann darauf hindeuten, dass ein Problem vorliegt und einige Spalten nicht mehr so gefüllt sind, wie das vor den Warnungen der Fall war.

Das Error-Handling ist in dieser Richtung strenger geworden. Das hat ggf. Anpassungen zur Folge, damit sie diese Meldungen wieder weg bekommen. Dieser Mehraufwand rechtfertigt sich aber dadurch, dass sie dann auch bemerken, wenn wirklich ein Fehler vorliegt wenn Quellspalten plötzlich verschwinden.

Warnung bei ungemappten Werten

Flows werden im Status "WARNUNG" beendet, wenn Sie in einem Mapper ein unvollständiges [Mappingset](#) verwenden. Bei Bedarf können Sie [diese Warnung auch deaktivieren](#), so dass der Flow nicht im Status "WARNUNG" endet.

! Wir raten Ihnen dies nur zu deaktivieren, wenn Sie sich sicher sind, dass die Warnung in Ordnung ist und es sich nicht um einen Fehler handelt. In der Regel deutet die Warnung darauf, dass Sie etwas vergessen haben, oder mit der Zeit neue Werte hinzu gekommen sind, die Sie im Mappingset nachtragen sollten.

[Lesen Sie mehr über diese Warnung und was Sie tun können.](#)

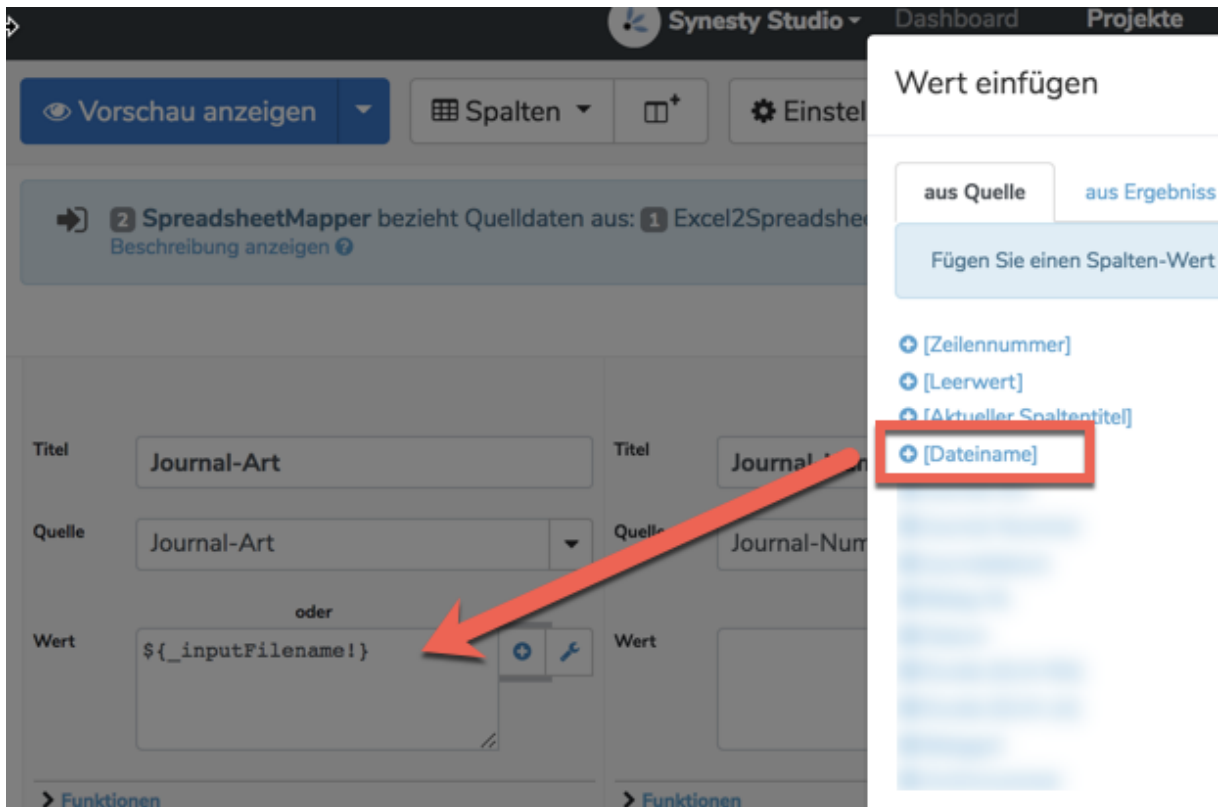
Wie kann ich den Dateinamen der eingelesenen Datei in einer Spalte ausgeben?

Beispiel:

Sie lesen eine Excel-Datei mit dem Step [Excel2Spreadsheet](#) ein, damit Sie ein SPREADSHEET bekommen. Sie möchten den Dateinamen der Excel-Datei in einer Spalte des Spreadsheets ausgeben. Grund könnte sein, dass dieser Dateiname dynamisch ist, und sich z.B. täglich ändert.

Im Mapper Step der den Excel2Spreadsheet Output konsumiert, sollte eine Variable **Dateiname** auswählbar sein. Darin ist der Dateiname der Excel-Datei enthalten.

```
$_inputFilename!
```



i Dieser Mechanismus um auf den Dateinamen zuzugreifen funktioniert auch in den Steps:

- SpreadsheetCSVReader
- JSON2Spreadsheet
- XML2Spreadsheet
- VisualJSON2Spreadsheet
- VisualXML2Spreadsheet