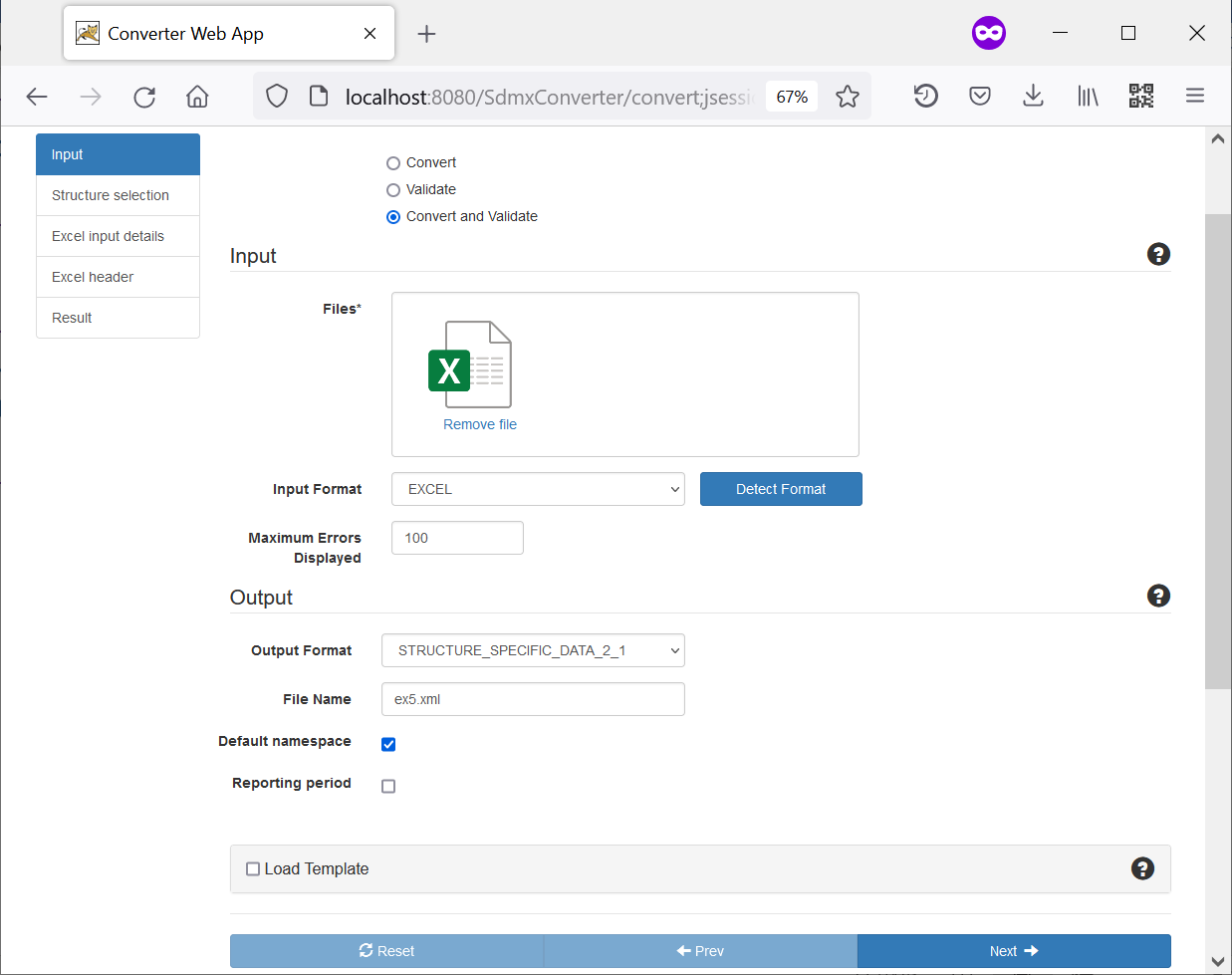
# Exercise 5

## Converting data using SDG dataflows and content constraints

In this exercise, you compare converting an Excel spreadsheet to SDMX using the Global SDG DSD vs the global SDG dataflows.

1. Open SDMX Converter at <http://dfs-desaiis-56.dpko.un.org:81/converter-webapp>
2. SDMX Converter opens. On the first screen, you select the input file containing the data (Oman.xlsx), the output file where the SDMX data will be written, and the format of the output file.



1. Select: **Convert and Validate**

3. Make sure input format is **Excel**

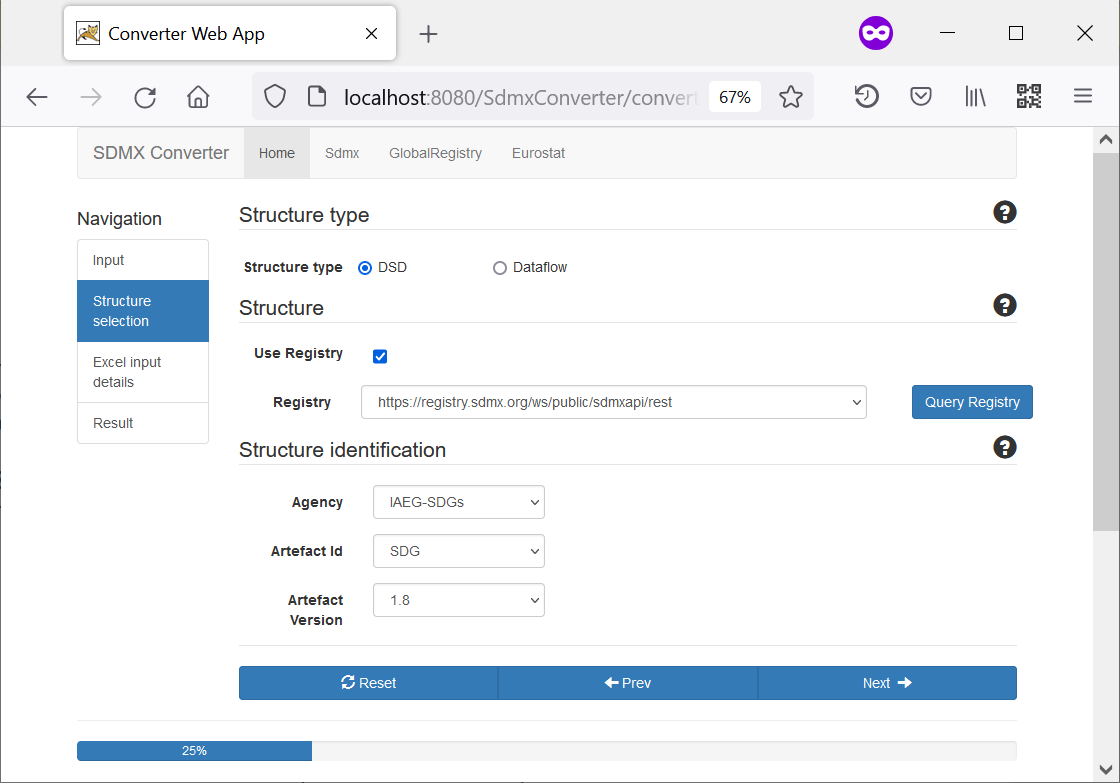
5. Enter output file name, e.g. **Ex5.xml**

2. Drag and drop file **Oman.xlsx**

6. Click **Next**

4. Select output format: **STRUCTURE\_SPECIFIC\_DATA\_2\_1**

1. On this screen, you load the SDG Data Structure Definition from the SDMX Global Registry.



7. Select: **DSD**

9. Select the Global Registry at **https://registry.sdmx.org/...**

8. Check **Registry**

12. Click **Next**

10. Click **Query** **Registry**

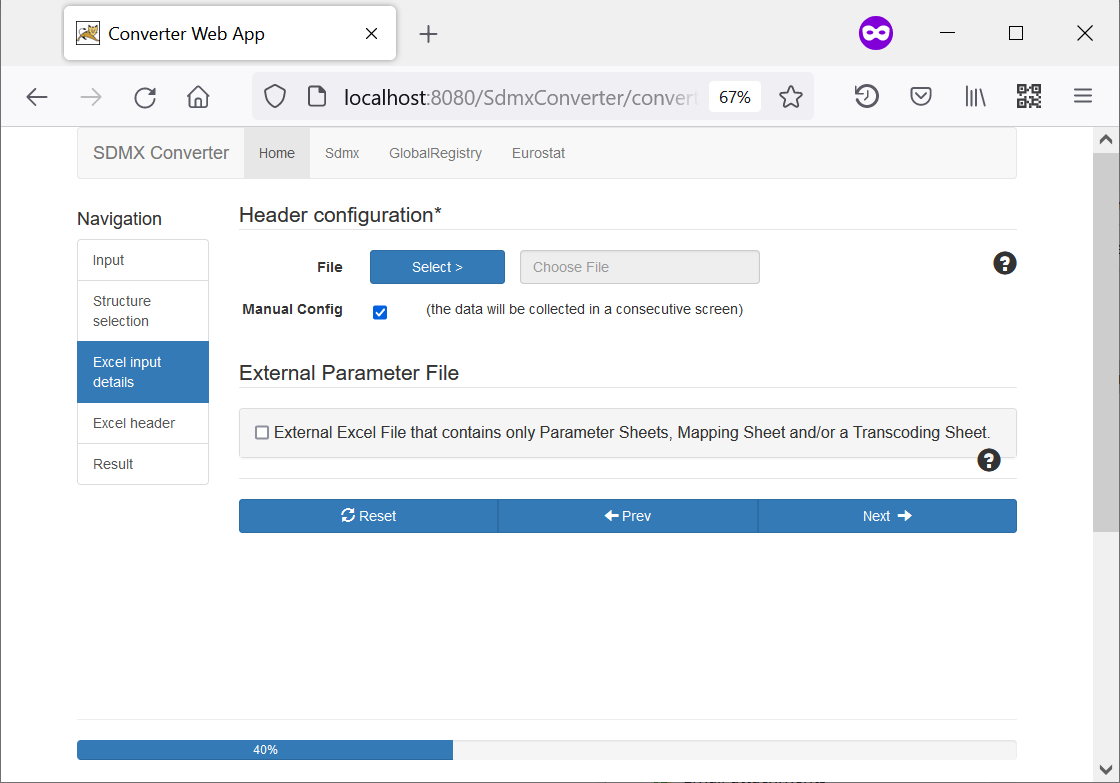
11. Select

Agency: **IAEG-SDGs**

Artefact ID: **SDG**

Version: latest version listed

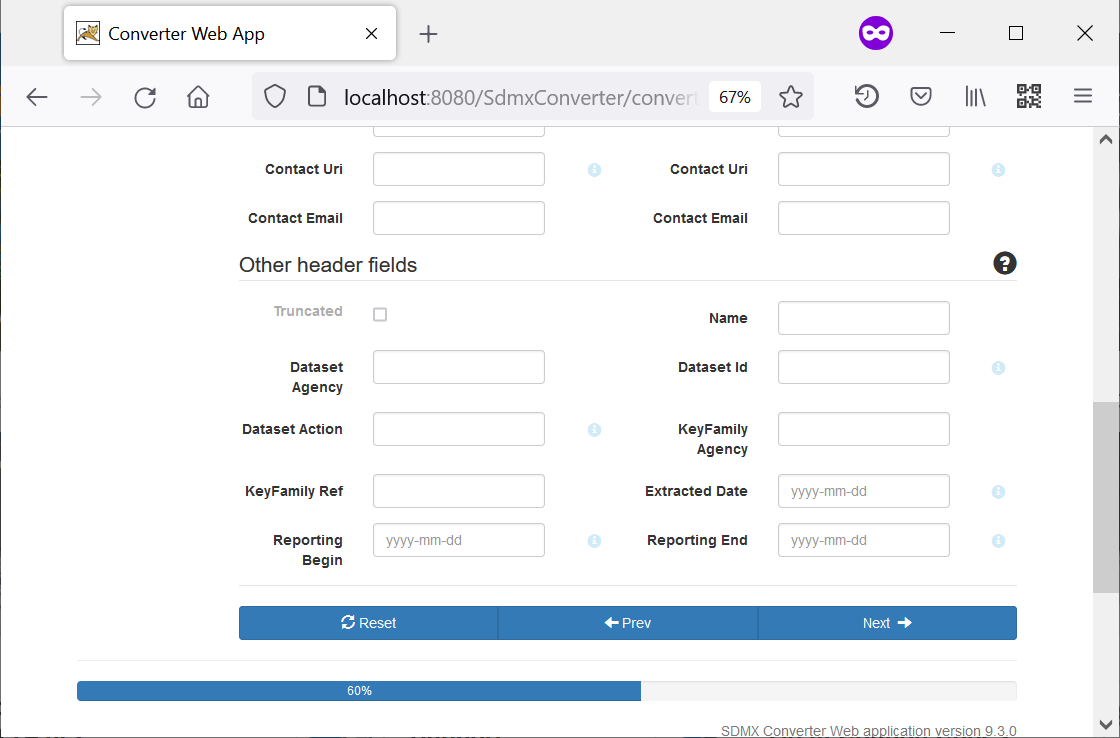
1. On this screen, you can load an SDMX message header and mapping parameters.   
   Every SDMX message must have a header in a defined format, but it is not used in this exercise. As for parameters, they are contained in the Excel file.



13. Check **Manual Config**

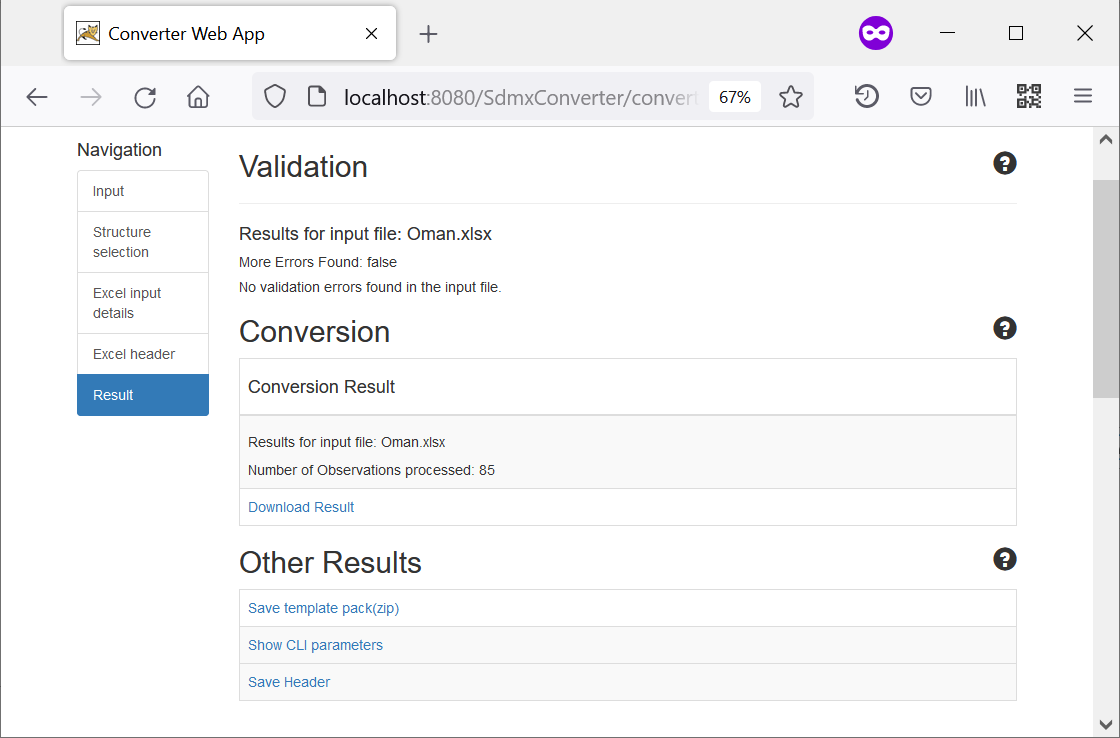
14. Click **Next**

1. On this screen, you can configure the header. You may wish to update the header fields, or simply leave the defaults because we are not using them.



15. Click **Next**

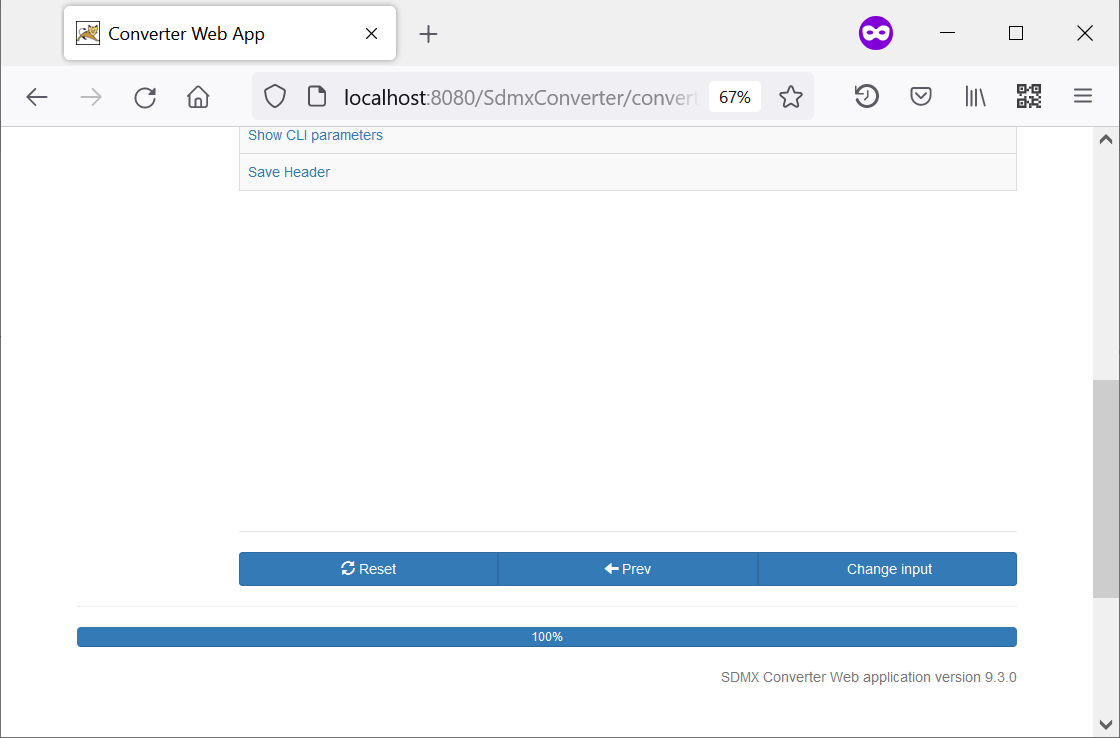
1. If the steps have been completed correctly, the data will be retrieved without errors.



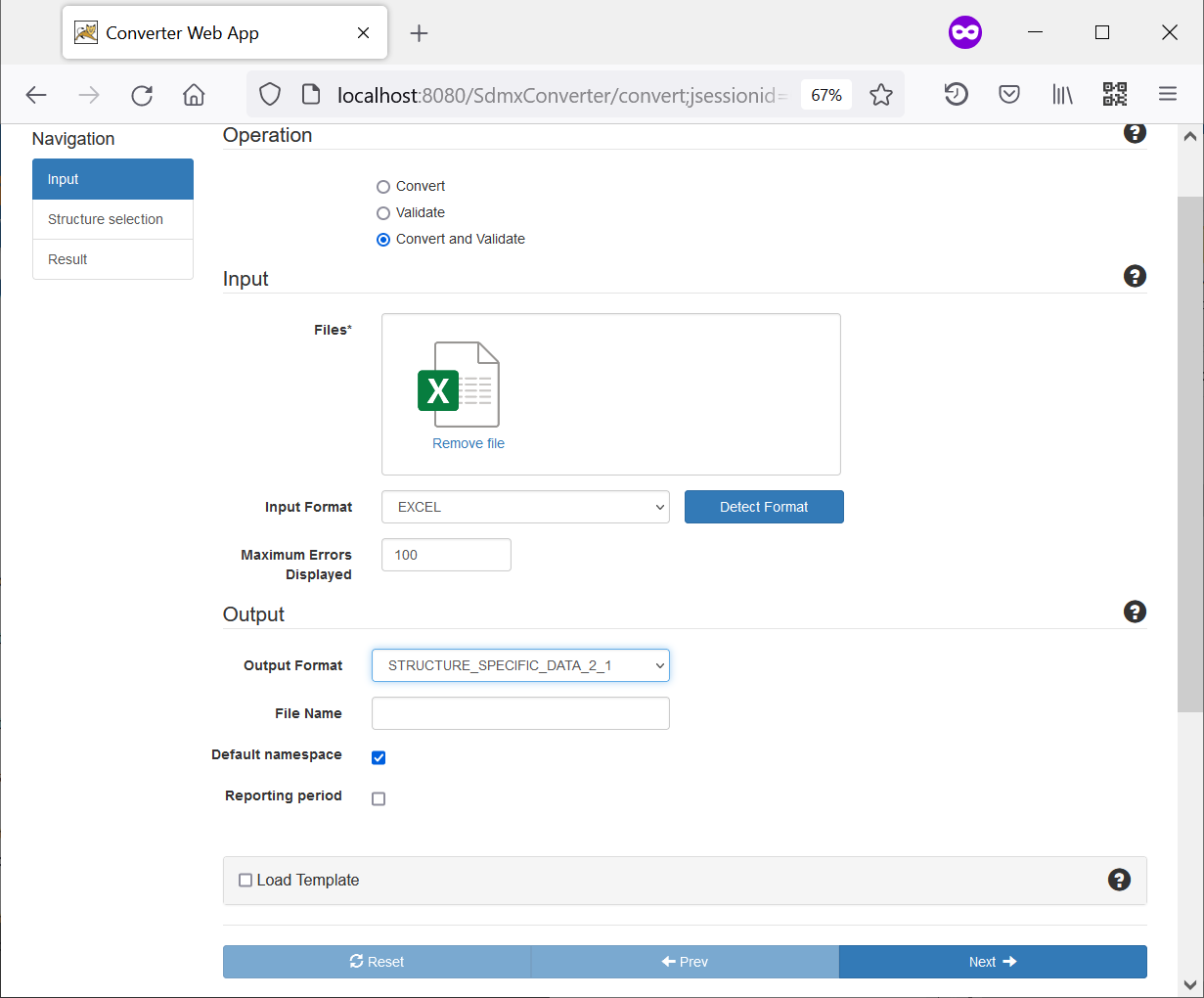
16. Click **Download Result**

1. Now, we will convert the same dataset using the SDG country global dataflow instead of the DSD. Click **Change input** to go back to the first screen.

17. Click **Change input**

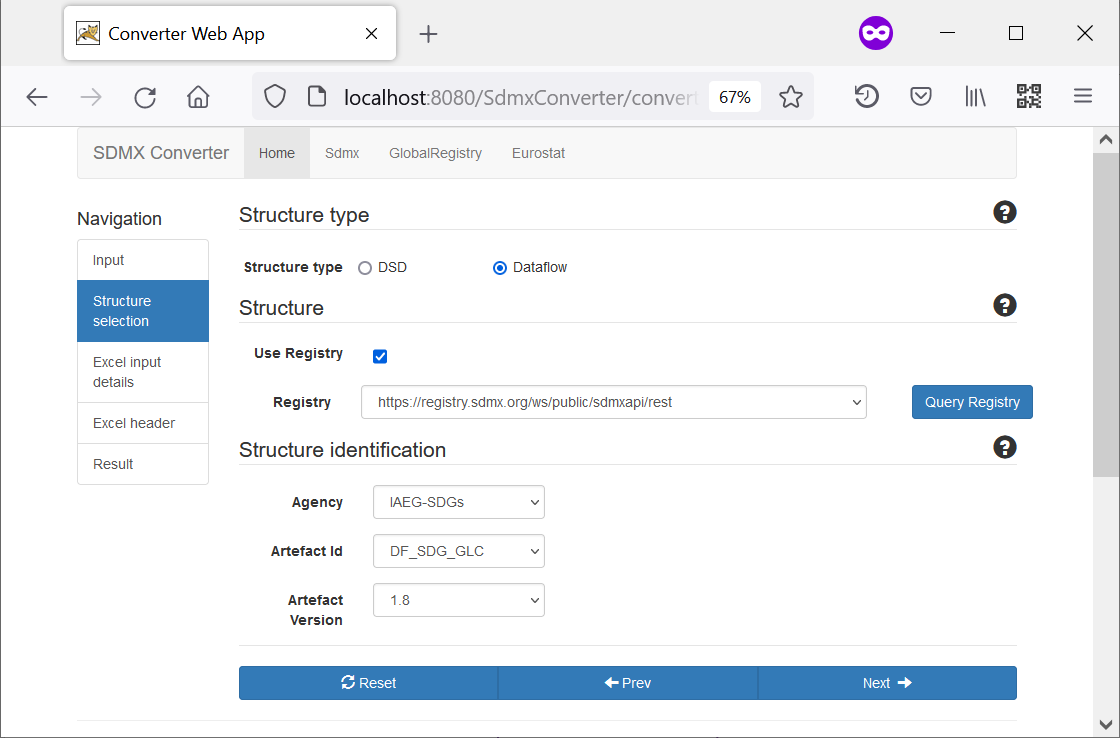


1. You are taken back to the Input Selection screen. Since the dataset has not been updated, no change is needed on the screen.



18. Click **Next**

1. Now, instead of the SDG DSD, select the SDG Country Global Dataflow (**DF\_SDG\_GLC**).



19. Select: **Dataflow**

21. Select the Global Registry at **https://registry.sdmx.org/...**

20. Check **Registry**

24. Click **Next**

22. Click **Query** **Registry**

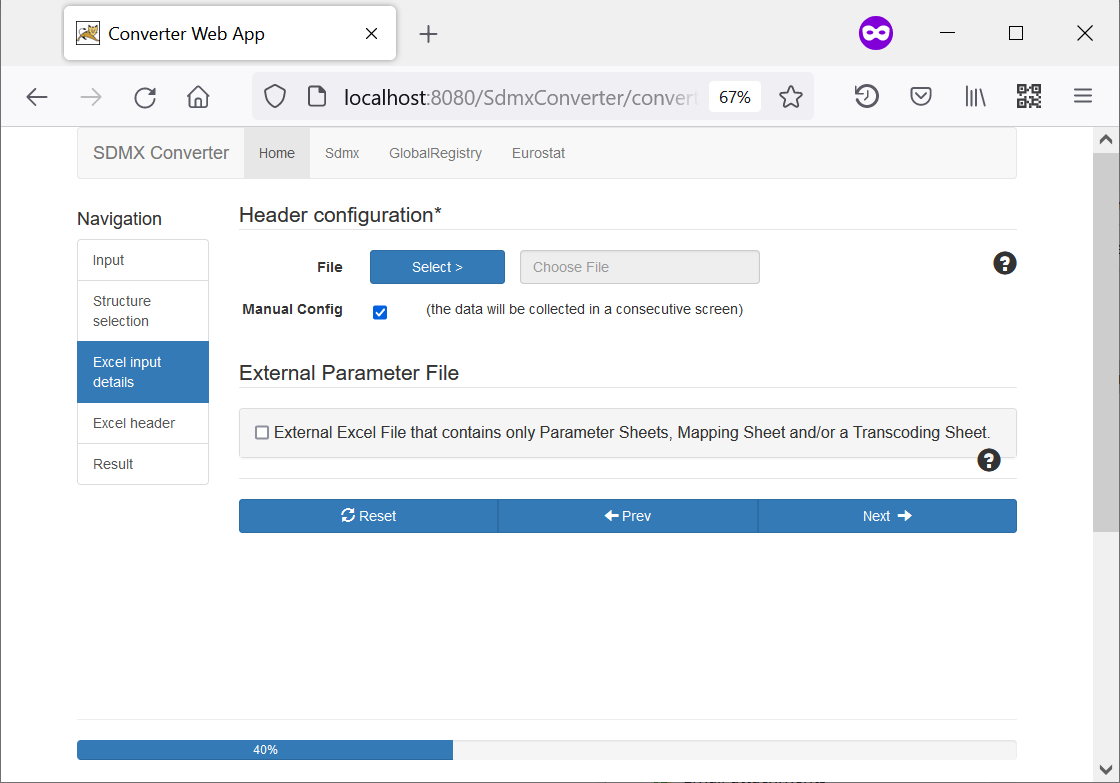
23. Select

Agency: **IAEG-SDGs**

Artefact ID: **DF\_SDG\_GLC**

Version: latest version listed

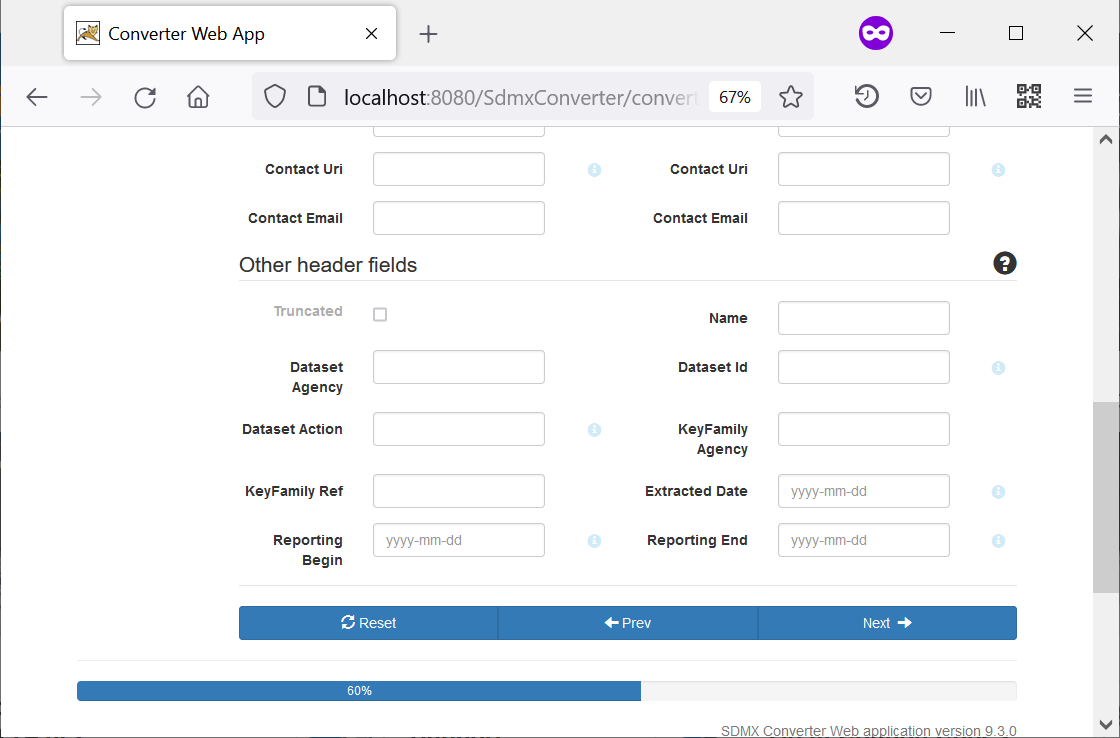
1. No change is required at this this screen.



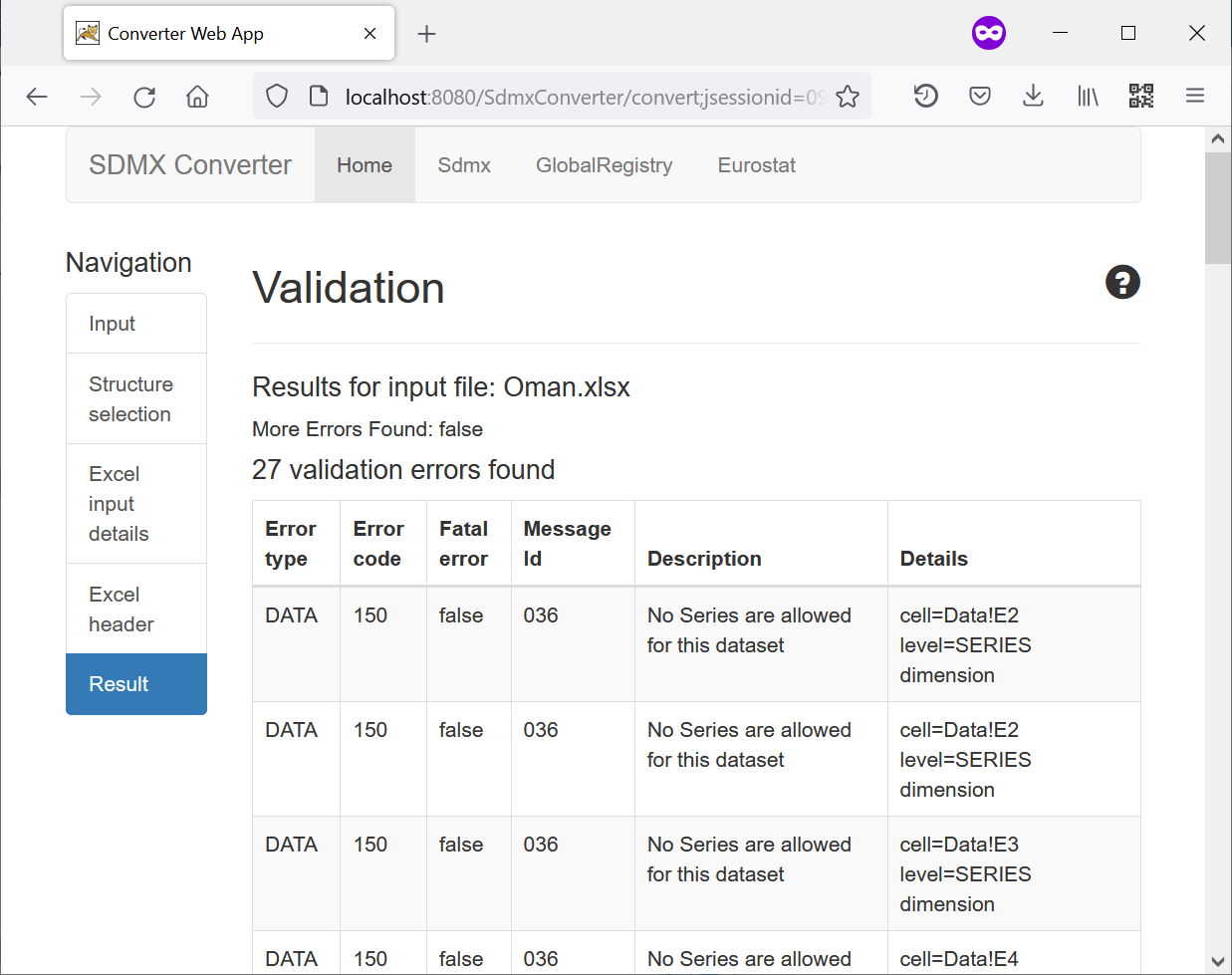
25. Check **Manual Config**

26. Click **Next**

1. No change is required at this screen.



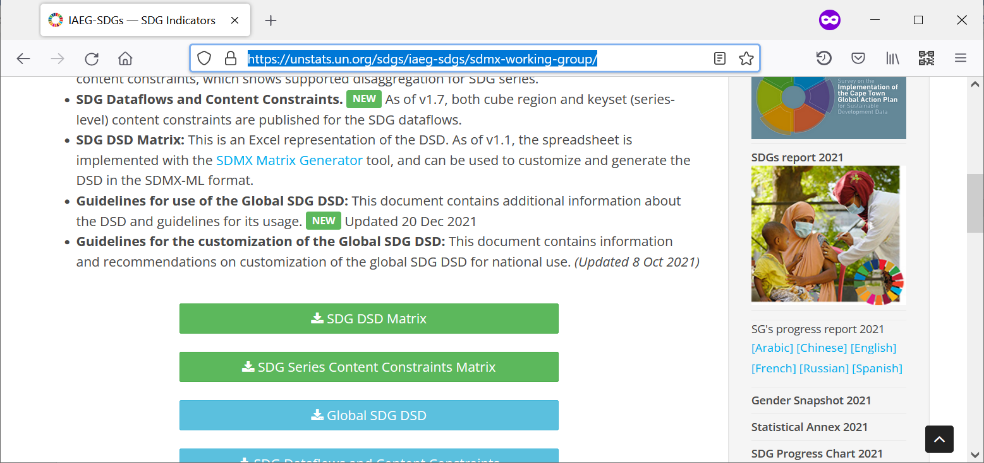
27. Click **Next**

1. Now, you can see that there are validation errors. This is because this time the dataset was validated using the global SDG dataflow, and content constraints took effect. The dataset in question is structurally valid but contains errors in relationships between dimensions.
2. Note that error descriptions say **No Series are allowed for this dataset**. This indicates that the series encoding is not allowed under Content Constraints.
3. Open the spreadsheet **Oman.xlsx** and examine the data. You will see that it contains data for 3 series:

* SH\_STA\_STNTN
* SH\_STA\_WAST
* SN\_STA\_OVWGT

We need to check the encoding of these series against the Content Constraints to see whether codes for all dimensions are allowed.

1. To download Excel/CSV representation of the SDG Content Constraints, open <https://unstats.un.org/sdgs/iaeg-sdgs/sdmx-working-group/> and download the **SDG Series Content Constraints Matrix**.



1. Open the Content Constraint matrix. You will see that for each series, it lists allowed codes for each dimension. Multiple allowed codes are separated with a semicolon (**;**), and if all codes are allowed for a dimension, **all** is used.
2. Locate the first series in the spreadsheet, **SH\_STA\_STNTN**. Compare age groups allowed for that series in the content constraints, with age groups used in the data spreadsheet. Update the age group in the data accordingly.
3. Repeat steps 10 and 12 for the other two series, **SH\_STA\_WAST** and **SN\_STA\_OVWGT**.
4. Save the data spreadsheet and convert it again using the **DF\_SDG\_GLC** dataflow. You should see no errors this time.
5. Congratulations! You used Content Constraints to validate the file, and fix issues with incorrect encoding of data.