FM

FLOWMATRIX.DEV

Premium Workflow

A practical guide – step by step

Data Enrichment Pipeline

A data enrichment pipeline: import raw data, enrich it via API, normalize it with AI and store it structured in Supabase.

DATA & AUTOMATION

EXPERT

ABOUT 4 HOURS SETUP

OVERVIEW

Data Enrichment Pipeline

A data enrichment pipeline: import raw data, enrich it via API, normalize it with AI and store it structured in Supabase.

CATEGORY	DIFFICULTY	DURATION
Data & Automation	**Expert**	**About 4 hours setup**

What you use it for

Data-driven products with large, inconsistent datasets.

Tools in this workflow

n8n · Supabase · OpenAI API · Mapbox Maps API

STEP BY STEP

How to implement it

1

Import raw data from CSV or APIs

In this step, raw data from CSV files or APIs will be imported and prepared for further processing.

WHAT YOU NEED TO DO

1. Open n8n and create a new workflow.
2. Add a 'Read Binary File' Node to import CSV files.
3. Configure the Node with the path of the CSV file.
4. Add an 'HTTP Request' Node to import data from an API.
5. Configure the HTTP Request Node with the API URL and necessary headers.

CHECK AT THE END

[ ] CSV file successfully imported into n8n

[ ] API data successfully imported into n8n

RESULT

Raw data is successfully imported into the system.

LINKS

n8n CSV Import: https://docs.n8n.io/integrations/builtin/applications/csv/

2

Enrich records via external APIs

The imported datasets are enriched by external APIs to obtain additional information such as location data or demographic information.

WHAT YOU NEED TO DO

1. Fulfill this step by adding an 'HTTP Request' node in n8n to use the Mapbox Maps API.
2. Configure the node with the Mapbox API URL and API key.
3. Fulfill this step by adding another 'HTTP Request' node to use the OpenAI API for text enrichment.
4. Configure the node with the OpenAI API URL and API key.
5. Connect the nodes to the imported data to perform the enrichment.

CHECK AT THE END

[ ] Mapbox API successfully configured

[ ] OpenAI API successfully configured

[ ] Datasets successfully enriched

RESULT

Datasets are enriched with additional information.

LINKS

OpenAI API: https://platform.openai.com/docs/api-reference

3 Normalize and classify fields with AI

The data is normalized and classified using AI technologies to ensure consistency and correct categorization.

WHAT YOU NEED TO DO

1. Fulfill a 'Function' Node in n8n to transform the data.
2. Use the OpenAI API to normalize text fields.
3. Configure the 'Function' Node to add classification logic.
4. Test normalization and classification with sample data.
5. Verify the output for consistency and correctness.

CHECK AT THE END

[ ] Normalization successfully performed
[ ] Classification successfully performed
[ ] Output checked for consistency

RESULT

Data is uniformly normalized and correctly classified.

LINKS

OpenAI AI Modelle: https://platform.openai.com/docs/models

4 Detect and merge duplicates

Duplicate records are identified and merged to ensure data integrity and increase processing efficiency.

WHAT YOU NEED TO DO

1. FAdd a 'Merge' node in n8n to identify duplicate records.
2. Configure the node to search for specific fields that represent duplicates.
3. Run the workflow to recognize duplicates.
4. Manually review the merged records for accuracy.
5. Save the cleaned data for the next step.

CHECK AT THE END

[ ] Duplicates successfully identified
[ ] Records correctly merged
[ ] Data integrity ensured

RESULT

Duplicate records are successfully recognized and merged.

LINKS

Supabase Duplicate Detection: https://supabase.com/docs/guides/database

5 Version results in Supabase

The enhanced and cleaned data will be versioned in Supabase to track changes and provide the latest data.

WHAT YOU NEED TO DO

1. Create a new table in Supabase for the cleaned data.
2. Use the 'Supabase' Node in n8n to transfer the data.
3. Configure the Node with the Supabase database URL and API key.
4. Run the workflow to store the data in Supabase.
5. Verify the data in Supabase for completeness and accuracy.

CHECK AT THE END

[ ] Table created in Supabase

[ ] Data successfully stored in Supabase

[ ] Data versioning checked

RESULT

Data is successfully versioned and stored in Supabase.

LINKS

Supabase Versioning: https://supabase.com/docs/guides/database/versioning

RECOMMENDED TOOLS & RESOURCES

Recommended Tools & Resources

All tools in this workflow with their official website.

n8n

https://n8n.io/?ref=flowmatrix

Official URL

Supabase

https://supabase.com/?ref=flowmatrix

Official URL

OpenAI API

https://platform.openai.com/?ref=flowmatrix

Official URL

Mapbox Maps API

https://www.mapbox.com/

Full version as Premium PDF

All steps, tools and resources as a PDF – €4.99.

Buy now – €4.99