

**ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ**



ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS

MSc in Business Analytics

Machine Learning & Content Analytics

JiraSQL: AI-Powered Ticket and Schema Generator from Docs

Pefanis Georgios (p2822427)

Kalogirou Pavlos (p2822412)

Pavlogeorgatos Emmanouil (p2822426)

Contents

Introduction.....	3
Goals.....	3
System walkthrough.....	3
Entry and configuration.....	3
Step 1 — Structure approval loop (Agents 1–2 + Function 1)	4
Step 2 — Ticket enrichment: labels (Agent 3) and optional groups (Agent 4)	4
Step 3 — SQL/GAP branch (Agents 5–8, optional but enabled here)	5
Step 4 — Priority inference (Agent 9).....	5
Step 5 — Final “values-only” approval (Function 3)	6
Step 6 — Create Jira issues and attach artefacts (Function 4)	6
Jira walkthrough	9
Streamlit UI walkthrough	21
Appendices	32
Appendix A — Case with Group Assignment	32
Appendix B — Run without SQL.....	35
Appendix C — Requirements Document (Input)	42
Appendix D — Database Schema (Input)	45
Contribution	48

Introduction

This project delivers an AI-powered Ticket & Schema Generator from Docs: a pipeline that reads a technical specification, turns it into a clean Jira backlog (Epic → Tasks → Subtasks), runs a schema gap analysis, generates the necessary SQL artefacts, and pushes everything to Jira with human approval gates. In the executed notebook we focus on the function `run_pipeline(doc_path, schema_text, include_sql=True, include_assignee_group=False)`, which orchestrates the end-to-end flow—from parsing a .docx to creating issues and attaching SQL files.

Throughout the project we used OpenAI's GPT-4o mini as the core LLM powering all agents (document parsing, structure refinement, labeling, optional group assignment, gap analysis, and SQL generation). We chose GPT-4o mini for its strong cost-latency-quality trade-off and ran it with deterministic settings (`temperature = 0`) and constrained prompts/templates to ensure reproducible, Jira-ready outputs.

Goals

The system aims to (i) understand arbitrary requirement documents and propose a hierarchical ticket structure, (ii) enrich tickets automatically with labels and (optionally) team assignment, (iii) detect schema gaps against a provided database schema and synthesize SQL for both schema changes and reporting queries, and (iv) ship production-ready tickets to Jira with the correct parent-child relationships and any generated SQL files attached, under explicit human-in-the-loop approvals. The central orchestrator is `run_pipeline`, whose signature and default behaviour are shown in the notebook; in this run we enabled SQL and disabled assignee-group enrichment.

System walkthrough

Entry and configuration

The notebook sets the document path and loads the schema text, then calls `run_pipeline(doc_path, schema_text, include_sql=True, include_assignee_group=False)`. This configuration means the SQL branch runs (because a schema is provided), while the assignee-group step is skipped. The call and outputs are visible in the “Run the pipeline” cells.

Step 1 — Structure approval loop (Agents 1–2 + Function 1)

The pipeline starts by reading the .docx and entering a loop that uses two LLM agents: Agent 1 (Document Structure) parses the document into a proposed Epic/Task/Subtask tree; Agent 2 (Review & Improve Structure) refines that structure. The loop then asks for structured user feedback and only proceeds once you approve. In code, this appears as the while not approved: block calling `document_structure_agent(...)`, `review_and_improve_structure(...)`, and finally `get_user_feedback_and_approve(...)`. When you approve, the proposal is materialised into a staging DataFrame that becomes the working backlog. The run prints “Here is the Jira ticketing structure:” and displays the DataFrame (e.g., rows for E1, T1, U1 ...).

Here is the Jira ticketing structure:

local_id	summary	description	issue_type	parent_local_id
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic	
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1
2	U1 Conduct GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2
5	U3 Transformation Logic	Implement data transformations and standardizations.	Subtask	T2
6	U4 Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2
7	T3 SQL Reporting Layer	Build views and procedures for data presentation.	Task	E1
8	T4 Account Portfolio Overview	Create a report in tabular format with specified fields.\nFields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1
9	T5 Account Transaction History View	Create a report in tabular format with specified fields.\nFields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1
10	T6 Customer Contact Information Overview	Create a report in tabular format with specified fields.\nFields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1
11	T7 Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task	E1
12	U5 Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask	T7
13	U6 Report Layout and Navigation	Create intuitive navigation and layout for reports in Power BI.	Subtask	T7
14	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for insights.	Subtask	T7
15	T8 Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1
16	U8 Data Validation	Compare report outputs with source records for accuracy.	Subtask	T8
17	U9 Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T8
18	T9 Documentation and Handover	Provide documentation for maintenance and user guidance.	Task	E1
19	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T9
20	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T9

Are you okay with this structure? (yes/no): no
Please provide feedback on the changes you would like to make.
Your feedback: do not include task named SQL Reporting Layer

The structure was refined after feedback.

Here is the Jira ticketing structure:

local_id	summary	description	issue_type	parent_local_id
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic	
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1
2	U1 Conduct GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2
5	U3 Transformation Logic	Implement data transformations and standardizations.	Subtask	T2
6	U4 Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2
7	T3 Account Portfolio Overview	Create a report in tabular format with specified fields.\nFields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1
8	T4 Account Transaction History View	Create a report in tabular format with specified fields.\nFields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1
9	T5 Customer Contact Information Overview	Create a report in tabular format with specified fields.\nFields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1
10	T6 Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task	E1
11	U5 Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask	T6
12	U6 Report Layout and Navigation	Create intuitive navigation and layout for reports in Power BI.	Subtask	T6
13	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for insights.	Subtask	T6
14	T7 Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1
15	U8 Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7
16	U9 Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T7
17	T8 Documentation and Handover	Provide documentation for maintenance and user guidance.	Task	E1
18	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8
19	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8

Are you okay with this structure? (yes/no): yes
The process continues.

Step 2 — Ticket enrichment: labels (Agent 3) and optional groups (Agent 4)

After approval, the pipeline enriches each row. Agent 3 (Add Labels) assigns one label per ticket

from a controlled list (QA, Reporting, ETL, Data, Development, Documentation, Business) using a small, deterministic LLM prompt; see the `labels_list` and `add_labels_to_dataframe(...)` definitions. In this execution we disabled Agent 4 via `include_assignee_group=False`, so the code explicitly prints “[INFO] Skipping assignee group enrichment.” (the function `add_groups_to_dataframe(...)` remains available if you enable it).

```
⌚ Adding labels with LLM...
[INFO] Skipping assignee group enrichment.
```

Step 3 — SQL/GAP branch (Agents 5–8, optional but enabled here)

Because `include_sql=True` and `schema_text` is provided, the SQL branch runs. The pipeline scans the approved tickets for missing data fields against the supplied schema with `find_missing_fields_df_from_staging(...)` and prints a “Suggested missing fields” table for review. Based on our confirmations, it generates SQL: DDL/ALTER statements for schema gaps and SELECT statements for tabular reports with `generate_select_sql_for_reports(...)`. The notebook shows both the identification of report-like tickets and the generation of query files.

```
⌚ Scanning for missing report fields against schema...
Suggested missing fields:
missing_field_name suggested_table parent_local_id
0 CustomerNumber Dim_Customer T3
1 AccountStatus Dim_Account T3
2 TransactionNumber Fact_Transaction T4
3 IsReversed Fact_Transaction T4
4 NationalPostalCode Dim_Contact T5

Confirm which ones to append as standardized 'Add field to table' subtasks...
Does field 'CustomerNumber' already exist in the database? (yes/no): no
Do you want to add 'CustomerNumber' to table 'Dim_Customer' as a subtask? (yes/no): yes
Does field 'AccountStatus' already exist in the database? (yes/no): no
Do you want to add 'AccountStatus' to table 'Dim_Account' as a subtask? (yes/no): yes
Does field 'TransactionNumber' already exist in the database? (yes/no): no
Do you want to add 'TransactionNumber' to table 'Fact_Transaction' as a subtask? (yes/no): yes
Does field 'IsReversed' already exist in the database? (yes/no): no
Do you want to add 'IsReversed' to table 'Fact_Transaction' as a subtask? (yes/no): yes
Does field 'NationalPostalCode' already exist in the database? (yes/no): no
Do you want to add 'NationalPostalCode' to table 'Dim_Contact' as a subtask? (yes/no): yes

⌚ Generating ALTER TABLE statements (one per subtask)...
📌 Identifying tabular report tasks...
📝 Generating SELECT SQL queries for tabular reports...
```

Step 4 — Priority inference (Agent 9)

A lightweight agent computes a priority suggestion per ticket (e.g., Low/Medium/High) so that downstream triage starts from a reasonable baseline. This is performed by the `add_priority_predictions(...)` step in the pipeline just before the final approval and creation.

▲ Predicting priority for each item...

Here is the final ticketing structure:

local_id	summary	description	issue_type	parent_local_id	label	assignee_group	priority
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic		Reporting	NaN	Medium
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1	Reporting	NaN	Medium
2	U1 Conduct GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1	Business	NaN	Medium
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1	ETL	NaN	High
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2	ETL	NaN	Medium
5	U3 Transformation Logic	Implement data transformations and standardizations.	Subtask	T2	ETL	NaN	Medium
6	U4 Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2	ETL	NaN	Medium
7	T3 Account Portfolio Overview	Create a report in tabular format with specified fields.\nFields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1	Reporting	NaN	Low
8	T4 Account Transaction History View	Create a report in tabular format with specified fields.\nFields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1	Reporting	NaN	Medium
9	T5 Customer Contact Information Overview	Create a report in tabular format with specified fields.\nFields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1	Reporting	NaN	Low
10	T6 Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task	E1	Development	NaN	Medium
11	U5 Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask	T6	ETL	NaN	Medium
12	U6 Report Layout and Navigation	Create intuitive navigation and layout for reports in Power BI.	Subtask	T6	Reporting	NaN	Low
13	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for insights.	Subtask	T6	Reporting	NaN	Medium
14	T7 Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1	QA	NaN	Medium
15	U8 Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	NaN	Medium
16	U9 Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T7	QA	NaN	Medium
17	T8 Documentation and Handover	Provide documentation for maintenance and user guidance.	Task	E1	Documentation	NaN	Low
18	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8	Documentation	NaN	Low
19	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	NaN	Low
20	U12 Add field "CustomerNumber" to table "Dim_Customer"	Implement changes to add the field "CustomerNumber" into table "Dim_Customer", including ETL and database updates.	Subtask	T3	Data	Medium	Medium
21	U13 Add field "AccountStatus" to table "Dim_Account"	Implement changes to add the field "AccountStatus" into table "Dim_Account", including ETL and database updates.	Subtask	T3	Data	Medium	Medium
22	U14 Add field "TransactionNumber" to table "Fact_Transaction"	Implement changes to add the field "TransactionNumber" into table "Fact_Transaction", including ETL and database updates.	Subtask	T4	Data	Medium	Medium
23	U15 Add field "IsReversed" to table "Fact_Transaction"	Implement changes to add the field "IsReversed" into table "Fact_Transaction", including ETL and database updates.	Subtask	T4	Data	Medium	Medium
24	U16 Add field "NationalPostalCode" to table "Dim_Contact"	Implement changes to add the field "NationalPostalCode" into table "Dim_Contact", including ETL and database updates.	Subtask	T5	Data	Medium	Medium

Step 5 — Final “values-only” approval (Function 3)

Before anything is pushed to Jira, the pipeline includes a final approval gate that lets you change values (summaries, descriptions, labels, priorities, etc.) without altering the parent-child structure. This is implemented by `ask_final_approval(...)`, which clearly distinguishes structural changes (go back to Step 1) from value-level edits (proceed).

Are you okay with this final outcome? (yes/no):
Please provide feedback on the changes you would like to make (values only, not structure).
Your feedback: set all priorities to medium

local_id	summary	description	issue_type	parent_local_id	label	assignee_group	priority
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic		Reporting	NaN	Medium
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1	Reporting	NaN	Medium
2	U1 Conduct GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1	Business	NaN	Medium
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1	ETL	NaN	Medium
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2	ETL	NaN	Medium
5	U3 Transformation Logic	Implement data transformations and standardizations.	Subtask	T2	ETL	NaN	Medium
6	U4 Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2	ETL	NaN	Medium
7	T3 Account Portfolio Overview	Create a report in tabular format with specified fields.\nFields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1	Reporting	NaN	Medium
8	T4 Account Transaction History View	Create a report in tabular format with specified fields.\nFields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1	Reporting	NaN	Medium
9	T5 Customer Contact Information Overview	Create a report in tabular format with specified fields.\nFields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1	Reporting	NaN	Medium
10	T6 Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task	E1	Development	NaN	Medium
11	U5 Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask	T6	ETL	NaN	Medium
12	U6 Report Layout and Navigation	Create intuitive navigation and layout for reports in Power BI.	Subtask	T6	Reporting	NaN	Medium
13	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for insights.	Subtask	T6	Reporting	NaN	Medium
14	T7 Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1	QA	NaN	Medium
15	U8 Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	NaN	Medium
16	U9 Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T7	QA	NaN	Medium
17	T8 Documentation and Handover	Provide documentation for maintenance and user guidance.	Task	E1	Documentation	NaN	Medium
18	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8	Documentation	NaN	Medium
19	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	NaN	Medium
20	U12 Add field "CustomerNumber" to table "Dim_Customer"	Implement changes to add the field "CustomerNumber" into table "Dim_Customer", including ETL and database updates.	Subtask	T3	Data	Medium	Medium
21	U13 Add field "AccountStatus" to table "Dim_Account"	Implement changes to add the field "AccountStatus" into table "Dim_Account", including ETL and database updates.	Subtask	T3	Data	Medium	Medium
22	U14 Add field "TransactionNumber" to table "Fact_Transaction"	Implement changes to add the field "TransactionNumber" into table "Fact_Transaction", including ETL and database updates.	Subtask	T4	Data	Medium	Medium
23	U15 Add field "IsReversed" to table "Fact_Transaction"	Implement changes to add the field "IsReversed" into table "Fact_Transaction", including ETL and database updates.	Subtask	T4	Data	Medium	Medium
24	U16 Add field "NationalPostalCode" to table "Dim_Contact"	Implement changes to add the field "NationalPostalCode" into table "Dim_Contact", including ETL and database updates.	Subtask	T5	Data	Medium	Medium

Step 6 — Create Jira issues and attach artefacts (Function 4)

On approval, the function `create_jira_from_df(...)` creates the Epic, then Tasks, then Subtasks, preserving the hierarchy by mapping each row's `parent_local_id` to the Jira key that was just created. If SQL generation was enabled, the function attaches the SQL files produced in the previous step to the relevant Task/Subtask, and prints the creation log (“Epic created ... Task created ... Attached ...”). In this run you can see multiple Attachedsql → <ISSUE> lines confirming the SQL artefacts were uploaded to Jira alongside the newly created issues.

Are you okay with this final outcome? (yes/no): yes

Creating Jira issues...

Here is the Jira ticketing structure, with the extra subtasks:

local_id	summary	description	issue_type	parent_local_id	label	assignee_group	priority
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic		Reporting	NaN	Medium
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1	Reporting	NaN	Medium
2	U1 Conduct GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1	Business	NaN	Medium
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1	ETL	NaN	Medium
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2	ETL	NaN	Medium
5	U3 Transformation Logic	Implement data transformations and standardizations.	Subtask	T2	ETL	NaN	Medium
6	U4 Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2	ETL	NaN	Medium
7	T3 Account Portfolio Overview	Create a report in tabular format with specified fields.\nFields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1	Reporting	NaN	Medium
8	T4 Account Transaction History View	Create a report in tabular format with specified fields.\nFields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1	Reporting	NaN	Medium
9	T5 Customer Contact Information Overview	Create a report in tabular format with specified fields.\nFields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1	Reporting	NaN	Medium
10	T6 Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task	E1	Development	NaN	Medium
11	U5 Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask	T6	ETL	NaN	Medium
12	U6 Report Layout and Navigation	Create intuitive navigation and layout for reports in Power BI.	Subtask	T6	Reporting	NaN	Medium
13	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for insights.	Subtask	T6	Reporting	NaN	Medium
14	T7 Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1	QA	NaN	Medium
15	U8 Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	NaN	Medium
16	U9 Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T7	QA	NaN	Medium
17	T8 Documentation and Handover	Provide documentation for maintenance and user guidance.	Task	E1	Documentation	NaN	Medium
18	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8	Documentation	NaN	Medium
19	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	NaN	Medium
20	U12 Add field "CustomerNumber" to table "Dim_Customer"	Implement changes to add the field "CustomerNumber" into table "Dim_Customer", including ETL and database updates.	Subtask	T3	Data	Medium	Medium
21	U13 Add field "AccountStatus" to table "Dim_Account"	Implement changes to add the field "AccountStatus" into table "Dim_Account", including ETL and database updates.	Subtask	T3	Data	Medium	Medium
22	U14 Add field "TransactionNumber" to table "Fact_Transaction"	Implement changes to add the field "TransactionNumber" into table "Fact_Transaction", including ETL and database updates.	Subtask	T4	Data	Medium	Medium
23	U15 Add field "IsReversed" to table "Fact_Transaction"	Implement changes to add the field "IsReversed" into table "Fact_Transaction", including ETL and database updates.	Subtask	T4	Data	Medium	Medium
24	U16 Add field "NationalPostalCode" to table "Dim_Contact"	Implement changes to add the field "NationalPostalCode" into table "Dim_Contact", including ETL and database updates.	Subtask	T5	Data	Medium	Medium

Created Epic: DOC2SQL-1162

Epic created: DOC2SQL-1162

Created Task: DOC2SQL-1163

Task created: DOC2SQL-1163

Created Task: DOC2SQL-1164

Task created: DOC2SQL-1164

Attached DOC2SQL-1165_account_portfolio_overview.sql + DOC2SQL-1165

Created Task: DOC2SQL-1165

Task created: DOC2SQL-1166

Attached DOC2SQL-1166_account_transaction_history_view.sql + DOC2SQL-1166

Created Task: DOC2SQL-1167

Task created: DOC2SQL-1167

Attached DOC2SQL-1167_customer_contact_information_overview.sql + DOC2SQL-1167

Created Task: DOC2SQL-1168

Task created: DOC2SQL-1168

Attached DOC2SQL-1169

Created Task: DOC2SQL-1170

Task created: DOC2SQL-1170

Attached Subtask: DOC2SQL-1171

Created Subtask: DOC2SQL-1172

Subtask created: DOC2SQL-1172

Attached Subtask: DOC2SQL-1173

Created Subtask: DOC2SQL-1173

Subtask created: DOC2SQL-1173

Attached Subtask: DOC2SQL-1174

Created Subtask: DOC2SQL-1174

Subtask created: DOC2SQL-1174

Attached Subtask: DOC2SQL-1175

Created Subtask: DOC2SQL-1175

Subtask created: DOC2SQL-1176

Subtask created: DOC2SQL-1176

Attached Subtask: DOC2SQL-1177

Subtask created: DOC2SQL-1177

Attached Subtask: DOC2SQL-1178

Subtask created: DOC2SQL-1178

Attached Subtask: DOC2SQL-1179

Created Subtask: DOC2SQL-1179

Subtask created: DOC2SQL-1180

Attached Subtask: DOC2SQL-1180

Created Subtask: DOC2SQL-1181

Subtask created: DOC2SQL-1181

Attached Subtask: DOC2SQL-1182

Created Subtask: DOC2SQL-1182

Subtask created: DOC2SQL-1182

Attached DOC2SQL-1182_add_field_customernumber_to_table_dim_customer.sql + DOC2SQL-1182

Created Subtask: DOC2SQL-1183

Subtask created: DOC2SQL-1183

Attached DOC2SQL-1183_add_field_accountstatus_to_table_dim_account.sql + DOC2SQL-1183

Created Subtask: DOC2SQL-1184

Subtask created: DOC2SQL-1184

Attached DOC2SQL-1184_add_field_transactionnumber_to_table_fact_transaction.sql + DOC2SQL-1184

Created Subtask: DOC2SQL-1185

Subtask created: DOC2SQL-1185

Attached DOC2SQL-1185_add_field_isreversed_to_table_fact_transaction.sql + DOC2SQL-1185

Created Subtask: DOC2SQL-1186

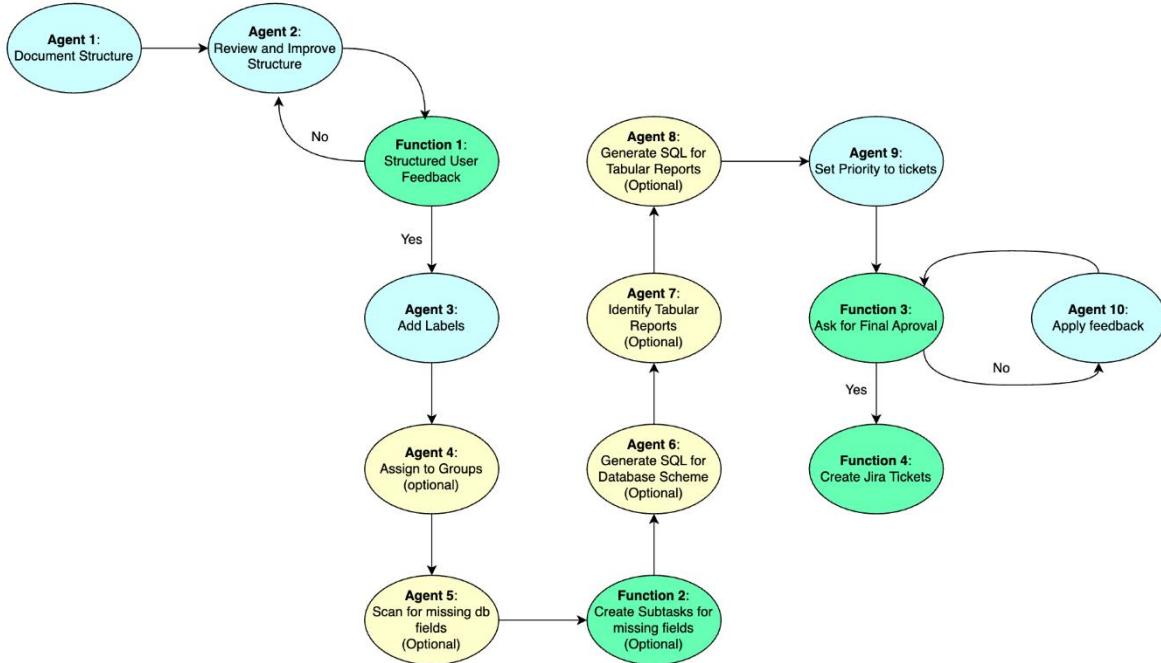
Subtask created: DOC2SQL-1186

Attached DOC2SQL-1186_add_field_nationalpostalcode_to_table_dim_contact.sql + DOC2SQL-1186

End-to-end, at a glance:

the pipeline (1) parses and iteratively improves a ticket hierarchy with your feedback until approved; (2) enriches tickets with labels (and optionally team groups); (3) performs schema gap analysis and synthesizes SQL for both schema updates and reporting queries; (4) infers priorities

and requests a final values-only approval; and (5) creates the Epic/Tasks/Subtasks in Jira and attaches the generated SQL files so that engineering can start work immediately.



This diagram summarizes the end-to-end, human-in-the-loop pipeline from a requirements document to Jira tickets. Blue ovals are LLM agents, yellow ovals are optional SQL/data agents, and green ovals are functions/human gates. The flow proceeds left-to-right: Agent 1 derives an initial Epic→Task→Subtask outline, Agent 2 refines it, and Function 1 collects structured feedback; a “No” loops back to Agent 2, while “Yes” advances the run. The approved plan is enriched by Agent 3 with labels, then Agent 4 can optionally assign delivery groups depending on configuration.

The center branch shows the optional data/SQL track. Agent 5 scans the approved items for missing database fields; Function 2 turns confirmed gaps into standardized “Add field ...” subtasks; Agent 6 generates DDL for those gaps. In parallel, Agent 7 flags tabular-report tickets and Agent 8 produces their SELECT queries. These artifacts are later attached to the corresponding tickets.

On the right, Agent 9 sets a priority for each item, after which Function 3 runs the final approval. If the reviewer requests changes, Agent 10 applies the value-level edits and returns to the gate; once approved, Function 4 creates the Jira issues in Epic→Task→Subtask order and uploads any generated SQL. Overall, the diagram makes explicit the two approval loops, the configurability of group assignment, and the independence of the optional SQL branch from the core ticketing flow.

Jira walkthrough

This section shows how the AI pipeline's output materializes inside Jira. Starting from an approved staging table, the system creates an Epic, its Tasks, and Subtasks in the correct parent-child order, populates each issue's Summary and Description, applies metadata (Labels, Priority, and—when enabled—Assigned Group), and attaches SQL artifacts where applicable. You'll see two kinds of tickets side-by-side: (i) reporting Tasks (no DDL; may include a SELECT when generated) and (ii) standardized "Add field ..." Subtasks with an attached ALTER TABLE ... script. For traceability, the app also returns a JSON map from the local IDs (e.g., E1, T3, U12) to the actual Jira keys (e.g., DOC2SQL-1291). The screenshots that follow were produced from a run with SQL generation on and group assignment disabled, so you can observe hierarchy, routing, and attachments exactly as they appear to end users in Jira.

Epic Data Warehouse and Reporting Specification

The image contains two screenshots of the Jira interface, illustrating the creation of an Epic and its child work items.

Screenshot 1: Epic Creation

This screenshot shows the Jira project navigation sidebar on the left, with the "DOC2SQL" project selected. The main area displays a single epic ticket titled "Data Warehouse and Reporting Specification". The ticket details are as follows:

Type	Key	Summary	Status	Comments	Assignee	Due date	Priority	Labels	Created	Updated	Reporter
✓	DOC2SQL-1162	Data Warehouse and Reporting Specification	To Do	Add comment			Medium	reporting	11 Sept 2025	11 Sept 2025	George Pefanis

Screenshot 2: Epic Details and Child Work Items

This screenshot shows the detailed view of the epic ticket. The "Description" field contains the text: "Comprehensive plan for data warehouse and reporting development." The "Child work items" section lists several tasks and subtasks, each with a progress bar indicating completion status. The tasks include:

- DOC2SQL-1163 Reporting Requirements Analysis (0% Done)
- DOC2SQL-1164 ETL Development (0% Done)
- DOC2SQL-1165 Account Portfolio Overview (0% Done)
- DOC2SQL-1166 Account Transaction History View (0% Done)
- DOC2SQL-1167 Customer Contact Information Overview (0% Done)
- DOC2SQL-1168 Power BI / Visualization Layer (0% Done)
- DOC2SQL-1169 Quality Assurance and Testing (0% Done)
- DOC2SQL-1170 Documentation and Handover (0% Done)

The right panel displays the "Details" and "To Do" tabs for the epic, showing fields like Assignee, Labels, Due date, Reporter, and Development status.

Task Reporting Requirements Analysis

The screenshot shows a Jira work item titled "DOC2SQL-1162 / DOC2SQL-1163". The work item is named "Reporting Requirements Analysis" and has a priority of Medium. The description is "Review reporting needs to ensure accurate data delivery." A child work item "DOC2SQL-1171 Conduct GAP Analysis" is listed under "Child work items". The "Details" panel shows the following information:

- Assignee: Unassigned (with a link to "Assign to me")
- Assigned Group: None
- Labels: reporting
- Parent: DOC2SQL-1162 Data Warehouse
- Due date: Add due date
- Team: Add team
- Start date: Add date
- Reporter: George Pefanis
- Development: Create branch, Create commit

The "Comments" tab is selected in the activity bar.

Subtask Conduct GAP Analysis

The screenshot shows a Jira work item titled "DOC2SQL-1163 / DOC2SQL-1171". The work item is named "Conduct GAP Analysis" and has a priority of Medium. The description is "Compare fields against the database schema and identify missing fields." The "Details" panel shows the following information:

- Assignee: Unassigned (with a link to "Assign to me")
- Assigned Group: None
- Labels: business
- Parent: DOC2SQL-1163 Reporting Req.
- Due date: Add due date
- Start date: Add date
- Reporter: George Pefanis
- Development: Create branch, Create commit

The "Comments" tab is selected in the activity bar.

Task ETL Development

The screenshot shows the Jira Work Item details for a task named "ETL Development". The work item key is DOC2SQL-1162, and it is assigned to George Pefanis. The priority is Medium. The description is "Update ETL processes for accurate data extraction and loading." The work section lists three child work items: "Source System Extraction" (key DOC2SQL-1172), "Transformation Logic" (key DOC2SQL-1173), and "Data Loading" (key DOC2SQL-1174). The status for all three children is To Do. The activity panel shows a comment from George Pefanis: "Looks good!".

Subtask Source System Extraction

The screenshot shows the Jira Work Item details for a subtask named "Source System Extraction". The work item key is DOC2SQL-1172, and it is assigned to George Pefanis. The priority is Medium. The description is "Identify source systems and develop extraction processes." The activity panel shows a comment from George Pefanis: "Looks good!".

Subtask Transformation Logic

The screenshot shows the Jira Work Item details for a task titled "Transformation Logic". The task has a priority of Medium and a description of "Implement data transformations and standardizations." The activity tab is selected, showing a comment input field and several reaction icons: "Looks good!", "Need help?", "This is blocked...", "Can you clarify...?", and "Tr". The right panel displays the "Details" section with fields for Assignee (Unassigned), Assigned Group (None), Labels (etl), Parent (DOC2SQL-1164 ETL Developm), Due date (Add due date), Start date (Add date), Reporter (George Pefanis), Development (Create branch, Create commit), and Automation (Rule executions). The status bar at the bottom indicates the work item was created and updated 3 days ago.

Subtask Data Loading

The screenshot shows the Jira Work Item details for a task titled "Data Loading". The task has a priority of Medium and a description of "Integrate transformed data into warehouse tables." The activity tab is selected, showing a comment input field and several reaction icons: "Looks good!", "Need help?", "This is blocked...", "Can you clarify...?", and "Tr". The right panel displays the "Details" section with fields for Assignee (Unassigned), Assigned Group (None), Labels (etl), Parent (DOC2SQL-1164 ETL Developm), Due date (Add due date), Start date (Add date), Reporter (George Pefanis), Development (Create branch, Create commit), and Automation (Rule executions). The status bar at the bottom indicates the work item was created and updated 3 days ago.

Task Account Portfolio Overview

Account Portfolio Overview

Priority: Medium

Description: Create a report in tabular format with specified fields.

Fields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status

Attachments: DOC2SQL_new.sql (3 days ago)

Child work items:

- DOC2SQL-1182 Add field "CustomerNumber" to table (Priority: Medium, Due date: 19.09.2025)
- DOC2SQL-1183 Add field "CustomerStatus" to table (Priority: Medium, Due date: 19.09.2025)

Linked work items: Add linked work item

Activity: Add a comment... (Comments, History, Work log)

SELECT Query for Reporting

```
DOC2SQL-1185.account_portfolio_overview.sql
sq - 342 B

1 SELECT
2     a.AccountNumber,
3     a.CustomerID AS CustomerNumber,
4     c.LastName,
5     c.FirstName,
6     a.AccountType AS TypeOfAccount,
7     CASE
8         WHEN a.CloseDate IS NULL THEN 'Active'
9         ELSE 'Inactive'
10    END AS AccountStatus
11 FROM
12     Dim_Account a
13 JOIN
14     Dim_Customers c ON a.CustomerID = c.CustomerID;
15
```

Subtask Add field "CustomerNumber" to table "Dim_Customer"

Add field "CustomerNumber" to table "Dim_Customer"

Priority: Medium

Description: Implement changes to add the field "CustomerNumber" into table "Dim_Customer", including ETL and database updates.

Attachments: DOC2SQL_new.sql (3 days ago)

Linked work items: Add linked work item

Activity: Add a comment... (Comments, History, Work log)

ALTER TABLE Query for Schema update

```
DOC2SQL-1182.add_field_customernumber_to_table_dim_customer.sql
sq - 63 B

1 ALTER TABLE Dim_Customers ADD CustomerNumber VARCHAR(50) NULL;
2
```

Task Account Transaction History View

Account Transaction History View

Priority: Medium

Description: Create a report in tabular format with specified fields.

Fields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, IsReversed

Attachments: DOC2SQL_low.sql (3 days ago)

Child work items:

- DOC2SQL-1184 Add field "TransactionNumber" to table Fact_Transaction (M 10:00)
- DOC2SQL-1185 Add field "IsReversed" to table Fact_Transaction (M 10:00)

Linked work items: Add linked work item

Activity: Add a comment... (Comments, History, Work log)

Automation: Rule executions

Created: 11 Sept 2025 | Updated: 11 Sept 2025 | Reporter: George Pefanis

SELECT Query for Reporting

```

1 SELECT
2     a.AccountNumber,
3     t.TransactionID AS TransactionNumber,
4     t.TransactionDate,
5     t.TransactionType,
6     t.Amount,
7     t.IsReversed
8 FROM
9     Fact_Transaction t
10 JOIN
11     Dim_Account a ON t.AccountID = a.AccountID;
12

```

Subtask Add field “TransactionNumber” to table “Fact_Transcation”

Add field "TransactionNumber" to table "Fact_Transcation"

Priority: Medium

Description: Implement changes to add the field "TransactionNumber" into table "Fact_Transcation", including ETL and database updates.

Attachments: DOC2SQL_low.sql (3 days ago)

Linked work items: Add linked work item

Activity: All (Comments, History, Work log)

Automation: Rule executions

Created: 11 Sept 2025 | Updated: 11 Sept 2025 | Reporter: George Pefanis

ALTER TABLE Query for Schema update

```

1 ALTER TABLE Fact_Transcation ADD TransactionNumber VARCHAR(100) NULL;
2

```

Subtask Add field “IsReversed” to table “Fact_Transcation”

The screenshot shows a Jira interface for a subtask under a project named 'DOC2SQL'. The subtask title is 'Add field "IsReversed" to table "Fact_Transcation"'. The priority is set to Medium. The due date is listed as '11 Sept 2025'. There is an attachment named 'DOC2SQL...ion.sql' from 3 days ago. The comment section contains a message: 'Looks good!' with a green checkmark icon, 'Need help?' with a yellow question mark icon, 'This is blocked...' with a red exclamation mark icon, and 'Can you clarify...?' with a blue info icon. A 'Automation' section is also visible.

ALTER TABLE Query for Schema update

The screenshot shows a terminal window with the command: 'ALTER TABLE Fact_Transcation ADD IsReversed BOOLEAN DEFAULT NULL;'. The session ID is 678. The command was run at approximately 07:08.

Task Customer Contact Information Overview

The screenshot shows a Jira interface for a task titled 'Customer Contact Information Overview'. The priority is set to Medium. The due date is listed as '11 Sept 2025'. A child work item titled 'Add field "NationalPostalCode" to table "Dim_Contact"' is listed under 'Child work items'. The comment section contains a message: 'Looks good!' with a green checkmark icon, 'Need help?' with a yellow question mark icon, 'This is blocked...' with a red exclamation mark icon, and 'Can you clarify...?' with a blue info icon. A 'Automation' section is also visible.

SELECT Query for Reporting

The screenshot shows a terminal window with the following SQL query:

```

1: SELECT
2:   a.AccountNumber AS CustomerNumber,
3:   c.LastName AS CustomerLastName,
4:   c.FirstName AS CustomerFirstName,
5:   co.ContactValue AS Address,
6:   co.ContactValue AS Country,
7:   co.ContactValue AS NationalPostalCode,
8:   co.IsPrimary AS IsPrimaryContact
9: FROM
10:   Dim_Customers c
11:   JOIN
12:     Dim_Account a ON c.CustomerID = a.CustomerID
13:   JOIN
14:     Dim_Contact co ON c.CustomerID = co.CustomerID
15: WHERE
16:   co.ContactType = 'Address';
17:

```

Subtask Add field “NationalPostalCode” to table “Dim_Contact”

The screenshot shows the Jira interface for creating a new issue. The project is set to 'DOC2SQL'. The issue type is 'Subtask' and the key is 'DOC2SQL-1182'. The title is 'Add field “NationalPostalCode” to table “Dim_Contact”'. Priority is set to 'Medium'. The description states: 'Implement changes to add the field “NationalPostalCode” into table “Dim_Contact”, including ETL and database updates.' An attachment named 'DOC2SQL-1182.xlsx' is attached. The 'Linked work items' section shows 'Add linked work item'. The 'Activity' section includes tabs for 'All', 'Comments', 'History', and 'Work log'. A comment box at the bottom says 'Looks good!'.

ALTER TABLE Query for Schema update

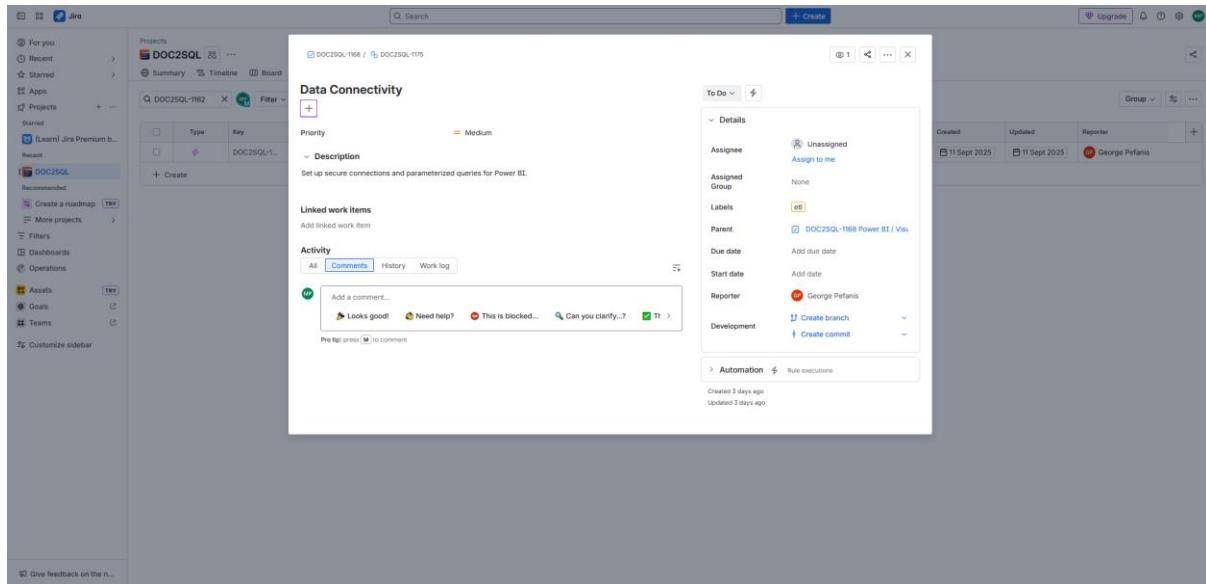
The screenshot shows a terminal window with the following command:

```
DOC2SQL-1186.add_field_nationalpostalcode_to_table_dim_contact.sql
sql -E B
1 ALTER TABLE Dim_Contact ADD NationalPostalCode VARCHAR(20) NULL;
```

Task Power BI / Visualization Layer

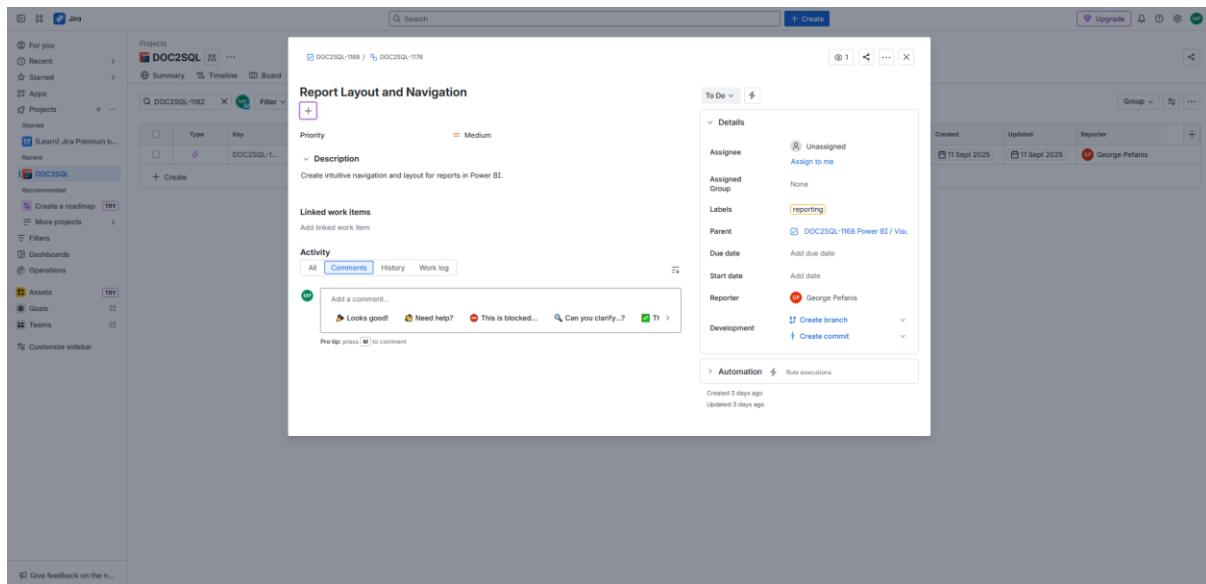
The screenshot shows the Jira interface for creating a new issue. The project is set to 'DOC2SQL'. The issue type is 'Task' and the key is 'DOC2SQL-1162'. The title is 'Power BI / Visualization Layer'. Priority is set to 'Medium'. The description states: 'Develop interactive dashboards for data visualization.' The 'Child work items' section lists three items: 'DOC2SQL-1175 Data Connectivity', 'DOC2SQL-1176 Report Layout and Navigation', and 'DOC2SQL-1177 KPIs and Summary Metrics'. The 'Linked work items' section shows 'Add linked work item'. The 'Activity' section includes tabs for 'All', 'Comments', 'History', and 'Work log'. A comment box at the bottom says 'Looks good!'.

Subtask Data Connectivity



This screenshot shows the Jira interface for a subtask titled "Data Connectivity". The sidebar on the left lists various project management options like "Recent", "Starred", and "Projects". The main content area displays the subtask details, including its priority (Medium), description ("Set up secure connections and parameterized queries for Power BI."), and activity feed. The "Details" panel on the right shows fields such as Assignee (Unassigned), Assigned Group (None), Labels (Power BI), Parent (DOC2SQL-1168 Power BI / Vis.), Due date (Add due date), Start date (Add date), Reporter (George Pefanis), Development (Create branch, Create commit), and Automation (Rule executions). The status bar at the bottom indicates the subtask was created and updated 3 days ago.

Subtask Report Layout and Navigation



This screenshot shows the Jira interface for a subtask titled "Report Layout and Navigation". The sidebar and main content area are identical to the previous screenshot, showing the subtask details and activity feed. The "Details" panel on the right shows fields such as Assignee (Unassigned), Assigned Group (None), Labels (reporting), Parent (DOC2SQL-1168 Power BI / Vis.), Due date (Add due date), Start date (Add date), Reporter (George Pefanis), Development (Create branch, Create commit), and Automation (Rule executions). The status bar at the bottom indicates the subtask was created and updated 3 days ago.

Subtask KPIs and Summary Metrics

KPIs and Summary Metrics

Priority: Medium

Description: Add calculated metrics and visualizations for insights.

Linked work items: DOC2SQL-1978, DOC2SQL-1979

To Do

Assignee: Unassigned

Assigned Group: None

Labels: reporting

Parent: DOC2SQL-198 Power BI / Visu.

Due date: Add due date

Start date: Add date

Reporter: George Pefanis

Development: Create branch, Create commit

Automation: Rule executions

Created: 11 Sept 2025 | Updated: 11 Sept 2025 | Reporter: George Pefanis

Task Quality Assurance and Testing

Quality Assurance and Testing

Priority: Medium

Description: Ensure accuracy and performance of data and reports.

Child work items:

- DOC2SQL-1178 Data Validation
- DOC2SQL-1179 Performance Testing

To Do

Assignee: Unassigned

Assigned Group: None

Labels: qa

Parent: DOC2SQL-196 Data Warehouse

Due date: Add due date

Team: Add team

Start date: Add date

Reporter: George Pefanis

Development: Create branch, Create commit

Automation: Rule executions

Created: 11 Sept 2025 | Updated: 11 Sept 2025 | Reporter: George Pefanis

Subtask Data Validation

The screenshot shows a Jira Software interface with the following details:

- Project:** DOC2SQL
- Subtask Key:** DOC2SQL-1169
- Title:** Data Validation
- Priority:** Medium
- Description:** Compare report outputs with source records for accuracy.
- Linked work items:** DOC2SQL-1178
- Activity:** Comments tab selected, showing a comment from George Pefanis: "Looks good!"
- Details:** Unassigned, None, q3, DOC2SQL-1169 Quality Assur, George Pefanis, Create branch, Create commit.
- Automation:** Rule executions.

Subtask Performance Testing

The screenshot shows a Jira Software interface with the following details:

- Project:** DOC2SQL
- Subtask Key:** DOC2SQL-1169
- Title:** Performance Testing
- Priority:** Medium
- Description:** Conduct stress tests on SQL queries and dashboards.
- Linked work items:** DOC2SQL-1178
- Activity:** Comments tab selected, showing a comment from George Pefanis: "Looks good!"
- Details:** Unassigned, None, q3, DOC2SQL-1169 Quality Assur, George Pefanis, Create branch, Create commit.
- Automation:** Rule executions.

Task Documentation and Handover

The screenshot shows a Jira task management interface. The main window displays a task titled "Documentation and Handover" under project "DOC2SQL". The task details include:

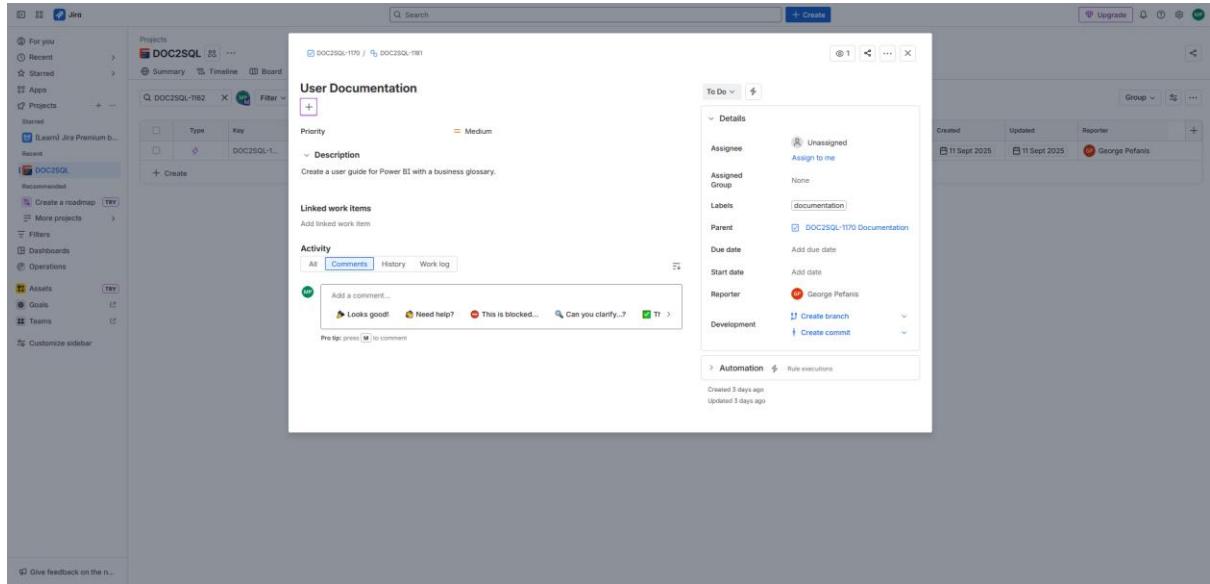
- Priority:** Medium
- Description:** Provide documentation for maintenance and user guidance.
- Child work items:** Two items listed: "DOC2SQL-1180 Technical Documentation" and "DOC2SQL-1181 User Documentation", both due at 10:00.
- Linked work items:** Add linked work item.
- Activity:** Options for All, Comments, History, and Work log.
- Comments:** A comment box with a green "Looks good!" button and other interaction options.
- Details panel:** Shows fields like Assignee (Unassigned), Assigned Group (None), Labels (documentation), Parent (DOC2SQL-1182 Data Warehouse), Due date (Add due date), Team (Add team), Start date (Add date), Reporter (George Pefanis), Development (Create branch, Create commit), and Automation (Rule executions).
- Timeline:** Shows creation and update dates (11 Sept 2025) and reporter (George Pefanis).

Subtask Technical Documentation

The screenshot shows a Jira task management interface. The main window displays a task titled "Technical Documentation" under project "DOC2SQL". The task details include:

- Priority:** Medium
- Description:** Update diagrams and document ETL job flows and SQL views.
- Linked work items:** Add linked work item.
- Activity:** Options for All, Comments, History, and Work log.
- Comments:** A comment box with a green "Looks good!" button and other interaction options.
- Details panel:** Shows fields like Assignee (Unassigned), Assigned Group (None), Labels (documentation), Parent (DOC2SQL-1170 Documentation), Due date (Add due date), Start date (Add date), Reporter (George Pefanis), Development (Create branch, Create commit), and Automation (Rule executions).
- Timeline:** Shows creation and update dates (11 Sept 2025) and reporter (George Pefanis).

Subtask User Documentation



Streamlit UI walkthrough

As part of our effort to make the application more user-friendly, we built a Streamlit front end that wraps the entire document → Jira pipeline into a guided, point-and-click experience. After authenticating to Jira and OpenAI, the user uploads the requirements document and (optionally) a database schema, selects run options (SQL branch, Assignee Group prediction, dry run, hierarchy tree direction), and launches the agents. The app then walks through the same human-in-the-loop flow as the notebook: it presents the proposed Epic → Task → Subtask structure for approval, shows the normalized staging table, and—when a schema is provided—runs gap analysis with per-field decisions and appends standardized “Add field ...” subtasks. A built-in graph view renders the hierarchy for quick visual validation. Finally, a values-only approval gate allows edits to labels, groups, priorities, and text before the user triggers Create Jira Issues; on success, the UI displays a clear ID map from local row IDs to actual Jira keys and attaches any generated SQL to the relevant tickets.

The process step by step:

The Streamlit app's login screen collects all required connections: Jira base URL and project key, user email and API token, plus an optional Assigned Group custom field id (for group routing). The bottom section captures the OpenAI API key. After completing these fields, the user clicks Login to authenticate and enable ticket creation and LLM-powered steps.

The image shows a Streamlit application window titled "Login".

Jira Credentials

- Jira Base URL (<https://<org>.atlassian.net>)
https://mlprojectworkspace.atlassian.net/
- Jira Project Key
DOC2SQL
- Jira Email
George.pef21@gmail.com
- Jira API Token
..... (redacted)

Optional

- Assigned Group custom field id (optional, e.g. customfield_10124)
..... (redacted)

OpenAI

- OpenAI API Key
..... (redacted)

Login

After login, the user uploads the project document (DOCX) and the database schema (TXT), and enables Include SQL (gap analysis + report SELECTs). The Predict Assignee Group (optional) toggle is also turned on so the pipeline assigns each ticket to a delivery team. The Show hierarchy tree option will render the interactive graph after parsing; pressing Run Agents starts the end-to-end run with these settings.

AI Agents → Jira Pipeline App

Welcome, [George.pef21@gmail.com](#) 

[Reset Process](#)

Upload Project Document (.docx)

 Drag and drop file here
Limit 200MB per file • DOCX

 Technical_Document_01.docx 21.4KB 

Upload Database Schema (.txt)

 Drag and drop file here
Limit 200MB per file • TXT

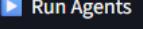
 Database_Schema.txt 2.6KB 

Include SQL (missing fields + report SELECTs)
 Predict Assignee Group (optional)
 Dry Run (only show payloads, don't call Jira)

Tree Layout Direction

UD 

Show hierarchy tree

 Run Agents

The app presents the proposed Jira hierarchy as a structured table and offers an explicit Yes/No decision. The Show Structure Text toggle exposes the raw, templated outline for full transparency. Choosing Yes locks the hierarchy for downstream enrichment; choosing No triggers the feedback loop to revise titles, grouping, or scope before continuing.

Agents finished. Review structure below.

Generated Structure ↗

› Show Structure Text

	local_id	summary	description
0	E1	Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting requirements.
1	T1	Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.
2	U1	GAP Analysis	Compare fields against the database schema and identify gaps.
3	T2	ETL Development	Update ETL processes for accurate data extraction.
4	U2	Source System Extraction	Identify source systems and develop extraction procedures.
5	U3	Transformation Logic	Implement data transformations and standardize formats.
6	U4	Data Loading	Integrate transformed data into warehouse tables.
7	T3	SQL Reporting Layer	Build views and procedures for data presentation.
8	T4	Account Portfolio Overview	Create a report in tabular format with specified fields.
9	T5	Account Transaction History View	Create a report in tabular format with specified fields.

Are you okay with this structure?

(Please choose)
 Yes
 No

👉 Please confirm if you are okay with the structure.

Expanding the section reveals the normalized staging DataFrame that will drive all subsequent steps. Each row represents an Epic, Task, or Subtask with local_id, summary, description, issue_type, and parent_local_id. Note how report Tasks (e.g., T4/T5) include explicit Fields: lists in the description, while Subtasks (e.g., U1–U11) hang from their parents. This topologically ordered table is the single source of truth for labeling, (optional) group assignment, SQL generation, and Jira creation.

local_id	summary	description	issue_type	parent_local_id
0 E1	Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic	
1 T1	Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1
2 U1	GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1
3 T2	ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1
4 U2	Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2
5 U3	Transformation Logic	Implement data transformations and standardize formats.	Subtask	T2
6 U4	Data Loading	Integrate transformed data into warehouse tables with incremental loading.	Subtask	T2
7 T3	SQL Reporting Layer	Build views and procedures for data presentation.	Task	E1
8 T4	Account Portfolio Overview	Create a report in tabular format with specified fields. Fields: Account Number, Customer Number, Cus...	Task	E1
9 T5	Account Transaction History View	Create a report in tabular format with specified fields. Fields: Account Number, Transaction Number, Tr...	Task	E1
10 T6	Customer Contact Information Overview	Create a report in tabular format with specified fields. Fields: Customer Number, Customer Last Name, Task...	Task	E1
11 T7	Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task	E1
12 U5	Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask	T7
13 U6	Report Layout and Navigation	Create intuitive tabs for account data and transaction history.	Subtask	T7
14 U7	KPIs and Summary Metrics	Add calculated metrics and visualizations for data insights.	Subtask	T7
15 T8	Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1
16 U8	Data Validation	Compare report outputs with source records for accuracy.	Subtask	T8
17 U9	Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T8
18 T9	Documentation and Handover	Provide comprehensive documentation for users and technical teams.	Task	E1
19 U10	Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T9
20 U11	User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T9

Here the reviewer selects No and provides concrete guidance (“Do not include SQL Reporting Layer”). Pressing Re-run Review Agent triggers the review agent to re-apply rules and regenerate the outline accordingly; the updated staging table will then be shown again for approval.

Show Structure Text

	local_id	summary	description
	11 T7	Power BI / Visualization Layer	Develop interactive dashboards for data visualization.
	12 U5	Data Connectivity	Set up secure connections and parameterized queries for Power BI.
	13 U6	Report Layout and Navigation	Create intuitive tabs for account data and transaction history.
	14 U7	KPIs and Summary Metrics	Add calculated metrics and visualizations for data insights.
	15 T8	Quality Assurance and Testing	Ensure accuracy and performance of data and reports.
	16 U8	Data Validation	Compare report outputs with source records for accuracy.
	17 U9	Performance Testing	Conduct stress tests on SQL queries and dashboards.
	18 T9	Documentation and Handover	Provide comprehensive documentation for users and technical teams.
	19 U10	Technical Documentation	Update diagrams and document ETL job flows and SQL views.
	20 U11	User Documentation	Create a user guide for Power BI with a business glossary.

Are you okay with this structure?

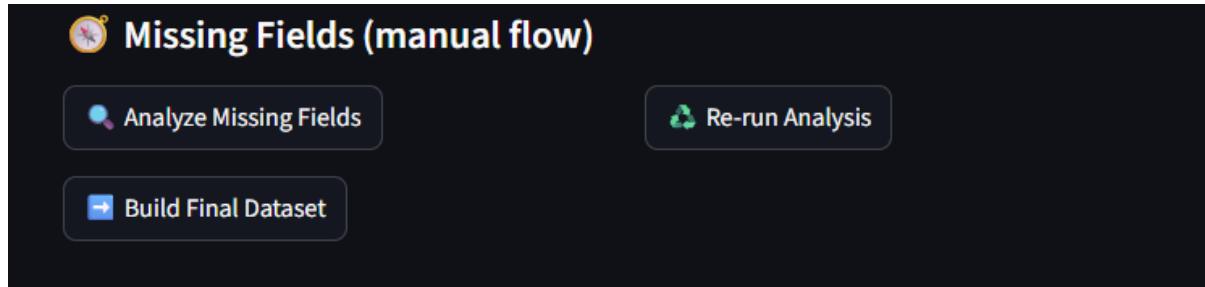
(Please choose)
 Yes
 No

Provide feedback for refinement

Do not include SQL Reporting Layer

[Re-run Review Agent](#)

This panel exposes the gap-analysis controls. Analyze Missing Fields runs the comparison of requested report fields against the uploaded schema and shows the candidate gaps; Re-run Analysis repeats the check after any edits/approvals; Build Final Dataset merges the confirmed gaps back into the staging table (creating standardized “Add field ...” subtasks) to prepare for DDL/SELECT generation and the final run.



After running Analyze Missing Fields, the app reports “Analysis complete” and lists candidate gaps with the suggested table and parent local_id. The selector lets the reviewer choose which fields to convert into standardized “Add field ...” subtasks. Clicking Append Selected Fields creates those subtasks, and Build Final Dataset merges them back into the staging table to prepare for DDL/SELECT generation.

A screenshot of the same dark-themed interface after the "Analyze Missing Fields" button was clicked. A green banner at the top states "Analysis complete. Found 4 candidate fields." Below this is a table with four rows of data:

	missing_field_name	suggested_table	parent_local_id
0	CustomerNumber	Dim_Customer	T3
1	AccountStatus	Dim_Account	T3
2	TransactionNumber	Fact_Transaction	T4
3	NationalPostalCode	Dim_Contact	T5

Select which fields to review/append as subtasks

Choose options

+ Append Selected Fields

Build Final Dataset

The interface includes a "Choose options" dropdown menu above the "Append Selected Fields" button, and a "Build Final Dataset" button at the bottom.

For each candidate gap, the reviewer explicitly decides whether the field already exists in the database and whether to append it as an “Add field ...” subtask to the suggested table. This guardrail prevents false positives and lets the user control which items become actionable work before clicking Append Selected Fields and Build Final Dataset.

Decisions per field

Does field 'CustomerNumber' already exist in the database?

(choose) ▾

Add 'CustomerNumber' to table 'Dim_Customer' as a subtask?

(choose) ▾

+ Append Selected Fields

➡ Build Final Dataset

Decisions per field

Does field 'CustomerNumber' already exist in the database?

(choose) ▾

(choose)

Yes

No

Add 'CustomerNumber' to table 'Dim_Customer' as a subtask?

(choose) ▾

(choose)

Yes

No

Here the reviewer selects CustomerNumber, marks “Does it already exist?” → No, and chooses “Add … as a subtask?” → Yes. Clicking Append Selected Fields creates the standardized “Add column CustomerNumber to Dim_Customer” subtask; Build Final Dataset then merges it into the staging table, ready for DDL/SELECT generation downstream.

Select which fields to review/append as subtasks

CustomerNumber ×

Decisions per field

Does field 'CustomerNumber' already exist in the database?

No ▾

Add 'CustomerNumber' to table 'Dim_Customer' as a subtask?

Yes ▾

+ Append Selected Fields

➡ Build Final Dataset

+ Append Selected Fields

Subtasks appended: 1. Now build the Final Dataset.

➡ Build Final Dataset

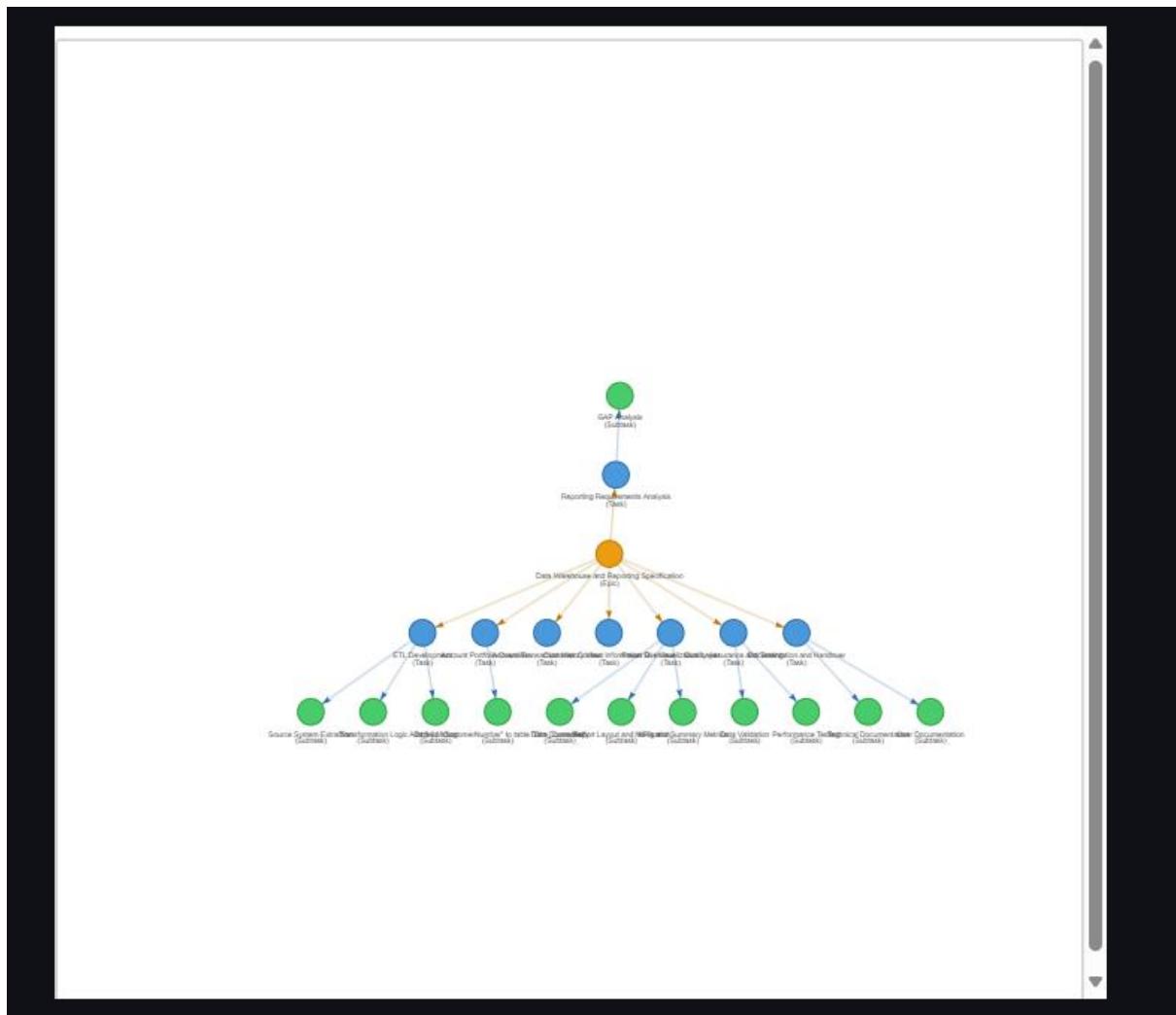
The final staging table after appending the gap as a standardized subtask (e.g., U12: “Add field ‘CustomerNumber’ to ‘Dim_Customer’”).

	local_id	summary	description
11	U5	Data Connectivity	Set up secure connections and parameterize queries.
12	U6	Report Layout and Navigation	Create intuitive tabs for account data and navigation.
13	U7	KPIs and Summary Metrics	Add calculated metrics and visualizations.
14	T7	Quality Assurance and Testing	Ensure accuracy and performance of data and reports.
15	U8	Data Validation	Compare report outputs with source records.
16	U9	Performance Testing	Conduct stress tests on SQL queries and data flows.
17	T8	Documentation and Handover	Provide comprehensive documentation for the solution.
18	U10	Technical Documentation	Update diagrams and document ETL job flows.
19	U11	User Documentation	Create a user guide for Power BI with a business glossary.
20	U12	Add field "CustomerNumber" to table "Dim_Customer"	Implement changes to add the field "CustomerNumber" to the table "Dim_Customer".

The same table expanded to reveal the operational columns used downstream: issue_type, parent_local_id, label (Agent 3), assignee_group (Agent 4, if enabled), and priority (Agent 9). This is the version the pipeline hands off to SQL generation, final approval, and Jira creation.

local_id	summary	description	issue_type	parent_local_id	label	assignee_group	priority	
0	E1	Data Warehouse and Reporting Specification	Epic		Reporting	Data Warehouse Team	Medium	
1	T1	Reporting Requirements Analysis	Task	E1	Reporting	Business Analysis Team	Medium	
2	U1	Conduct GAP Analysis	Subtask	T1	Documentation	Business Analysis Team	Medium	
3	T2	ETL Development	Task	E1	ETL	Data Warehouse Team	High	
4	U2	Source System Extraction	Subtask	T2	ETL	Data Warehouse Team	Medium	
5	U3	Transformation Logic	Subtask	T2	ETL	Data Warehouse Team	Medium	
6	U4	Data Loading	Subtask	T2	ETL	Data Warehouse Team	Medium	
7	T3	Account Portfolio Overview	Create a report in tabular format with specified fields: Account Number Customer Number City.	Task	E1	Reporting	Business Intelligence Team	Medium
8	T4	Account Transaction History View	Create a report in tabular format with specified fields: Account Number Transaction Number.	Task	E1	Reporting	Business Intelligence Team	Medium
9	T5	Customer Contact Information Overview	Create a report in tabular format with specified fields: Customer Number Customer Last Name.	Task	E1	Reporting	Business Intelligence Team	Medium
10	T6	Power BI / Visualization Layer	Develop dashboards for interactive data visualization.	Task	E1	Development	Business Intelligence Team	High
11	U5	Data Connectivity	Set up secure connections and parameterized queries.	Subtask	T6	ETL	Data Warehouse Team	High
12	U6	Report Layout and Navigation	Create intuitive navigation and layout for reports.	Subtask	T6	Reporting	Business Intelligence Team	Medium
13	U7	KPIs and Summary Metrics	Add calculated metrics and clear visualizations.	Subtask	T6	Reporting	Business Intelligence Team	Medium
14	T7	Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1	QA	Quality Assurance Team	High
15	U8	Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	Quality Assurance Team	Medium
16	U9	Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T7	QA	Quality Assurance Team	High
17	T8	Documentation and Handover	Deliver documentation for maintenance and user guidance.	Task	E1	Documentation	Business Analysis Team	Medium
18	U10	Technical Documentation	Update diagrams and document ETL job flows.	Subtask	T8	Documentation	Business Analysis Team	Medium
19	U11	User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	Business Analysis Team	Medium
20	U12	Add field "CustomerNumber" to table "Dim_Customer"	Implement changes to add the field "CustomerNumber" into table "Dim_Customer", including logic and validation.	Subtask	T3	ETL	Data Warehouse Team	Medium

An interactive graph of the staging table right before the final review. Nodes are colored by issue type (Epic at the center, Tasks in blue, Subtasks in green); edges encode parent–child links in a top-down layout. This view lets the reviewer quickly validate structure and relationships before confirming the final approval.



The last checkpoint before ticket creation. The reviewer confirms or rejects the final outcome—labels, groups, priorities, and short text—without altering the parent-child structure. Selecting No activates the feedback box.

✓ Final Approval

Are you okay with this *final* outcome (labels/groups/priorities/text)?

(Please choose)

Yes

No

Provide feedback (values only: labels, groups, priority, summary/description)

Press Ctrl+Enter to apply

Apply Feedback

Create Jira Issues

Provide feedback (values only: labels, groups, priority, summary/description)

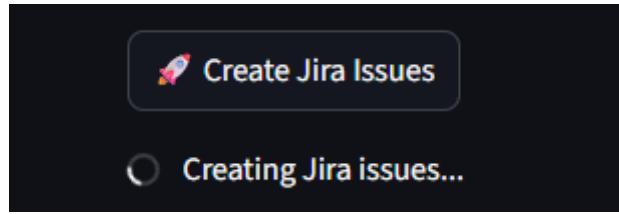
Press Ctrl+Enter to apply

Apply Feedback

An example of values-only feedback: the reviewer requests “Set all priorities equal to Low.” Pressing Apply Feedback updates the staging table accordingly; the screen then refreshes for a final confirmation prior to creating the Jira issues.

local_id	summary	description	issue_type	parent_local_id	label	assignee_group	priority
0 E1	Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic		Reporting	Data Warehouse Team	Low
1 T1	Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1	Reporting	Business Analysis Team	Low
2 U1	Conduct GAP Analysis	Compare fields against the database schema to identify gaps.	Subtask	T1	Documentation	Business Analysis Team	Low
3 T2	ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1	ETL	Data Warehouse Team	Low
4 U2	Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2	ETL	Data Warehouse Team	Low
5 U3	Transformation Logic	Implement data transformations and standardizations.	Subtask	T2	ETL	Data Warehouse Team	Low
6 U4	Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2	ETL	Data Warehouse Team	Low
7 T3	Account Portfolio Overview	Create a report in tabular format with specified fields: Account Number Customer Number City.	Task	E1	Reporting	Business Intelligence Team	Low
8 T4	Account Transaction History View	Create a report in tabular format with specified fields: Account Number Transaction Number.	Task	E1	Reporting	Business Intelligence Team	Low
9 T5	Customer Contact Information Overview	Create a report in tabular format with specified fields: Customer Number Customer Last Name.	Task	E1	Reporting	Business Intelligence Team	Low
10 T6	Power BI / Visualization Layer	Develop dashboards for interactive data visualization.	Task	E1	Development	Business Intelligence Team	Low
11 U5	Data Connectivity	Set up secure connections and parameterized queries.	Subtask	T6	ETL	Data Warehouse Team	Low
12 U6	Report Layout and Navigation	Create intuitive navigation and layout for reports.	Subtask	T6	Reporting	Business Intelligence Team	Low
13 U7	KPIs and Summary Metrics	Add calculated metrics and clear visualizations.	Subtask	T6	Reporting	Business Intelligence Team	Low
14 T7	Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1	QA	Quality Assurance Team	Low
15 U8	Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	Quality Assurance Team	Low
16 U9	Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T7	QA	Quality Assurance Team	Low
17 T8	Documentation and Handover	Deliver documentation for maintenance and user guidance.	Task	E1	Documentation	Business Analysis Team	Low
18 U10	Technical Documentation	Update diagrams and document ETL job flows.	Subtask	T8	Documentation	Business Analysis Team	Low
19 U11	User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	Business Analysis Team	Low
20 U12	Add field "CustomerNumber" to table "Dim_Customer"	Implement changes to add the field "CustomerNumber" into table "Dim_Customer", includin	Subtask	T3	ETL	Data Warehouse Team	Low

After the final values-only approval, the user clicks Create Jira Issues. A progress indicator confirms the handoff is in progress while the app creates the Epic, then Tasks, then Subtasks in the correct parent-child order and uploads any generated SQL as attachments.



On success, the app displays a green “Jira issues created” banner and a JSON-style map from the local IDs (e.g., E1, T1, U1 ...) to the actual Jira keys (e.g., DOC2SQL-1288, DOC2SQL-1289, ...). This serves as an audit trail and makes it easy to navigate to each ticket.

```
{
    "E1" : "DOC2SQL-1288",
    "T1" : "DOC2SQL-1289",
    "T2" : "DOC2SQL-1290",
    "T3" : "DOC2SQL-1291",
    "T4" : "DOC2SQL-1292",
    "T5" : "DOC2SQL-1293",
    "T6" : "DOC2SQL-1294",
    "T7" : "DOC2SQL-1295",
    "T8" : "DOC2SQL-1296",
    "U1" : "DOC2SQL-1297",
    "U2" : "DOC2SQL-1298",
    "U3" : "DOC2SQL-1299",
    "U4" : "DOC2SQL-1300",
    "U5" : "DOC2SQL-1301",
    "U6" : "DOC2SQL-1302",
    "U7" : "DOC2SQL-1303",
    "U8" : "DOC2SQL-1304",
    "U9" : "DOC2SQL-1305",
    "U10" : "DOC2SQL-1306",
    "U11" : "DOC2SQL-1307",
    "U12" : "DOC2SQL-1308"
}
```

Appendices

Appendix A — Case with Group Assignment

In this run we enable routing by teams by calling:

```
run_pipeline(doc_path, schema_text, include_sql=True, include_assignee_group=True).
```

With `include_assignee_group=True`, the pipeline activates Agent 4 (Assign to Groups) after structure approval and labeling. Each ticket is mapped—deterministically and from a controlled list—to exactly one delivery team, and that value is carried through to Jira in the Assigned Group field. In the example execution attached, the system produced assignments to the following teams: Data Warehouse Team, Business Analysis Team, Business Intelligence Team, and Quality Assurance Team. The rest of the flow (gap analysis, SQL generation, priority, approvals, and Jira creation) remains unchanged.

The execution results step by step:

# 1) Full run WITH SQL (requires schema_text) run_pipeline(doc_path, schema_text, include_sql=True, include_assignee_group=True)			
Here is the Jira ticketing structure:			
local_id	summary	description	issue_type
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task
2	U1 GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask
5	U3 Transformation Logic	Implement data transformations and standardize formats.	Subtask
6	U4 Data Loading	Integrate transformed data into warehouse tables with incremental loading.	Subtask
7	T3 SQL Reporting Layer	Build views and procedures for data presentation.	Task
8	T4 Account Portfolio Overview	Create a report in tabular format with specified fields.\nFields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task
9	T5 Account Transaction History View	Create a report in tabular format with specified fields.\nFields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task
10	T6 Customer Contact Information Overview	Create a report in tabular format with specified fields.\nFields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task
11	T7 Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task
12	U5 Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask
13	U6 Report Layout and Navigation	Create intuitive tabs for account data and transaction history.	Subtask
14	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for data insights.	Subtask
15	U8 Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task
16	U9 Data Validation	Compare report outputs with source records for accuracy.	Subtask
17	U9 Performance Testing	Conduct stress tests on SQL queries and Power BI dashboards.	Subtask
18	T9 Documentation and Handover	Provide comprehensive documentation for users and technical teams.	Task
19	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask
20	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask
Are you okay with this structure? (yes/no): no Please provide feedback on the changes you would like to make. Your feedback: do not include task named SQL Reporting Layer			
The structure was refined after feedback.			
Here is the Jira ticketing structure:			
local_id	summary	description	issue_type
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task
2	U1 GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask
5	U3 Transformation Logic	Implement data transformations and standardize formats.	Subtask
6	U4 Data Loading	Integrate transformed data into warehouse tables with incremental loading.	Subtask
7	T3 Account Portfolio Overview	Create a report in tabular format with specified fields.\nFields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task
8	T4 Account Transaction History View	Create a report in tabular format with specified fields.\nFields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task
9	T5 Customer Contact Information Overview	Create a report in tabular format with specified fields.\nFields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task
10	T6 Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task
11	U5 Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask
12	U6 Report Layout and Navigation	Create intuitive tabs for account data and transaction history.	Subtask
13	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for data insights.	Subtask
14	U8 Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task
15	U9 Data Validation	Compare report outputs with source records for accuracy.	Subtask
16	U9 Performance Testing	Conduct stress tests on SQL queries and Power BI dashboards.	Subtask
17	T8 Documentation and Handover	Provide comprehensive documentation for users and technical teams.	Task
18	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask
19	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask
Are you okay with this structure? (yes/no): yes			
The process continues.			

- Adding labels with LLM...
- Adding assignee group with LLM...
- Scanning for missing report fields against schema...

Suggested missing fields:

missing_field_name	suggested_table	parent_local_id
CustomerNumber	Dim_Customer	T3
AccountStatus	Dim_Account	T3
TransactionNumber	Fact_Transaction	T4
IsReversed	Fact_Transaction	T4
NationalPostalCode	Dim_Contact	T5

Confirm which ones to append as standardized 'Add field to table' subtasks...

Does field 'CustomerNumber' already exist in the database? (yes/no): no
 Do you want to add 'CustomerNumber' to table 'Dim_Customer' as a subtask? (yes/no): yes
 Does field 'AccountStatus' already exist in the database? (yes/no): no
 Do you want to add 'AccountStatus' to table 'Dim_Account' as a subtask? (yes/no): yes
 Does field 'TransactionNumber' already exist in the database? (yes/no): no
 Do you want to add 'TransactionNumber' to table 'Fact_Transaction' as a subtask? (yes/no): yes
 Does field 'IsReversed' already exist in the database? (yes/no): no
 Do you want to add 'IsReversed' to table 'Fact_Transaction' as a subtask? (yes/no): yes
 Does field 'NationalPostalCode' already exist in the database? (yes/no): no
 Do you want to add 'NationalPostalCode' to table 'Dim_Contact' as a subtask? (yes/no): yes

Generating ALTER TABLE statements (one per subtask)...

Identifying tabular report tasks...

Generating SELECT SQL queries for tabular reports...

Predicting priority for each item...

Here is the final ticketing structure:

local_id	summary	description	issue_type	parent_local_id	label	assignee_group	priority
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic		Reporting	Data Warehouse Team	Medium
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1	Reporting	Business Analysis Team	Medium
2	U1 GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1	Reporting	Business Analysis Team	Medium
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1	ETL	Data Warehouse Team	High
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2	ETL	Data Warehouse Team	Medium
5	U3 Transformation Logic	Implement data transformations and standardize formats.	Subtask	T2	ETL	Data Warehouse Team	Medium
6	U4 Data Loading	Integrate transformed data into warehouse tables with incremental loading.	Subtask	T2	ETL	Data Warehouse Team	Medium
7	T3 Account Portfolio Overview	Create a report in tabular format with specified fields.\nFields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1	Reporting	Business Intelligence Team	Low
8	T4 Account Transaction History View	Create a report in tabular format with specified fields.\nFields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1	Reporting	Business Intelligence Team	Medium
9	T5 Customer Contact Information Overview	Create a report in tabular format with specified fields.\nFields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1	Reporting	Business Intelligence Team	Low
10	T6 Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task	E1	Development	Business Intelligence Team	Medium
11	U5 Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask	T6	ETL	Data Warehouse Team	Medium
12	U6 Report Layout and Navigation	Create intuitive tabs for account data and transaction history.	Subtask	T6	Reporting	Business Intelligence Team	Low
13	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for data insights.	Subtask	T6	Reporting	Business Intelligence Team	Medium
14	T7 Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1	QA	Quality Assurance Team	Medium
15	U8 Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	Quality Assurance Team	Medium
16	U9 Performance Testing	Conduct stress tests on SQL queries and Power BI dashboards.	Subtask	T7	QA	Quality Assurance Team	Medium
17	T8 Documentation and Handover	Provide comprehensive documentation for users and technical teams.	Task	E1	Documentation	Business Analysis Team	Low
18	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8	Documentation	Data Warehouse Team	Low
19	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	Business Analysis Team	Low
20	U12 Add field "CustomerNumber" to table "Dim_Customer"	Implement changes to add the field 'CustomerNumber' into table 'Dim_Customer', including ETL and database updates.	Subtask	T3	Data	Data Warehouse Team	Medium
21	U13 Add field "AccountStatus" to table "Dim_Account"	Implement changes to add the field 'AccountStatus' into table 'Dim_Account', including ETL and database updates.	Subtask	T3	Data	Data Warehouse Team	Medium
22	U14 Add field "TransactionNumber" to table "Fact_Transaction"	Implement changes to add the field 'TransactionNumber' into table 'Fact_Transaction', including ETL and database updates.	Subtask	T4	Data	Data Warehouse Team	Medium
23	U15 Add field "IsReversed" to table "Fact_Transaction"	Implement changes to add the field 'IsReversed' into table 'Fact_Transaction', including ETL and database updates.	Subtask	T4	Data	Data Warehouse Team	Medium
24	U16 Add field "NationalPostalCode" to table "Dim_Contact"	Implement changes to add the field 'NationalPostalCode' into table 'Dim_Contact', including ETL and database updates.	Subtask	T5	Data	Data Warehouse Team	Medium

Are you okay with this final outcome? (yes/no): no

Please provide feedback on the changes you would like to make (values only, not structure).

Your feedback: set all priorities to low

Applying your feedback. Please wait...

Here is the final ticketing structure:

local_id	summary	description	issue_type	parent_local_id	label	assignee_group	priority
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic		Reporting	Data Warehouse Team	Low
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1	Reporting	Business Analysis Team	Low
2	U1 GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1	Reporting	Business Analysis Team	Low
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1	ETL	Data Warehouse Team	Low
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2	ETL	Data Warehouse Team	Low
5	U3 Transformation Logic	Implement data transformations and standardize formats.	Subtask	T2	ETL	Data Warehouse Team	Low
6	U4 Data Loading	Integrate transformed data into warehouse tables with incremental loading.	Subtask	T2	ETL	Data Warehouse Team	Low
7	T3 Account Portfolio Overview	Create a report in tabular format with specified fields.\nFields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1	Reporting	Business Intelligence Team	Low
8	T4 Account Transaction History View	Create a report in tabular format with specified fields.\nFields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1	Reporting	Business Intelligence Team	Low
9	T5 Customer Contact Information Overview	Create a report in tabular format with specified fields.\nFields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1	Reporting	Business Intelligence Team	Low
10	T6 Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task	E1	Development	Business Intelligence Team	Low
11	U5 Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask	T6	ETL	Data Warehouse Team	Low
12	U6 Report Layout and Navigation	Create intuitive tabs for account data and transaction history.	Subtask	T6	Reporting	Business Intelligence Team	Low
13	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for data insights.	Subtask	T6	Reporting	Business Intelligence Team	Low
14	T7 Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1	QA	Quality Assurance Team	Low
15	U8 Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	Quality Assurance Team	Low
16	U9 Performance Testing	Conduct stress tests on SQL queries and Power BI dashboards.	Subtask	T7	QA	Quality Assurance Team	Low
17	T8 Documentation and Handover	Provide comprehensive documentation for users and technical teams.	Task	E1	Documentation	Business Analysis Team	Low
18	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8	Documentation	Data Warehouse Team	Low
19	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	Business Analysis Team	Low
20	U12 Add field "CustomerNumber" to table "Dim_Customer"	Implement changes to add the field 'CustomerNumber' into table 'Dim_Customer', including ETL and database updates.	Subtask	T3	Data	Data Warehouse Team	Low
21	U13 Add field "AccountStatus" to table "Dim_Account"	Implement changes to add the field 'AccountStatus' into table 'Dim_Account', including ETL and database updates.	Subtask	T3	Data	Data Warehouse Team	Low
22	U14 Add field "TransactionNumber" to table "Fact_Transaction"	Implement changes to add the field 'TransactionNumber' into table 'Fact_Transaction', including ETL and database updates.	Subtask	T4	Data	Data Warehouse Team	Low
23	U15 Add field "IsReversed" to table "Fact_Transaction"	Implement changes to add the field 'IsReversed' into table 'Fact_Transaction', including ETL and database updates.	Subtask	T4	Data	Data Warehouse Team	Low
24	U16 Add field "NationalPostalCode" to table "Dim_Contact"	Implement changes to add the field 'NationalPostalCode' into table 'Dim_Contact', including ETL and database updates.	Subtask	T5	Data	Data Warehouse Team	Low

Are you okay with this final outcome? (yes/no): yes

Creating Jira issues...

Here is the Jira ticketing structure, with the extra subtasks:

local_id	summary	description	issue_type	parent_local_id	label	assignee_group	priority
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting development.	Epic		Reporting	Data Warehouse Team	Low
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1	Reporting	Business Analysis Team	Low
2	U1 GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1	Reporting	Business Analysis Team	Low
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1	ETL	Data Warehouse Team	Low
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2	ETL	Data Warehouse Team	Low
5	U3 Transformation Logic	Implement data transformations and standardize formats.	Subtask	T2	ETL	Data Warehouse Team	Low
6	U4 Data Loading	Integrate transformed data into warehouse tables with incremental loading.	Subtask	T2	ETL	Data Warehouse Team	Low
7	T3 Account Portfolio Overview	Create a report in tabular format with specified fields.\nFields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1	Reporting	Business Intelligence Team	Low
8	T4 Account Transaction History View	Create a report in tabular format with specified fields.\nFields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1	Reporting	Business Intelligence Team	Low
9	T5 Customer Contact Information Overview	Create a report in tabular format with specified fields.\nFields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1	Reporting	Business Intelligence Team	Low
10	T6 Power BI / Visualization Layer	Develop interactive dashboards for data visualization.	Task	E1	Development	Business Intelligence Team	Low
11	U5 Data Connectivity	Set up secure connections and parameterized queries for Power BI.	Subtask	T6	ETL	Data Warehouse Team	Low
12	U6 Report Layout and Navigation	Create intuitive tabs for account data and transaction history.	Subtask	T6	Reporting	Business Intelligence Team	Low
13	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for data insights.	Subtask	T6	Reporting	Business Intelligence Team	Low
14	T7 Quality Assurance and Testing	Ensure accuracy and performance of data and reports.	Task	E1	QA	Quality Assurance Team	Low
15	U8 Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	Quality Assurance Team	Low
16	U9 Performance Testing	Conduct stress tests on SQL queries and Power BI dashboards.	Subtask	T7	QA	Quality Assurance Team	Low
17	T8 Documentation and Handover	Provide comprehensive documentation for users and technical teams.	Task	E1	Documentation	Business Analysis Team	Low
18	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8	Documentation	Data Warehouse Team	Low
19	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	Business Analysis Team	Low
20	U12 Add field "CustomerNumber" to table "Dim_Customer"	Implement changes to add the field "CustomerNumber" into table "Dim_Customer", including ETL and database updates.	Subtask	T3	Data	Data Warehouse Team	Low
21	U13 Add field "AccountStatus" to table "Dim_Account"	Implement changes to add the field "AccountStatus" into table "Dim_Account", including ETL and database updates.	Subtask	T3	Data	Data Warehouse Team	Low
22	U14 Add field "TransactionNumber" to table "Fact_Transaction"	Implement changes to add the field "TransactionNumber" into table "Fact_Transaction", including ETL and database updates.	Subtask	T4	Data	Data Warehouse Team	Low
23	U15 Add field "IsReversed" to table "Fact_Transaction"	Implement changes to add the field "IsReversed" into table "Fact_Transaction", including ETL and database updates.	Subtask	T4	Data	Data Warehouse Team	Low
24	U16 Add field "NationalPostalCode" to table "Dim_Contact"	Implement changes to add the field "NationalPostalCode" into table "Dim_Contact", including ETL and database updates.	Subtask	T5	Data	Data Warehouse Team	Low

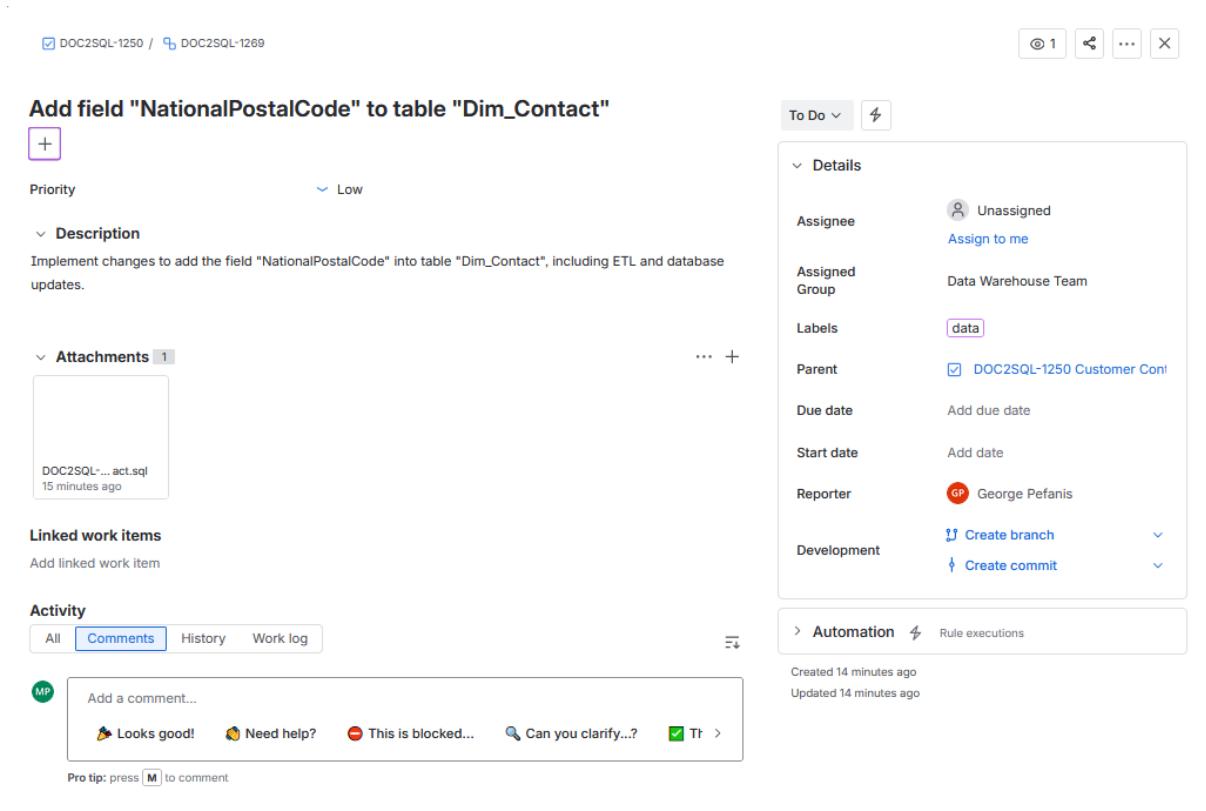
```

Epic created: DOC2SQL-1245
Task created: DOC2SQL-1246
Task created: DOC2SQL-1247
Task created: DOC2SQL-1248
Attached DOC2SQL-1248_account_portfolio_overview.sql + DOC2SQL-1248
Task created: DOC2SQL-1249
Attached DOC2SQL-1249_account_transaction_history_view.sql + DOC2SQL-1249
Task created: DOC2SQL-1250
Attached DOC2SQL-1250_customer_contact_information_overview.sql + DOC2SQL-1250
Task created: DOC2SQL-1251
Task created: DOC2SQL-1252
Task created: DOC2SQL-1253
Subtask created: DOC2SQL-1254
Subtask created: DOC2SQL-1255
Subtask created: DOC2SQL-1256
Subtask created: DOC2SQL-1257
Subtask created: DOC2SQL-1258
Subtask created: DOC2SQL-1259
Subtask created: DOC2SQL-1260
Subtask created: DOC2SQL-1261
Subtask created: DOC2SQL-1262
Subtask created: DOC2SQL-1263
Subtask created: DOC2SQL-1264
Subtask created: DOC2SQL-1265
Attached DOC2SQL-1265_add_field_customernumber_to_table_dim_customer.sql + DOC2SQL-1265
Subtask created: DOC2SQL-1266
Attached DOC2SQL-1266_add_field_accountstatus_to_table_dim_account.sql + DOC2SQL-1266
Subtask created: DOC2SQL-1267

Attached DOC2SQL-1267_add_field_transactionnumber_to_table_fact_transaction.sql + DOC2SQL-1267
Subtask created: DOC2SQL-1268
Attached DOC2SQL-1268_add_field_isreversed_to_table_fact_transaction.sql + DOC2SQL-1268
Subtask created: DOC2SQL-1269
Attached DOC2SQL-1269_add_field_nationalpostalcode_to_table_dim_contact.sql + DOC2SQL-1269

```

Indicatively, for subtask DOC2SQL-1269 (“Add field *NationalPostalCode* to table *Dim_Contact*”), the generated Jira ticket sets the Assigned Group to Data Warehouse Team, as shown below.



Appendix B — Run without SQL

This appendix documents the document-to-Jira flow when the SQL/data branch is disabled. With the call `run_pipeline(doc_path, schema_text=None, include_sql=False)`, the system focuses on the core pipeline: extracting a clean Epic → Task → Subtask structure from the requirements document, refining it via a human-in-the-loop approval, building the staging DataFrame, automatically assigning labels (and, if enabled, assignee groups), inferring priority, collecting a final values-only approval, and finally creating the Jira issues with the correct hierarchy. No schema gap analysis is performed and no DDL/SELECT SQL is generated—this run isolates the fundamental understanding and ticketing capabilities for cases where a database schema is unavailable or out of scope. This appendix documents the document-to-Jira flow when the SQL/data branch is disabled. With the call `run_pipeline(doc_path, schema_text=None, include_sql=False)`, the system focuses on the core pipeline: extracting a clean Epic → Task → Subtask structure from the requirements document, refining it via a human-in-the-loop approval, building the staging DataFrame, automatically assigning labels (and, if enabled, assignee groups), inferring priority, collecting a final values-only approval, and finally creating the Jira issues with the correct hierarchy. No schema gap analysis is performed and no DDL/SELECT SQL is generated—this run isolates the fundamental understanding and ticketing capabilities for cases where a database schema is unavailable or out of scope.

Here is the Jira ticketing structure:

local_id	summary	description	issue_type	parent_local_id
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting processes.	Epic	
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1
2	U1 Conduct GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2
5	U3 Transformation Logic	Implement data transformations and standardizations.	Subtask	T2
6	U4 Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2
7	T3 SQL Reporting Layer	Build views and procedures for data presentation.	Task	E1
8	T4 Account Portfolio Overview	Create a report in tabular format with specified fields. In Fields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1
9	T5 Account Transaction History View	Create a report in tabular format with specified fields. In Fields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1
10	T6 Customer Contact Information Overview	Create a report in tabular format with specified fields. In Fields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1
11	T7 Power BI / Visualization Layer	Create dashboards for interactive data presentation.	Task	E1
12	U5 Data Connectivity	Set up secure connection between Power BI and data warehouse.	Subtask	T7
13	U6 Report Layout and Navigation	Develop intuitive navigation and layout for reports.	Subtask	T7
14	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for insights.	Subtask	T7
15	T8 Quality Assurance and Testing	Verify accuracy and performance of data and reports.	Task	E1
16	U8 Data Validation	Compare report outputs with source records for accuracy.	Subtask	T8
17	U9 Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T8
18	T9 Documentation and Handover	Deliver documentation for maintenance and user guidance.	Task	E1
19	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T9
20	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T9

Are you okay with this structure? (yes/no): no
 Please provide feedback on the changes you would like to make.
 Your feedback: do not include task named SQL Reporting Layer

The structure was refined after feedback.

Here is the Jira ticketing structure:

local_id	summary	description	issue_type	parent_local_id
0	E1 Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting processes.	Epic	
1	T1 Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1
2	U1 Conduct GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1
3	T2 ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1
4	U2 Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2
5	U3 Transformation Logic	Implement data transformations and standardizations.	Subtask	T2
6	U4 Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2
7	T3 Account Portfolio Overview	Create a report in tabular format with specified fields. In Fields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1
8	T4 Account Transaction History View	Create a report in tabular format with specified fields. In Fields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1
9	T5 Customer Contact Information Overview	Create a report in tabular format with specified fields. In Fields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1
10	T6 Power BI / Visualization Layer	Create dashboards for interactive data presentation.	Task	E1
11	U5 Data Connectivity	Set up secure connection between Power BI and data warehouse.	Subtask	T6
12	U6 Report Layout and Navigation	Develop intuitive navigation and layout for reports.	Subtask	T6
13	U7 KPIs and Summary Metrics	Add calculated metrics and visualizations for insights.	Subtask	T6
14	T7 Quality Assurance and Testing	Verify accuracy and performance of data and reports.	Task	E1
15	U8 Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7
16	U9 Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T7
17	T8 Documentation and Handover	Deliver documentation for maintenance and user guidance.	Task	E1
18	U10 Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8
19	U11 User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8

Are you okay with this structure? (yes/no): yes

The process continues.

👉 Adding labels with LLM...

👤 Adding assignee group with LLM...

⚠️ Predicting priority for each item...

Here is the final ticketing structure:

local_id	summary		description	issue_type	parent_local_id	label	assignee_group	priority
0	E1	Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting processes.	Epic		Reporting	Data Warehouse Team	Medium
1	T1	Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1	Reporting	Business Analysis Team	Medium
2	U1	Conduct GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1	Business	Business Analysis Team	Medium
3	T2	ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1	ETL	Data Warehouse Team	High
4	U2	Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2	ETL	Data Warehouse Team	Medium
5	U3	Transformation Logic	Implement data transformations and standardizations.	Subtask	T2	ETL	Data Warehouse Team	Medium
6	U4	Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2	ETL	Data Warehouse Team	Medium
7	T3	Account Portfolio Overview	Create a report in tabular format with specified fields. Fields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1	Reporting	Business Intelligence Team	Low
8	T4	Account Transaction History View	Create a report in tabular format with specified fields. Fields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1	Reporting	Business Intelligence Team	Low
9	T5	Customer Contact Information Overview	Create a report in tabular format with specified fields. Fields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1	Reporting	Business Intelligence Team	Low
10	T6	Power BI / Visualization Layer	Create dashboards for interactive data presentation.	Task	E1	Reporting	Business Intelligence Team	Medium
11	U5	Data Connectivity	Set up secure connection between Power BI and data warehouse.	Subtask	T6	ETL	Data Warehouse Team	High
12	U6	Report Layout and Navigation	Develop intuitive navigation and layout for reports.	Subtask	T6	Reporting	Business Intelligence Team	Low
13	U7	KPIs and Summary Metrics	Add calculated metrics and visualizations for insights.	Subtask	T6	Reporting	Business Intelligence Team	Medium
14	T7	Quality Assurance and Testing	Verify accuracy and performance of data and reports.	Task	E1	QA	Quality Assurance Team	Medium
15	U8	Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	Quality Assurance Team	Medium
16	U9	Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T7	QA	Quality Assurance Team	Medium
17	T8	Documentation and Handover	Deliver documentation for maintenance and user guidance.	Task	E1	Documentation	Business Analysis Team	Low
18	U10	Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8	Documentation	Data Warehouse Team	Low
19	U11	User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	Business Intelligence Team	Low

Are you okay with this final outcome? (yes/no): no
 Please provide feedback on the changes you would like to make (values only, not structure).
 Your feedback: set all priorities to medium

Applying your feedback. Please wait...

Here is the final ticketing structure:

local_id	summary		description	issue_type	parent_local_id	label	assignee_group	priority
0	E1	Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting processes.	Epic		Reporting	Data Warehouse Team	Medium
1	T1	Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1	Reporting	Business Analysis Team	Medium
2	U1	Conduct GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1	Business	Business Analysis Team	Medium
3	T2	ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1	ETL	Data Warehouse Team	Medium
4	U2	Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2	ETL	Data Warehouse Team	Medium
5	U3	Transformation Logic	Implement data transformations and standardizations.	Subtask	T2	ETL	Data Warehouse Team	Medium
6	U4	Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2	ETL	Data Warehouse Team	Medium
7	T3	Account Portfolio Overview	Create a report in tabular format with specified fields. Fields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1	Reporting	Business Intelligence Team	Medium
8	T4	Account Transaction History View	Create a report in tabular format with specified fields. Fields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1	Reporting	Business Intelligence Team	Medium
9	T5	Customer Contact Information Overview	Create a report in tabular format with specified fields. Fields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1	Reporting	Business Intelligence Team	Medium
10	T6	Power BI / Visualization Layer	Create dashboards for interactive data presentation.	Task	E1	Reporting	Business Intelligence Team	Medium
11	U5	Data Connectivity	Set up secure connection between Power BI and data warehouse.	Subtask	T6	ETL	Data Warehouse Team	Medium
12	U6	Report Layout and Navigation	Develop intuitive navigation and layout for reports.	Subtask	T6	Reporting	Business Intelligence Team	Medium
13	U7	KPIs and Summary Metrics	Add calculated metrics and visualizations for insights.	Subtask	T6	Reporting	Business Intelligence Team	Medium
14	T7	Quality Assurance and Testing	Verify accuracy and performance of data and reports.	Task	E1	QA	Quality Assurance Team	Medium
15	U8	Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	Quality Assurance Team	Medium
16	U9	Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T7	QA	Quality Assurance Team	Medium
17	T8	Documentation and Handover	Deliver documentation for maintenance and user guidance.	Task	E1	Documentation	Business Analysis Team	Medium
18	U10	Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8	Documentation	Data Warehouse Team	Medium
19	U11	User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	Business Intelligence Team	Medium

Are you okay with this final outcome? (yes/no): yes

Creating Jira issues...

Here is the Jira ticketing structure, with the extra subtasks:

local_id	summary		description	issue_type	parent_local_id	label	assignee_group	priority
0	E1	Data Warehouse and Reporting Specification	Comprehensive plan for data warehouse and reporting processes.	Epic		Reporting	Data Warehouse Team	Medium
1	T1	Reporting Requirements Analysis	Review reporting needs to ensure accurate data delivery.	Task	E1	Reporting	Business Analysis Team	Medium
2	U1	Conduct GAP Analysis	Compare fields against the database schema and identify missing fields.	Subtask	T1	Business	Business Analysis Team	Medium
3	T2	ETL Development	Update ETL processes for accurate data extraction and loading.	Task	E1	ETL	Data Warehouse Team	Medium
4	U2	Source System Extraction	Identify source systems and develop extraction processes.	Subtask	T2	ETL	Data Warehouse Team	Medium
5	U3	Transformation Logic	Implement data transformations and standardizations.	Subtask	T2	ETL	Data Warehouse Team	Medium
6	U4	Data Loading	Integrate transformed data into warehouse tables.	Subtask	T2	ETL	Data Warehouse Team	Medium
7	T3	Account Portfolio Overview	Create a report in tabular format with specified fields. Fields: Account Number, Customer Number, Customer Last Name, Customer First Name, Type of Account, Account Status	Task	E1	Reporting	Business Intelligence Team	Medium
8	T4	Account Transaction History View	Create a report in tabular format with specified fields. Fields: Account Number, Transaction Number, Transaction Date, Transaction Type, Transaction Amount, Is Reversed	Task	E1	Reporting	Business Intelligence Team	Medium
9	T5	Customer Contact Information Overview	Create a report in tabular format with specified fields. Fields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact	Task	E1	Reporting	Business Intelligence Team	Medium
10	T6	Power BI / Visualization Layer	Create dashboards for interactive data presentation.	Task	E1	Reporting	Business Intelligence Team	Medium
11	U5	Data Connectivity	Set up secure connection between Power BI and data warehouse.	Subtask	T6	ETL	Data Warehouse Team	Medium
12	U6	Report Layout and Navigation	Develop intuitive navigation and layout for reports.	Subtask	T6	Reporting	Business Intelligence Team	Medium
13	U7	KPIs and Summary Metrics	Add calculated metrics and visualizations for insights.	Subtask	T6	Reporting	Business Intelligence Team	Medium
14	T7	Quality Assurance and Testing	Verify accuracy and performance of data and reports.	Task	E1	QA	Quality Assurance Team	Medium
15	U8	Data Validation	Compare report outputs with source records for accuracy.	Subtask	T7	QA	Quality Assurance Team	Medium
16	U9	Performance Testing	Conduct stress tests on SQL queries and dashboards.	Subtask	T7	QA	Quality Assurance Team	Medium
17	T8	Documentation and Handover	Deliver documentation for maintenance and user guidance.	Task	E1	Documentation	Business Analysis Team	Medium
18	U10	Technical Documentation	Update diagrams and document ETL job flows and SQL views.	Subtask	T8	Documentation	Data Warehouse Team	Medium
19	U11	User Documentation	Create a user guide for Power BI with a business glossary.	Subtask	T8	Documentation	Business Intelligence Team	Medium

Epic created: DOC2SQL-1354
Task created: DOC2SQL-1355
Task created: DOC2SQL-1356
Task created: DOC2SQL-1357
Task created: DOC2SQL-1358
Task created: DOC2SQL-1359
Task created: DOC2SQL-1360
Task created: DOC2SQL-1361
Task created: DOC2SQL-1362
Subtask created: DOC2SQL-1363
Subtask created: DOC2SQL-1364
Subtask created: DOC2SQL-1365
Subtask created: DOC2SQL-1366
Subtask created: DOC2SQL-1367
Subtask created: DOC2SQL-1368
Subtask created: DOC2SQL-1369
Subtask created: DOC2SQL-1370
Subtask created: DOC2SQL-1371
Subtask created: DOC2SQL-1372
Subtask created: DOC2SQL-1373

Indicatively, for task DOC2SQL-1359 (“Customer Contact Information Overview”)

The screenshot shows the Jira task details page for task DOC2SQL-1359, which is a subtask of DOC2SQL-1354. The task title is "Customer Contact Information Overview".

Priority: Medium

Description: Create a report in tabular format with specified fields.
Fields: Customer Number, Customer Last Name, Customer First Name, Address, Country, National Postal Code, Is Primary Contact

Subtasks: Add subtask

Linked work items: Add linked work item

Activity: All, Comments (selected), History, Work log

Comments: Add a comment...
↳ Looks good! ⚡ Need help? 🛡 This is blocked... 🔍 Can you clarify...? 📋 Tl >

Details:

- Assignee: Unassigned (button: Assign to me)
- Assigned Group: Business Intelligence Team
- Labels: reporting
- Parent: DOC2SQL-1354 Data Warehouse
- Due date: Add due date
- Team: Add team
- Start date: Add date
- Reporter: George Pefanis (button: Create branch)
- Development: Create commit (button: Create branch)

Automation: Rule executions

Created 13 hours ago
Updated 13 hours ago

Appendix C — Requirements Document (Input)

The document used during execution is the source technical specification that drives the entire pipeline. It contains business requirements, functional descriptions, and—where present—explicit field lists for reports. The system uses it to derive the Epic → Task → Subtask hierarchy (Agents 1–2), enrich item descriptions and labels (Agent 3), and extract requested report fields that feed the optional gap analysis and SQL generation.

Data Warehouse and Reporting Specification

1. Reporting Requirements Analysis

Conduct a thorough review of the reporting needs to ensure the business receives exactly the data and insights it requires. This includes understanding the context of each report, the end users, and the decisions they will support.

1.1. GAP Analysis

Compare the validated list of fields against the existing database schema. Identify all missing fields and determine the best approach for adding them to the data warehouse — whether through schema changes, ETL logic, or derived calculations. Document any required upstream changes in source systems.

2. ETL Development

The goal is to update and extend the ETL processes to extract, transform, and load all required data into the data warehouse, ensuring it is accurate, complete, and up to date.

2.1. Source System Extraction

Identify all source systems, such as the core banking platform and CRM. Develop extraction processes — SQL queries, API calls, or flat file imports — that retrieve the required data fields without unnecessary overhead.

2.2. Transformation Logic

Implement transformations to normalize the data that was extracted, map transaction types to standardized codes. Ensure that all date and currency formats are standardized across the warehouse.

2.3. Data Loading

Integrate the transformed data into the warehouse's fact and dimension tables, updating tables such as Fact_Transactions and Dim_Accounts. Implement an incremental load strategy to handle daily updates efficiently without reloading historical data unnecessarily.

3. SQL Reporting Layer

Build database views or stored procedures that consolidate and present the required data in a format ready for reporting and visualization.

3.1. Account Portfolio Overview

The first report concerns Account Portfolio Overview. The report must be in tabular format, and it must contain the following fields.

- Account Number
- Customer Number
- Customer Last Name
- Customer First Name
- Type of Account
- Account Status

3.2. Account Transaction History View

The second report displays detailed records of all transactions per account. This report also must be in tabular format. It must contain the following fields.

- Account Number
- Transaction Number
- Transaction Date
- Transaction Type
- Transaction Amount
- Is Reversed

3.3. Customer Contact Information Overview

The report must be in tabular format, and it must contain the following fields:

- Customer Number
- Customer Last Name
- Customer First Name
- Address
- Country
- National Postal Code
- Is Primary Contact

4. Power BI / Visualization Layer

Create user-friendly dashboards in Power BI that present the data from the SQL layer in a clear, interactive, and actionable way.

4.1. Data Connectivity

Set up a secure connection between Power BI and the data warehouse. Where appropriate, use parameterized queries to enable dynamic filtering at the source, minimizing data load times.

4.2. Report Layout and Navigation

Develop a dedicated Account Portfolio Overview tab that displays account data with filters for account status and type. Create a Transaction History tab with slicers for date ranges, transaction types, and other relevant criteria. Ensure navigation between tabs is intuitive.

4.3. KPIs and Summary Metrics

Add calculated metrics such as total transaction amounts per period, average transaction values per account, and counts of reversed transactions. Use clear visualizations — cards, line charts, bar charts — to make trends and outliers easy to spot.

5. Quality Assurance and Testing

Verify that all new data, reports, and dashboards are correct, performant, and aligned with business needs.

5.1. Data Validation

Compare the output from the reports with source system records to ensure accuracy. Perform random sampling across different accounts and transactions to validate both summary and detail-level data.

5.2. Performance Testing

Run stress tests on SQL queries and Power BI dashboards to ensure they can handle expected data volumes without degradation. Fine-tune indexing, joins, and filtering strategies as needed to maintain responsiveness.

6. Documentation and Handover

Deliver comprehensive documentation so that both technical teams and business users can maintain and use the solution effectively.

6.1. Technical Documentation

Update the entity relationship diagrams to reflect schema changes. Document ETL job flows, including schedules and dependencies. Provide detailed definitions of SQL views, including business logic applied in each.

6.2. User Documentation

Create a Power BI user guide with screenshots and instructions on using filters, slicers, and navigation. Include a business glossary explaining each data field and KPI to ensure consistent interpretation across the organization.

Appendix D — Database Schema (Input)

The attached database schema serves as the reference for data alignment. It lists table/column names and data types and is used to: (i) compare requested fields from the document against what exists (gap analysis), (ii) generate DDL scripts (e.g., ALTER TABLE ... ADD ...) for missing columns, and (iii) compose SELECT queries for tabular reports using real table/column names.

-- Dim_Customer

```
CREATE TABLE Dim_Customer (
    CustomerID INT PRIMARY KEY AUTO_INCREMENT,
    FirstName VARCHAR(50),
    LastName VARCHAR(50),
    DateOfBirth DATE,
    Gender ENUM('Male', 'Female', 'Other'),
    CreatedDate DATETIME DEFAULT CURRENT_TIMESTAMP
);
```

-- Dim_Account

```
CREATE TABLE Dim_Account (
    AccountID INT PRIMARY KEY AUTO_INCREMENT,
    CustomerID INT,
    AccountNumber VARCHAR(20) UNIQUE,
    AccountType ENUM('Checking', 'Savings', 'Credit', 'Loan'),
    OpenDate DATE,
    CloseDate DATE,
    FOREIGN KEY (CustomerID) REFERENCES Dim_Customer(CustomerID)
);
```

-- Dim_Contact

```
CREATE TABLE Dim_Contact (
    ContactID INT PRIMARY KEY AUTO_INCREMENT,
    CustomerID INT,
    ContactType ENUM('Email', 'Phone', 'Address'),
    ContactValue VARCHAR(255),
);
```

```

    IsPrimary BOOLEAN DEFAULT FALSE,
    FOREIGN KEY (CustomerID) REFERENCES Dim_Customer(CustomerID)
);

-- Dim_Service

CREATE TABLE Dim_Service (
    ServiceID INT PRIMARY KEY AUTO_INCREMENT,
    ServiceName VARCHAR(100) NOT NULL,
    ServiceCategory VARCHAR(50),
    ServiceDescription TEXT
);

-- Dim_Branch

CREATE TABLE Dim_Branch (
    BranchID INT PRIMARY KEY AUTO_INCREMENT,
    BranchName VARCHAR(100),
    BranchType ENUM('Physical', 'Online'),
    Address VARCHAR(255),
    City VARCHAR(50),
    State VARCHAR(50),
    Country VARCHAR(50)
);

-- Dim_Currency

CREATE TABLE Dim_Currency (
    CurrencyID INT PRIMARY KEY AUTO_INCREMENT,
    CurrencyCode CHAR(3) NOT NULL,
    CurrencyName VARCHAR(50),
    Symbol VARCHAR(5),
    UNIQUE KEY uq_currency_code (CurrencyCode)
);

-- Fact_FXRate (daily rates or as-granular-as-you-load)

CREATE TABLE IF NOT EXISTS Fact_FXRate (
    RateDate DATE NOT NULL,

```

```
FromCurrencyID INT NOT NULL,  
ToCurrencyID INT NOT NULL,  
Rate DECIMAL(18,8) NOT NULL,  
PRIMARY KEY (RateDate, FromCurrencyID, ToCurrencyID),  
FOREIGN KEY (FromCurrencyID) REFERENCES Dim_Currency(CurrencyID)  
    ON UPDATE CASCADE ON DELETE RESTRICT,  
FOREIGN KEY (ToCurrencyID) REFERENCES Dim_Currency(CurrencyID)  
    ON UPDATE CASCADE ON DELETE RESTRICT  
);  
-- Fact_Transaction  
CREATE TABLE Fact_Transaction (  
    TransactionID BIGINT PRIMARY KEY AUTO_INCREMENT,  
    AccountID INT NOT NULL,  
    ServiceID INT NULL,  
    BranchID INT NULL,  
    TransactionDate DATETIME NOT NULL,  
    Amount DECIMAL(15, 2) NOT NULL,  
    CurrencyID INT NOT NULL,  
    AmountBase DECIMAL(18, 2) NULL,  
    BaseCurrencyID INT NULL,  
    FXRateUsed DECIMAL(18,8) NULL,  
    FXRateDate DATE NULL,  
    TransactionType ENUM('Credit', 'Debit') NOT NULL,  
    Description TEXT,  
    FOREIGN KEY (AccountID) REFERENCES Dim_Account(AccountID)  
);
```

Contribution

Business Plan/ Analysis: all team members

Georgios Pefanis: main coder of Agents and Functions

Pavlos Kalogirou: main coder of Streamlit app

Manolis Pavlogiorgatos: report/ presentation, QA – testing/ bug fixing/ validation