

2026-02-11 – Phase 4 Pipeline Cohesion + YAML-Driven Task Config

Alcance implementado:

- Cohesión visual/funcional entre overview, detail y pro sidebar para contexto DAG.
- Configuración de task sensible a etapa/tipo alineada al contrato de `configfile_proposal.yml`.
- Agrupación dinámica de pipelines (`by tag` / `by integration`) con columna de estado.
- En Pipelines Pro , file tree expandido por default, sin creación libre de archivos y handoff limpio sin subtab lateral de pipeline.

Qué cambió

Overview: agrupación y estado de pipeline

- `PipelineOverview` soporta toggle real `by tag` y `by integration` :
 - `by integration` : agrupa por carpeta/integración.
 - `by tag` : agrupa por tags funcionales (ej. `snapshot` , `incremental`), no por folder.
- Se agregó columna `Status` con `StatusBadge` calculado por pipeline.
- Nuevo helper `getPipelineStatus(...)` para derivar estado agregado en base a tasks no-DDL.
- Se agregó `Status legend` + botón `cycle` por fila para testear badges en scaffold sin backend.

Badges: significado y disparadores en scaffold

- `Draft` : estado inicial o edición pendiente.
- `Submitted` : simula push/submit a dev.
- `Pending` : simula envío pendiente de aprobación.
- `Approved` : simula aprobación final.
- Para probarlas en demo: usar botón `cycle` de cada fila (rota `Draft` → `Submitted` → `Pending` → `Approved`).

Detail: estado agregado en header

- `PipelineDetail` muestra badge de estado del pipeline junto al tipo (`snapshot` / `incremental`).
- El badge usa la misma lógica agregada que el overview para evitar desalineaciones.

Task Config: campos por etapa/tipo

- `TaskConfigPanel` ahora diferencia explícitamente etapas:
 - `extract` : muestra `Expected Workload` + `Target Table Name` .
 - `transform` : no muestra target connection.
 - `load` y `dqa` : mantienen campos/variantes relevantes.
- Se agregó `targetTableName` al tipo `TaskConfig` .
- Se agrega pie de contexto `Stage` + `Task type` para lectura rápida.

Pipelines Pro: contexto DAG desde folder root (handoff)

- Se corrigió la resolución de contexto DAG desde `selectedFolder` aun cuando no hay `selectedFile` .
- Se quitó la subtab lateral `Pipeline` .
- En su lugar se muestra una tarjeta compacta de contexto con CTA `Open pipeline handoff` .

Guardrail de scaffold en Pro

- Se quitó creación libre de SQL desde el file tree en Pipelines Pro .
- Mensaje explícito: creación de archivos solo vía flujos de pipeline.

What's New híbrido (checklist + tour)

- Se agregó botón What's new en header con modal híbrido:
 - checklist del plan (done/current/next/out),
 - CTA Start guided tour .
- Se integró tour guiado con react-joyride sobre puntos clave de F4.
- Se agregó acceso rápido al PDF actual en header (Current PDF) y en ruta /changelog/current .

Evidencia visual

Overview grouped by tag + status lab (dark)

The screenshot shows the Data Explorer interface in Pipeline submode. At the top, there are tabs for SQL Explorer and Pipelines, with Pipelines selected. The header includes status indicators (Dev, Prod), a theme switch, and user roles (Data Engineer, Engineer). A search bar at the top left allows searching by name, integration, or tag. Below the search bar, a status legend indicates Draft (blue), Submitted (orange), Pending (yellow), and Approved (green). A note says "Use status button per row to cycle in scaffold".

The main area is divided into two sections: INCREMENTAL and SNAPSHOT. Each section has a table with columns: PIPELINE, TYPE, SCHEDULE, TASKS, STATUS, and TAGS.

INCREMENTAL Section:

PIPELINE	TYPE	SCHEDULE	TASKS	STATUS	TAGS
tr_GameTransaction	incremental	Every hour in 2m	3	Draft	cyclic@incremental
tr_DailyTransactionAmount	incremental	Every hour in 2m	3	Draft	cyclic@incremental

SNAPSHOT Section:

PIPELINE	TYPE	SCHEDULE	TASKS	STATUS	TAGS
dbo_AccountReference	snapshot	Every 3h in 2h 27m	3	Draft	cyclic@snapshot
gc_Game	snapshot	Every 3h in 2h 25m	3	Draft	cyclic@snapshot
dbo_AccountLogType	snapshot	Daily at 06:00 in 28h 28m	2	Draft	cyclic@snapshot

At the bottom left, there is a red button with a white exclamation mark containing the text "2 Issues X". On the right side, there is a purple circular icon with a white question mark.

Qué mirar:

- Toggle by tag activo y grupos distintos (snapshot / incremental).
- Status legend visible con significados.
- Botones cycle por fila para probar estados en scaffold.

Overview grouped by integration + status (dark)

SQL Explorer < Pipelines

Dev Prod DE Data Engineer Engineer Save all

Pipeline submode persisted

Simple Pro

Search by name, integration, or tag... 5 of 5 DAGs by tag **by integration**

STATUS LEGEND • Draft • Submitted • Pending • Approved Use status button per row to cycle in scaffold.

BEATS_INTEGRATION 1

PIPELINE	TYPE	SCHEDULE	TASKS	STATUS	TAGS
dbo_AccountLogType	snapshot	Daily at 06:00 in 20h 20m	2	• Draft	cycle:snapshot

CRM_INTEGRATION 3

PIPELINE	TYPE	SCHEDULE	TASKS	STATUS	TAGS
dbo_AccountReference	snapshot	Every 3h in 2h 27m	3	• Draft	cycle:snapshot
gc_Game	snapshot	Every 3h in 2h 25m	3	• Draft	cycle:snapshot
tr_GameTransaction	incremental	Every hour in 20m	3	• Draft	cycle:incremental

DATA_SOURCES 1

PIPELINE	TYPE	SCHEDULE	TASKS	STATUS	TAGS
tr_DailyTransactionAmount	incremental	Every hour in 20m	3	• Draft	cycle:incremental

N 2 Issues X

Qué mirar:

- Toggle `by integration` activo.
- Columna Status visible por pipeline.
- Agrupación cambia de tags a integración sin romper conteos.

Status cycle en scaffold (dark)

The screenshot shows the Pipelines overview page in Data Studio. It displays several pipelines grouped by integration:

- BEATS_INTEGRATION** (1 pipeline):

Pipeline	Type	Schedule	Tasks	Status	Tags
dbo_AccountLogType	snapshot	Daily at 06:00 in 20h 20m	2	Submitted	snapshot
- CRM_INTEGRATION** (3 pipelines):

Pipeline	Type	Schedule	Tasks	Status	Tags
dbo_AccountReference	snapshot	Every 3h in 2h 27m	3	Draft	cyclicsnapshot
gc_Game	snapshot	Every 3h in 2h 25m	3	Draft	cyclicsnapshot
tr_GameTransaction	incremental	Every hour in 20m	3	Draft	cyclicincremental
- DATA_SOURCES** (1 pipeline):

Pipeline	Type	Schedule	Tasks	Status	Tags
tr_DailyTransactionAmount	incremental	Every hour in 20m	3	Draft	cyclicincremental

At the bottom left, there is a red box with the text "N 2 Issues X". At the bottom right, there is a purple circular icon with a white question mark.

Qué mirar:

- Pipeline de ejemplo cambia de Draft a Submitted .
- Cambio visible en badge de fila sin salir del overview.

Pipeline detail con status badge (dark)

The screenshot shows the detailed view for the pipeline **CRM_integration_dbo_AccountReference**. The pipeline is currently in Draft status, indicated by a badge. The pipeline details are as follows:

- CRM_integration_dbo_AccountReference** • Draft • snapshot
- CRM_integration • Every 3h • 3 tasks • owner: javier • US/Eastern • Next in 2h 13m

The pipeline consists of three tasks:

- #1 extract_crm_accounts (EXTRACT)
- #2 transform_account_to_datamodel (TRANSFORM)
- #3 cleanup_old_account_snapshots (DQA)

On the right side, there is a sidebar with the following sections:

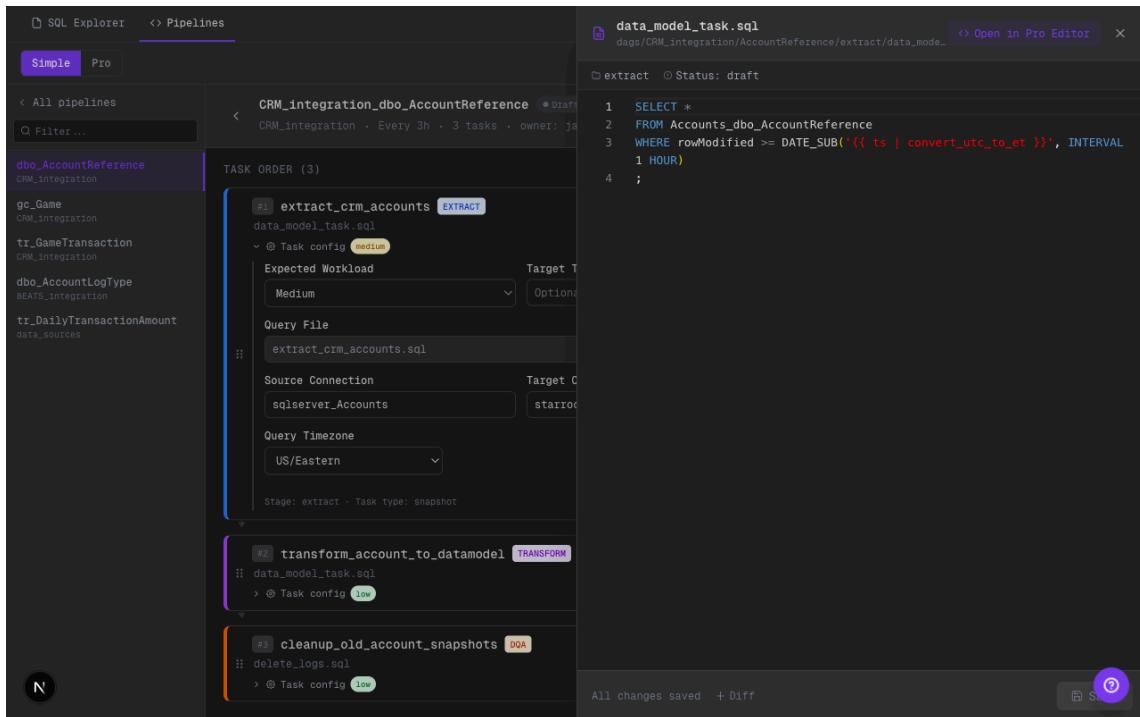
- Basic Info**: Owner (javier), Start Date (15/01/2024), Timezone (US/Eastern).
- Schedule**
- Tags**
- Channels**

At the bottom left, there is a black circular icon with a white letter 'N'. At the bottom right, there is a purple circular icon with a white question mark.

Qué mirar:

- En el header del detalle aparece badge de estado agregado.
- Estado visible junto a tipo y metadata DAG.

Task config extract (dark)



Qué mirar:

- Task EXTRACT con Expected Workload y Target Table Name .
- Target Connection visible para extract.

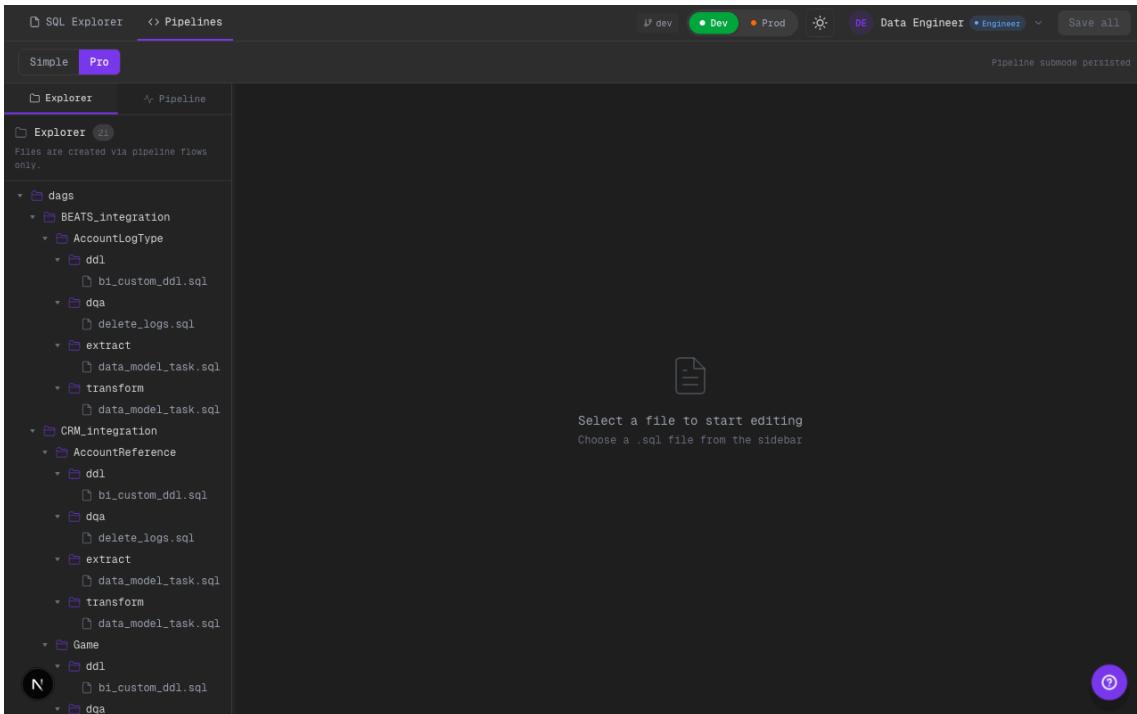
Task config dqa (dark)

The screenshot shows the Databricks SQL Explorer interface. On the left, there's a sidebar with a tree view of pipelines and data sources. The main area displays a pipeline named 'CRM_integration_dbo_AccountReference' in draft mode. The pipeline has three tasks: 'extract_crm_accounts' (EXTRACT), 'transform_account_to_datamodel' (TRANSFORM), and 'cleanup_old_account_snapshots' (DATA LOAD). The 'extract_crm_accounts' task is currently selected, showing its configuration details: Expected Workload (Medium), Target Table ('starrof'), Query File ('extract_crm_accounts.sql'), Source Connection ('sqlserver_Accounts'), Target Connection ('starrof'), and Query Timezone ('US/Eastern'). The status bar at the bottom indicates 'All changes saved'.

Qué mirar:

- Task `DQA` expone opciones de validación/alerta.
- Pie `Stage` + `Task type` visible para contexto rápido.

Pro explorer expandido (dark)



Qué mirar:

- Tree aparece expandido por default en primera carga.
- Stages `extract/transform/load/dqa/ddl` quedan visibles sin navegación extra.

Pro handoff context (dark)

The screenshot shows the Databricks SQL Explorer interface in dark mode. The top navigation bar includes tabs for 'SQL Explorer' and 'Pipelines', and status indicators for 'Dev' (green), 'Prod' (orange), and 'Data Engineer' (blue). The main area displays a pipeline named 'BEATS_integration_dbo_AccountLogType'. The pipeline tree is expanded, showing stages like 'dags', 'extract', 'transform', and 'dqa'. A specific task named 'data_model_task.sql' is selected, showing its SQL code:

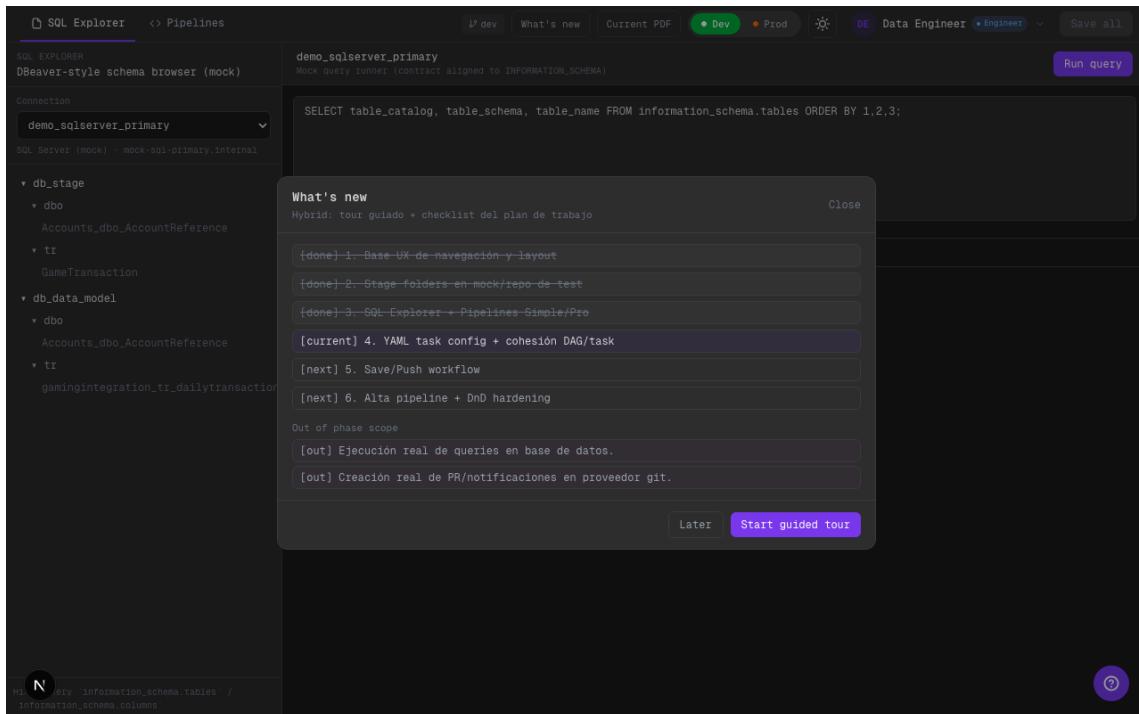
```
1 TRUNCATE TABLE db_data_model.Accounts_dbo_AccountLogType
2 ;
3
4 INSERT INTO db_data_model.Accounts_dbo_AccountLogType
5 SELECT
6     accountLogTypeID,
7     saga_hash,
8     saga_real_run_ts,
9     typeName,
10    rowCreated,
11    rowModified
12 FROM db_stage.Accounts_dbo_AccountLogType
13 WHERE saga_logical_run_ts = '{{ ts | convert_utc_to_et("US/Eastern") }}'
14 ORDER BY saga_real_run_ts ASC
15 ;
```

Below the code editor, there's a 'CHANGES' section with a note: 'No changes to display'. At the bottom left, a red box highlights the '2 Issues' button. On the right side, there's a purple circular icon with a question mark.

Qué mirar:

- No existe subtab lateral `Pipeline`.
- Aparece tarjeta `Pipeline context` con botón `Open pipeline handoff` para saltar al detalle en `Simple`.

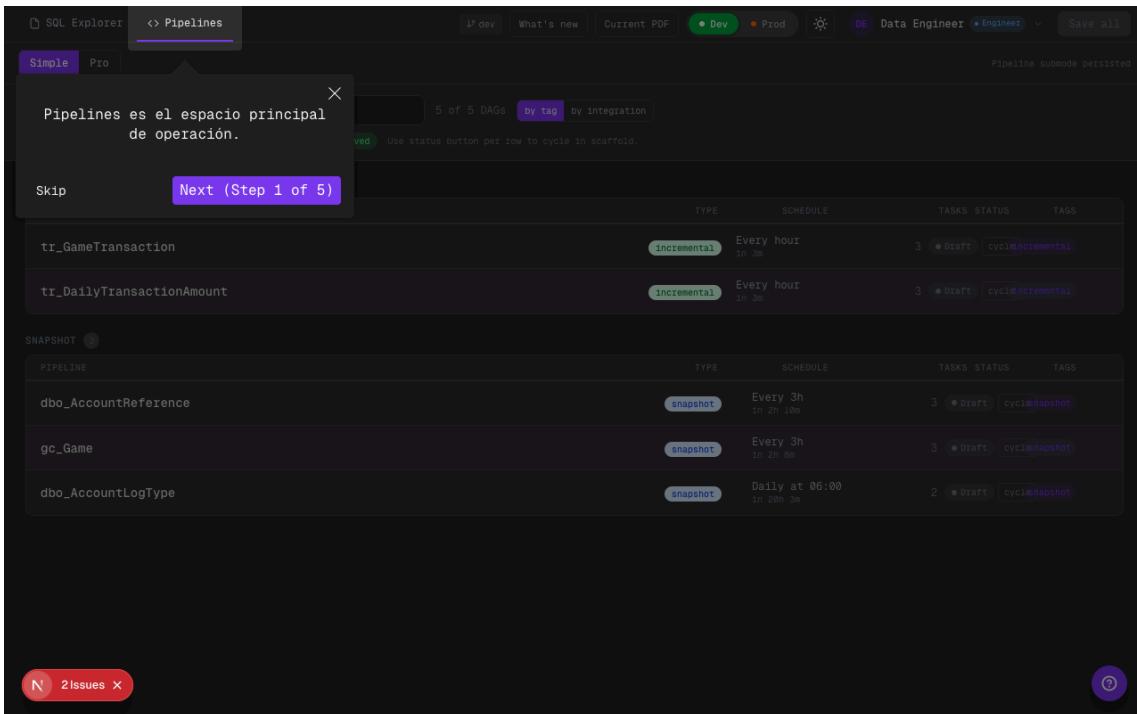
What's New checklist (dark)



Qué mirar:

- Modal con checklist del plan (done/current/next/out-of-scope).
- CTA `Start guided tour`.
- Link `Current PDF` visible en header.

What's New guided tour (dark)



Qué mirar:

- Tooltip de Joyride con navegación de pasos.
- Spotlight sobre elementos clave (Pipelines , agrupación, badges, PDF).

Límites scaffold

- Contrato YAML aplicado a nivel UI/configuración, sin validación semántica profunda runtime.
- El status agregado usa estado local de archivos/tasks mock.

TODO hooks

- Conectar validación de TaskConfig a parser real del YAML de contrato.
- Reusar status agregado en futuras vistas de auditoría/approvals para evitar lógica duplicada.

Calidad

- `cd ui && npm run lint` (ok, warnings existentes no bloqueantes).
- `cd ui && npm run build` (ok).
- Validación visual en Chrome (Playwright) con screenshots dark mode.