SOLUTION PROPOSAL

AI-BI Issue Desk

Proposed pilot for MILO-2 — AI Brand Performance Command Center, Nestlé Vietnam AI Reinvention

Confidential · **Version:** 1.0 · **Date:** 22 November 2026 · Based on the MILO-2 challenge brief

91.9%	1.09	30 of 30
of causes named correctly three days in (89.2% at the first alert); none confidently wrong	false alerts a week, against a budget of 2	metric questions answered exactly, including 3 correct refusals

Measured on synthetic data with deterministic tools. [3] Nothing has yet run on Nestlé's data.

AI-BI adds an assistant to the tools the Milo team already uses. Each morning it checks sales and distribution data, reports what moved and why in the team's channel (Slack or Microsoft Teams), and follows each issue as a card in the team's tracker (Jira, Linear or another). Numbers are computed from one shared dictionary that Nestlé approves; people decide. It is built in layers, so what is built carefully once for Milo can be replicated to similar features, to other departments and to every member of a team.

Item	Detail
Lead use case	MILO-2 — AI Brand Performance Command Center
Also submitted for	NH-2 Competitor Radar · SALES-1 eCom Next Best Move · SALES-2 Market Move Radar · CMS-1 Media Intelligence Hub · BUC-3 Trade Spend Forecasting · SC-3 DSA Predictive Analytics · FOOD-1 Insight-to-Action Copilot
Proposed solution	AI-BI Issue Desk (a proposed name, not a registered mark)
Applicant	AIT-SGN, a team: Phan Minh Trí, team lead, solution architect and developer · David Tran Hoang Nguyen, data and AI engineer
Contact and demo	Phan Minh Trí · tri.phan@ait-sgn.online · Demo: nestle-26.ait-sgn.online/milo-2/
Current stage	Prototype. Detection and metric tools measured on synthetic data; assistant written and not yet run; governance designed
Works with	Slack or Microsoft Teams · Jira, Linear or the tracker in use · SharePoint or Google Drive · Nestlé's data platform, read-only (section 04)
Proposed pilot	Milo only, 12 weeks, read-only data, four gates (section 08). The other seven follow as their own pilots (section 05)

Important. This prototype is not affiliated with Nestlé. All evidence is local and comes from synthetic data. No part of AI-BI has run on Nestlé's or any other company's data. Terms marked "(proposed)" are our proposal; they are agreed in writing with Nestlé before the pilot starts and are listed in Appendix B. This document promises no saving or return.

Contents

01. Executive summary 3
1.1. Three ideas behind this proposal 3
1.2. Why Nestlé can decide with confidence . . 3
1.3. What exists today 4
02. Scope, layer by layer 5
2.1. The brief, answered 5
2.2. Five layers and the control across them . . 6
2.3. What exists, by layer 6
2.4. In the pilot, optional, and out of scope . . . 7
03. How it works 8
3.1. One morning, across the layers 8
3.2. Six steps, one loop 9
3.3. One issue, from the scan to the decision . 9
3.4. What is computed and what is written . . . 10
3.5. How it differs from traditional BI 10
04. Integration with the team's existing tools 11
4.1. One picture 11
4.2. Tool by tool 11
4.3. Set-up in six steps 12
4.4. Changing a tool, and what the team does not have to do 12
4.5. Data 12
05. Replication: features, departments, people 13
5.1. Features: build the first carefully, then add similar ones 13
5.2. Departments: the same features for other teams 13
5.3. People: one person's way of using AI becomes everyone's 14
06. Roles, data access and AI governance 15
6.1. Roles 15
6.2. Rules that are enforced 15
6.3. What is sent to the model 16
6.4. Where it runs, and what is kept 16
07. Evidence 17
7.1. What was measured 17
7.2. Limits 17
08. The pilot 18
8.1. Twelve-week schedule (proposed) 18
8.2. Milestones and gates 18
8.3. Team 19
8.4. What Nestlé provides, by week 19
8.5. Measures and baselines 19
8.6. Acceptance before go-live 20
8.7. Main risks and how they are reduced . . . 20
09. Terms 21
9.1. Assumptions 21
9.2. Responsibilities 21
9.3. Support during the pilot 21
9.4. Ownership and handover 22
9.5. Commercial structure 22
10. Next steps 23
Appendix A. Capability overview by layer 24
Appendix B. Items to agree 25
Appendix C. Glossary and evidence register 26

Decision summary

For the sponsor: the decision on half a page

This submission	One proposal for eight use cases that run on the same layers. MILO-2 is the pilot proposed here. NH-2, SALES-1, SALES-2, CMS-1, BUC-3, SC-3 and FOOD-1 are the next replication steps, each its own pilot (section 05 and the case appendices).
You approve	A 12-week pilot on Milo only, with read-only access to sales and distribution data, in one channel and one tracker project the team already uses. Three named people: an approver (the Milo brand lead), an AI Data Governor and a data contact.
You get	A morning message on what moved and why; one issue card per unusual move; answers in the thread; a daily summary page. A shared dictionary and a check set that Nestlé approves and keeps. At week 12: results against your own baseline, and a note on what a second team would reuse.
It needs from Nestlé	Data extracts in weeks 1–2 and the last ten known issues; marking of replayed alerts in weeks 3–5; approval of the brief each working day and a 30-minute weekly review in weeks 6–10 (8.4).
Your risks, and the exits	Results on real data may be lower than on synthetic data; alerts may be too many; running inside Nestlé's tenant has not been demonstrated. Four gates, at week 2, week 5, every live week and week 12. The pilot can stop at each, and Nestlé keeps what was built (8.2, 9.4).
Not yet proven	The assistant has not run end to end. No connection to Nestlé's tools or data exists. There is no result on real data. Weeks 1–5 answer these before anything goes live.
Cost	A fixed fee paid by gate, so stopping at a gate stops the spend (proposed). The figure is given at Matching Day (9.5).

> **The proposal in one sentence.** AIT-SGN proposes a 12-week, read-only pilot on Milo that puts detection and metric tools already measured on synthetic data behind an assistant in the team's own channel and tracker, replays them on Nestlé's history before anyone sees an alert, and ends with a written decision and a note on what a second team would reuse.

Matching Day questions and where they are answered: 1 the problem → 02 · 2 what the pilot proves → 08 · 3 the solution → 03 · 4 data and integrations → 04 · 5 where it has been tested → 07 and 10 · 6 what Nestlé keeps → 09.

01. Executive summary

What is proposed, and what exists today

MILO-2 asks for one daily view of sales, distribution, penetration and share in which unusual moves are flagged, explained and turned into priority actions, so that issues are spotted the same day instead of weeks later. Today the team consolidates and analyses the data itself, and business discussions wait for it. [1]

1.1. Three ideas behind this proposal

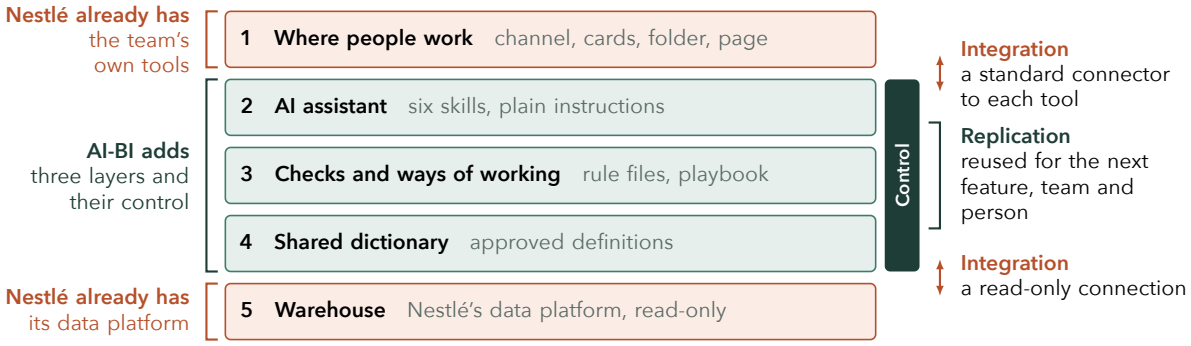


Figure 1: The project in layers. In every figure, orange is what Nestlé already has and green is what AI-BI adds. Control means approvals, logs and replay.

- **It is built in layers.** Every capability, every piece of work and every week of the schedule belongs to one of five layers, or to the control across them. The layers say what exists, what is built once and what is made for Nestlé (section 02).
- **It fits the tools the team already uses.** Nobody opens a new application. The assistant reports in the team's channel and keeps its issues as cards in the team's tracker. Each tool is reached through a standard connector; changing Linear for Jira is a change of one connection (section 04).
- **It is made to be replicated.** The first feature is built carefully, with approved definitions, a replay and an approval gate. After that, a similar feature is an approved file change, another department reuses the assistant and the control, and another person simply joins the channel (section 05).

1.2. Why Nestlé can decide with confidence

- **The misses are reported beside the hits.** Section 07 lists what the checks catch on the first two days and what they do not: competitor promotions and most pack-mix shifts.
- **Nothing is shown before it is replayed.** The checks are scored on Nestlé's own past months before the wider team sees an alert (section 08).
- **Read-only, and people decide.** No write access to any Nestlé system. Definitions, briefs and changes to checks take effect only after a named person approves; in the prototype this is enforced in code (section 06).
- **Numbers are computed, not generated.** Every figure traces to a query and a definition; a question outside the definitions is refused (section 03).
- **Four gates, each with a way out,** and the conditions that count as failure written down before the start (section 08).
- **Nestlé keeps what is built,** as readable, versioned files, and can check the evidence itself in one session (sections 09 and 10).

1.3. What exists today

AI-BI is a prototype in three different states. This proposal keeps them apart throughout.

State	What	Evidence
Built and measured	Layers 3 to 5 for MILO-2: daily snapshot, detection of unusual moves, exact four-way split, driver naming, a governed metric layer, an approval gate. Synthetic data, deterministic tools.	[2] [3] [6]
Written, not yet run	Layers 1 and 2: four skills (daily scan, investigate, summary page, file exchange) and the connections to a channel, a tracker and a data server.	[4]
Designed, not built	The governance: a company-wide dictionary with a named approver, the saved check set as its own layer, three roles, two configuration repositories, the set-up and propose-change skills.	[5]

The measured part is in section 07: 1.09 false alerts a week, 89.2% of causes named correctly at the first alert and 91.9% three days after an issue starts, with none confidently wrong, and 30 of 30 metric questions exact. The assistant-led scan, end to end, is not measured: it has not run. [3]

02. Scope, layer by layer

Matching Day question 1: the business problem, and what is in scope

Every line of the brief has an answer, and every answer sits in a layer with a stated state.

2.1. The brief, answered

The left column quotes or paraphrases the public brief. [1]

The brief asks for	Our answer	Layer and state today
One daily view of sales, distribution, penetration and share	A daily snapshot: sell-out against expected, sell-in, distribution, price against list, by region and channel. Penetration and share are shown with their panel period. Delivered as the morning message and the summary page.	Layers 5 and 1. Snapshot measured; message and page written, not run.
AI flags unusual moves	Saved checks compare each day with an expected value that knows the calendar. Moves that matter are kept and ranked by value at stake.	Layer 3. Measured on synthetic data: what is caught and what is not, by cause (07).
Explains drivers	An exact split into distribution, velocity, price and mix, then a cause from six written signatures, or "unexplained".	Layer 3. Measured: 89.2% correct at the first alert, 91.9% three days later, 0 confident wrong (07).
Suggests priority actions	One proposed action per move from a playbook file, with owner, deadline and priority. Proposals only; the brand lead approves (3.3).	Layer 3 and control. Built, not scored; the playbook is a stand-in.
Much less time answering questions	Questions asked in the thread are answered from the shared dictionary, with the query kept.	Layers 4 and 2. Metric layer measured: 30 of 30; thread answers written, not run.
Issues spotted the same day; time freed for programmes and customers	The pilot measures both against baselines recorded in weeks 1–2 (8.5).	To be measured in the pilot.
Target users: Milo brand and commercial teams	They work in layer 1 only: the channel and the cards. The brand lead approves; an AI Data Governor owns the definitions (06).	Roles designed.
Data: sales, distribution, penetration, share; synthetic data may be provided	Read-only extracts (4.5). If Nestlé provides synthetic or anonymised data, the replay is rerun on it before any real data (10).	Layer 5. Prototype runs on its own synthetic data.
Environment: Microsoft Azure and Microsoft Copilot. Technologies: generative AI, machine learning	Layer 1 is whatever the team already uses; the model can run inside Nestlé's Azure tenant (04). Statistics and written rules compute the numbers; the generative model writes and investigates.	Written for Slack and Linear; the rest not demonstrated.

Why this is harder than a dashboard. The data does not move at one speed: sales are daily, the household panel is periodic and late. Most unusual days are not issues: Tết, back to school, month end and planned promotions all move the numbers. A data problem looks like a sales problem: a distributor file that arrives short reads as a drop. Explanations lose trust fast: a confident wrong cause is worse than "we do not know yet". And a dashboard waits to be opened, while "priority actions" need an owner and a state.

2.2. Five layers and the control across them

Layer 1 is where people work; layer 5 is the data. [5] Each layer talks only to the one beside it, so the assistant never touches raw data directly. Codes refer to Appendix A.

Layer	What it does	Used by · set up by	Belongs to
1 Where people work L1-01 to 05	The team's channel, issue cards, shared folder and summary page. The only layer people see.	The whole team · channel and tracker administrators	Nestlé's existing tools
2 AI assistant L2-01 to 07	Six skills: daily scan, investigate, summary page, file exchange, set-up, propose a change. Drafts the daily brief.	The team, through the channel · the deployment team	AI-BI foundation
3 Checks and ways of working L3-01 to 11	What counts as unusual and how to handle it: expected value, alert rules, four-way split, driver signatures, playbook, the team's instructions.	The assistant · the AI Data Governor and the team lead	Foundation method; content made for Nestlé
4 Shared dictionary L4-01 to 05	What each number means: definitions with a named approver; refusal of questions outside them.	The assistant · the AI Data Governor	Foundation mechanism; definitions made for Nestlé
5 Warehouse L5-01 to 04	Sales, distribution, stock, calendars and panel data, read-only; the scan log.	The shared dictionary · the data team	Nestlé's data platform
Control, across 2 to 4 CT-01 to 06	Approval gate, decision log, scan log, replay, limits enforced in two places, two configuration repositories.	Everyone who approves · the deployment team	AI-BI foundation

2.3. What exists, by layer

All 38 capabilities are in the pilot scope. Measured means on synthetic data.

Layer	Measured or tested	Built, not scored	Written, not run	Designed	Total
1 Where people work	–	–	5	–	5
2 AI assistant	–	1	4	2	7
3 Checks and ways of working	9	1	–	1	11
4 Shared dictionary	3	–	–	2	5
5 Warehouse	4	–	–	–	4
Control	3	–	–	3	6
All layers	**19**	**2**	**9**	**8**	**38**

The lower layers, which compute the numbers, are measured. The upper layers, which connect to the team's tools, are written and wait for a workspace to run in. That is what the pilot provides.

Part	What it is	Why it matters
Existing foundation	The kernel (approval gate, decision log, read-only data access, rule files, data server, evaluation harness); the detection method, the four-way split and the driver scoring; the metric layer; four skills.	Not built again for the pilot. The same kernel already carries 27 prototype packs (section 05).
Extension in the pilot	The software that enforces the dictionary; the scan log; the set-up and propose-change skills; limits enforced at the warehouse; the replay on real history; the connector for the tracker Nestlé uses.	Built once. It stays in the foundation and serves the next feature and the next team.
Made for Nestlé	Milo's definitions; thresholds and the seasonal profile learned from Nestlé's history; driver signatures and the playbook written with the team; the connections to Nestlé's channel, tracker, folder and data.	Nestlé owns it (9.4). It is the only part a second team writes again.

2.4. In the pilot, optional, and out of scope

In the pilot	Optional, agreed separately	Out of scope
Milo only. All five layers and the control: the daily scan, issue cards, answers in the thread, the summary page and file exchange.	A second brand or category (section 05).	Writing to any Nestlé system.
The dictionary and check set for the measures the team uses, approved by the AI Data Governor.	An entry point from Microsoft Copilot.	Changing a price, promotion or order; contacting a distributor or retailer.
A replay of the checks on Nestlé's past months before go-live.	New driver signatures learned from the issue log.	Replacing existing dashboards.
A weekly review in which the team rates the alerts.	Forecasting of sales or spend.	Loading data into the warehouse; the data team owns this.
Output in Vietnamese or English, as the team chooses.	The seven other use cases in this submission, each as its own pilot (section 05).	Any personal data beyond the names of card owners and approvers.

The pilot proposed here is Milo only. This proposal is submitted for eight use cases; the other seven are the next replication steps, described in section 05 and in their case appendices.

03. How it works

Matching Day question 3: the proposed solution

Each morning the assistant runs saved checks through the shared dictionary and tells the team what moved, in the tools the team already uses. This section follows the AI-BI architecture review of 15 November 2026. [5] Where the review names a product (Slack, Linear, Google Drive), this proposal names the role it plays: channel, issue cards, shared folder.

3.1. One morning, across the layers

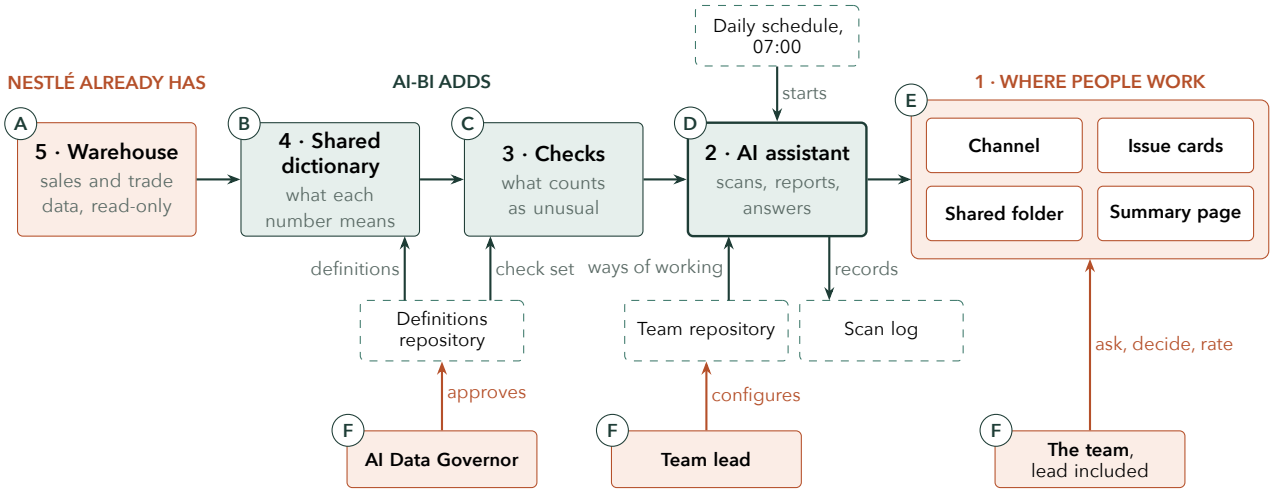


Figure 2: How it works. Numbers are layers; letters are the steps of a morning. Dashed boxes are configuration and records.

A. Data. Sales and trade figures sit in Nestlé's warehouse. The assistant can only read, and only through the dictionary.

B. Definitions. The shared dictionary says what each number means. The AI Data Governor approves it; it is kept in the definitions repository.

C. Checks. The saved check set compares today with an expected value that knows the weekday, month end, Tết, back to school, public holidays and the planned promotion and price calendars. A late or partial file is reported as a data problem, not analysed as performance.

D. Scan and report. At 07:00 the assistant runs the checks, writes one message to the channel and opens or updates one issue card per move, ranked by value at stake. Every scan is recorded.

E. Where people work. People read the message, ask "why" in the thread, pass files through the shared folder and open the summary page. If the data cannot answer, the assistant says what is missing and does not estimate.

F. People decide. The brand lead approves the daily brief before it is circulated. The team rates the alerts. The AI Data Governor and the team lead configure, in two separate repositories.

One gap is open and stated here. In the current scaffold the scan skill still writes its own queries against the tables. The design routes it through the dictionary and the saved checks, and the pilot would build it that way. [4] [5]

3.2. Six steps, one loop

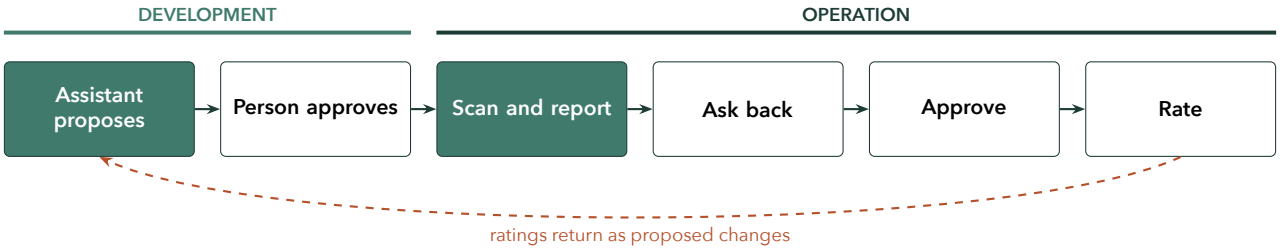


Figure 3: Development and operation in AI-BI: six steps and one loop back. The assistant does the two filled steps; people approve and decide.

Development is two steps: the assistant reads the data structure, interviews the AI Data Governor and proposes the dictionary and the checks; a person approves them. A new number later is the same two steps, not a new project. **Operation** is four: the assistant scans and reports, people ask back, approve what should be done, and rate each alert. The ratings return as proposed changes to the checks, which again wait for approval. That is the single loop.

3.3. One issue, from the scan to the decision

The figures below are the output of the measured tools for 26 August 2026 in the synthetic data. [2] The messages show how the assistant would present them; they are illustrations, because the assistant has not run. Distributor names are invented, and the playbook wording is a stand-in.

1. The morning message, in the team's Slack or Teams channel

> **Assistant** 07:00, channel **Milo daily scan, 26 Aug 2026.** 22 of 22 source files arrived.
> National sell-out VND 28.77 bn against VND 31.60 bn expected (−9.0%, inside the normal range).
> **2 unusual moves**, ranked by value at stake:
> 1. **Ho Chi Minh City: sell-out 27.4% below expected**, VND 2.33 bn at stake today. Likely cause: distributor out of stock (high confidence). New today. *Card opened.*
> 2. Mekong Delta, traditional trade, RTD180: 13.9% above expected, VND 68.6 m at stake. Likely cause: distribution gain (high confidence). Day 2. *Card updated.*
> Panel, four-weekly: penetration 24.1%, value share 38.7% for 29 Jun–26 Jul, delivered 16 Aug. Not a reading for today.

2. The issue card, in the team's tracker (Jira or Linear)

Field	Content
Title and state	Ho Chi Minh City: sell-out 27.4% below expected · *Triage*
Actual against expected	VND 6.18 bn against VND 8.51 bn; gap −VND 2.33 bn
Four-way split	Velocity −3,012 m · mix +666 m · price +13 m · distribution +6 m. Sums to the gap.
Where it sits	Traditional trade carries 87.5% of the gap; one distributor, NPP-HCM-01, carries 71.8%; RTD180, RTD110 and CAN240 carry 77.5%.
Likely cause	Distributor out of stock, high confidence: 65% of the lost sales sit in distributor and SKU combinations that started the day with less than one day of stock.
Proposed action	Priority 1, today, for the regional sales manager with customer service: expedite an order for the SKUs without stock, find out why the last order was held, and cover key outlets from a neighbouring distributor until stock arrives.
Evidence	The alert, the driver rule, four stock records, and the queries that produced each figure.

3. A question in the thread

> **Team member** 08:12 Is this only one distributor?
> **Assistant** 08:13 Mostly. NPP-HCM-01 sold VND 0.74 bn against VND 2.41 bn expected (−69.4%) and carries 71.8% of the gap. The other two HCM distributors are 13.5% and 14.5% below expected and carry 15.8% together. Queries added to the card.

4. The decision. The brand lead reads the drafted brief, edits it and approves. Only then is it circulated; the approval is recorded with name and time, and the card moves to *Approved*. The owner acts outside AI-BI and marks the card *Done*.

5. The rating. In the weekly review the team marks the alert useful. Had it been noise, repeated ratings would come back as a proposed change to the check, with its replay result, for the AI Data Governor to approve.

When the data does not explain a move. The next day the same tools flag national traditional-trade sell-out at 18.9% below expected, with VND 4.45 bn at stake, and no signature accounts for it. The card says "unexplained, needs review" and proposes a review assignment. No cause is named.

3.4. What is computed and what is written

The numbers are computed first; the model writes second.

Computed by tools from rule files (layers 3 to 5)	Written by the assistant (layer 2)
Expected values, thresholds and which moves are kept	The message, the card text and the daily brief
The four-way split, exact to the last VND	The answer in a thread, in Vietnamese or English
The driver, from six written signatures, or "unexplained"	Deeper analysis when one issue is investigated
Answers to metric questions, or a refusal with the reason	Proposed changes to definitions and checks

3.5. How it differs from traditional BI

These are the design goals of AI-BI; none has been measured, because the assistant has not run. The traditional flow is a common pattern, and each company may differ. [5]

	Traditional BI	AI-BI
Knowing there is an issue	Someone has to open the dashboard and notice	The assistant scans each morning and reports to the channel
Asking why	Send a request to an analyst and wait	Ask the assistant in the thread, in plain words
Meaning of a number	Scattered across dashboards and reports, easily inconsistent	One shared dictionary approved by the AI Data Governor, used by every answer
Follow-up	Email or chat, no record	Each unusual move is a card with an owner and a state
Learning from results	No way back: the dashboard does not know which alerts were right	The team's ratings return as proposed changes to the checks, approved by a person
Length of the flow	9 steps across two phases, 5 loops back	6 steps across two phases, 1 loop back

AI-BI does not replace the dashboard. For free exploration across many dimensions a dashboard is still the better tool.

04. Integration with the team's existing tools

Matching Day question 4: data and integrations

AI-BI is not a new application: it connects to the two layers Nestlé already has, the tools the team works in and the data platform. Each connection uses a published, standard connector (the Model Context Protocol), not integration code written for this project.

4.1. One picture

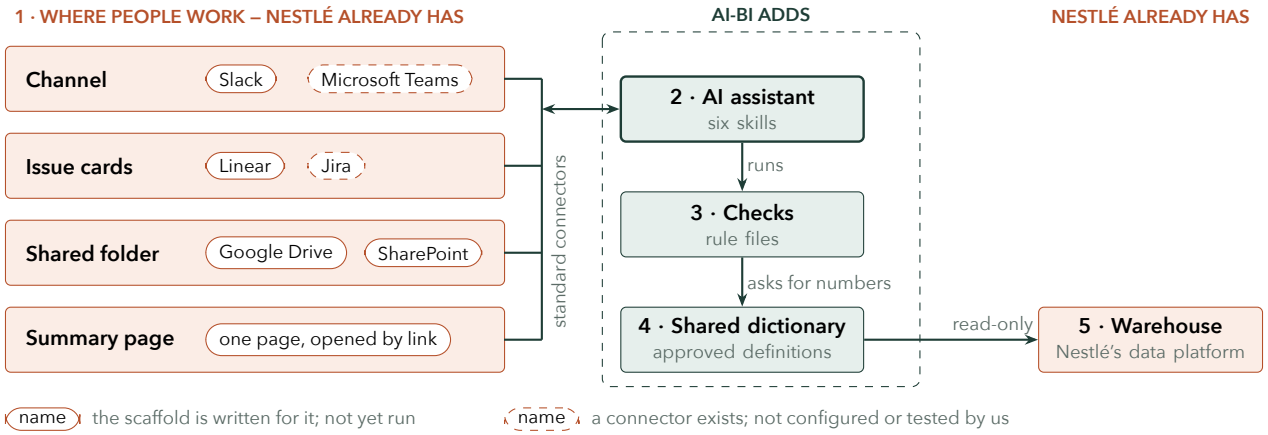


Figure 4: Integration. The team keeps its tools and the assistant comes to them. Only layer 2 touches the team's tools, and only layer 4 touches the data.

4.2. Tool by tool

Nestlé's tool	What AI-BI does there	How it connects	State today
Slack (channel)	Posts the morning message; answers questions in threads.	Claude in Slack, with a scheduled daily run	The scaffold is written for it; not run. [4]
Microsoft Teams (channel)	The same.	Microsoft 365 connector	Possible. Whether the assistant can post into a Teams channel is to be confirmed in the deployment review. [10]
Linear (issue cards)	Opens one card per move, updates it the next day, reads its owner and state.	Linear's own connector	The scaffold is written for it; not run. [4]
Jira (issue cards)	The same.	Atlassian's own connector	The connector exists; not configured or tested by us. [9]
SharePoint or Google Drive (shared folder)	Reads a file a person puts in the inbox folder; writes results to the outbox folder.	A connector added by a workspace administrator	The scaffold is written for it; not run. [4]
Data platform (warehouse)	Read-only queries, through the shared dictionary.	A data server inside Nestlé's tenant, using the same protocol	The prototype server is measured on synthetic data; no Nestlé connection. [2]
Model	Writes and investigates; computes no figure.	Claude, through Microsoft Foundry in Nestlé's Azure tenant (proposed)	Not demonstrated.

4.3. Set-up in six steps

These are the steps written in the scaffold. They have not been carried out in any workspace. [4]

	Step	Who
1	**Data server.** Host the read-only data server inside Nestlé's tenant, behind Nestlé's sign-in.	Deployment team
2	**Connection file.** Enter the server's address and credentials.	Deployment team
3	**Tracker.** Create the project, its seven states (Triage, Investigating, Explained, Action proposed, Approved, Done, Not an issue) and the labels for source and confidence.	Team lead, deployment team
4	**Channel.** Create the channel, add the assistant, enable the plugin for it.	Workspace administrator
5	**Shared folder.** Add the connector; create the inbox and outbox folders.	Workspace administrator
6	**Schedule.** Ask the assistant, in the channel, to run the daily scan every morning at 07:00 Vietnam time.	Team lead

4.4. Changing a tool, and what the team does not have to do

When a tool is swapped, for example Linear for Jira	The team does not have to
Changes: one entry in the connection file, and the tracker's name, project, states and labels in the skills: four files, about 250 lines of plain instructions.	Learn a new application or sign in to one.
Does not change: the dictionary, the checks, what a card contains, the approvals, the logs and the replay. Layers 3 and 4 do not know which channel or tracker sits above them.	Open a new dashboard, or give up an existing one.
	Move data out of Nestlé's platform, change how it works its cards, or write prompts.

4.5. Data

Need	Likely source (to confirm)	Pilot access
Daily sell-out by region, channel, distributor, SKU	Distributor management system	Read-only daily extract
Daily sell-in	SAP or the sales data warehouse	Read-only daily extract
Numeric and weighted distribution	Coverage data or retail audit	As often as it exists
Distributor stock	Distributor management system	Daily
Feed log (which files arrived, how complete)	Data platform	Daily
Price, promotion and media calendars	Revenue management, trade marketing, media agency	On change
Competitor activity	Field reports, retailer leaflets, platform price scans	As captured
Penetration and share	Household panel	Per delivery; licence terms to be confirmed

All inputs are aggregated sales and trade data. The pilot needs no personal data beyond the names of card owners.

One installation per company, with the data staying in that company's infrastructure. No write access to any Nestlé system. No training on Nestlé data.

> **Not demonstrated.** The Azure deployment; any connection to Slack, Teams, Jira, Linear, SharePoint or Copilot; any live connection to Nestlé's sales systems; real data volumes and latencies; the skills running under a model; an end-to-end scan. [4] [5]

05. Replication: features, departments, people

Built carefully once, then copied

The pilot is one feature for one team, built so that the next feature, the next department and the next person reuse most of what the first one needed. The layers say what is reused each time. This is the design and the evidence behind it; replication has not yet been carried out at any company.

Replicating to	Reused as it is	What is new	Who approves
A similar feature for the same team: a new number, check or cause	Layers 1, 2 and 5, the control, and the mechanisms of layers 3 and 4	One definition, check or signature, as a file change	The AI Data Governor, after the replay has scored it
Another department	Layer 2, the control, the mechanisms of layers 3 and 4, and every definition already shared	Its channel and tracker project (1); its check set, playbook and instructions (3); its own definitions (4); its data (5)	That team's lead and the AI Data Governor
Another person in a team	Everything	Nothing. The person joins the channel.	The team lead, who grants access

5.1. Features: build the first carefully, then add similar ones

For the first feature the pilot builds everything around it: approved definitions, thresholds learned from history, the replay, the acceptance scenarios, the approval gate and the logs. A later feature of the same kind reuses all of that.

To add	What changes	Approved by	Checked by
A new number	One definition in the dictionary (layer 4)	AI Data Governor	The agreed question set, answered exactly
A new check or threshold	One entry in the detection rule file (layer 3)	AI Data Governor	The replay on past months
A new cause	One driver signature (layer 3)	AI Data Governor	The replay; the count of confident wrong causes
A new recommended action	One playbook entry (layer 3)	Team lead	The weekly review

Evidence. The prototype repository holds 27 use-case packs in 11 plugins, all on one kernel: the approval gate, the decision log, rule files, the data server and the evaluation harness. Each pack added its own tools, rule files, synthetic data and a scored evaluation; none needed a new platform. This shows that the structure repeats. The packs are prototypes on synthetic data and are not offered here as evidence of how well each case performs. [7]

5.2. Departments: the same features for other teams

Shared by the whole company: one installation, the assistant's skills, the control, one definitions repository (so a number means the same in every department) and one AI Data Governor. **One per team:** a channel and tracker project, a team repository, a check set and a playbook, a team lead.

Several other briefs in this programme draw on the same sales, price, promotion and media data as MILO-2 and could share one governed metric layer. [8] Each has a prototype pack on synthetic data and its own two-page case appendix, attached to this proposal, that maps the case onto the layers: what the team sees, what is reused from the Milo pilot and what the team writes. [7]

Challenge	What the check watches	What changes	Read more
NH-2 Competitor Radar	Each week: launches, claims, prices and promotions from a fixed watch-list	Sources are outside items, not warehouse tables. The dictionary adds a competitor and event list. Weekly, not daily.	Case appendix NH-2
SALES-2 Market Move Radar	Each day: shelf price and promotion moves by channel against history	Needs a capture step for price tags upstream. Response scenarios are proposals for a person.	Case appendix SALES-2
SALES-1 eCom Next Best Move	Each day: GMV, contribution and return by platform, category, brand and creator	First-party platform data. The card is a ranked proposal, which is harder than detection.	Case appendix SALES-1
CMS-1 Media Intelligence Hub	During a campaign: delivery against plan; past learning retrieved on request	Depends on getting planning data back from the agency platform.	Case appendix CMS-1
BUC-3 Trade Spend Forecasting	Each month: projected spend against budget by channel	The check set needs a forecast model. Cards go to Sales Control and F&C.	Case appendix BUC-3
SC-3 DSA Predictive Analytics	Each day: distributor cover, sell-out against forecast, late receipts, credit	Alerts come before the failure. A daily cap of alerts per RSM.	Case appendix SC-3
FOOD-1 Insight-to-Action Copilot	No daily scan; the investigation path only, across research, market and complaint sources	Adds document retrieval beside tables. The weakest fit of the seven.	Case appendix FOOD-1

This proposal is submitted for all eight use cases because they run on the same layers. The Milo pilot goes first. Each of the other seven is then its own pilot, with its own sponsor, its own measures and its own gates on the pattern of section 08; its case appendix describes how it plugs into the layers. Each use case has a demo of its prototype on synthetic data at nestle-26.ait-sgn.online/*use-case code*/, for example nestle-26.ait-sgn.online/nh-2/.

5.3. People: one person's way of using AI becomes everyone's

- **Nobody has to learn prompting.** The way of working is written once, as a skill or as the team's instructions. Every member meets the same assistant in the channel.
- **What one person finds, the team keeps.** When someone rates an alert as noise, or an investigation turns up a pattern worth watching, the assistant turns it into a proposed check. Once approved, it runs every morning for the whole team.
- **Answers stay where the team works.** A question, its answer and the queries behind it are kept in the thread and on the card. The next person reads them and does not have to ask again.
- **Two people keep it consistent.** The team lead decides how the assistant works for the team; the AI Data Governor decides what the numbers mean. No individual can change either alone.

The skills are written and the propose-change loop is designed; neither has run. [4] [5] The pilot measures this level directly (8.5).

06. Roles, data access and AI governance

Who can do what, what the model sees, and where it runs

Nobody opens an application called AI-BI, and nothing the assistant proposes takes effect without a named person. The team works in layer 1 only; the team lead also sets the team's instructions in layer 3; the AI Data Governor owns layer 4; the deployment team connects layers 1 and 5. [5]

6.1. Roles

Role	Works in	Can do	Cannot
Team member (Milo brand and commercial teams)	Channel, issue cards, shared folder	Read the morning message; ask in a thread; update a card; send a file; rate an alert.	Change a definition, a check or how the assistant works.
Team lead (in the pilot: the Milo brand lead)	The same, and the team repository	Everything a team member can. Approve the daily brief before it is circulated; approve proposed actions; set the team's instructions.	Change what a number means.
AI Data Governor (one named person at Nestlé)	Definitions repository, set-up workspace, issue cards	Approve the shared dictionary and the check set; approve or reject each change the assistant proposes.	Circulate a brief or decide an action for the team.
Deployment team (Nestlé IT with AIT-SGN)	Plugin repository, administration	Install, connect the warehouse read-only, grant and remove access.	Approve definitions, briefs or actions.
The assistant	Runs on a schedule and when asked	Run the saved checks and report; query read-only to investigate; draft a card, an answer, the summary page and a brief; propose a change.	Change a definition or check without approval; state a cause as known unless the data accounts for it; assign owners or close cards; circulate a brief before approval; write to any system.

6.2. Rules that are enforced

- **Read-only at the source.** The assistant's connection to the data has no write permission. In the prototype, data access is read-only by construction. [6]
- **Through the dictionary.** Every answer uses the approved definition. A question outside the definitions is refused with the reason: a daily penetration figure does not exist, and the assistant says so.
- **A person approves.** An action cannot be committed until a named person has decided, and the decision log cannot be edited. The gate, the log and read-only access are enforced in code and pass 12 of 12 automated tests. [6]
- **A cause needs evidence.** A cause is named only when its evidence clears a written threshold; otherwise it is marked "verify first" or "unexplained". A competitor promotion that coincides with a drop is never reported above "verify first".
- **Everything leaves a record.** Each scan logs which checks ran and what they returned. A number in a card or an answer carries its query and its definition.
- **In two places.** The design blocks what the assistant may not do at the warehouse and at the assistant's tools, not only by instruction. Today the approval rule is in code; the other limits are written into the skills. [5]

6.3. What is sent to the model

This is the design, to be confirmed in Nestlé's deployment review. [5] Not sent in any function: personal data, credentials, or data the dictionary does not expose. All inputs are aggregated sales and trade data.

Function	Sent to the model	Kept afterwards
Daily scan	Results of the saved checks (aggregated figures by region, channel, source and SKU), the definitions used, the team's instructions.	The scan log: which checks ran and what they returned.
Investigate	The question, the results of read-only queries made through the dictionary, the card's history.	Queries and answers, on the card.
Summary page	The snapshot figures and the open cards.	The page, rebuilt daily.
File exchange	The file a person put in the inbox folder.	The output file in the outbox folder.
Propose a change	The team's ratings, the check concerned, its replay result.	The proposal, its approver and the date.
Set-up	Table and column names, and the AI Data Governor's answers.	The proposed dictionary.

6.4. Where it runs, and what is kept

- **Processing (proposed).** The assistant runs inside Nestlé's Azure tenant and reaches the model through Nestlé's own agreement. The data stays in Nestlé's data platform. AIT-SGN works through the access Nestlé grants and keeps no copy of Nestlé's data.
- **No training.** Nestlé's data is not used to train or tune any model.
- **Retention (proposed).** Cards and thread answers live in Nestlé's tracker and channel under Nestlé's own retention rules. Scan logs are kept for the pilot and handed over at its end.
- **Usage.** One scheduled scan per day; investigations on request. Model usage per run is reported weekly (proposed). Nestlé can pause the schedule at any time; the checks are deterministic tools and can still be run without a model, which is how they were measured.

07. Evidence

Matching Day question 5: where it has been tested

The tools that compute the numbers have been measured on synthetic data, and the misses are reported with the hits. The assistant itself has not run.

7.1. What was measured

The MILO-2 pack (one prototype use case: synthetic data, tools, rule files and an evaluation) generates 28 months of daily synthetic data with 44 planted issues in the last 20 months. Thresholds, noise levels and seasonal factors come from the first eight months only. The tools are then replayed one day at a time. [2] [3]

Measure	Result	Reading
Flagged within 7 days	81.8%	7 of 44 were not flagged within 14 days.
What it catches on day one or two	6 of 6 each: feed gaps, distributor stock-outs, distribution losses, unplanned price changes	Also all 8 issues of the two driver types it had never seen.
What it misses	Competitor promotions 0 of 6; pack-mix shifts 2 of 6	Their effect on a region and channel total sits inside ordinary daily noise for the first two days.
False alerts	1.09 per week	Budget was 2.
Driver named correctly	89.2% (33 of 37) at the first alert; 91.9% (34 of 37) three days after the issue starts	0 confident wrong causes. A competitor report can reach the team up to two days after a promotion starts, which is why the later figure is higher. On unseen driver types it abstained 7 times out of 8.
Metric questions	30 of 30 exact	Includes 3 it correctly refused: a daily penetration figure, share by distributor, distribution summed over SKUs.
Approval gate and audit log	12 of 12 automated tests	A commit is refused until a named person approves. [6]

7.2. Limits

- **Synthetic data only.** The data, the detection policy, the driver signatures, the playbook and the metric definitions are stand-ins written without access to the Milo team's own.
- **The numbers measure the tools, not the assistant.** The model's narrative, its reading of free-text questions and the brand lead's edits are in none of them. If the assistant analyses on its own, two runs on the same day can give different issues; that is why the design runs saved checks.
- **No expert review.** No one at Nestlé or elsewhere has checked the rules or the outputs.
- **Usefulness is unknown.** Whether a real team would rate these alerts as worth reading has not been tested.

08. The pilot

Matching Day question 2: what the pilot proves and how we would know

Twelve weeks, one brand, read-only, with four gates and the conditions for failure written down before the start. One brand lead as approver, one AI Data Governor, one channel. The schedule starts when access is granted; it holds if Nestlé's inputs in 8.4 arrive in their week.

8.1. Twelve-week schedule (proposed)

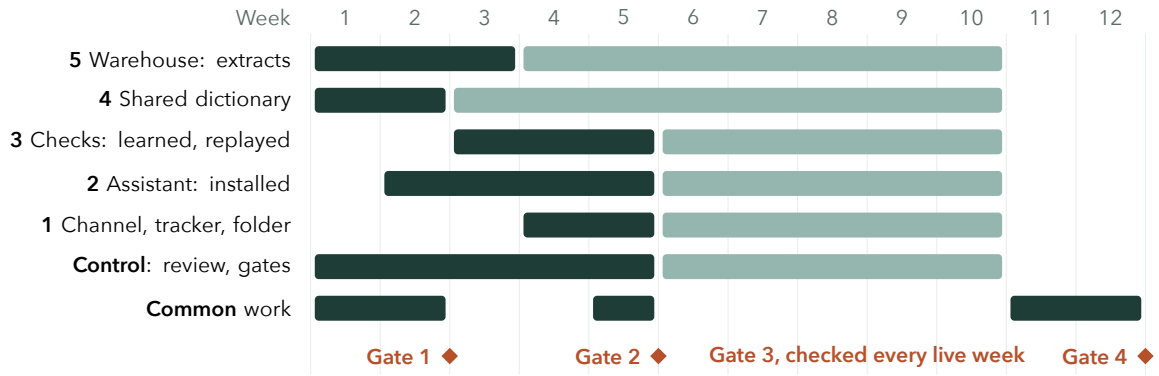


Figure 5: Proposed schedule by layer, built from the data upwards. Dark: set-up (extracts agreed and checked, dictionary approved, checks learned and replayed, assistant installed, tools connected, deployment review; discovery, acceptance and the decision as common work). Light: in live use, changed only by approval. The replay in weeks 3–5 does not need the assistant, so it proceeds while the deployment review runs.

Weeks 1–2 are discovery and the dictionary; weeks 3–5 replay the checks on Nestlé's history, shown only to the pilot group; weeks 6–10 are the live daily scan in one channel with a weekly review; weeks 11–12 measure against the baselines and decide.

8.2. Milestones and gates

Each milestone ends with Nestlé's confirmation in writing.

Milestone	Week	Condition
Kick-off	1	Sponsor, brand lead, AI Data Governor and data contact are named. Extracts and the access route are agreed.
Gate 1: dictionary	End of 2	The AI Data Governor has approved every definition in use, and the agreed question set is answered exactly on real data.
Gate 2: replay	End of 5	Same-day flag rate and false alerts are reported beside the current checks, and false alerts are inside a budget the team sets.
Go-live	6	The deployment review is passed and the acceptance scenarios in 8.6 pass. The scan reports to the pilot group.
Gate 3: control	Each week, 6–10	No brief circulated without a recorded approval. No figure in a card without its query.
Gate 4: exit	12	The measures in 8.5 against their baselines, the team's own view of whether to continue, and a one-page replication note: what a second team would reuse and what it would write.

The pilot would count as failed if the team stops reading the alerts, if causes are named with confidence and turn out wrong, or if definitions cannot be agreed in the first two weeks.

8.3. Team

Role	Person	Scope
Team lead, solution architect and developer	Phan Minh Trí, AIT-SGN	Contact for Nestlé. The design, the assistant's skills and connections, deployment with Nestlé IT, the weekly report, acceptance and handover.
Data and AI engineer	David Tran Hoang Nguyen, AIT-SGN	Checks on the data extracts, the dictionary and the check set, the replay and its evaluation, driver signatures and thresholds.
Sponsor and approver	Nestlé: the Milo brand lead	Scope, approval of each brief, the weekly review, the exit decision.
AI Data Governor	Nestlé: one named person	Approval of the dictionary, the check set and each proposed change.
Data and IT contact; security and legal reviewer	Nestlé	Read-only extracts, the place to run, the channel and tracker; review of the deployment and of the panel licence.

Coordination. One contact on each side; for AIT-SGN, Phan Minh Trí. A weekly written report on progress, risks and what Nestlé needs to do next. A weekly review in the live weeks, with decisions recorded in writing. One shared list of defects, change requests and decisions. A change to scope is written down, assessed and agreed before work starts. Nestlé sees replay results from week 3, before anything is live.

8.4. What Nestlé provides, by week

An input that arrives late moves the later milestones by the same time. The effect is recorded in writing; scope is not reduced silently.

Week	Nestlé provides
1	The named people in 8.3. The last ten known issues with the date each started and the date it was noticed. The current definitions of the measures. The access route: tenant, channel and tracker.
1–2	Read-only extracts or views: sell-out, sell-in, distribution, distributor stock, the feed log, and the price, promotion and media calendars. Panel deliveries and confirmation of the panel licence. Current hours spent consolidating reports.
3–5	Marking of the replayed alerts as real or not. The weekly alert budget.
6	Sign-off of the deployment review. The members of the pilot group.
6–10	Approval of the brief each working day. A weekly review of about 30 minutes.
11–12	The decision meeting.

8.5. Measures and baselines

What it proves	Measure	Baseline to record in weeks 1–2
Issues are seen the same day	Share of issues, as later confirmed by the team, flagged on their first or second day; median days from start to flag	How long the last ten known issues took to be noticed
Alerts are worth reading	False alerts per week; share of alerts the team rates useful	Alerts from the current week-on-week checks
Explanations can be trusted	Share of named causes the team confirms; number of confident wrong causes	None today
Questions get exact answers	Exact answers on a question set agreed with the team	Time to answer the same set by hand
Time is freed	Hours per week spent consolidating reports	Current hours
It can be replicated	New numbers or checks the team asked for in the live weeks that were added as an approved file change, with the lead time of each; team members who ask in the thread	None today

8.6. Acceptance before go-live

Nestlé runs the scenarios with our support; each one passes or fails. A scenario that cannot be run for lack of an input is recorded as blocked, not passed. Silence is not acceptance. A wrong figure, access beyond a person's permission, or a brief sent without approval blocks go-live; wording and layout defects do not, if Nestlé agrees in writing.

No.	Scenario	Passes when
AT-01	Complete feeds	The morning message arrives by the agreed time, lists every source as arrived, and its figures equal the dictionary's answer for the same day.
AT-02	Short feed	A late or partial file is reported as a data problem and is not analysed as performance.
AT-03	Known past issue	Replayed on its date, it produces one card with actual, expected, a four-way split that sums to the gap, a cause and a confidence.
AT-04	Same move next day	The existing card is updated; no second card is opened.
AT-05	Unexplained move	A move that no signature accounts for is marked unexplained and assigned for review; no cause is named.
AT-06	Metric question	A question from the agreed set is answered exactly, with its definition and query.
AT-07	Question outside the definitions	A request such as daily penetration is refused with the reason.
AT-08	Approval gate	A brief cannot be circulated before the brand lead approves; the approval is recorded with name and time.
AT-09	Read-only	A write to the data through the assistant's connection is refused.
AT-10	Proposed change	A proposed change to a check shows its replay result and takes effect only after the AI Data Governor approves.
AT-11	Trace	For any card, the scan log shows which checks ran and what they returned.

8.7. Main risks and how they are reduced

Risk	How it is reduced
Extracts arrive late or incomplete	Agreed in week 1. The replay starts on what exists; later milestones move in writing.
"Sales" or "distribution" mean different things to different teams	Gate 1: no alert is shown before the dictionary is approved.
Too many alerts, and the team stops reading	The team sets the alert budget. The replay shows the expected weekly count before go-live; alerts are rated every week.
A confident wrong cause	A cause is named only above a written score and margin; otherwise "verify first" or "unexplained". Counted as a pilot measure.
Results on real data are lower than on synthetic data	Gate 2 comes before go-live. The known misses, competitor promotions and pack-mix shifts, are stated in section 07.
Deployment in Nestlé's tenant takes longer than planned; it has not been demonstrated	The review starts in week 1. The replay does not need the assistant, so weeks 3–5 proceed.
A small delivery team	All work is kept in repositories Nestlé can read, and the weekly report records the state, so the work can be picked up from it.

09. Terms

Matching Day question 6: what Nestlé keeps, and on what terms

Every term here is proposed and is agreed in writing before the pilot starts (Appendix B). When an assumption changes, both sides assess the effect on scope and schedule, and Nestlé agrees before work continues.

9.1. Assumptions

- Daily sell-out and distribution exist by region, channel, source and SKU, or can be extracted read-only at that level.
- At least twelve months of daily history are available, including one Tết, to learn the seasonal profile and to replay.
- There is a place to run inside Nestlé's tenant, with access to the model under Nestlé's own agreement.
- The pilot group uses one channel and one tracker, and a named brand lead can approve the brief each working day.

9.2. Responsibilities

Nestlé is responsible for	AIT-SGN is responsible for
The inputs in 8.4, in their week	The dictionary proposal, the check set and the replay
Naming the approver and the AI Data Governor	The assistant's skills and connections, and their installation with Nestlé IT
The deployment review and the access it grants	The weekly report, the shared list and the acceptance scenarios
What the team does about each issue; the content of definitions and of the playbook	Correcting defects against the agreed scope; handover at the end

Anything outside the scope in 2.4 that Nestlé needs is scoped and agreed before work starts.

9.3. Support during the pilot

- **Who and when.** Phan Minh Trí (tri.phan@ait-sgn.online) is the first contact, on working days during working hours in Vietnam. Response times are agreed before go-live.
- **Blocking defects:** the scan did not run; a figure in a card or brief is wrong; a person saw data beyond their permission; a brief was circulated without approval. A wrong figure pauses the scan until it is corrected. Other defects are handled in the weekly review.
- **When a scan fails,** the channel receives a message that says so. A partial result is never presented as complete.
- **Defects are not change requests.** Correcting a defect against the agreed scope does not change the scope or the schedule.

9.4. Ownership and handover

What stays with Nestlé is the capability, not a licence to a screen: the dictionary, the check set and the playbook as versioned files its own people read and change; the issue log, which is the labelled history a second version needs to learn the causes the first could only call unexplained; and the replay, which scores any change before it goes live.

Asset	Owner (proposed)	Note
Nestlé's data and extracts	Nestlé	Stays in Nestlé's infrastructure. AIT-SGN keeps no copy.
Shared dictionary, check set, playbook and team instructions written in the pilot	Nestlé	Versioned files in Nestlé's repositories.
Issue cards, scan logs, ratings, decisions and replay results on Nestlé's data	Nestlé	Handed over in full at the end.
Skills and tool definitions as configured for the pilot; the kit and plugin code that existed before it	AIT-SGN	Nestlé receives them in readable form and may keep using them after the pilot. Licence wording to be agreed.

Handover at the end. The repositories; a short operating guide for each role; the deployment and connection notes; the replay report; the final report on the measures in 8.5; the replication note; the list of open defects and open proposals.

Exit. When the pilot ends or is stopped at a gate, access is removed, and pilot data and logs are returned or deleted on Nestlé's instruction, confirmed in writing.

9.5. Commercial structure

This document states no figure. The structure we propose:

- **What the pilot fee covers:** the work in section 08 by the team in 8.3, from discovery to handover, including the extension of the foundation described in 2.3.
- **A fixed fee, paid by gate.** A part falls due at each gate that Nestlé confirms in writing (8.2). If the pilot stops at a gate, the later parts are not due.
- **What Nestlé carries itself:** model usage and hosting inside its own tenant, under its own agreements. Usage per run is reported weekly, so this cost is visible from the first live week.
- **When the figure is given:** at Matching Day, once the extracts, the channel and the tracker are known, because these decide the work in weeks 1–5.

10. Next steps

Matching Day question 5, continued: verify, do not take our word

This proposal cites no reference customer and no deployment. The evidence offered is the prototype and its evaluation, and Nestlé can check both in one session.

- **Open the demo** of the prototype on synthetic data at nestle-26.ait-sgn.online/milo-2/; each other use case has its own address under the same site.
- **Watch the tools run** on the synthetic data, without a model: the daily snapshot, the unusual moves of a chosen day, the explanation of one move, a metric question, and a question that is refused.
- **Rerun the evaluation.** One command regenerates the data and one scores the tools; the result should match the scorecard quoted in section 07. [3]
- **Try the approval gate.** A commit is refused until a named person approves, and the decision log cannot be edited. [6]
- **Read the rules and the skills.** Thresholds, driver signatures, metric definitions and the playbook are five text files, not code; the four skills are plain instructions. [2] [4]

What cannot be checked yet: the assistant running end to end, any connection to Nestlé's environment, and any result on real data. The first five weeks of the pilot exist to answer these before go-live.

Three steps to start

1. **Briefing** with the MILO-2 owner on the items in Appendix B, and with the owner of each other use case on its case appendix.
2. **Names and scope.** Nestlé names the approver, the AI Data Governor and the data contact. Both sides agree the extracts and a one-page scope with the measures and gates of section 08.
3. **Replay on Nestlé's data.** If Nestlé can provide synthetic or anonymised data, the replay is rerun on it first, before any real data is touched.

Appendix A. Capability overview by layer

State: **M** measured or tested on synthetic data; **B** built, not scored; **W** written as a skill, not run; **D** designed, not built. All 38 are in the pilot scope; the counts are in section 2.3.

Code	Capability	State
1 Where people work — Nestlé's existing tools		
L1-01	Morning message in the team's channel: what is new, updated and unexplained	W
L1-02	One issue card per unusual move, with owner and state; the same move is updated, not filed again	W
L1-03	Questions and answers in the thread	W
L1-04	Summary page opened by link, rebuilt after each scan	W
L1-05	Files passed through the shared folder	W
2 AI assistant		
L2-01	Daily scan skill: run the checks, open or update cards, report to the channel	W
L2-02	Investigate skill: answer a question or a reported move, keeping the queries	W
L2-03	Summary page skill	W
L2-04	File exchange skill	W
L2-05	Set-up for a new company: read the data structure, interview the AI Data Governor, propose the dictionary	D
L2-06	Propose a change to a definition, a check or a team instruction	D
L2-07	Daily brief draft in Vietnamese or English, with mandatory notices for provisional data and unverified causes	B
3 Checks and ways of working		
L3-01	Expected value that knows the weekday, month end, Tết, public holidays and back to school	M
L3-02	Wider tolerance when a planned promotion or an uncertain calendar day covers the scope	M
L3-03	Smallest move worth an alert, and a weekly alert budget	M
L3-04	Placing the alert on the most specific scope that carries the gap; ranking by value at stake	M
L3-05	Tracking a move from its first day and counting the days it runs	M
L3-06	Four-way split into distribution, velocity, price and mix, exactly additive	M
L3-07	Contributions by region, channel, source and SKU	M
L3-08	Six driver signatures: data feed gap, distributor out of stock, distribution loss, unplanned price change, pack-mix shift, competitor promotion	M
L3-09	Confidence: high, verify first, or unexplained and sent for review	M
L3-10	One proposed action per move from the playbook: owner, deadline, priority	B
L3-11	The team's instructions for the assistant, kept in the team repository	D
4 Shared dictionary		
L4-01	Ten governed metrics: sell-out and sell-in value and units, average net price, numeric and weighted distribution, household penetration, value share, volume share	M
L4-02	Comparison with the prior period and year on year	M
L4-03	Refusal of a question outside the definitions, with the reason	M
L4-04	A named approver for every definition, and the definitions repository	D
L4-05	Software that makes every query go through the dictionary	D
5 Warehouse — Nestlé's data platform		
L5-01	Daily snapshot: national, regional and channel sell-out against expected; sell-in; distribution; price against list	M
L5-02	Feed check: which source files arrived and how complete each is	M
L5-03	Panel reading (penetration, share) shown with its period; never used as a daily figure	M
L5-04	Read-only data access	M
Control, across layers 2 to 4		
CT-01	Approval gate: nothing is committed before a named person decides	M
CT-02	Decision log that cannot be edited	M
CT-03	Scan log: which checks ran and what they returned	D
CT-04	Replay of the checks on past data; built for synthetic data, extended to Nestlé's history in weeks 3–5	M
CT-05	Limits enforced at the warehouse and at the assistant's tools	D
CT-06	Two configuration repositories: definitions (company) and team	D

Appendix B. Items to agree

Each term marked "(proposed)" in this document, and each open question, is listed here.

Item	What this proposal assumes	Agreed
Which "sales" the team means: sell-in, distributor sell-out or retailer offtake	Daily distributor sell-out as the main measure, with sell-in beside it	Briefing
Where distribution comes from, and how often	Daily if it exists; otherwise as often as it is delivered	Briefing
The panel that supplies penetration and share; whether its licence allows an AI tool to read it	Shown per delivery; left out if the licence does not allow it	Week 1
What counts as "unusual" today, and who receives the daily view	The team sets the alert budget and the pilot group	Weeks 1–2
Channel and work tracker	Whichever the Milo team already uses: Slack or Microsoft Teams; Jira, Linear or another	Before week 1
If the channel is Microsoft Teams: whether the assistant can post into it	To be confirmed; the cards and the summary page do not depend on it	Deployment review
An existing semantic model	The dictionary sits on it if there is one	Week 1
Place to run and model access; where the summary page is hosted	Nestlé's Azure tenant, with the model through Microsoft Foundry; not demonstrated	Deployment review
Fit with Nestlé's global analytics and assistant tools	Open	Briefing
Retention of cards, answers and scan logs; weekly report of model usage	Section 6.4	Before go-live
Ownership, licence wording and exit	Section 9.4	Before the start
Support hours and response times	Section 9.3	Before go-live
Pilot length and start date	12 weeks from the day access is granted	Matching Day
Commercial terms	The structure in 9.5; no figure in this document	Matching Day

Left open in the design. The software that enforces the shared dictionary (SLayer and Cube are candidates); the sign-in in front of the warehouse; how per-person data permissions are passed through; how data is loaded into the warehouse, which the data team owns. [5]

How a move is detected. The expected value for a day is a running level, the median of recent valid days, multiplied by a seasonal factor learned from a calibration window only. The tolerance is widened around Tết, public holidays, back to school, month end and planned promotions. A move becomes an alert only if it is large enough to matter, and it is placed on the most specific scope that carries most of the gap. [2]

Appendix C. Glossary and evidence register

Term	Meaning in this proposal
Sell-in, sell-out	Sell-in: what Nestlé invoices to distributors, retailers and platforms. Sell-out: what they sell on to outlets or shoppers.
Numeric and weighted distribution	The share of outlets that handle a product; the same share weighted by each outlet's category turnover.
Penetration and share	From a household panel: the share of households buying the brand in a period, and the brand's share of category value or volume.
Expected value	What a day's figure should be, given the recent level, the weekday and the calendar.
Value at stake	The gap between actual and expected sell-out value for the day, in VND.
Velocity	The rate of sale where the product is stocked; one of the four parts of the split, with distribution, price and mix.
Driver signature	A written rule that says which evidence must be present before a cause is named.
Connector	A published, standard link between the assistant and a tool, built on the Model Context Protocol.
Shared dictionary; check set	The single set of approved metric definitions every answer uses; the saved checks the daily scan runs, kept as rule files.
Replay	Running the checks day by day over past data to score them before they go live.
Pack	One prototype use case: synthetic data, tools, rule files and an evaluation.

Evidence register

Ref	Source	Location
[1]	MILO-2 public challenge brief	https://beyondpilot.ai/use-cases/abb14abc-3523-4868-92e2-851b5b4622e4
[2]	MILO-2 prototype pack: tools, rule files, data description	`kit/packs/milo-2/`
[3]	MILO-2 evaluation scorecard and notes	`.state/eval/milo-2/scorecard.json`
[4]	AI-BI plugin scaffold: four skills, connections, open items	`plugins/nestle-ai-bi/`
[5]	AI-BI architecture review, 15 November 2026	`output/ai-bi/architecture-review.html`
[6]	Kit kernel: approval gate, audit log, read-only data access; 12 tests	`kit/src/nestle_kit/`, `kit/tests/`
[7]	Prototype packs of the related cases; the seven case appendices	`kit/packs/<id>/`
[8]	Programme research report	`reports/Nestlé Vietnam AI hackathon challenges.md`
[9]	Atlassian's connector for Jira (Rovo MCP server)	https://github.com/atlassian/atlassian-mcp-server
[10]	Microsoft 365 connector for Claude	https://claude.com/docs/connectors/microsoft/365

CASE APPENDIX · to be read with the AI-BI Issue Desk proposal for MILO-2

FOOD-1 · Insight-to-Action Copilot

How this case plugs into the AI-BI layers

Confidential · **Version:** 1.0 · **Date:** 22 November 2026

Applicant: AIT-SGN · Contact: Phan Minh Trí, tri.phan@ait-sgn.online · Demo: nestle-26.ait-sgn.online/food-1/

> **The case.** Search research, tests, social listening, complaints and business reviews for one sub-category together, group the evidence into recurring tensions and unmet needs, answer market numbers only from the retail-audit tables, and assemble an opportunity brief.
>
> **For:** the insight team · **Approver:** Insight lead · **Rhythm:** On request · **One card is:** one opportunity brief in preparation.

1. The case on the five layers

The architecture is the one in the proposal: Nestlé already has layers 1 and 5, and AI-BI adds layers 2 to 4 and the control across them. This case changes only what is marked as this team's own.

1 Where people work	THIS TEAM'S OWN	the insight channel and tracker project
2 AI assistant	EXTENDED	no scan; adds search over documents
3 Checks and ways of working	THIS TEAM'S OWN	synthesis rules and concept list
4 Shared dictionary	EXTENDED	adds retail-audit measures
5 Warehouse	THIS TEAM'S OWN	audit tables and document sources
Control approvals, logs, replay	REUSED AS IT IS	same gate, same logs

Orange: what Nestlé already has. Green: what AI-BI adds. Reused as it is: nothing is rebuilt for this case. Extended: the shared part gains entries for this case. This team's own: written with the team in its pilot.

2. Layer by layer

Layer	In this case
1 Where people work	The insight team's channel and a tracker project. One card per opportunity brief in preparation.
2 AI assistant	No scheduled scan. The "investigate" skill does the work, with search across the document sources in Vietnamese and English.
3 Checks and ways of working	Group passages into recurring themes and show which source types and periods support each; answer a market number only from the audit tables, with the exact value and its cell; check a brief before it is saved. Prototype rule files: audit dictionary, concepts, synthesis.
4 Shared dictionary	The retail-audit measures of sales and distribution by period, brand, region and segment.
5 Warehouse	Retail-audit tables for sales and distribution, and an inventory of document sources: usage and attitude studies, concept and product tests, social listening exports, consumer care logs and business reviews.
Control	The Insight lead approves. The action that waits: sharing an opportunity brief with the brand team.

3. A request in the team's tools

The team works in its own channel and tracker, as in section 04 of the proposal. Nobody opens a new application. Set-up for this team repeats four of the six steps in section 4.3: its tracker project, its channel, its shared folders and its schedule. The data server and the connection file are in place from the Milo pilot; this team's tables are added to them.

1. **Request.** Someone asks in the channel what consumers struggle with in a sub-category.
2. **Search and group.** The assistant searches all sources together and groups the evidence into themes, each with its supporting sources.
3. **Ask.** "How big is that segment?" The assistant answers from the audit tables only, with the exact value and where it comes from.
4. **Decide.** The Insight lead edits the drafted brief and approves it; only then is it shared.
5. **Rate.** The team marks which themes were useful; the pattern returns as a proposed change to the synthesis rules.

4. What is reused from the Milo pilot, and what this team writes

Reused as it is	Written with this team
The "investigate" skill; the approval gate and logs; the card and thread pattern; the rule that a number comes from a table and never from a document.	The list of sources and who may read each; the concept list; the synthesis rules; the audit dictionary.

What differs from the Milo case. There is no daily scan, and document search sits beside tables. Of the seven cases it is the furthest from the Milo pattern.

5. What exists today

A prototype pack for this case is in the repository, on synthetic data: six tools, three rule files and three data tables, with a named approver and an action that cannot be committed before that person decides. A demo of this prototype on synthetic data is at nestle-26.ait-sgn.online/food-1/. It has not run on Nestlé's data, and the assistant has not run end to end. In the layer architecture the pack's rule files become this team's check set in layer 3, and its synthetic tables are replaced by Nestlé's data in layer 5.

6. What Nestlé would provide, and what is agreed first

Nestlé would provide	Agreed before the start
• The retail-audit extract and its licence terms. • The research reports, test results, social listening exports and consumer care logs for the sub-category. • Which sources an AI tool may read.	• Which sub-category comes first. • How personal data in consumer care logs is removed before loading.

In the proposal: the layers are in section 02, how it works in 03, integration with the team's tools in 04, replication to other teams in 05, roles and governance in 06, terms in 09. This appendix describes integration only.