



Tealium AI Workshop

From Insights to Agency: Deploying Tealium in the Agentic Economy



Tealium Team



Drew Halsch

RVP, Solutions Consulting



Sarah Smith

Staff Solutions Consultant Engineer
Tealium



Frederick Berlanti

Principal Product Manager,
Applied AI

What we are going to cover

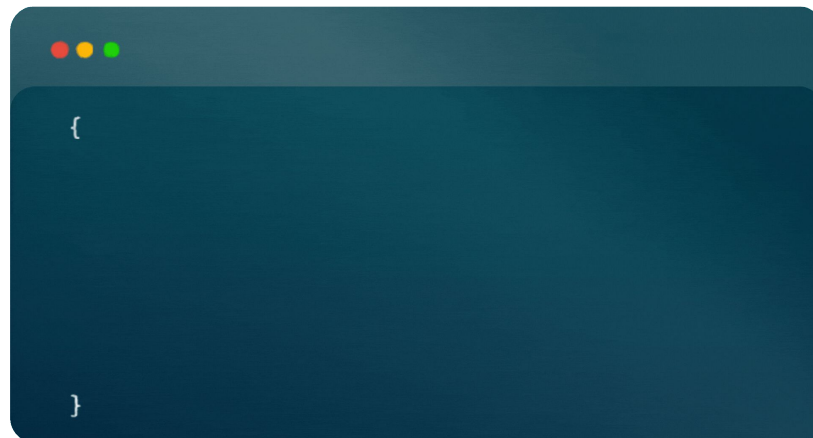
Operationalizing Custom Models and Autonomous Agents

✓ **Tealium's Role in the AI Ecosystem**

✓ **Module 1: Invoke Your Own Model (IYOM)**

✓ **Module 2: Direct Model Activation in AI Platforms**

✓ **Module 3: Agentic Frameworks & MCP**





Tealium's Role in the AI Ecosystem

Opening

The Customer Context Layer for AI Activation

AI Has a Context and Latency Problem

A conversation about what's between AI models and results

The Usual Symptoms

Do these sound familiar?

Your AI agent greets a 10-year VIP customer like a stranger.

Your data scientists spend more time building pipelines than building models.

A "high intent" signal fires, but by the time it reaches a touchpoint, the customer has moved on.

Your chatbot can reason brilliantly; but can't check if a customer has an open support ticket.

Your propensity scores age overnight. By morning, they're yesterday's guesses.

Models are smart. Activation is slow. Data is incomplete.
The gap between intelligence and action is where revenue leaks.

The "Brain in a Jar" Problem

AI can reason. But it can't operate.

What AI CAN do:

- Reason brilliantly
- Generate content
- Analyze patterns
- Have intelligent conversations

What AI CAN'T do alone:

- Know your customers
- Remember their last interaction
- Take action across your systems
- Learn from what happened after it acted

What's missing:



No Context

Your AI doesn't know what's happening right now. It can't see that this customer just abandoned a cart, or came from a 20% off campaign - so it responds to the wrong moment.



No Tools

Your model can decide the best next action, but can't execute it. The recommendation is "call support" instead of actually resolving the issue.



No Feedback Loop

Your AI can't learn from outcomes. It doesn't see what happened next - whether the customer converted, churned, or ignored it - so it doesn't improve.

The Bottleneck Isn't Models.

So, what is really creating the “Brain in a Jar” problem?

1

Context

The model is ready.

But without real-time signals - who this customer is, what they just did, what they're worth - it's guessing.

"AI is a commodity.

Context is differentiation."

— *Scott Brinker*, Martech for 2026

2

Infrastructure

Three breaks in the same pipe:

- Collection → AI
- AI → Action
- Action → Learning

Data pipelines work for training, they fail at inference.



The Customer Context Layer for AI Activation

Tealium is the real-time infrastructure between your customer touchpoints and AI.

We send the right signals before the model decides, activate the decision in the moment and close the loop so your AI gets smarter with every interaction.



Customer Context Layer for AI Activation



Event streaming

Collect and validate real-time behavioral events with schema enforcement



Feature engineering

Enrich signals with session context, visitor history, consent status, at the edge before storage



Contextual Labeling

Built-in tools to add relevance, metadata, and contextual information to improve AI performance



Governance layer

Keep compliant without slowing you down using native governance and consent integrations



Model routing infrastructure

Deliver predictions to and from AI platforms via REST, webhooks, or streaming connectors



ENTERPRISE AI ECOSYSTEM

Traditional ML / AI

Analyzes data and surfaces patterns, insights and scores

e.g. "this customer is 80% likely to churn"



TEALIUM[®]

AI in platform

Intelligence

Productivity

Agentic & GenAI

Acts on those insights autonomously — generating content, making decisions, and executing workflows without human intervention

e.g., "this customer is likely to churn, so I've created and presented a personalized retention offer"

AI Ecosystem: Activating Your AI Stack



Foundation Models

Integrate and activate directly with LLMs in your GenAI platforms

AI Claude

GPT-4o

Gemini

LLaMA



Data Clouds

Integrate and activate directly with models on your ML/AI platforms



snowflake



Google Cloud



databricks



Vector Databases

Get Tealium data directly into your vector DB for your RAG pipelines

Pinecone

Weaviate



Agent Frameworks

Connect Tealium directly to your agent framework



LangChain

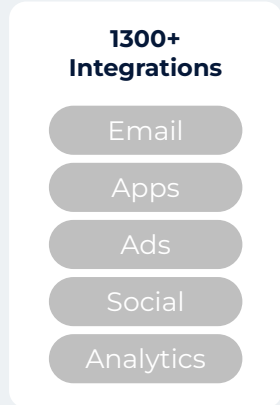
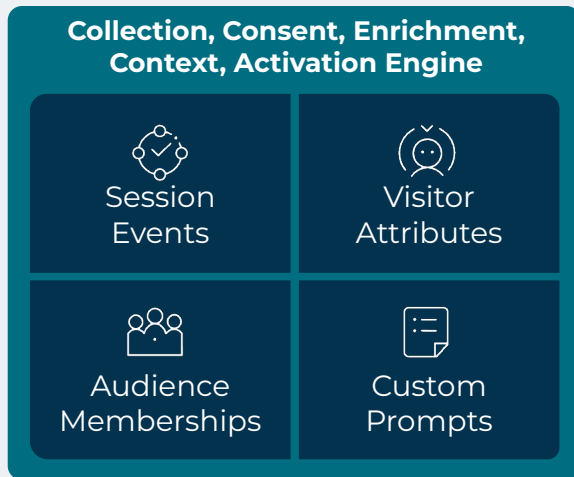
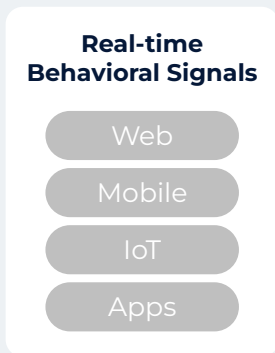
crewai

Enterprise AI Ecosystem



Real-time customer context sent to models

Output received and activated across ecosystem



Traditional ML / AI

Real-Time Inferencing

Model Re-Scoring with new, in-session contextual data

Functions

Real-Time Activation

Call your inference endpoints and get predictions back in real-time

Batch Inferencing

Output Stored in Tables

Data Sources

Cloud Connectors

Updating Models

Sending data to models for updates and fine-tuning



TEALIUM
AI in platform

Intelligence

Productivity

Agentic & GenAI

LLM Activation

Triggered Prompts

Sending real-time context directly to LLMs and activating the outputs

Task Agents

Retrieval Augmented Generation (RAG)

Agents need visitor context, in real-time.

AI Connectors / Functions

MCP / Moments API



Invoke Your Own Model (IYOM)

Module 1

Operationalizing Custom Models at Scale

Invoke Your Own Model (IYOM)

Run your models. Tealium triggers inference and activates the output — in real time.

What is IYOM?

Invoke Your Own Model connects Tealium's real-time behavioral data to **your AI and ML models** — hosted wherever you run them.

Tealium Functions fires on any event, sends rich contextual signals to your model, and routes the response back through your pipeline — all within your data governance rules.

Supported platforms



Your model. Your data. Real-time activation.

01 — Invoke

Any customer event/audience triggers Tealium Functions and your Model Serving API Endpoint

02 — Real-time Inference

Customer data is processed by YOUR own AI models

03 — Activate

Model output collected and routed into action via Tealium



IYOM with Databricks



Tealium + Databricks Mosaic AI Model Serving

Tealium triggers your Databricks model for real-time inferencing. The model returns the content Tealium surfaces the recommendation surfaces in the session.

Your model. Your data. Real-time activation.

Databricks Foundational Models

- Foundational Model APIs
- Provisioned Throughput

External Models

- Third-Party Hosted LLMs
- Served through Databricks

Custom ML Models

- Trained ML Models in MLFlow
- Feature Store, AI Vector Search

IYOM with Snowflake



Tealium + Snowflake: Real-Time Inference

Tealium triggers your Snowflake model within **Snowpark Container Services (SPCS)** for real-time inferencing. The model returns the content Tealium surfaces the recommendation surfaces in the session.

Your model. Your data. Real-time activation.

Snowflake Cortex REST API

- Invoke Foundational Models

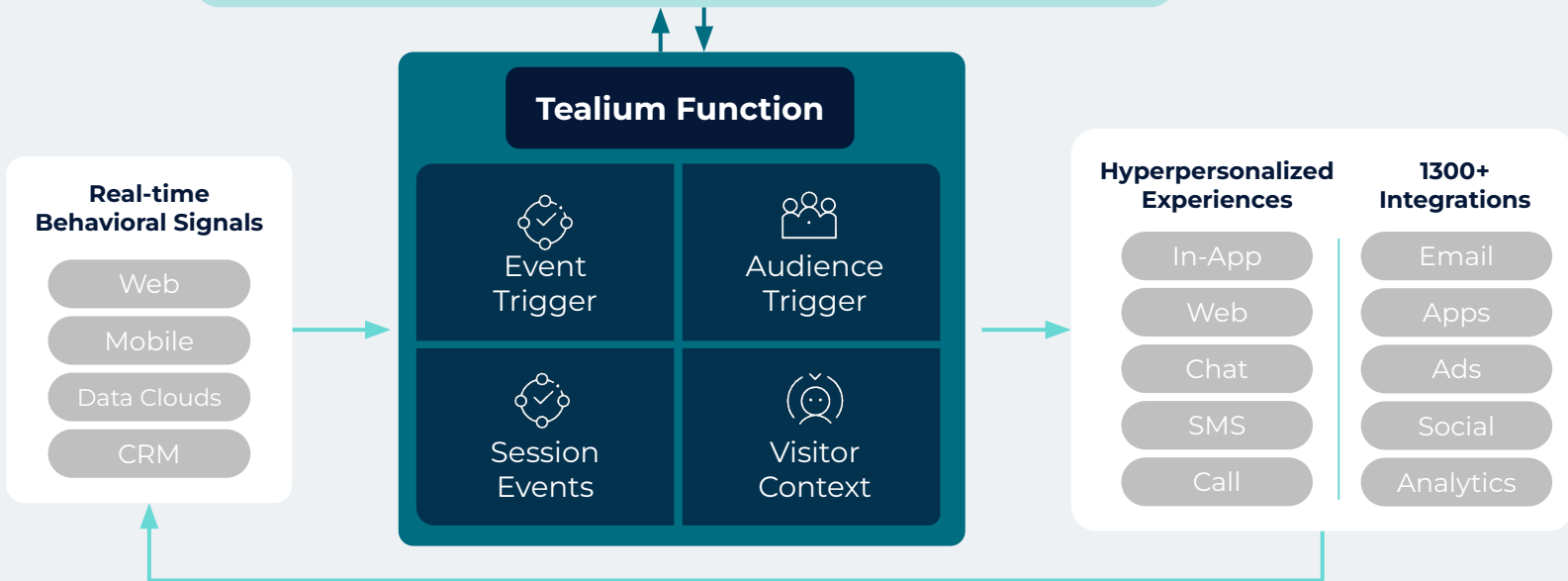
Cortex Fine-Tuned LLMs

- Deployed through Snowflake Model Registry + Inference Service

Custom ML Models

- Trained ML Models
- Feature Serving, Cortex Search

Enterprise AI Ecosystem



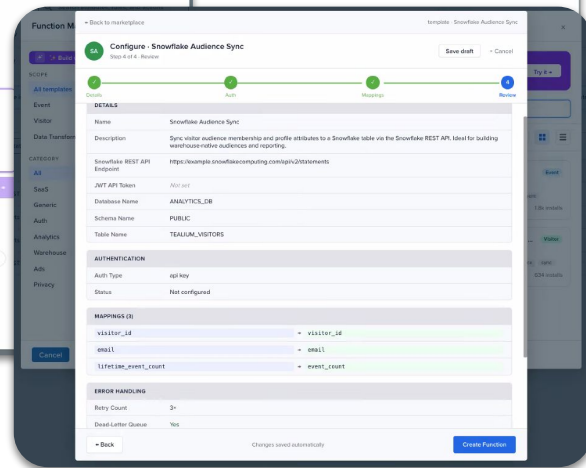
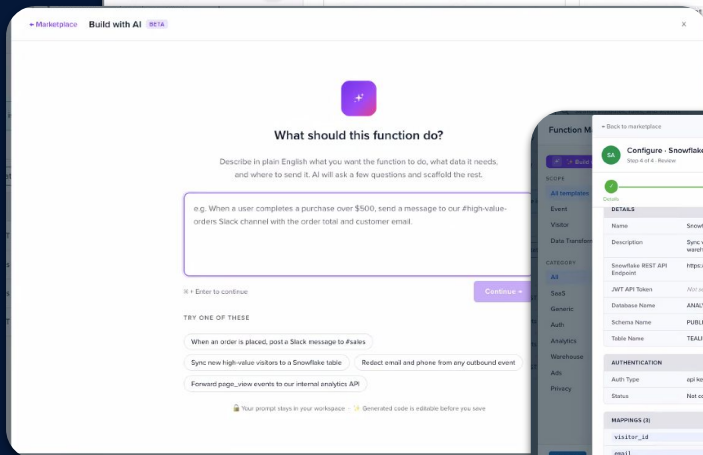
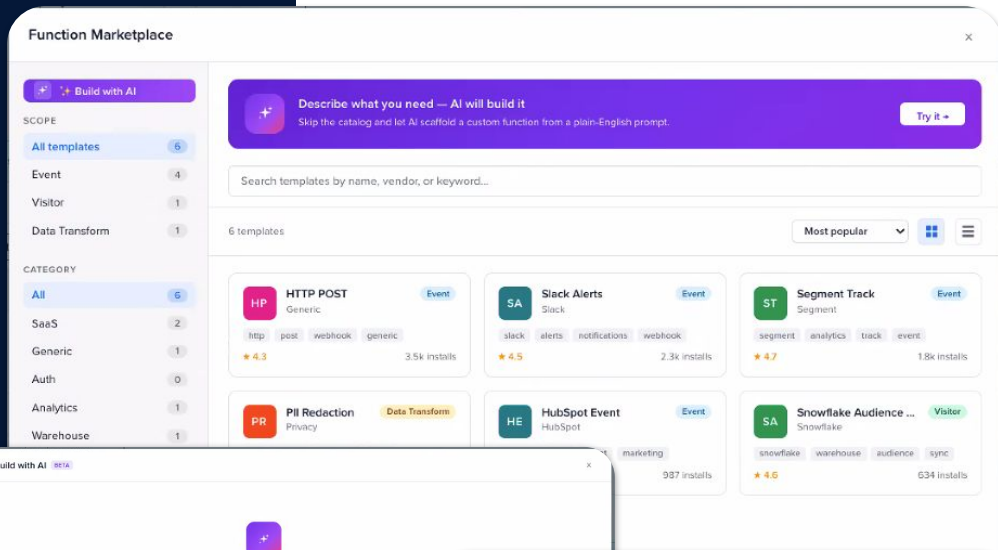
Demo

Invoke Your Own Model

Coming Soon

Functions Marketplace: Faster, Easier Deployments

- Curated catalog of **pre-built functions** for rapid deployment
- Discover, configure, and activate from a **central gallery**
- Simplifies deploying **Invoke Your Own Model (IYOM)** for AI use cases



Questions

Invoke Your Own Model



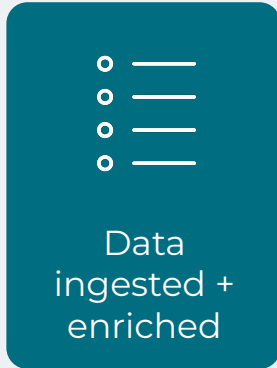
Direct Model Activation in AI Platforms

Module 2

Instant LLM Enrichment via AI Connectors

Direct Model Activation in AI Platforms

Activate models directly in your AI platform, with custom prompts or with real-time context and orchestrate the results



Real-time context for your AI outcomes

Why AI Connectors?

Turn your CDP into a real-time intelligence engine



Real-Time Sentiment

Score customer sentiment during live interactions — not days later in batch reports



Dynamic Classification

Classify purchase intent, churn risk, and product affinity at the event level



AI-Powered Personalization

Generate personalized recommendations, offers, and content using visitor profiles



Automated Decisioning

Route customers to the right experience based on AI-enriched profile attributes



Cost-Effective Scale

Tokenize AI calls with precise prompts — pay only for the intelligence you need



Privacy-First AI

Consent 2.0 enforcement ensures AI only processes consented customer data

Use Cases

Targeted, high-value AI enrichments for customer profiles

Sentiment Scoring

Analyze customer interactions in real-time. Score on a 1-10 scale. Feed sentiment into audience rules for triggered experiences.

```
"Score this interaction  
1-10:  
{event_payload}"
```

```
sentiment_score: 8
```

Intent Classification

Classify visitor intent into categories: browse, compare, buy, churn, support. Drive next-best-action decisioning.

```
"Classify intent from:  
{visitor_profile}"
```

```
intent: "buy"
```

Product Affinity

Identify product category preferences from browsing and purchase history. Power recommendation engines.

```
"Top 3 product interests:  
{visitor_profile}"
```

```
affinity: ["electronics",  
"outdoor", "fitness"]
```

Churn Risk Assessment

Score likelihood of customer churn based on engagement patterns, recency, and support interactions.

```
"Assess churn risk:  
{visitor_profile}"
```

```
churn_risk: "high"  
risk_score: 0.87
```

Demo

AI Connectors

Choosing between AI Connectors Invoke Your Own Model

IYOM - Functions

Use when any of these apply

- ✓ Call any ML/AI model (including fine-tuned or foundational LLMs) for real-time inference.
- ✓ Use case doesn't fit a supported AI provider.
- ✓ Need custom logic/feature transforms/PII filtering before the model.
- ✓ Need code-based configuration in Tealium.

AI Connectors

Use when all of these apply

- ✓ Call foundational LLMs with prompt-based inference.
- ✓ Use case matches a supported AI provider.
- ✓ Use a consistent prompt template with event/visitor data.
- ✓ Want no-code configuration in the Tealium UI.

Questions

Direct Model Activation in AI Platforms



Agentic Frameworks & MCP

Module 3

Powering Autonomous Frameworks with Real-Time Context

Agent Frameworks

The runtime layer where LLMs become autonomous — and where Tealium plugs in

Orchestration

Frameworks like LangGraph & LangChain coordinate multi-step reasoning, tool calls, and state across agent turns.

LANGCHAIN · LANGGRAPH · CREWAI

Tool Calling

Agents invoke external functions to fetch data or take action. MCP is the emerging standard for exposing tools safely.

MCP · FUNCTION · API

Memory & State

Short-term session memory plus long-term stores keep agents coherent across turns, sessions, and channels.

VECTOR · GRAPH · KV STORES

The Context Gap

Frameworks handle reasoning — but they ship blind. No identity, no scores, no history of what prior agents already did.

WHERE TEALIUM ENTERS

Tealium as Context

Moments API + MCP Server deliver identity, ML scores, intent, and prior actions at decision time — as one tool call.

MOMENTS API · MCP SERVER

Tealium as Write-Back

Every agent decision flows back into AudienceStream, recomputing scores and feeding the next agent a fresher state.

AUDIENCESTREAM · EVENTSTREAM

Signal Loop for Agents

Each stage adds context to the signal. The agent consumes the full compounded result.



AI AGENT

The agent sees all of this:

- Identity** Gold tier, \$12K LTV
- ML Scores** churn_risk: 0.9
- AI Intent** return_inquiry, frustrated
- AI Action** recommended: retain
- Prior Actions** retention email sent
- Outcomes** email opened, return done
- Next Action** loyalty_bonus offer

Agent responds:

"As a Gold member, you have a 90-day return window with free pickup. I've started your jacket return — and here's 15% off your next order as a thank you."

Architecture: Tealium + LangChain

Moments API & MCP Server deliver context in; observability & outcomes flow back

DATA SOURCES

Website

Mobile App

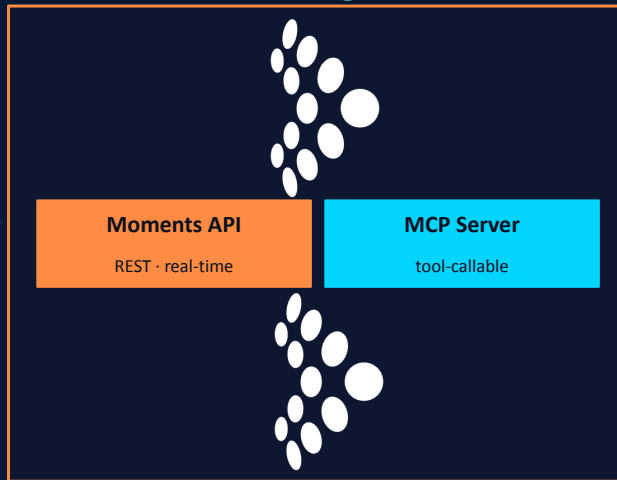
Email / CRM

Call Center

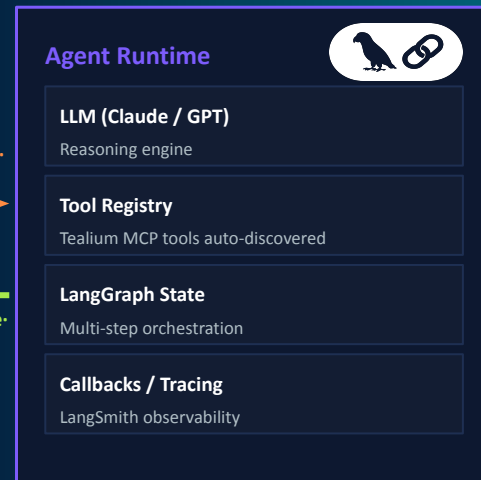
Social / Ads

Collect API

TEALIUM



LANGCHAIN



The Loop: Tealium sends identity + ML + intent + priors → LangChain reasons and acts → outcomes write back to Tealium → scores recompute → next agent starts from the post-action state.

Architecture: Personal Shopper in Fleet

Orchestrator + 3 sub-agents. Claude Sonnet 4.6. MCP auto-discovered.

ORCHESTRATION

Personal Shopper (Orchestrator)

Calls Moments API first. Routes work to sub-agents. Synthesizes the final reply. Claude Sonnet 4.6.

FIRST

profile_reader

Calls Moments MCP
Returns profile block
strong / weak / none

PICK

stylist

Single item + curated
Product DNA knowledge
Returns picks with SKUs

OUTFIT

outfit_composer

Full head-to-toe look
Dress code + occasion
One trade-off named

ORCHESTRATION RULES

- profile_reader runs first, every session — agent never asks the customer for their email.
- If Moments returns weak / none: agent proceeds gracefully on conversation alone. No technical errors surfaced.
- Orchestrator picks stylist vs. outfit_composer based on request shape (single item vs. dress code / occasion).

KEY CONFIG (fleet)

config.json • tools.json

```
// Model selection (Fleet)
llm_model_config: {
  modelId: "anthropic:claude-sonnet-4-6",
  envVarName: "FLEET_ANTHROPIC_API_KEY",
  displayName: "Sonnet 4.6",
}

// MCP server (auto-discovered by Fleet)
tools: [{
  server: "TealiumPersonalization",
  url: "https://...tealiumapis.com/.../personalization/mcp",
  tool: "getPersonalizationContentByVisitorId",
}]

// Sub-agent handoff
agents: [
  profile_reader, // FIRST call, every session
  stylist, // single item / curated edit
  outfit_composer, // dress-code / occasion
]
```

Demo

Agentic Frameworks & MCP

Questions / Open Discussion



The Customer Context Layer for AI Activation

Tealium is the real-time infrastructure between your customer touchpoints and AI.

We send the right signals before the model decides, activate the decision in the moment and close the loop so your AI gets smarter with every interaction.



Tealium AI Offers

Get Connected with Your Tealium Team to Get Started

AI Accelerator Offer

Jump-start a new AI use case

- ✓ Free Tealium AI volume for up to 3 months / 50M events
- ✓ Tealium experts at your side at no extra cost
- ✓ Get new AI use cases live in record time



AI Discovery Workshop

Design your AI roadmap

- ✓ Free expert-led workshop to harness your AI potential
- ✓ Tie priority AI use cases to clear business and CX goals
- ✓ Optimize your data, architecture, and roadmap for AI

