Day 5: Google AI Studio — The AI Control Panel

UBUS 670 | AI for Business Leaders Day 5 · Week 2 · Wednesday, March 18, 2026

Learning Objectives

By the end of today, you will be able to:

1. **Configure** model parameters (temperature, top-p, top-k, max output tokens) and explain their effect on AI output
2. **Write** effective system prompts that define consistent AI behavior for a business use case
3. **Estimate** token costs for a real business deployment scenario
4. **Build** a reusable AI configuration in Google AI Studio for Beacon's email triage

---

Quick Recap: Days 1–4 → Day 5

You've been talking to AI. Now you learn to configure it.

- Day 1: What is AI?
- Day 2: Prompt Engineering (RCTFC framework)
- Day 3: Context Engineering
- Day 4: Multimodal AI
- **Day 5: Configure AI (AI Studio)**

Section 1: Behind the Curtain — What is AI Studio?

Definition

Google AI Studio is a browser-based tool that lets you configure, test, and prototype with Google's Gemini models. It exposes the parameters and settings that consumer chatbots like Gemini Chat hide from you.

Gemini Chat vs. AI Studio

Gemini Chat — The Showroom	AI Studio — The Workshop
Simple conversational interface	Full parameter controls
Pre-set parameters (hidden)	Custom system prompts
No cost visibility	Real-time token counter
No system prompt access	Model comparison tools
Great for quick tasks	Export to code / API

The AI Studio Interface

The AI Studio interface has three key zones: - **System Instructions** — at the top of the central workspace (where you define persistent behavior) - **Chat Area** — the central workspace below system instructions (where you test with messages) - **Run Settings** — right panel (temperature, top-p, top-k, max output tokens, token counter)

The **model selector** sits at the top of the page (e.g., Gemini 3 Flash).

Why Configurers > Users

The AI skill pyramid: - **Builder** (top) — Writes code + deploys - **Configurer** (middle) — Sets system prompts + parameters ← Day 5 goal - **User** (bottom) — Types prompts in chatbot

Day 5 moves you from User to Configurer. You won't need to write code, but you will control the settings that determine how AI behaves in your organization.

Section 2: The Dials — Model Parameters

Temperature: The Creativity Dial

Temperature controls how random or deterministic the model's output is. Low temperature = more predictable; high temperature = more creative and varied.

Range	Zone	Use Case	Example Output
0.0–0.3	Low / Focused	Classification & triage, data extraction, sentiment analysis	"The email is a complaint about shipping delays."
0.4–0.7	Medium / Balanced	Customer responses, summaries, recommendations, professional writing	"This frustrated customer is unhappy with delivery."
0.8–1.0	High / Creative	Brainstorming, marketing copy, product naming, taglines	"A tempest of discontent brews over shipping woes!"

Rule of Thumb: The more you need *consistency*, the lower the temperature. The more you need *variety*, the higher.

Top-p and Top-k: Fine-Tuning Randomness

Temperature is the big knob. Top-p (nucleus sampling) and top-k are fine adjustments. *Nucleus sampling* means the model only considers the smallest set of tokens whose combined probability exceeds p. You rarely need to change these, but understanding what they do makes you a better configurer.

- **Top-p (Nucleus Sampling):** "Consider tokens until probabilities sum to p." A lower top-p restricts choices to higher-probability tokens.
- **Top-k:** "Only consider the top k most likely tokens." A hard limit on the number of candidate tokens.

Max Output Tokens

Max Output Tokens sets a ceiling on how long the model's response can be. One token is roughly 3/4 of a word in English. This parameter controls response length and cost.

- Short (~50 tokens): Email classification label (~37 words)
- Medium (~200 tokens): Email response draft (~150 words)
- Long (~2,000 tokens): Report generation (~1,500 words)

Set this to match your use case — longer is not always better.

Parameter Interaction

Parameters and prompts interact. A perfect prompt with wrong parameters still produces wrong results. **Configuration is as important as the question you ask.**

The Business Takeaway

Parameters aren't just technical settings — they're business decisions.

Wrong example: Beacon sets temperature to 1.2 for complaint classification → inconsistent routing, broken SLAs, angry customers. Right choice: Temperature 0.1 for classification → same input produces same classification every time → reliable routing, predictable SLAs, trust.

Section 3: The AI Job Description — System Prompts

What is a System Prompt?

A **system prompt** is a set of persistent instructions that shapes every response the AI generates. Unlike a regular prompt that is sent once, the system prompt is active for the entire conversation.

Think of it as the **Day 2 RCTFC framework made persistent**: - **R**ole → System prompt Role section - **C**ontext → System prompt Rules & Boundaries - **T**ask → Defined once in system instructions, applied to every email - **F**ormat → System prompt Format section (JSON, categories, etc.) - **C**onstraints → System prompt Escalation, Tone & Boundaries

Anatomy of a Great System Prompt

Six components radiate from the center:

1. **Role** — Who the AI is. "You are Beacon's email triage specialist."
2. **Rules** — Behavioral constraints. "Always classify into exactly one of 7 categories."
3. **Format** — Output structure. "Return JSON with category, priority, and summary."
4. **Escalation** — When to involve humans. "Flag legal threats and safety issues for manager."
5. **Tone** — Voice and style. "Professional, empathetic, solution-oriented."
6. **Boundaries** — What AI must NOT do. "Never promise refunds or share internal pricing."

System Prompt: Beacon Email Triage

```
# Beacon Retail Group - Email Triage System Prompt

## Role
You are Beacon Retail Group's email triage specialist.
You classify incoming customer emails and draft responses.

## Rules
- Classify every email into exactly ONE category:
  Complaint, Product Question, Order Status,
  Return/Exchange, Billing, Partnership Inquiry,
  Feedback/Praise
- Assign a priority: Low, Medium, High, Critical

## Format
Always respond with:
  Category: [category]
  Priority: [priority]
  Summary: [1-sentence summary]
  Suggested Response: [draft reply]

## Escalation
- Legal threats → flag for Legal team
- Safety concerns → flag for Safety Manager
```

```
- Requests over $500 → flag for Supervisor

## Tone
Professional, empathetic, solution-oriented.
Use the customer's name when available.

## Boundaries
- Never promise specific refund amounts
- Never share internal pricing or margins
- Never make up product information
```

Good vs. Bad System Prompts

Bad (Vague): "You are a helpful assistant. Answer customer emails politely." → Inconsistent format, no classification, no escalation rules.

Good (Detailed Beacon Prompt): Role + Rules + Format + Escalation + Tone + Boundaries → Structured output every time, clear categories, automatic escalation, safe boundaries.

A vague system prompt is like a job description that says "do stuff." The more specific you are, the more reliable the AI becomes.

System Prompts as Governance

A well-crafted system prompt is your **first line of AI governance**: - **Scope Control** — What topics the AI can discuss and what it must decline - **Escalation Paths** — When to route to a human, which team gets flagged - **Data Protection** — How to handle sensitive info: PII, financials, legal matters

System prompts are governance you can write today. In Day 6, we'll test how well they hold up under adversarial conditions (red-teaming).

Iterating System Prompts

System prompts are living documents. Version them, test them, and improve them.

Iteration Process: Write → Test typical inputs → Test edge cases → Refine → Version

Edge Cases to Test: Email in Spanish? Profanity? Products Beacon doesn't sell? Competitor pricing requests? Blank emails? Multiple issues in one email?

Common Mistake: Writing a system prompt once and never testing it. Every prompt has blind spots.

Section 4: What Does It Cost? — Token Economics

Tokens (Refresher from Day 1)

Tokens are the units AI models use to process text. ~1.3 tokens per English word. Your system prompt is sent **with EVERY message**.

Token Breakdown per email: System Prompt (~400 tokens) + Email Input (~300 tokens) + Output (~300 tokens) = ~1,000 tokens

Token Pricing: Model Tiers

Illustrative, based on early 2026 rates. Check current pricing before deployment.

Model	Input (per 1M tokens)	Output (per 1M tokens)
Gemini 3 Flash — The Toyota	~$0.50	~$3.00
Gemini 3 Pro — The Mercedes	~$2.00	~$12.00

Pro costs ~4x more per token than Flash. Choose based on task difficulty.

The Beacon Email Math

Volume: 200 emails/day × 30 days = 6,000/month - Input per email: ~700 tokens (400 system prompt + 300 email) - Output per email: ~300 tokens - Monthly input: 4.2M tokens - Monthly output: 1.8M tokens

Gemini 3 Flash: Input: $2.10 + Output: $5.40 = ~**$7.50/month Gemini 3 Pro:** Input: $8.40 + Output: $21.60 = ~**$30.00/month**

Cost vs. Value

Monthly Cost Comparison: - **Human Triage:** $600/month (1 hr/day @ $20/hr) - **AI (Flash):** ~$7.50/month - **AI (Pro):** ~$30/month

With Flash at ~$7.50/month vs. $600/month manual cost, that's roughly **80x savings**.

Real deployments have additional costs: development time, testing, monitoring, and human oversight. AI triage augments human work — it doesn't eliminate it.

Building Reusable Configurations

Combine System Prompt + Parameters + Model Choice → **Saved Template** → Deploy to Email Triage (production), Team Training (internal demo), API Integration (exported to code).

Key Takeaways

1. **Parameters Control Behavior** — Temperature is the key dial. Low for consistency, high for creativity. Top-p and top-k are fine adjustments.
2. **System Prompts = Persistent RCTFC** — A system prompt is your Day 2 RCTFC framework set once. Six components: Role, Rules, Format, Escalation, Tone, Boundaries.
3. **Pennies vs. Dollars** — Beacon's email triage costs ~$7.50/month with Flash vs. $600/month manual — roughly 80x savings. At enterprise scale, model choice matters.
4. **AI Studio: User → Configurer** — System prompts + parameters + model choice = your AI configuration toolkit.

What's Next: Day 6

You've *built* AI systems in Days 4–5. Day 6: we *break* them.

- **Red-Teaming** — Deliberately trying to make AI fail
- **Testing** — Systematic evaluation under edge cases
- **Guardrails** — Building defenses that keep AI in bounds

Save the system prompt you build in today's lab. We'll stress-test it on Day 6.