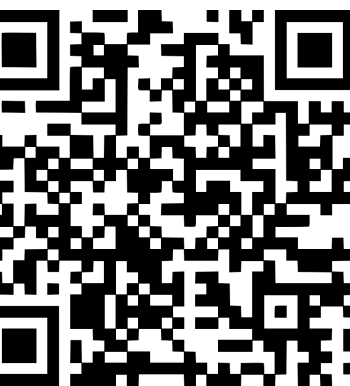# Classifying Emails

For Fun and Non-Profit
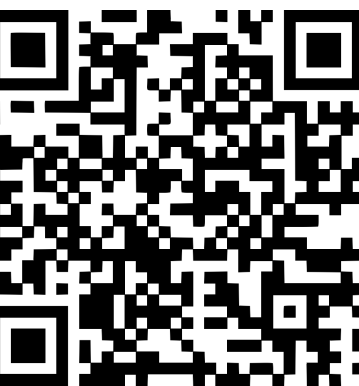
Knut & Marty

Knut Graf

Marty Spellerberg

Opportunity:
Inbox Management is a Constant Chore
with an Impact on Visitor Relations.

- A good email response can happen if it is written by the right person (language, specialization, expertise, availability.)

- Manually sorting incoming email and getting messages to the right people is a burden.

- What if we could automate that?

Solution:
Classify Emails Automatically,
Route to the Best Person to Respond

Classification Criteria for Emails:

- Segments of Senders

- Categories & Subcategories of Topics

- Urgency Levels

(or: something else)

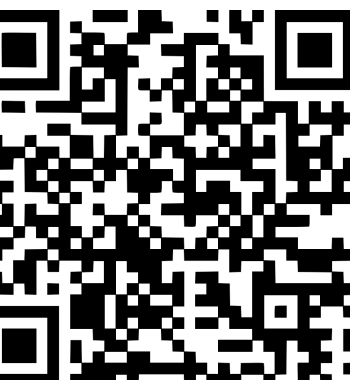*Based on the results, route emails to the right people to respond.*

Impact:
Better Visitor Relations

Simplified inbox management becomes an opportunity for the best people to create the best responses, building differentiated customer relationships.

Better response times, consistent high quality responses: happy visitors!
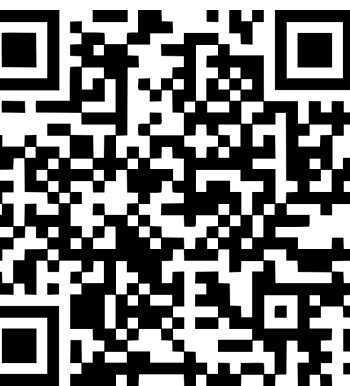
What This Is Not

- A chatbot creating agentic slop responses (those hurt relationships)

- Replacing people in customer relationships (people are valued)

- Selling out emails to someone's big tech cloud (ick factor)

- Another CRM (solved problem, integration opportunity)

- Another email system (solved problem, integration opportunity)

Under the Hood

Under The Hood

- Definitions: which segments, categories, subcategories?

- A model: DistilBERT

- Training (with training data)

- Monitoring (to make sure it's working well)

- Corrections (easy, low friction)

- Administration (setting up routing etc)

- Integration (email connector or workflow platform)

- Robust scalable architecture (vs. a throw-away architecture)

Under The Hood

- **Definitions: which segments, categories, subcategories?**

- **A model: DistilBERT** (small enough to run locally)

- Training (with training data)

- Monitoring (to make sure it's working well)

- Corrections (easy, low friction)

- Administration (setting up routing etc)

- Integration (email connector or workflow platform)

- Robust maintainable architecture (vs. a throw-away prototype)

"Category Dictionary" - A Starting Point

Segments

- Domestic Tourists
- International Tourists
- Local New Yorkers
- Researchers and Professionals
- Special Event Attendees
- Students and Educators

Urgency

- standard
- medium
- high

Categories & Subcategories

- **Access**: Entry, admission, and special access requests
  Same-day access, future access, special accommodations

- **Booking**: Reservations and scheduling
  Reservations, modifications, cancellations

- **Information**: Museum information and policy questions
  General info, specific program info, policy info

- **Emergency**: Urgent situations requiring immediate attention
  Safety, security, medical, facility issues

- **Administrative**: Documentation and institutional processes
  Documentation, permissions, registrations

- **Lost and found**: Lost or found item inquiries
  Lost items, found items, item retrieval

1-Minute-LLM-101

2 "flavors" of LLMs:

GPT: "*decoder*-only"

generates text by predicting the next word in a sequence based on the words that came before it

Best at text generation
(creative writing , chatbots)

BERT: "*encoder*-only"

analyzes text by considering the context of a word from both its left and right sides simultaneously

Best at text classification
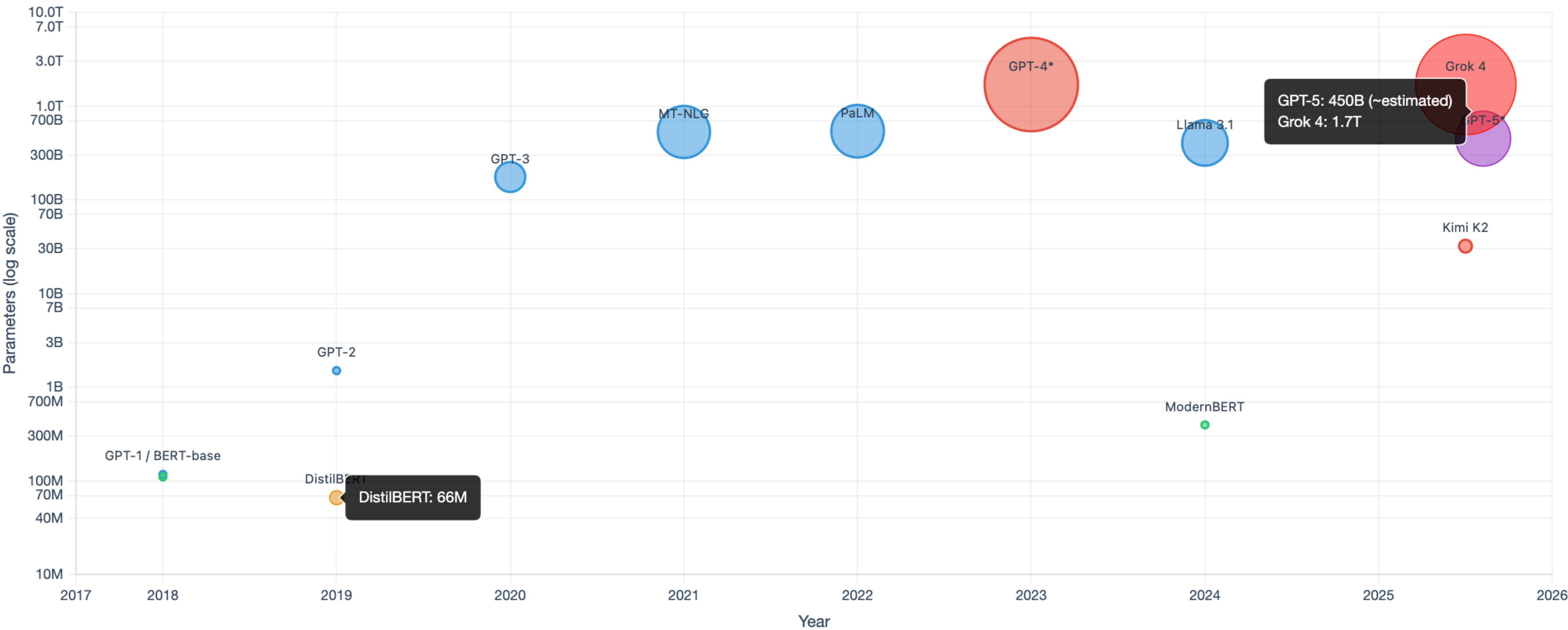(question answering, relationships & entities)

Un-quantized vs. Quantized Models:

• "All the Parameters" vs "The important ones"

• "Datacenter required" vs "runs locally, almost as good"

• "Rent someone's LLM" vs. "Run your own"

Parameters (weights, biases) shape connections between **neurons**, trained with **tokens** represented as **embedding vectors** which capture semantic meaning & relationships.

Evolution of Language Model Sizes

Parameters (log scale)

10.0T
7.0T
3.0T — GPT-4*   Grok 4
1.0T — GPT-5: 450B (~estimated)
700B — MT-NLG   PaLM   Llama 3.1   Grok 4: 1.7T   GPT-5*
300B — GPT-3
100B
70B
30B — Kimi K2
10B
7B
3B — GPT-2
1B
700M — ModernBERT
300M
100M — GPT-1 / BERT-base
70M — DistilBERT   DistilBERT: 66M
40M
10M

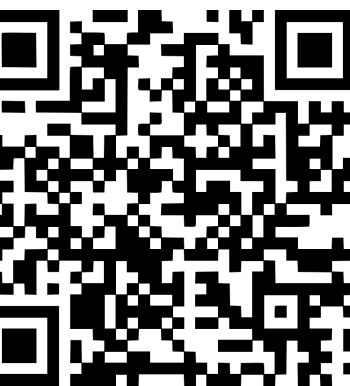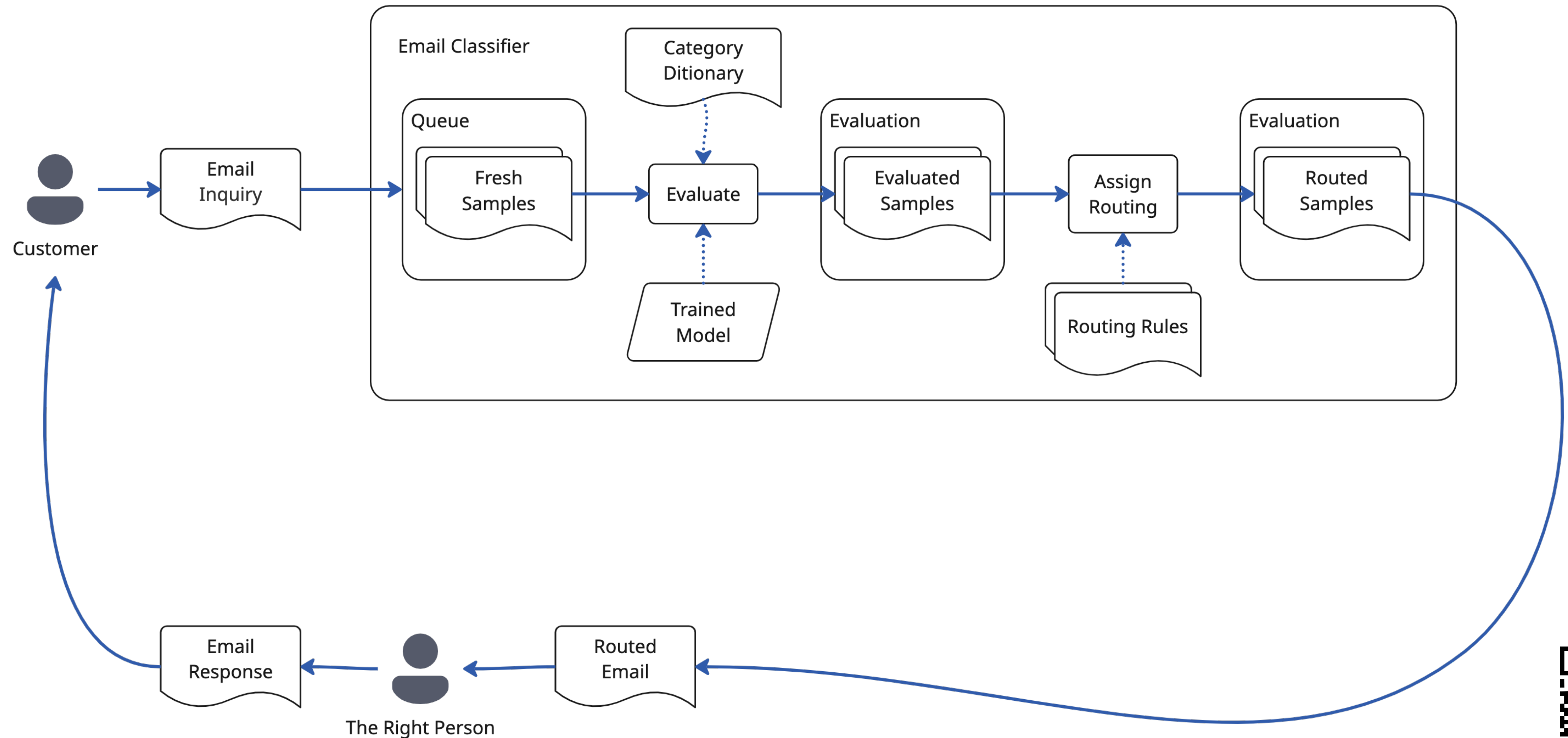2017   2018   2019   2020   2021   2022   2023   2024   2025   2026

Year

Dense Models   MoE/Sparse Models   Encoder Models   Distilled Models   Unified System

Estimated values marked with (~). MoE models show active parameters where applicable. Circle size represents parameter count when "Size by Parameters" is enabled.
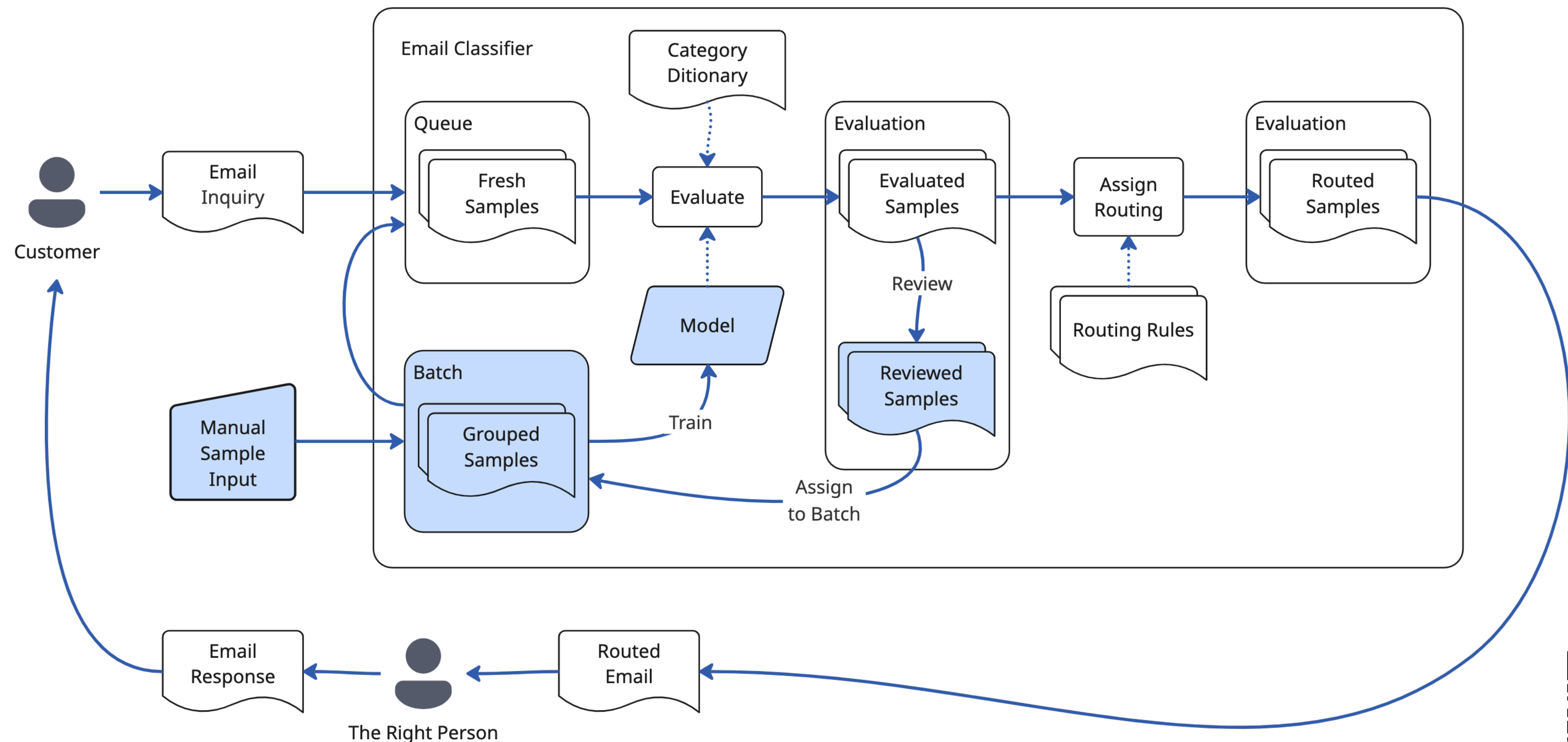
Basic Flow:
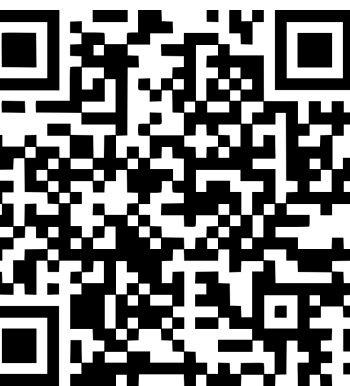# Customer Email, Routed to Right Person

Manual Review, Sample Input
for Quality Control & Model Training

The Front End

Email Classifier Utility

Sample Queue
Museum Email Classifier - Crawl Version

API Connected

Queue

All | Queue (7) | Batch ∨ | Evaluation ∨ | Batch | Refresh

1 selected ✕ | ⋯ Sample Actions | + Add Sample | + Create Batch

Keep Samples ⬤ | Model: Model Using New Scheduled Hierarchical Loss ∨ | Evaluate ▶

	Text	Evaluated	Label Sets	Created	Actions
☐	The New Jersey Pine Barrens Preservation Trust (22 members) would like to arrange a specialized tour foc…	Yes: 7	⊘	2025-06-26 10:00	🗍 🗑
☐	Good morning, I'm interested in booking a group visit for my book club (around 8 people) to see the 'Ma…	Yes: 2	⊘	2025-06-26 10:20	🗍 🗑
☐	Hello, I need to reserve a timed entry slot for myself and my daughter for this Saturday, May 18th, prefera…	Yes: 2	⊘	2025-06-26 10:20	🗍 🗑
☐	Singapore National Heritage Board requests formal documentation of your Southeast Asian textile collecti…	Yes: 7	⊘	2025-06-26 10:10	🗍 🗑
☐	URGENT - Young visitor having severe seizure in Egyptian gallery. We're from Wisconsin. Child approximat…	Yes: 7	⊘	2025-06-26 10:05	🗍 🗑
☐	URGENT - Young visitor having severe seizure in Egyptian gallery. We're from Wisconsin. Child approximat…	Yes: 3	⊘	2025-08-17 03:49	🗍 🗑
☐	Lost important medication in Modern Art gallery - small green case with Norwegian pharmacy labels. Con…	Yes: 7	⊘	2025-06-26 10:15	🗍 🗑

‹ | 1 | ›

Inspector Panel

▼ Queue Context

Queue View: samples in the evaluation queue

Samples in the queue are ready for batch evaluation

Evaluating... Museum Email Classifier - Crawl Version

API Connected

Queue

All | Queue (7) | Batch ⌄ | Evaluation ⌄ | Batch | Refresh

1 selected ✕ | ⋯ Sample Actions | + Add Sample | + Create Batch | Keep Samples | Model: Model Using New Scheduled Hierarchical Loss ⌄ | Evaluate ▶

	Text	Evaluated	Label Sets	Created	Actions
☐	The New Jersey Pine Barrens Preservation Trust (22 members) would like to arrange a specialized tour foc...	Yes: 7	⊘	2025-06-26 10:00	📋 🗑
☐	Good morning, I'm interested in booking a group visit for my book club (around 8 people) to see the 'Ma...	Yes: 2	⊘	2025-06-26 10:20	📋 🗑
☐	Hello, I need to reserve a timed entry slot for myself and my daughter for this Saturday, May 18th, prefera...	Yes: 2	⊘	2025-06-26 10:20	📋 🗑
☐	Singapore National Heritage Board requests formal documentation of your Southeast Asian textile collecti...	Yes: 7	⊘	2025-06-26 10:10	📋 🗑
☐	URGENT - Young visitor having severe seizure in Egyptian gallery. We're from Wisconsin. Child approximat...	Yes: 7	⊘	2025-06-26 10:05	📋 🗑
☐	URGENT - Young visitor having severe seizure in Egyptian gallery. We're from Wisconsin. Child approximat...	Yes: 3	⊘	2025-08-17 03:49	📋 🗑
☐	Lost important medication in Modern Art gallery - small green case with Norwegian pharmacy labels. Con...	Yes: 7	⊘	2025-06-26 10:15	📋 🗑

‹ 1 ›

Inspector Panel

● Queue Context

Queue View: samples in the evaluation queue

Samples in the queue are ready for batch evaluation

● Evaluation Progress

Model: d9ec0335-53eb-47dd-807d-08d800fa6693

Run ID: a43004ec-9c58-4501-9614-52916513cff5

Started: 10/4/2025, 1:31:15 AM

Progress: 7 / 7 samples

100%

⣿ Processing samples...

✕ Cancel | Created evaluation batch for results.

Evaluating... Museum Email Classifier - Crawl Version

Queue ⚙

All | Queue (7) | Batch ⌄ | Evaluation ⌄ | Batch | 🗁 | ⟳ Refresh

1 selected ✕ | ⋯ Sample Actions | + Add Sample | + Create Batch

Keep Samples ⬤ | **Model: Model Using New Scheduled Hierarchical Loss** ⌄ | Evaluate ▶

	Text	Evaluated	Label Sets	Created	Actions
☐	The New Jersey Pine Barrens Preservation Trust (22 members) would like to arrange a specialized tour foc...	Yes: 7	⊘	2025-06-26 10:00	⧉ 🗑
☐	Good morning, I'm interested in booking a group visit for my book club (around 8 people) to see the 'Ma...	Yes: 2	⊘	2025-06-26 10:20	⧉ 🗑
☐	Hello, I need to reserve a timed entry slot for myself and my daughter for this Saturday, May 18th, prefera...	Yes: 2	⊘	2025-06-26 10:20	⧉ 🗑
☐	Singapore National Heritage Board requests formal documentation of your Southeast Asian textile collecti...	Yes: 7	⊘	2025-06-26 10:10	⧉ 🗑
☐	URGENT - Young visitor having severe seizure in Egyptian gallery. We're from Wisconsin. Child approximat...	Yes: 7	⊘	2025-06-26 10:05	⧉ 🗑
☐	URGENT - Young visitor having severe seizure in Egyptian gallery. We're from Wisconsin. Child approximat...	Yes: 3	⊘	2025-08-17 03:49	⧉ 🗑
☐	Lost important medication in Modern Art gallery - small green case with Norwegian pharmacy labels. Con...	Yes: 7	⊘	2025-06-26 10:15	⧉ 🗑

< 1 >

‹ Inspector Panel › ✕

● Queue Context

Queue View: samples in the evaluation queue

Samples in the queue are ready for batch evaluation

● Evaluation Progress

Model: d9ec0335-53eb-47dd-807d-08d800fa6693

Run ID: 86555510-c7cd-4107-9e12-3bd0638a7d89

Started: 10/4/2025, 1:21:11 AM

Completed: 10/4/2025, 1:21:31 AM

Created evaluation batch for results. | 👁 View Results

Evaluation Batch Museum Email Classifier - Crawl Version

API Connected

Evaluation: New Model Test ⚙ Choose Evaluation ⌄

| All | Queue (7) | Batch ⌄ | Evaluation ⌄ | Batch | ↻ Refresh |

··· Sample Actions   + Create Batch

	Text	Evaluated	Validated	Routing	Created
☐	URGENT - Young visitor having severe seizure in Egyptian gallery. We're fro...	⊘	✓ Validated	Front Office Queue ⊘	2025-08-17 03:49
☐	Good morning, I'm interested in booking a group visit for my book club (ar...	⊘	✓ Validated	Communications Queue ⏱	2025-06-26 10:20
☐	Hello, I need to reserve a timed entry slot for myself and my daughter for t...	⊘	✓ Validated	Communications Queue ⏱	2025-06-26 10:20
☐	Lost important medication in Modern Art gallery - small green case with N...	⊘	✓ Validated	Communications Queue ⏱	2025-06-26 10:15
☐	Singapore National Heritage Board requests formal documentation of your...	⊘	ⓘ Validated	No routing —	2025-06-26 10:10
☐	URGENT - Young visitor having severe seizure in Egyptian gallery. We're fro...	⊘	✓ Validated	Octavia Jones (Outreach) ⊘	2025-06-26 10:05
☐	The New Jersey Pine Barrens Preservation Trust (22 members) would like to...	⊘	ⓘ Validated	No routing —	2025-06-26 10:00

‹ 1 ›

Inspector Panel

▼ Evaluation Context

Batch name	Evaluation Run - 2025-10-03 20:31
Batch notes	Model: Model Using New Scheduled Hierarchical Loss, Samples: 7
Evaluation Timestamp	2025-10-04 01:31:35
Model name	Model Using New Scheduled Hierarchical Loss
Category dictionary	Category Dictionary with Segments and Urgency
Samples evaluated	7

▼ Predictions   ↻ Refresh

Segment Predictions

Domestic Tourists	(3)	✓ 3
International Tourists	(2)	✓ 1 ✗ 1 (Lowest Accuracy Segment)
Local New Yorkers	(2)	✓ 2
Researchers and Professionals	(0)	
Special Event Attendees	(0)	
Students and Educators	(0)	

Category / Subcategory Predictions

access /	(0)	
same-day access	(0)	
future access	(0)	
special accommodations	(0)	
booking /	(3)	✓ 2 ✗ 1
reservations	(3)	✓ 2 ✗ 1
modifications	(0)	
cancellations	(0)	
information /	(0)	
general info	(0)	
specific program info	(0)	

Routing Page
Museum Email Classifier - Crawl Version

Routing ⚙

| | All | Active | Rule Sets | ↻ Refresh |

| ··· Rule Actions | + Add Rule | + Add Ruleset |

	Priority	Rule Name	Destination	Execution Timing	Status	Today	Actions
☐	1	Emergencies during opening hours	👤 Mary Gonzales (Visitor Services)	Immediately	✓	2	···
☐	2	Urgent Booking, Cancellations	👤 Mary Gonzales (Visitor Services)	Immediately	✓	15	···
☐	3	Other Ticketing	👤 Visitor Services Queue	Every 15 Minutes	✓	236	···
☐	4	Researchers / Urgent	👤 Zara Dubois (Curatorial)	Immediately	✓	4	···
☐	⌃ 5 ⌄	International Tourists / Program Info	👤 Zara Dubois (Curatorial)	Every Hour	✓	46	···
☐	6	Special Event Attendees / Booking	👤 Freya Schmidt (Development)	Hold for Review	✓	28	···
☐	7	Facility Issues	👤 Alistair Khan (Technical)	Immediately	✓	7	···
☐	8	Researchers / Administrative	👤 Freya Schmidt (Development)	Twice Daily	⊗	41	···
☐	9	Students & Educators	👤 Octavia Jones (Outreach)	Every 15 Minutes	✓	89	···
☐	10	VIP Inquiries / Special Access	👤 Freya Schmidt (Development)	Immediately	✓	3	···
☐	11	Media & Press Requests	👤 Communications Queue	Every 30 Minutes	✓	12	···
☐	12	Lost & Found Items	👤 Visitor Services Queue	Every Hour	✓	18	···

‹ Routing Rule Inspector ›   ✕

▾ Details   ✎ Edit

Name International Tourists / Program Info

Status Active Working   Created   2025-06-26 10:10

Use 3 Rule Sets   Modified 2025-09-28 16:56

Traffic Medium   Unique ID 142-45-g2

▾ Rule Term   ✎ Edit

Segment		Category
International Tourists		Information
And	Or	And
Categories		Subcategory
Booking Or Information		Specific Program Info
Any Urgency		Any Segment, Any Urgency

▾ Destination   ✎ Edit

	PRIMARY		SECONDARY
Name	Zara Dubois Online	Name	Renata Volkov Online
Role	Outreach Coordinator	Role	Assistant Coordinator
Department	Curatorial	Department	Curatorial

📅 Coordinate Schedules...

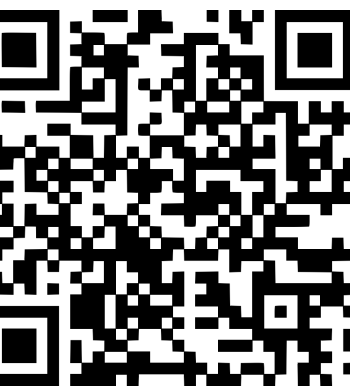### ▾ Execution Timing   ✎ Edit

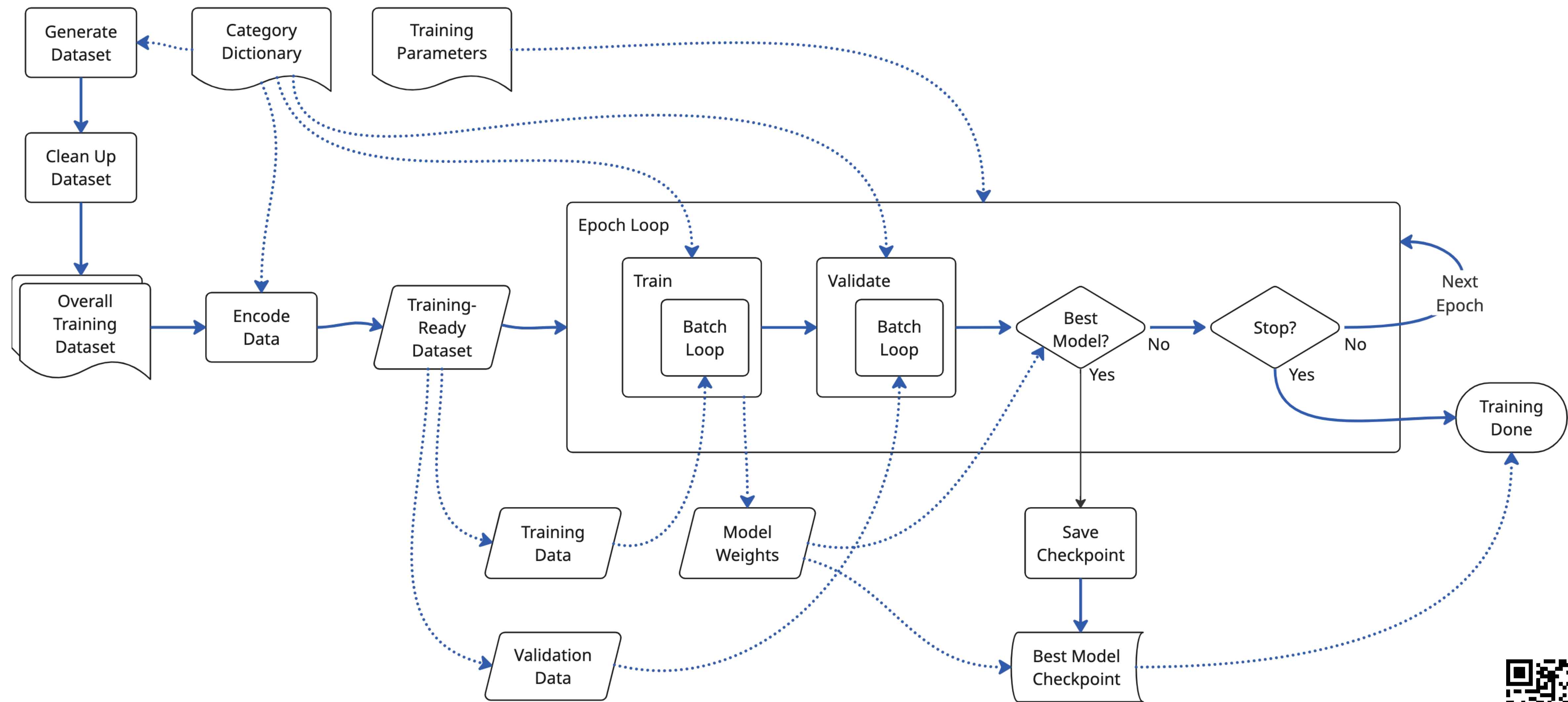Timing Type Interval - Hourly

When On the hour
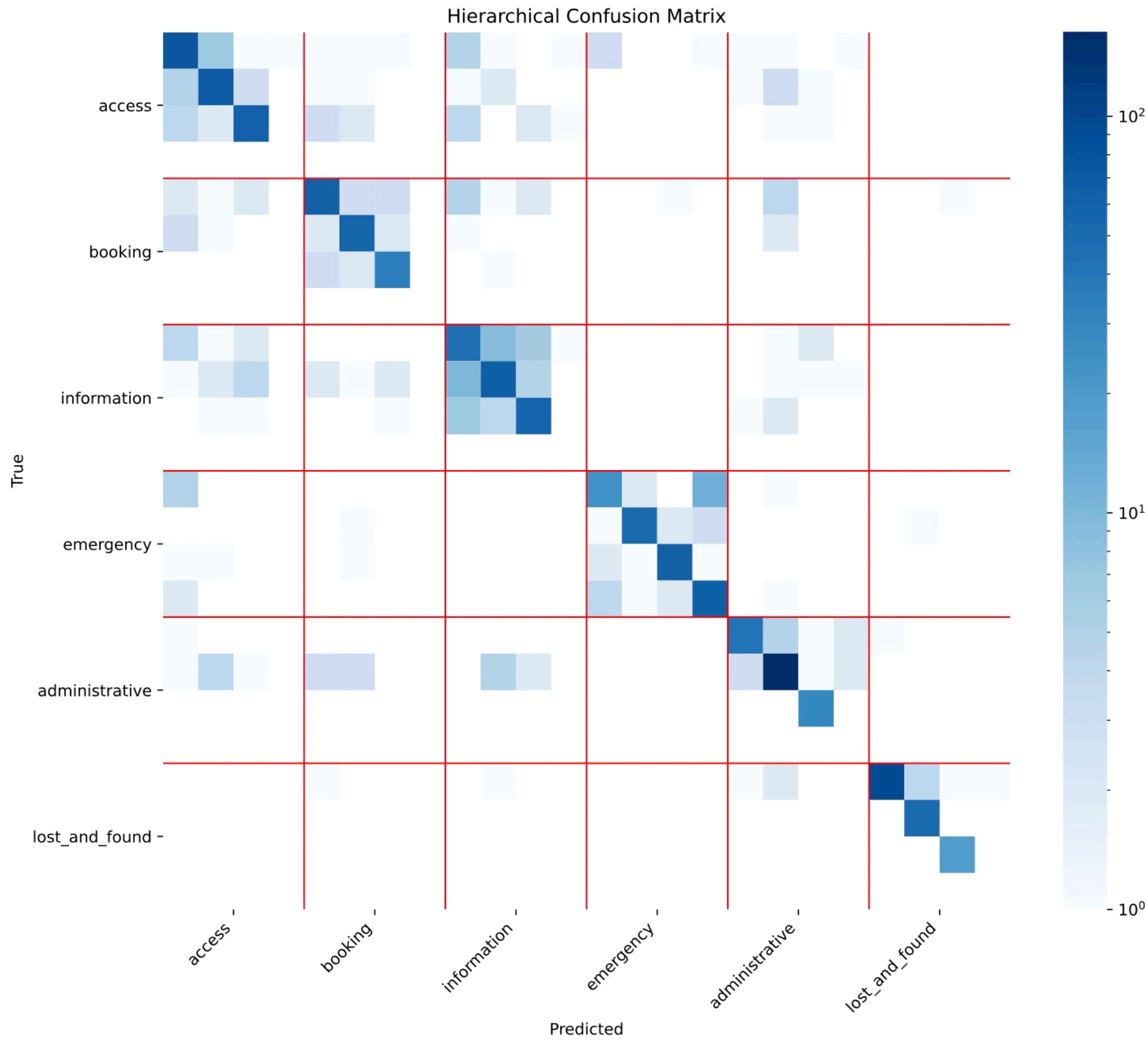
Training Pipeline

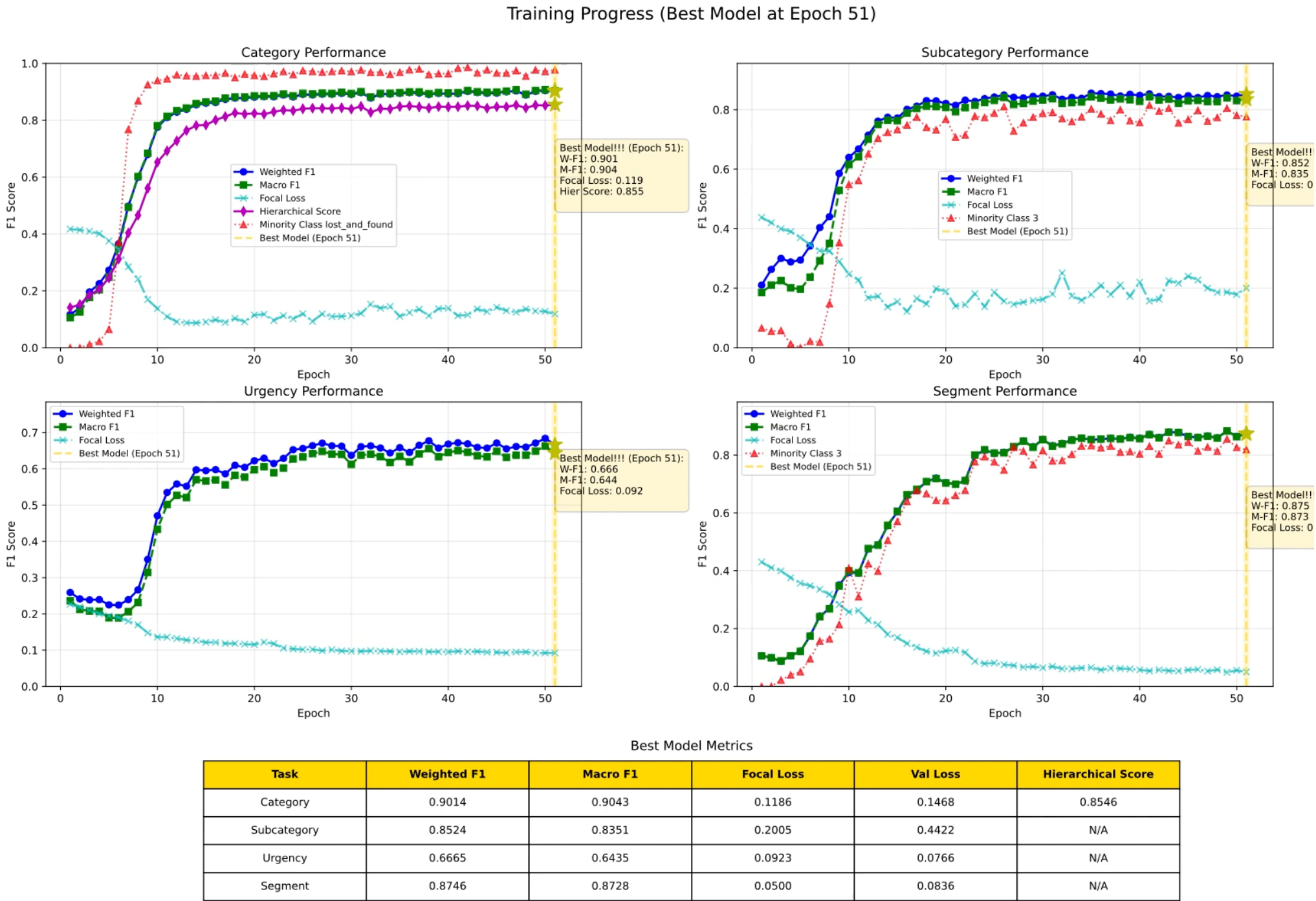To Train The Model

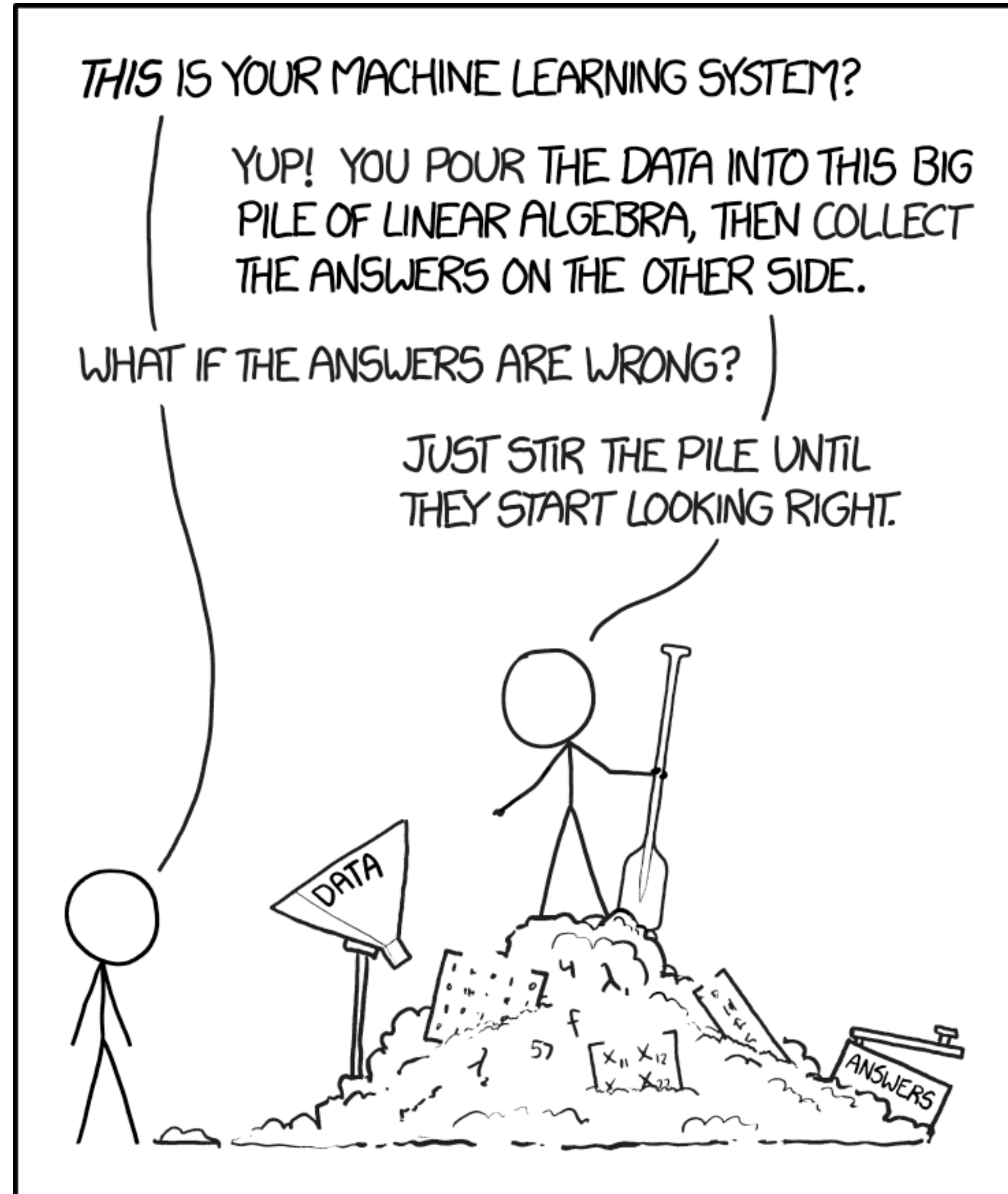# How Training Works

Training Visualizations

Hierarchical Confusion Matrix

Training Progress Tracker
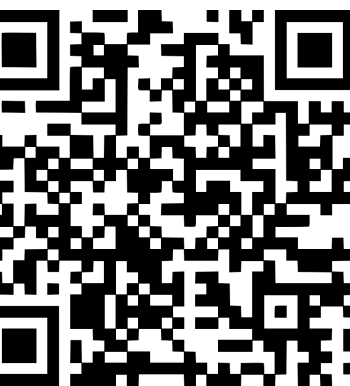
Modify Training Parameters

(xkcd.com)

Product Gestalt

Email Classifier Utility
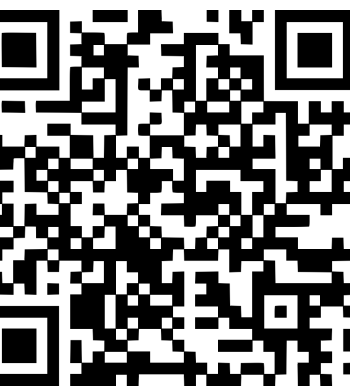
"Product Gestalt"

The "vibe-coding joke" has a reality behind it:
Desire and Opportunity to Express Product Ideas

- Non-tech people can create tech

  - Upside: room for product perspective to become relevant

  - Everyone can express product ideas

  - Constraint-driven people do it better

Engineering Beauty

Email Classifier Utility

Engineering Beauty, or Efficiency vs Convenience …Modes of Working, Choices

Big Model

- Effort: Quick solution

- Provides Convenience

- Generic toolset

- User Stance: Consumer

- Longevity: Replace
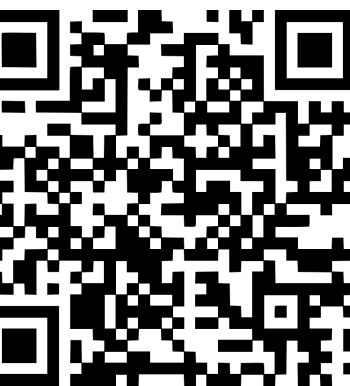
- Resources: Wasteful

- Result: Hit the Spot

or

Small Model

- Effort: Bigger lift

- Requires Expertise

- Specific toolset

- User Stance: Producer

- Longevity: Maintain

- Resources: Efficient

- Result: Scale

Economic Reality?

Email Classifier Utility

Economic Reality?

Museum tech as "small" industry might not provide a scaling opportunity

- Seems like an opportunity for a utility-type offering (true for a lot of tech)

- Are museums "customer engagement forward"?

- Also, "Small" is relative

Thank You.

Would love to hear from you.

Website, Contact Form

MCN Presentation Survey