Synthetic Experts

I. The Current Landscape: AI’s Environmental and Economic Challenges

1.1 AI Energy Consumption and Emissions

- **Environmental Cost**:
  - A single ChatGPT query = electricity to power a light bulb for 20 minutes (Allen Institute for AI, 2024).
  - Google’s greenhouse gas emissions **+48%** since 2019 due to AI.
  - Microsoft’s emissions **+29%** since 2020, linked to building more data centers for AI.
  - In 2023, Google **abandoned operational carbon neutrality**.
- **E-Waste**:
  - Rapid model obsolescence leads to significant electronic waste (Technology Review, 2024).

1.2 AI Industry and Costs

- **Training Large Models**:
  - **Meta’s Llama 3 405B** needed ~31 million GPU hours using **16,000 H100 GPUs**.
  - Training emissions equivalent to powering **15,000 households**.
- **Inference and Infrastructure**:
  - Even running queries on large models demands significant capital and operational expense.
  - Cost structure: Rent and own GPUs, electricity, cooling, infrastructure.
- **Hype Cycle**:

- **Edge AI** (smaller, localized models) and **Generative AI** at peak hype.

II. Rethinking AI: Specialized vs. Generalized Models

2.1 General AI (e.g., GPT-4, Gemini)

Feature	Characteristics
Scope	Wide — can tackle many tasks
Resource Usage	Very high
Accessibility	Largely centralized, expensive
Examples	GPT-4, Gemini, Claude-3

2.2 Specialized AI (e.g., Custom Classifiers)

Feature	Characteristics
Scope	Narrow — specific task focus
Resource Usage	Low (lighter models, faster inference)
Accessibility	Local, open-source, reproducible
Examples	Toxicity classifiers, MMX specialists

Analogy:

> "Why buy the whole candy store when you just need a lollypop?"

III. Specialized AI in Marketing: MMX on Social Media

3.1 Marketing Problem

- Marketing managers must **understand marketing mix effectiveness**.
- The 4 Ps of Marketing Mix:
    - **Product**
    - **Place**
    - **Price**
    - **Promotion**
- **Challenge**:

  Manual human analysis of millions of tweets/posts is **slow**, **expensive**, and **subjective**.
- **Opportunity**:

  Train specialized AI classifiers to **automatically categorize consumer comments** by MMX elements.

3.2 The Classification Task

- Example Tweet:

  > "Their new online shop is a catastrophe! Took me an hour to enter a billing address."

- **Correct Classification**:

  → *Place* (distribution/channel experience)
- **Simple Metrics Fail**:
    - Sentiment alone ("positive" vs. "negative") doesn't reveal *what* went wrong.

IV. Building Synthetic Specialists: A New Approach

4.1 Traditional vs. New Labeling Approaches

Method	Pros	Cons
Crowdsourcing (e.g., MTurk)	Fast, cheap for simple labels	Poor quality for complex constructs
Human Experts	High-quality labels	Expensive, time-intensive
Generative AI Labeling	Fast, low human input	Proprietary, costly per API call, privacy concerns

4.2 Synthetic Specialists Concept

- **Definition**:

  Specialized AI models **fine-tuned** on a specific classification task using **open-source architectures**.

- **Process**:
  1. Use **Generative AI** (GPT-4o, Llama 3) to **label** a large dataset.
  2. **Fine-tune** an **open-source model** (e.g., BERTweet) on this labeled data.
  3. Deploy this fine-tuned model **locally** for fast, cheap, private inference.

V. Methodology for Building a Synthetic Specialist

5.1 Step-by-Step Pipeline

1. **Data Collection**:
   - 699 brands
   - Tweets mentioning brands (N = 30,000 for training)
2. **Labeling with GenAI**:
   - 3 LLMs label each Tweet
   - Majority voting ensures label reliability

3. **Preprocessing**:
    - Replace URLs
    - Normalize whitespace
    - Anonymize users
4. **Model Fine-tuning**:
    - Base Model: **BERTweet-Large** (trained on 873M tweets)
    - Add classification layer for MMX categories.
    - Fine-tune with low learning rate, small batch sizes.
5. **Validation**:
    - Compare model output to **expert-labeled** tweets.
    - Use **Krippendorff's α** for agreement measure.

5.2 Fine-tuning Details

- **Training Loss**: Binary Cross Entropy
- **Optimizer**: Adam
- **Batch Size**: 16
- **Learning Rate**: 1e-5
- **Epochs**: 4 with early stopping

VI. Results and Evaluation

6.1 Evaluation Metrics

- **Krippendorff's α**:
    - $\alpha > 0.8$ = Reliable
    - α 0.67-0.8 = Acceptable
    - $\alpha < 0.67$ = Insufficient

- **Performance**:
  - Fine-tuned Specialist achieved **reliable agreement** with expert labels.
  - Performance **comparable to GenAI** but at a fraction of the cost.

6.2 Efficiency Gains

Metric	GenAI API	Synthetic Specialist
Tweets per second	~0.8	**114**
Cost per Tweet	~$0.002	**$0.0000002**
CO2 Emission	Significant	**Negligible**
Speed	Slow (API)	**Near real-time (local)**

VII. Applications and Implications

7.1 Insights Enabled by MMX Specialist

- Differentiate **Product, Place, Price, Promotion** feedback across brands.
- Detect **pain points** (e.g., Abercrombie's e-commerce issues).
- Segment and **cluster consumer discussions**.
- Perform **topic modeling** and **sentiment analysis** by MMX category.

7.2 Broader Organizational Impact

- **Business**:
  - Faster, cheaper, greener AI deployments.
  - No dependency on external GenAI providers.
- **Research**:
  - Synthetic specialists allow **reproducibility** and **greater transparency**.
- **Society**:
  - **Local AI**: democratization of powerful tools without ecological disaster.

VIII. Conclusion: Why Synthetic Specialists Matter

Advantage	Impact
8% more accurate	vs. direct GenAI labeling
150x faster	inference than GenAI
8,400x cheaper	per inference
99.99% lower CO2 emissions	vs. GenAI API