The Environmental Cost of AI: Why Efficiency Matters
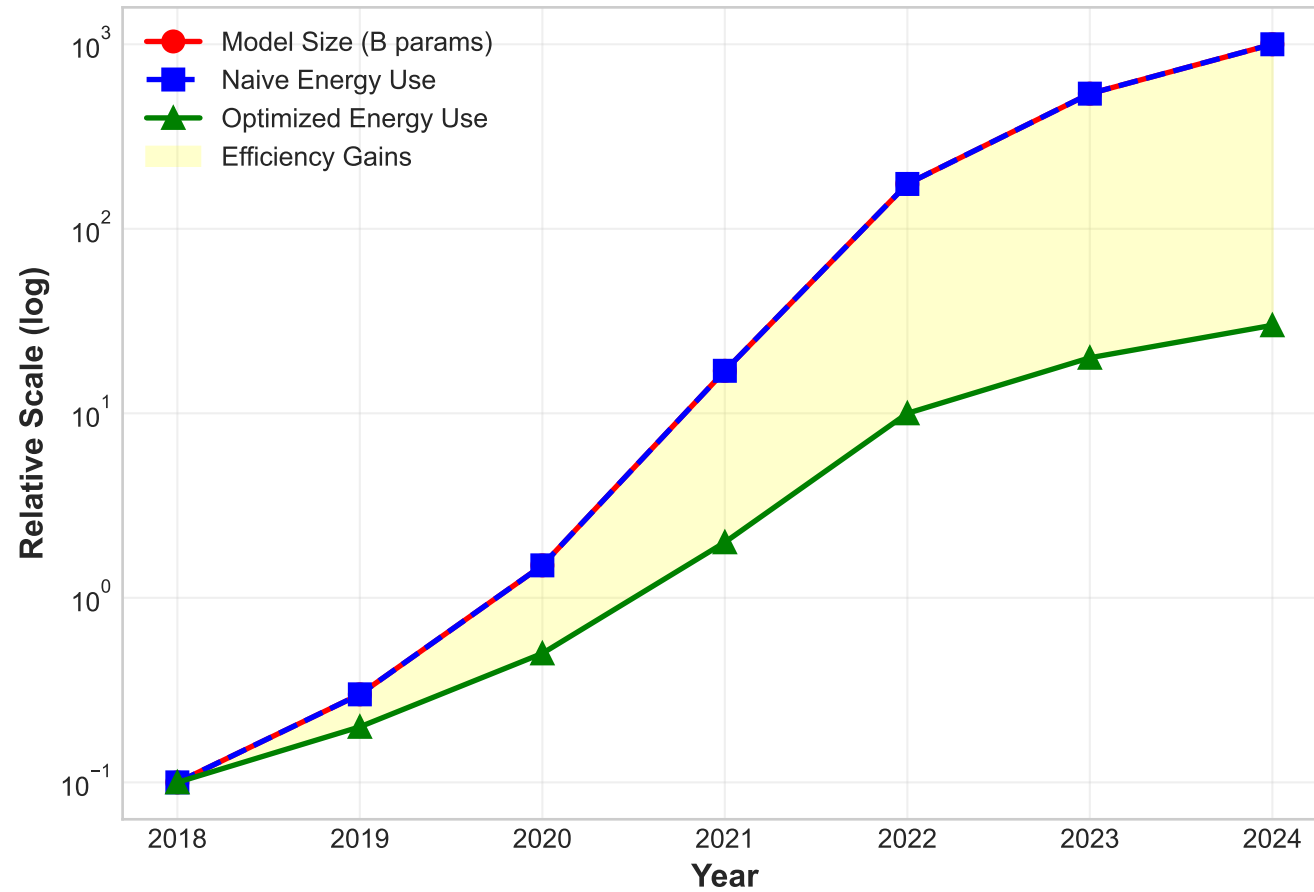
Model Size vs Energy Consumption Trends

Legend:
- Model Size (B params)
- Naive Energy Use
- Optimized Energy Use
- Efficiency Gains

X-axis: Year (2018, 2019, 2020, 2021, 2022, 2023, 2024)
Y-axis: Relative Scale (log)

Carbon Footprint of Different Models

Y-axis: CO_2 Emissions (kg)

- GPT-3 Training: 552k kg
- GPT-3 Inference (1M queries): 150.0 kg
- BERT Training: 1k kg
- DistilBERT Training: 144.0 kg
- Mobile Model Training: 14.0 kg
- Edge Inference (1M queries): 0.5 kg

Car trip NYC-SF