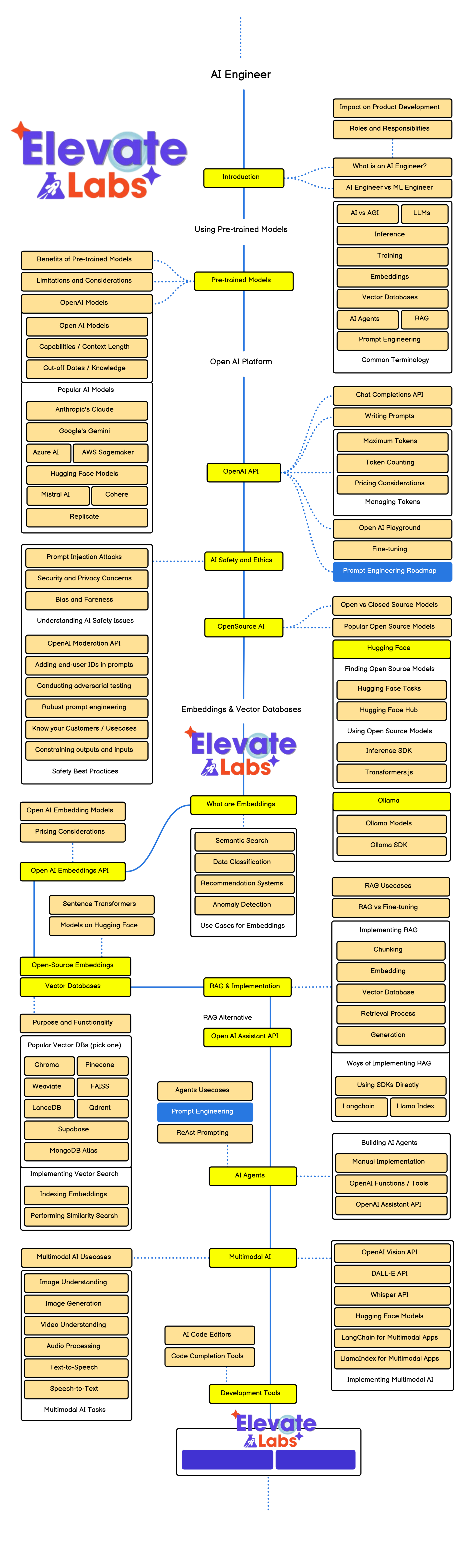
AI Engineer
Introduction
Impact on Product Development
Roles and Responsiblities
What is an AI Engineer?
AI Engineer vs ML Engineer
AI vs AGI
LLMs
Inference
Training
Embeddings
Vector Databases
AI Agents
RAG
Prompt Engineering
Common Terminology
Using Pre-trained Models
Pre-trained Models
Benefits of Pre-trained Models
Limitations and Considerations
OpenAI Models
Open AI Models
Capabilities / Context Length
Cut-off Dates / Knowledge
Popular AI Models
Anthropic's Claude
Google's Gemini
Azure AI
AWS Sagemaker
Hugging Face Models
Mistral AI
Cohere
Replicate
Open AI Platform
OpenAI API
Chat Completions API
Writing Prompts
Maximum Tokens
Token Counting
Pricing Considerations
Managing Tokens
Open AI Playground
Fine-tuning
Prompt Engineering Roadmap
AI Safety and Ethics
Prompt Injection Attacks
Security and Privacy Concerns
Bias and Fareness
Understanding AI Safety Issues
OpenAI Moderation API
Adding end-user IDs in prompts
Conducting adversarial testing
Robust prompt engineering
Know your Customers / Usecases
Constraining outputs and inputs
Safety Best Practices
OpenSource AI
Open vs Closed Source Models
Popular Open Source Models
Hugging Face
Finding Open Source Models
Hugging Face Tasks
Hugging Face Hub
Using Open Source Models
Inference SDK
Transformers.js
Ollama
Ollama Models
Ollama SDK
Embeddings & Vector Databases
What are Embeddings
Semantic Search
Data Classification
Recommendation Systems
Anomaly Detection
Use Cases for Embeddings
Open AI Embedding Models
Pricing Considerations
Open AI Embeddings API
Sentence Transformers
Models on Hugging Face
Open-Source Embeddings
Vector Databases
Purpose and Functionality
Popular Vector DBs (pick one)
Chroma
Pinecone
Weaviate
FAISS
LanceDB
Qdrant
Supabase
MongoDB Atlas
Implementing Vector Search
Indexing Embeddings
Performing Similarity Search
RAG & Implementation
RAG Usecases
RAG vs Fine-tuning
Implementing RAG
Chunking
Embedding
Vector Database
Retrieval Process
Generation
Ways of Implementing RAG
Using SDKs Directly
Langchain
Llama Index
RAG Alternative
Open AI Assistant API
Agents Usecases
Prompt Engineering
ReAct Prompting
AI Agents
Building AI Agents
Manual Implementation
OpenAI Functions / Tools
OpenAI Assistant API
Multimodal AI Usecases
Image Understanding
Image Generation
Video Understanding
Audio Processing
Text-to-Speech
Speech-to-Text
Multimodal AI Tasks
Multimodal AI
OpenAI Vision API
DALL-E API
Whisper API
Hugging Face Models
LangChain for Multimodal Apps
LlamaIndex for Multimodal Apps
Implementing Multimodal AI
AI Code Editors
Code Completion Tools
Development Tools