

# Agentic AI Developer Roadmap (2026)

A practical, hands-on learning path for building LLM-powered autonomous agents with TypeScript & Node.js

🚩 You Are Here

## 1. LLM Foundations

### LLM Fundamentals

What Is a Large Language Model? [↗](#)  
The Illustrated Transformer [↗](#)  
Andrej Karpathy: Intro to LLMs (Video) [↗](#)  
3Blue1Brown: Neural Networks (Video) [↗](#)

### Reasoning Models

OpenAI o1 System Card [↗](#)  
Learning to Reason with LLMs [↗](#)  
DeepSeek R1 Technical Report [↗](#)  
Reasoning Tokens and Effort Controls [↗](#)

### Prompt Engineering

OpenAI Prompt Engineering Guide [↗](#)  
Anthropic Prompt Engineering Docs [↗](#)  
Brex's Prompt Engineering Guide [↗](#)  
ChatGPT Prompt Engineering (DeepLearning.AI) [↗](#)  
Prompt Engineering Guide [↗](#)

### Structured Output

OpenAI Structured Outputs [↗](#)  
Instructor Library [↗](#)  
Outlines (Constrained Generation) [↗](#)

### APIs & SDKs

OpenAI Node.js SDK [↗](#)  
OpenAI Responses API [↗](#)  
Anthropic TypeScript SDK [↗](#)  
Google Gemini API [↗](#)  
Vercel AI SDK [↗](#)  
GitHub Copilot SDK [↗](#)  
Claude Agent SDK [↗](#)  
OpenRouter (Multi-Provider Gateway) [↗](#)

### Tokenization

OpenAI Tokenizer [↗](#)  
tiktoken (JS port) [↗](#)  
gpt-tokenizer [↗](#)

### Extended Context & Caching

Prompt Caching (Anthropic) [↗](#)  
Context Caching (OpenAI) [↗](#)  
Long Context Prompting Strategies [↗](#)  
Gemini 2M Context Window [↗](#)

### Local & Open-Source Models

Ollama [↗](#)  
Ollama.js [↗](#)  
LM Studio [↗](#)  
DeepSeek Models [↗](#)  
Hugging Face Transformers.js [↗](#)

### Courses

Vercel AI SDK Docs & Tutorials [↗](#)  
AI Engineer Roadmap (dswh Video Series) [↗](#)

## 2. Single-Agent Systems

### ReAct Pattern

ReAct: Synergizing Reasoning and Acting (Paper) [↗](#)  
ReAct Prompting Explained [↗](#)

### Agent Fundamentals

LLM Powered Autonomous Agents (Lilian Weng) [↗](#)  
Building Effective Agents (Anthropic) [↗](#)  
What Are AI Agents? (OpenAI) [↗](#)  
GitHub Copilot Agentic Capabilities [↗](#)  
Functions, Tools and Agents with LangChain (DeepLearning.AI) [↗](#)

### Function Calling

OpenAI Function Calling Guide [↗](#)  
Anthropic Tool Use Docs [↗](#)  
Claude Agent Skills Overview [↗](#)

### Agent Frameworks

LangGraph.js Docs [↗](#)  
Vercel AI SDK [↗](#)  
OpenAI Agents SDK [↗](#)  
Claude Agent SDK [↗](#)  
GitHub Copilot SDK [↗](#)  
Mastra [↗](#)  
OpenAI Swarm (Python) [↗](#)  
CrewAI (Python) [↗](#)  
AutoGen (Python) [↗](#)

### Output Validation

Zod [↗](#)  
TypeBox [↗](#)  
Instructor-JS [↗](#)

### Testing Agents

Testing AI Applications [↗](#)  
LangSmith Testing Guide [↗](#)  
Mock Service Worker (msw) [↗](#)

### Tracing

LangSmith Docs [↗](#)  
Arize Phoenix [↗](#)  
Braintrust [↗](#)

## 3. Tool Design & External Integrations

### MCP (Model Context Protocol) & A2A

MCP Specification [↗](#)  
MCP TypeScript SDK [↗](#)  
Building MCP Servers (Anthropic) [↗](#)  
A2A Protocol (Agent-to-Agent) [↗](#)

### Tool Design

OpenAI Function Calling Best Practices [↗](#)  
Anthropic Tool Use Best Practices [↗](#)  
Claude Agent Skills Best Practices [↗](#)

### Sandboxing

E2B (Code Execution Sandbox) [↗](#)  
gVisor (Container Sandboxing) [↗](#)  
Firecracker (MicroVMs) [↗](#)  
WebAssembly System Interface (WASI) [↗](#)

### Browser Automation

Playwright [↗](#)  
Puppeteer [↗](#)  
Browserbase [↗](#)

### Circuit Breakers & Resilience

Circuit Breaker Pattern (Martin Fowler) [↗](#)  
Cockatiel (Node.js Resilience) [↗](#)  
p-retry [↗](#)

## 4. Retrieval, Planning & Advanced Reasoning

### RAG

RAG From Scratch (LangChain series) [↗](#)  
Retrieval-Augmented Generation (original paper) [↗](#)  
Pinecone: What is RAG? [↗](#)

### Vector Databases

Pinecone [↗](#)  
Weaviate [↗](#)  
Qdrant [↗](#)  
pgvector [↗](#)  
Chroma [↗](#)

### Embeddings

OpenAI Embeddings Guide [↗](#)  
MTEB Leaderboard (Model Comparisons) [↗](#)  
Transformers.js (Run Embeddings in Node) [↗](#)

### Chunking Strategies

Chunking Strategies for LLM Apps (Pinecone) [↗](#)  
Five Levels of Chunking (Greg Kamradt) [↗](#)

### Planning & Reflection

Reflexion: Language Agents with Verbal Reinforcement [↗](#)  
Plan-and-Solve Prompting (Paper) [↗](#)

### Code Execution

E2B Docs [↗](#)  
Modal [↗](#)

### Fine-Tuning

OpenAI Fine-Tuning Guide [↗](#)  
When to Fine-Tune LLMs [↗](#)

### Multi-Modal

OpenAI Vision API [↗](#)  
Anthropic Vision Docs [↗](#)  
Vercel AI SDK: Multi-Modal [↗](#)

### Advanced RAG

Building & Evaluating Advanced RAG (DeepLearning.AI) [↗](#)  
Advanced RAG Cheat Sheet [↗](#)  
GraphRAG (Microsoft) [↗](#)  
Evaluating RAG Components (HF Cookbook) [↗](#)

## 5. Memory & State

### Memory Concepts

LLM Powered Agents — Memory (Lilian Weng) [↗](#)  
LangGraph.js: Memory Concepts [↗](#)

### Memory Systems

Mem0 (Memory Layer for AI) [↗](#)  
Zep (Long-Term Memory) [↗](#)  
LangGraph.js Persistence [↗](#)

### Context Management

Managing Conversation History (OpenAI Cookbook) [↗](#)  
Vercel AI SDK: Managing Conversations [↗](#)

### Checkpointing

LangGraph.js Checkpointers [↗](#)  
Temporal TypeScript SDK: Workflow State [↗](#)

## 6. Multi-Agent Orchestration

### Multi-Agent Patterns

Multi-Agent Systems (LangGraph.js) [↗](#)  
OpenAI Swarm Framework [↗](#)  
CrewAI: Multi-Agent Workflows [↗](#)  
Mastra: Multi-Agent Workflows [↗](#)  
AutoGen: Multi-Agent Conversation [↗](#)

### Supervisor & Router

LangGraph.js Supervisor Tutorial [↗](#)  
Build a Customer Support Bot (LangGraph.js) [↗](#)

### DAG Pipelines

LangGraph.js: How-to Guides [↗](#)  
LangGraph Cloud [↗](#)  
Inngest (Durable Functions) [↗](#)  
Temporal (TypeScript SDK) [↗](#)

### Human-in-the-Loop

LangGraph.js: Human-in-the-Loop [↗](#)  
Inngest: Wait for Event [↗](#)

### Case Studies

How Replit Built Their Agent [↗](#)  
Devin Architecture Analysis [↗](#)  
Cursor Agent Technical Overview [↗](#)

## 7. Production & Operations

### Observability & Tracing

LangSmith Docs [↗](#)  
Langfuse (Open-Source LLM Engineering) [↗](#)  
Arize Phoenix [↗](#)  
Braintrust [↗](#)  
Helicone (LLM Observability) [↗](#)  
OpenLIT [↗](#)  
OpenTelemetry [↗](#)

### Evaluation

Promptfoo (LLM Eval Framework) [↗](#)  
Ragas (RAG Evaluation) [↗](#)  
DeepEval [↗](#)  
OpenAI Evals [↗](#)

### Guardrails

Guardrails AI [↗](#)  
NeMo Guardrails [↗](#)  
Lakera Guard [↗](#)

### Security

OWASP Top 10 for LLM Applications [↗](#)  
Prompt Injection Primer (Simon Willison) [↗](#)  
Rebuff (Prompt Injection Detection) [↗](#)

### Deployment

Docker [↗](#)  
Kubernetes [↗](#)  
AWS Bedrock Agents [↗](#)  
Vercel AI [↗](#)  
Cloudflare Workers AI [↗](#)  
Modal [↗](#)  
Fly.io [↗](#)  
Railway [↗](#)

### Cost Optimization

OpenAI API Pricing [↗](#)  
Anthropic Pricing [↗](#)  
Prompt Caching (Anthropic) [↗](#)  
LLM Pricing Calculator [↗](#)

### LLMOps

LLMOps Guide (Databricks) [↗](#)  
LLMOps (DeepLearning.AI) [↗](#)  
Anthropic Workbench [↗](#)  
Inngest (Durable Agent Workflows) [↗](#)  
Weights & Biases (Prompts) [↗](#)

### Red Teaming

Red Teaming LLM Applications (DeepLearning.AI) [↗](#)  
Automated Testing for LLMOps (DeepLearning.AI) [↗](#)  
AI Risk Resources (OWASP) [↗](#)

## 8. Capstone Project

🔧 Expert Agentic AI Developer