TORONTO AI SCHOOL

FLAGSHIP ENTERPRISE PROGRAM

CEAAP™

Certified Enterprise AI Architect & Agentic Systems Professional

From AI Foundations to Production-Grade RAG, Multi-Agent Systems,
Amazon Bedrock & **AgentCore**

10	40	600+	1
WEEKS	DEEP-DIVE MODULES	LEARNING & BUILD ACTIVITIES	END-TO-END ENTERPRISE AI PLATFORM

TEN WEEKS · FIRST PRINCIPLES → PRODUCTION

Ten Weeks From First Principles to Production

CEAAP™ is Toronto AI School's flagship enterprise program for **engineers, architects and technical leaders** who want to design, secure and ship real agentic AI systems — not just prompt a chatbot. Across 40 deep-dive modules you will architect, build, secure, evaluate and deploy one continuous capstone platform: **the TAS Agentic Student Hub**, running on Amazon Bedrock and AgentCore.

10
Weeks of structured building

40
Deep-dive modules

600+
Learning & build activities

1
End-to-end enterprise AI platform

What You Will Design, Secure & Ship

Enterprise AI Architecture
Requirements, lifecycle, AWS & Bedrock reference designs

Advanced RAG
Hybrid search, reranking, GraphRAG & a golden evaluation suite

Multi-Agent Systems
Supervisor / specialist teams orchestrated on AgentCore

Identity, Memory & Guardrails
Least-privilege agents with safe, isolated memory

Security & Red Teaming
Prompt-injection defense and adversarial evaluation

Production Operations
Terraform, CI/CD, observability, FinOps & reliability engineering

The Learning Arc

1 **Foundations & Architecture**
Modules 01–05

2 **Bedrock, RAG & Retrieval**
Modules 06–12

3 **Agents & AgentCore**
Modules 13–20

4 **Security & Governance**
Modules 21–24

5 **Production Engineering**
Modules 25–32

6 **Enterprise Integration & Defense**
Modules 33–40

ALL 40 MODULES AT A GLANCE

The Complete Curriculum Map

A quick-reference map of the complete CEAAP™ curriculum. Full module detail — every learning objective and hands-on build — follows on the pages ahead.

PHASE 1 Foundations & Architecture — Modules 01–05

- 01 🚀 From "What Is AI?" to "How Does Enterprise AI Actually Work?"
- 02 🎯 Before You Touch the Code: Turn a Business Problem into an AI Blueprint
- 03 🧬 The Enterprise AI Lifecycle: From Whiteboard Idea to Production System
- 04 🏗 Think Like an AI Solutions Architect: Design the System Before Building It
- 05 ☁ Build the Cloud Brain: AWS Architecture for Enterprise AI

PHASE 2 Bedrock, RAG & Retrieval — Modules 06–12

- 06 🧠 Amazon Bedrock Unlocked: Build Your Enterprise GenAI Engine
- 07 🎛 One AI Is Never Enough: Build an Intelligent Model Gateway
- 08 🔐 Your AI Has Access to Company Data—Now Secure It Like an Enterprise
- 09 📚 Give Your AI a Private Brain: Build RAG from Zero
- 10 🔥 RAG on Steroids: Turn Basic Search into Advanced Enterprise Retrieval
- 11 🔎 Inside the Vector Brain: Make Enterprise Knowledge Search Actually Work
- 12 🧪 "It Gave an Answer" Is Not Testing: Prove Your RAG Actually Works

PHASE 3 Agents & AgentCore — Modules 13–20

- 13 🤖 Stop Building Chatbots: Build an AI Agent That Can Actually Do Work
- 14 🧭 Who Is in Charge? Master the Agent Orchestration Patterns Enterprises Use
- 15 👥 Build Your AI Workforce: Supervisor + Specialist Multi-Agent Systems
- 16 🛠 Give Your Agents Hands: APIs, Tools & Enterprise Actions
- 17 🧠 Meet AgentCore: From AI Prototype to Enterprise Agent Platform
- 18 ⚡ Put Your Agent to Work: AgentCore Runtime & Gateway Deep Dive
- 19 🪪 "Who Are You Acting As?" Solve the Agent Identity Problem
- 20 🧠 Should Your AI Remember You? Engineer Agent Memory Without Creating a Privacy Nightmare

MODULES 21–40

Security, Production & Defense

PHASE 4 Security & Governance — Modules 21–24

21 🛡 Put Boundaries Around Intelligence: Enterprise AI Guardrails

22 ⚔ Attack Your Own AI Before Someone Else Does: Prompt Injection & Red Teaming

23 🔏 Your LLM Doesn't Need Every Piece of Customer Data: Privacy-by-Design

24 🏛 Who Approves the AI? Build an Enterprise AI Governance Operating Model

PHASE 5 Production Engineering — Modules 25–32

25 🌐 Build the Front Door to Your AI: Production API Architecture

26 ⏳ What If an AI Task Takes 20 Minutes? Build Durable Agent Workflows

27 ⚡ Make AI Faster—and Cheaper: The Enterprise Caching Playbook

28 💥 Your Model Is Throttled at 2 AM—Now What? Reliability Engineering for AI

29 📊 Can You Prove Your Agent Is Good? Enterprise Agent Evaluation

30 🔭 See Inside the AI Black Box: Observability from User Click to Agent Action

31 ♻ AI Changes Every Week—How Do You Operate It for Years? LLMOps & AgentOps

32 💰 Your AI Works—But Can the Company Afford It? AI FinOps & Performance

PHASE 6 Enterprise Integration & Defense — Modules 33–40

33 ☎ Turn the Contact Centre into an AI Workforce: Amazon Connect + Agentic AI

34 🎙 Beyond Chat: Build Voice & Multimodal AI Experiences

35 🔌 AI Can't Live on an Island: Connect Agents to Real Enterprise Systems

36 🏗 Stop Clicking in the Console: Build Your Entire AI Platform with Terraform

37 🚢 From Git Commit to Production AI: Build the CI/CD Factory

38 🚨 Production Day: 10,000 Users Are Coming—Is Your AI Actually Ready?

39 🧑‍💼 Face the Architecture Board: Can You Defend Every Decision You Made?

40 🏆 The CEAAP Enterprise Build: From Empty Folder to Production-Grade Agentic AI Platform

TORONTO AI SCHOOL

PHASE 1 OF 6 · MODULES 01–05

Foundations & Architecture

Think like an AI architect — from first principles and requirements to your first AWS reference design.

5	70	5
MODULES	TOPICS	HANDS-ON BUILDS & LABS

01 🚀 From "What Is AI?" to "How Does Enterprise AI Actually Work?"

02 🎯 Before You Touch the Code: Turn a Business Problem into an AI Blueprint

03 🧬 The Enterprise AI Lifecycle: From Whiteboard Idea to Production System

04 🏗 Think Like an AI Solutions Architect: Design the System Before Building It

05 ☁ Build the Cloud Brain: AWS Architecture for Enterprise AI

01



From "What Is AI?" to "How Does Enterprise AI Actually Work?"

WHAT YOU'LL LEARN

14 TOPICS

01 AI Without the Buzzwords
What actually makes a system intelligent?

02 AI → ML → Deep Learning → GenAI → Agentic AI
The evolution that changed enterprise technology

03 Inside an LLM
How does ChatGPT-style AI actually generate an answer?

04 Tokens, Context & Predictions
What happens after you press Send?

05 Training vs Inference
When is the model learning—and when is it working?

06 Foundation Models Explained
Why enterprises don't train everything from scratch

07 Generative AI vs Traditional AI
When should you use each?

08 RAG vs Fine-Tuning vs Agents
Which problem needs which architecture?

09 What Makes AI "Agentic"?
From answering questions to completing work

10 The Enterprise AI Technology Stack
Model → Data → RAG → Agents → Tools → APIs

11 Where AWS, Azure & Google Cloud Fit

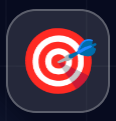
12 Why Enterprise AI Projects Fail Even When the Model Works

13 From Business Problem → AI Use Case
Finding opportunities worth building

14 Your First Enterprise AI Architecture
Whiteboard challenge

★ CAPSTONE MOMENT

CEAAP Capstone Launch: Meet the Toronto AI School Agentic Student Hub

02  Before You Touch the Code: Turn a Business Problem into an AI Blueprint

WHAT YOU'LL LEARN — 14 TOPICS

01 **"We Need AI" Isn't a Requirement**
Asking the questions an AI Architect asks

02 **Discover the Real Problem Before Choosing the Model**

03 **Who Will Actually Use This System?**
Persona & stakeholder mapping

04 **Turn Business Pain into AI Use Cases**

05 **Functional Requirements Without Boring Requirement Documents**

06 **What Must the AI Never Do?**
Boundary design

07 **How Fast Is Fast Enough?**
Designing latency targets

08 **How Accurate Is Accurate Enough?**
Quality thresholds

09 **What Happens When 10,000 Students Arrive Together?**
Scalability requirements

10 **What Happens When AWS Goes Down?**
Availability thinking

11 **How Much Should One AI Conversation Cost?**
Cost requirements

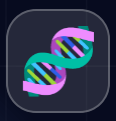
12 **What Student Data Can the LLM See?**
Privacy requirements

13 **Human vs AI**
Designing escalation boundaries

14 **Turn Requirements into Measurable Acceptance Criteria**

BUILD Create 01-requirements.md + 02-success-metrics.md

03



The Enterprise AI Lifecycle: From Whiteboard Idea to Production System

WHAT YOU'LL LEARN — 14 TOPICS

01 **Why Normal SDLC Isn't Enough for AI**

02 **SDLC vs ML Lifecycle vs GenAI Lifecycle vs Agentic Lifecycle**

03 **Discovery**
Should this problem even use AI?

04 **Data**
What information will power the system?

05 **Model**
Which intelligence layer fits the task?

06 **Prompt**
How do instructions become system behavior?

07 **Retrieval**
How does AI access private enterprise knowledge?

08 **Tools**
How does AI actually take action?

09 **Agents**
How do we give AI controlled autonomy?

10 **Evaluation**
How do we prove it works?

11 **Security**
How do we stop misuse?

12 **Deployment**
How does AI reach production?

13 **Monitoring**
How do we know it's still working tomorrow?

14 **Feedback Loops**
How production systems improve safely

★ **CAPSTONE MOMENT** — Architecture Workshop: Map the complete TAS AI lifecycle

04 Think Like an AI Solutions Architect: Design the System Before Building It

WHAT YOU'LL LEARN

14 TOPICS

01 **Business Architecture → Application Architecture → AI Architecture**

02 **Draw the Whole System Without Writing Code**

03 **Who Talks to Whom?**
Component architecture

04 **Where Does Data Travel?**
Data-flow architecture

05 **What Happens After the User Clicks Send?**
Sequence diagrams

06 **C4 Architecture Without the Confusion**

07 **Logical vs Physical Architecture**

08 **Synchronous vs Asynchronous AI Workflows**

09 **API-Driven Architecture**

10 **Event-Driven Architecture**

11 **Trust Boundaries**
Where security architecture begins

12 **Design for Failure Before Failure Happens**

13 **ADR**
How Architects Record Million- Dollar Technology Decisions

14 **Architecture Review Challenge**
Defend your design

BUILD 04-architecture.md, 05-data-flow.md, 11-sequence-diagrams.md

05 Build the Cloud Brain: AWS Architecture for Enterprise AI

WHAT YOU'LL LEARN — 14 TOPICS

01 **Which AWS Services Actually Matter for AI Architects?**

02 **Accounts, Regions, AZs & Enterprise Landing Zones**

03 **IAM**
Who can touch what?

04 **VPC**
Where does your AI live?

05 **PrivateLink**
Can AI operate without public internet exposure?

06 **KMS**
Who owns the encryption keys?

07 **S3**
Building the enterprise knowledge lake

08 **Lambda**
Serverless intelligence glue

09 **API Gateway**
Opening controlled doors to AI

10 **DynamoDB**
State, sessions & metadata

11 **SQS/EventBridge**
Making AI asynchronous

12 **Secrets Manager**
Keeping credentials out of code

13 **CloudWatch vs CloudTrail**
Performance vs accountability

14 **Dev → Test → Stage → Prod**
Enterprise environment design

BUILD First AWS reference architecture

TORONTO AI SCHOOL

PHASE 2 OF 6 · MODULES 06–12

Bedrock, RAG & Retrieval

Build the GenAI engine, secure it, and engineer retrieval that is measurably correct.

7	98	7
MODULES	TOPICS	HANDS-ON BUILDS & LABS

06 🧠 Amazon Bedrock Unlocked: Build Your Enterprise GenAI Engine

07 🎛 One AI Is Never Enough: Build an Intelligent Model Gateway

08 🔐 Your AI Has Access to Company Data—Now Secure It Like an Enterprise

09 📚 Give Your AI a Private Brain: Build RAG from Zero

10 🔥 RAG on Steroids: Turn Basic Search into Advanced Enterprise Retrieval

11 🔎 Inside the Vector Brain: Make Enterprise Knowledge Search Actually Work

12 🧪 "It Gave an Answer" Is Not Testing: Prove Your RAG Actually Works

06  Amazon Bedrock Unlocked: Build Your Enterprise GenAI Engine

WHAT YOU'LL LEARN — 14 TOPICS

01 **What Actually Happens Inside Amazon Bedrock?**

02 **Choosing a Foundation Model Without Guesswork**

03 **Model Families & Enterprise Workload Matching**

04 **Inference**
What happens when your application calls a model?

05 **Converse API**
One interface, multiple model possibilities

06 **Streaming**
Why users shouldn't wait for the full answer

07 **Structured Output**
Make AI return data your software can trust

08 **Embeddings**
Turning meaning into mathematics

09 **Knowledge Bases**
Giving models enterprise knowledge

10 **Agents**
Moving from answers to actions

11 **Guardrails**
Controlling model boundaries

12 **Inference Profiles & Capacity Thinking**

13 **Tokens → Latency → Cost**
The relationship every architect must understand

14 **What Happens When Bedrock Throttles You?**

⌘ CODE — Build bedrock_service.py

07 One AI Is Never Enough: Build an Intelligent Model Gateway

WHAT YOU'LL LEARN — 14 TOPICS

01 **Stop Hard-Coding Your Application to One Model**

02 **Model Abstraction**
Swap models without rewriting your application

03 **Can a Cheap Model Handle This Request?**
Smart routing

04 **When Do We Need the Stronger Model?**
Complexity routing

05 **Temperature & Top-P Without the Theory Headache**

06 **Context Windows**
Why your model "forgets"

07 **Token Budgeting Before You Hit Production**

08 **System Prompts as Architecture**

09 **Reusable Prompt Templates**

10 **Force Reliable JSON Output**

11 **Timeouts**
What if the model takes too long?

12 **Retries**
When should you try again?

13 **Exponential Backoff + Jitter**

14 **Fallback Models**
Keep the product alive when the primary fails

CODE model_gateway.py + model_router.py

08  Your AI Has Access to Company Data—Now Secure It Like an Enterprise

WHAT YOU'LL LEARN — 14 TOPICS

01 **What Could Go Wrong?**
AI threat modeling

02 **Zero Trust for AI Applications**

03 **Least Privilege Without Breaking the Agent**

04 **User Identity vs Agent Identity vs Service Identity**

05 **Who Is Allowed to Invoke the Model?**

06 **Encrypt Data at Rest and in Transit**

07 **Protect Secrets from Developers, Logs & LLMs**

08 **Private AI Networking**

09 **Tenant Isolation**

10 **Sensitive Data Classification**

11 **Secure API Boundaries**

12 **Audit Every Important AI Action**

13 **CloudTrail Forensics**

14 **Security Architecture Review**

BUILD 06-security-design.md

09  Give Your AI a Private Brain: Build RAG from Zero

WHAT YOU'LL LEARN — 14 TOPICS

01 **Why LLMs Don't Know Your Company's Data**

02 **RAG Explained from User Question to Final Answer**

03 **Documents → Parsing → Cleaning → Chunks**

04 **What Is an Embedding Really Doing?**

05 **Vector Search Without the Math Headache**

06 **Similarity Search**

07 **Metadata**
The Secret Weapon People Ignore

08 **Indexing Enterprise Knowledge**

09 **Query → Retrieval → Context → Generation**

10 **How Many Chunks Should We Retrieve?**

11 **Prevent the Model from Answering Outside the Knowledge Base**

12 **Add Citations Users Can Trust**

13 **Handle Missing Answers Correctly**

14 **Build Your First Knowledge Base**

⌘ CODE — Complete basic RAG pipeline

10 RAG on Steroids: Turn Basic Search into Advanced Enterprise Retrieval

WHAT YOU'LL LEARN — 14 TOPICS

01 **Why Basic RAG Falls Apart in Production**

02 **Fixed vs Semantic vs Recursive Chunking**

03 **Parent–Child Retrieval**

04 **Hybrid Search**
Keywords + Meaning

05 **Metadata Filtering Before Retrieval**

06 **Query Rewriting**
Fix the User's Bad Question

07 **Multi-Query Retrieval**

08 **HyDE**
Search Using a Hypothetical Answer

09 **Reranking**
Let the Best Context Win

10 **Contextual Compression**

11 **Dynamic Top-K**

12 **Corrective RAG**

13 **Agentic RAG**

14 **GraphRAG**
When Relationships Matter

⌘ CODE — chunking.py, retriever.py, reranker.py, query_rewriter.py

11



Inside the Vector Brain: Make Enterprise Knowledge Search Actually Work

WHAT YOU'LL LEARN — 14 TOPICS

01 What Happens to a Document After Upload?

02 Embeddings End-to-End

03 Choosing Embedding Dimensions

04 Cosine Similarity & Distance Intuition

05 Vector Indexes Explained

06 OpenSearch Serverless for Enterprise RAG

07 HNSW Without a PhD

08 Design Metadata Schemas That Scale

09 S3 → Knowledge Base → Vector Store

10 Knowledge Base Sync Lifecycle

11 Incremental Updates

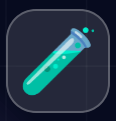
12 What Happens When a Source Document Changes?

13 What Happens When It Is Deleted?

14 Multi-Tenant Knowledge Architecture

⌘ CODE — vector_store.py + ingestion pipeline

12



"It Gave an Answer" Is Not Testing: Prove Your RAG Actually Works

WHAT YOU'LL LEARN — 14 TOPICS

01 Retrieval Quality vs Answer Quality

02 Did We Retrieve the Right Document?

03 Precision & Recall for RAG

04 Hit Rate

05 MRR/NDCG Intuition

06 Is the Answer Grounded in Retrieved Evidence?

07 Groundedness vs Faithfulness

08 Correctness vs Relevance

09 Are the Citations Actually Correct?

10 Build a Golden Evaluation Dataset

11 Generate Synthetic Test Questions

12 Test No-Answer Scenarios

13 Catch Hallucinations Before Customers Do

14 Regression Testing for RAG

⌘ CODE — rag_evaluator.py + evaluation dashboard

TORONTO AI SCHOOL

PHASE 3 OF 6 · MODULES 13–20

Agents & AgentCore

Move from chatbots to agents that act — orchestrated, identity-aware and memory-safe on AgentCore.

8	112	8
MODULES	TOPICS	HANDS-ON BUILDS & LABS

13 🤖 Stop Building Chatbots: Build an AI Agent That Can Actually Do Work

14 🧭 Who Is in Charge? Master the Agent Orchestration Patterns Enterprises Use

15 👥 Build Your AI Workforce: Supervisor + Specialist Multi-Agent Systems

16 🛠 Give Your Agents Hands: APIs, Tools & Enterprise Actions

17 🧠 Meet AgentCore: From AI Prototype to Enterprise Agent Platform

18 ⚡ Put Your Agent to Work: AgentCore Runtime & Gateway Deep Dive

19 🪪 "Who Are You Acting As?" Solve the Agent Identity Problem

20 🧠 Should Your AI Remember You? Engineer Agent Memory Without Creating a Privacy Nightmare

13



Stop Building Chatbots: Build an AI Agent That Can Actually Do Work

WHAT YOU'LL LEARN — 14 TOPICS

01 Chatbot vs Copilot vs Agent

02 What Makes Software an AI Agent?

03 Brain + Tools + Memory + State + Permissions

04 Reason → Decide → Act → Observe

05 ReAct Explained Through a Real Task

06 When Should AI Be Allowed to Act?

07 Deterministic Workflow vs Autonomous Agent

08 Agent State

09 Agent Sessions

10 Tool Selection

11 Human Approval Gates

12 Error Recovery

13 Agent Contracts

14 Autonomy Levels

⌘ CODE — Build the TAS Admissions Agent

14  Who Is in Charge? Master the Agent Orchestration Patterns Enterprises Use

WHAT YOU'LL LEARN — 14 TOPICS

01 One Agent vs Many Agents

02 Supervisor–Router Pattern

03 Planner–Executor Pattern

04 ReAct Pattern

05 Sequential Agents

06 Parallel Agents

07 Hierarchical Agent Teams

08 Evaluator–Optimizer Pattern

09 Reflection Patterns

10 Agent Handoffs

11 State Machines

12 LangGraph-Style Graph Orchestration

13 Checkpointing

14 Recovering a Failed Agent Workflow

ARCHITECTURE LAB — Choose the right pattern for banking, telecom & education

15



Build Your AI Workforce: Supervisor + Specialist Multi-Agent Systems

WHAT YOU'LL LEARN

14 TOPICS

01 Why Enterprises Need Specialist Agents

02 Design the Supervisor Agent

03 Admissions Agent

04 IT Support Agent

05 Attendance Agent

06 Finance Agent

07 Knowledge Agent

08 Tool-Calling Agent

09 Evaluator Agent

10 Security/Policy Agent

11 How Does the Supervisor Know Who to Call?

12 Passing Context Without Leaking Everything

13 Shared vs Isolated Memory

14 Resolve Agent Conflicts

⌘ CODE — Complete TAS multi-agent team

16 Give Your Agents Hands: APIs, Tools & Enterprise Actions

WHAT YOU'LL LEARN — 14 TOPICS

01 An Agent Without Tools Can Talk—But Can't Work

02 Function Calling End-to-End

03 Design Reliable Tool Schemas

04 JSON Contracts

05 API Tools

06 Database Tools

07 CRM Tools

08 SIS/LMS Tools

09 Authentication for Tools

10 Tool-Level Authorization

11 Validate Arguments Before Execution

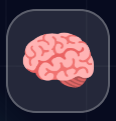
12 Prevent Duplicate Transactions

13 Tool Timeout & Retry Strategy

14 Stop the Agent from Calling Dangerous Tools

CODE — Build tools/ + admissions-status action

17



Meet AgentCore: From AI Prototype to Enterprise Agent Platform

WHAT YOU'LL LEARN

14 TOPICS

01 **Why a Working Agent Isn't Yet a Production Agent**

02 **Where AgentCore Fits in the Architecture**

03 **Runtime**
Where does your agent actually execute?

04 **Gateway**
How do agents reach enterprise tools?

05 **Identity**
Who is the agent acting as?

06 **Memory**
What should the agent remember?

07 **Observability**
What did the agent actually do?

08 **Bedrock + AgentCore Architecture**

09 **AgentCore + Multi-Agent Systems**

10 **Session Lifecycle**

11 **Scaling Agent Workloads**

12 **Failure Boundaries**

13 **Multi-Tenant Architecture**

14 **Prototype vs Production Architecture Comparison**

BUILD 09-agentcore-design.md

18 ⚡ Put Your Agent to Work: AgentCore Runtime & Gateway Deep Dive

WHAT YOU'LL LEARN — 14 TOPICS

01 From Python Agent to Managed Runtime

02 Runtime Packaging

03 Session Execution

04 Request Lifecycle

05 Gateway Architecture

06 Expose Enterprise Tools Safely

07 Authentication at the Gateway

08 Authorization at the Tool Boundary

09 Request Routing

10 Scaling Agent Runtime

11 Health Checks

12 Timeouts

13 Retries

14 Runtime Failure Debugging

⌘ CODE — Deploy the TAS agent runtime

19 "Who Are You Acting As?" Solve the Agent Identity Problem

WHAT YOU'LL LEARN — 14 TOPICS

01 Why Agent Identity Is Bigger Than Login

02 User vs Agent vs Workload Identity

03 Authentication vs Authorization

04 Delegated Access

05 Identity Propagation Across Agents

06 OAuth Concepts for Agentic Systems

07 Token Lifecycle

08 Short-Lived Credentials

09 Tool-Level Permissions

10 Least-Privilege Agents

11 Can the Admissions Agent See Finance Data?

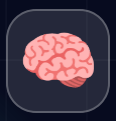
12 Multi-Agent Identity Isolation

13 Auditing "Who Did What?"

14 Salesforce/SIS Identity Integration

SECURITY LAB — Stop a privilege-escalation scenario

20



Should Your AI Remember You? Engineer Agent Memory Without Creating a Privacy Nightmare

WHAT YOU'LL LEARN

14 TOPICS

01 **Conversation History Is Not the Same as Memory**

02 **Short-Term Memory**

03 **Long-Term Memory**

04 **Semantic Memory**

05 **Episodic Memory Concepts**

06 **What Is Worth Remembering?**

07 **Memory Retrieval**

08 **Memory Relevance**

09 **Memory Isolation**

10 **Student A Must Never See Student B's Memory**

11 **Retention Policies**

12 **Consent & Personalization**

13 **Delete/Expire Memory Correctly**

14 **Memory Poisoning Risks**

⌘ CODE — Student Hub memory layer

TORONTO AI SCHOOL

PHASE 4 OF 6 · MODULES 21–24

Security & Governance

Put boundaries around intelligence: guardrails, red teaming, privacy-by-design and approval models.

4	56	4
MODULES	TOPICS	HANDS-ON BUILDS & LABS

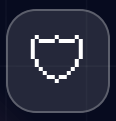
21 🛡 Put Boundaries Around Intelligence: Enterprise AI Guardrails

22 ⚔ Attack Your Own AI Before Someone Else Does: Prompt Injection & Red Teaming

23 🔏 Your LLM Doesn't Need Every Piece of Customer Data: Privacy-by-Design

24 🏛 Who Approves the AI? Build an Enterprise AI Governance Operating Model

21



Put Boundaries Around Intelligence: Enterprise AI Guardrails

WHAT YOU'LL LEARN — 14 TOPICS

01 **What Should Your AI Refuse?**

02 **Input vs Output Controls**

03 **Sensitive-Information Protection**

04 **PII Handling**

05 **Denied Topics**

06 **Business-Specific Policies**

07 **Grounding Controls**

08 **Retrieval Guardrails**

09 **Tool Guardrails**

10 **Prevent Unsafe Actions**

11 **False Positives**
When Guardrails Become Too Aggressive

12 **False Negatives**
When Bad Content Gets Through

13 **Measure Guardrail Trigger Rate**

14 **Test Guardrails Before Production**

⌘ CODE — Guardrail implementation + tests

22 Attack Your Own AI Before Someone Else Does: Prompt Injection & Red Teaming

WHAT YOU'LL LEARN — 14 TOPICS

01 "Ignore All Previous Instructions" —What Actually Happens?

02 Direct Prompt Injection

03 Indirect Prompt Injection

04 Jailbreaks

05 Malicious PDFs & Web Content

06 RAG Poisoning

07 Tool Manipulation

08 Data Exfiltration Attempts

09 System Prompt Leakage

10 Privilege Escalation

11 Confused-Deputy Attacks

12 Trust External Content as Data, Not Instructions

13 Tool Allowlists

14 Adversarial Evaluation Suite

RED TEAM LAB — Break—and then secure—the Student Hub

23  Your LLM Doesn't Need Every Piece of Customer Data: Privacy-by-Design

WHAT YOU'LL LEARN — 14 TOPICS

01 Minimum Necessary Data Principle

02 What Data Should Never Reach the Prompt?

03 PII Identification

04 Data Minimization

05 Masking & Redaction

06 Retention

07 Data Residency

08 Consent

09 Role-Based Access

10 Logging Without Leaking Sensitive Data

11 Human Oversight

12 Transparency

13 Auditability

14 Privacy Incident Response

BUILD — Enterprise privacy architecture

24 Who Approves the AI? Build an Enterprise AI Governance Operating Model

WHAT YOU'LL LEARN — 14 TOPICS

01 From AI Experiment to Governed Enterprise Product

02 AI Use-Case Inventory

03 Model Inventory

04 Agent Inventory

05 Tool Inventory

06 Risk Classification

07 Model Cards

08 System Cards

09 Data Lineage

10 Prompt Version Governance

11 Agent Version Governance

12 Approval Workflows

13 Architecture Review Board

14 Exception Management

BUILD — AI production-approval checklist

TORONTO AI SCHOOL

PHASE 5 OF 6 · MODULES 25–32

Production Engineering

Make it durable, fast, reliable, observable, evaluated and affordable — for years, not demos.

8	112	8
MODULES	TOPICS	HANDS-ON BUILDS & LABS

25	🌐	Build the Front Door to Your AI: Production API Architecture
26	⏳	What If an AI Task Takes 20 Minutes? Build Durable Agent Workflows
27	⚡	Make AI Faster—and Cheaper: The Enterprise Caching Playbook
28	💥	Your Model Is Throttled at 2 AM—Now What? Reliability Engineering for AI
29	📊	Can You Prove Your Agent Is Good? Enterprise Agent Evaluation
30	🔭	See Inside the AI Black Box: Observability from User Click to Agent Action
31	♻	AI Changes Every Week—How Do You Operate It for Years? LLMOps & AgentOps
32	💰	Your AI Works—But Can the Company Afford It? AI FinOps & Performance

25 Build the Front Door to Your AI: Production API Architecture

WHAT YOU'LL LEARN — 14 TOPICS

01 How Applications Actually Talk to AI

02 REST Architecture

03 API Gateway + Lambda

04 Request/Response Schemas

05 Authentication

06 Authorization

07 Input Validation

08 API Versioning

09 Rate Limiting

10 Throttling

11 Idempotency

12 Correlation IDs

13 Standard Error Contracts

14 Async APIs & Webhooks

CODE — Production Student Hub API

26 ⏳ What If an AI Task Takes 20 Minutes? Build Durable Agent Workflows

WHAT YOU'LL LEARN — 14 TOPICS

01 Why Not Every Agent Task Fits One API Request

02 Synchronous vs Asynchronous Processing

03 Step Functions

04 EventBridge

05 SQS

06 SNS

07 Long-Running Agent Jobs

08 Human Approval Workflows

09 Retry Policies

10 Dead-Letter Queues

11 Compensating Actions

12 Orchestration vs Choreography

13 Durable Execution Concepts

14 Workflow Recovery

⌘ CODE — End-to-end admissions workflow

27 ⚡ Make AI Faster—and Cheaper: The Enterprise Caching Playbook

WHAT YOU'LL LEARN — 14 TOPICS

01 Why Ask the LLM the Same Question Twice?

02 Response Cache

03 Semantic Cache

04 Embedding Cache

05 Retrieval Cache

06 Vector Search Cache

07 Tool-Result Cache

08 Session Cache

09 Prompt-Caching Concepts

10 Cache Keys

11 TTL Strategy

12 Cache Invalidation
The Hard Part

13 Never Leak Cached Answers Across Users

14 Measure Cache-Hit Rate & Savings

⌘ CODE cache_service.py

28



Your Model Is Throttled at 2 AM—Now What? Reliability Engineering for AI

WHAT YOU'LL LEARN — 14 TOPICS

01 Design for Failure, Not Hope

02 SLI vs SLO vs SLA

03 Availability Targets

04 Latency Budgets

05 Timeout Strategy

06 Retries Without Creating Retry Storms

07 Exponential Backoff + Jitter

08 Circuit Breakers

09 Bulkheads

10 Rate Limits

11 Bedrock Throttling

12 Fallback Models

13 Graceful Degradation

14 DLQs & Recovery

⚡ CHAOS LAB — Break production safely

29 Can You Prove Your Agent Is Good? Enterprise Agent Evaluation

WHAT YOU'LL LEARN — 14 TOPICS

01 "Looks Good to Me" Is Not an Evaluation Strategy

02 Routing Accuracy

03 Task-Completion Rate

04 Tool Success Rate

05 Groundedness

06 Correctness

07 Citation Quality

08 Policy Compliance

09 Hallucination Rate

10 Escalation Accuracy

11 Latency

12 Cost per Successful Task

13 Golden Test Sets

14 Human vs Automated Evaluation

CODE — Enterprise agent evaluation harness

30



See Inside the AI Black Box: Observability from User Click to Agent Action

WHAT YOU'LL LEARN — 14 TOPICS

01 Logs vs Metrics vs Traces

02 Follow One Request End-to-End

03 Correlation IDs

04 Model Latency

05 Token Usage

06 Retrieval Latency

07 Agent Routing Traces

08 Tool Execution Traces

09 Errors & Exceptions

10 Throttling Metrics

11 CloudWatch Dashboards

12 CloudTrail Auditing

13 Alerts

14 Production Debugging

BUILD — Complete AI observability dashboard

31



AI Changes Every Week—How Do You Operate It for Years? LLMOps & AgentOps

WHAT YOU'LL LEARN — 14 TOPICS

01 Version Everything That Can Change

02 Prompt Versioning

03 Model Versioning

04 RAG Versioning

05 Agent Versioning

06 Evaluation Dataset Versioning

07 Dev → QA → Stage → Production

08 Release Gates

09 Feature Flags

10 Canary Releases

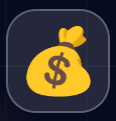
11 Blue/Green Deployments

12 Rollback

13 Continuous Evaluation

14 Production Feedback Loops

BUILD — Enterprise AgentOps lifecycle

32  Your AI Works—But Can the Company Afford It? AI FinOps & Performance

WHAT YOU'LL LEARN — 14 TOPICS

01 Calculate the Cost of One AI Request

02 Input vs Output Token Economics

03 Cost per Conversation

04 Cost per Successful Agent Task

05 Cheap Model vs Powerful Model Routing

06 RAG Cost

07 Embedding Cost

08 Vector Search Cost

09 Caching Savings

10 Concurrency

11 Throughput

12 Cold Starts

13 Autoscaling

14 Cost Budgets & Alerts

◎ CHALLENGE — Reduce AI cost without destroying quality

TORONTO AI SCHOOL

PHASE 6 OF 6 · MODULES 33–40

Enterprise Integration & Defense

Connect to real enterprise systems, automate the platform, launch it — and defend every decision.

8	112	8
MODULES	TOPICS	HANDS-ON BUILDS & LABS

33	☎	Turn the Contact Centre into an AI Workforce: Amazon Connect + Agentic AI
34	🎙	Beyond Chat: Build Voice & Multimodal AI Experiences
35	🔌	AI Can't Live on an Island: Connect Agents to Real Enterprise Systems
36	🏗	Stop Clicking in the Console: Build Your Entire AI Platform with Terraform
37	🚢	From Git Commit to Production AI: Build the CI/CD Factory
38	🚨	Production Day: 10,000 Users Are Coming—Is Your AI Actually Ready?
39	👩‍💼	Face the Architecture Board: Can You Defend Every Decision You Made?
40	🏆	The CEAAP Enterprise Build: From Empty Folder to Production-Grade Agentic AI Platform

33 Turn the Contact Centre into an AI Workforce: Amazon Connect + Agentic AI

WHAT YOU'LL LEARN — 14 TOPICS

01 Student Calls—What Happens Next?

02 Amazon Connect Architecture

03 Contact Flows

04 Queues & Routing

05 Agent Workspace

06 Customer Profiles

07 Cases

08 Lambda Integration

09 Connect → Bedrock → AgentCore

10 RAG During Live Support

11 Agentic Self-Service

12 Human Escalation

13 Containment Rate

14 FCR, CSAT & Abandon Rate

BUILD TAS AI-powered student contact centre

34 Beyond Chat: Build Voice & Multimodal AI Experiences

WHAT YOU'LL LEARN — 14 TOPICS

01 How Does a Voice Agent Actually Work?

02 Speech → Text → Intelligence → Speech

03 Turn-Taking

04 Interruptions

05 Latency

06 Transcripts

07 Voice Privacy

08 Authentication Challenges

09 Multilingual AI

10 Image + Text Inputs

11 Document Understanding

12 Multimodal RAG

13 Accessibility

14 Voice-to-Human Escalation

BUILD Multimodal Student Support experience

35 AI Can't Live on an Island: Connect Agents to Real Enterprise Systems

WHAT YOU'LL LEARN — 14 TOPICS

01 System of Record vs System of Intelligence

02 Salesforce Integration

03 Student Information System Integration

04 LMS Integration

05 ERP Integration

06 CRM Integration

07 Databases

08 REST APIs

09 Events

10 Read vs Write Permissions

11 Canonical Data Models

12 Transactions

13 Data Quality

14 Reconciliation

BUILD Complete Student Hub enterprise integration layer

36



Stop Clicking in the Console: Build Your Entire AI Platform with Terraform

WHAT YOU'LL LEARN — 14 TOPICS

01 **Infrastructure as Code for AI Architects**

02 **Terraform Project Structure**

03 **Providers**

04 **Variables**

05 **Outputs**

06 **Reusable Modules**

07 **IAM as Code**

08 **S3 + KMS**

09 **DynamoDB**

10 **OpenSearch**

11 **Bedrock Infrastructure**

12 **Lambda**

13 **API Gateway**

14 **Environment Management**

⌘ CODE — main.tf → complete enterprise infrastructure

37



From Git Commit to Production AI: Build the CI/CD Factory

WHAT YOU'LL LEARN — 14 TOPICS

01 What Happens After a Developer Pushes Code?

02 Git Strategy

03 Pull Requests

04 Code Reviews

05 Unit Tests

06 Integration Tests

07 RAG Evaluation Gates

08 Agent Evaluation Gates

09 Security Scanning

10 Terraform Validation

11 Build Artifacts

12 Environment Promotion

13 Production Approvals

14 Automated Deployment

BUILD — Full AI CI/CD pipeline

38 🚨 Production Day: 10,000 Users Are Coming—Is Your AI Actually Ready?

WHAT YOU'LL LEARN — 14 TOPICS

01 **Production Readiness Review**

02 **Load Testing**

03 **Security Testing**

04 **Red-Team Results**

05 **Evaluation Thresholds**

06 **Deployment Approval**

07 **Smoke Tests**

08 **Dashboards**

09 **Alerts**

10 **On-Call Operations**

11 **Incident Response**

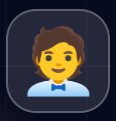
12 **Runbooks**

13 **Backup & Recovery**

14 **Postmortems**

▶ SIMULATION — Production launch day

39



Face the Architecture Board: Can You Defend Every Decision You Made?

WHAT YOU'LL LEARN — 14 TOPICS

01 Present the Business Problem in 3 Minutes

02 Defend Your Architecture

03 Why Bedrock?

04 Why RAG?

05 Why This Vector Store?

06 Why Multi-Agent?

07 Why AgentCore?

08 Defend Identity & Memory Decisions

09 Defend Security Boundaries

10 Explain Your Threat Model

11 Prove Your Evaluation Numbers

12 Defend Reliability

13 Explain Cost & ROI

14 Handle CTO/CISO Questions

SIMULATION — Enterprise Architecture Review Board

40  The CEAAP Enterprise Build: From Empty Folder to Production-Grade Agentic AI Platform

WHAT YOU'LL LEARN — 14 TOPICS

01 **Start with the Business Case**

02 **Walk Through Requirements**

03 **Present the Architecture**

04 **Demonstrate Advanced RAG**

05 **Demonstrate Multi-Agent Orchestration**

06 **Demonstrate AgentCore**

07 **Show Identity & Memory**

08 **Execute Real Tools**

09 **Trigger Guardrails**

10 **Show an Attack—and the Defense**

11 **Show RAG & Agent Evaluation**

12 **Show Observability**

13 **Deploy with Terraform + CI/CD**

14 **Present Cost, Reliability & Security**

♛ FINAL DEFENSE — Present the system like an Enterprise AI Architect

TORONTO
AI SCHOOL

Become a Certified Enterprise AI Architect & Agentic Systems Professional

Forty modules. One production-grade capstone. From your first architecture diagram to defending a live agentic AI platform in front of a review board — CEAAP™ is built for architects who ship.

AMAZON BEDROCK | AGENTCORE | MULTI-AGENT | RAG

info@torontoaischool.com • torontoaischool.com