RAVINEL FIELD GUIDE · NO. 01

The Language of AI.

Thirty terms every operator should know.
Five minutes to read.
A lifetime of clarity.

Read in five minutes. Refer for years.

www.ravinel.ai · info@ravinel.ai

HOW TO USE THIS GUIDE

A field guide, not a glossary.

Each term explains what it is in plain language, and what it means for the person who has to make a decision about it.

Thirty terms across six clusters. Test yourself at the end.

Read straight through in about five minutes. Come back to it whenever a term shows up in a meeting and you need a clean, honest definition.

SIX CLUSTERS

01 The Foundations
02 How It Learns
03 How It Talks
04 How It Behaves
05 When It Acts
06 How It's Governed

THE FOUNDATIONS

The Foundations.

What AI actually is.

THE FOUNDATIONS · 01 OF 05

Artificial Intelligence

Systems that perform tasks that would require human intelligence. An umbrella term covering everything else in this guide.

FOR AN OPERATOR

The word AI alone tells you nothing about capability, cost, or risk. Always ask which type.

THE FOUNDATIONS · 02 OF 05

Machine Learning

A branch of AI where systems learn patterns from data instead of being explicitly programmed.

FOR AN OPERATOR

ML is what powers most predictive AI in your organisation today.

THE FOUNDATIONS · 03 OF 05

Deep Learning

Machine learning based on layered neural networks. What made modern AI possible.

FOR AN OPERATOR

When someone says we use deep learning, ask what task and what data they trained on.

THE FOUNDATIONS · 04 OF 05

Foundation Model

A large AI model trained on broad data that can be adapted to many downstream tasks.

FOR AN OPERATOR

The economics changed when you stopped building custom models and started renting foundation ones.

THE FOUNDATIONS · 05 OF 05

Predictive vs Generative AI

Predictive AI forecasts a value or a class. Generative AI creates content.

FOR AN OPERATOR

Different data needs, different KPIs, different failure modes. Do not manage them as one budget line.

HOW IT LEARNS

How It Learns.

The process behind the model.

HOW IT LEARNS · 01 OF 05

Training

Feeding data to a model so it learns statistical patterns.

FOR AN OPERATOR

Training is expensive and slow. Most enterprises use pre-trained models rather than training their own.

HOW IT LEARNS · 02 OF 05

Fine-tuning

Additional training on your own data to specialise a general model.

FOR AN OPERATOR

Fine-tuning is not always needed. Try prompting and RAG first. They are cheaper and often good enough.

HOW IT LEARNS · 03 OF 05

RAG

Retrieval-Augmented Generation. A pattern where the model retrieves relevant information at query time and uses it to answer.

FOR AN OPERATOR

RAG is how you make a generic model speak your organisation's language without retraining.

HOW IT LEARNS · 04 OF 05

Embeddings

Numerical representations of text that capture meaning. Two similar sentences produce two similar embeddings.

FOR AN OPERATOR

Embeddings are the plumbing behind semantic search, RAG, and recommendation.

HOW IT LEARNS · 05 OF 05

Vector Database

A database optimised for storing and searching embeddings.

FOR AN OPERATOR

If you are building RAG, you will need one. Not all enterprise data platforms handle this well.

HOW IT TALKS

How It Talks.

The mechanics of language models.

HOW IT TALKS · 01 OF 05

LLM

Large Language Model. A model trained to predict the next word, or token, given a context of previous words.

FOR AN OPERATOR

LLMs are pattern-matchers, not reasoners. Their fluency can mask that.

HOW IT TALKS · 02 OF 05

Token

The unit an LLM reads and writes. Roughly three-quarters of an English word.

FOR AN OPERATOR

You pay per token. Long prompts and long answers are both cost drivers.

HOW IT TALKS · 03 OF 05

Context Window

The maximum amount of text a model can consider at once.

FOR AN OPERATOR

Bigger context windows unlock new use cases like whole-document analysis, but they raise cost and latency.

HOW IT TALKS · 04 OF 05

Prompt

The input you send to a model.

FOR AN OPERATOR

The prompt is the primary product surface of most enterprise AI. Design it deliberately.

HOW IT TALKS · 05 OF 05

Prompt Engineering

The practice of crafting prompts to get reliable results.

FOR AN OPERATOR

Prompt engineering is real work. It is what happens between buying the licence and the model actually helping.

HOW IT BEHAVES

How It Behaves.

What to expect, what to guard against.

HOW IT BEHAVES · 01 OF 05

Inference

When a trained model is used to generate an output. Every user query is an inference.

FOR AN OPERATOR

Inference cost is your ongoing operating expense. Training is a one-off. Inference never stops.

HOW IT BEHAVES · 02 OF 05

Hallucination

When a model produces plausible-sounding but factually wrong output.

FOR AN OPERATOR

Hallucination risk is highest when the model has no source to check. Give it a source and the risk drops sharply.

HOW IT BEHAVES · 03 OF 05

Grounding

Techniques that force the model to answer only from a specific set of sources.

FOR AN OPERATOR

Grounding is your primary defence against hallucination in enterprise use cases.

HOW IT BEHAVES · 04 OF 05

Guardrails

Rules that constrain what a model can say or do.

FOR AN OPERATOR

Guardrails are the difference between a demo and a production system. Do not launch without them.

HOW IT BEHAVES · 05 OF 05

Latency

The time between sending a prompt and receiving an output.

FOR AN OPERATOR

Latency matters. A 4-second answer is fine for research. Unacceptable for a call-centre agent.

WHEN IT ACTS

When It Acts.

From assistant to agent.

WHEN IT ACTS · 01 OF 05

Agent

A system that uses an LLM to plan, decide, and take actions in a loop.

FOR AN OPERATOR

Agents are AI that does things, not just AI that says things. That is why they need different governance.

WHEN IT ACTS · 02 OF 05

Tool Use

When a model can call external functions like APIs, databases, or code as part of its reasoning.

FOR AN OPERATOR

Tool use is what makes agents useful and dangerous. Every tool is a potential attack surface.

WHEN IT ACTS · 03 OF 05

Autonomy Levels

Assist, augment, autopilot, autonomous. A spectrum of how much control the model has.

FOR AN OPERATOR

Most enterprise agents should live at augment. Autopilot needs explicit governance. Autonomous needs board approval.

WHEN IT ACTS · 04 OF 05

Orchestration

Coordinating multiple models, tools, and agents to complete a complex workflow.

FOR AN OPERATOR

Orchestration frameworks are where AI infrastructure is consolidating fast. Expect this vocabulary to change monthly.

WHEN IT ACTS · 05 OF 05

MCP

Model Context Protocol. An open standard that lets models connect to tools and data sources in a consistent way.

FOR AN OPERATOR

MCP is emerging as the USB standard for AI integration. Watch which of your vendors support it.

HOW IT'S GOVERNED

How It's Governed.

Keeping AI safe to use.

HOW IT'S GOVERNED · 01 OF 05

Model Risk

The risk that a model produces incorrect, biased, or harmful outputs.

FOR AN OPERATOR

Model risk is a governance discipline that pre-dates AI. It just now applies to every function.

HOW IT'S GOVERNED · 02 OF 05

Data Sovereignty

Legal control over where your data is processed and stored.

FOR AN OPERATOR

In MENA, data sovereignty is a regulatory expectation, not a preference. Verify where your AI vendor runs inference.

HOW IT'S GOVERNED · 03 OF 05

Human-in-the-loop

A design where a human reviews or approves model outputs before they take effect.

FOR AN OPERATOR

Human-in-the-loop is the safest way to deploy AI in regulated processes. It also caps your automation ceiling.

HOW IT'S GOVERNED · 04 OF 05

Evals

Automated tests that measure a model's or agent's quality on your specific tasks.

FOR AN OPERATOR

If you cannot evaluate an AI system on your own data, you do not know what you have deployed.

HOW IT'S GOVERNED · 05 OF 05

AI Governance

The framework of policies, roles, and controls that ensure AI is used responsibly and effectively.

FOR AN OPERATOR

AI governance is not compliance theatre. It is how you keep AI deployments alive past 12 months.

SECTION 07 · QUIZ

Test Yourself.

Ten questions. Answers on the following pages.

TEST YOURSELF · QUESTIONS 01 AND 02 OF 10

01 RAG stands for:

a Retrieval-Augmented Generation

b Randomised Agent Grading

c Recursive Alignment Governance

d Real-time API Gateway

02 A token in an LLM is closest to:

a A user account credential

b About three-quarters of an English word

c A cryptocurrency

d An API rate limit

TEST YOURSELF · QUESTIONS 03 AND 04 OF 10

03 The primary defence against hallucination in enterprise use cases is:

a Using bigger models

b Better-worded prompts

c Grounding the model in specific sources

d Adding a human reviewer

04 Fine-tuning a foundation model means:

a Making it faster

b Additional training on your own data

c Reducing its size

d Adding new UI features

TEST YOURSELF · QUESTIONS 05 AND 06 OF 10

05 Predictive AI and Generative AI should be governed:

a Under one unified AI policy

b Only if regulators require it

c Differently. Different data, KPIs, and failure modes

d The same as traditional analytics

06 In MENA, data sovereignty is:

a A best practice

b A regulatory expectation

c Optional if you use encryption

d Only relevant for financial services

TEST YOURSELF · QUESTIONS 07 AND 08 OF 10

07 The difference between an LLM and an agent is:

- **a** There is none
- **b** An LLM says things. An agent does things
- **c** An LLM is bigger than an agent
- **d** An agent is open source

08 Your ongoing operating cost, not a one-off, is:

- **a** Training
- **b** Fine-tuning
- **c** Inference
- **d** Prompt engineering

TEST YOURSELF · QUESTIONS 09 AND 10 OF 10

09 Autopilot as an autonomy level means:

a The model runs without governance

b The model executes routine decisions with human oversight on exceptions

c The model is banned in the EU

d The model has memory

10 Evals matter because:

a They reduce compute cost

b They are required by GDPR

c They tell you whether your AI system actually works on your tasks

d They speed up inference

SECTION 08 · ANSWERS

Answer Key.

How you did, and why each answer matters.

ANSWER KEY · ANSWERS 01 AND 02 OF 10

01 a) Retrieval-Augmented Generation.

WHY

RAG retrieves relevant sources at query time and grounds the model's answer in them. It is the single most effective anti-hallucination pattern in enterprise use.

02 b) About three-quarters of an English word.

WHY

Tokens are the unit models read and write. You pay per token, so verbosity in both prompt and answer costs money.

ANSWER KEY · ANSWERS 03 AND 04 OF 10

03 c) Grounding the model in specific sources.

WHY

Bigger models still hallucinate. Grounding, via RAG or provided context, is the primary defence.

04 b) Additional training on your own data.

WHY

Fine-tuning specialises a general model. Try prompting and RAG first. They are cheaper and often enough.

ANSWER KEY · ANSWERS 05 AND 06 OF 10

05 c) Differently.

WHY

Predictive AI and generative AI have different data needs, KPIs, and failure modes. Bundling them under one AI policy under-governs the riskier one.

06 b) A regulatory expectation.

WHY

MENA regulators, including CBUAE, expect binding governance around where data is processed, not encryption alone.

ANSWER KEY · ANSWERS 07 AND 08 OF 10

07 b) An LLM says things. An agent does things.

WHY

Agents plan, decide, and act using tools. That is why they need different governance from a chatbot.

08 c) Inference.

WHY

Training is a one-off cost. Every user query is an inference, so cost accumulates continuously.

ANSWER KEY · ANSWERS 09 AND 10 OF 10

09 b) The model executes routine decisions with human oversight on exceptions.

WHY

Autopilot removes humans from the routine path but keeps them in the loop for edge cases.

10 c) They tell you whether your AI system actually works on your tasks.

WHY

Without evals, you have no idea if the system is drifting, degrading, or hallucinating on your data.

THE RAVINEL DIFFERENCE

Start from the outcome. Not from the model.

Most AI advisory begins with the technology and hopes a business case appears. That is why most programs stall.

Ravinel begins with the P&L. The operating model. The customer outcome. The risk position. Only then do we choose the AI paradigm, the model, the deployment pattern, the governance framework.

BUSINESS OUTCOME FIRST.

AI STRATEGY SECOND.

SERVICES FOLLOW.

HOW WE DELIVER

Where Ravinel fits.

AI Readiness Audit

Where you stand across data, decision rights, governance, and fluency. So you know what to invest in and in what order.

AI Fluency Program

Modular training for board, SMT, and functional leaders. So decisions about AI get made by people who understand it.

AI Operating Model Design

Turning strategy into owned workflows, KPIs, and accountability lines. So AI reaches the P&L.

AI Governance Framework

Model risk, decision rights, and evals for regulated environments. So AI stays deployed past twelve months.

RAVINEL FIELD GUIDE · NO. 01

The language is the starting line.

The discipline is where the value lives.

Book a 30-minute AI Readiness conversation.

www.ravinel.ai · info@ravinel.ai

Volume 02 arrives next month. Follow Ravinel.AI on LinkedIn.