How AI Works and How to Work With It

GenAI Foundations

IDSC 4590: Agentic AI for Business Applications | Week 1, Session 1

Minnesota Carlson



About Your Instructor



PROFESSOR

Dr. De Liu

Xian Dong Eric Jing Professor, Information and Decision Sciences. IDS PhD Coordinator and former Academic Director, MS in Business Analytics

FOUNDER AND CEO

SmartPAL AI Inc.

AI that helps learners and instructors, and does not replace the learning

RESEARCH

Gen AI, Gamification, EduTech, Platform Economics, User Generated Content

Learning objectives

By the end of this session you will be able to:

1

Explain, in plain language, why the same prompt can give different outputs and why a fluent answer can still be wrong.

2

Distinguish rule-based, predictive ML, generative and agentic AI by what each needs and how each fails.

3

Rebuild a weak prompt using at least two of three techniques: spell out the context and standard, ground and check, and break it down with gates.

Today's Outline: Know the Tool Before You Trust It

- Most AI mistakes in business come from using a tool without knowing what it does
- You do not need the math behind AI. You do need a clear picture of how it works, so you can tell where it will go wrong

1 · HOW AI WORKS
Where it came from, how a model works

2 · HOW TO WORK WITH IT
Principles, loop, techniques

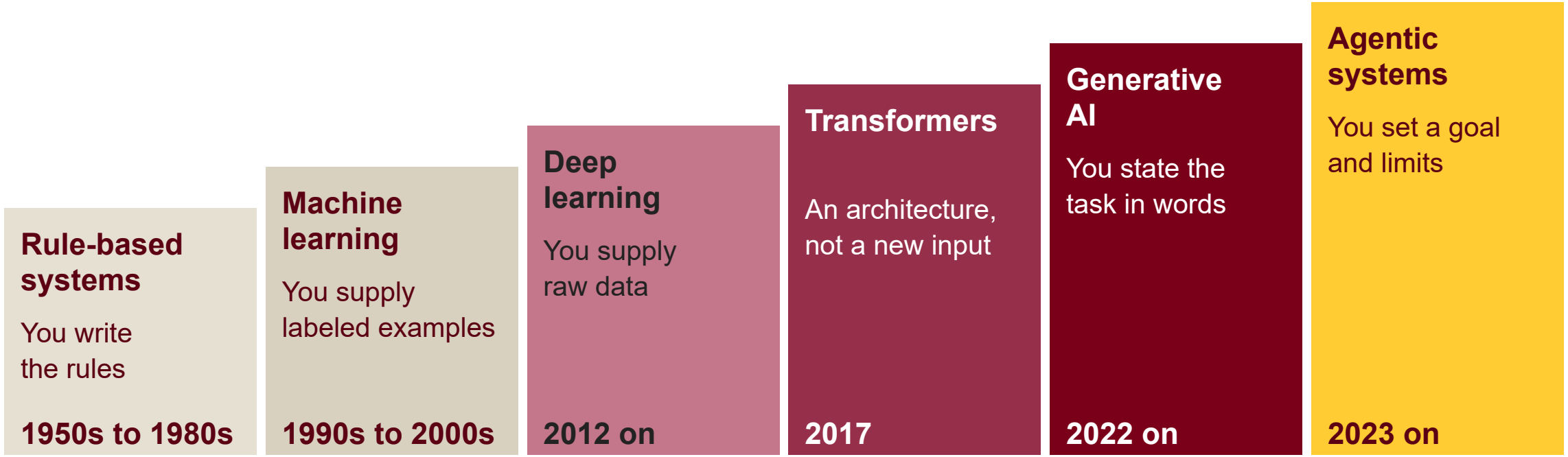
3 · PROMPT AUDIT
Apply it to a weak prompt

4 · COURSE ORIENTATION
Goals, learning, grading

How AI Works

Where AI came from, the four kinds of AI, how a language model works, and a live demo.

How We Got Here: Each Step Changed What You Supply to the System



Four Kinds of AI, Four Management Problems

RULE-BASED

Follows rules a person wrote. Fails on cases nobody wrote a rule for.

You manage the rule owner.

PREDICTIVE ML

Learns patterns from past data to score or forecast. Repeats the past, including its bias, and degrades when the world changes.

You manage the data and the model.

GENERATIVE AI

Produces new text, images or code from a prompt. Fluent and wrong; different answers each run.

You manage the person reviewing the output.

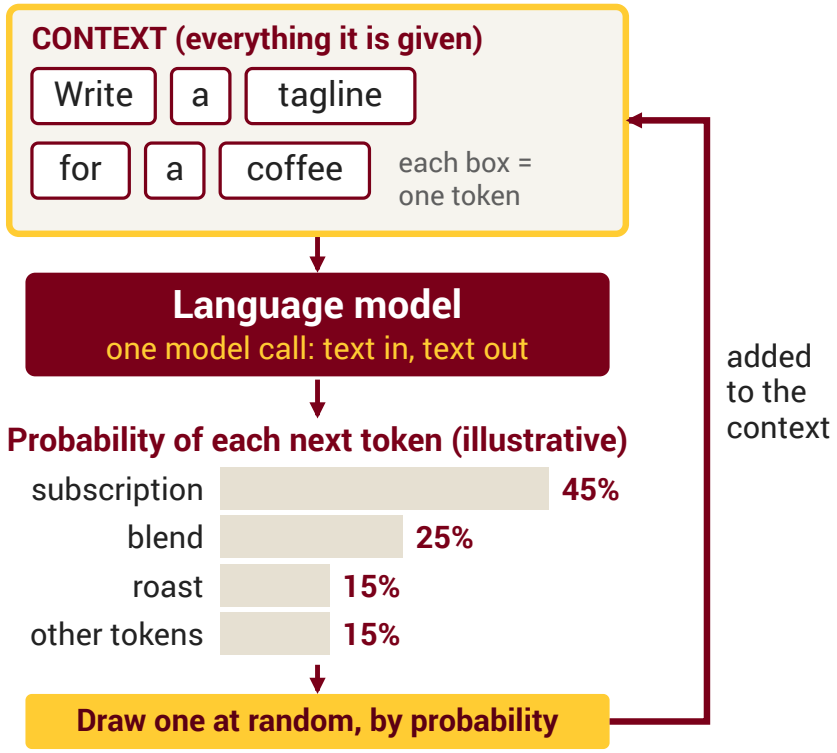
AGENTIC AI

Pursues a goal by calling a model repeatedly and using tools. Errors compound across steps.

You manage the permissions, limits and approvals around the loop.

Real products are often **several at once.**

What a Language Model Actually Does



- It reads the **context** and predicts the next **token** (about a word) at a time. Each **model call** is text in, text out: no lookup, no memory
- Each step, **every possible next token gets a probability**, then it **draws one at random** by those probabilities
- Likely tokens come up often, others sometimes. So the same input can give different outputs, and **fluent does not mean true**

Bonus: temperature tilts the draw. At 0, always the top token (nearly the same output every run). Higher: more variety.

Live Demo: Variation, Fluency and Grounding

1. **Variation:** the same tagline prompt, five runs, side by side. The answers differ.

2. **Fluent, not verified:** ask for the findings of a report from a company that does not exist, then ask again with search. A fluent answer is **not a checked answer.**

3. **Grounded:** paste a short excerpt and ask again. The answer now comes from the text, and the model says "not stated" for what the text lacks.

> Nothing is broken. This is how the tool works. The manager's question is **what to do about it.**

Three Insights to Use All Term

1

Outputs are sampled

Results vary, so you test with many runs and constrain the format.

Returns in Week 2, when workflows break if the format varies

2

Fluent is not verified

Check before you rely on it; ask for sources or reasoning.

Returns in today's techniques and in Week 6

3

Context is the only working memory

It is billed in tokens: the model forgets what is outside the context, and long contexts cost more.

Returns in Weeks 4 and 5, with agent memory and cost

From How It Works to How to Work With It

Now what you do about it: a way of working with AI.

Four Principles for Good Prompting

1

Clarify the task

Does the AI understand the goal and the standard?

2

Delegate deliberately

What should the AI do, and what should you do?

3

Make it checkable

How will you know if the answer is good?

4

Manage the human bottleneck

Can you actually process and supervise what the AI produces?

The Five-Step Loop

1. Define the question
2. Provide context
3. Delegate one bounded task
4. Check the result
5. Decide and record

YOU

Decide the standard

You decide the goal and what a good answer is, and supply background, data, constraints and an example.

THE AI

Does one bounded task

One subtask whose output you can inspect.

YOU

Check, decide, record

Check against the source, the data, a calculation or a test. Keep a short log of the prompt, the output and what you changed.

Technique 1: Spell Out the Context and Standard

Principle: clarify the task

WEAK

"Analyze this company's strategy."

STRONGER

"I am preparing a 5-minute brief for a regional sales director. Using the excerpt below, identify the two biggest strategic risks. A good answer names the risk, the evidence in the text, and one question to ask management. Format: a table with those three columns, max 6 rows."

Give the **context**, say what a good answer looks like, add one worked example, and fix the output format. Both Google and Anthropic recommend this.

Technique 2: Ground It and Check It

Principle: make it checkable

ASK

Ask me first

"Ask me clarification questions instead of making assumptions."

SURFACE

Assumptions and sources

"List every assumption you made. For each factual claim, name the part of the document it comes from."

FACTS

Search grounding

"Search for current information, and give a source link for each fact."

Trust what you can trace to a source, not how confident the answer sounds.

Technique 3: Break It Down, With Gates

Principle: delegate deliberately



> "First write a plan with numbered steps and wait. After I approve it, do step 1 only and show me the result."

- Each gate is a place where a person can **stop, redirect or approve**
- **Run in phases:** small test, inspect, revise, then scale up

From Lecture to Activity

Now apply the three techniques to a weak business prompt.

Activity: Prompt Audit

You get a weak business prompt and its real output. **Diagnose why it fails, then rebuild it** using at least two of the three techniques.

4 MIN Diagnose → **13 MIN** Rebuild and run twice → **4 MIN** Check one claim → **4 MIN** Document → **10 MIN** Share out

What is missing?

What is missing from the weak prompt: context, standard, grounding, check or gates?

What changed?

Use at least two of the three techniques. Run it twice and note what changed.

What did you check?

Verify one claim against a source. Add what was wrong, what you changed and what improved to the class doc. A few pairs present.

Prompt Audit Card

1. **Weak prompt and what was missing:** context, standard, grounding, check or gates
2. **Rebuilt prompt:** which two or more techniques did you use?
3. **Run twice:** what changed between runs?
4. **One claim checked:** what did you find at the source?
5. **What improved,** and what would you still want a person to verify?

What Changed, and What's Next

NEXT SESSION

Personal productivity with AI (chat, workspaces, connectors), then how companies use AI: RAG, governance and the enterprise.

TAKE-HOME BEFORE NOV 4

A personal AI experiment, and Reflection Post 1.

BEFORE MONDAY

Nothing to prepare. Bring a laptop and one task from your own life or work.

Course Orientation

How the course is organized: what it is for, how you will learn, and how it is graded.

The Goal: Informed, Confident AI Fluency

UNDERSTAND

Know the tool

Explain how AI works and why it errs. Pick the right kind of AI for a problem.

BUILD

Build with no code

Design and build simple AI workflows.

JUDGE

Evaluate and decide

Evaluate agentic AI: how it fails, how much autonomy to give. Weigh risk, ethics and vendors.

PITCH

Pitch a solution

Scope and pitch an AI solution to a real business problem.

For people who will **manage AI tools, commission AI projects and decide about AI adoption**, not build AI from scratch. No coding required.

How You Will Learn: From Class to Capstone

1. **Class sessions** (lectures, demos, activities): the week's topic, with something to try each time.
2. **Build sprints** (three): the week's skill, by doing. Make, n8n, then a red team.
3. **Reflection posts** (four): your own thinking. What you did, how you checked it, what you would change.
4. **Weekly check-ins** (six): the course itself. What worked for you and what would help.
5. **Team capstone:** everything together, pitched at the Showcase on Dec 14. Teams form Nov 16.

Grading, Late Work and AI Use

GRADING

What counts

Capstone 25%.
Final 20%.
Reflections 20%.
Midterm 15%.
Sprints 15%.
Check-ins 5%.

WHERE

Canvas and Chat

Everything due is in Canvas. Questions go in the Google Chat space.

LATE WORK

Five flex days

Defer a reflection or check-in up to 2 days each. Sprints, Showcase and exams are fixed.

AI USE

You stay answerable

Reflections are your own thinking. Exam Part A is closed-book; Part B allows AI.

I used AI to help build this course, and **I own any errors.** If you spot one, tell me.

Terms

Not presented in class. These are plain-language explanations, not rigorous definitions. Use them to recall what each term is about and what we talked about in class.

Terms: A–B

Agentic AI A system that works toward a goal instead of answering once. In a loop, it decides the next step, uses a tool (search, email, a spreadsheet), looks at the result and decides again, carrying what it has learned from step to step, until it finishes or hits a limit. Each step builds on the last, so one wrong step can carry through the rest.

Slide 7: Four Kinds of AI, Four Management Problems

Bounded task A small, well-defined piece of work whose output you can inspect and judge. *Example: "list the two biggest risks in this excerpt, with the supporting quote," not "analyze this company."*

Slide 12: The Five-Step Loop

Build sprint A short, time-boxed stretch of hands-on work aimed at one concrete result that you build, test and show. *In this course: a workflow in Make, an agent in n8n, and a red team of another team's agent.*

Slide 20: How You Will Learn: From Class to Capstone

Terms: C–F

Context Everything the model is given in a single call: your prompt, any text or documents you paste in, and any earlier messages that are sent along with it. It is the only information the model has to work from in that call, and it has a size limit, the *context window*, which works like the model's working memory.

Slide 8: What a Language Model Actually Does

Deep learning Machine learning that uses many-layered neural networks to learn its own features from raw data such as images, speech or text, instead of relying on features people design. It needs large amounts of data and computing power.

Slide 6: How We Got Here: Each Step Changed What You Supply to the System

Fluent answer An answer that reads smoothly and confidently. Fluency says nothing about accuracy: the model produces text that sounds right whether or not the claims have been checked. NIST calls confidently stated but false output *confabulation*, and notes that it follows from how generative models work.

Slide 9: Live Demo: Variation, Fluency and Grounding

Terms: G

Gate A checkpoint where a person reviews and can stop, redirect or approve before work continues.
Example: approving an AI's written plan before it carries out step 1.

Slide 15: Technique 3: Break It Down, With Gates

Generative AI A model that produces new text, images or code in response to a prompt, resembling the data it was trained on. A language model does this by predicting what plausibly comes next. Its output reads well whether or not it is correct, and it can differ from one run to the next.

Slide 7: Four Kinds of AI, Four Management Problems

Grounded Based on source text that is supplied with the question, such as a pasted excerpt, instead of on what the model absorbed in training. Each claim can then be traced back to the supplied text.

Slide 9: Live Demo: Variation, Fluency and Grounding

Terms: H–M

Human bottleneck The limit on how much AI output a person can actually read, verify and act on. The model can produce text far faster than a person can review it, so the person's capacity sets the pace.

Slide 11: Four Principles for Good Prompting

Machine learning Software that learns patterns from example data instead of following rules a person wrote, then applies them to new cases. *Examples: lead scores, fraud flags, demand forecasts.*

Slide 6: How We Got Here: Each Step Changed What You Supply to the System

Model call One request to a model: text goes in, text comes out. The model does not look anything up, and it remembers an earlier call only if that text is sent again as part of the new context.

Slide 8: What a Language Model Actually Does

Terms: P–R

Predictive ML A machine learning model trained on past data to score or forecast an outcome, such as which leads will buy or which transactions are fraudulent. It repeats the patterns of the past, including past bias, and its accuracy drops when the world changes.

Slide 7: Four Kinds of AI, Four Management Problems

Prompt audit A structured review of a prompt: diagnose what it lacks (context, a standard, grounding, a check or gates), rebuild it, run the new version more than once, and verify at least one claim in the output against a source.

Slide 16: Activity: Prompt Audit

Rule-based AI A system that decides by applying explicit if-then rules that people wrote. It is predictable and easy to audit, but it fails on any case nobody wrote a rule for. *Example: approve any refund under $50.*

Slide 7: Four Kinds of AI, Four Management Problems

Terms: S–T

Sampling At each step the model gives every possible next token a probability, then draws one at random, weighted by those probabilities. Likely tokens come up often, but not every time, so the same input can give different outputs.

Slide 8: What a Language Model Actually Does

Search grounding A feature that lets the model run a web search and base its answer on the results instead of on memory alone, which connects the answer to sources that can be checked. It helps with facts that change, such as prices or news, and the answer is only as reliable as the sources it finds.

Slide 14: Technique 2: Ground It and Check It

Temperature A setting that controls the randomness of the draw. Lower values favor the most likely token and give more consistent output; higher values give more variety. Even at 0, output is not guaranteed to be identical from call to call. For some newer models providers recommend leaving it at the default. Many chat apps do not show it.

Slide 8: What a Language Model Actually Does

Terms: T

Token The unit of text a language model reads and writes: a word, part of a word, a character or even a single byte, depending on the model's tokenizer. The same text can be a different number of tokens in different models and languages. Usage limits and cost are measured in tokens.

Slide 8: What a Language Model Actually Does

Transformer The neural-network design introduced in 2017 ("Attention Is All You Need", Vaswani et al.) behind today's chat models. It is built on *attention*: to make sense of each word, the model weighs how much every other word in the context matters, all at once. In "The bank approved the loan," the words "approved" and "loan" tell it which meaning of "bank" is meant. Because it processes the whole context in parallel instead of one word at a time, it trains faster and could be trained on far more text than earlier designs.

Slide 6: How We Got Here: Each Step Changed What You Supply to the System