AI in Society and Public Services

Session 1: AI Fundamentals & Context

Mário Antunes

January 23, 2026

Universidade de Aveiro

Table of Contents i

AI in Society and Public Services

Session 1: AI Fundamentals & Context

Duration: 3 Hours

Instructor: Mário Antunes

Part 1: Course Introduction

Course Objectives

- **Demystify AI:** Move beyond the “magic” to understand the mechanics.
- **Hands-on Experience:** Gain practical intuition through examples (Notebooks).
- **Critical Thinking:** Learn to evaluate when (and how) to use AI and when *not* to.
- **Future Proofing:** Understand the trajectory from current LLMs to future agentic systems.

Part 2: A Brief History of Intelligence

“Can Machines Think?”

The Genesis (1950s)

- **Alan Turing (1950):** Published *“Computing Machinery and Intelligence”*.
- **The Imitation Game:** Proposed a test where a machine tries to pass as a human in a text-based conversation.
- **Dartmouth Conference (1956):** The term “Artificial Intelligence” is coined.
    - *Optimism:* “Every aspect of learning... can be so precisely described that a machine can be made to simulate it.”

Symbolic AI (1950s - 1980s)

"Good Old-Fashioned AI" (GOFAI)

- **The Logic:** AI as a set of logical rules and symbols.
- **If-Then Systems:**
  - *Example:* Expert Systems (Medical diagnosis based on strict symptom rules).
- **Successes:** Chess (Deep Blue), Algebra word problems.
- **The Failure:** The "Polanyi's Paradox"—we know more than we can tell.
  - *Hard to explain:* How do you define a "cat" using only if-then rules? (It has ears? So does a dog.)

The AI Winters

Cycles of Hype and Disillusionment

1. **First Winter (1974–1980):** Funding cuts due to lack of progress in translation and perception.
2. **Second Winter (1987–1993):** Collapse of the Lisp machine market and failure of Expert Systems to scale.

- **Lesson:** Overpromising leads to funding droughts.
- **Requirement:** We needed more data and more compute.

The Machine Learning Paradigm Shift

From “Programming” to “Learning”

- **Traditional Programming:**
  - Data + Rules = **Output**
- **Machine Learning:**
  - Data + Output = **Rules**
- **Concept:** Instead of writing the code, we feed the computer examples, and it figures out the statistical correlation.

The Machine Learning Landscape i

Taxonomy by Architecture (Depth)

- **Shallow Learning (Classic ML):**
  - Simple models (e.g., Linear Regression, Decision Trees, SVM).
  - Best for structured/tabular data (Excel files).
  - Requires manual “Feature Engineering” (human tells AI what to look for).
- **Deep Learning:**
  - Neural Networks with many layers (e.g., CNNs, Transformers).
  - Best for unstructured data (Images, Audio, Text).
  - **Feature Learning:** Automatically discovers useful patterns.

The Machine Learning Landscape ii

Taxonomy by Supervision (Method)

1. **Supervised Learning:** Learning with a teacher (Labeled Data).
   - *Task:* "Here is an image, it is a Cat."
2. **Unsupervised Learning:** Discovery (Unlabeled Data).
   - *Task:* "Here are piles of photos, sort them by similarity."
3. **Reinforcement Learning:** Trial and Error (Reward System).
   - *Task:* "Play this game; +1 point for winning, -1 for losing."

Deep Learning & The Big Bang (2012)

- **Neural Networks:** Inspired by the human brain (neurons and synapses).
- **The Catalyst:** The ImageNet Competition (2012).
  - **AlexNet:** A Deep Convolutional Neural Network (CNN) crushed the competition.
- **The Ingredients:**
  1. **Big Data:** The internet provided millions of labeled images.
  2. **GPUs:** NVIDIA hardware allowed massive parallel processing.
  3. **Algorithms:** Backpropagation allowed networks to learn from mistakes efficiently.

The Era of Generative AI (2017 - Present)

- **The Transformer Architecture (2017):**
  - Paper: *"Attention Is All You Need"* (Google).
  - Allowed models to process massive amounts of text in parallel (unlike previous RNNs).
- **LLMs (Large Language Models):**
  - GPT-1 to GPT-4, Claude, Llama.
  - **Scaling Laws:** More data + more parameters = emerging reasoning capabilities.

Part 3: Capabilities & Limitations

Defining the Terms: It's Not All "Smart"

The Hierarchy of Cognition

1. **Intelligence (The "Brain"):**
    - The ability to acquire knowledge, solve problems, and achieve goals.
    - *Example:* A calculator is intelligent; AlphaGo is highly intelligent.
    - *Current AI:* **High.**
2. **Sentience (The "Heart"):**
    - The capacity to *feel* and experience subjective states (qualia)—pain, joy, the "redness" of a rose.
    - *Example:* A dog is sentient.
    - *Current AI:* **None (Debated, but generally accepted as zero).**
3. **Sapience (The "Soul/Self"):**
    - Wisdom, self-awareness, and the ability to reflect on one's own existence ("I think, therefore I am").
    - *Example:* Humans are sapient (*Homo sapiens*).

The Turing Test vs. Sapience

Why “Passing” Isn’t Enough

- **The Turing Test (1950):** Measures **Imitation**.
  - If a human cannot distinguish the machine from another human, it passes.
  - *Flaw:* It tests *performance*, not *internal reality*.
- **The Chinese Room Argument (Searle):**
  - Imagine a person inside a room with a rulebook. They receive Chinese characters, look up the rule, and output new characters.
  - To the outside, they speak Chinese perfectly.
  - Inside, they understand **nothing**.
- **Conclusion:** Modern LLMs are the person in the room—manipulating symbols with statistical mastery, but lacking the *sapience* to understand meaning.

Separating Science from Sci-Fi

The AI Taxonomy

1. **ANI (Artificial Narrow Intelligence):**
   - Expert at ONE thing (Chess, Protein Folding, Excel).
   - *Status:* **Current Reality.**
2. **AGI (Artificial General Intelligence):**
   - Human-level ability to learn any intellectual task.
   - *Status:* **The Goal (Timeline debated: 5-20 years).**
3. **ASI (Artificial Super Intelligence):**
   - Intellect far smarter than the best human brains in practically every field.
   - *Status:* **Theoretical.**

What AI Does Well (Strengths)

- **Pattern Recognition:** Detecting tumors in X-rays, spotting credit card fraud.
- **Optimization:** Logistics, server load balancing, traffic routing.
- **Prediction:** Next-word prediction (writing code, email drafts).
- **Creative Synthesis:** Combining existing styles (image generation).

Where AI Fails (Limitations)

- **Common Sense:** Lacks a "World Model." It doesn't know physics intuitively like a toddler does.
- **Causality:**
  - AI sees: *People with umbrellas get wet.*
  - AI concludes: *Umbrellas cause wetness.*
- **Context Window:** Models have limited "short-term memory."
- **Math & Logic:** LLMs are probabilistic, not deterministic (though getting better with tools).

1. **Hallucination**
    - Confidently stating facts that are wrong.
    - *Cause:* The model is predicting the most *probable* next word, not checking a database of facts.

The Pitfalls ii

2. **Bias**
   - "Garbage In, Garbage Out."
   - If training data contains historical biases (e.g., CEOs are mostly men), the model will replicate that bias.

The Pitfalls iii

3. **The “Black Box” Problem**
    - We know *that* it works, but strictly speaking, we don’t always know *how* it reached a specific conclusion inside the hidden layers.

Hype vs. Reality

The Gartner Hype Cycle: 1. **Innovation Trigger:** ChatGPT launch. 2. **Peak of Inflated Expectations:** “AI will replace everyone next year.” 3. **Trough of Disillusionment:** “AI makes too many mistakes.” 4. **Slope of Enlightenment:** “Here is how we actually use this useful tool.”

Conclusion: AI is a Co-pilot, not an Autopilot.

Q&A

Next: Example Notebooks