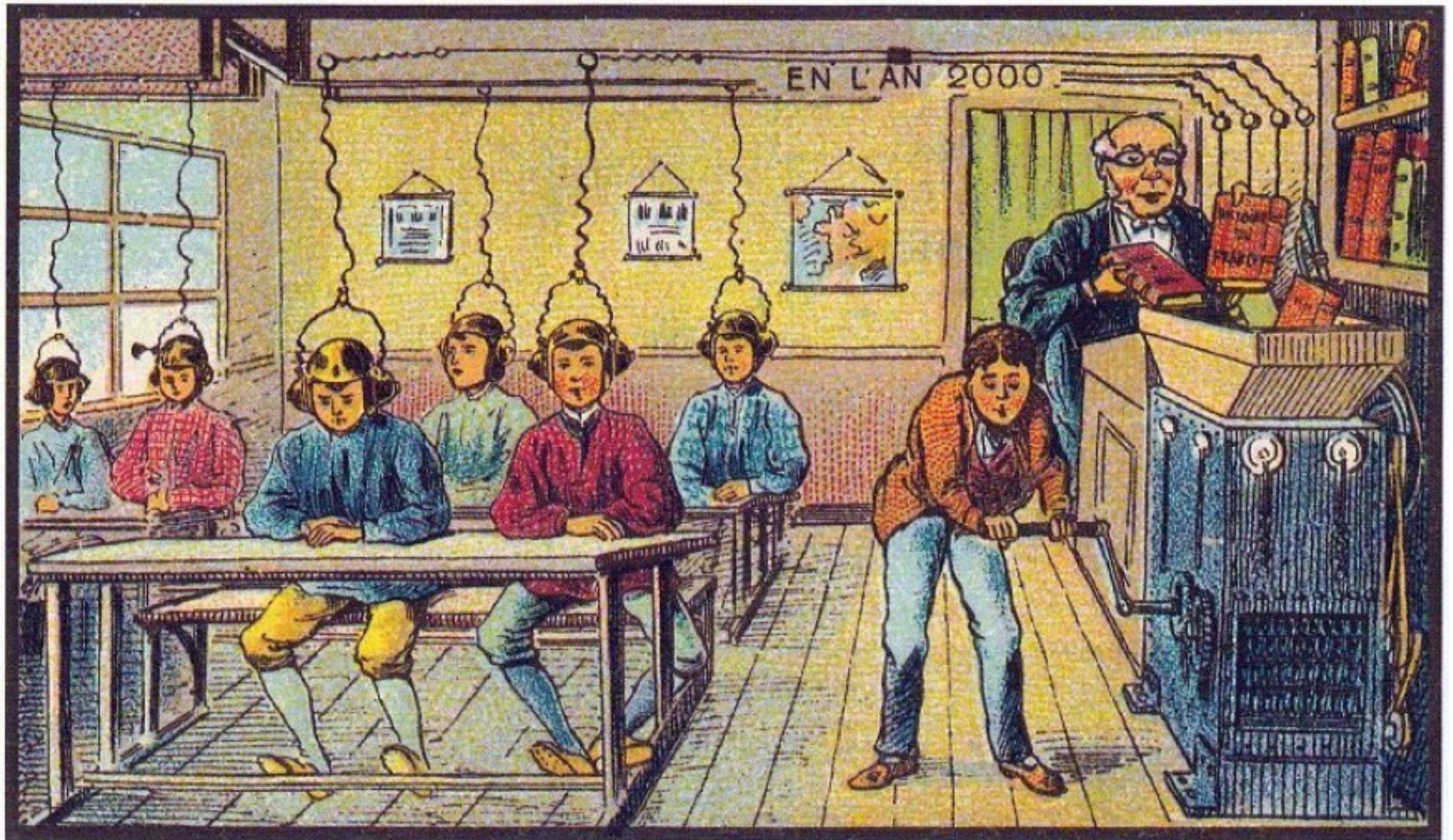# Without Ontology LLMs are Clueless

John F. Sowa

Fellow, Permion.AI

Ontology Summit

28 February 2024

A 1900 prediction about generative AI in the year 2000.

A classroom in the year 2000, as imagined in 1900 *

Large Language Models (LLMs)

Powerful technology remarkably similar to a joke in 1900.

- Dump books in a machine, turn a crank, and expect a stream of knowledge to flow through the wires.

The results are sometimes good and sometimes disastrous.

- LLM methods are excellent for translation, useful for search, but unreliable in generating new combinations.
- A lawyer used it to find precedents for a legal case.
- It generated an imaginary precedent and created a citation that seemed to be legitimate.
- But the opposing lawyer found that the citation was false.

Any results found or generated must be treated as hypotheses.

- They are abductions that must be tested by deduction.
- An ontology of the subject matter is the basis for the test.

Multiple Formal Ontologies

Detailed formal ontologies are essential for precision.

Banks, for example, transmit messages that use a generic ontology. But the details are so specific to each bank that no two banks have identical detailed ontologies. The results must be precise to a fraction of a cent. Generic messages must always be interpreted and verified in terms of each bank's ontology for each account type.

Similar precision is essential for every business. Recent issues with Boeing show that the ontology for each model of each plane has a different ontology for generic words that are used to describe airplanes. Statistical methods such as LLMs cannot be trusted without verification with the ontology for each model and even each instance of a plane.

Summary: Mistakes in any industry can be costly -- to the extent of bankruptcy or deaths. 99% certainty is unacceptable. Nobody will fly a plane that has a 1% chance of crashing.

1. Semantics of Language and Thought

Human languages are derived from the way people talk about everything they see, hear, feel, and do.

But thinking is intimately integrated with perception and action.

The semantics and pragmatics of a language
- Are distributed in the brains of every speaker of the language,
- Are dynamically generated and interpreted in terms of a constantly evolving context,
- Relate internal neural processes to external communication.

Large Language Models (LLMs) are based on written language.
- But language maps a continuum of sensations to a finite vocabulary.
- Written language is isolated from the perception, feelings, actions, and reactions of people in a dynamically changing world.
- How much of human intelligence is lost in a mapping to LLMs?