Epistemebox:
A formalism for the Knowledge Base of Corporate AI Systems

By
G. B. Prabhat

Abstract of the lecture given by the author at
Gen AI-ML Summit 2026 conducted by the Chennai Mathematical Institute, Chennai, India
on January 10, 2026

The mystic power of Artificial Intelligence (AI) requires to be harnessed in practical frameworks and formalisms for it be applied to the task of corporate performance improvement. Such frameworks will have 2 entities: the methods and the episteme. The *episteme* is the knowledge base that underlies the framework which is the wellspring of all decisions that AI would take while the *methods* are the procedures—LLM prompts, Agents, Rule-based Inference Engines, Blackboard Systems and other analytics techniques —that act upon the episteme to deliver organization-specific advice, decisions, actions and results.

The focus of this article is the episteme.

The venerated homily that applied to the data management age, “Garbage in, garbage out,” is only more inexorably relevant to AI-based systems. The quality of decisions or the work performed by such systems will depend completely on the quality of the episteme, i.e., the knowledge that the methods act upon. Over decades of experience, we know how difficult it has been to certify the integrity of data in enterprise systems. Certifying the integrity of the episteme is orders of magnitude more complex. Witness how misleading (and perhaps deceitful?) hallucinations are. Organized methodologies, specifically tailored to build AI-based systems, are necessary; painful adaptations of data management methodologies would fail sooner rather than later.

The *Epistemebox* is an all-encompassing structural formalism for the knowledge base that would underlie non-trivial corporate AI applications. It is only part of a larger lifecycle management methodology for AI-based enterprise applications.

The Epistemebox supplies the structure and nomenclature to the episteme. As the name indicates, it is three dimensional. Refer to Figure 1. (below) for the representation of the Epistemebox.

Continued.

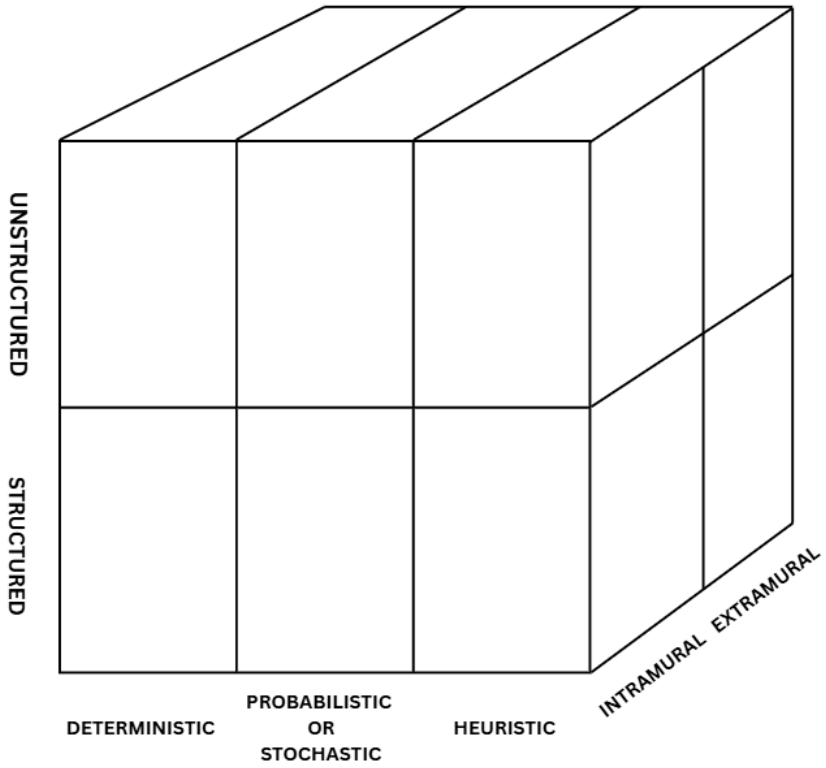

Figure 1. The Epistemebox – 3 dimensional

Dimension 1: Extramural and Intramural Knowledge
A business organization takes decisions or implements actions considering:

- the circumstances outside its boundaries—industry, geography, macroeconomic environment and culture, and
- internal knowledge or knowledge specific to the organization—business goals, operating capacities, business processes and costs.

Knowledge specific to a business organization such as its business goals, current processes, history and expertise gathered over the years is called *Intramural Knowledge*. Knowledge extraneous to the organization that is relevant to the problem-solving task is called *Extramural Knowledge*. Extramural Knowledge is all the public knowledge on the internet but may also be found in textbooks and industry handbooks.

Techniques like Context Engineering and Agentic RAG help include Intramural Knowledge as part of the episteme.

Dimension 2: Structured and Unstructured Knowledge
Knowledge that the methods reason with could be *Structured* (databases, spreadsheets, tables, decision trees) or *Unstructured* (text, images, voice, movies).

Dimension 3: Categories of knowledge
The episteme could contain three types of knowledge:

- *Deterministic*, which offers a definite solution to a problem. Examples are determining the velocity of a body or the roots of a quadratic equation. A non-mathematical example is confirming a diagnosis of tuberculosis if the sputum of the patient tests positive for mycobacterium tuberculosis.

- *Probabilistic or Stochastic*, which offer solutions with uncertainty or incorporating the effects of randomness. A well-known example is that the probability of getting either heads or tails on the toss of a coin is 0.5.
- *Heuristic*, which offers solutions based on thumb rules which are largely empirical and experiential. Such knowledge may not stand scientific scrutiny but is practically useful. Heuristic knowledge also tends to be hyperlocal. Cultural considerations to be borne in mind when the same product is launched in different geographies is an example of Heuristic knowledge.

When the Epistemebox is reduced to 2 dimensions, by flattening it against the dimension of Intramural or Extramural knowledge, 12 quadrants of knowledge emerge, 6 for Intramural and 6 for Extramural. See Figure 2. (below).

Continued.

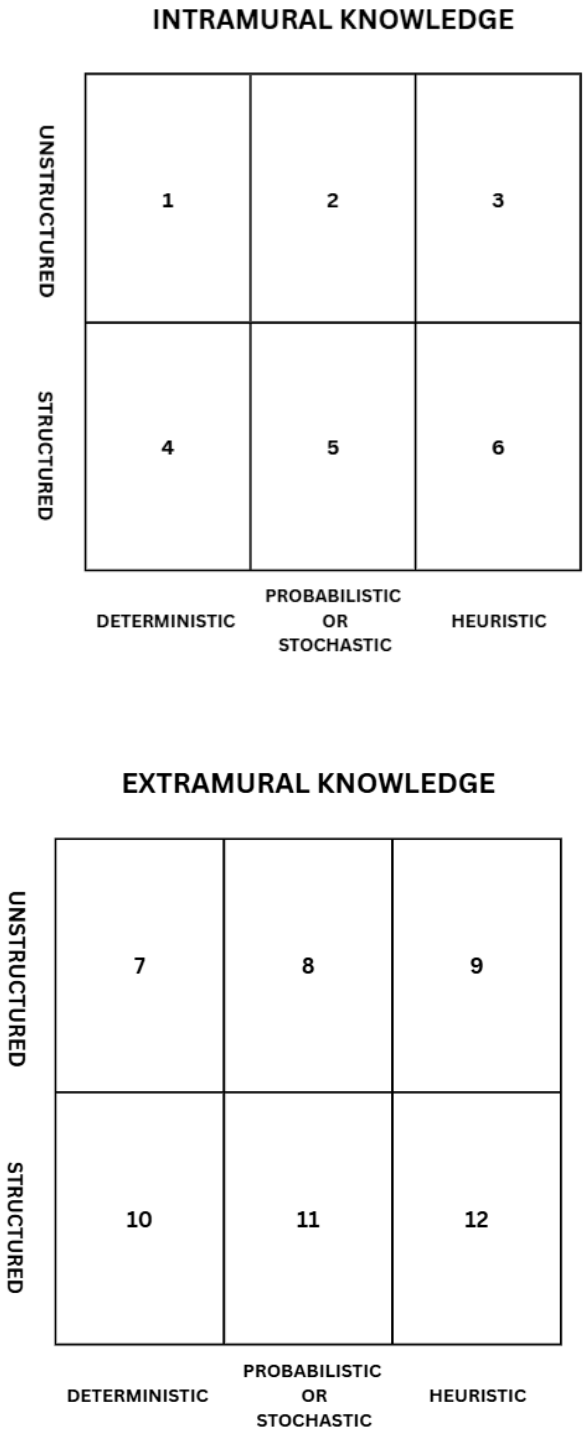

Figure 2. The Epistemebox – 2 dimensional

Examples of select quadrants: Quadrants 3 and 6 could contain the high-value expertise that resides in the heads of multiple veterans in the company that has been gathered over years of experience. Expert knowledge has seldom been capitalized by organizations in the form of assets and walks out when the employee does. The *Fortune 500* list and other lists published by research companies are examples of knowledge pertaining to quadrant 10. Stories of what worked and what didn't for people outside the company could populate quadrant 9.

None of the quadrants should be construed as watertight compartments; porosity exists at their boundaries. Knowledge in each of the quadrants—more as a rule rather than as an exception—will be hybrid. For example, a Microsoft Word document that embeds a spreadsheet could be Unstructured and Structured. The Epistemebox provides a useful taxonomy to ensure that no quadrant is overlooked during the lifecycle management process. Each quadrant represents the dominant type of knowledge it contains rather than exclusivity. The knowledge in different quadrants requires to be managed using different techniques. The Epistemebox establishes the basis for the differential management.

All the quadrants may not be necessary for any one application. The Epistemebox is a useful lens to view the hardly static episteme. Along with other lifecycle methodology elements, such a view would enable the certification of the episteme. A whole host of lifecycle methodology elements is required to complement the Epistemebox for the orderly commissioning and sustenance of AI for corporate performance improvement.

The author is the founder of Anantara Solutions Private Limited and may be reached at *Prabhat_GB@anantsol.com.* **Epistemebox is the trademark of G. B. Prabhat**