Introduction: Evidence Gaps in a Rapidly Expanding Field

The Free Energy Principle and Active Inference Framework

The Free Energy Principle (FEP), introduced by Karl Friston, proposes that self-organizing systems maintain their structural and functional integrity by minimizing variational free energy—an upper bound on sensory surprise friston2006free, friston2010free. Under this principle, living systems are cast as approximate Bayesian inference engines that build generative models of their environment and act to reduce the discrepancy between predicted and observed states. Active Inference (AIF) extends this picture from passive perception to goal-directed behavior: agents select actions that bring about observations consistent with their preferred states, unifying perception, learning, and decision-making within a single variational framework parr2022active, friston2017active. Since its initial formulation for sensorimotor control, AIF has been applied to navigation, visual foraging, language comprehension, social cognition, and multi-agent coordination.

Challenges Posed by Rapid Literature Growth

The active inference literature has expanded exponentially over the past two decades, sustaining peak publication volumes into the late 2020s. While early research concentrated almost exclusively on theoretical neuroscience, the field has since diversified across biology (C5), robotics (C2), computational psychiatry (C4), algorithm scaling (B), and formal mathematics (A1). This rapid, multi-disciplinary growth creates three interrelated challenges. First, tracking which core theoretical claims—such as FEP universality or the physical realism of Markov blankets—are deeply supported, contested, or merely assumed becomes intractable. Second, because the relationship between mathematical formalisms and empirical evidence remains frequently implicit, systematic evidence synthesis demands prohibitive manual labor. Third, new entrants must navigate a literature heavily weighted toward broad qualitative philosophy (A2), interspersed with rapidly accelerating, highly specialized applied pockets.

Traditional narrative reviews attempt to address these challenges but are inherently static, subjective, and quickly outdated.

Related Work and Prior Meta-Analyses

Several prior efforts have surveyed aspects of the Active Inference landscape. Sajid et al. sajid2021active compare active inference with alternative decision-making frameworks; Da Costa et al. dacosta2020active synthesize the discrete-state-space formulation; Lanillos et al. lanillos2021active survey robotics applications; Smith et al. smith2021computational provide a tutorial bridging theory and empirical data; and Millidge et al. millidge2021understanding examine information-theoretic foundations of exploration behavior. Ramstead et al. ramstead2018answering extend the FEP to questions of biological self-organization, while Pezzulo et al. pezzulo2015active connect active inference to homeostatic regulation.

Closest to our work, Knight, Cordes, and Friedman knight2022fep conducted a systematic literature analysis of publications using the terms "Free Energy Principle" or "Active Inference," with an emphasis on works by Karl J. Friston. Their analysis—maintained by the Active Inference Institute—combined manual annotation of structural, visual, and mathematical features with automated

Synergizing Knowledge Graphs and LLMs

Recent systematic literature initiatives underscore a powerful reciprocal synergy between Large Language Models (LLMs) and Knowledge Graphs: LLMs parse unstructured text to rapidly extract semantic claims, efficiently populating the structured, queryable architecture of the graph quevedo2025combining, li2024unifying. We adopt the *nanopublication* groth2010anatomy—a minimal, machine-readable unit of scientific evidence comprising a core assertion bound to explicit provenance metadata—as the fundamental serialization format for this extracted knowledge.

This Study: Approach and Overview

This paper presents a computational meta-analysis of the Active Inference literature ($N = 1208$). Rather than relying exclusively on bibliometric metadata or slow manual coding, we deploy a Large Language Model (LLM) to "read" each paper's abstract and assess its relationship to eight core hypotheses within the FEP paradigm. We serialize these assessments as nanopublications—each encoding an assertion ("Paper X supports Hypothesis Y") coupled with the LLM's natural-language reasoning and confidence score. The resulting knowledge graph aggregates these nanopublications and links them to paper metadata, citation networks, subfield classifications, and hypothesis definitions. A citation-weighted scoring formula quantifies the net evidence for or against each hypothesis, producing scores in $[-1, 1]$ that reflect both the direction and strength of published evidence.

Research Questions

This meta-analysis addresses four primary research questions:

1. **RQ1 (Field Structure):** What is the disciplinary structure and growth trajectory of the Active Inference literature, and how are papers distributed across the three domains—Core Theory (A), Tools & Translation (B), and Application Domains (C)?
2. **RQ2 (Growth Dynamics):** What are the temporal growth dynamics of the field, and which subfields are experiencing the most rapid expansion?
3. **RQ3 (Hypothesis Evidence):** What is the current balance of evidence for and against the eight standard hypotheses, and how has this balance evolved over time? (See hypothesis dashboard and assertion figures in §4.)
4. **RQ4 (Tooling Readiness):** What is the state of software tooling and infrastructure for Active Inference research, and what gaps remain?

Scope and Delimitations

This study focuses on the English-language peer-reviewed and preprint literature retrievable from arXiv, Semantic Scholar, and OpenAlex. We do not include book chapters or monographs not indexed by these sources, software documentation, or non-English publications. Domain classification uses keyword matching rather than expert annotation—a deliberate trade-off favoring reproducibility over precision, whose consequences we quantify in Section 3. Hypothesis scoring relies on LLM-extracted assertions; the fidelity and limitations of this approach are examined in Section 4a. The hypothesis definitions and domain taxonomy are informed by, but not identical to, the Active Inference Ontology used by Knight et al. knight2022fep; future alignment would enable direct comparison with that earlier analysis.

Principal Contributions

This work makes five contributions:

1. **A multi-source retrieval and deduplication pipeline** for Active Inference literature, using a canonical identifier hierarchy across three academic databases.

2. **A nanopublication-based knowledge graph schema** encoding directed, confidence-scored assertions about eight core hypotheses with full provenance tracking.

3. **A quantitative field overview** characterizing the growth, domain distribution (A/B/C taxonomy), citation topology, and latent topic structure of the Active Inference literature.

4. **An LLM-based hypothesis scoring dashboard** that produces differentiated evidence profiles with temporal trend visualization.

5. **A tooling assessment** of the software ecosystem supporting Active Inference research, including the implemented extraction pipeline, existing software (pymdp, SPM