

# The Discovery Engine: A Framework for AI-Driven Synthesis and Navigation of Scientific Knowledge Landscapes

Vladimir Baulin\*

*Active Inference Institute, Crescent City, California, 95531, USA and  
Universitat Rovira i Virgili, Tarragona, Spain*

Austin Cook, Daniel Friedman, Janna Lumiruusu,

Andrew Pashea, Shagor Rahman, and Benedikt Waldeck

*Active Inference Institute, Crescent City, California, 95531, USA*

(Dated: May 26, 2025)

arXiv:2505.17500v1 [cond-mat.soft] 23 May 2025

# Abstract

The prevailing model for disseminating scientific knowledge relies on individual publications dispersed across numerous journals and archives. This *legacy* system is ill suited to the recent exponential proliferation of publications, contributing to insurmountable information overload, issues surrounding reproducibility and retractions. We introduce the Discovery Engine, a framework to address these challenges by transforming an array of disconnected literature into a unified, computationally tractable representation of a scientific domain. Central to our approach is the LLM-driven distillation of publications into structured "knowledge artifacts," instances of a universal conceptual schema, complete with verifiable links to source evidence. These artifacts are then encoded into a high-dimensional Conceptual Tensor. This tensor serves as the primary, compressed representation of the synthesized field, where its labeled modes index scientific components (concepts, methods, parameters, relations) and its entries quantify their interdependencies. The Discovery Engine allows dynamic "unrolling" of this tensor into human-interpretable views, such as explicit knowledge graphs (the CNM graph) or semantic vector spaces, for targeted exploration. Crucially, AI agents operate directly on the graph using abstract mathematical and learned operations to navigate the knowledge landscape, identify non-obvious connections, pinpoint gaps, and assist researchers in generating novel knowledge artifacts (hypotheses, designs). By converting literature into a structured tensor and enabling agent-based interaction with this compact representation, the Discovery Engine offers a new paradigm for AI-augmented scientific inquiry and accelerated discovery.

## I. INTRODUCTION

Scientific progress relies on the effective accumulation, synthesis, and critical evaluation of knowledge. Traditionally, the well-documented, peer reviewed publication served as the primary standard for filtering and disseminating credible findings within the scientific community. Recently, however, we are witnessing an unprecedented acceleration in research output, a veritable explosion of scientific publications across all disciplines [1]. Yet, this very abundance creates a paradox: the sheer volume threatens to overwhelm the mechanisms designed for its assimilation and synthesis. Researchers, even within highly specialized subfields, face

---

\* [vbaulin@activeinference.institute](mailto:vbaulin@activeinference.institute)

an almost insurmountable challenge in keeping abreast of relevant developments, integrating disparate findings, and identifying the truly novel signals amidst the noise [2]. This information overload contributes to disciplinary fragmentation, hindering the cross-pollination of ideas essential for disruptive innovation [3]. Furthermore, persistent concerns regarding "reproducibility crisis" [2], predatory journals, inflation of research areas[4], growing retractions and the potential influences of bibliometrics on research direction [5] highlight systemic challenges in validating and prioritizing scientific contributions to fundamental knowledge.

We argue that a significant contributing factor to these challenges lies in the **persistent reliance on a document-centric model** of scientific communication, a legacy format optimized for human reading but ill-suited for large-scale machine search, analysis and synthesis. A scientific paper, typically disseminated as a PDF or typeset documents, presents findings embedded within a linear narrative. This narrative necessarily intertwines background research, core factual claims, detailed methodological parameters, quantitative results, and underlying theoretical assumptions with authorial interpretation, contextual framing, and rhetorical choices [6]. While vital for conveying the scientific journey, this format makes the automatic extraction and integration of the underlying, reusable knowledge components, *e.g.* the precise definitions, the specific experimental conditions, the verifiable parameter values, the explicitly tested relationships, an arduous, often manual, and inherently lossy process.

Extracting the verifiable, relational "atomic facts" or knowledge components from millions of such documents remains an intractable challenge for purely automated systems without significant human intervention or the adoption of more structured representations. Key relationships, parameter dependencies, methodological nuances, and implicit assumptions often remain locked within the prose, lacking the standardization, granularity, and explicit relational structure required for effective machine processing and deep synthesis across the corpus.

To address this bottleneck and unlock the next level of scientific progress, particularly in an era increasingly shaped by artificial intelligence (AI), we need to transition towards a knowledge infrastructure that is fundamentally machine-readable and designed for automatic extraction and synthesis into new narratives. This requires embracing principles like FAIR (Findable, Accessible, Interoperable, Reusable) [7, 8], ensuring that the core components of scientific knowledge are structured not just for human consumption, but for computational

analysis and integration by intelligent agents [9]. We need to move beyond viewing the literature as a digital library of static documents towards conceptualizing it as a dynamic, interconnected graph of verifiable knowledge components.

Here we introduce the Discovery Engine (DE) as a methodology and conceptual platform designed to architect this transition. The DE moves away from the analysis of isolated publications towards the automated synthesis of entire scientific fields into dynamic, structured knowledge repositories. The central construct generated and maintained by the DE is the **Conceptual Nexus Model (CNM)**. This CNM is envisioned as an evolving, AI-curated knowledge graph that serves as a cognitive "World Model" for a specific research domain [10], capturing not just individual facts in human language but the intricate web of relationships between concepts, methods, findings, and theories. The Discovery Engine is not a static model, but an interactive ecosystem incorporating AI agents designed to assist researchers in navigating, analyzing, and extending this synthesized knowledge landscape.

It is noteworthy, that LLMs are employed not for unconstrained generation, but as tools for *Structured Knowledge Distillation*, guided by adaptive, formally-defined templates (the structured and self-consistent schema for processing scientific documents, specific for each field and dynamically adapting). This process extracts granular, verifiable knowledge components from source literature, ensuring consistency and traceability. These extracted components are then represented using formalisms designed to capture both semantic meaning and structural relationships. This might involve hybrid approaches combining dense vector embeddings for semantic similarity [11, 12] with techniques capable of encoding explicit compositionality and relational structure, potentially drawing inspiration from Vector Symbolic Architectures (VSA) [13, 14] or the abstract relational language of Category Theory (CT) [15–17].

The resulting machine-readable CNM provides the essential substrate for a new generation of AI-driven scientific tools that are optimized for machines, while also being useful for computational multimodal generation and human accessibility. AI agents operating on this structured knowledge graph can perform tasks far beyond simple information retrieval [9]. They can navigate complex conceptual landscapes, identify deep structural analogies between different research areas [18, 19], detect subtle inconsistencies or contradictions across studies [20], systematically pinpoint under-explored regions or knowledge gaps [21], and assist researchers in formulating novel, data-grounded hypotheses [22]. The DE, therefore,

aims to foster a synergistic relationship between human researchers and AI systems, where AI handles the large-scale synthesis and structural analysis, freeing human intellect to focus on interpretation, creativity, and critical evaluation.

## II. THE CONCEPTUAL NEXUS MODEL: A STRUCTURED, COMPUTABLE REPRESENTATION OF SCIENTIFIC KNOWLEDGE

One of the aims of the Discovery Engine methodology is the creation and update of the Conceptual Nexus Model (CNM). It serves as a dynamic, structured representation of a scientific field, moving beyond the limitations of traditional, document-based knowledge storage. Its construction integrates AI-driven extraction processes with formal representation schemas, enabling the transformation of textual knowledge into a computable and interconnected network designed for both human exploration and automatic AI-based analysis. The aspiration is to construct a CNM of such fidelity and completeness that it begins to mirror the intrinsic logical structure of the scientific domain itself. In such a scenario, analogous to how physical systems compute their own evolution according to underlying natural laws [23], an AI agent could interact with this validated "World Model" in a zero-shot fashion, inferring new knowledge or validating hypotheses through computational exploration of the CNM's structure without requiring specific human feedback for each instance. The observation that fundamental physical laws are often expressible in remarkably compact mathematical forms [24] suggests that the parameter space of core scientific principles within a domain might be learnable and navigable by AI, provided a sufficiently structured and verified knowledge representation like the CNM.

A fundamental challenge in synthesizing scientific knowledge lies in transforming the heterogeneous, often narrative-driven content of publications into such a structured, computable format, Fig. 1. The traditional publication format, while effective for detailed human communication, inherently embeds core findings, methods, parameters, and conceptual relationships within natural language, making large-scale, automated extraction and comparison extremely difficult [6]. However, one of the requirements for reproducibility of results, is that the logical structure of experiments must be articulated with precision to allow for replication, a process that requires unambiguous language or unified schema in the description of steps [25].

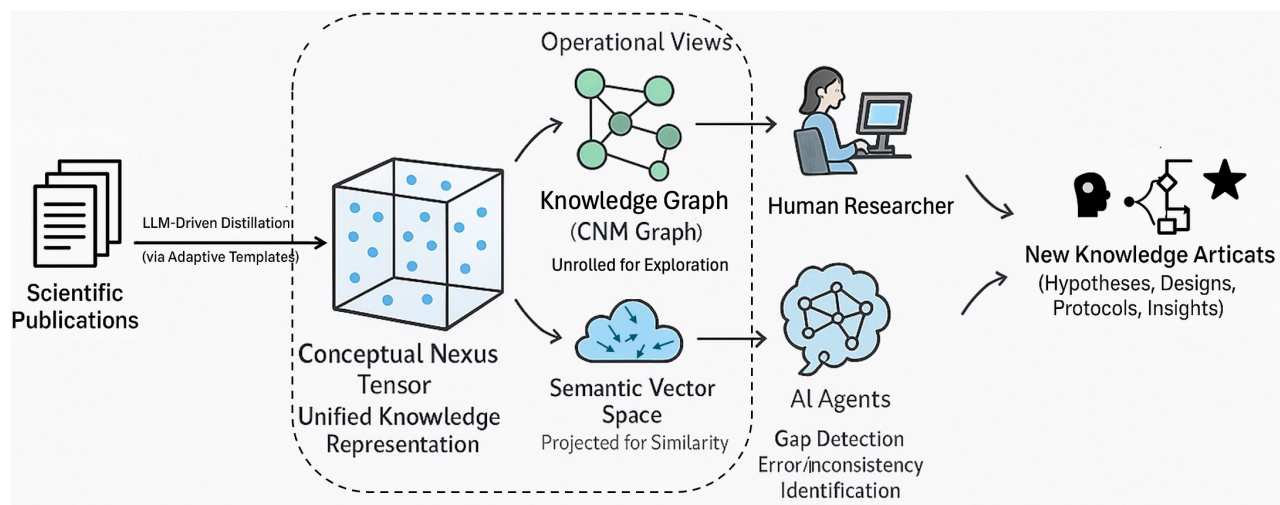


Figure 1. Conceptual Nexus Model for distillation of the knowledge into machine-readable format ready for human and agent exploration and machine-facilitated discoveries.

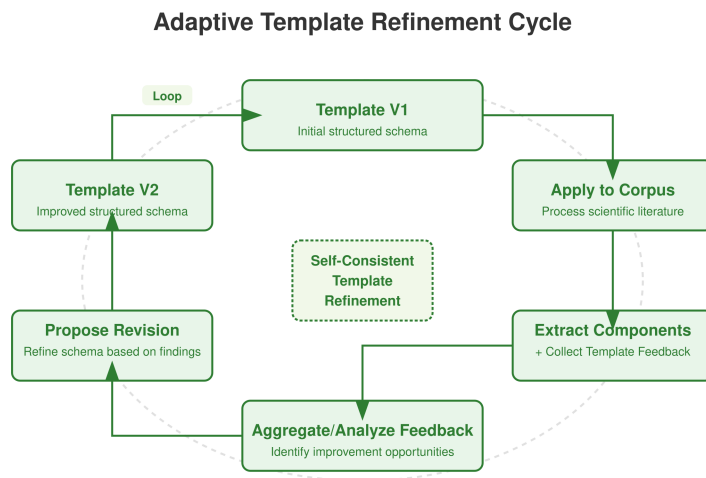


Figure 2. The self-consistent template refinement cycle in the DE until the template and the corpus of literature become consistent.

Instead of utilizing traditional bibliometrics as measures of the credibility of a publication, DE implements a robust yet flexible FAIR-aligned Verifiability and Robustness Scoring system for evaluating each knowledge artifact in the CNM. These scores could evaluate not the perceived importance of a finding, but its inherent scientific utility and trustworthiness in a dynamic and multi-faced way. For example, a "Findability & Accessibility" score would

reflect clearly defined components and well-documented data/methods, "Interoperability" would be gauged by the artifact's ability to connect with and be understood in the context of diverse concepts within the CNM, facilitated by standardized mapping to the Universal Concept Schema, "Reusability" (and overall robustness) is assessed through metrics like: Feasibility (e.g., are methods practically implementable, are parameters within realistic ranges?), Confirmation Strength (e.g., how many independent studies or lines of evidence within the CNM support this artifact or its constituent claims?), Evidence Linkage (e.g., directness and quality of provenance trails to primary data and explicit experimental validation), and Predictive Consistency (e.g., how well does this artifact align with, or get corroborated by, analogous findings or structurally similar systems identified elsewhere in the CNM?). These dynamic and adaptive scores would provide a multi-faceted, evidence-grounded measure of an artifact's potential for reliable integration and reuse in future discovery, rather than a problematic singular and fixed metric of academic influence.

Recognizing the limitations of purely statistical NLP methods for capturing the precise semantics and relationships critical to scientific discourse, this process employs Large Language Models (LLMs) as sophisticated analytical engines, carefully guided by formally specified, adaptive templates (details in Appendix and template schema, Fig. 5). These templates act as dynamic schemas, defining the types of knowledge components (nodes) and relationships (edges) to be extracted, thereby ensuring consistency and structure across diverse sources. Critically, the process mandates that LLMs provide justifications and link extracted components directly to evidence within the source publication, fostering verifiability and traceability essential for addressing reproducibility concerns [2]. These templates (drawing conceptual inspiration from Category Theory's focus on structure and mappings [15, 16], and potentially analyzable using techniques inspired by Graph Isomorphism Networks [26]) define the explicit schema for knowledge extraction. LLMs are tasked with instantiating this schema for each publication, identifying specific entities and relationships and providing auditable links back to the source text, thus ensuring verifiability and provenance. Furthermore, the template framework incorporates a *self-consistent refinement loop* (Fig. 2), whereby feedback on the template's applicability gathered during corpus processing informs its subsequent evolution [21], ensuring the schema dynamically adapts to the specific nuances and evolving structure of the knowledge within a given field [27].

The conceptual architecture of the Discovery Engine (Fig. 3), outlines the structured

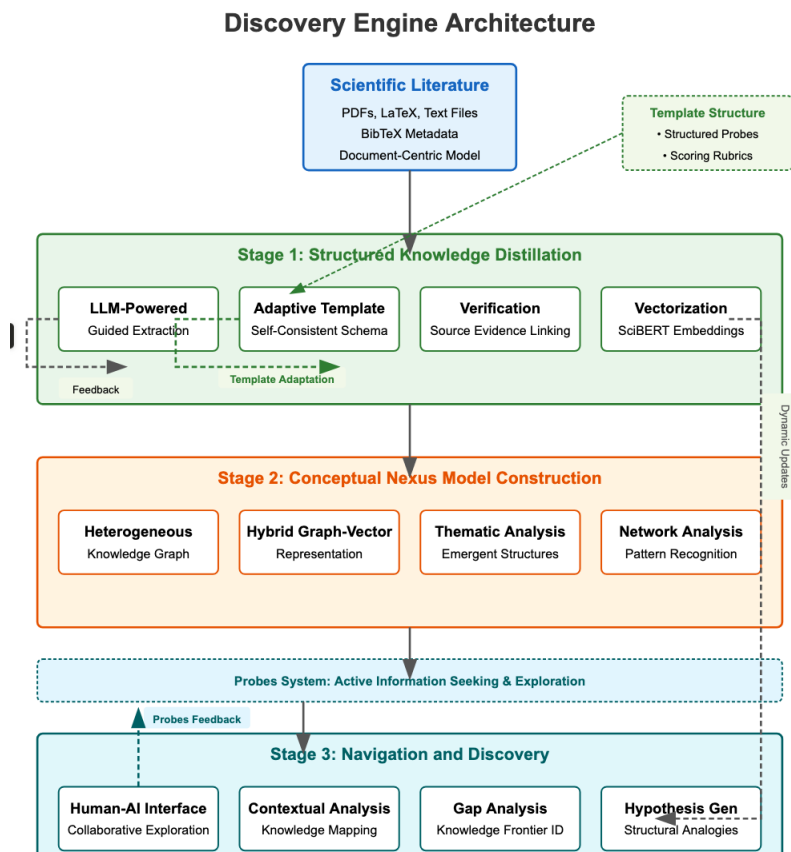


Figure 3. Conceptual architecture of the Discovery Engine framework.

knowledge distillation process, the hybrid graph-vector representation forming the CNM, and the methodologies for synthesizing and analyzing the resulting knowledge landscape. This framework enables systematic identification of knowledge gaps and inconsistencies, supports the generation of novel hypotheses grounded in synthesized evidence, and provides a foundation for next level human-AI collaboration in science. We argue that adopting such computable knowledge models is not merely an incremental improvement but a necessary evolutionary step for navigating the complexity of modern science, fostering deeper understanding, enhancing research integrity, and ultimately accelerating the engine of discovery by transforming the way we interact with and build upon our collective knowledge.

The extracted, attributed components populate the CNM, structured as a heterogeneous and hierarchical knowledge graph (Fig. 3). This graph utilizes distinct node types for di-

verse scientific entities and richly typed edges to represent their specific interrelations. While the graph forms the primary knowledge structure, complementary dense vector representations (Fig. 3) are derived to enable efficient semantic similarity search, machine learning integration, and analogy finding [12, 22].

### A. Structured Knowledge Distillation via Guided AI and Adaptive Templates

The objective is to move beyond simple document summaries or keywords extraction towards isolating granular, verifiable *knowledge components*. These components serve as the *building blocks* for representing the core concepts of a field. What constitutes a "concept" is itself a complex issue debated in philosophy and cognitive science [11, 28]. Within the DE framework, concepts are operationalized through the structured set of components: nodes and edges (see Appendix B) defined by a given appropriate kind of template. The template, see Figure 5, essentially defines which aspects, such as specific mechanisms, underlying principles, key quantitative parameters, methodological details, reported outcomes, limitations, and explicit relationships to other concepts, collectively constitute the machine-readable representation of a concept or finding within the synthesized Conceptual Nexus Model. From this standpoint, conceptual meaning arises from both its local document-centric context as well as its structured role and relations as a contribution to a larger knowledge framework [11, 29, 30].

The distillation process is initiated by defining the scope of interest, which can inform the initial structure or focus of the template. This template, implemented as a structured plain text document (e.g., Markdown), serves as a dynamic schema and an analytical lens. It contains a detailed set of probes, ranging from requests for qualitative descriptions and justifications to specific fields for quantitative parameters (requiring values and units) and comparative scores based on predefined rubrics. The modular design (e.g., sections covering aspects like system scope, objectives, implementation, etc.) ensures a systematic interrogation of each publication in a standardized form.

The extraction itself is performed by an LLM. Unlike open-ended generative tasks, the LLM's role here is carefully constrained: it is instructed to act as a rigorous analytical agent, tasked with populating the provided template based *solely* on the content of a single input publication. The prompts emphasize precision, requiring the LLM to extract specific and

well-defined values such as parameters and values of experimental techniques, adhere strictly to the template format, and critically, provide direct textual justifications from the source document for every extracted component and assigned score. This focus on verifiable extraction linked to source evidence is paramount for building a trustworthy knowledge base [2] and mitigating potential LLM hallucination or fabrication. While LLMs have known limitations [31], their ability to follow complex instructions and structured information extraction makes them suitable engines for this distillation task when properly guided and constrained by the template.

An innovative aspect of this methodology is the *self-consistent refinement loop* designed to iteratively adapt the extraction template towards an optimal representation of the scientific domain (Fig. 2). This adaptive process draws conceptual parallels with evolutionary approaches in AI-driven discovery, such as AlphaEvolve [32], where solutions (in our case, the template schema) are iteratively improved based on evaluative feedback. Here, the "evaluation" is not against a predefined fitness function but against the collective structure of knowledge expressed across the entire processed corpus. Specifically, the template incorporates meta-probes that instruct the LLM to reflect on its own capacity to capture the full scope of information within each processed publication—identifying knowledge components or relational nuances present in the text that the current template iteration fails to adequately represent. This feedback, systematically aggregated across numerous documents, reveals statistical patterns of misfit or recurring representational gaps. Periodically, this aggregated feedback drives AI-assisted (or human-curated) revisions to the template structure, such as adding new descriptive probes, refining existing component definitions, or expanding relationship taxonomies. This iterative cycle [21] enables the template schema to converge toward dynamical alignment, not with an initial user guess, but with the inherent structure and conceptual distinctions manifest in the broader scientific literature. The goal is to achieve a template that is comprehensive, minimally ambiguous, and acts as a dynamically reconciled consensus model of how the scientific field itself organizes and articulates its findings, becoming maximally fit for distilling the essence of the corpus [27].

This feedback, collected systematically across the entire corpus of processed publications, provides a rich dataset reflecting the alignment between the current analytical framework (the template) and the actual structure of knowledge presented in the literature. This aggregated feedback is then used to inform periodic revisions of the template itself, a process

potentially also assisted by LLMs tasked with synthesizing the feedback and proposing specific modifications (e.g., adding new probes, refining definitions, adjusting scoring rubrics).

This iterative cycle of extracting components is based on the current template, collecting feedback on the template’s adequacy, and using aggregated feedback to refine the template for the next iteration and establishes a self-consistent loop [21]. The justification for this approach draws parallels with self-consistent field methods in physics or iterative refinement algorithms in machine learning. The core idea is that a good representation schema (the template) should accurately and consistently capture the phenomena it aims to describe (the knowledge in the literature). If the template consistently fails to capture important aspects or generates ambiguity across many documents, it indicates a mismatch. By using the feedback derived from applying the template to the data (the literature corpus), we iteratively adjust the template itself. This process aims to minimize the "tension" or discrepancy between the template’s structure and the structure inherent in the data of the corpus of the publications (consensus). It converges, ideally, towards a state where the template provides a stable, comprehensive, and minimally ambiguous schema for distilling the knowledge within the target corpus, achieving a form of self-consistency between the analytical framework and the knowledge base [27]. This dynamic adaptation is vital for fields where concepts and methodologies evolve.

This process of guided distillation and iterative template refinement is hypothesized to effectively "extract" concepts from the literature. By forcing the analysis into a hierarchical structured format defined by the template, and then refining that template based on how well it fits the corpus, the system learns to identify and delineate the key recurring informational units (parameters, mechanisms, methods, relationships) that constitute the conceptual building blocks of the field. The converged template implicitly defines the relevant concepts and their constituent features as understood from the corpus. For example, repeated difficulties in distinguishing two types of control mechanisms using the initial template probes would lead to refined probes that better capture the distinguishing features, effectively solidifying the representation of those distinct concepts within the framework.

The distilled knowledge components, extracted according to the converged template, form the input for the subsequent representation stage. Here, techniques inspired by Vector Symbolic Architectures (VSA) and Hyperdimensional Computing (HDC) [13, 14, 33] offer a promising approach for compressing this structured information into a hierarchical vector

database. This VSA/HDC approach offers several advantages relevant to the DE’s goals: (1) *Compression*: Complex structured information is compressed into fixed-size vectors. (2) *Robustness*: High-dimensional vectors are inherently robust to noise and partial information, reflecting the often incomplete or noisy nature of scientific data extraction [13]. (3) *Computation*: Vector operations (addition, multiplication, permutation, similarity calculation via dot product/cosine/Hamming distance) provide a basis for querying, comparing, and reasoning about the structured knowledge directly in the vector space [33]. A concrete example is the challenge of representing recurring methodological descriptions. Thousands of papers might describe variations of a standard protocol (e.g., PCR, Western Blot, specific simulation setup). Instead of storing redundant or incomplete textual descriptions, the DE uses the template to extract key parameters and variables. VSA/HDC could then represent the core protocol as a base vector, and specific variations (e.g., different annealing temperatures, antibody concentrations, simulation time steps) could be encoded by binding parameter-value vectors to the base vector or modifying specific components. This creates a highly compressed, structured representation where similarities and differences between protocols can be computed via vector operations, automatically identifying clusters of similar methods or specific deviations.

This entire process relies on several key assumptions and faces potential challenges:

- **LLM Fidelity Assumption:** The process assumes that the guided LLM can accurately extract the specified components and provide faithful justifications without significant hallucination or misinterpretation, especially concerning quantitative data and nuanced relationships. Ongoing LLM development and rigorous validation are crucial [34].
- **Template Expressiveness Assumption:** It assumes that a structured template, even an evolving one, can adequately capture the richness and complexity of scientific concepts and arguments, which often involve subtle nuances, implicit assumptions, and complex logical structures potentially challenging to fit into predefined probes [28].
- **Convergence Assumption:** The iterative template refinement process is assumed to converge towards a stable and useful state. However, convergence is not guaranteed. Oscillations, divergence, or convergence to a suboptimal local minimum are possible, depending on the quality of feedback, the nature of the literature, and the template

update mechanism. Continuous monitoring and potential human oversight are likely necessary.

- **Bias Amplification Risk:** Biases present in the literature corpus or the initial template design could potentially be amplified through the iterative process if not carefully managed. The feedback mechanism needs to be robust against reinforcing systematic omissions or misrepresentations.
- **Scalability:** Processing millions of papers and managing the feedback loop for template evolution presents significant computational and logistical challenges, requiring efficient algorithms and infrastructure.

These challenges represent the validity limitations of this stage. For instance, if an LLM consistently fails to extract specific quantitative data accurately despite template refinements, or if the template structure fundamentally cannot represent a newly emerging paradigm in the field, the distillation process will be compromised, impacting the quality of the resulting CNM. Similarly, if the iterative refinement loop fails to stabilize or converges to a poorly representative template, the self-consistency goal is not met.

Despite these challenges, the proposed methodology of guided AI distillation coupled with adaptive template refinement offers a principled and potentially powerful approach for converting the vast scientific literature into a structured, computable knowledge base, paving the way for the synthesis and discovery stages of the Discovery Engine. The structured output adheres inherently to FAIR principles: (i) being findable (via metadata), (ii) accessible (machine-readable format), (iii) interoperable (through the consistent schema), and (iv) reusable (as granular components). It serves as the verified, componentized input for subsequent representation within the pipeline.

## B. Unified Knowledge Representation: The Conceptual Nexus Tensor

Following structured distillation (Sec. II A), the heterogeneous collection of extracted knowledge artifacts—each a rich instantiation of the Universal Concept Schema (UCS, Appendix B) comprising textual descriptions, quantitative parameters, symbolic types, and explicit relational links—is encoded into a unified, high-dimensional mathematical object:

the **Conceptual Nexus Tensor** ( $T_{\text{CNM}}$ ). This tensor serves as the primary, computationally tractable representation of the synthesized scientific knowledge field within the Discovery Engine. It moves beyond a simple collection of linked data points to embody a latent space capturing the multifaceted interdependencies of scientific knowledge.

The construction of  $T_{\text{CNM}}$  is a critical encoding step. Its modes (axes) are rigorously defined by the UCS archetypes and key metadata dimensions extracted from the source literature. For instance, distinct modes could index:

- **UCS Node Archetypes** (e.g., ‘Concept/Principle’, ‘Process/Method’, specific instances like ‘natural-selection’ or ‘PCR’).
- **UCS Relationship Archetypes** (e.g., ‘CAUSES’, ‘USES\_INPUT’, ‘EVIDENCE\_FOR’).
- **Contextual Dimensions** derived from knowledge artifact attributes (e.g., discretized parameter values, experimental conditions, temporal markers, semantic features from textual descriptions via pre-trained embeddings [11, 12]).
- **Provenance Indicators** (e.g., clusters of source publications, research sub-domains).

An entry  $T_{i,j,k,\dots}$  in this tensor would quantify the existence, strength, probability, or information-theoretic measure (e.g., mutual information) of the relationship between the elements indexed by  $i, j, k, \dots$  across their respective modes. For example,  $T_{\text{ConceptA}, \text{ProcessB}, \text{CAUSES}}$  could store a value indicating the strength of evidence that Concept A causes Process B. This explicit indexing transforms the initial array of disconnected publications into a structured and compact tensor that inherently models n-ary relationships and interdependencies. Methods for populating such a tensor could range from direct encoding of extracted structured relations to learning procedures that map graph structures and component features into this tensorial space, potentially using techniques inspired by tensor factorization for knowledge graph completion [35] or graph neural networks that learn to populate tensor entries [36].

While  $T_{\text{CNM}}$  itself is a dense, high-dimensional object primarily intended for machine operation, the Discovery Engine provides mechanisms to "unroll" or project it into human-interpretable views:

- **Knowledge Graph View (The CNM Graph):** A primary projection renders  $T_{\text{CNM}}$  as the explicit, heterogeneous Conceptual Nexus Model graph discussed previ-

ously (Fig. 3). Nodes correspond to UCS artifact instances (derived from tensor mode indices or specific tensor fibers), and edges are instantiated based on significant tensor entries representing explicit relationships. This view supports human navigation, qualitative exploration, and symbolic graph algorithms.

- **Semantic Vector Space Views:** Other projections can generate task-specific vector embeddings for concepts or artifacts by, for example, slicing or contracting  $T_{\text{CNM}}$  along relevant modes. These support semantic similarity searches, clustering, and analogy finding [18, 22].

This dual nature—a core computational tensor and derivable interpretable views—is fundamental to the Discovery Engine. AI agents (Sec. IV C) primarily interact with and reason upon  $T_{\text{CNM}}$  using operations from tensor algebra, geometric deep learning [37], or learned transformations [13, 14, 33]. For instance, identifying novel relationships might involve detecting unexpected non-zero entries in  $T_{\text{CNM}}$  or applying tensor completion techniques. Synthesizing a new hypothesis (‘KnowledgeArtifact’) could correspond to constructing a new subtensor pattern based on existing tensor components and desired properties. The insights or artifacts generated through these abstract computations are then translated back into the graph view or natural language for human comprehension and validation. This transformation of narrative scientific literature into a compact, computationally amenable tensor, from which various structured representations can be derived, forms the core of the Discovery Engine’s approach to knowledge synthesis.

### C. Interoperability and FAIR Principles

The structured, component-based nature of the CNM aligns fundamentally with the FAIR guiding principles[7, 8] for scientific data management and stewardship: Findable, Accessible, Interoperable, and Reusable. Knowledge components are *Findable* through rich metadata and graph querying. They are *Accessible* via standardized formats and potentially APIs built upon the CNM. *Interoperability* is fostered by the use of controlled vocabularies (within the evolving template), potential links to standard ontologies, and the goal of using universal formalisms (like CT-inspired structures) for representation. Finally, the granular, verifiable, and context-rich nature of the components makes them highly *Reusable* for sub-

sequent analysis, synthesis, modeling, or integration into new research contexts. By design, the CNM aims to be a FAIR representation of scientific knowledge, moving beyond the limitations of isolated, opaque documents.

### III. SYNTHESIZING THE FIELD: REVEALING THE KNOWLEDGE LANDSCAPE

Constructing the CNM involves more than just aggregating individual knowledge components; it requires synthesizing these elements to reveal the emergent structure, consensus, conflicts, and overall landscape of the research field. This synthesis process, facilitated by the pipeline component of the Discovery Engine, transforms the collection of processed publications into a coherent, analyzable model.

#### A. Aggregation and Principled Integration in Research Field Synthesis

Following the structured distillation (Sec II A), knowledge components from individual publications are integrated into the unified CNM graph. This process, central to field synthesis, moves beyond simple concatenation to intelligently combine information from diverse sources, resolve redundancies, align related concepts, and establish a consistent structure.

The aggregation begins by loading the structured knowledge components—representing node instances defined by the schema with their associated metadata and preliminary relationships—into the graph framework. These components form the initial "point cloud" derived from the literature corpus according to the evolved schema.

A primary challenge during this phase is **Entity Resolution and Alignment**. The same underlying scientific concept (TheoreticalConceptNode), method (MethodNode), material, etc., may be described differently across publications. The DE employs a multi-pronged strategy:

- *Semantic Alignment*: Utilizes derived vector embeddings associated with textual descriptions. Similarity searches (e.g., k-NN in embedding space [12]) group potentially equivalent or related components, suggesting candidates for merging or linking.
- *Structural Alignment*: Leverages the structured nature of the extraction. Components extracted using the same template probe (representing a consistent aspect or

definition) are considered structurally aligned. Metadata comparison (e.g., units for ParameterNode, category for MechanismNode) provides further alignment evidence.

- *Identifier Resolution*: Links components to standardized external identifiers (e.g., DOIs for PublicationNode, ontology terms, chemical identifiers) where available, providing unambiguous anchors.

Combining these evidence streams allows for robust identification and linking of components referring to the same entity, forming consolidated nodes within the CNM.

Once related components are aligned, **Principled Integration** determines how information from multiple sources is represented:

- *Quantitative Aggregation*: Multiple reported values for a parameter, ParameterNode can be aggregated statistically (mean, median, distribution fitting) to provide a synthesized estimate and uncertainty.
- *Evidence Triangulation and Weighting*: Information prominence can be modulated based on source attributes (e.g., study type, sample size extracted via template) or network properties (e.g., citation impact of the PublicationNode, analogous to precision weighting in predictive processing architectures).
- *Explicit Conflict Representation*: Contradictory findings are not ignored but explicitly captured, for instance by linking conflicting nodes or evidence via a dedicated KnowledgeGapNode, thereby highlighting areas requiring further investigation [20].
- *Strict Provenance Tracking*: All synthesized information retains explicit links back to the originating source publications, ensuring full traceability.
- *Formal Integration (Conceptual)*: Principles from Category Theory, such as the colimit construction [38], offer a formal blueprint for merging structured information based on shared components while ensuring mathematical consistency, providing theoretical guidance for principled integration algorithms [16].

This aggregation and integration process yields a unified, coherent knowledge graph where redundancies are minimized, related concepts interconnected, conflicts highlighted, and all assertions remain traceable to their evidentiary sources. This synthesized CNM forms the foundation for subsequent discovery-oriented analyses.

## B. Discovering Emergent Structures within the Synthesized Knowledge

The Discovery Engine leverages the synthesized CNM structure to automatically discover latent organizational patterns within the research field, reflecting its intrinsic structure rather than imposing rigid external classifications. Key techniques include:

- *Thematic Clustering*: Applying methods like BERTopic [39] to semantic vector representations derived from node descriptions (*e.g.*, mechanism definitions, publication abstracts) identifies dominant research themes and subfields, providing a semantic map of the knowledge landscape.
- *Relational Network Analysis*: Employing algorithms from network science on the CNM graph structure identifies influential nodes (*e.g.*, central concepts, foundational papers), bridge nodes connecting disparate thematic clusters, cohesive research communities, and recurring structural motifs (*e.g.*, common experimental workflows involving sequences). Analyzing the temporal evolution of these structures, potentially using methods from scientometrics [3] or complex systems [27], reveals dynamic shifts in research focus and paradigm emergence, potentially uncovering dynamics analogous to self-organized criticality.
- *Quantitative Landscape Synthesis*: Statistical analysis of aggregated quantitative data reveals typical value ranges, correlations between parameters and outcomes (*e.g.*, linking material properties to system performance), identifies outliers requiring scrutiny, and establishes field-wide empirical distributions that contextualize individual findings.

These synthesis steps collectively transform the aggregated data into an interpretable model of the field’s structure, dynamics, and quantitative landscape.

## C. The CNM as a Dynamic, Evolving Knowledge Structure

Crucially, the CNM is conceptualized as a dynamic entity, distinct from static knowledge bases. Its evolution mirrors the ongoing scientific process [21, 27] through several mechanisms:

- *Continuous Integration:* New findings from ongoing publications are processed by the pipeline and integrated into the CNM graph, keeping the model current.
- *Schema Adaptation:* The underlying extraction template is periodically refined via the self-consistent feedback loop (Fig. 2), allowing the representational framework itself to adapt to evolving concepts and improve its fidelity.
- *Dynamic Emergent Structures:* Thematic clusters, influential concepts, and network topology identified via synthesis are not fixed but evolve over time [3, 39], providing insights into the meta-level dynamics of the scientific field.

This inherent dynamism, managed through appropriate versioning, ensures the CNM remains a relevant and adaptive foundation for ongoing scientific inquiry, transforming the knowledge base into a living model that reflects the research front.

#### IV. NAVIGATING THE KNOWLEDGE LANDSCAPE: INTERACTION AND DISCOVERY

The synthesized Conceptual Nexus Model (CNM) serves not only as a dynamic knowledge repository but, more importantly, as an interactive landscape for exploration and discovery. The Discovery Engine provides mechanisms for users, both human researchers and AI agents [9], to navigate this landscape, analyze information in context, and leverage the structure of the CNM to identify novel research opportunities.

##### A. Contextual Placement and Analysis

A fundamental capability enabled by the CNM is the contextualization of new information. When a researcher introduces a new query, a draft manuscript abstract, a set of experimental results, or references a recent publication, the DE processes this input using the same distillation and representation pipeline used to build the CNM itself. The resulting structured, vectorized components are then mapped onto the existing CNM landscape. Algorithms compute semantic proximity (using embeddings) and structural relationships (using graph connectivity or potentially VSA/CT formalisms) to identify the most relevant existing nodes, thematic clusters, and established relationships within the model [12]. This

mapping provides immediate context, indicating how the new information relates to the established body of knowledge: Does it align with existing consensus? Does it fall within a known thematic area? Does it address an identified gap? Does it potentially conflict with previous findings? This contextual placement transforms information retrieval into a deeper analytical process.

## B. Structure-Driven Exploration and Conceptual Traversal

The Discovery Engine facilitates modes of knowledge exploration fundamentally different from traditional linear document reading or keyword-based search paradigms. Users interact with the CNM by navigating its inherent relational structure, often visualized through interactive interfaces such as dynamic graph browsers or hyperlinked knowledge bases (conceptually similar to applications developed for specific databases [40]). This structure-driven navigation allows researchers to follow diverse conceptual pathways through the synthesized landscape. For example, one can trace the implementation of specific scientific mechanisms (MechanismNode) across different experimental systems (SystemNode) via ImplementsMechanismEdge. Users can compare methodologies (MethodNode) applied to investigate related phenomena (PhenomenonBehaviorNode) by exploring neighboring nodes within identified thematic clusters (Sec. III) or connected via specific relation types. Foundational concepts (TheoreticalConceptNode) or seminal publications (PublicationNode) can be identified by examining nodes with high centrality metrics derived from graph analysis. Critically, users can explore the frontiers of knowledge by investigating sparsely connected regions of the graph or by focusing on identified KnowledgeGapNode instances and their associated contextual nodes. The ability to systematically track the evolution of concepts or the usage of methods over time, by incorporating temporal metadata associated with nodes and edges, adds another dimension to exploration [3]. This traversal, guided by the CNM’s explicit structure and augmented by semantic proximity information, fosters a more holistic and nuanced understanding of a field’s topology than is possible through fragmented document access [41].

### C. Agent-Assisted Knowledge Synthesis and Generative Exploration on the CNM

The explicit, machine-readable, and richly interconnected structure of the Conceptual Nexus Model (Sec. II) serves as an ideal operational environment for sophisticated AI agents, transforming the knowledge graph from a passive repository into a dynamic substrate for discovery [9]. Within the Discovery Engine ecosystem, these agents are designed not merely to retrieve information, but to actively assist researchers in synthesizing existing knowledge and generating novel insights by leveraging the CNM’s structural and semantic integrity.

AI agents within the Discovery Engine can be tasked with several advanced functions that transcend traditional search and analysis capabilities:

- **Automated Evidence Aggregation and Structured Synthesis:** Agents can systematically traverse the Conceptual Nexus Model, following defined relational pathways (e.g., chains of ‘CAUSES’ edges, or compositions of ‘IMPLEMENTS\_METHOD’ and ‘PRODUCES\_OBSERVATION’ links) to gather diverse yet interconnected knowledge artifacts pertaining to a specific scientific query or hypothesis. Beyond simple collection, these agents can employ summarization techniques or structured reasoning (akin to "thought graph" traversal [42]) to synthesize these components into coherent, evidence-backed narratives or structured summaries. For instance, an agent could construct a mechanistic explanation for a phenomenon by assembling a subgraph of relevant ‘MechanismNode’, supporting ‘PublicationNode’, and associated ‘Property/Parameter’, ensuring all links are justified by the CNM’s provenance data [43].
- **Complex Pattern Recognition and Inconsistencies Detection:** The CNM’s graph structure is amenable to advanced analytical techniques. AI agents equipped with graph neural networks (GNNs) [26] or other pattern recognition algorithms can identify non-obvious recurring motifs (e.g., common methodological sequences, characteristic mechanistic patterns), complex correlations between disparate knowledge components, or deviations from established patterns (anomalies) that might signify emerging research fronts or inconsistencies requiring further investigation [3]. Such capabilities enable a shift from hypothesis-driven queries to data-driven discovery of latent structures within the entire scientific field as represented by the CNM [27].

- **Agent-Facilitated Generation of Novel Knowledge Artifacts:** A core innovation of the Discovery Engine lies in its support for AI agents to assist in the creation of new scientific knowledge (see Sec. V for hypothesis generation). This moves beyond mere exploration to active construction. For example:
  - *Analogical Transfer:* An "Analogy Agent," leveraging both semantic similarity from vector embeddings and structural isomorphism identified within the CNM (potentially inspired by category-theoretic notions of functors [15, 16]), could propose adapting a successful 'Process/Method' from one 'Entity/System' context to another, generating a novel experimental design artifact [18, 22].
  - *Compositional Design:* Users, interacting with agents on a conceptual workbench (further detailed in design-focused documentation for the Discovery Engine platform), can combine existing CNM components (e.g., 'MechanismNode's, 'MaterialNode's) in novel ways. AI agents would provide real-time feedback on the structural validity (adherence to the Universal Concept Scheme, Appendix B) and potential plausibility (based on known constraints or similar existing patterns in the CNM) of these new compositions, assisting in the assembly of new 'KnowledgeArtifactNode' instances representing complex hypotheses or system designs [44].
  - *Targeted Gap Filling:* When a KnowledgeGapNode is identified, AI agents can be tasked to search the CNM for components that, if combined or modified, could plausibly address that gap, effectively proposing research directions [21].

The operational premise is that these AI agents are tools whose reasoning processes (especially those involving graph traversal or component assembly) can be made transparent by virtue of operating on the explicit and verifiable structure of the CNM [45]. Human researchers interact with these agents, guiding their exploration, refining their suggestions, and ultimately validating the new knowledge artifacts generated. This symbiotic human-AI interaction, grounded in a shared, structured understanding of the scientific domain (the CNM), is what the Discovery Engine framework aims to foster, moving scientific inquiry towards a more synthesized, computationally augmented, and generative paradigm [10]. The specifics of user interaction with these agents, including user-configurable agent behaviors

and collaborative interfaces, are elaborated in complementary work focusing on the Discovery Engine platform design.

## V. AUTOMATED GAP ANALYSIS AND HYPOTHESIS GENERATION

Beyond facilitating understanding of existing knowledge, a primary goal of the Discovery Engine is to actively catalyze the discovery of new knowledge by supporting the iterative cycle of scientific inquiry. This cycle can be viewed as a process of refining models of the world based on evidence – starting with prior understanding, gathering informative data, updating beliefs (models), and using that updated understanding to generate new hypotheses and guide further investigation [46]. By leveraging the synthesized structure of the CNM, the platform aims to systematically identify promising research opportunities (areas where belief updating is most needed) and generating scientifically verifiable hypotheses grounded in the current state of knowledge.

### A. Systematic Identification of Knowledge Gaps

The explicit representation of what is known within the CNM simultaneously highlights what is unknown or poorly understood. The DE employs algorithmic techniques to automatically detect various types of knowledge gaps:

*Component Completeness Gaps* arise when specific types of information, defined as essential by the template (e.g., specific parameters, control mechanisms, quantitative performance metrics), are consistently missing for certain classes of systems or concepts represented in the CNM.

*Structural Holes* refer to regions within the knowledge graph where expected connections between related concepts, thematic clusters, or steps in a process are sparse or absent [21]. Identifying these "missing links" points towards unexplored relationships or required intermediate steps.

*Inconsistency Clusters* emerge when the CNM reveals groups of structurally similar components (e.g., experiments using comparable methodologies) that report contradictory findings or support conflicting theoretical interpretations. These represent areas where the current knowledge is contested and requires resolution through further investigation [20].

*Predictive Gaps* can be identified by leveraging the structural regularities or theoretical principles captured within the CNM. For example, based on established patterns or analogies identified through graph analysis or CT-inspired reasoning [15, 26], the system might predict the existence of a certain mechanism or relationship that has not yet been reported in the literature, thus identifying a specific target for empirical confirmation.

By systematically surfacing these diverse types of gaps, the DE provides researchers with a data-driven map of the scientific frontier, highlighting areas where new research is most needed or likely to be impactful.

## **B. Principled Hypothesis Generation**

The identification of gaps and structural patterns within the CNM serves as a foundation for generating novel, testable hypotheses. The DE aims to move beyond simple correlation-finding towards true abductive inference: generating hypotheses grounded in mechanistic understanding and structural analogy:

*Bridging Gaps with Existing Knowledge:* Hypotheses can be formulated to directly address identified gaps. For instance, if a crucial parameter is missing for a class of systems, the hypothesis might involve designing an experiment to measure it. If a structural hole exists between two related concepts, the hypothesis might propose a specific mechanism linking them. Specific informative experiments can be proposed, either heuristically (e.g. choosing experimental parameters not yet empirically investigated), or more formally (e.g. with an explicit information gain term).

*Exploiting Structural Analogies:* By identifying non-obvious structural similarities between components from different thematic clusters or even different disciplines represented within the CNM (e.g., using graph isomorphism techniques [26] or CT functors [15]), the DE can generate hypotheses based on analogy. For example, it might suggest that a successful control strategy [47]) from one system type could be adapted to address a limitation in a structurally analogous system from a different cluster [18, 22]. This structured analogy finding is potentially more powerful than relying solely on semantic similarity.

*Resolving Conflicts through Synthesis:* Identified inconsistencies can spur hypotheses that propose novel mechanisms, theories, or experimental conditions capable of reconciling the conflicting observations represented within the CNM.

Noteworthy, hypotheses generated via these mechanisms are not unconstrained speculations but are directly derived from, and justifiable by, the synthesized structure of existing scientific knowledge captured within the CNM. This grounding aims to increase the relevance and potential impact of computationally generated research suggestions, positioning the DE as a genuine tool for augmenting scientific creativity and strategic research planning.

## **VI. IMPLEMENTATION, VALIDATION, AND FUTURE VISION**

The conceptual framework of the Discovery Engine requires practical implementation and validation to demonstrate its feasibility and utility. This section briefly outlines the implementation approach, summarizes validation strategies, and discusses the broader vision for the DE as an evolving infrastructure for scientific knowledge.

### **A. Case studies**

## **VII. CASE STUDIES: APPLYING THE DISCOVERY ENGINE FRAMEWORK**

To illustrate the capabilities and potential impact of the Discovery Engine framework, we present two distinct case studies. The first demonstrates the Discovery Engine’s application to synthesize a nascent scientific field and collaboratively construct a forward-looking perspective. The second highlights how the principles underlying the Discovery Engine were meta-applied to assist in the conceptual design of an interactive platform for knowledge exploration itself.

### **A. Case Study 1: Synthesizing and Shaping the Field of Intelligent Soft Matter**

The emerging field of "Intelligent Soft Matter" lies at the dynamic intersection of materials science, physics, chemistry, biology, and cognitive science. It aims to create materials with life-like cognitive capabilities such as perception, learning, memory, and adaptive decision-making, moving beyond traditional passive or simply responsive materials [48]. Given its interdisciplinary nature and rapid evolution, this field presents an ideal testbed for the Discovery Engine’s ability to synthesize knowledge and identify a consensual research trajectory.

An initial corpus of key publications relevant to intelligent soft matter was processed using the Discovery Engine methodology (Sec. II). This involved:

1. **Initial Template Design and Distillation:** A preliminary extraction template, based on the Universal Concept Schema (Appendix B) and tailored with probes specific to soft matter and embodied intelligence, was used to guide LLMs in distilling knowledge artifacts from the selected papers.
2. **Iterative Template Refinement and CNM Construction:** A diverse group of researchers active in fields contributing to intelligent soft matter engaged with the initial distilled components and the template itself. Through a series of iterative feedback cycles, managed within a collaborative environment built on Discovery Engine principles, the template was refined. Specialized AI agents, trained on subsets of the evolving Conceptual Nexus Model pertaining to specific material classes (e.g., hydrogels, LCEs) or mechanisms (e.g., self-organization, memory encoding), assisted in identifying ambiguities in the template or inconsistencies in the extracted data. For example, an LLM agent trained on materials papers might flag that the template inadequately captured parameters related to swelling kinetics critical for adaptive responses, prompting a template revision.
3. **CNM Synthesis and Gap Analysis:** As the template converged towards a consensus representation reflecting the shared understanding of the involved experts and the literature corpus, a Conceptual Nexus Model for intelligent soft matter was constructed. AI agents then analyzed this CNM to identify overarching themes (using methods similar to those in Sec. III B), key conceptual hubs, and critical knowledge gaps (e.g., lack of robust methods for quantifying material 'learning').
4. **Collaborative Perspective Generation:** The synthesized CNM, along with the identified themes and gaps, served as the structured foundation for a collaborative effort involving domain experts and AI. This led to the generation of a forward-looking perspective on the field, outlining key challenges, promising research directions, and a conceptual roadmap for realizing materials with true intelligent behavior [48]. The AI agents assisted in drafting sections based on specific CNM subgraphs, ensuring claims were grounded in the synthesized evidence, and helping to maintain consistency across

contributions from multiple human authors.

This case study demonstrates how the Discovery Engine, through its iterative, AI-assisted, and collaborative approach to template refinement and knowledge synthesis, can facilitate the consolidation of an emerging scientific field. It transformed an array of individual publications into a structured Conceptual Nexus Model that not only represents existing knowledge but also serves as a generative substrate for defining the field’s future trajectory. The resulting perspective [48] showcases a collaboratively constructed understanding, richer and more systematically grounded than what might emerge from individual expert reviews alone.

## **B. Case Study 2: AI-Assisted Design of the DE Platform Interaction Framework**

The principles of structured knowledge synthesis and agent-assisted generation inherent in the DE framework were meta-applied to inform the conceptual design of the DE platform itself—the interactive environment for human-AI collaboration detailed in complementary work [49].

- 1. Distillation of HCI and KG Interaction Literature:** A corpus of relevant research papers focusing on human-computer interaction (HCI) for complex data, knowledge graph visualization, explainable AI (XAI), and collaborative systems (including those cited throughout this manuscript, e.g., [43, 44, 50–54]) was processed using an early version of the Discovery Engine distillation pipeline. The extraction template focused on identifying core interaction principles, user interface patterns, described user challenges, proposed AI assistance roles, and evaluation methodologies.
- 2. Synthesis of Design Principles and Feature Requirements:** The distilled components were synthesized into a CNM focused on "KG Interaction Design." AI agents were then used to analyze this specialized CNM.
  - *Pattern Identification:* Agents identified recurring successful interaction patterns (e.g., multi-modal views, focus+context, provenance tracking) and common usability challenges (e.g., query complexity, information overload).
  - *Principle Extraction:* Based on this analysis, core design principles for the DE platform were formulated (e.g., "Support Multi-Modal Exploration," "Ensure

Verifiability and Provenance," "Facilitate Human-AI Co-Creation," "Manage Cognitive Load").

3. **AI-Assisted Generation of Platform Concepts and UI Mockups:** Working from these principles and the synthesized interaction patterns, designers collaborated with generative AI tools (LLMs prompted with the distilled requirements and design principles). This collaboration yielded conceptual designs for key platform modules, such as the multi-modal exploration interface, the "Knowledge Card" summaries, and the "Hypothesis Workbench." AI was used to generate initial textual descriptions of module functionalities and even draft visual concepts for the user interface (an example conceptual UI schema generated through this process is shown in Fig. 4).

4. **Iterative Refinement:** These AI-generated concepts were then iteratively refined by human designers and HCI experts, ensuring alignment with user needs and established HCI best practices. This mirrored the ‘process.md’ workflow where AI provides initial drafts that humans then curate and enhance.

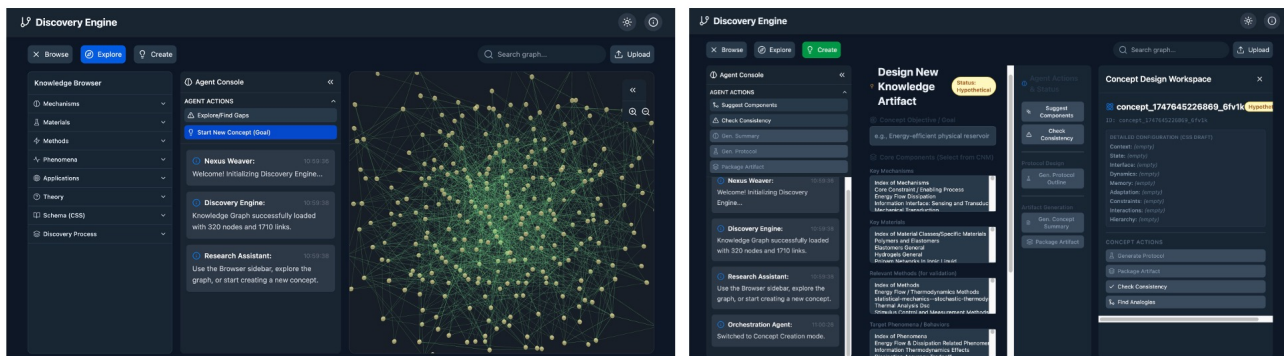


Figure 4. User Interface (UI) for the DE platform, fully generated through an AI-assisted design process. The core modules (e.g., graph visualization, knowledge browser, agent interaction panel, hypothesis workbench) and their relationships were initially outlined by DE (with Gemini Pro 2.5) synthesizing best practices from HCI and KG interaction literature, then refined by a human.

The Discovery Engine (DE) frontend is implemented as an interactive React/TypeScript application (see <https://github.com/ActiveInferenceInstitute/Research-Discovery-Engine>). It processes Markdown (.md) files from a KG/ directory via a `cnmBuilder.ts` utility to construct a client-side graph representation—the Conceptual Nexus Model (CNM). Key

UI components include: an AgentConsole for simulated agent interactions and workflow guidance; a GraphVisualization module using 3d-force-graph for exploring the CNM; a KnowledgeBrowserSidebar for navigating the content hierarchy of the source .md files; and a NodeView for displaying detailed content of selected knowledge graph nodes. A Concept-Designer module allows users to define new hypothetical SystemNode configurations based on components from the CNM, with future integration points for AI-driven suggestions and validation. The current system demonstrates the client-side parsing, graph construction, and interactive visualization pipeline, with persistent knowledge artifacts intended to be generated from the design process. While advanced LLM-driven extraction and algorithmic synthesis (like topic modeling) are conceptualized as backend processes in the DE system, the frontend focuses on structuring and presenting the user-guided discovery workflow.

### C. The CNM Structure as an Evolution of Scientific Communication

The DE methodology offers a potential pathway to address the systemic challenges facing scientific communication and progress in an era of exponential information growth. By shifting the focus from individual documents to synthesized, verifiable knowledge components organized within a dynamic CNM, it aims to create a more robust, navigable, and computationally accessible representation of collective scientific understanding.

This structured synthesis provides crucial context that is often lost in fragmented literature. It enables researchers to quickly situate new findings, objectively assess consensus and conflict, and identify the implicit structure and trajectory of their field. The ability to compare components across studies based on a consistent schema facilitates deeper understanding than narrative summaries alone can provide.

Its gap analysis and hypothesis generation features directly support the crucial steps of identifying where new evidence is most needed and proposing informative experiments, aligning with formal frameworks of rational inquiry like Bayesian reasoning and potentially Active Inference[46, 55].

The challenges of information overload, reproducibility concerns [2], the proliferation of potentially unreliable AI-generated content, and potentially declining research disruptiveness [1] necessitate a fundamental rethinking of how scientific knowledge is structured, disseminated, and utilized. The traditional narrative publication, often distributed as a static PDF,

faces limitations in this new environment.

The DE methodology and the resulting CNM structure offer a potential evolution, acting as a complementary layer built upon the primary literature. The CNM’s structure directly addresses several key challenges:

Its synthesized nature combats *information overload* by providing structured overviews. Its focus on verifiable components linked to source evidence acts as a filter against low-quality or fabricated content, including potential *LLM spam*. The explicit capture of methods and parameters enhances *reproducibility* assessment. The graph structure facilitates deep *synthesis* beyond individual document summaries. Crucially, the CNM is designed to be inherently machine-readable and structured, aligning perfectly with the principles of *FAIR data* (Findable, Accessible, Interoperable, Reusable) [7, 8]. This computational accessibility is vital for leveraging AI tools effectively and responsibly within the scientific process.

In future similar machine-readable structures can replace narrative papers entirely, however, for past knowledge reflected in numerous PDFs the CNM could become the primary interface for computational analysis, synthesis, and discovery, providing a dynamic, validated, and navigable map derived from the underlying literature. This represents a shift towards a more interconnected, computable, and potentially more efficient scientific knowledge ecosystem.

#### D. Limitations and Future Directions

Significant challenges remain. The DE’s effectiveness is contingent on the continued advancement of LLMs for nuanced scientific text understanding [31, 34] and the careful design and governance of the evolving template. Representing the full complexity of scientific knowledge, including uncertainty, causality, and temporal dynamics, within computationally tractable frameworks requires ongoing research, potentially integrating probabilistic methods [20, 56], Category Theory [15–17, 40, 57–59], or VSA techniques [13, 14]. However, scalability to disciplines with millions of publications poses substantial engineering hurdles [21].

Future work will focus on several key areas: enriching the CNM representation to capture deeper semantic and causal relationships; enhancing the data processing pipeline for greater robustness and efficiency; developing more powerful AI agent capabilities for synthesis, gap analysis, and hypothesis generation, perhaps drawing inspiration from Active

Inference [10, 60] or multi-agent systems [9, 42, 44, 45, 61]; and creating intuitive, interactive interfaces for human researchers to collaborate with the Discovery Engine [52–54, 62–64]. A long-term vision includes integrating the DE with automated experimental platforms or simulation engines, creating a closed loop from knowledge synthesis to hypothesis generation to empirical testing, thereby truly accelerating the cycle of scientific discovery.

## VIII. CONCLUSION

The Discovery Engine framework addresses the urgent need for more effective mechanisms to synthesize and navigate the ever-expanding body of scientific knowledge. Our methodology moves beyond document-centric approaches by first distilling publications into verifiable, structured knowledge artifacts using LLMs guided by adaptive templates. The core innovation lies in encoding this rich, heterogeneous information into a unified Conceptual Nexus Tensor ( $T_{\text{CNM}}$ ). This tensor serves as a compact, machine-operable "World Model" of a scientific domain, capturing complex interdependencies within its high-dimensional structure.

From this central tensorial representation, human-interpretable views, such as explicit knowledge graphs (the CNM graph) and semantic vector spaces, can be dynamically generated, enabling researchers to explore the synthesized landscape. More profoundly,  $T_{\text{CNM}}$  provides a substrate for specialized AI agents to perform complex reasoning, identify structural patterns, uncover latent analogies, and pinpoint knowledge gaps through abstract mathematical operations. These agents then collaborate with human researchers, assisting in the construction of novel knowledge artifacts—hypotheses, experimental designs, or theoretical models—that are grounded in the synthesized evidence contained within the tensor.

While the full realization of dynamic tensorial knowledge bases presents significant research and engineering challenges, the Discovery Engine framework outlines a principled pathway. By transforming disparate scientific narratives into a structured, computable tensor, and by enabling AI-assisted interaction with this representation, the Discovery Engine aims to create a generative ecosystem for scientific discovery.

By revealing the topology of knowledge such as clusters, connections, and, significantly, its gaps and inconsistencies, the Discovery Engine aims to transform discovery from a process often reliant on serendipity into a more systematic exploration guided by the structure of what is known and unknown in a more informed way. It provides a framework for

identifying non-obvious connections, formulating targeted questions, and generating novel hypotheses grounded in the synthesized evidence. While challenges remain in implementation, scalability, and validation, the Discovery Engine represents a powerful vision for augmenting human intellect, fostering deeper understanding, and accelerating the pace of scientific breakthroughs in an increasingly complex world. It endeavors to build not just better search engines, but genuine engines of discovery, potentially reshaping the future of scientific communication and collaboration.

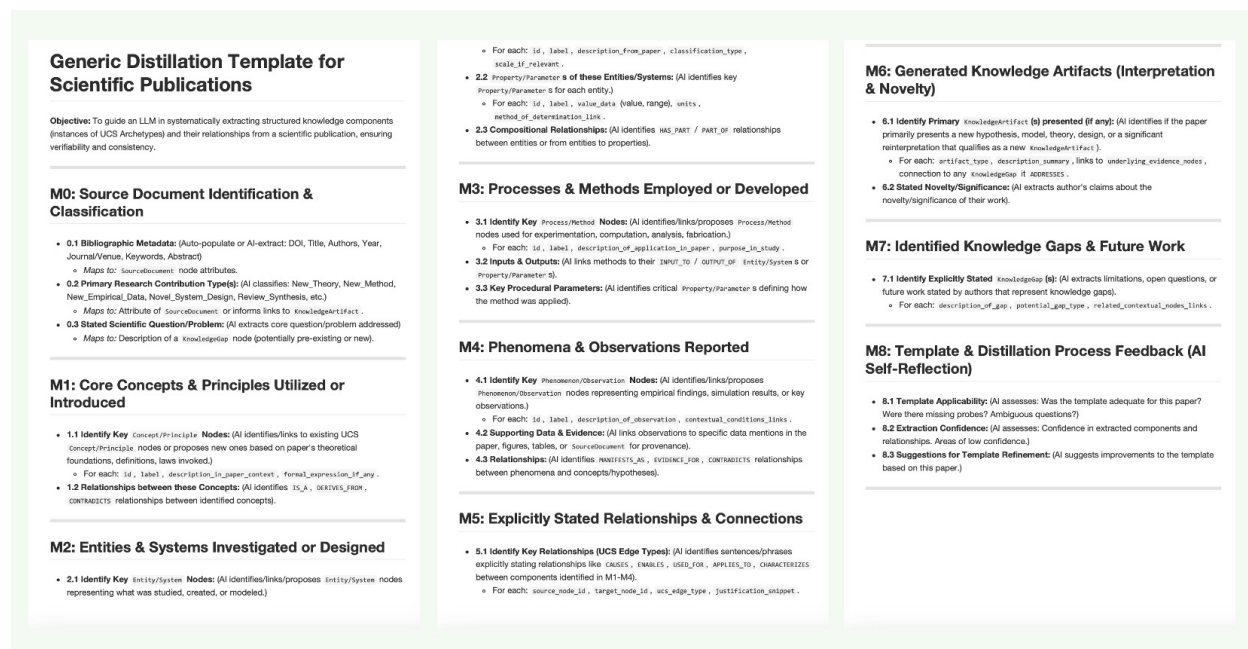


Figure 5. Conceptual architecture of the Distillation Template.

## Appendix A: General Structure of a Distillation Template

This template is designed to be a structured set of instructions and placeholders used by an AI agent (specifically a Large Language Model - LLM) to systematically extract key information from a single scientific publication. The goal is to transform the unstructured (or semi-structured) text of a paper into a collection of discrete, well-defined "knowledge components" and their relationships. These extracted components are then used to populate your Conceptual Nexus Model (CNM), which is built according to the Universal Concept Schema (UCS).

The template is divided into major modules (*e.g.*, M0: Meta-Information, M1: Core

Claims, M2: Entities & Systems). This organization helps break down the complex task of analyzing a whole paper into more manageable sub-tasks for the LLM. It also ensures that different categories of information are systematically considered. Within each module, there are specific "probes" or fields. These are direct instructions or questions to the LLM.

## Appendix B: A Universal Concept Schema for General Scientific Knowledge Representation and Synthesis

Node Type	Definition	Key Attributes	Connections
<b>System Node</b>	Represents an integrated system (experimental, theoretical, etc.)	system_id, system_type, description, design_rationale, domain, primary_focus, scale, status	Links to all other node types
<b>Mechanism Node</b>	Describes fundamental process enabling system functionality	mechanism_id, description, category, timescale_qualitative, energy_domain	Implemented/enabled/underpinned by systems, methods, materials, concepts
<b>Method Node</b>	Technique for fabrication, control, measurement, or modeling	method_id, description, category, output_type	Utilized by systems, publications, artifacts; linked to materials, parameters, metrics
<b>Material Node</b>	Describes a material or class used in a system	material_id, description, category, state_of_matter, key_properties	Used in systems, enabled mechanisms, processed via methods
<b>Parameter Node</b>	Variable or setting defining or characterizing a system	parameter_id, value_range, units, role, source	Constrains systems, mechanisms; measured by methods; quantified by metrics
<b>Phenomenon Node</b>	Observable emergent behavior or property	phenomenon_id, description, scale, emergent_level	Exhibited by systems; underpinned by mechanisms; enables applications; linked to gaps, metrics
<b>Application Node</b>	Functional or technological target of a system	application_id, description	Targeted by systems, phenomena, materials, mechanisms
<b>Theoretical Concept Node</b>	Abstract model or framework for interpretation	concept_id, description, field	Applied by systems, publications; used to model mechanisms or phenomena
<b>Publication Node</b>	Represents a scientific publication	publication_id, title, authors, year, journal, doi, paper_type, key_findings (optional)	Describes other nodes; may suggest gaps or propose artifacts
<b>Metric Node</b>	Quantitative measure of performance or characteristics	metric_id, description, units, optimization_goal	Quantifies parameters or phenomena; used by methods; reported in publications
<b>Energy Type Node</b>	Domain or type of energy required for operation	energy_type_id, description	Required by systems, mechanisms, methods
<b>Knowledge Gap Node</b>	Open question, inconsistency, or missing element in the graph	gap_id, description, gap_type, related_nodes, priority, source	Suggested by publications/agents; addressed by systems, artifacts
<b>Knowledge Artifact Node</b>	Synthesized proposal/design from graph-based exploration	artifact_id, description (implied)	Typically proposes a system and a method sequence to address a gap

Figure 6. CNM Nodes, their definitions and key attributes

The primary objective is to transform disparate scientific information into a dynamic, interconnected network of "knowledge components" that align with FAIR principles, that can represent the analogy of a "world model" for scientific discoveries. This network, far from being a static archive, is designed as an active substrate for computational reasoning, enabling the identification of knowledge gaps and, crucially, facilitating the generation of novel "Knowledge Artifacts". These artifacts may represent new hypotheses, proposed

experimental designs to test specific theories, innovative theoretical models, or conceptual blueprints for new systems or methods. Its instances that populate the CNM.

Edge Type	Definition	Connects
<b>Implements Mechanism</b>	System uses a specific mechanism	System → Mechanism
<b>Utilizes Method</b>	System or publication uses a method	System/Publication → Method
<b>Composed Of Material</b>	System includes a material	System → Material
<b>Exhibits Phenomenon</b>	System exhibits a phenomenon	System → Phenomenon
<b>Targets Application</b>	System or mechanism targets an application	System/Mechanism → Application
<b>Requires Energy</b>	Node requires a specific energy domain	System/Mechanism/Method → Energy Type
<b>Enables Mechanism</b>	Material/method enables a mechanism	Material/Method → Mechanism
<b>Constrained By Parameter</b>	Node is constrained or characterized by a parameter	System/Mechanism/Phenomenon → Parameter
<b>Measures Property</b>	Method measures a property (phenomenon/metric)	Method → Phenomenon/Metric/Parameter
<b>Models System</b>	Method or concept models a system	Method/Concept → System
<b>Validated By</b>	Phenomenon or mechanism is validated by a method or publication	Phenomenon/Mechanism → Method/Publication
<b>Based On Concept</b>	Node is modeled or informed by a concept	Any Node → Theoretical Concept
<b>References Publication</b>	Node cites a publication	Any Node → Publication
<b>Addresses Knowledge Gap</b>	Node addresses an identified gap	System/Artifact/Publication → Gap
<b>Suggests Knowledge Gap</b>	Node identifies or suggests a knowledge gap	Publication/Agent → Gap
<b>Packaged In Artifact</b>	Design or process is encapsulated in an artifact	System/Method → Artifact
<b>Related To</b>	Indicates general semantic relatedness	Any Node ↔ Any Node
<b>PartOf</b>	Node is a component/subtype of a broader category	Any Node → Parent Node

Figure 7. CNM Edges, their definitions and key attributes

## I. Core Node Archetypes (Conceptual Classes)

We propose a minimal set of fundamental node archetypes to represent the universal building blocks of scientific knowledge. These are conceptual classes; specific instances derived from the literature or generative processes would populate the actual knowledge graph.

## II. Core Relationship Archetypes (Conceptual Edge Types)

These define the fundamental ways in which the node archetypes connect, forming the relational fabric of the scientific knowledge graph.

## III. Interaction between Node Archetypes

The interaction between these archetypes can be visualized as a dynamic graph where the core process involves reasoning over the existing knowledge landscape (comprising Concepts, Entities, Properties, Methods, and Observations) to identify Knowledge Gaps and, critically,

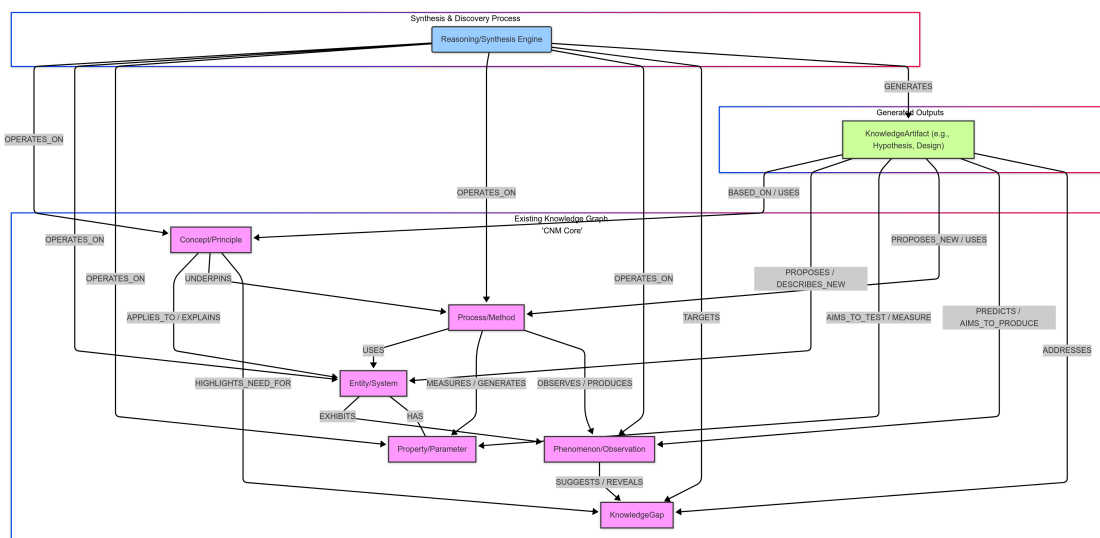


Figure 8. A conceptual representation of the core knowledge graph structure emphasizing the generation of Knowledge Artifacts. Existing knowledge components (Concepts, Entities, etc.) form a network that is analyzed by a Reasoning/Synthesis Engine. This engine, by identifying and targeting Knowledge Gaps, generates new Knowledge Artifacts (e.g., hypotheses, designs, models), which are themselves composed of, based on, or aim to validate existing and new components.

to synthesize novel Knowledge Artifacts designed to address these gaps or explore new frontiers (Fig. 8).

- 
- [1] M. Park, E. Leahey, and R. J. Funk, Papers and patents are becoming less disruptive over time, [Nature](#) **613**, 138 (2023).
  - [2] M. Baker, 1,500 scientists lift the lid on reproducibility, [Nature](#) **533**, 452 (2016).
  - [3] G. Rong, Y. Chen, F. Ma, and T. Koch, [40 Years of Interdisciplinary Research: Phases, Origins, and Key Turning Points \(1981-2020\)](#) (2025).
  - [4] D. Kang, R. S. Danziger, J. Rehman, and J. A. Evans, Limited diffusion of scientific knowledge forecasts collapse, [Nat Hum Behav](#) **9**, 268 (2025), publisher: Nature Publishing Group.
  - [5] A. Tyson and B. Kennedy, [Public Trust in Scientists and Views on Their Role in Policymaking](#), Tech. Rep. (Pew Research Center, 2024).
  - [6] G. Mongillo and M. Tsodyks, [Synaptic Theory of Working Memory for Serial Order](#) (2024).

- [7] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, J. Bouwman, A. J. Brookes, T. Clark, M. Crosas, I. Dillo, O. Dumon, S. Edmunds, C. T. Evelo, R. Finkers, A. Gonzalez-Beltran, A. J. G. Gray, P. Groth, C. Goble, J. S. Grethe, J. Heringa, P. A. C. 't Hoen, R. Hooft, T. Kuhn, R. Kok, J. Kok, S. J. Lusher, M. E. Martone, A. Mons, A. L. Packer, B. Persson, P. Rocca-Serra, M. Roos, R. van Schaik, S.-A. Sansone, E. Schultes, T. Sengstag, T. Slater, G. Strawn, M. A. Swertz, M. Thompson, J. van der Lei, E. van Mulligen, J. Velterop, A. Waagmeester, P. Wittenburg, K. Wolstencroft, J. Zhao, and B. Mons, The FAIR Guiding Principles for scientific data management and stewardship, [Sci Data 3, 160018 \(2016\)](#), publisher: Nature Publishing Group.
- [8] A. Jacobsen, R. De Miranda Azevedo, N. Juty, D. Batista, S. Coles, R. Cornet, M. Courtot, M. Crosas, M. Dumontier, C. T. Evelo, C. Goble, G. Guizzardi, K. K. Hansen, A. Hasnain, K. Hettne, J. Heringa, R. W. Hooft, M. Imming, K. G. Jeffery, R. Kaliyaperumal, M. G. Kersloot, C. R. Kirkpatrick, T. Kuhn, I. Labastida, B. Magagna, P. McQuilton, N. Meyers, A. Montesanti, M. Van Reisen, P. Rocca-Serra, R. Pergl, S.-A. Sansone, L. O. B. Da Silva Santos, J. Schneider, G. Strawn, M. Thompson, A. Waagmeester, T. Weigel, M. D. Wilkinson, E. L. Willighagen, P. Wittenburg, M. Roos, B. Mons, and E. Schultes, FAIR Principles: Interpretations and Implementation Considerations, [Data Intelligence 2, 10 \(2020\)](#).
- [9] M. Wooldridge, 1. WHAT ARE SOFTWARE AGENTS?, IEE Review (1996).
- [10] P. A. Tsividis, J. Loula, J. Burga, N. Foss, A. Campero, T. Pouncy, S. J. Gershman, and J. B. Tenenbaum, [Human-Level Reinforcement Learning through Theory-Based Modeling, Exploration, and Planning \(2021\)](#), version Number: 1.
- [11] S. T. Piantadosi, D. C. Muller, J. S. Rule, K. Kaushik, M. Gorenstein, E. R. Leib, and E. Sanford, Why concepts are (probably) vectors, [Trends in Cognitive Sciences 28, 844 \(2024\)](#).
- [12] N. Kriegeskorte, M. Mur, and P. A. Bandettini, Representational similarity analysis - connecting the branches of systems neuroscience, [Front. Syst. Neurosci. 2, 10.3389/neuro.06.004.2008 \(2008\)](#), publisher: Frontiers.
- [13] D. Kleyko, M. Davies, E. P. Frady, P. Kanerva, S. J. Kent, B. A. Olshausen, E. Osipov, J. M. Rabaey, D. A. Rachkovskij, A. Rahimi, and F. T. Sommer, Vector Symbolic Architectures as a Computing Framework for Emerging Hardware, [Proc. IEEE 110, 1538 \(2022\)](#), arXiv:2106.05268 [cs].

- [14] P. S. Jr, D. Summers-Stay, and Y. Aloimonos, [A Computational Theory for Life-Long Learning of Semantics](#) (2018), arXiv:1806.10755 [cs].
- [15] S. Phillips, What is category theory to cognitive science? Compositional representation and comparison, [Front. Psychol.](#) **13**, 1048975 (2022).
- [16] Y. Yuan, [A Categorical Framework of General Intelligence](#) (2023), arXiv:2303.04571 [cs].
- [17] W. Pan, [Token Space: A Category Theory Framework for AI Computations](#) (2024), arXiv:2404.11624 [math].
- [18] H. Lu, Y. N. Wu, and K. J. Holyoak, Emergence of analogy from relation learning, [Proceedings of the National Academy of Sciences](#) **116**, 4176 (2019), publisher: Proceedings of the National Academy of Sciences.
- [19] D. A. Friedman, [FieldSHIFT-2: Fully synthetic dissertations for all-by-all shifted domains](#) (2024).
- [20] D. Sejdinovic, A. Gretton, B. Sriperumbudur, and K. Fukumizu, [Hypothesis testing using pairwise distances and associated kernels \(with Appendix\)](#) (2012), arXiv:1205.0411 [cs].
- [21] M. J. Buehler, [Agentic Deep Graph Reasoning Yields Self-Organizing Knowledge Networks](#) (2025), arXiv:2502.13025 [cs].
- [22] T. O'Brien, J. Stremmel, L. Pio-Lopez, P. McMillen, C. Rasmussen-Ivey, and M. Levin, Machine learning for hypothesis generation in biology and medicine: exploring the latent space of neuroscience and developmental bioelectricity, [Digital Discovery](#) **3**, 249 (2024), publisher: RSC.
- [23] S. Lloyd, Ultimate physical limits to computation, [Nature](#) **406**, 1047 (2000), publisher: Nature Publishing Group.
- [24] E. P. Wigner, The Unreasonable Effectiveness of Mathematics in the Natural Sciences, in [Philosophical Reflections and Syntheses](#), edited by J. Mehra (Springer, Berlin, Heidelberg, 1995) pp. 534–549.
- [25] E. O. Buzbas, B. Devezer, and B. Baumgaertner, The logical structure of experiments lays the foundation for a theory of reproducibility, [R. Soc. open sci.](#) **10**, 221042 (2023).
- [26] M. J. Buehler, [Graph-Aware Isomorphic Attention for Adaptive Dynamics in Transformers](#) (2025), arXiv:2501.02393 [cs].
- [27] M. J. Buehler, [Self-Organizing Graph Reasoning Evolves into a Critical State for Continuous Discovery Through Structural-Semantic Dynamics](#) (2025), arXiv:2503.18852 [cs].

- [28] E. Margolis and S. Laurence, Concepts, in *The Blackwell Guide to Philosophy of Mind*, edited by S. P. Stich and T. A. Warfield (Wiley, 2003) 1st ed., pp. 190–213.
- [29] J. A. Fodor and Z. W. Pylyshyn, Connectionism and cognitive architecture: A critical analysis, *Cognition* **28**, 3 (1988).
- [30] T. S. Kuhn, *The structure of scientific revolutions*, 2nd ed., International encyclopedia of unified science No. 2,2 (Univ. of Chicago Press, Chicago, 1994).
- [31] A. Rogers, O. Kovaleva, and A. Rumshisky, A Primer in BERTology: What We Know About How BERT Works, *Transactions of the Association for Computational Linguistics* **8**, 842 (2021).
- [32] A. Novikov, N. Vu, M. Eisenberger, E. Dupont, P.-S. Huang, A. Z. Wagner, S. Shirobokov, B. Kozlovskii, F. J. R. Ruiz, A. Mehrabian, M. P. Kumar, S. Chaudhuri, G. Holland, A. Davies, S. Nowozin, P. Kohli, and M. Balog, AlphaEvolve: A coding agent for scientific and algorithmic discovery (2025).
- [33] E. P. Frady, D. Kleyko, C. J. Kymn, B. A. Olshausen, and F. T. Sommer, Computing on Functions Using Randomized Vector Representations (in brief), in *Neuro-Inspired Computational Elements Conference* (ACM, Virtual Event USA, 2022) pp. 115–122.
- [34] W. Lu, R. K. Luu, and M. J. Buehler, Fine-tuning large language models for domain adaptation: exploration of training strategies, scaling, model merging and synergistic capabilities, *npj Comput Mater* **11**, 1 (2025), publisher: Nature Publishing Group.
- [35] I. Balazevic, C. Allen, and T. Hospedales, TuckER: Tensor Factorization for Knowledge Graph Completion, in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, edited by K. Inui, J. Jiang, V. Ng, and X. Wan (Association for Computational Linguistics, Hong Kong, China, 2019) pp. 5185–5194.
- [36] H. Ren, M. Galkin, M. Cochez, Z. Zhu, and J. Leskovec, *Neural Graph Reasoning: Complex Logical Query Answering Meets Graph Databases* (2023), version Number: 1.
- [37] M. M. Bronstein, J. Bruna, T. Cohen, and P. Veličković, *Geometric Deep Learning: Grids, Groups, Graphs, Geodesics, and Gauges* (2021), arXiv:2104.13478 [cs].
- [38] M. J. Healy and T. P. Caudell, Episodic memory: A hierarchy of spatiotemporal concepts, *Neural Networks* **120**, 40 (2019).

- [39] M. Grootendorst, [BERTopic: Neural topic modeling with a class-based TF-IDF procedure](#) (2022).
- [40] P. Koupil and I. Holubová, A unified representation and transformation of multi-model data using category theory, [J Big Data](#) **9**, 61 (2022).
- [41] R. N. Shepard and S. Chipman, Second-order isomorphism of internal representations: Shapes of states, [Cognitive Psychology](#) **1**, 1 (1970).
- [42] M. Besta, F. Memedi, Z. Zhang, R. Gerstenberger, G. Piao, N. Blach, P. Nyczyk, M. Copik, G. Kwaśniewski, J. Müller, L. Gianinazzi, A. Kubicek, H. Niewiadomski, A. O’Mahony, O. Mutlu, and T. Hoefler, [Demystifying Chains, Trees, and Graphs of Thoughts](#) (2025), arXiv:2401.14295 [cs].
- [43] J. Sequeda, D. Allemang, and B. Jacob, Knowledge Graphs as a source of trust for LLM-powered enterprise question answering, [Journal of Web Semantics](#) **85**, 100858 (2025).
- [44] Z. He, N. Xiong, H. Li, P. Shen, G. Zhu, and L. Zhang, [The two-way knowledge interaction interface between humans and neural networks](#) (2024), arXiv:2401.05461 [cs].
- [45] E. Manzoor, J. Tong, S. Vijayaraghavan, and R. Li, [Expanding Knowledge Graphs with Humans in the Loop](#) (2023), arXiv:2212.05189 [cs].
- [46] C. Howson and P. Urbach, *Scientific reasoning: The Bayesian approach*, 3rd ed. (Open Court, Chicago, IL, 2006).
- [47] W. Bechtel and L. Bich, Grounding cognition: heterarchical control mechanisms in biology, [Phil. Trans. R. Soc. B](#) **376**, 20190751 (2021).
- [48] V. A. Baulin, A. Giacometti, D. A. Fedosov, S. Ebbens, N. R. Varela-Rosales, N. Feilich, M. Chowdhury, M. Hu, R. Fuchsli, M. Dijkstra, M. Mussel, R. Van Roij, D. Xie, V. Tzanov, M. Zu, S. Hidalgo-Caballero, Y. Yuan, L. Cocconi, C.-M. Ghim, C. Cottin-Bizonne, M. C. Miguel, M. J. Esplandiu, J. Simmchen, W. J. Parak, M. Werner, G. Gompfer, and M. M. Hanczyc, Intelligent soft matter: towards embodied intelligence, [Soft Matter](#) **10.1039/D5SM00174A** (2025).
- [49] J. Lumiruusu, D. Friedman, S. Rahman, V. Baulin, and A. Pashea, [ResNei: Solution Design Document](#) (2025), publisher: Active Inference Institute.
- [50] H. Li, G. Appleby, C. D. Brumar, R. Chang, and A. Suh, Knowledge Graphs in Practice: Characterizing their Users, Challenges, and Visualization Opportunities, [IEEE Trans. Visual. Comput. Graphics](#) **30**, 584 (2024), arXiv:2304.01311 [cs].

- [51] B. Sarrafzadeh, A. Roegiast, and E. Lank, [Hierarchical Knowledge Graphs: A Novel Information Representation for Exploratory Search Tasks](#) (2020), arXiv:2005.01716 [cs] version: 1.
- [52] S. Zhang, Z. Wang, C. Chen, Y. Dai, L. Ye, and X. Sun, Patterns for Representing Knowledge Graphs to Communicate Situational Knowledge of Service Robots, in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (2021) pp. 1–12, arXiv:2101.10602 [cs].
- [53] S. Meier and K. Glinka, To Classify is to Interpret: Building Taxonomies from Heterogeneous Data through Human-AI Collaboration, in *Mensch und Computer 2023* (2023) pp. 395–401, arXiv:2307.16481 [cs].
- [54] S. Rahman, F. Choi, H. Kim, D. Zhang, and E. Hruschka, [Knowledge Acquisition and Integration with Expert-in-the-loop](#) (2024), arXiv:2402.03291 [cs].
- [55] K. J. Friston, T. FitzGerald, F. Rigoli, P. Schwartenbeck, and G. Pezzulo, Active inference: a process theory, *Neural computation* **29**, 1 (2017).
- [56] D. Sejdinovic, B. Sriperumbudur, A. Gretton, and K. Fukumizu, Equivalence of distance-based and RKHS-based statistics in hypothesis testing, *The Annals of Statistics* **41**, 2263 (2013), publisher: Institute of Mathematical Statistics.
- [57] J. Hefford, V. Wang, and M. Wilson, [Categories of Semantic Concepts](#) (2020), arXiv:2004.10741 [cs] version: 1.
- [58] E. Purvine, C. Joslyn, and M. Robinson, [A Category Theoretical Investigation of the Type Hierarchy for Heterogeneous Sensor Integration](#) (2016), arXiv:1609.02883 [math].
- [59] L. Christino, S. Rezaeipourfarsangi, E. Milios, and F. V. Paulovich, [A Theoretical Approach for Structuring and Analysing Knowledge Provenance for Visual Analytics](#) (2023), arXiv:2204.00585 [cs].
- [60] A. Ciaunica, M. Levin, F. E. Rosas, and K. Friston, Nested Selves: Self-Organization and Shared Markov Blankets in Prenatal Development in Humans, *Topics in Cognitive Science* **00**, 1 (2023), \_eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/tops.12717>.
- [61] X. Zhao, M. Blum, R. Yang, B. Yang, L. M. Carpintero, M. Pina-Navarro, T. Wang, X. Li, H. Li, Y. Fu, R. Wang, J. Zhang, and I. Li, [AGENTiGraph: An Interactive Knowledge Graph Platform for LLM-based Chatbots Utilizing Private Data](#) (2024), arXiv:2410.11531 [cs].

- [62] H. Li, G. Appleby, and A. Suh, [A Preliminary Roadmap for LLMs as Assistants in Exploring, Analyzing, and Visualizing Knowledge Graphs](#) (2024), arXiv:2404.01425 [cs].
- [63] B. Sarrafzadeh, A. Vtyurina, E. Lank, and O. Vechtomova, Knowledge Graphs versus Hierarchies: An Analysis of User Behaviours and Perspectives in Information Seeking, in *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*, CHIIR '16 (Association for Computing Machinery, New York, NY, USA, 2016) pp. 91–100.
- [64] M. G. Skjæveland, K. Balog, N. Bernard, W. Łajewska, and T. Linjordet, An Ecosystem for Personal Knowledge Graphs: A Survey and Research Roadmap, *AI Open* **5**, 55 (2024), arXiv:2304.09572 [cs].