Toward Imaginative AI: Domain Transitivity for Cross-Domain Imagination

Virach Sornlertlamvanich [a,1],
Pinphong Ruangraweenukit [b], Ponlawat Chophuk [b], Waranrach Viriyavi [b] ,
and Thatsanee Charoenporn [a]
[a] *Asia AI Institute (AAII), Faculty of Data Science, Musashino University, Japan.*
[b] *Faculty of Informatics, Burapha University, Thailand.*
ORCID ID: Virach Sornlertlamvanich https://orcid.org/0000-0002-6918-8713,
Thatsanee Charoenporn https://orcid.org/0000-0002-9577-9082
64160205@go.buu.ac.th, ponlawat.ch@go.buu.ac.th, waranrach.vi@go.buu.ac.th

Abstract. Current large language models (LLMs) excel at pattern generation but lack explicit reasoning mechanisms, which limits genuine understanding, and creative knowledge formation, often leading to hallucinations and confinement to learned data distributions. This paper proposes a novel imaginative AI framework that integrates three core components to advance beyond generative AI: (1) the Semantically Aware Reasoning (SAR) model, which couples LLMs with a logical reasoning layer to improve reliability; (2) the Five States of Mind Enlightenment, formulated as a Buddhist-inspired computational framework (data, knowledge, consciousness, awareness, and enlightenment) that models metacognitive progression from perception to insight creation that supports structured imagination through latent concept transformation; and (3) a cross-domain imagination mechanism that projects latent talent structures between domains using cross-SVD (singular value decomposition) to generate novel ideas. Together, these elements form a structured pathway toward a computational framework enabling reasoned and structured cross-domain knowledge synthesis. The framework addresses key limitations of existing generative AI and outlines a research direction for achieving reliable and imaginative AI.

Keywords. Imaginative AI, Conscious AI, Five States of Mind Enlightenment, Semantically Aware Reasoning (SAR), Cross-Domain Imagination, Cross-SVD, Large Language Models

1. Introduction

Artificial intelligence (AI) systems capable of structured knowledge synthesis beyond pattern generation remain a key research challenge. While modern AI systems achieve strong performance in prediction and generation tasks, their ability to produce novel and meaningful concepts remains limited. Contemporary deep learning-based AI has demonstrated remarkable progress toward realizing this vision, achieving capabilities

---

[1] Corresponding Author: Virach Sornlertlamvanich, Asia AI Institute (AAII), Faculty of Data Science, Musashino University, 3-3-3 Ariake Koto-ku, Tokyo, 135-8181, Japan; E-mail: virach@musashino-u.ac.jp

once thought unsolvable. These systems enhance human abilities in areas such as visual recognition, classification, natural language generation, memorization, summarization, and even natural human-computer interaction. However, their operation is largely driven by the consumption of massive datasets through deep learning modeling. Internally, the underlying mechanism relies on predictive modeling from training data, generating outputs along statistically probable pathways derived from learned patterns [1]. While sophisticated fine-tuning and architectural innovations—such as long-term reference mechanisms and forgetting strategies—can mitigate issues related to gradient descent and improve information retention, these remain fundamentally data-dependent approaches [2]. Notably, the attention mechanism in Transformers has proven highly effective in resolving long-standing reference ambiguity problems in Natural Language Processing (NLP) by enabling dynamic focus on relevant contextual information.

The remarkable success of large language models (LLMs) has reshaped the landscape of artificial intelligence, demonstrating unprecedented capabilities in natural language generation, translation, and task completion [3]. However, their operational core remains fundamentally generative and statistical. They predict the next most probable token based on patterns learned from vast training corpora [4]. As noted in critiques, LLMs “lack explicit mechanisms for structured reasoning over grounded knowledge representations” [5]. This limitation leads to well-documented issues such as hallucinations, logical inconsistencies, and limited generalization beyond training distributions.

Generative AI, in its current form, operates analogously to System 1 thinking as described by Kahneman (2011)—fast, intuitive, and pattern-matching. What is missing is the complementary System 2—slow, deliberate, and logical reasoning [6]. The Semantically Aware Reasoning (SAR) model was proposed to address this gap by coupling an LLM (System 1) with a structured knowledge base and a reasoning layer (System 2), thereby enhancing answer reliability and factual grounding [7]. While SAR mitigates hallucinations and improves factual accuracy, it still operates within the bounds of existing knowledge, limiting its capacity for novel idea creation.

Beyond the current study of cognitive frameworks, there has been growing interest in incorporating insights from Buddhist philosophy into AI research. Concepts such as *sati* (meta-knowledge or mindfulness; that is “knowing what one knows”) and *sampajañña* (clear comprehension; that is “knowing how to get access to the knowledge”) emphasize awareness of mental states and intentional action, aligning naturally with metacognitive and reflective AI architectures [17].

Recent studies in Buddhist AI explore how mindfulness-inspired mechanisms can support self-monitoring, ethical awareness, and intentional decision-making in artificial agents. Rather than treating cognition as a purely computational process, these approaches emphasize the transformation of knowledge into wisdom through awareness and context sensitivity, thereby enhancing the capabilities of LLMs.

The present work extends this line of research by formalizing a Five States of Mind Enlightenment, mapping philosophical concepts onto concrete computational stages, that is, from data and knowledge to consciousness, awareness, and ultimately knowledge creation. This framework provides a conceptual scaffold for imaginative AI to reach the stage of mind enlightenment, grounding cross-domain inspiration in both cognitive science and contemplative epistemology.

True human-like creativity and innovation often arise from cross-domain inspiration—the ability to transfer insights, patterns, or talents from one field (e.g., music, art) to solve problems or generate ideas in another (e.g., engineering, cuisine) [8].

This process is not merely an extension of learned data but an expansion into conceptually adjacent yet unexplored territories.

This paper proposes a computational framework for cross-domain knowledge synthesis. The key idea is to model imagination as a structured transformation of latent concept representations across domains. We define imaginative AI as a system capable of forming novel, feasible, and valuable concepts by integrating semantic reasoning with cross-domain latent concept projection. To ground this process in a coherent model of cognitive progression, we introduce the Five States of Mind Enlightenment (data, knowledge, consciousness, awareness, and enlightenment) as a framework for modeling the AI's internal progression from raw data perception to the "creation" of new knowledge.

The contributions of this paper are threefold:

1. A formal computational framework for modeling cross-domain imagination through structured latent transformation.
2. The cross-SVD method for interpretable and structure-preserving concept projection.
3. A cross-domain imaginative model that uses matrix factorization e.g., cross-SVD (singular value decomposition) to map and project talent vectors between domains.

This paper is structured as follows: Section 2 reviews related works on reasoning AI, cognitive architectures, and knowledge transfer. Section 3 establishes the framework of the Five States of Mind Enlightenment. Section 4 details the proposed integrated framework for imaginative modelling. Section 5 discusses the formal model for cross-domain imagination. Section 6 exemplifies the cross-domain imagination based on cross-SVD projection. Section 7 concludes with future research directions.

2. Related Works

Numerous efforts have been made to advance current generative AI toward reasoning-based AI. The dominance of transformer-based LLMs has highlighted a critical need for systems capable of reasoning rather than mere pattern reproduction. While techniques such as Chain-of-Thought (CoT) prompting attempt to scaffold logical inference, and Retrieval-Augmented Generation (RAG) enriches models with external knowledge, both remain fundamentally generative in nature [9]. RAG, as argued by Sornlertlamvanich (2025), is essentially a "knowledge-adding process" that retrieves and vectorizes external content for token prediction [7]. However, this process does not inherently understand or reason about the retrieved content; it merely incorporates it into the generative pathway. Consequently, the outputs—while more informed—are still products of statistical generation, and cannot guarantee logical consistency or factual correctness. The Semantically Aware Reasoning (SAR) model addresses this limitation by decoupling intuitive generation (handled by the LLM as System 1) from deliberate verification (performed by a structured knowledge base and reasoning engine as System 2). This dual-process architecture ensures that generated responses are not only fluent but also logically sound and factually validated, moving AI beyond generation toward reasoned understanding.

The emergence of multimodal foundation models has significantly advanced the integration of heterogeneous data modalities. CLIP (Contrastive Language–Image Pretraining) represents a milestone in this direction by jointly learning aligned text and image embedding spaces through contrastive learning [10]. By encoding text using Transformer-based language models (e.g., BERT-like architectures) and images using Vision Transformers (ViT) [11], CLIP enables semantic alignment across modalities without task-specific supervision.

In CLIP, textual descriptions are encoded into dense semantic vectors using Transformer-based text encoders, while visual inputs are encoded using ViT architectures that model images as sequences of patches. The contrastive objective enforces that semantically corresponding text–image pairs are mapped close to each other in the shared embedding space. This alignment has been shown to support zero-shot classification, cross-modal retrieval, and flexible semantic reasoning.

From the perspective of talent modeling, CLIP-style multimodal embeddings provide a powerful mechanism to represent achievements expressed across different modalities, such as textual descriptions, visual artifacts, or symbolic representations. In this work, such multimodal vectors can serve as inputs to Persona–Achievement Matrices (PAM), enabling the aggregation of diverse evidence of expertise into a unified latent talent space. However, while CLIP excels at multimodal alignment, it does not explicitly model concept strength, intentional transfer, or reasoning. Our approach complements CLIP by applying structured decomposition and cross-SVD transformation on top of multimodal representations.

Relating to the cognitive architectures and metacognition, human cognition is often modeled hierarchically. The Five States of Mind Enlightenment offers a Buddhist-inspired framework aligning with cognitive science models: data/information (sensation), knowledge (understanding), consciousness (*sati*) (meta-knowledge, mindfulness), awareness (*sampajañña*) (clear comprehension of action), and enlightenment (novel insight creation). This progression mirrors the cognitive load theory's emphasis on moving from extraneous to germane processing [12] and Baddeley's episodic buffer for integrating information [13]. Integrating such a framework into AI aims to instill a form of metacognition—"knowing what one knows" and "knowing how to get access to the knowledge."

Concerning cross-domain learning and creativity, techniques such as Latent Semantic Analysis (LSA) and related methods have long been employed to uncover latent semantic structures in textual data [14]. In machine learning, transfer learning and domain adaptation allow models to apply knowledge from a source-domain to a target-domain. Our work extends this by focusing not on task performance, but on creative idea generation by projecting the conceptual talent structure of a human expert from one domain to another, using matrix alignment methods like Procrustes analysis.

Latent concept extraction has a long history in information retrieval and representation learning. SVD forms the mathematical foundation of LSA, where high-dimensional co-occurrence matrices are decomposed to uncover latent semantic structures underlying observed data [15,16]. By projecting observations into a lower-dimensional latent space, SVD captures global structural regularities that are not apparent at the surface level. In natural language processing, SVD-based approaches have been used to reveal latent topics, semantic associations, and conceptual similarity beyond explicit word overlap. Unlike neural embeddings learned through backpropagation, SVD yields closed-form, interpretable decompositions, where singular values explicitly encode the relative importance of latent concepts. This property makes SVD particularly suitable for modeling concept strength rather than merely similarity.

Recent work has revisited matrix factorization methods as complementary to deep learning, emphasizing their interpretability and stability. In contrast to stochastic embedding spaces, SVD-based latent spaces preserve linear structure and are amenable to principled transformations, which is essential for controlled cross-domain reasoning. The present work builds upon this line of research by extending SVD into a cross-SVD formulation, where singular value structures are intentionally transformed across domains to enable imagination knowledge creation.

To bridge the gap between generative AI and creativity-oriented AI, a framework for imaginative AI has been modeled. While SAR addresses reliability and RAG addresses knowledge scope, neither facilitates true imagination. Existing cross-domain methods focus on adaptation, not creation. Our framework is the first to combine a reasoning layer (SAR), a cross-domain projection modelling, and a metacognitive framework in mind enlightenment to aim for imaginative AI.

3. The Five States of Mind Enlightenment

The proposed framework of the Five States of Mind Enlightenment is constructed by incorporating selected aspects of recognition derived from Buddhist cognitive theory—specifically perception and recognition (*saññā*), mindfulness and recollection of past events (*sati*), clear comprehension (*sampajañña*), consciousness (*viññāṇa*), and wisdom (*paññā*)—as systematically theorized in Buddhadhamma [17]. The framework is proposed as a modeling tool for representing the thought process and its progressive stages toward clearer cognition, with the specific aim of supporting the explanation and formalization of the proposed computational imaginative model. It does not attempt to present Buddhist cognitive theory in its entirety; rather, it selectively adopts relevant cognitive principles to enhance conceptual clarity and computational interpretability.

Within this framework, the progressive transformation of cognition—hereafter termed cogniton—is described as a process evolving from sensory input to genuine knowledge creation. A cogniton is defined as the minimal functional unit of thought that encapsulates meaning, intentionality, and potential action, and whose degree of activation changes across distinct cognitive stages.

The Five Stages of Mind Enlightenment are defined as a progressive sequence describing the transformation of cognition from raw sensory input to genuine knowledge creation. Each stage represents a distinct mode of cognitive activation and integration, reflecting increasing levels of structure, awareness, and intentionality.

1. Data and Information
   This stage corresponds to the direct sensing of observable phenomena without structured interpretation. Cognitive processing at this level is limited to the reception of stimuli and basic registration of events. Cognitons remain weakly activated and fragmented, as observations have not yet been organized into coherent meaning. This stage provides the raw material upon which higher cognitive processes operate.
2. Knowledge and recognition (*saññā*)
   At the knowledge stage, observed data are organized, interpreted, and structured into meaningful representations. This involves recognizing patterns, relationships, and semantic content, allowing understanding of what has been observed. Cognitons acquire internal structure and stability but remain largely implicit and context-dependent. In computational terms, this stage corresponds to structured knowledge representations such as feature matrices or latent embeddings, which enable reuse of interpreted

information but do not guarantee situational awareness or appropriate application.

3. Consciousness and mindfulness (*sati*)
   Consciousness, aligned with *sati* (mindfulness), represents the explicit awareness of one's current cognitive state. Functionally, this stage is characterized by knowing what one knows. Cognitons are reflexively activated as meta-knowledge, enabling recognition of available internal resources in relation to the present context. *Sati* plays a critical role in selecting relevant knowledge and preventing unreflective or habitual responses. Without this stage, knowledge remains dormant and cannot be effectively mobilized for problem-solving.
4. Awareness and clear comprehension (*sampajañña*)
   Awareness corresponds to *sampajañña*, or clear comprehension, which governs access to and application of activated knowledge. This stage involves understanding how to retrieve, integrate, and apply what is known in a precise and contextually appropriate manner. While *sati* identifies relevant knowledge, *sampajañña* determines the correct method of execution. It ensures alignment between intention, action, and situational constraints, thereby enabling effective and purposeful cognition.
5. Enlightenment and wisdom (*paññā*)
   Enlightenment represents the culmination of the cognitive process, where integrated knowledge gives rise to genuinely new understanding. At this stage, multiple cognitons are synthesized through reflective and imaginative processes, resulting in knowledge creation rather than mere recall or application. Enlightenment thus marks a qualitative transformation of cognition, characterized by insight, coherence, and the emergence of novel conceptual structures that did not previously exist in explicit form.

At the initial levels of data and information, cognitons remain weakly activated, reflecting raw sensory observation without structured interpretation. As cognition advances to the stage of knowledge, cognitons acquire semantic organization and internal structure, yet remain largely implicit and context-dependent. In computational terms, this stage corresponds to knowledge representations such as feature matrices or latent embeddings in machine learning models, which mediate interpreted information into utilizable resources for downstream tasks but do not, by themselves, guarantee situational awareness or purposeful application.

The transition from knowledge to consciousness corresponds to the emergence of *sati* in Buddhist cognitive theory, referring to mindfulness or present-moment awareness. Functionally, *sati* represents the capacity of the system or subject to explicitly recognize the current cognitive state—namely, knowing what one knows. At this stage, cognitons are reflexively activated as meta-knowledge, allowing the subject to survey available internal resources in relation to the present situation. For example, in a daily decision-making scenario such as selecting a payment method at a cashier counter, success depends not on possessing multiple payment options, but on being aware of one's current resource availability (e.g., cash, credit card, mobile payment) and their relevance to the context. This illustrates the operational role of *sati* as a prerequisite for effective action. Contemporary large language models (LLMs), despite their extensive encoded knowledge, generally lack this form of endogenous mindfulness; they do not autonomously identify which internal knowledge states are relevant to a given task without explicit human instruction or prompting, indicating the absence of intrinsic *sati*.

Beyond consciousness lies the stage of awareness, aligned with *sampajañña*, or clear comprehension. Whereas *sati* concerns recognition of available knowledge, *sampajañña* governs access—the ability to retrieve, integrate, and apply cognitons appropriately and efficiently. Continuing the cashier example, even after selecting a payment method, successful execution requires understanding the correct operational procedure (e.g., swiping, inserting, or tapping a card depending on the terminal). *sampajañña* thus enables procedural alignment between knowledge and action, effectively answering the question of how to reach what one knows. In the context of LLMs, this limitation is likewise evident: even when relevant information is identified through prompting (an external proxy for *sati*), models still require carefully structured guidance to apply that information correctly. This indicates that current LLMs do not autonomously perform *sampajañña* and instead rely on human intelligence to supply both awareness and execution pathways.

Prior to the final stage of enlightenment, cognition is characterized by an intensive imagination process in which multiple hypothetical interpretations, analogies, and transformations of cognitons are generated and evaluated. This imagination activity is not purely stochastic; rather, it is constrained by internal consistency, self-understanding, and coherence with perceived reality or existence. Only those imagination constructions that satisfy these constraints stabilize and persist within the cognitive system. Enlightenment, therefore, is not merely the culmination of accumulated understanding, but the emergence of genuinely new knowledge through the successful integration and survival of coherent imagination structures. In this sense, Enlightenment represents the transformation of cognition from recognition and access into knowledge creation, marking a qualitative shift beyond retrieval-based intelligence toward creative and self-consistent cognitive synthesis.

4. Proposed Integrated Framework for Imaginative model Creation

To operationalize imagination intelligence within large language models while maintaining epistemic reliability, we propose an integrated, multi-layered architectural framework that unifies perception, reasoning, metacognition, and creative generation. The framework is grounded in a progressive state transition—from data acquisition to knowledge construction, reflective reasoning, and ultimately novel concept creation—aligning computational processes with cognitive and contemplative theories. By embedding Semantically Aware Reasoning (SAR) as a central regulatory mechanism, the architecture enables the system not only to understand and verify information but also to become aware of its own reasoning processes and knowledge limitations. This structured progression provides a principled pathway for extending LLM capabilities beyond pattern reproduction toward controlled cross-domain imagination, ensuring that creativity emerges from validated understanding rather than hallucination. The overall architecture is decomposed into three interacting layers, as detailed below.

Let x denote an input query or stimulus, $K = \{K_1, \dots, K_n\}$ a structured knowledge base, and R a set of reasoning paths.

1. The Perception & Foundation Layer (States 1 & 2: Data to Knowledge)
   Raw input is encoded into a structured representation:
   $$z = f_{enc}(x)$$

where f_{enc} denotes a multimodal encoder (e.g., CLIP or LLM embeddings).

This layer ingests multimodal data (e.g., text and images) and employs a large language model (e.g., GPT-4) or multimodal encoder (e.g., CLIP) to transform raw data into structured knowledge representations. This process corresponds to moving from unprocessed data to interpretable knowledge. At this stage, the SAR knowledge base is populated, where incoming data are assimilated to enrich existing semantic structures. Coupled with System 1, which is trained through embedding-based representations, SAR consolidates factual knowledge beyond the generative capacity of the LLM acting as a language expert. As a result, the system produces fluent responses that are grounded in structured and validated knowledge.

2. The Reasoning & Metacognition Layer (States 3 & 4: Consciousness to Awareness)
   We define consciousness as a relevance-based selection function over available knowledge:
   $$C(x) = \underset{K_i \in K}{argmax}\, sim(z, K_i)$$
   where sim denotes semantic similarity. This corresponds to identifying "what the system knows" in relation to the input.

   Awareness is defined as the selection of an appropriate reasoning path:
   $$A(x) = \underset{r \in R}{argmax}\, P(r|C(x), g)$$
   where g denotes the task goal. This models "how to use the selected knowledge."

   This is the core SAR module. The LLM's initial, intuitive response (System 1) is passed to the Reasoning Module (System 2). This module, equipped with the knowledge base and logical rules, evaluates, verifies, and refines the response. Consciousness (*sati*) is modeled as the system's monitoring of this process—its awareness of its own internal states and knowledge gaps. Awareness (*sampajañña*) is the system's clear comprehension of why a particular reasoning path is chosen and how to act upon it to achieve a goal (e.g., answering a query correctly, applying knowledge across domains). In cross-domain imagination, the source- and target-domains are explicitly defined, and a latent concept matrix is derived by applying SVD to each domain-specific PAM.
3. The Imagination & Creation Layer (State 5: Enlightenment)
   We define imagination as a structured transformation followed by validation:
   $$E(x) = SAR(CrossSVD(z))$$
   where CrossSVD generates candidate cross-domain representations and SAR validates them for consistency and feasibility.

   When the task requires innovation beyond existing knowledge (e.g., "design a novel dish inspired by Vincent van Gogh's art work"), the system activates the cross-domain imaginative modelling. It retrieves the latent talent vectors (derived from PAM) for the source-domain (Vincent van Gogh's art) and, using the projection mechanism described in Section 4, generates candidate concepts in the target-domain (food). These candidates are then evaluated and refined by the Reasoning Layer for feasibility and novelty. The successful generation of a validated, novel concept represents the Enlightenment state—the creation of new knowledge.

Imaginative modelling begins with the construction of a persona's talent representation, derived from the individual's past well-recognized achievements. These achievements are encoded using multimodal vectorization methods, enabling the system to capture distinctive, high-fidelity latent features across heterogeneous data sources (e.g., text, images, artifacts, and descriptions). Through this process, raw data and information are transformed into structured knowledge representations, reducing interpretative ambiguity and preserving the uniqueness of the persona's creative capability.

To formalize talent, we introduce PAM, which aggregate and organize achievement-derived features within a focused domain. The selection and weighting of these features are guided by *sati* (conscious monitoring) and *sampajañña* (clear comprehension), ensuring that only salient, domain-relevant characteristics are retained. This metacognitive focusing mechanism allows the framework to distill the essence of a persona's creative uniqueness rather than relying on diffuse or averaged representations.

Building upon these matrices, the proposed framework performs cross-domain concept mapping via a cross-SVD formulation between source and target PAMs, whereby talent representations encoded by left singular vectors in the source-domain are transformed into the target-domain's conceptual space through target-specific concept weighting and right singular vectors. This imagination process is not stochastic but is instead constrained by structural analogy and continuously assessed by an SAR-based reasoning layer, ensuring that generated concepts remain grounded in validated talent structures and domain-aware constraints. Cross-SVD enables controlled imagination transfer by re-expressing validated talent structures across domain-specific conceptual bases rather than relying on unconstrained generation.

By supporting multi-talent PAM as a shared platform, the framework facilitates imagination transfer across diverse domains while preserving creative authenticity. This guarantees that imagination outputs remain rooted in identifiable talents and prior excellence, thereby enabling controlled, explainable, and reproducible creativity that reflects the belief that each talent is inherently and creatively unique.

This framework ensures that imagination is not a random process but is guided by reasoned structural analogy and grounded in metacognitive awareness. We believe in one's talent which is creatively unique. Widely recognized works of talent can serve as guides for creativity and imagination across various domains.

5. Formal Model for Cross-Domain Imagination

Given a source-domain (S) e.g., art, and a target-domain (T) e.g., food, and a set of P personas (e.g., experts, artists, chefs) with recorded achievements in both domains, we aim to generate novel, plausible achievements in T for a persona, based solely on their achievement pattern in S.

To construct Persona-Achievement Matrices (PAM), we define two matrices:

$M_S \in \mathbb{R}^{P \times S}$ (P personas × S achievements) where an entry indicates the association strength of a persona with a source-domain achievement.

$M_T \in \mathbb{R}^{P \times T}$ (P personas × T achievements) defined similarly for target-domain achievements.

5.1. Build Domain Matrices: Singular Value Decomposition (SVD) for Latent Concept Extraction

Applying SVD to each matrix reveals latent conceptual spaces:

$$M_S = U_S \Sigma_S V_S^T \tag{1}$$

$$M_T = U_T \Sigma_T V_T^T \tag{2}$$

In Eq. (1), U_S encodes persona-to-concept associations, V_S encodes achievement-to-concept associations. U_S and V_S are orthogonal matrices, repressing the vector in a new coordinate system that is rotated relative to the original one. Σ_S is a diagonal matrix, represents the relative strengths of latent source-domain concepts and carries domain-specific emphasis. These singular values are interpreted as a compact, ordered representation of the dominant conceptual forces underlying artistic achievement. This is also the same for the target-domain matrix (M_T) decomposition in Eq. (2), where U_T encodes persona-to-concept associations, V_T encodes achievement-to-concept associations, and Σ_T represents the relative strengths of latent target-domain concepts.

5.2. Align Concept Spaces: Define the Cross-Domain Concept Mapping

To project from art concepts to food concepts, we find a transformation matrix T that aligns the two concept spaces as expressed in Eq. (3).

$$T = U_S^T U_T \tag{3}$$

T is a composition (change of basis) matrix between two domain frames. This can be achieved via orthogonal Procrustes analysis to find how concepts in art relate to concepts in food. If columns of U_S and U_T are concept directions, then T is their alignment.

$$T = \underset{\mathrm{T}}{\operatorname{argmin}} \|U_T T - U_S\| \text{ subject to } T^T T = I \tag{4}$$

This finds the optimal rotation of the source-domain concept space to align with the target-domain concept space. Rotate the talent space of domain S so it matches the talent space of domain T as closely as possible, without destroying its internal structure. Orthogonal Procrustes analysis is employed to align latent talent spaces across domains by finding the optimal structure-preserving rotation, ensuring that cross-domain imagination transfers conceptual relationships to preserve relationships between talents and avoid hallucinated distortion rather than surface-level representations.

5.3. Transform Concepts: Structure-Preserving Cross-Domain Imaginative modelling

The core imagination mechanism in the proposed framework is realized through a cross-SVD–based projection, which enables the transfer of latent conceptual strength from a source-domain to a target-domain while preserving the internal structure of talent representations. Unlike conventional transfer learning or embedding alignment, cross-SVD explicitly operates on the singular value structure of domain-specific decompositions, thereby ensuring interpretability and structural consistency.

To enable cross-domain imagination, we introduce a cross-SVD transformation, which operates directly on the singular value space. Specifically, the transformation matrix T is learned via orthogonal Procrustes alignment between the latent concept spaces of the source- and target-domains. This ensures that conceptual strength is reinterpreted rather than distorted.

The transformed singular values are defined in Eq. (5).

$$\Sigma_T' = T\Sigma_T \tag{5}$$

In this formulation, T serves as a structure-preserving conceptual rotation operator, mapping source-domain latent concept strengths into the target-domain conceptual space while preserving relative structure. The orthogonality constraint on T ensures that relative conceptual structure, including ordering and proportional relationships among concept strengths, is preserved, thereby preventing arbitrary distortion such as uncontrolled amplification or attenuation of talent features. As a result, Σ_T' represents target-domain–interpretable concept strengths inferred purely from source-domain expertise.

5.4. Cross-Domain Imagination Matrix

Using this transformed singular value matrix, we reconstruct a hypothetical cross-domain imagination matrix $M_{T|S}$ based on cross-SVD transformation:

$$M_{T|S} = U_S\Sigma_T'V_T^T \tag{6}$$

The matrix represents the imagined target-domain achievement space conditioned on source-domain talent. Each row of $M_{T|S}$ corresponds to a persona, while each column corresponds to a target-domain achievement. Importantly, the reconstruction combines persona-level conceptual tendencies from the source-domain (U_S) with achievement-level semantic structure from the target-domain (V_T), mediated by the cross-SVD–transformed concept strengths Σ_T'.

As a result, $M_{T|S}$ encodes the hypothetical strength of association between personas and target-domain achievements, derived not from observed target-domain data but from cross-domain reinterpretation of source talent. Rows of this matrix suggest which novel target-domain achievements a persona may plausibly excel at, given their latent source-domain conceptual profile.

The numerical output of $M_{T|S}$ is subsequently interpreted by the Semantically Aware Reasoning (SAR) layer, which translates latent associations into structured, human-interpretable ideas. For example, if dominant art-domain concepts encoded in Σ_S correspond to notions such as abstract geometry, color minimalism, and textural contrast, the cross-SVD transformation may yield target-domain analogues such as modular plating, monochromatic sauces, and crisp–tender juxtaposition.

A concrete imagination instance could be the cross-domain reinterpretation of Vincent van Gogh’s principles, such as expressive texture, chromatic intensity, and rhythmic repetition into a culinary concept, such as a vividly colored vegetable purée plated in thick, swirling ridges, with repeated crisp elements arranged in flowing patterns, prioritizing tactile contrast and emotional intensity over conventional symmetry.

6. Cross-SVD based Cross-Domain Imagination

Cross-domain imagination in the proposed framework is realized by constructing domain-specific conceptual spaces derived from renowned personas who exemplify expertise within their respective domains. Given any pair of domains, a latent conceptual space can be formed by encoding representative achievements of domain experts. Within this setting, a multi-talent persona—one who exhibits meaningful engagement across domains—can be utilized as a structural bridge to derive a cross-domain transformation matrix.

Once the latent concepts and the transformation matrix are established, expertise from a source-domain can be projected into a target-domain while preserving internal conceptual structure. This process ensures that cross-domain imagination is not arbitrary, but grounded in measurable and transferable latent representations.

6.1. Domain Expert and Conceptual Alignment

As a concrete conceptual study, we select two renowned personas from distinct domains: culinary arts and visual arts.

Jacques Pépin is renowned for his mastery of culinary technique and pedagogical clarity in gastronomy. His work is characterized by procedural precision, structural simplicity, and a deep respect for ingredients. In addition to his culinary expertise, Pépin has engaged in visual art through painting, making him a multi-talent persona whose conceptual structures span both food and art domains. This dual engagement positions Pépin as an ideal bridge for cross-domain transformation.

Vincent van Gogh, by contrast, is renowned for his mastery of visual expression, characterized by bold chromatic intensity, dynamic brushwork, rhythmic repetition, and emotionally driven composition. His work prioritizes perceptual and affective communication over representational realism, establishing enduring principles that define modern visual art.

Table 1. Aspect-wise comparison between two domain experts, Jacques Pépin and Vincent van Gogh

Aspect	Jacques Pépin	Vincent van Gogh
Domain	Culinary arts	Visual arts
Core mastery	Technique, structure, pedagogy	Expression, perception, emotion
Knowledge type	Procedural and applied	Perceptual and expressive
Transferable principles	Precision, simplicity, material respect	Color, texture, rhythm

For clarity, the conceptual correspondence between the two domain experts is summarized as in Table 1. This comparison highlights that, while their domains differ substantially, both experts exhibit coherent latent principles that can be abstracted and transformed across domains.

6.2. Cross-SVD Construction and CLIP-based Encoding

To construct the latent conceptual spaces, Pépin's original culinary works and visual artworks are encoded using CLIP, producing aligned multimodal embeddings. Similarly, Vincent van Gogh's original artworks are encoded to form the visual-art conceptual space. These embeddings are then decomposed using singular value decomposition to extract latent concepts and their corresponding strengths.

Using Pépin's art–food correspondence, a cross-SVD transformation matrix T is derived. This matrix captures how conceptual strength shifts between visual and culinary domains for a single multi-talent persona, thereby encoding a structure-preserving cross-domain mapping.

6.3. Imagination Results and Interpretation

Figure 1 illustrates the experimental results of applying the Pépin-derived transformation matrix across multiple cases.

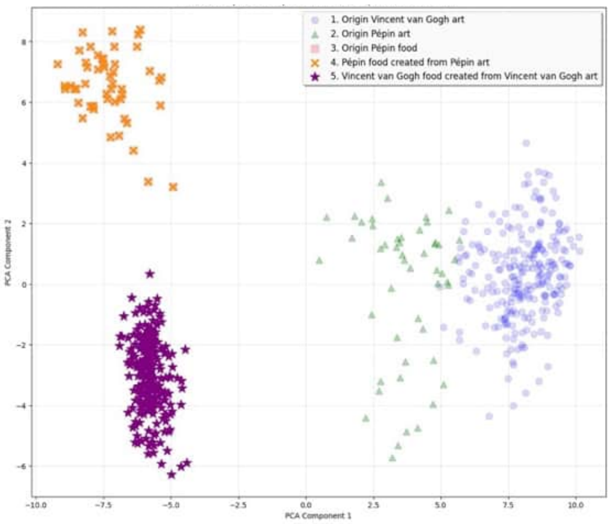

Figure 1. Cross-SVD–based cross-domain imagination (×, ⋆) using the Pépin bridge

This figure visualizes the latent conceptual spaces of culinary and visual arts constructed using CLIP embeddings and singular value decomposition. The transformation matrix T, derived solely from Jacques Pépin's art–food correspondence, is applied across multiple cases. Pépin's food works reconstructed from his own art align with his original culinary region, validating the structure-preserving property of Eq. (6).

When Vincent van Gogh's artworks are projected into the culinary domain using the same transformation, a distinct and previously unoccupied region emerges, representing novel food concepts inspired by artistic principles rather than culinary precedent. This example illustrates the qualitative behavior of the proposed method.

The key observations are as follows:

1. Self-consistency of the bridge persona

   Pépin's culinary works generated from his own visual artworks are projected into the same region as his original culinary works. This confirms that the reconstructed matrix $M_{T|S}$, as defined in Eq. (6), preserves domain consistency and validates the correctness of the cross-SVD formulation.

   From Eq. (6), Σ_T' is replaced by $T\Sigma_T$, as defined in Eq. (5), yielding Eq. (7).

   $$M_{T|S} = U_S T \Sigma_T V_T^T \tag{7}$$

   Next, T in Eq. (7) is replaced by $U_S^T U_T$, as defined in Eq. (3), resulting in Eq. (8).

   $$M_{T|S} = U_S U_S^T U_T \Sigma_T V_T^T \tag{8}$$

   Since U_S is an orthogonal matrix, therefore,

   $$U_S U_S^T = I \tag{9}$$

where I denotes the identity matrix. Substituting Eq. (9) into Eq. (8) leads to Eq. (10).

$$M_{T|S} = U_T \Sigma_T V_T^T \tag{10}$$

Eq. (10) demonstrates that Pépin's culinary works generated from his own visual artworks are identical to the matrix representation of Pépin's original culinary works.

2. Cross-domain imagination emergence
   When Vincent van Gogh's artworks are projected into the Pépin-derived culinary conceptual space, the resulting points occupy a distinct and previously unpopulated region. This region does not overlap with Pépin's original culinary works, indicating the emergence of novel culinary concepts grounded in Vincent van Gogh's artistic principles rather than culinary precedent. This result is confirmed by applying Eq. (6), in which U_S captures the latent artistic structure of Vincent van Gogh's artworks, V_T defines the culinary feature space of Pépin's works, and Σ_T' serves as the transformed food concept matrix that mediates cross-domain imagination between visual art and cuisine.
3. Interpretation of the novel region
   The appearance of a new region represents hypothetical culinary creations inspired by Vincent van Gogh's latent artistic structure—such as chromatic intensity, rhythmic composition, and expressive texture—reinterpreted through the culinary domain. These outputs are not random extrapolations but are structurally constrained by the Pépin bridge, ensuring conceptual plausibility.

6.4. Implications for Imaginative AI

This experiment demonstrates that imagination can be computationally realized as a structure-preserving transformation of latent talent across domains. By relying on a single multi-talent persona to establish the transformation matrix, the proposed cross-SVD method enables scalable and controllable cross-domain inspiration. The resulting creative space is effectively unbounded, offering substantial potential for innovation not only in creative industries but also in scientific discovery and invention.

Crucially, this approach reframes imagination as neither generative randomness nor simple analogy, but as a principled reinterpretation of expertise, grounded in linear algebra and semantic structure. This positions cross-SVD as a foundational mechanism for imaginative AI capable of producing meaningful, novel, and explainable cross-domain knowledge.

7. Conclusion

This paper introduced imaginative AI, a structured framework for cross-domain knowledge creation grounded in latent concept transformation rather than unconstrained generation. At the core of the proposed approach is cross-SVD, a novel extension of singular value decomposition that enables structure-preserving imagination transfer across heterogeneous domains by explicitly transforming singular value representations of conceptual strength. Unlike existing multimodal or generative approaches that rely primarily on similarity matching or probabilistic sampling, cross-SVD operates at the level of latent concept intensity. By decomposing domain-specific PAM and applying an orthogonality-constrained transformation to the singular value space, the proposed method reinterprets expertise from a source-domain into a target-domain without collapsing semantic structure. This formulation provides a mathematically transparent mechanism for imagination, where novelty emerges from principled conceptual rotation rather than stochastic extrapolation.

The integration of cross-SVD with Semantically Aware Reasoning (SAR) further ensures that imagination outputs remain coherent, interpretable, and domain-consistent. While cross-SVD generates hypothetical cross-domain associations, SAR evaluates and grounds these associations against structured knowledge, mitigating hallucination and enabling verifiable reasoning. Together, these components establish a clear separation between imagination projection and semantic validation, which is essential for controllable and trustworthy AI systems. Beyond technical contributions, this work also advances the conceptual understanding of imagination in artificial intelligence. By framing imagination as a transformation of latent conceptual structure—rather than the generation of surface-level artifacts—the proposed framework bridges representation learning, reasoning, and consciousness-inspired AI. The incorporation of mindfulness-informed concepts, such as intentional awareness and reflective evaluation, provides a principled pathway toward AI systems capable of moving beyond data and information to meaningful knowledge creation, through a cognitive process aligned with the Five States of Mind Enlightenment.

The current study has several limitations. First, the experimental evaluation is conducted on a limited set of domains and personas, which may not fully capture the diversity of real-world knowledge structures. Second, the Cross-SVD transformation assumes linear relationships between latent concept spaces, which may not hold in more complex domains. Third, human evaluation remains subjective and may introduce bias.

Future work will address these limitations by expanding datasets, exploring nonlinear transformations, and incorporating more rigorous evaluation protocols.

References

[1] Lake BM, Salakhutdinov R, Tenenbaum JB. Human-level concept learning through probabilistic program induction. Science. 2015;350(6266):1332–1338. doi:10.1126/science.aab3050

[2] Marcus G. Deep learning: A critical appraisal. arXiv [Preprint]. 2018. arXiv:1801.00639. Available from: https://arxiv.org/abs/1801.00639

[3] Shao M, Basit A, Karri R, Shafique M. Survey of different large language model architectures: Trends, benchmarks, and challenges. IEEE Access. 2024;12:188664–188706. doi:10.1109/ACCESS.2024.3482107

[4] Sun L, Shi Z. Prompt learning under the large language model. In: Proceedings of the International Seminar on Computer Science and Engineering Technology (SCSET); 2023.

[5] Sornlertlamvanich V, Doi R, Charoenporn T. AI model deficiency in knowledge insufficiency. In: Kiyoki Y, Sornlertlamvanich V, Tropmann-Frick M, Jaakkola H, Yoshida N, editors. Information Modelling and Knowledge Bases XXXVI. Frontiers in Artificial Intelligence and Applications, vol. 399. Amsterdam: IOS Press; 2025. p. 19–33. doi:10.3233/FAIA241572

[6] Kahneman D. Thinking, fast and slow. New York (NY): Farrar, Straus and Giroux; 2011.

[7] Sornlertlamvanich V. Enhancing large language models: Alleviating knowledge deficiency with external knowledge and semantically aware reasoning (SAR). Knowl Inf Syst. 2025. doi:10.1007/s10115-025-02485-4

[8] Sornlertlamvanich V, Aksorn N, Charoenporn T. Multilingual support in connective learning schemes for refining and connecting open educational videos. In: Proceedings of the Language Technologies for All (LT4All). Paris: UNESCO; 2019.

[9] Lewis P, Perez E, Piktus A, Petroni F, Karpukhin V, Goyal N, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. In: Advances in Neural Information Processing Systems (NeurIPS). 2020;33:9459–9474.

[10] Radford A, Kim JW, Hallacy C, Ramesh A, Goh G, Agarwal S, et al. Learning transferable visual models from natural language supervision. In: Proceedings of the 38th International Conference on Machine Learning (ICML); 2021.

[11] Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An image is worth 16×16 words: Transformers for image recognition at scale. In: International Conference on Learning Representations (ICLR); 2021.

[12] Sweller J. Cognitive load theory, learning difficulty, and instructional design. Learn Instr. 1994;4:295–312. doi:10.1016/0959-4752(94)90003-5

[13] Baddeley A. The episodic buffer: A new component of working memory? Trends Cogn Sci. 2000;4(11):417–423. doi:10.1016/S1364-6613(00)01538-2

[14] Landauer TK, Dumais ST. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychol Rev. 1997;104(2):211–240. doi:10.1037/0033-295X.104.2.211

[15] Deerwester S, Dumais ST, Furnas GW, Landauer TK, Harshman R. Indexing by latent semantic analysis. J Am Soc Inf Sci. 1990;41(6):391–407. doi:10.1002/(SICI)1097-4571(199009)41:6<391::AID-ASI1>3.0.CO;2-9

[16] Golub GH, Reinsch C. Singular value decomposition and least squares solutions. Numerische Mathematik. 1970;14:403–420. doi:10.1007/BF02163027

[17] Payutto PA (Phra Brahmagunabhorn). Buddhadhamma: The laws of nature and their benefits to life. Bangkok: Buddhadhamma Foundation; 2021.