

# Il vincolo che dà forma all'AI | The Constraint That Shapes AI

## *Dimensioni ortogonali come preconditione della riferibilità | Orthogonal Dimensions as a Precondition for Referentiality*

---

Autore / Author: Pasquale Papanice

Affiliazione / Affiliation: Uniark SRLU — SYNTO® (<https://synto.it>)

Ruolo / Role: CEO & Founder

Data / Date: Gennaio / January 2026

Versione / Version: 3.0

---

### Indice / Table of Contents

- [Versione Italiana](#)
  - [English Version](#)
  - [Bibliografia / References](#)
- 

### Sommario

Il vincolo che dà forma all'AI   The Constraint That Shapes AI.....	1
<i>Dimensioni ortogonali come preconditione della riferibilità   Orthogonal Dimensions as a Precondition for Referentiality.....</i>	1
Indice / Table of Contents.....	1
Bibliografia / References.....	3
VERSIONE ITALIANA.....	3
Abstract.....	3
1. Introduzione: Il problema non è l'AI.....	4
2. Il triangolo di Kanizsa: forma emergente dal vincolo.....	4
3. Kant e le condizioni a priori della conoscenza .....	5

4. Il grounding non basta.....	5
5. Dimensioni ortogonali: lo spazio della riferibilità.....	6
6. Emergenza per sottrazione: quando filtrare è creare.....	6
7. Il frattale della classificazione: dal rumore alla forma.....	7
8. Implicazioni per l'AI: il compito umano nell'era delle macchine.....	8
9. Oltre la riferibilità: verso un'ontologia del confine.....	9
Il confine come unità fondamentale.....	9
Il segno come confine persistente.....	9
Il sentire prima del segno.....	9
Dal rumore alla materia: saturazione di confini.....	10
10. Conclusione: Strumento minimo, pura classe.....	10
ENGLISH VERSION.....	12
Abstract.....	12
1. Introduction: The Problem Is Not the AI.....	12
2. The Kanizsa Triangle: Form Emerging from Constraint.....	12
3. Kant and the A Priori Conditions of Knowledge.....	13
4. Grounding Is Not Enough.....	14
5. Orthogonal Dimensions: The Space of Referentiality.....	14
6. Emergence by Subtraction: When Filtering Is Creating.....	15
7. The Fractal of Classification: From Noise to Form.....	15
8. Implications for AI: The Human Task in the Age of Machines.....	16
9. Beyond Referentiality: Toward an Ontology of the Boundary.....	17
The Boundary as Fundamental Unit.....	17
The Sign as Persistent Boundary.....	18
Feeling Before the Sign.....	18
From Noise to Matter: Saturation of Boundaries.....	18
10. Conclusion: Minimal Tool, Pure Class.....	19

## Bibliografia / References

- Bellante, M. (2025). RAG come paradigma di riferibilità / RAG as a paradigm of referentiality. [Working paper].
  - Es, S., James, J., Espinosa-Anke, L., & Schockaert, S. (2023). RAGAS: Automated Evaluation of Retrieval Augmented Generation. arXiv:2309.15217. <https://arxiv.org/abs/2309.15217>
  - Farquhar, S., Kossen, J., Kuhn, L., & Gal, Y. (2024). Detecting hallucinations in large language models using semantic entropy. *Nature*, 630, 625–630. <https://www.nature.com/articles/s41586-024-07421-0>
  - Hedden Information Management. (n.d.). Faceted Classification and Faceted Taxonomies. <https://www.hedden-information.com/faceted-classification-and-faceted-taxonomies/>
  - Illusory contour. (n.d.). In Wikipedia. [https://en.wikipedia.org/wiki/Illusory\\_contour](https://en.wikipedia.org/wiki/Illusory_contour)
  - Kanizsa, G. (1955). Margini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia*, 49(1), 7–30.
  - Kant, I. (1781). *Kritik der reinen Vernunft* [Critica della Ragion Pura / Critique of Pure Reason].
  - Spencer-Brown, G. (1969). *Laws of Form*. Allen & Unwin.
  - Stanford Encyclopedia of Philosophy. (n.d.). Categories. <https://plato.stanford.edu/entries/categories/>
  - The Illusions Index. (n.d.). Kanizsa Triangle. <https://www.illusionsindex.org/i/kanizsa-triangle>
  - Theory of Informational Emergence (TIE). (2024). ResearchGate. <https://www.researchgate.net/publication/389490033>
  - The Information-Based Model of Emergent Structures. (2024). ResearchGate. <https://www.researchgate.net/publication/391653709>
  - arXiv. (2025). A Systematic Review of Key Retrieval-Augmented Generation (RAG) Systems: Progress, Gaps, and Future Directions. arXiv:2507.18910. <https://arxiv.org/html/2507.18910v1>
- 

## VERSIONE ITALIANA

### Abstract

L'integrazione dei Large Language Models in contesti operativi ha evidenziato un problema fondamentale: l'allucinazione. La risposta tecnica dominante — il Retrieval-Augmented Generation (RAG) — propone di ancorare la generazione a documenti recuperati dinamicamente. Questo articolo sostiene che il grounding, per quanto necessario, non è sufficiente: presuppone che i documenti esistano in uno spazio navigabile deterministicamente. Attingendo alla teoria delle categorie kantiane, alla psicologia della Gestalt e alla Theory of Informational Emergence, proponiamo che le dimensioni classificatorie ortogonali non siano metadati accessori ma condizioni a priori della riferibilità. Come nel triangolo di Kanizsa la forma emerge dai vincoli percettivi, così il significato emerge dall'intersezione di dimensioni ortogonali. Il vincolo non limita l'AI — la rende possibile.

## 1. Introduzione: Il problema non è l'AI

L'adozione diffusa dei Large Language Models (LLM) ha portato alla luce un fenomeno che la letteratura tecnica ha battezzato "allucinazione": la generazione di contenuti sintatticamente coerenti ma fattualmente errati, privi di ancoraggio alla realtà (Farquhar et al., 2024).

La risposta ingegneristica a questo problema è stata il Retrieval-Augmented Generation (RAG), un'architettura che vincola il processo generativo a un contesto recuperato dinamicamente da fonti controllate (Bellante, 2025). Il RAG trasforma l'LLM da "oracolo opaco" ad "analista referenziato", permettendo la tracciabilità delle risposte alle fonti documentali.

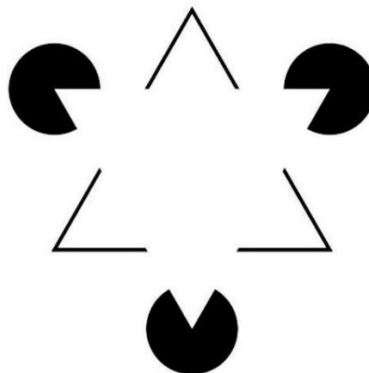
Tuttavia, questa soluzione presuppone qualcosa che raramente viene esplicitato: che i documenti da recuperare esistano in uno spazio navigabile. Che abbiano coordinate. Che siano classificati secondo dimensioni che permettano il retrieval deterministico.

La tesi di questo articolo è che il problema dell'allucinazione non risiede primariamente nell'AI, ma nell'assenza di struttura nel dato. Non servono LLM migliori — serve lo spazio in cui possano operare con riferibilità.

## 2. Il triangolo di Kanizsa: forma emergente dal vincolo

Nel 1955, lo psicologo italiano Gaetano Kanizsa descrisse un'illusione ottica destinata a diventare paradigmatica per la psicologia della Gestalt: il triangolo di Kanizsa (Kanizsa, 1955).

L'osservatore percepisce un triangolo bianco sovrapposto a tre cerchi neri e a un triangolo invertito. Ma questo triangolo bianco non esiste. Non c'è alcun contorno fisico che lo delimiti. Ciò che esiste sono tre "Pac-Man" — cerchi con un settore mancante — e tre angoli di un triangolo incompleto.



Il triangolo emerge dalla configurazione dei vincoli. È un esempio di quello che la letteratura chiama "reificazione": il sistema percettivo costruisce informazione spaziale esplicita che non è presente nello stimolo sensoriale (Illusory contour, Wikipedia).

Questo fenomeno illustra un principio profondo: la forma non è nei componenti, ma nella loro relazione. Il triangolo non è nei cerchi. Non è negli angoli. È nell'intersezione dei vincoli imposti da questi elementi.

La stessa dinamica, sosteniamo, governa l'emergenza del significato nell'informazione classificata.

### **3. Kant e le condizioni a priori della conoscenza**

Immanuel Kant, nella Critica della Ragion Pura (1781), pose una domanda fondamentale: come è possibile la conoscenza sintetica a priori? Come possiamo conoscere qualcosa del mondo prima di sperimentarlo?

La risposta kantiana è che esistono categorie — strutture della mente — che sono condizioni di possibilità di qualsiasi esperienza. Queste categorie (quantità, qualità, relazione, modalità) non derivano dall'esperienza ma la rendono possibile (Stanford Encyclopedia of Philosophy, Categories).

Kant scrive: "I pensieri senza contenuto sono vuoti, le intuizioni senza concetti sono cieche". Non c'è nulla da conoscere senza contenuto empirico; ma senza strutture a priori, non abbiamo modo di dare forma intelligibile a qualsiasi contenuto possiamo avere.

Questa intuizione si applica direttamente al problema dell'informazione digitale. Un documento senza classificazione è come un'intuizione senza concetto: esiste, ma non può essere conosciuto nel senso di essere navigato, filtrato, recuperato deterministicamente.

Le dimensioni classificatorie — chi produce, chi può accedere, dove si colloca geograficamente, di cosa tratta — non sono metadati accessori. Sono condizioni a priori della riferibilità.

### **4. Il grounding non basta**

La letteratura recente sul RAG ha identificato con precisione i limiti del grounding. Bellante (2025) nota che "la traceability non è un attributo emergente, ma una caratteristica ingegnerizzata deliberatamente attraverso il grounding". Ma il grounding presuppone che esista qualcosa a cui ancorare.

Una survey sistematica del 2025 rileva che "RAG breaks at scale because organizations treat it like a feature of LLMs rather than a platform discipline. The real challenges emerge in ingestion, retrieval optimization, metadata management, versioning, indexing, evaluation, and long-term governance" (arXiv:2507.18910).

Il problema non è il modello. È l'infrastruttura classificatoria.

Farquhar et al. (2024) hanno proposto l'entropia semantica come metrica per rilevare le allucinazioni: se l'LLM, variando la temperatura, produce risposte con significati radicalmente

diversi, l'incertezza epistemica è alta. Ma questa metrica presuppone che esista un ground truth rispetto al quale misurare la divergenza.

RAGAS (Retrieval Augmented Generation Assessment) fornisce metriche come Faithfulness e Context Precision (Es et al., 2023). Ma Faithfulness misura la coerenza della risposta rispetto al contesto recuperato — e se il contesto recuperato è sbagliato perché la classificazione era inadeguata?

Il grounding è necessario ma non sufficiente. Prima del grounding, serve lo spazio in cui il grounding possa operare.

## 5. Dimensioni ortogonali: lo spazio della riferibilità

In fisica, tempo e spazio sono le dimensioni ortogonali entro cui qualsiasi evento può essere localizzato. Un evento senza coordinate spazio-temporali non è un evento — è un'astrazione priva di riferibilità.

Analogamente, proponiamo che l'informazione riferibile richieda dimensioni ortogonali di classificazione. Non una gerarchia singola (cartella → sottocartella → file), ma un sistema di coordinate multiple e indipendenti:

- Chi produce l'informazione (provenienza, attribuzione)
- Chi può accedere (perimetro di visibilità, controllo)
- Livello di riservatezza (gradiente pubblico/privato)
- Collocazione geografica (dove è rilevante)
- Dominio semantico (di cosa tratta)
- Argomento specifico (sotto-classificazione tematica)

Queste dimensioni sono ortogonali nel senso matematico: indipendenti l'una dall'altra. Un documento può essere prodotto da A, accessibile a B, riservato, rilevante per l'Italia, nel dominio legale, sul tema contratti. Ogni dimensione aggiunge informazione non derivabile dalle altre.

La letteratura sulle tassonomie a faccette conferma questa intuizione: "It has become more common to design faceted taxonomies than purely hierarchical taxonomies. Faceted taxonomies integrate better with search, serve for filters and sorting, and provide a good user experience" (Hedden Information Management).

Ma il valore delle dimensioni ortogonali va oltre l'usabilità. È ontologico.

## 6. Emergenza per sottrazione: quando filtrare è creare

La Theory of Informational Emergence (TIE) propone che "neither information nor structure is primary; instead, they are two inseparable aspects of a dynamic process" (ResearchGate, 2024).

L'Information-Based Model of Emergent Structures (IBMES) avanza ulteriormente: "emergent phenomena result not from the properties of matter per se, but from patterns of information dynamically organizing and interacting under constraints" (ResearchGate, 2024).

Ma cosa significa "under constraints"? I vincoli non aggiungono — sottraggono. Come lo scultore che toglie il marmo in eccesso per rivelare la forma, le dimensioni ortogonali non costruiscono l'informazione: la filtrano dal rumore. Ogni classificazione è un setaccio che lascia passare meno.

Questa prospettiva illumina il valore delle dimensioni ortogonali. Non sono contenitori passivi. Sono vincoli sottrattivi. L'intersezione di dimensioni multiple genera informazione che non è presente in nessuna dimensione singola — esattamente come il triangolo di Kanizsa emerge dai vincoli percettivi senza essere presente in nessun elemento individuale.

Un documento classificato come [prodotto da X] + [accessibile a Y] + [rilevante per Z] + [sul tema W] acquisisce una posizione nello spazio informativo. Questa posizione permette:

- Retrieval deterministico: trovare esattamente ciò che serve, non approssimativamente
- Filtering combinatorio: intersecare criteri multipli
- Attribuzione: sapere chi ha prodotto cosa
- Controllo accessi: garantire che solo chi deve vedere, veda
- Relazioni emergenti: scoprire connessioni non esplicite tra documenti che condividono coordinate

L'ultima proprietà è cruciale. Quando due documenti condividono alcune dimensioni ma differiscono in altre, emerge una relazione che nessun sistema potrebbe dedurre dai documenti in isolamento.

## **7. Il frattale della classificazione: dal rumore alla forma**

Le dimensioni ortogonali non sono statiche. Ogni dimensione può contenere sotto-dimensioni, che a loro volta possono contenere ulteriori livelli. È una struttura frattale: auto-similare a scale diverse.

Un punto cruciale: lo stato originario non è il vuoto — è il rumore. Nel rumore, tutto è sovrapposto: aria, acqua, dati, segnali, potenziale indifferenziato. Non "nulla", ma "tutto insieme". Aria e acqua non sono sostanze primarie — sono categorie, tagli nel continuum. Prima del taglio che definisce "aria", non esiste aria. Esiste solo il rumore che contiene ogni possibilità.

L'output stesso del sistema classificatorio — le relazioni emergenti, i pattern scoperti — può diventare nuovo input per ulteriore classificazione. Il sistema non solo organizza l'informazione; filtra rumore per rivelare forma.

Questo processo ricorsivo può essere visualizzato come un'implosione dal rumore verso il significato:

```
RUMORE (tutto sovrapposto, indifferenziato)
  ↓ [primo taglio: distinzione]
QUALCOSA ≠ ALTRO
  ↓ [classificazione: dimensioni ortogonali]
DATO CLASSIFICATO (coordinate nello spazio)
  ↓ [intersezione: relazioni emergenti]
PATTERN (informazione di ordine superiore)
  ↓ [meta-classificazione]
PATTERN CLASSIFICATO
  ↓ [ricorsione: filtri su filtri]
FORMA (ciò che resta dopo n tagli)
```

Ad ogni livello, il vincolo classificatorio non aggiunge informazione — la rivela sottraendo rumore. Come i tre Pac-Man non riducono il campo visivo ma generano un triangolo, così le dimensioni ortogonali non limitano il dato ma rivelano significato.

## 8. Implicazioni per l'AI: il compito umano nell'era delle macchine

Se l'analisi precedente è corretta, le implicazioni per i sistemi AI sono profonde.

Il problema dell'allucinazione non è primariamente un problema di modelli. È un problema di assenza di coordinate. Un LLM che opera su un corpus non strutturato è come un sistema percettivo senza vincoli: può generare qualsiasi forma, ma nessuna sarà ancorata.

Il RAG è necessario ma presuppone la classificazione. Il retrieval funziona se e solo se esistono criteri di retrieval. Questi criteri sono le dimensioni ortogonali. Senza di esse, il retrieval è aleatorio.

La reference-free evaluation ha limiti intrinseci. RAGAS e framework simili misurano la coerenza interna (faithfulness al contesto, relevance alla query). Ma non possono misurare se il contesto giusto è stato recuperato in primo luogo. La valutazione completa richiede ground truth — che presuppone classificazione.

L'AI non ha sentire. Un Large Language Model processa segni — miliardi di segni, con velocità e pattern-matching che nessun umano può eguagliare. Ma non sente dove tagliare. Entra in scena dopo che qualcuno ha già sentito, già distinto, già classificato. L'AI è potenza di elaborazione sul già-tagliato.

Questo ridefinisce il compito umano nell'era dell'AI:

```
ERA PRE-AI:    human sente, human taglia, human elabora
ERA AI:        human sente, human taglia, AI elabora
```

Il compito umano non scompare — si concentra. L'elaborazione la cediamo alle macchine. Ma il sentire? Il primo taglio? Insostituibile. Chi non taglia, sarà tagliato. Chi delega anche il sentire, diventa rumore per l'AI altrui.

## 9. Oltre la riferibilità: verso un'ontologia del confine

Le sezioni precedenti hanno trattato le dimensioni ortogonali come condizioni della riferibilità — la capacità di trovare, verificare, attribuire. Ma la reificazione nel triangolo di Kanizsa suggerisce qualcosa di più radicale.

Nel fenomeno di Kanizsa, il sistema percettivo non scopre un triangolo — lo genera. I vincoli percettivi non rivelano una forma preesistente; la producono. Questa è la definizione stessa di reificazione: la creazione di qualcosa che non c'era.

Se estendiamo questa logica, emerge una domanda: qual è l'elemento minimo della reificazione? Cosa accomuna il digitale e il materiale?

### ***Il confine come unità fondamentale***

La risposta non è il bit inteso come 0 o 1. Il bit è il confine tra 0 e 1. Non rappresenta uno stato — rappresenta l'esistenza di una separazione. Prima che esistano "0" e "1" come valori, deve esistere la distinzione che li rende possibili.

George Spencer-Brown, in *Laws of Form* (1969), formalizzò questo principio: "Draw a distinction". La distinzione precede ciò che distingue. Il confine è ontologicamente primo; i lati sono derivati.

```
0 | 1
  ↑
  IL BIT È QUI
  (la distinzione stessa)
```

Questo ribalta la visione classica:

```
VISIONE CLASSICA:   Sostanze → Confini (i confini separano sostanze)
VISIONE RIBALTATA: Confini → Sostanze (le sostanze emergono dai
confini)
```

La sostanza è illusione derivata. Il confine è realtà primaria.

### ***Il segno come confine persistente***

Se il confine è l'elemento minimo, il segno è il confine reso persistente. Prima del segno "aria", non esiste aria come categoria. Il segno non descrive una distinzione preesistente — la crea.

Il prefisso DIS- (latino) cattura questo gesto primordiale: dis-cernere (separare col setaccio), distinguere (marcare la differenza), dis-crimen (ciò che separa, da cui: crimine, crisi, criterio).

Vedere è già separare. Percepire è già tagliare.

### ***Il sentire prima del segno***

Ma c'è qualcosa che precede anche il segno: il sentire. Prima della pennellata c'è l'intuizione del pittore. Prima del colpo di scalpello c'è la visione dello scultore. Prima del segno c'è il sentire di dove il confine vuole essere.

Il segno cristallizza il sentire. Lo rende trasmissibile, persistente. Ma il taglio originario avviene nel sentire stesso — pre-verbale, pre-formale, pre-tutto.

```
SENTIRE (distinzione pre-verbale, intuizione)
  ↓ [cristallizza]
SEGNO (distinzione persistente, trasmissibile)
  ↓ [accumula]
TESTO (geometria di distinzioni)
  ↓ [sistema]
CLASSIFICAZIONE (spazio navigabile)
```

### ***Dal rumore alla materia: saturazione di confini***

La fisica contemporanea offre indizi convergenti. John Archibald Wheeler propose il principio "It from bit": l'informazione non descrive la realtà — la costituisce. Ma più precisamente: It from cut. La realtà emerge dal taglio primordiale nel rumore.

Il frattale della classificazione può essere riformulato come saturazione progressiva di confini:

```
RUMORE (zero confini: tutto sovrapposto)
  ↓ [primo taglio: ≠]
DUE (qualcosa si distingue da altro)
  ↓ [secondo taglio]
QUATTRO
  ↓ [n tagli]
COMPLESSITÀ (geometria di confini)
  ↓ [saturazione: non si può più tagliare]
MATERIA (densità massima di distinzioni)
```

In questa prospettiva, la materia non è una sostanza speciale — è ciò che resta quando i gradi di libertà sono esauriti. È rumore così tanto tagliato da non poter essere altro. La materia è densità di confini.

L'allucinazione, in questa luce, non è semplicemente un errore. È informazione con confini insufficienti. Può assumere qualsiasi forma perché non ha abbastanza tagli a fissarla. Resta "gassosa".

Il grounding RAG è un tentativo di aggiungere confini. Ma se i confini sono pochi, mal strutturati, o non ortogonali, l'informazione non si condensa. Un sistema con dimensioni ortogonali ricorsive è, in questo senso, una macchina di reificazione: prende rumore e lo spinge attraverso tagli successivi fino a renderlo solido, riferibile, reale.

Questa non è una metafora. È una proposta ontologica: il confine non separa sostanze — le genera.

## **10. Conclusione: Strumento minimo, pura classe**

Questo articolo ha proposto un ribaltamento di prospettiva a tre livelli.

Al primo livello, tecnico: l'allucinazione non è un bug dell'AI da correggere con modelli migliori. È il sintomo di un'assenza strutturale — l'assenza di dimensioni ortogonali che permettano all'informazione di esistere in uno spazio navigabile.

Al secondo livello, operativo: l'AI non ha sentire. Processa segni, non li genera. Il compito umano nell'era delle macchine si concentra: sentire dove tagliare, imporre il primo confine. L'elaborazione la cediamo. Il taglio originario resta nostro.

Al terzo livello, ontologico: il confine non è una limitazione imposta alla sostanza — è la condizione stessa della sua esistenza. La forma emerge dal confine. Il significato emerge dall'intersezione di confini. La materia, al limite, emerge dalla saturazione di confini.

Come il triangolo di Kanizsa emerge dai vincoli percettivi — non nonostante essi ma grazie ad essi — così ogni livello di realtà emerge dal precedente attraverso l'imposizione di tagli.

Un sistema di classificazione ortogonale non è un contenitore di informazioni. È uno strumento minimo — la zappa senza cui il contadino non può coltivare. È pura classe — kantiano fino all'osso: non contenuto, ma condizione di possibilità del contenuto.

Le categorie kantiane erano condizioni a priori della conoscenza empirica. Le dimensioni ortogonali sono condizioni a priori della riferibilità informativa. E forse, spingendo il ragionamento, i confini ricorsivi sono condizioni a priori della realtà stessa.

Il futuro dell'AI affidabile non passa primariamente per LLM più potenti o RAG più sofisticati. Passa per l'infrastruttura classificatoria che rende possibile l'uno e l'altro. Passa per chi sa sentire dove tagliare.

Il vincolo dà forma all'AI. E forse, dà forma a tutto il resto.

# ENGLISH VERSION

## Abstract

The integration of Large Language Models in operational contexts has highlighted a fundamental problem: hallucination. The dominant technical response — Retrieval-Augmented Generation (RAG) — proposes to anchor generation to dynamically retrieved documents. This article argues that grounding, while necessary, is not sufficient: it presupposes that documents exist in a deterministically navigable space. Drawing on the theory of Kantian categories, Gestalt psychology, and the Theory of Informational Emergence, we propose that orthogonal classificatory dimensions are not accessory metadata but a priori conditions of referentiality. Just as in the Kanizsa triangle the form emerges from perceptual constraints, so meaning emerges from the intersection of orthogonal dimensions. The constraint does not limit AI — it makes it possible.

## 1. Introduction: The Problem Is Not the AI

The widespread adoption of Large Language Models (LLMs) has brought to light a phenomenon that technical literature has dubbed "hallucination": the generation of syntactically coherent but factually incorrect content, lacking anchorage to reality (Farquhar et al., 2024).

The engineering response to this problem has been Retrieval-Augmented Generation (RAG), an architecture that constrains the generative process to a context dynamically retrieved from controlled sources (Bellante, 2025). RAG transforms the LLM from an "opaque oracle" to a "referenced analyst," enabling traceability of responses to documentary sources.

However, this solution presupposes something rarely made explicit: that the documents to be retrieved exist in a navigable space. That they have coordinates. That they are classified according to dimensions that permit deterministic retrieval.

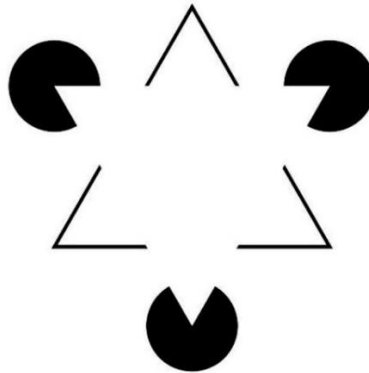
The thesis of this article is that the hallucination problem does not primarily reside in AI, but in the absence of structure in the data. We don't need better LLMs — we need the space in which they can operate with referentiality.

## 2. The Kanizsa Triangle: Form Emerging from Constraint

In 1955, Italian psychologist Gaetano Kanizsa described an optical illusion destined to become paradigmatic for Gestalt psychology: the Kanizsa triangle (Kanizsa, 1955).

The observer perceives a white triangle superimposed on three black circles and an inverted triangle. But this white triangle does not exist. There is no physical contour delimiting it. What exists are three "Pac-Men" — circles with a missing sector — and three corners of an incomplete

triangle.



The triangle emerges from the configuration of constraints. It is an example of what the literature calls "reification": the perceptual system constructs explicit spatial information that is not present in the sensory stimulus (Illusory contour, Wikipedia).

This phenomenon illustrates a profound principle: form is not in the components, but in their relationship. The triangle is not in the circles. It is not in the corners. It is in the intersection of the constraints imposed by these elements.

The same dynamic, we argue, governs the emergence of meaning in classified information.

### **3. Kant and the A Priori Conditions of Knowledge**

Immanuel Kant, in the Critique of Pure Reason (1781), posed a fundamental question: how is synthetic a priori knowledge possible? How can we know something about the world before experiencing it?

The Kantian answer is that there exist categories — structures of the mind — that are conditions of possibility for any experience. These categories (quantity, quality, relation, modality) do not derive from experience but make it possible (Stanford Encyclopedia of Philosophy, Categories).

Kant writes: "Thoughts without content are empty, intuitions without concepts are blind." There is nothing to know without empirical content; but without a priori structures, we have no way of giving intelligible form to whatever content we may have.

This insight applies directly to the problem of digital information. A document without classification is like an intuition without a concept: it exists, but cannot be known in the sense of being navigated, filtered, retrieved deterministically.

Classificatory dimensions — who produces, who can access, where it is geographically located, what it is about — are not accessory metadata. They are a priori conditions of referentiality.

## 4. Grounding Is Not Enough

Recent literature on RAG has precisely identified the limits of grounding. Bellante (2025) notes that "traceability is not an emergent attribute, but a characteristic deliberately engineered through grounding." But grounding presupposes that something exists to which to anchor.

A systematic survey from 2025 finds that "RAG breaks at scale because organizations treat it like a feature of LLMs rather than a platform discipline. The real challenges emerge in ingestion, retrieval optimization, metadata management, versioning, indexing, evaluation, and long-term governance" (arXiv:2507.18910).

The problem is not the model. It is the classificatory infrastructure.

Farquhar et al. (2024) proposed semantic entropy as a metric to detect hallucinations: if the LLM, varying the temperature, produces responses with radically different meanings, epistemic uncertainty is high. But this metric presupposes that there exists a ground truth against which to measure divergence.

RAGAS (Retrieval Augmented Generation Assessment) provides metrics such as Faithfulness and Context Precision (Es et al., 2023). But Faithfulness measures the coherence of the response with respect to the retrieved context — and what if the retrieved context is wrong because the classification was inadequate?

Grounding is necessary but not sufficient. Before grounding, we need the space in which grounding can operate.

## 5. Orthogonal Dimensions: The Space of Referentiality

In physics, time and space are the orthogonal dimensions within which any event can be located. An event without space-time coordinates is not an event — it is an abstraction devoid of referentiality.

Analogously, we propose that referable information requires orthogonal dimensions of classification. Not a single hierarchy (folder → subfolder → file), but a system of multiple and independent coordinates:

- Who produces the information (provenance, attribution)
- Who can access (visibility perimeter, control)
- Level of confidentiality (public/private gradient)
- Geographic location (where it is relevant)
- Semantic domain (what it is about)
- Specific topic (thematic sub-classification)

These dimensions are orthogonal in the mathematical sense: independent of each other. A document can be produced by A, accessible to B, confidential, relevant to Italy, in the legal

domain, on the topic of contracts. Each dimension adds information not derivable from the others.

The literature on faceted taxonomies confirms this intuition: "It has become more common to design faceted taxonomies than purely hierarchical taxonomies. Faceted taxonomies integrate better with search, serve for filters and sorting, and provide a good user experience" (Hedden Information Management).

But the value of orthogonal dimensions goes beyond usability. It is ontological.

## **6. Emergence by Subtraction: When Filtering Is Creating**

The Theory of Informational Emergence (TIE) proposes that "neither information nor structure is primary; instead, they are two inseparable aspects of a dynamic process" (ResearchGate, 2024).

The Information-Based Model of Emergent Structures (IBMES) advances further: "emergent phenomena result not from the properties of matter per se, but from patterns of information dynamically organizing and interacting under constraints" (ResearchGate, 2024).

But what does "under constraints" mean? Constraints do not add — they subtract. Like the sculptor who removes excess marble to reveal the form, orthogonal dimensions do not construct information: they filter it from noise. Every classification is a sieve that lets less pass through.

This perspective illuminates the value of orthogonal dimensions. They are not passive containers. They are subtractive constraints. The intersection of multiple dimensions generates information that is not present in any single dimension — exactly as the Kanizsa triangle emerges from perceptual constraints without being present in any individual element.

A document classified as [produced by X] + [accessible to Y] + [relevant to Z] + [on topic W] acquires a position in informational space. This position enables:

- Deterministic retrieval: finding exactly what is needed, not approximately
- Combinatorial filtering: intersecting multiple criteria
- Attribution: knowing who produced what
- Access control: ensuring that only those who should see, see
- Emergent relations: discovering non-explicit connections between documents sharing coordinates

The last property is crucial. When two documents share some dimensions but differ in others, a relation emerges that no system could deduce from the documents in isolation.

## **7. The Fractal of Classification: From Noise to Form**

Orthogonal dimensions are not static. Each dimension can contain sub-dimensions, which in turn can contain further levels. It is a fractal structure: self-similar at different scales.

A crucial point: the original state is not void — it is noise. In noise, everything is superimposed: air, water, data, signals, undifferentiated potential. Not "nothing," but "everything together." Air and water are not primary substances — they are categories, cuts in the continuum. Before the cut that defines "air," air does not exist. Only the noise containing every possibility exists.

The output itself of the classificatory system — the emergent relations, the discovered patterns — can become new input for further classification. The system does not just organize information; it filters noise to reveal form.

This recursive process can be visualized as an implosion from noise toward meaning:

```
NOISE (everything superimposed, undifferentiated)
  ↓ [first cut: distinction]
SOMETHING ≠ OTHER
  ↓ [classification: orthogonal dimensions]
CLASSIFIED DATA (coordinates in space)
  ↓ [intersection: emergent relations]
PATTERN (higher-order information)
  ↓ [meta-classification]
CLASSIFIED PATTERN
  ↓ [recursion: filters on filters]
FORM (what remains after n cuts)
```

At each level, the classificatory constraint does not add information — it reveals it by subtracting noise. Just as the three Pac-Men do not reduce the visual field but generate a triangle, so orthogonal dimensions do not limit data but reveal meaning.

## 8. Implications for AI: The Human Task in the Age of Machines

If the preceding analysis is correct, the implications for AI systems are profound.

The hallucination problem is not primarily a model problem. It is a problem of absence of coordinates. An LLM operating on an unstructured corpus is like a perceptual system without constraints: it can generate any form, but none will be anchored.

RAG is necessary but presupposes classification. Retrieval works if and only if retrieval criteria exist. These criteria are the orthogonal dimensions. Without them, retrieval is random.

Reference-free evaluation has intrinsic limits. RAGAS and similar frameworks measure internal coherence (faithfulness to context, relevance to query). But they cannot measure whether the right context was retrieved in the first place. Complete evaluation requires ground truth — which presupposes classification.

AI does not feel. A Large Language Model processes signs — billions of signs, with speed and pattern-matching no human can equal. But it does not feel where to cut. It enters the scene after someone has already felt, already distinguished, already classified. AI is processing power on the already-cut.

This redefines the human task in the AI era:

```
PRE-AI ERA:    human feels, human cuts, human processes
AI ERA:       human feels, human cuts, AI processes
```

The human task does not disappear — it concentrates. We cede processing to machines. But feeling? The first cut? Irreplaceable. Those who do not cut will be cut. Those who delegate even feeling become noise for others' AI.

## 9. Beyond Referentiality: Toward an Ontology of the Boundary

The preceding sections have treated orthogonal dimensions as conditions of referentiality — the ability to find, verify, attribute. But reification in the Kanizsa triangle suggests something more radical.

In the Kanizsa phenomenon, the perceptual system does not discover a triangle — it generates one. Perceptual constraints do not reveal a preexisting form; they produce it. This is the very definition of reification: the creation of something that was not there.

If we extend this logic, a question emerges: what is the minimal element of reification? What do the digital and the material have in common?

### *The Boundary as Fundamental Unit*

The answer is not the bit understood as 0 or 1. The bit is the boundary between 0 and 1. It does not represent a state — it represents the existence of a separation. Before "0" and "1" exist as values, the distinction that makes them possible must exist.

George Spencer-Brown, in *Laws of Form* (1969), formalized this principle: "Draw a distinction." The distinction precedes what it distinguishes. The boundary is ontologically primary; the sides are derived.

```
0 | 1
  ↑
THE BIT IS HERE
(the distinction itself)
```

This inverts the classical view:

```
CLASSICAL VIEW:    Substances → Boundaries (boundaries separate
substances)
INVERTED VIEW:    Boundaries → Substances (substances emerge from
boundaries)
```

Substance is derived illusion. The boundary is primary reality.

### ***The Sign as Persistent Boundary***

If the boundary is the minimal element, the sign is the boundary made persistent. Before the sign "air," air does not exist as a category. The sign does not describe a preexisting distinction — it creates it.

The prefix DIS- (Latin) captures this primordial gesture: dis-cernere (to separate with a sieve), dis-tinguere (to mark the difference), dis-crimen (that which separates, hence: crime, crisis, criterion). To see is already to separate. To perceive is already to cut.

### ***Feeling Before the Sign***

But something precedes even the sign: feeling. Before the brushstroke there is the painter's intuition. Before the chisel stroke there is the sculptor's vision. Before the sign there is the feeling of where the boundary wants to be.

The sign crystallizes feeling. It makes it transmissible, persistent. But the original cut happens in feeling itself — pre-verbal, pre-formal, pre-everything.

```
FEELING (pre-verbal distinction, intuition)
  ↓ [crystallizes]
SIGN (persistent, transmissible distinction)
  ↓ [accumulates]
TEXT (geometry of distinctions)
  ↓ [systemizes]
CLASSIFICATION (navigable space)
```

### ***From Noise to Matter: Saturation of Boundaries***

Contemporary physics offers convergent clues. John Archibald Wheeler proposed the principle "It from bit": information does not describe reality — it constitutes it. But more precisely: It from cut. Reality emerges from the primordial cut in noise.

The fractal of classification can be reformulated as progressive saturation of boundaries:

```
NOISE (zero boundaries: everything superimposed)
  ↓ [first cut: ≠]
TWO (something distinguishes itself from other)
  ↓ [second cut]
FOUR
  ↓ [n cuts]
COMPLEXITY (geometry of boundaries)
  ↓ [saturation: no more cutting possible]
MATTER (maximum density of distinctions)
```

In this perspective, matter is not a special substance — it is what remains when degrees of freedom are exhausted. It is noise so thoroughly cut that it cannot be otherwise. Matter is density of boundaries.

Hallucination, in this light, is not simply an error. It is information with insufficient boundaries. It can assume any form because it does not have enough cuts to fix it. It remains "gaseous."

RAG grounding is an attempt to add boundaries. But if the boundaries are few, poorly structured, or not orthogonal, information does not condense. A system with recursive orthogonal dimensions is, in this sense, a reification machine: it takes noise and pushes it through successive cuts until it becomes solid, referable, real.

This is not a metaphor. It is an ontological proposal: the boundary does not separate substances — it generates them.

## **10. Conclusion: Minimal Tool, Pure Class**

This article has proposed a reversal of perspective on three levels.

At the first level, technical: hallucination is not an AI bug to be corrected with better models. It is the symptom of a structural absence — the absence of orthogonal dimensions that allow information to exist in a navigable space.

At the second level, operational: AI does not feel. It processes signs, it does not generate them. The human task in the age of machines concentrates: feeling where to cut, imposing the first boundary. We cede processing. The original cut remains ours.

At the third level, ontological: the boundary is not a limitation imposed on substance — it is the very condition of its existence. Form emerges from the boundary. Meaning emerges from the intersection of boundaries. Matter, at the limit, emerges from the saturation of boundaries.

Just as the Kanizsa triangle emerges from perceptual constraints — not despite them but because of them — so every level of reality emerges from the preceding one through the imposition of cuts.

An orthogonal classification system is not a container of information. It is a minimal tool — the hoe without which the farmer cannot cultivate. It is pure class — Kantian to the bone: not content, but condition of possibility of content.

Kantian categories were a priori conditions of empirical knowledge. Orthogonal dimensions are a priori conditions of informational referentiality. And perhaps, pushing the reasoning, recursive boundaries are a priori conditions of reality itself.

The future of reliable AI does not primarily pass through more powerful LLMs or more sophisticated RAG. It passes through the classificatory infrastructure that makes both possible. It passes through those who know how to feel where to cut.

The constraint shapes AI. And perhaps, it shapes everything else.

Paper prepared for publication on Academia.edu

January 2026