Mapping the Matrix of Meaning: A Method for Measuring Information Translation in Science Communication

Author: Sophia Toubian
Department of Information Studies, University of California, Los Angeles
300 Charles E. Young Dr N, Los Angeles, CA 90095, USA

Corresponding Author: Sophia Toubian
Email: toubian@ucla.edu
Institutional Address: Department of Information Studies, UCLA, 300 Charles E. Young Dr N, Los Angeles, CA 90095

Acknowledgments

This research was conducted as part of doctoral work in the Department of Information Studies at UCLA. I am grateful to my committee and colleagues for their support.

Statements and Declarations

Funding

The author received financial support from the UCLA Graduate Division.

Declaration of Conflicting Interest

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical Approval

This study did not involve human participants or identifiable personal data. Ethical approval was therefore not required in accordance with institutional guidelines.

Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Availability of Data and Materials

The corpora used in this study consist of publicly available academic articles and media texts. Processed datasets and analysis materials are available from the corresponding author upon reasonable request.

Code Availability
Not applicable.

Authors' Contributions
This article was solely conceptualized, written, and revised by Sophia Toubian.

Mapping the Matrix of Meaning: A Method for Measuring Information Translation in Science Communication

Abstract: This paper introduces a mixed-methods framework for analytically measuring the transformation of scientific knowledge as it circulates from specialized to public contexts. The integration of computational tools (natural language processing, semantic similarity modeling, network analysis) with qualitative approaches (discourse analysis, rhetorical analysis, visual rhetoric) enables multidimensional analysis of how scientific information changes across domains. At the framework's core is the Accessibility-Accuracy Matrix, which quantifies linguistic accessibility, factual accuracy, and epistemic stability as scientific concepts traverse different media environments. Through case studies of epigenetics and Maslow's hierarchy of needs, the methodology demonstrates how scientific concepts undergo distinct transformation trajectories. The framework reveals that visual elements play a crucial role in stabilizing meaning, while linguistic simplification often correlates with epistemic drift. This approach advances empirical rigor in science communication research, providing researchers with systematic tools for tracing meaning as it moves between specialized and public spheres.

Keywords: Science Communication Metrics, Knowledge Transformation, Computational Discourse Analysis, Epistemic Stability in Media

I. Introduction

As scientific knowledge circulates from specialized domains like academic journals to public discourse, the systematic analysis of how information transforms becomes an increasingly urgent priority in science communication. The translation of scientific concepts across platforms, genres, and audiences involves linguistic changes as well as deeper shifts in meaning, authority, and epistemic function (Bucchi & Trench, 2014; Latour, 1987). Existing science communication frameworks often emphasize the tension between accessibility and accuracy, yet typically lack the methodological specificity required to measure these transformations rigorously (Schäfer & Metag, 2021). Without empirical tools capable of evaluating how scientific concepts are simplified, reframed, or repurposed in public discourse, science communication risks remaining intuitively directed and inconsistently applied (Neuendorf, 2017). This paper addresses this gap by introducing a mixed-methods framework designed explicitly to measure the transformation of scientific knowledge across academic and public domains.

The framework integrates computational tools—natural language processing (NLP), semantic similarity modeling, and network analysis—with qualitative methodologies such as discourse analysis, rhetorical analysis, and visual rhetoric. Combining these approaches allows analysis across multiple dimensions, from changes in word choice and rhetorical strategies to shifts in broader conceptual structures and narrative framings. Together, these combined techniques generate both diagnostic insights and transferable metrics that advance empirical research in science communication, offering an evidence-based methodology where intuitive assessments previously dominated.

At its core, this approach involves three methodological components. Computational NLP tools, including word embeddings (Mikolov et al., 2013), sentence similarity modeling (Reimers & Gurevych, 2019), and n-gram analysis, detect patterns in the restructuring of scientific discourse as it moves through diverse media environments. These computational methods provide scalable measures of semantic overlap, linguistic complexity, and framing strategies. Concurrently, qualitative methods drawn from discourse and conversation analysis (Fairclough, 1995; Gee, 2014) offer interpretive depth by examining the rhetorical production and contestation of authority, uncertainty, and controversy. Network analysis complements these techniques by mapping the evolving conceptual structures of scientific discourse, identifying patterns in how key terms cluster, fragment, or realign across academic and public contexts (Borgatti et al., 2009). As van Dijck (2013) notes, such networked architectures are not neutral; they embed social, technological, and epistemic dynamics that shape how scientific knowledge circulates and gains traction within broader media ecosystems.

This triangulated methodology addresses the inherent multidimensionality of information transformation. Quantitative methods alone identify structural changes but struggle to illuminate the underlying motivations or interpretive consequences (Neuendorf, 2017). Qualitative methods provide essential contextual insights yet often face scalability limitations. Network analysis, in turn, highlights structural patterns but requires interpretative grounding to explain shifts in meaning. Van Dijck (2013) emphasizes that media infrastructures and algorithmic logics actively condition how information is networked, surfaced, or suppressed, underscoring the need for critical frameworks that situate network patterns within sociotechnical systems. The iterative interplay among these methods forms a methodological ecology that balances computational detection with qualitative interpretation, enhancing both validity and depth of findings (Creswell & Clark, 2017; Tashakkori & Teddlie, 2003).

Central to this methodological blueprint is the Accessibility-Accuracy (A/A) Matrix, conceptualized as a composite metric integrating accessibility, accuracy, and epistemic stability. Accessibility quantifies linguistic legibility for non-specialist audiences; accuracy evaluates fidelity to original scientific claims; epistemic stability assesses the preservation of core meanings across transformations. The matrix thus operationalizes these dimensions, facilitating empirical comparison and analysis. The study applies this approach through two case studies, epigenetics and Maslow's hierarchy of needs—concepts selected specifically for their differences in disciplinary origin, transformation trajectories, and public communication dynamics.

The paper advances three core objectives: first, constructing an empirically robust methodological tool capable of measuring scientific knowledge transformation; second, evaluating the framework's utility and robustness through detailed case analyses; and third, establishing a replicable methodological template applicable across diverse scientific domains and communication contexts. These objectives guide the following research questions: How can the transformation of scientific knowledge between academic and public formats be measured systematically and replicably? Which computational and qualitative methods best capture shifts in accessibility, accuracy, and epistemic stability across domains? Can a composite metric like the A/A Matrix reliably reflect the nuanced nature of transformation in empirical analyses?

By systematically operationalizing these concepts, the proposed framework significantly enhances existing methodological rigor in science communication research, offering researchers practical tools for tracing meaning as it traverses between scientific and public spheres.

II. Literature Review: Existing Measurement Approaches

A. Approaches to Information Transformation and Loss

Scientific knowledge transforms significantly as it moves through platforms, genres, and diverse audiences. Scholars conceptualize these changes through varying lenses, including simplification, reframing, and epistemic drift. McLuhan (2003) argues that medium-specific forms influence the epistemic and affective dimensions of communicated knowledge. Innis (1991) further highlights how the temporal and spatial properties of communication media condition which knowledge is prioritized and preserved, thereby shaping public understanding.

From a science and technology studies (STS) perspective, Latour (1987, 1999) emphasizes that information does not passively transfer across contexts; instead, it is actively reshaped within sociotechnical systems. Documents, databases, institutions, and other infrastructural elements function as actors in this process, co-producing rather than merely conveying meaning (Bowker & Star, 1999). In the context of visual knowledge production, Drucker (2010, 2017) critically challenges assumptions of transparency and neutrality, illustrating that visualizations actively shape epistemic understandings, particularly as scientific concepts move from specialized contexts into broader public visual cultures.

Collectively, these perspectives highlight information transformation as a multidimensional, active process involving linguistic, visual, epistemic, and rhetorical dimensions, thus necessitating an integrative analytic approach.

B. Quantitative Methods in Information Studies

Quantitative methods effectively detect structural patterns in information movement, utilizing bibliometrics and scientometrics to map scientific concept diffusion within institutional networks (Leydesdorff, 2001; Moed, 2005). Content analysis systematically captures textual shifts across different contexts (Neuendorf, 2017), enabling tracking of linguistic transformations via tools like Voyant and n-gram modeling. However, these methods primarily focus on formal linguistic changes, often failing to sufficiently capture deeper rhetorical shifts or contextual nuances underlying meaning transformations.

Recent advances in computational linguistics, particularly natural language processing (NLP), offer sophisticated metrics for semantic similarity and textual comparison. Methods like word embeddings and transformer-based semantic similarity modeling (Mikolov et al., 2013; Reimers & Gurevych, 2019) quantify shifts in conceptual meanings, enabling precise, scalable measurement of how scientific content changes in accessibility and epistemic clarity across platforms. These methods build on foundational work in computational linguistics that models language as a probabilistic and structured system, integrating syntactic parsing, distributional semantics, and discourse-level representations (Jurafsky & Martin, 2023). Despite their analytical power, these computational tools often neglect rhetorical or ideological framing, emphasizing the necessity for integrative frameworks incorporating interpretive analysis.

C. Qualitative Methods in Science Communication

Qualitative research provides critical insights into the rhetorical, cultural, and epistemic dimensions often overlooked by quantitative analyses. Discourse analysis systematically explores how linguistic features construct social and epistemic realities, elucidating the rhetorical shifts in authority, uncertainty, and audience orientation as scientific content moves into public contexts (Fairclough, 1995; Gee, 2014).

Visual discourse analysis further examines how images and visualizations function rhetorically, structuring knowledge differently depending on context (Drucker, 2010; Latour, 1995). For instance, scientific diagrams or infographics do not simply represent concepts but actively argue or persuade, framing public perceptions of complex topics.

Rhetorical analysis complements these approaches by critically examining persuasive strategies used when communicating scientific concepts to diverse audiences. Drawing from classical rhetoric (Perelman & Olbrechts-Tyteca, 1969), this method illuminates how transformations in scientific communication frequently involve emotional appeals, ethos-building, and shifts in logical framing. Qualitative analyses thus provide essential interpretative depth that quantitative methodologies alone lack.

D. Mixed-Methods Frameworks and Integrative Designs

Neither quantitative nor qualitative methods alone adequately address the complexity inherent in tracking scientific knowledge transformations across domains. Mixed-method approaches combine these strengths to create comprehensive analytical frameworks. Triangulation (Denzin, 1978) is fundamental in mixed-method research, ensuring robust validation by enabling cross-verification between computational detection and qualitative interpretation.

Integrative mixed-method designs, including concurrent integrative frameworks (Creswell & Clark, 2017; Tashakkori & Teddlie, 2003), facilitate simultaneous quantitative precision and qualitative depth, effectively managing complexity without oversimplification. In applying such integrative logic, this study systematically combines quantitative tools like NLP and semantic similarity metrics with qualitative rhetorical and discourse analysis, creating a dynamic feedback loop that ensures analytical rigor.

E. Limitations of Current Approaches

Despite methodological advancements, current approaches exhibit several limitations. Computational methods frequently struggle to capture rhetorical nuances and context-dependent meanings. Even sophisticated NLP models often miss subtle rhetorical transformations critical to understanding shifts in public interpretation (Neuendorf, 2017).

Qualitative methods, although contextually insightful, suffer scalability limitations, restricting their generalizability. Additionally, existing methods typically adopt binary frameworks, oversimplifying relationships between accessibility and accuracy. Transformation processes involve complex interactions across multiple axes of accessibility, accuracy, and epistemic stability—dimensions inadequately represented in simpler binary models.

These methodological shortcomings justify the development of a robust integrative framework, capable of empirically mapping nuanced transformations in scientific meaning comprehensively.

III. Methodological Framework: Quantifying Scientific Transformation

This study seeks to understand what it means to measure how scientific knowledge morphs as it moves across publics, platforms, and formats. How can we trace the unraveling of a concept's epistemic coherence over time? What signals—linguistic, structural, or affective—might reveal the presence of transformation? What metrics can capture not only what is said, but what is reshaped, repurposed, and reframed? The methodological challenge, then, is to develop tools capable of capturing these processes not only descriptively but evaluatively, producing a framework that can visualize, compare, and ultimately guide communicative strategies in science translation.

To address these questions, this study constructs a composite methodological framework that integrates both quantitative and qualitative analysis. This system includes natural language processing (NLP), semantic similarity modeling, and network analysis, alongside discourse analysis, rhetorical theory, and visual modeling. At its center is the Accessibility-Accuracy (A/A) Matrix, expanded to include a third axis: epistemic stability. Each of these axes is operationalized through distinct analytical strands, allowing for the detection, modeling, and visualization of transformation across scientific and public discourse.

Rather than assuming a linear or reductive model of communication, this framework starts from the premise that meaning is not simply transmitted but produced. Transformation is not merely an effect of translation but a discursive and epistemic act. As scientific concepts migrate across media, institutions, and audiences, they undergo shifts in voice, scale, evidentiary framing, and rhetorical function. This framework does not seek to freeze these shifts but to make them legible—charting where, how, and to what epistemic consequence they unfold.

The following subsections detail the analytic procedures used to track transformation across texts. These methods include readability and linguistic compression measures, the identification of transformation typologies, semantic drift modeling, the implementation of NLP tools, and the design of case study corpora. Each method works in recursive relation with the others, contributing to a system that treats transformation not as a distortion, but as a complex process of epistemic reconfiguration.

A. Linguistic Legibility: Composite Readability Scoring

To operationalize the accessibility axis of the A/A Matrix, this study employs a composite approach to readability measurement. Accessibility here refers to the degree of linguistic legibility a text offers to non-specialist audiences. Rather than relying on a single index, this study uses four well-established readability formulas: the Flesch-Kincaid Grade Level, SMOG Index, Gunning Fog Index, and Coleman-Liau Index. These metrics are selected for their broad applicability in English-language corpora and their frequent use in science communication (DuBay, 2004; McLaughlin, 1969; Wang et al., 2013).

Each of these formulas estimates readability by measuring sentence length, syllable count, or character count, all of which contribute to the complexity or simplicity of the text. By aggregating the results from these indices, this study constructs a composite accessibility score that mitigates the biases of any one formula and enhances comparability across genres. For instance, a text may contain long sentences but use familiar vocabulary, or short sentences with dense technical terms. The composite score captures this interplay between different readability features, offering a robust indicator of linguistic accessibility (Friedman & Hoffman-Goetz, 2006).

This multi-metric approach also allows us to identify shifts in readability across versions of a scientific text, marking early evidence of translational movement. Significant changes in readability may indicate when epistemic content is compressed, reframed, or repositioned in response to shifting publics.

B. Typologies of Transformation: Translation, Transmutation, Transposition

While readability metrics capture linguistic legibility, they do not address the deeper epistemic reorientations that occur when scientific knowledge is adapted across publics, genres, and institutional contexts. These reconfigurations involve not just simplifying or altering language but shifting the very moral grammars, affective regimes, and institutional logics that underlie scientific discourse. To capture these shifts, this study applies a typological coding scheme grounded in the transformation modes outlined in the companion theory paper: translation, transmutation, and transposition.

Translation refers to the process by which scientific knowledge is adapted for new audiences, often through simplification, metaphor, or narrative restructuring. This mode involves reducing technical language and reframing concepts to make them accessible to non-expert audiences. Indicators of translation include a reduction in technical vocabulary, the increased use of metaphors, and shifts in readability scores.

Transmutation involves reframing scientific information to serve normative, commercial, or emotional agendas. This mode often amplifies or distorts uncertainty to align with public narratives or agendas. Key indicators include affective language, causal overstatements, and the introduction of moral or value-laden terms.

Transposition refers to the migration of scientific knowledge into new institutional or genre contexts. This transformation is evident when scientific content moves from academic settings to more public-facing formats like marketing materials, educational content, or policy documents. Indicators of transposition include genre shifts, changes in authorial voice, and the use of institutional language.

Rather than treating these modes as discrete categories, this study treats them as overlapping vectors. Their interaction produces layered transformations, capturing both linguistic and epistemic shifts as scientific knowledge circulates through different domains.

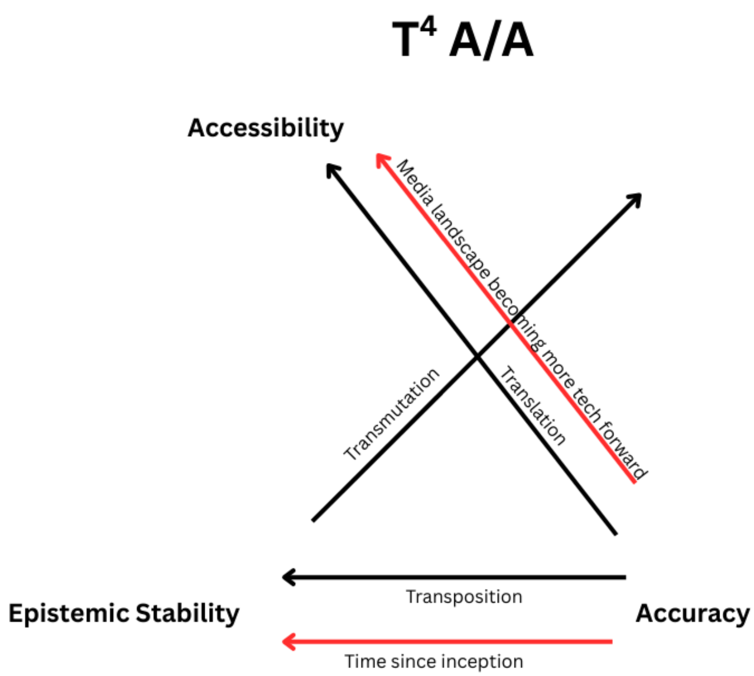

Figure 1. Conceptual diagram of the T⁴ A/A spectrum, showing three transformation pathways—translation, transmutation, and transposition—mapped against accessibility, accuracy, and epistemic stability. Temporal and media dynamics are represented by red directional vectors.

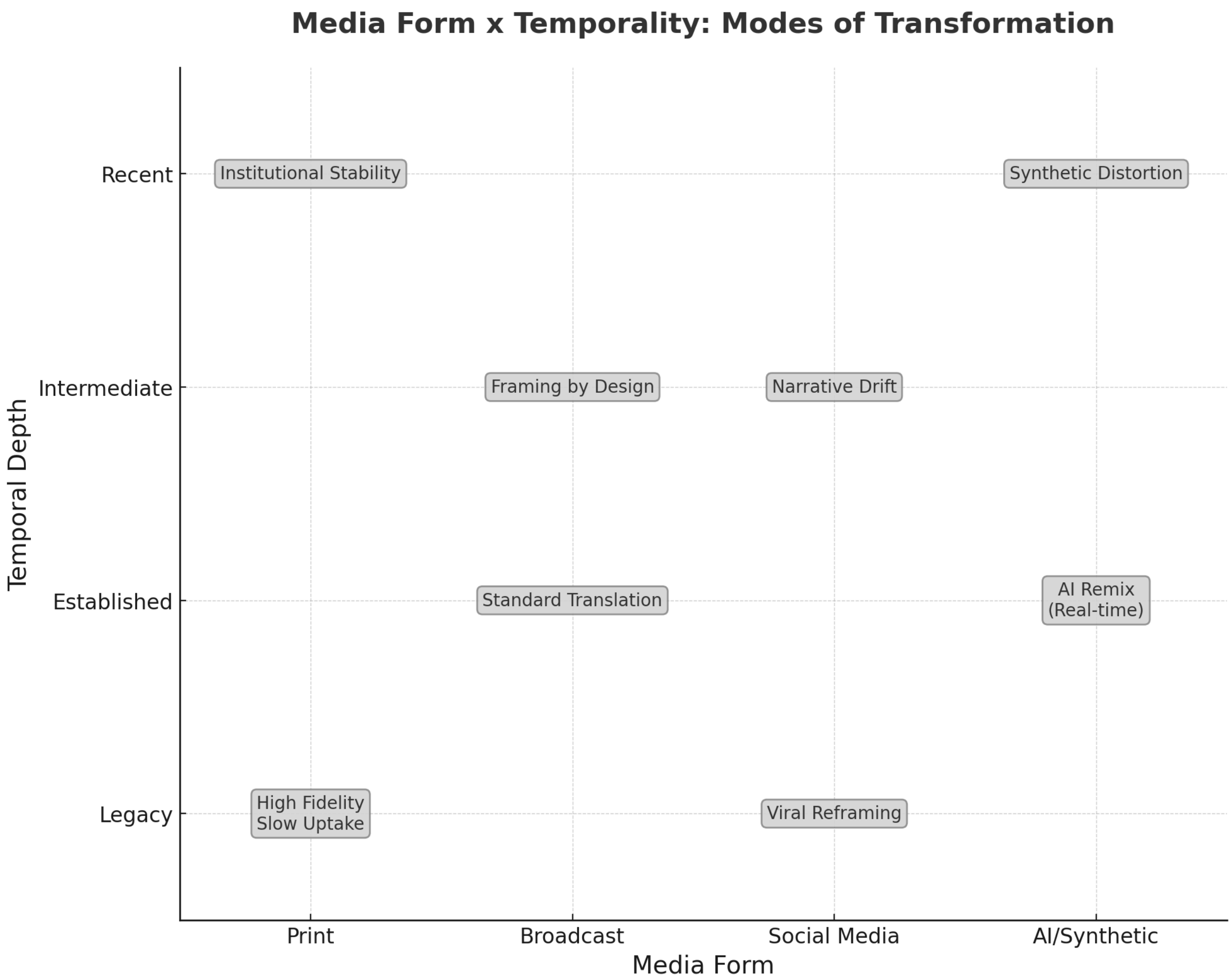

Figure 2. Matrix of media form by temporal depth, identifying dominant modes of transformation across communicative settings. The chart situates forms such as "viral reframing" or "synthetic distortion" within the T⁴ framework.

C. Tracing Stability: Semantic Drift and Visual Persistence

The final axis of the Accessibility-Accuracy Matrix, epistemic stability, presents the most elusive challenge in measuring scientific transformation. Epistemic stability refers to the resilience of core meanings as scientific knowledge circulates through different media, genres, and formats. Does the concept retain its foundational epistemic commitment as it is adapted and reframed, or does it acquire unintended meanings or shifts that affect its epistemic orientation?

To trace epistemic stability, this study uses a two-pronged approach: semantic similarity modeling and visual/rhetorical tracking. Using models like SBERT (Sentence-BERT) or doc2vec, this study computes cosine similarity between the original scientific text and its transformed versions. The cosine similarity score measures how similar the vectors representing the original and transformed texts are. A significant drop in similarity indicates conceptual drift, suggesting that the core idea has undergone substantial change or loss in meaning. Conversely, stable similarity values across transformations suggest that the core concept has maintained its integrity despite rhetorical or structural adjustments.

Alongside textual analysis, this study also tracks the visual and rhetorical components that may signal the persistence of epistemic stability. This includes the recurrent use of specific metaphors, diagrams, and conceptual markers that help stabilize meaning across transformations. For example, the use of Maslow's hierarchy as a pyramid in various public-facing media suggests that visual persistence supports the epistemic stability of the concept, even when the surrounding textual framing may shift. This dual approach allows us to capture both the linguistic and visual elements that contribute to the persistence or alteration of scientific meaning.

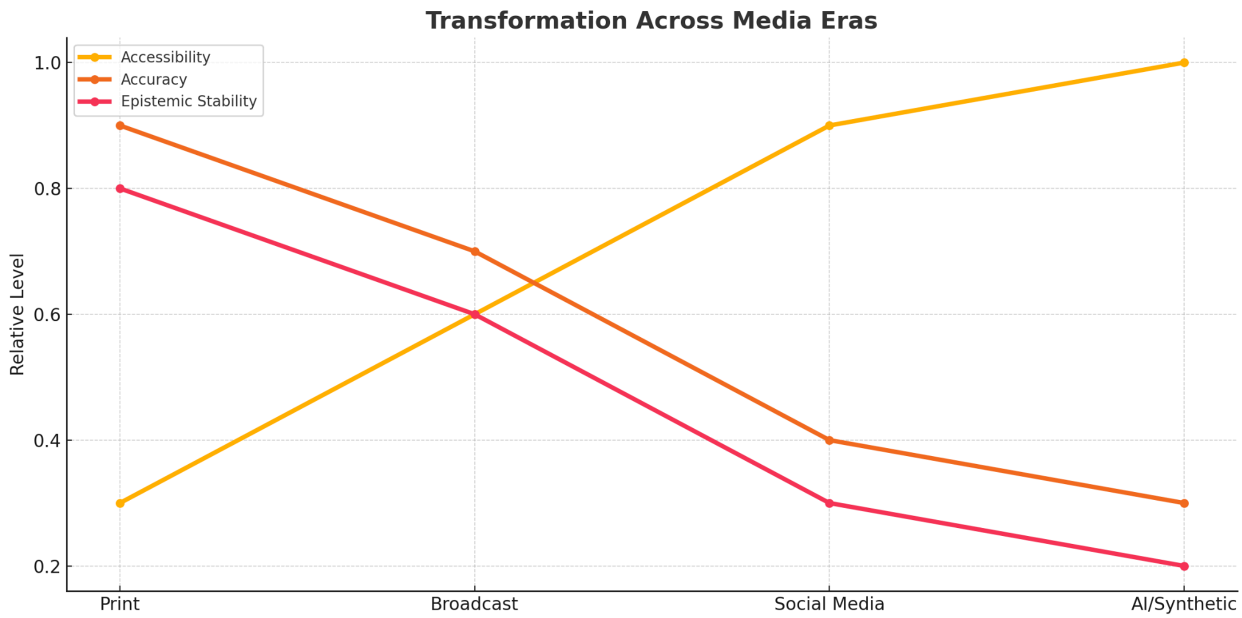

Figure 3. Relative levels of accessibility, accuracy, and epistemic stability across four media eras: print, broadcast, social media, and AI/synthetic platforms. The figure highlights a shift toward accessibility and away from epistemic rigor as platform logics change.

By combining these tools, this study can identify critical moments where epistemic stability shifts significantly, marking the points at which meaning moves from being faithful to the original to becoming ideologically or rhetorically reoriented. This method recognizes that transformation is not a linear process but a dynamic and layered shift in meaning.

D. Computational Implementation Using AI

The Accessibility-Accuracy Matrix relies heavily on computational tools to scale the analysis of scientific transformation. AI plays a pivotal role in this system, not by replacing human judgment but by structuring and expediting the identification of key transformations. AI helps identify patterns in large datasets, supporting qualitative interpretation, and assisting in quantifying how scientific knowledge shifts across contexts.

The process begins with the generation of composite accessibility scores using readability formulas. These scores, calculated using four established metrics (Flesch-Kincaid, SMOG, Gunning Fog, and Coleman-Liau), provide a measure of how accessible a given text is to non-specialist audiences. By aggregating the results from these formulas, a unified accessibility measure is created that facilitates comparison across texts, domains, and formats.

To evaluate accuracy, this study employs transformer-based language models to detect epistemic drift. Specifically, these models analyze changes in causal language (e.g., shifts from "may be associated with" to "causes"), the presence of hedging terms, and the loss of methodological context. These linguistic features are essential for determining whether scientific uncertainty is overstated or downplayed in public-facing texts.

For epistemic stability, this study uses Sentence-BERT (SBERT) or doc2vec models to compute semantic similarity scores between original texts and their transformed counterparts. This method provides a quantitative measure of conceptual fidelity, highlighting sections where core meanings have diverged significantly. Network analysis further enhances this by mapping the co-occurrence of scientific concepts across platforms, helping us trace the persistence of key terms and ideas as they circulate in different contexts.

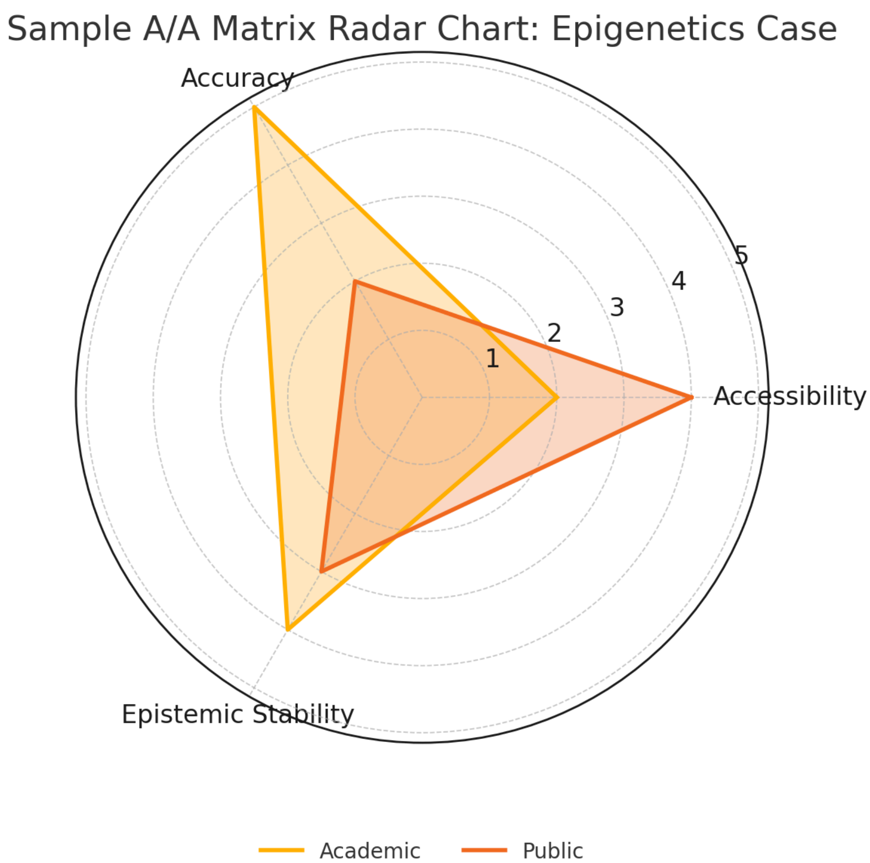

Figure 4. Comparative A/A matrix chart of academic and public presentations of epigenetics. Academic renderings maintain high accuracy and epistemic stability, whereas public forms emphasise accessibility with reduced fidelity.

Throughout this process, AI functions as a tool for organizing and scaling analysis, but its outputs are not taken as definitive answers. Instead, they serve as heuristics, guiding human interpretation and refinement. This iterative process ensures that AI remains a complementary tool rather than a replacement for human insight.

Table: Agents of Scientific Transformation

Agent	Primary Transformation	Typical Target Field	Example
Academic Scientists	Translation	Peer-reviewed journals	Weaver et al. (2004) study on maternal care
Science Journalists	Translation / Transposition	Popular science magazines, blogs	*Scientific American*, TED Talks
PR / Communications Teams	Transposition	Press releases, university websites	Institutional announcement of findings
Educators / Curriculum	Transposition	Textbooks, slides, LMS platforms	Maslow pyramid in classroom materials
Influencers / Content Creators	Transmutation	Social media, podcasts, YouTube	Trauma TikToks referencing epigenetics
Therapists / Coaches	Transmutation / Transposition	Self-help, wellness, UX design	Epigenetics in coaching or trauma healing
Policymakers / Funders	Transposition	Policy briefs, funding narratives	Childhood trauma funding initiatives

Table 1. Table of transformation agents showing how various actors (scientists, journalists, educators, influencers, therapists, policymakers) engage with and reshape scientific information, with examples of typical target fields and formats.

E. From Dimensions to Diagrams: Mapping the A/A T^4 Matrix

To synthesize the measurements obtained through accessibility, accuracy, and epistemic stability, each communicative instance is plotted on a triangular radar chart representing the

Accessibility-Accuracy (A/A) Matrix. This visualization captures how scientific meaning shifts across the three axes: accessibility, accuracy, and epistemic stability.

Each axis of the A/A Matrix is derived from specific metrics: Accessibility is quantified through the composite readability score, which combines results from four readability formulas. Accuracy is assessed through the retention of core claims, including methodological qualifiers and hedging terms. Epistemic Stability is modeled through semantic similarity scores, which compare the original text to its transformed versions.

Each communicative artifact is represented as a point on the radar chart, with its position reflecting its performance across these axes. For example, a public-facing text that simplifies scientific content for broader audiences might score high on accessibility but lower on accuracy and epistemic stability. A technical journal article, in contrast, might score highly on accuracy and epistemic stability but remain less accessible to general audiences.

This method not only visualizes transformation but also highlights the trade-offs inherent in each communicative choice. By mapping these transformations, the radar chart provides a clear, comparative view of how meaning evolves across various media and formats.

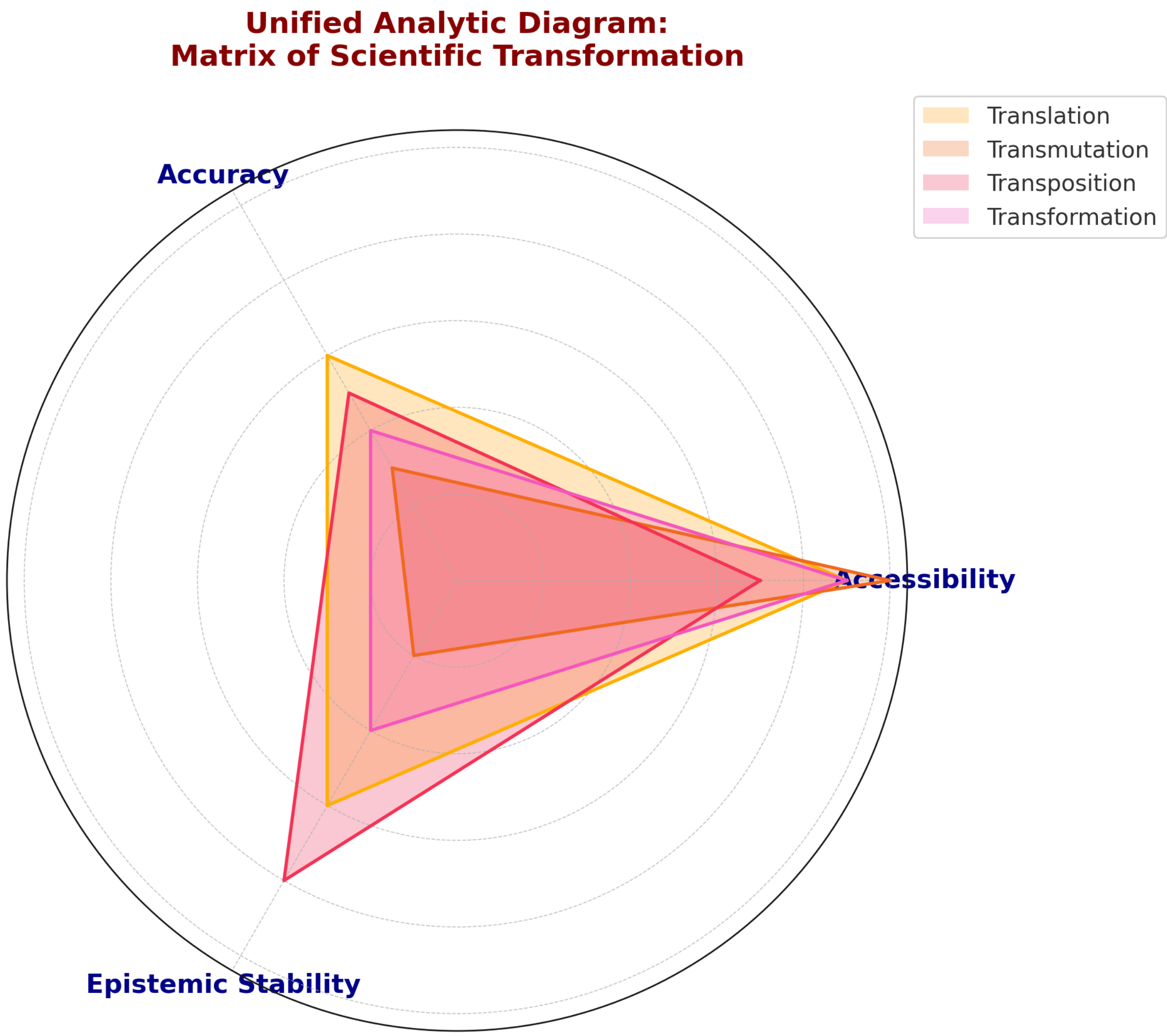

Figure 5. Unified radar chart visualising the accessibility-accuracy-epistemic stability profile of each transformation type: translation, transmutation, transposition, and transformation. Serves as a methodological template for evaluating science communication across forms.

F. Corpus Design: Case Selection, Construction, and Media Tracing

A central component of this study is the construction of matched corpora for the case studies on epigenetics and Maslow's hierarchy. The corpora are designed to reflect both the epistemic foundations of each concept and their recontextualized forms as they circulate through public discourse. The goal is not exhaustive coverage but representative diversity—capturing a range of rhetorical forms, institutional functions, and discursive genres that enable the analysis to detect meaningful patterns of transformation.

The academic corpus for each case includes 50 to 75 texts, primarily drawn from peer-reviewed journals and foundational works in the field. These texts are selected to reflect the disciplinary and methodological origins of the concept, including key studies, literature reviews, and meta-analyses.

The public corpus consists of 75 to 100 texts, including news articles, science blogs, infographics, TED Talks, YouTube videos, and social media posts. These texts are chosen for their relevance to the case study concepts and their engagement with the core scientific claims, whether through explanatory journalism, advocacy, or emotional reframing. The corpus spans a range of genres and platforms, reflecting the diverse ways in which scientific knowledge is transformed and communicated to different audiences.

By creating matched corpora across academic and public texts, this study ensures that it captures both the initial conceptualization of the scientific idea and its recontextualization in broader discourse. This dual-corpus design allows for cross-domain comparison, highlighting both the internal consistency of academic knowledge and the ways it evolves in public-facing formats.

Corpus	Sources	Estimated Size	Time Frame	Selection Criteria
Academic	Peer-reviewed journals, literature reviews, meta-analyses, conference proceedings	50–75 texts	2004–2024	Scientific rigor, conceptual relevance, internal debates
Public	News articles, blogs, TED Talks, YouTube videos, Instagram posts, wellness media	75–100 texts	2004–2024	Engagement with epigenetic concepts, accessibility, rhetorical framing

Table 2. Overview of corpus sources used in the case studies on epigenetics and Maslow's hierarchy of needs. The table details corpus type, estimated size, time frame, and selection criteria, distinguishing academic and public texts to enable comparative analysis of transformation across scientific and public domains.

G. Multi-Strand Analysis Procedure: Quantitative and Qualitative Layers

The analysis of each text in the corpus follows a multi-stage procedure that combines quantitative and qualitative approaches. First, texts are preprocessed using standard NLP pipelines, including tokenization, sentence segmentation, and syllable or character counts. Readability scores are calculated using the four established indices, which are normalized and averaged to produce a composite accessibility score.

Next, the accuracy of each public-facing text is assessed by comparing it to the original scientific sources. Key shifts in language—such as the removal of hedging terms or the simplification of causal language—are identified and analyzed. Accuracy is evaluated on a scale, with higher scores indicating greater fidelity to the original claims.

In the third stage, qualitative coding is applied to identify transformation modes: translation, transmutation, and transposition. These modes are operationalized through a set of linguistic and rhetorical markers, such as the simplification of technical vocabulary (translation), the introduction of emotional or moral framing (transmutation), and genre shifts (transposition).

Finally, epistemic stability is assessed using both semantic similarity modeling and visual persistence tracking. The stability of core concepts is evaluated by comparing semantic vectors and by examining the recurrence of specific diagrams or metaphors that anchor the meaning of the concept.

This multi-strand approach allows for a nuanced analysis of transformation, combining the precision of quantitative methods with the interpretive depth of qualitative analysis.

H. Evaluation and Validation

1. Reliability Assessments

Reliability is a key concern in mixed-methods research. To ensure the consistency of the analysis, multiple coders independently annotate the transformation modes—translation, transmutation, and transposition—using a structured rubric. Inter-coder agreement is measured using Cohen's kappa, with disagreements resolved through collaborative adjudication. This process ensures that the application of codes is systematic and reproducible.

On the computational side, reliability is maintained by cross-validating the outputs across different tools and metrics. For example, readability scores are measured using four standardized indices, and semantic similarity is verified through manual alignment of sentence pairs. These consistency checks ensure that model variance does not obscure significant patterns in the data.

2. Validity Considerations

The validity of the proposed framework is assessed through both face validity and construct validity. Face validity is tested by comparing the outputs of the model—such as the radar charts and similarity scores—with expert judgments about whether a transformation has occurred and in what form. Construct validity is supported by the coherence of the A/A Matrix itself, ensuring that each axis—accessibility, accuracy, and epistemic stability—corresponds to the theoretical constructs outlined in the companion theory paper.

Finally, triangulation is embedded throughout the analytic process. Quantitative findings are cross-checked with qualitative interpretations to ensure alignment and detect any contradictions. This triangulation ensures that no single method or tool dominates the analysis, allowing for a comprehensive and rigorous evaluation of scientific transformation.

IV. Case Study: Epigenetics and the Transformation of Scientific Meaning

A. Rationale for Selecting Epigenetics

Epigenetics provides a particularly rich context for studying the transformation of scientific knowledge. This interdisciplinary field bridges molecular biology, environmental science, and psychology, exploring how environmental factors influence gene expression and how these changes can be inherited across generations. Epigenetic research has profound implications for understanding health, behavior, and development, making it highly relevant in both academic and public spheres (Weaver et al., 2004; Lappé, 2016).

The complexity of epigenetics—spanning molecular mechanisms, developmental biology, and social determinants of health—makes it an ideal case study for examining how scientific knowledge is transformed across domains. The field is also characterized by significant conceptual slippage: while scientific discourse maintains a high degree of precision, public discussions often simplify, reframe, or even distort key concepts to suit broader cultural or political agendas.

By focusing on epigenetics, this case study offers a window into how scientific knowledge moves from technical, experimental contexts to broad public consumption. The transformation of epigenetic knowledge highlights shifts in authority, epistemic stability, and conceptual clarity as it is translated for non-expert audiences and recontextualized within social, political, and media narratives.

B. Corpus Construction and Design

The epigenetics case study uses a dual-corpus design, constructed to reflect both the disciplinary origins of epigenetics and its recontextualized forms in public discourse. The academic corpus includes 50 to 75 key texts, sourced from journals, foundational studies, literature reviews, and meta-analyses. This selection captures the scientific framing of epigenetics, focusing on peer-reviewed articles that emphasize experimental methods, genetic mechanisms, and theoretical foundations.

The public corpus consists of 75 to 100 texts spanning a range of formats and platforms: news articles, science blogs, TED Talks, YouTube videos, social media posts, and wellness websites. These texts represent how epigenetics is communicated to the general public, often simplifying complex scientific concepts, invoking metaphors, and addressing public concerns such as the environment’s effect on health and behavior. The corpus spans from 2004 to the present, tracking the evolution of epigenetic discourse as it gains visibility and mainstream relevance.

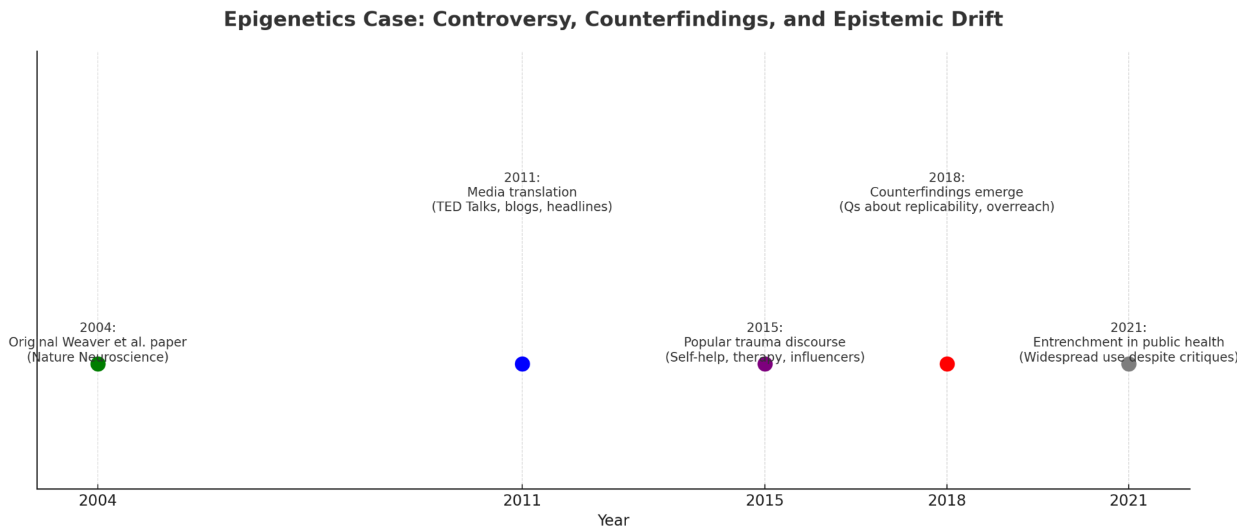

Figure 6. Timeline tracing the discursive trajectory of epigenetics from 2004 to 2021. The figure illustrates narrative drift, public uptake, and the emergence of controversy around replicability and epistemic overreach.

By structuring the corpora this way, this study is able to compare the scientific discourse with its public-facing counterparts, enabling the identification of transformation patterns in both content and framing.

C. Analytical Implementation

Both corpora undergo systematic NLP preprocessing (tokenization, lemmatization, sentence segmentation). Composite readability scores (Flesch-Kincaid, SMOG, Gunning Fog, Coleman-Liau) quantify linguistic accessibility for comparative analysis. Semantic similarity modeling (SBERT) compares public-facing texts to their academic counterparts, quantifying epistemic drift by identifying shifts in core conceptual meanings.

Qualitative coding systematically identifies transformation typologies (translation, transmutation, transposition). For instance, translation typologies appear through simplified language and metaphorical reframing (gene as “blueprints”); transmutation emerges when texts amplify causal claims beyond original scientific evidence; and transposition occurs in shifts from academic genres to public-facing media formats.

Visual and rhetorical analysis specifically evaluates visual metaphors, diagrams, and imagery, assessing how visual representations stabilize or alter conceptual meanings across the corpus.

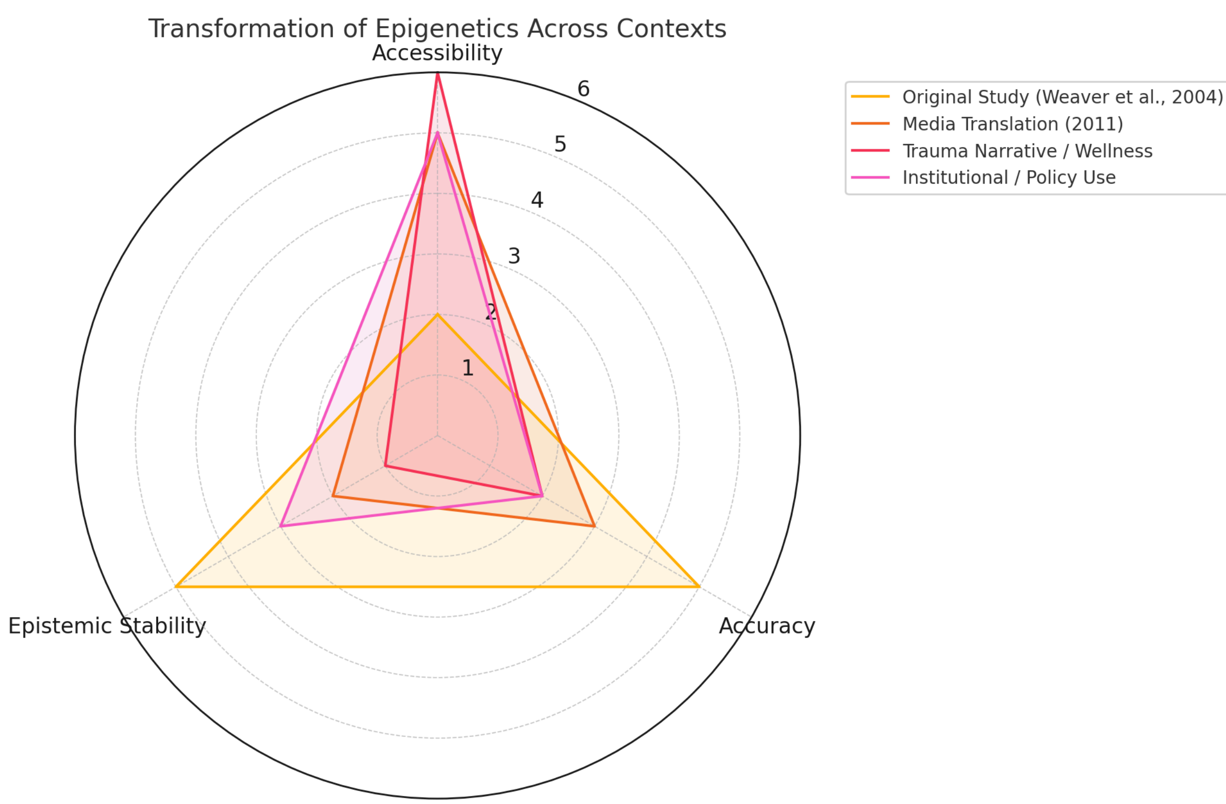

Figure 7. Transformation of epigenetics across four interpretive contexts (original study, media translation, trauma/wellness discourse, and policy use), visualised using the A/A matrix radar chart. Higher accessibility is associated with declines in accuracy and epistemic stability.

V. Case Study: Maslow's Hierarchy and the Visual Stabilization of Psychological Knowledge

A. Rationale for Selecting Maslow's Hierarchy

Maslow's hierarchy of needs is a compelling example of how scientific knowledge can become visually stabilized and persist across time. Originally proposed in 1943, Maslow's model has become one of the most recognizable frameworks in psychology, often represented as a pyramid with basic physiological needs at the base and self-actualization at the top. While the hierarchy has been critiqued and revised within academic psychology, it remains widely used in public discourse, particularly in education, business, and marketing (Kenrick et al., 2010; Wahba & Bridwell, 1976).

Maslow's hierarchy offers a distinct type of transformation compared to epigenetics. Instead of rapid conceptual shifts, the transformation of Maslow's hierarchy is characterized by its visual persistence. The pyramid diagram, widely reproduced in textbooks, infographics, and self-help materials, serves as a visual anchor that maintains the core meaning of the hierarchy, even as the specific details of the theory are simplified or reframed for different audiences.

This case study focuses on how Maslow's hierarchy is transformed through visual and rhetorical recontextualization, examining how its meaning persists across different domains despite evolving interpretations.

B. Corpus Construction and Design

The corpus for Maslow's hierarchy is divided into two parts: an academic corpus and a public corpus. The academic corpus consists of 50 to 75 texts, including foundational works by Maslow, as well as more recent critiques and adaptations of the hierarchy within psychology and related fields. This corpus focuses on the theoretical underpinnings of Maslow's model, exploring how the hierarchy has been refined and debated in academic settings.

The public corpus includes 75 to 100 texts sourced from a variety of platforms: business leadership books, educational resources, self-help articles, corporate training materials, and online infographics. The public-facing representations of Maslow's hierarchy often simplify the theory, using the pyramid diagram as a tool for explaining motivation, personal development, and organizational behavior. The visual form of the pyramid is a critical aspect of its transformation, and these texts offer insight into how the model is adapted to suit different cultural and institutional needs.

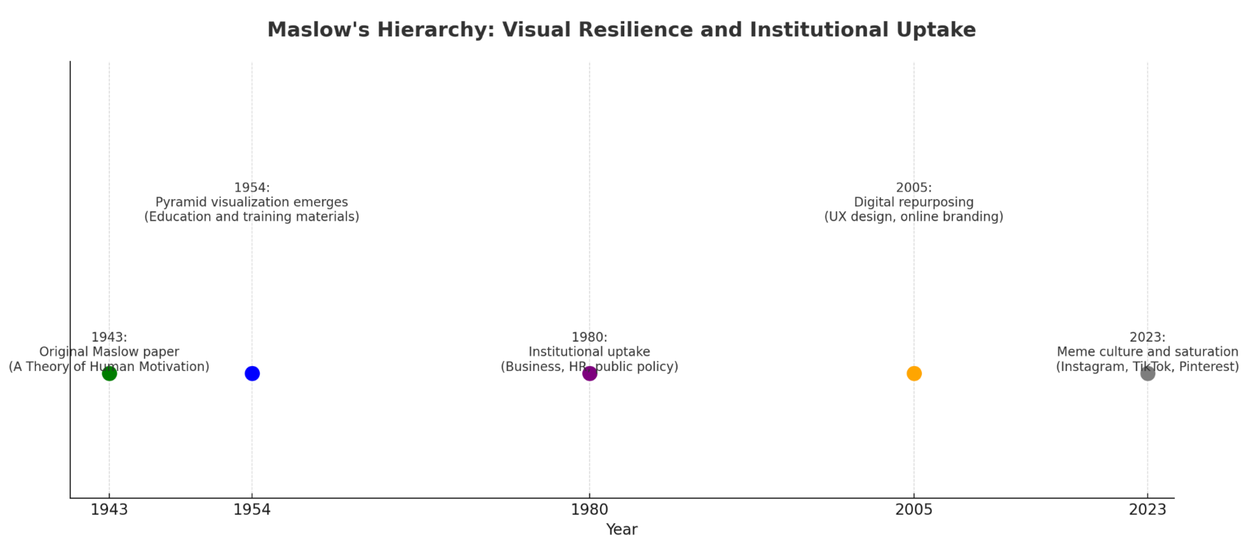

Figure 8. Timeline of Maslow's hierarchy of needs from its 1943 inception to its memeified and institutionalised forms in 2023. The chart visualises its enduring uptake through visual simplification and digital repurposing.

The temporal frame for the Maslow corpus spans from 1943 to the present, with a focus on the period after 2000, when the pyramid became widely used in corporate and educational settings. By comparing these academic and public corpora, this study can track how Maslow's hierarchy is recontextualized as it moves from academic psychology to broader public and institutional applications.

C. Analytical Implementation

Analytical procedures mirror the epigenetics case study. NLP preprocessing enables composite readability assessments, facilitating comparisons of linguistic accessibility between academic and public texts. Semantic similarity modeling quantifies epistemic drift by directly comparing textual representations in public versus academic corpora.

Qualitative coding categorizes transformations through the typological framework. Translation manifests as simplified or motivational language within corporate or educational contexts. Transmutation emerges through exaggerated claims emphasizing personal achievement or organizational efficiency. Transposition occurs clearly in genre shifts from psychological theory toward management and self-help discourses.

Visual analysis specifically tracks pyramid representations across texts, assessing visual persistence. Analysis identifies instances where stable visual metaphors anchor meaning, buffering conceptual drift occurring within textual discourse, thus highlighting visual rhetoric's critical stabilizing function.

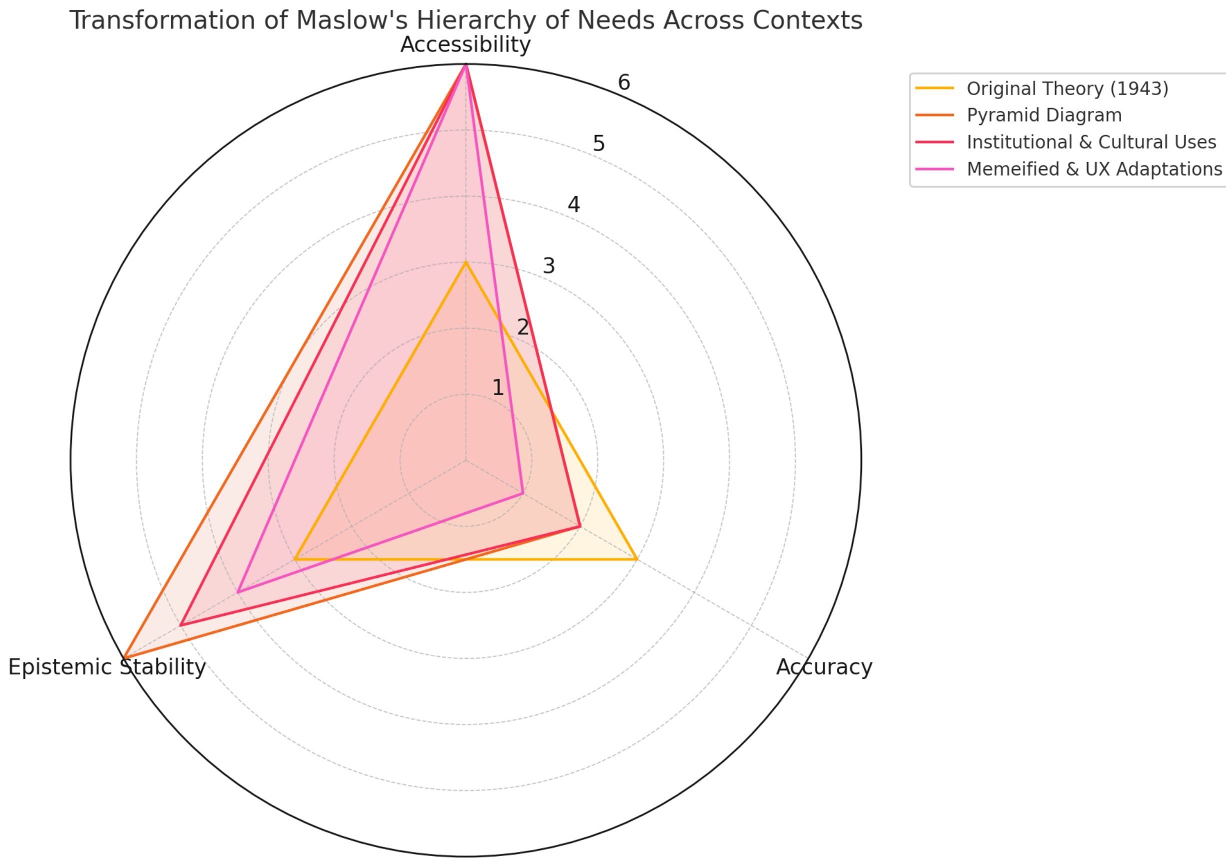

Figure 9. A/A matrix radar chart showing how Maslow's hierarchy transforms across modalities: from theory to pyramid, to institutional discourse, to meme and UX design. Each phase shows a trade-off between accessibility and either accuracy or epistemic stability.

VI. Discussion

A. Interpretation of Results and Methodological Insights

The case studies of epigenetics and Maslow's hierarchy demonstrate that scientific knowledge transformation involves complex interactions between linguistic, rhetorical, and visual dimensions. In epigenetics, significant epistemic drift arises from public discourses simplifying probabilistic language into deterministic or causal claims (e.g., stating environmental factors "cause" genetic changes rather than merely influencing them). Such linguistic shifts consistently correlate with increases in readability, revealing a clear trade-off between accuracy and accessibility.

In contrast, Maslow's hierarchy exemplifies visual stabilization despite textual simplifications or distortions. Although public texts frequently simplify or modify Maslow's original psychological theory, the persistent pyramid diagram effectively stabilizes the concept's epistemic core, reducing conceptual drift even amidst significant textual variability. Visual rhetoric thus emerges as an integral epistemic anchor, influencing public understanding more effectively than textual precision alone.

The mixed-methods framework used here proves methodologically robust, successfully integrating quantitative NLP-based analyses with qualitative rhetorical and visual insights. The iterative interplay between computational detection and qualitative interpretation provides multidimensional clarity essential for accurately assessing scientific transformations.

B. Challenges in Measurement

While the mixed-methods framework offers a comprehensive approach to measuring transformation, several challenges emerged during the analysis. One of the primary challenges was ensuring consistency across the different methods. The computational tools, while powerful, are not without their limitations. For example, semantic similarity modeling can detect structural shifts in meaning but may not capture the subtleties of rhetorical framing or emotional resonance that can be crucial in understanding how meaning is transformed. In contrast, qualitative analysis provides rich contextual insights but is often more subjective and harder to scale.

Additionally, the readability formulas used to assess accessibility are valuable but not foolproof. These formulas rely on formal linguistic features such as sentence length and syllable count, which may not fully capture the complexities of accessibility in science communication. For instance, a text with shorter sentences and simpler vocabulary may still be conceptually challenging or ideologically loaded, presenting a challenge for straightforward readability assessments.

Another challenge is the integration of visual elements, especially in the case of Maslow's hierarchy. While visual persistence plays a key role in stabilizing the meaning of the hierarchy, quantifying this persistence can be difficult. Visual elements like diagrams or metaphors are harder to analyze computationally than textual features, and their impact on epistemic stability is often more subtle and contextual.

These challenges highlight the importance of iterative refinement in the application of the framework, ensuring that each method compensates for the limitations of the others and providing a more holistic understanding of transformation.

C. Interdisciplinary Application and Transferability

The developed methodological framework's interdisciplinary flexibility makes it widely transferable to various scientific domains where public communication significantly reshapes meaning. Fields such as public health communication, environmental policy, and science education benefit directly from understanding epistemic stability and drift across discursive contexts. For example, applying this methodology to climate change or vaccination discourses could elucidate critical rhetorical transformations impacting public beliefs and actions.

Transferability depends heavily on careful corpus selection and typological customization to context-specific rhetorical transformations, but the framework's inherent flexibility ensures broad applicability across diverse communicative environments.

D. Strengths, Limitations, and Integration with Existing Methods

The study's primary strength lies in its systematic integration of quantitative precision and qualitative interpretive depth. The triangulated approach substantially enhances analytical rigor, surpassing simpler single-method approaches.

However, limitations include potential computational biases inherent in semantic modeling and readability indices, occasionally misrepresenting nuanced epistemic shifts. Additionally, although effective, the dual-case study approach may not fully represent transformations occurring across all scientific disciplines, suggesting future methodological refinements and expanded corpus designs.

Despite limitations, this approach significantly extends current methodological frameworks in science communication literature, offering a robust model for future studies to follow and refine.

E. Initial Findings and Visual Media Observations

One of the most striking findings from the case studies was the role of visual media in stabilizing the meaning of scientific concepts. In both epigenetics and Maslow's hierarchy, visual elements such as diagrams, charts, and metaphors played a crucial role in maintaining epistemic stability. The pyramid diagram of Maslow's hierarchy, for example, was visually reproduced across a wide range of media, helping to solidify the concept in public consciousness. Similarly, in the case of epigenetics, the use of visual metaphors—such as "genetic blueprints"—served to simplify complex scientific ideas, making them more accessible while simultaneously reframing the epistemic function of the concept.

These observations highlight the power of visual media in science communication and underscore the importance of including visual analysis in any comprehensive study of transformation. Visual elements not only simplify but also structure meaning, guiding how audiences interpret and engage with scientific content.

VII. Conclusion and Future Directions

This study presents a new mixed-methods framework for analyzing the transformation of scientific knowledge across academic and public domains. By combining computational tools like natural language processing and semantic similarity modeling with qualitative methods such as discourse and rhetorical analysis, the framework provides a comprehensive approach to understanding how scientific concepts are adapted, simplified, and reframed in public discourse.

The case studies of epigenetics and Maslow's hierarchy illustrate the utility of the framework in capturing both structural shifts and cultural recontextualizations of scientific knowledge. However, several challenges remain, particularly in the integration of visual media and the quantification of epistemic stability. Future research should refine the methodological tools to address these challenges, particularly by expanding the range of case studies to include more diverse scientific fields and media formats.

Further work could also explore the application of the framework to specific communication contexts, such as public health campaigns, environmental advocacy, or science-policy interfaces. The framework's potential to trace transformation in real-time across different media could offer valuable insights into how scientific ideas evolve and influence public attitudes and behaviors.

Ultimately, this research contributes to the growing understanding of how scientific knowledge is communicated, transformed, and received, providing a more systematic and replicable method for studying these processes.

References

Bender E and Gebru T (2021) On the dangers of stochastic parrots: Can language models be too big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21), Virtual Event, Canada, 3–10 March 2021, pp. 610–623. New York: ACM.

Borgatti SP, Everett MG and Johnson J (2009) Analyzing Social Networks. London: Sage.

boyd d (2010) Social network sites as networked publics: Affordances, dynamics, and implications. In: Papacharissi Z (ed) Networked Self: Identity, Community, and Culture on Social Network Sites. New York: Routledge, pp. 39–58.

Bucchi M and Trench B (2014) Science Communication: A Handbook for Researchers. London: Routledge.

Callon M (1986) The sociology of an actor-network: The case of the electric vehicle. In: Callon M, Law J and Rip A (eds) Mapping the Dynamics of Science and Technology. London: Macmillan, pp. 19–34.

Coleman M and Liau T (1975) A computer readability formula designed for machine scoring of text. Journal of Applied Psychology 60(2): 283–284.

Crawford K (2016) Artificial Intelligence as Cultural Technology. Cambridge, MA: MIT Press.

Davies S and Horst M (2016) The Role of Science in the Media: How Scientific Knowledge Interacts with Public Discourse. London: Routledge.

Drucker J (2010) Graphesis: Visual Forms of Knowledge Production. Cambridge, MA: Harvard University Press.

Drucker J (2011) Design and the rhetoric of data visualization. Information Design Journal 19(4): 128–146.

Drucker J (2017) SpecLab: Digital Arts and Humanities. Chicago: University of Chicago Press.

DuBay WH (2004) The Principles of Readability. Costa Mesa, CA: Impact Information.

Entman RM (1993) Framing: Toward clarification of a fractured paradigm. Journal of Communication 43(4): 51–58.

Fairclough N (1995) Critical Discourse Analysis: The Critical Study of Language. London: Longman.

Fischhoff B (2013) The importance of uncertainty in public communication about science. Science Communication 35(3): 345–355.

Foss SK (2004) Rhetorical Criticism: Exploration and Practice. Long Grove, IL: Waveland Press.

Friedman DB and Hoffman-Goetz L (2006) A review of cancer communication interventions: Do they work? Journal of Health Communication 11(1): 127–147.

Gibbons M (2015) The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies. London: Sage.

Gee JP (2014) How to Do Discourse Analysis: A Toolkit. London: Routledge.

Gunning R (1952) The Technique of Clear Writing. New York: McGraw-Hill.

Haraway D (1988) Situated knowledges: The science question in feminism and the privilege of partial perspective. Feminist Studies 14(3): 575–599.

Hall S (1997) Representation: Cultural Representations and Signifying Practices. London: Sage.

Hess DJ (1997) Science Studies: An Advanced Introduction. New York: New York University Press.

Jurafsky D and Martin JH (2023) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. 3rd edn. Stanford, CA: Draft version available at https://web.stanford.edu/~jurafsky/slp3/.

Kenrick DT, Griskevicius V, Neuberg SL, et al. (2010) The evolution of Maslow's hierarchy of needs: Human nature in the context of culture. American Psychologist 65(5): 449–459.

Kincaid JP, Fishburne RP, Rogers RL and Chissom BS (1975) Derivation of New Readability Formulas (For Research and Teaching). Millington, TN: Naval Air Station Memphis, Educational Testing Service.

Kress G and van Leeuwen T (2006) Reading Images: The Grammar of Visual Design. London: Routledge.

Latour B (1987) Science in Action: How to Follow Scientists and Engineers through Society. Cambridge, MA: Harvard University Press.

Latour B (1995) The Politics of Nature: How to Bring the Sciences into Democracy. Cambridge, MA: Harvard University Press.

Latour B (1999) Pandora's Hope: Essays on the Reality of Science Studies. Cambridge, MA: Harvard University Press.

Lappé FM (2016) Diet for a Small Planet. New York: Ballantine Books.

Leydesdorff L (2001) The challenge of measuring the social impact of scientific research. Scientometrics 50(1): 1–22.

McLaughlin GH (1969) SMOG grading: A new readability formula. Journal of Reading 12(8): 639–646.

Mikolov T, Chen K, Corrado G, et al. (2013) Efficient estimation of word representations in vector space. In: Bengio Y and LeCun Y (eds) Proceedings of the International Conference on Learning Representations (ICLR 2013), Scottsdale, AZ, USA, 2–4 May 2013.

Moed HF (2005) Citation Analysis in Research Evaluation. Dordrecht: Springer.

Neuendorf KA (2017) The Content Analysis Guidebook. London: Sage Publications.

Nida EA (1964) Toward a Science of Translation: With Special Reference to Principles and Procedures Involved in Bible Translating. Leiden: Brill.

Perelman C and Olbrechts-Tyteca L (1969) The New Rhetoric: A Treatise on Argumentation. Notre Dame, IN: University of Notre Dame Press.

Prlic A, Yates A, Bliven SE, et al. (2012) BioJava: The open-source, multi-platform toolkit for bioinformatics. Bioinformatics 28(3): 367–368.

Reimers N and Gurevych I (2019) Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing, Hong Kong, 3–7 November 2019, pp. 3982–3992.

Schäfer MS and Metag J (2021) Science Communication in a Digital World. Cham: Springer.

Tufte ER (1983) The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press.

Van Dijck J (2013) The Culture of Connectivity: A Critical History of Social Media. Oxford: Oxford University Press.

Venuti L (1995) The Translator's Invisibility: A History of Translation. London: Routledge.

Wahba MA and Bridwell LG (1976) Maslow's hierarchy of needs: A critique. Psychological Bulletin 83(2): 572–582.

Waggoner J and Uller T (2015) Epigenetics: From Science to Practice. Cham: Springer.