

Philosophical Lenses on Data Visualization

Stefan Cobeli*

University of Illinois Chicago

Gustavo Moreira*

University of Illinois Chicago

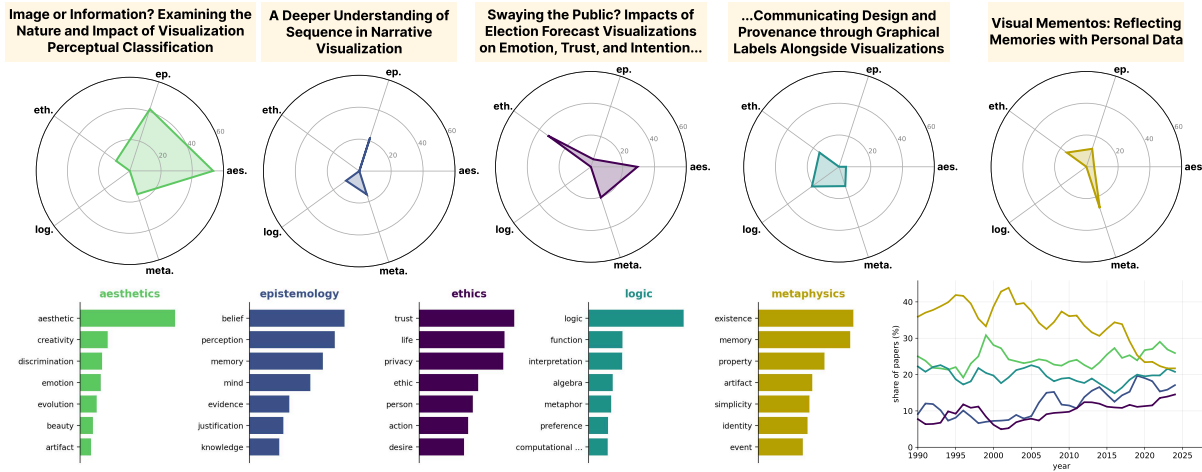


Figure 1: Philosophical vocabulary across 4,645 IEEE VIS abstracts (1990–2024). Top: five example per-paper fingerprints, one representative paper per branch. Bottom left: the terms driving each branch's score. Bottom right: the dominant-branch share over time (three-year average), showing metaphysics receding while epistemology and ethics grow.

ABSTRACT

Data visualization treats visual insight as its fundamental research object, drawing on computer science, perceptual studies, and cognition. But producing insight means producing truth and knowledge, which philosophy has studied for millennia. Despite this, only a fraction of work at IEEE VIS and TVCG draws explicitly on philosophical sources. In his 2025 VIS keynote, Hans-Christian Hege argued that elevating visualization to a science requires a philosophical foundation. We note that philosophical commitments already permeate VIS papers implicitly. Some of these commitments are revealed by the language these papers use. As an exploratory study, we build a lexicon for the five primary branches of philosophy from the Internet Philosophy Ontology (InPhO) and score every IEEE VIS abstract from 1990 to 2024 using BM25. We then examine the distribution of papers across the branches split by VIS submission area and its progression year by year. Metaphysics leads and aesthetics follows, though largely through borrowed data-modeling vocabulary, and over three decades the field drifts toward epistemology and ethics.

Index Terms: Philosophy of visualization, foundations of visualization, meta-research.

1 INTRODUCTION

Data visualization is fundamentally concerned with knowledge. Its purpose is to turn data into insight that people can reason about and act on. This purpose is not only a technical matter. It carries assumptions about what counts as data in the first place, and about how a viewer comes to know something by looking at an image. Philosophy has examined questions of this sort for millennia, under

headings such as epistemology, metaphysics, and aesthetics. Johanna Drucker pressed the issue when she argued that what the field calls data is better described as *capta*, something actively “taken” rather than simply “given”, with direct consequences for the claims a visualization makes about the world [3]. As a discipline, visualization has built strong foundations, though largely empirical and methodological ones. Much less of this groundwork is explicitly philosophical. When philosophical questions do appear, they usually spread under other names. Correll draws attention to the ethical weight carried by ordinary design decisions [2]. This, for example, engages a branch of philosophy without naming it. Leaving these commitments implicit has a cost, because what goes unnamed is rarely examined. In his 2025 IEEE VIS keynote, Hans-Christian Hege argued that elevating visualization to a science depends in part on giving it a philosophical foundation [4]. We read this as an invitation and take a small first step toward it.

We intend this poster as the opening of a conversation about the philosophy of data visualization. The experiments pursue one narrow question: **as VIS papers write about data, models, perception, and value, which philosophical commitments do they reach for?** We answer this question using looking to VIS papers using the lens of language commitments. To approach this, we assemble a lexicon for the five primary branches of philosophy from an existing ontology, keep only the terms that separate one branch from another, and measure how strongly each VIS abstract speaks in each branch's language. The output is a description of shared vocabulary, which we read as a trace of tacit philosophical ideas.

2 METHODS

The goal in this work is interpretative rather than taxonomic, and classifications are used as a tool to kick-start discussion. To make this measurable, a corpus and a philosophy vocabulary were used. The corpus is the IEEE VIS literature from 1990 to 2024, assembled from a public dataset of IEEE VIS metadata [5], yielding 4,645 papers.

*e-mail: {scobel2, gmorei3}@uic.edu

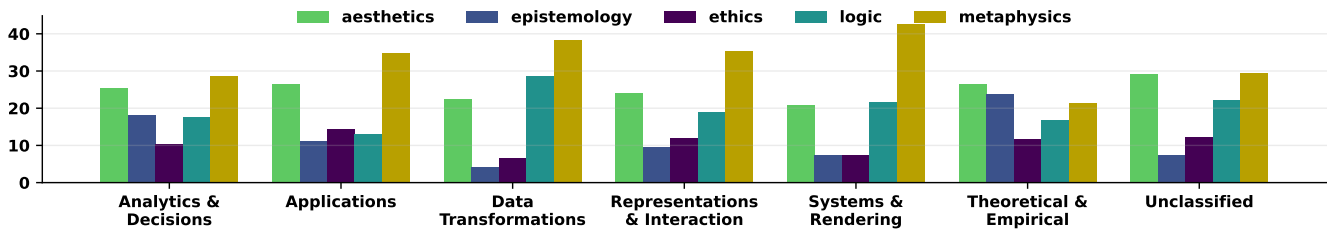


Figure 2: Share of each VIS area's papers by strongest branch (1990–2024).

Building the philosophy vocabulary. We first compiled, for each of the five branches, a list of terms characteristic of that branch. The terms come from the Internet Philosophy Ontology (InPhO) [1], an ontology of philosophy curated from the Stanford Encyclopedia of Philosophy. Beginning from the five branch roots, we gathered terms two ways and merged them: from InPhO's curated OWL taxonomy (each root's descendant classes and their instances), and by walking its idea graph outward along hyponym and instance links, deliberately avoiding related links, which cross-connect the branches and leak terms between them. A naïve vocabulary, however, would be dominated by words such as *value* or *analysis* that belong to every branch and to every VIS paper alike. We therefore refined each list with a discriminative comparison ("Fightin' Words" [6]), retaining only terms that genuinely separate one branch from the others and weighting them by how sharply they do so.

Analyzing shared vocabulary. With this vocabulary in hand, we treat each branch's weighted term list as a search query and score every paper (title and abstract) against it with Okapi BM25, the standard ranking function for information retrieval. A term's weight is discounted by how common it is within the VIS corpus itself: words that appear in nearly every paper contribute almost nothing, so domain-specific generic vocabulary is suppressed, and multi-word concepts are matched as units. Each paper thus receives a five-number philosophical fingerprint on a 0–100 scale, following a reference-percentile convention. In the figures we summarize each paper by its strongest branch. Additionally, we enrich each paper's metadata with its VIS submission area (we infer it with a manually created keyword classifier).

3 RESULTS & DISCUSSION

Summarized by each abstract's strongest branch, the philosophical vocabulary of the 4,645 VIS papers (1990–2024) is unevenly distributed: ethics (10.4%) and epistemology (13.5%) are least represented, logic about a fifth (19.2%), and aesthetics (24.6%) and metaphysics (32.2%) together cover half. The lead of metaphysics is, at first, surprising, since visualization rarely discusses substance or existence in the philosopher's sense. The terms behind the score (Figure 1, bottom-left) explain it: existence, property, identity, artifact. These are the words of data modeling. We read this as a sign that visualization is, structurally, applied ontology: representing data means committing to what exists in it, what its parts are, and how they relate. That does not mean authors are aware of it, but the inherited vocabulary is worth noting.

The per-area structure (Figure 2) reinforces this. Metaphysical and logical vocabulary concentrates in the most technical areas, Systems & Rendering and Data Transformations, exactly where property, identity, and function are data-structure terms. Epistemology, also known as the theory of knowledge, peaks in the reflective areas, Theoretical & Empirical and Analytics & Decisions. In these areas, words like perception, evidence, and uncertainty appear in something close to their epistemological sense.

Regarding temporal trends (Figure 1, bottom-right), metaphysics recedes from about 40% of papers in the 1990s to roughly 24% recently, while epistemology nearly doubles (about 9% to 18%) and ethics grows (about 10% to 13%). One possible explanation is the field's language drifting away from the ontology of data objects toward how visualizations produce knowledge and what that knowledge owes its viewers, echoing the rise of trust, uncertainty, and perception research.

These preliminary results suggest that visualization has never been philosophically neutral. The shared vocabulary we report may be the surface of commitments the field holds without yet making them explicit. The question we leave open is whether visualization can move from carrying these commitments tacitly to examining them deliberately, turning a shared vocabulary into shared, and arguable, ideas.

4 CONCLUSION

In this work, we used a single lens of five philosophy branches over abstracts as a first step to elevate visualization as a science through philosophical foundations. However, several other lenses could be involved in future work. As epistemological lenses papers could be read by how they justify their claims. Examples of justifications include experiments, formal proofs, or appeals to authority. Another set of lenses are metaphysical, analyzing the recurring visual metaphors that VIS reaches for. These include concepts like tree, flow, or journey, which span a spectrum of abstraction, from data structures to novel frameworks.

We see this as an early map of the philosophical vocabulary already present in VIS writing, where metaphysics and logic dominate the technical areas while epistemology rises in the reflective ones. We offer it not as a judgment of any single paper but as a starting point for making visualization's implicit philosophical commitments explicit enough to discuss and build on.

REFERENCES

- [1] C. Buckner, M. Niepert, and C. Allen. From encyclopedia to ontology: Toward dynamic representation of the discipline of philosophy. *Synthese*, 182(2):205–233, 2011. doi: 10.1007/s11229-009-9659-9 2
- [2] M. Correll. Ethical dimensions of visualization research. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI)*, pp. 1–13, 2019. doi: 10.1145/3290605.3300418 1
- [3] J. Drucker. *Visualization and interpretation: Humanistic approaches to display*. MIT Press, 2020. 1
- [4] H.-C. Hege. The science of visual insight: Transforming data visualization into a rigorous discipline. IEEE VIS 2025 Keynote, 2025. 1
- [5] P. Isenberg, F. Heimerl, S. Koch, T. Isenberg, P. Xu, C. Stolper, M. Sedlmair, J. Chen, T. Möller, and J. Stasko. vispubdata.org: A metadata collection about IEEE visualization (VIS) publications. *IEEE Transactions on Visualization and Computer Graphics*, 23(9):2199–2206, 2017. doi: 10.1109/TVCG.2016.2615308 1
- [6] B. L. Monroe, M. P. Colaresi, and K. M. Quinn. Fightin' words: Lexical feature selection and evaluation for identifying the content of political conflict. *Political Analysis*, 16(4):372–403, 2008. doi: 10.1093/pan/mpn018 2