

Warranting Words: Empirical Verification in Written Description Doctrine After Generative AI

Andrew Chin

University of North Carolina School of Law

Generative artificial intelligence systems can now draft full-length patent specifications in mere minutes, blending boilerplate prose, hypothetical examples, and apparently authoritative data. Their outputs mimic experienced patent drafting, yet large language models (LLMs) prioritize linguistic plausibility over empirical accuracy. This creates significant risks for patent doctrine, introducing detailed yet fictitious embodiments, molecular sequences, and algorithmic data capable of eluding detection by patent examiners and practitioners. Although prior scholarship addresses such AI-generated errors primarily through the enablement doctrine, emphasizing undue experimentation, this Article argues the doctrinal threat lies equally in patent law's written-description requirement—particularly the Federal Circuit's insistence that disclosures allow skilled artisans to “visualize or recognize” each member of a claimed genus. To fulfill this requirement without demanding working prototypes, patent disclosures must include what this Article terms “warranting words”: empirical teachings sufficient to render patentability criteria testable and falsifiable.

Interpreting written description through an ontological lens, the Article clarifies that genus claims inherently rely on a minimal empirical threshold. At filing, patentees must provide textual information sufficient to support the empirical reality of the claimed genus—such as defined assay protocols, calibration evidence demonstrating reliability above inherent statistical hallucination rates, and rationales linking specific tests to broader claims. Without these empirical warrants, generative AI disclosures risk producing patent claims that become mere linguistic abstractions disconnected from any real-world anchors.

The Kalai–Vempala theorem establishes that calibrated language models must inevitably produce a statistically predictable rate of fictitious outputs (“hallucinations”), bounded below by the prevalence of “monofacts”—facts encountered exactly once—in their training data. Empirical evidence from large biochemical and materials databases demonstrates monofact prevalences commonly exceeding 15%, implying substantial minimal hallucination rates in patent contexts. Coupled with Janet Freilich's research, which finds that nearly half of biotechnology patents already rely heavily on unverifiable prophetic examples, this creates profound vulnerabilities in the patent system. Generative AI threatens to dramatically amplify these vulnerabilities, promoting claims that superficially comply with current doctrine while lacking empirical substance.

To address these risks, this Article proposes that the phrase “reasonably conveys” from *Ariad Pharmaceuticals v. Eli Lilly* be read as requiring evidence that “reasonably warrants” the patent system's belief in the disclosed invention's empirical reality. Under this clarified standard, generative AI-produced sequences, algorithms, or structural motifs must be accompanied by verifiable evidence—laboratory data, validated simulations, or other robust empirical anchors—adequate to meet the minimal falsifiability threshold without requiring full reduction to practice. This modest doctrinal refinement integrates seamlessly with existing written-description precedent, reinforcing patent law's essential empirical foundation without necessitating legislative change or complex new legal frameworks.