Classifying Examiner Behavior Using Claim Scope Reduction as a Primary Feature

David Andrews[1]

Abstract

This paper presents a classification system for United States patent examiners that incorporates claim scope reduction during prosecution as a central dimension. While prior examiner analytics have focused on grant rates and prosecution events, this framework introduces relative claim narrowing -- the percentage change in claim score from filing to grant -- as an additional measure of examiner behavior. Through PCA and correlation analysis, we identified three orthogonal behavioral dimensions with minimal redundancy. Using k-means clustering on 15,014 examiners, we identify four behavioral archetypes: The Skeptic, The Adversary, The Collaborator, and The Ally. The Claim Score algorithm enables numerical analysis of claim scope reduction that directly correlates with infringement rates.

Introduction

Patent prosecution outcomes depend substantially on examiner assignment. Prior research has documented significant variation in grant rates across examiners, even within the same art unit. However, grant rate alone provides an incomplete picture of examiner behavior. Two examiners with identical grant rates may differ substantially in the scope of claims they allow, the prosecution burden they impose, and the likelihood that applicants will retain their originally-filed claim scope.

A challenge in measuring examiner claim scope behavior is that absolute claim breadth varies systematically by technology area. An examiner who allows claims that are narrow relative to their technology peers may appear permissive when compared to examiners in a different field where claims are inherently more specific.

This analysis addresses this challenge by measuring relative claim narrowing: the percentage change in claim scope from the published application to the granted patent. This metric isolates the examiner's impact on claim scope from technology-specific baselines.

---

[1] David Andrews, cofounder of Novelty Hill IP. David is a former Microsoft software developer, former Microsoft patent attorney, and currently runs a patent consulting and technology business in Duvall, WA. David received his law degree from Seattle University in 2007 and was admitted to the Washington State bar after serving as a law clerk in the Western District of Washington federal court. davida@noveltyhillip.com

ClaimScore 2: Quantifying Claim Breadth

This analysis relies on ClaimScore 2, a computational linguistics system that quantifies patent claim breadth by analyzing the structure and language of claim text. The score ranges from 1 to 100, where higher scores indicate broader claims and lower scores indicate narrower claims. The metric has been validated against 8,197 district court infringement decisions, showing a 13.4 percentage point spread in infringement rates between the broadest and narrowest score categories (see Appendix C for methodology and validation).

For this analysis, ClaimScore 2 enables measurement of how much examiners narrow claims during prosecution.

Prosecution Burden: Quantifying Examination Intensity

Traditional examiner analytics report individual prosecution events separately -- office actions, RCEs, appeals -- that are highly correlated with each other. This analysis introduces a composite prosecution burden metric that combines these events into a single weighted score reflecting the cumulative effort required to obtain a patent.

The prosecution burden formula assigns increasing weights to events based on their cost and procedural significance:

$$\textit{Prosecution Burden = (Non-Final Rejections × 1) + (Final Rejections × 2) +\ (RCEs × 3) + (Appeals × 4)}$$

Event	Weight	Rationale
Non-Final Rejection	1	Standard examination step; lowest cost response
Final Rejection	2	Requires more substantive response; limits amendment options
RCE (Request for Continued Examination)	3	Significant fee and timeline impact; restarts prosecution
Notice of Appeal	4	Highest cost; requires brief preparation and PTAB engagement

These weights are assumed rather than empirically calibrated, reflecting approximate relative cost and effort of responding to each event type. The ordinal ranking is well-established in practice, though reasonable alternative weightings could produce somewhat different burden scores. The archetype assignments are robust to moderate perturbations in these weights, as clustering is driven primarily by large differences in event counts across examiner types rather than by precise weight values.

The metric enables direct comparison of examination intensity across examiners. An examiner with average prosecution burden of 7.4 (as seen in The Adversary archetype) imposes approximately double the procedural demands of the average examiner.

Data and Methodology

Dataset

The analysis draws on prosecution data for 20,525 USPTO examiners. To ensure statistical reliability and data quality, the clustering analysis was limited to 15,014 examiners meeting the following criteria:

- Complete data for all clustering variables

- Minimum of 50 decided applications

- Claim score data available for both published applications and granted patents

- Non-positive claim narrowing percentage (see Data Quality Exclusions below)

Feature Selection: From 21 Variables to 3

The initial dataset contained 21 numeric behavioral metrics. We conducted PCA and correlation analysis to reduce these to the most informative features.

Correlation Analysis

We identified highly correlated feature pairs (|r| > 0.7) and grouped them into five clusters:

Group 1 - Volume Metrics (r > 0.85):

- total_decided_applications, granted_count, broad_claims_granted, total_with_claim_score

Group 2 - Claim Quality Metrics (r > 0.70):

- avg_claim_score_granted, avg_claim_score_all, broad_claim_success_rate, claim_score_diff, claim_narrowing_pct

Group 3 - Prosecution Burden Metrics (r > 0.70):

- avg_prosecution_burden_granted, avg_office_actions_granted, avg_appeals_granted, avg_rce_granted, avg_prosecution_burden

Standalone Features:

- grant_rate, burden_differential, avg_prosecution_burden_abandoned, abandoned_count, avg_days_to_first_action, avg_claim_score_published

Component	Variance Explained	Cumulative	Interpretation
PC1	40.5%	40.5%	Prosecution burden vs claim broadness
PC2	21.1%	61.6%	Volume/experience dimension
PC3	12.5%	74.1%	Claim score levels
PC4	6.8%	80.9%	Grant rate vs abandonment

Four components captured 81% of variance.

Selected Variables

Based on PCA loadings, correlation structure, and interpretability, we selected three features representing orthogonal behavioral dimensions:

Variable	Description	Rationale
Grant Rate	Percentage of decided applications resulting in grant	Core outcome metric; standalone feature with low redundancy
Claim Narrowing (%)	Percentage reduction in ClaimScore 2 from published to granted	Novel metric capturing scope dynamics; represents claim quality cluster
Avg Prosecution Burden (Granted)	Average prosecution burden score for granted applications (see Prosecution Burden section above)	Composite metric capturing examination intensity; represents burden cluster

This feature reduction improved clustering quality. Silhouette scores increased from approximately 0.25 with all 21 features to 0.33 with the three selected features.

Claim Scope Measurement

Using ClaimScore 2 as described above, relative claim narrowing was computed as:

Claim Narrowing (%) = ((Granted Score - Published Score) / Published Score) × 100

Because ClaimScore 2 assigns higher scores to broader claims, a negative result indicates narrowing during prosecution. A value of -28.8% means claim language in a way that increased the claim score thus decreasing the patent's chances of being found infringed. This percentage-based measure enables comparison across technology areas with different baseline claim breadths.

Data Quality Exclusions

Records with positive claim narrowing values (189 examiners, 0.9% of the dataset) were excluded. Positive values indicate granted claims broader than published claims, which

can occur through preliminary amendments in continuation applications and other ways in which the examined claims are broader than the claims as-published. While legitimate, these represent atypical examination dynamics. We verified that including these examiners does not materially change the clustering results.

Clustering Approach

Variables were standardized to zero mean and unit variance prior to clustering. K-means clustering was applied with k=4, selected based on elbow method, silhouette analysis, and interpretability. K-means assumes approximately spherical clusters with similar variance; prosecution burden is right-skewed and grant rate is bounded, violating these assumptions. We considered Gaussian Mixture Models and hierarchical clustering; k-means produced the most interpretable cluster assignments, and results were consistent across methods. Future work with hierarchical clustering may reveal sub-archetypes that might be useful for more granular understanding of examiner behavior. The current size of the clusters is useful for an initial read on examiner behavior but the large size of The Ally cluster in particular hides divergent practices. Patent prosecutors will need to examine the position within the cluster to fully determine a prosecution strategy.

In addition, it is possible technology-specific clustering may yield different results, however when specific technologies were tested, they showed similar characteristics to the overall population. Further work using an examiner similarity approach may yield more specialized results for patent prosecutors practicing in specific technical areas.

Cluster Count Selection

The choice of four clusters reflects a deliberate tradeoff between statistical quality and practical utility.

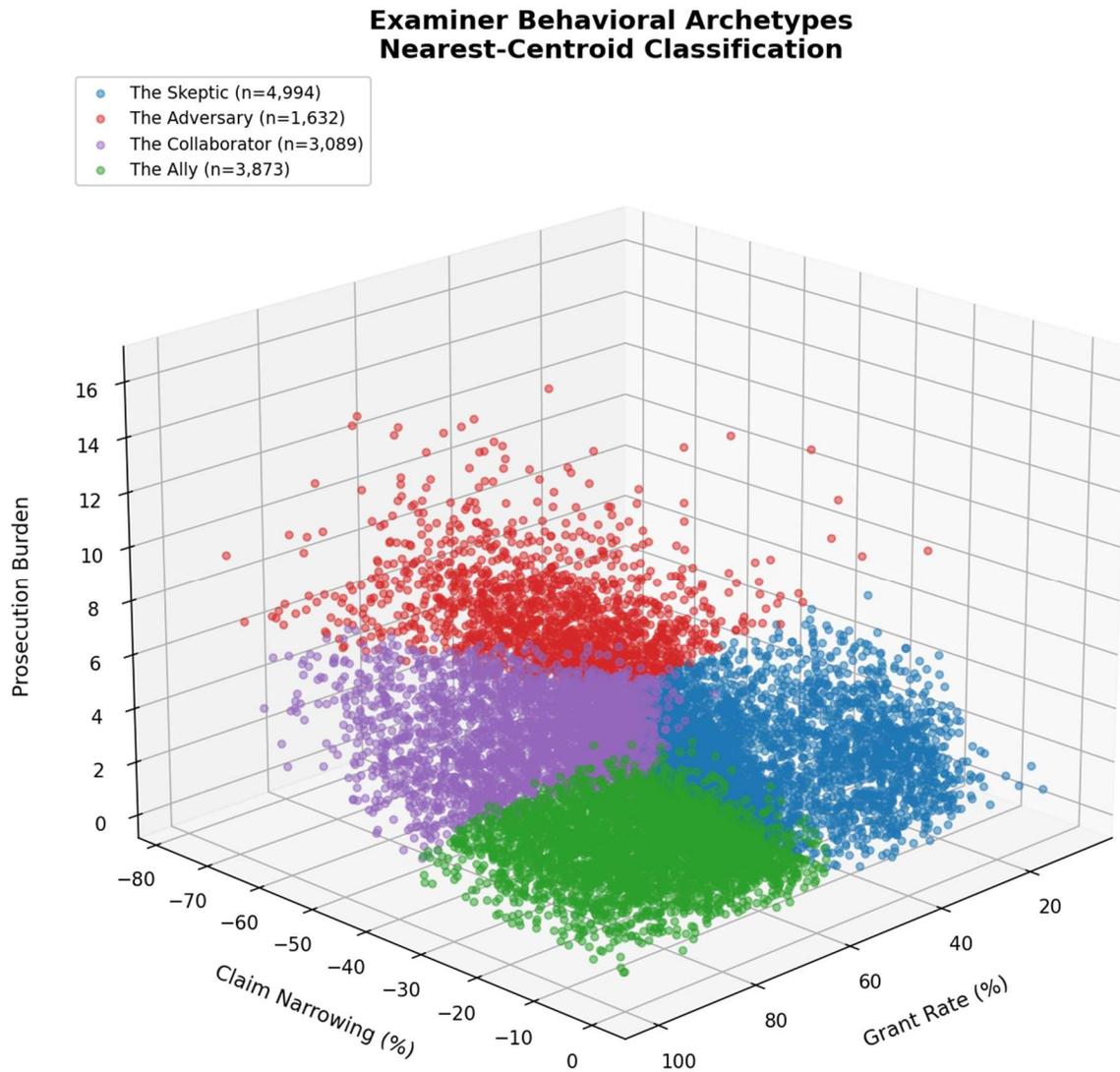| K	Silhouette Score	Assessment
2	0.385	Best statistical separation but too coarse
3	0.354	Elbow point; good separation
4	0.328	Good balance of separation and granularity
5	0.265	Declining returns

Higher k values collapse strategically important distinctions, such as the differences between The Skeptic and the Adversary. For prosecution strategy purposes, the behavioral distinctions preserved by k=4 are more valuable than the statistical efficiency gained by k=3.

Results

Population Characteristics

Metric	Mean	Description
Grant Rate	70.3%	Percentage of applications granted
Claim Narrowing (%)	-28.8%	Average claim scope reduction from filed to granted
Prosecution Burden	3.70	Average burden for granted applications

Identified Archetypes

Examiner Behavioral Archetypes
Nearest-Centroid Classification

- The Skeptic (n=4,994)
- The Adversary (n=1,632)
- The Collaborator (n=3,089)
- The Ally (n=3,873)

Archetype	% Population	n	Grant Rate	Prosecution Burden	Claim Narrowing (%)
The Skeptic	18.5%	2,785	45.6%	3.57	-24.4%
The Adversary	14.8%	2,215	50.4%	7.38	-45.2%
The Collaborator	29.1%	4,367	78.2%	4.13	-35.7%
The Ally	37.6%	5,647	84.0%	1.99	-19.2%

The Skeptic (18.5% of population)

This archetype represents examiners with low grant rates but average prosecution burden. The 45.6% grant rate is 35% below population average, yet the prosecution burden of 3.57 is only slightly below average. Claim narrowing at -24.4% is less severe than other archetypes.

The pattern suggests these examiners reach negative conclusions relatively efficiently -- applications are rejected without extended prosecution. The moderate claim narrowing for the minority that do grant indicates that scope is not the primary barrier; rather, these examiners appear to have fundamental patentability concerns.

Strategic implication: Before The Skeptic, early assessment of whether the application addresses the examiner's likely concerns may be more efficient than extended prosecution.

The Adversary (14.8% of population)

This archetype presents the most challenging prosecution environment across all dimensions. These examiners combine low grant rates (50.4%, 28% below average) with prosecution burden nearly double the population average (7.38 vs. 3.70). The -45.2% claim narrowing means that claims allowed by these examiners are significantly narrower than the originally filed scope and may not be worth pursuing.

Applicants who obtain grants do so through persistent amendment, accepting meaningful scope reduction. This is the only archetype where all three metrics are substantially adverse.

An important caveat: the -45.2% narrowing figure reflects only applications that reached grant. Applications abandoned before grant -- which represent roughly half of filings before these examiners -- may have faced even greater narrowing demands. The observed narrowing figure should therefore be understood as a floor rather than a ceiling for the scope concessions these examiners require. Data constraints make it difficult to measure claim scope "as abandoned", but that data may become easier to obtain with advances in LLMs and reductions in LLM prices.

Strategic implication: Before The Adversary, budget for extended prosecution, prepare multiple fallback claim positions, and consider whether the likely claim scope justifies the prosecution investment. For some examiners, an early interview may reveal the claim scope reduction necessary to achieve allowance. Following with aggressive amendment and a broadening continuation strategy may reduce overall prosecution burden and achieve reasonable claim scope. Consider early appeals for well positioned claims.

The Collaborator (29.1% of population)

This archetype represents a balanced prosecution environment. These examiners exhibit above-average grant rates (78.2%, 11% above average) with moderate prosecution burden (4.13, 12% above average). Claim narrowing at -35.7% is moderate -- claims are reduced but not severely.

The pattern suggests these examiners conduct thorough examination that typically results in allowance after reasonable prosecution.

Strategic implication: Before The Collaborator, standard prosecution approaches with prepared claim amendments will likely yield favorable outcomes.

The Ally (37.6% of population)

This archetype represents the largest cluster and the most favorable prosecution environment relative to the examiner population. These examiners grant applications at high rates (84.0%, 20% above average) while requiring minimal prosecution burden (1.99, 46% below average). Claim narrowing at -19.2% is the least severe of all archetypes, though it still represents meaningful scope reduction -- applicants retain approximately 81% of their originally-filed claim breadth on average.

Strategic implication: Before The Ally, efficient prosecution with original claim scope will likely succeed. Extended negotiation is rarely necessary. Additionally, as the largest cluster, individual examiners may differ more from the archetype average than smaller clusters. Reviewing the examiner's stats in light of the archetype averages will inform a more nuanced approach to this particular examiner.

Discussion

Feature Selection Improves Cluster Quality

The reduction from 21 to 3 features improved silhouette scores from 0.25 to 0.33 (32% improvement), reflecting the removal of redundant features that introduced noise without adding discriminative information.

Claim Narrowing as a Differentiating Dimension

The inclusion of claim narrowing reveals behavioral patterns not captured by grant-rate-focused analytics. Both The Skeptic and The Adversary have similar grant rates, but The Adversary imposes far greater prosecution burden and claim narrowing -- an analytics

framework using only grant rate would classify them identically. The same applies to The Collaborator and The Ally.

Dimension	The Skeptic	The Adversary
Grant Rate	45.6%	50.4%
Prosecution Burden	3.57	7.38
Claim Narrowing	-24.4%	-45.2%

The Skeptic resolves applications with average burden, suggesting early rejection on fundamental concerns. The Adversary engages in extended prosecution where applicants fight for allowance through multiple rounds. Before Skeptic examiners, addressing core concerns early is productive. Before Adversary examiners, persistent prosecution with gradual narrowing may yield allowance, but at substantial cost. The survivorship bias in narrowing metrics means the actual scope demands imposed by Adversary examiners may exceed the -45.2% figure -- applications abandoned during prosecution likely faced even steeper narrowing requirements.

Limitations

Several limitations warrant acknowledgment, in addition to those mentioned above:

Initial Claim Breadth Variation: Relative claim narrowing measures change from filed to granted claims, but initial claim breadth reflects applicant drafting philosophy rather than examiner behavior. Some practitioners file broad claims expecting to narrow; others file closer to expected allowable scope.

Continuation Practice Effects: This analysis does not distinguish between first-filed applications and continuations or applications benefiting from foreign prosecution, both of which may be drafted with knowledge of allowable claim scope.

Art Unit Effects: While relative narrowing addresses technology-specific claiming conventions, systematic art unit effects beyond claiming conventions may persist. Future work will incorporate examiner-within-art-unit comparisons to further isolate individual examiner behavior using examiner technology similarity techniques.

Temporal Stability: Examiner behavior may change over time due to experience, policy changes, or supervisory feedback. Archetype assignments reflect aggregate historical behavior and may not predict future behavior with precision.

Survivorship in Narrowing Metrics: Claim narrowing metrics apply only to granted applications. Abandoned applications may have faced even greater narrowing requirements. This is particularly relevant for The Adversary archetype, where the -45.2% narrowing reflects only the approximately 50% of applications that achieved allowance; the true narrowing pressure is likely greater.

Prosecution Burden Weights: The event weights (1, 2, 3, 4) are assumed based on practitioner understanding of relative cost, not empirically derived. The ordinal ranking is well-established and archetype assignments are robust to moderate perturbations, but alternative weightings could produce different burden scores.

Conclusion

This framework classifies examiner behavior using claim scope dynamics through relative narrowing metrics, isolating examiner-specific behavior from technology-specific claiming conventions. Reducing 21 metrics to 3 orthogonal dimensions improved clustering quality by 32%.

The four archetypes capture variation that cannot be reduced to grant rate alone. The Skeptic/Adversary distinction reveals that similar grant rates can mask fundamentally different prosecution dynamics. The Ally archetype (37.6% of examiners) provides relatively favorable prosecution environments when properly identified. The framework adds claim scope -- a dimension with direct commercial relevance -- to examiner analytics that have traditionally focused on grant rates.

Examiner Examples: The Skeptic vs. The Adversary

To illustrate how claim scope metrics reveal differences that grant rate alone cannot, consider two examiners with nearly identical grant rates but fundamentally different claim scope behavior.

The Skeptic: Alton Nathaniel Pryor

Grant Rate: 49.1% (307 grants / 625 decided applications)

Claim Narrowing: -26.3%

- Average Published ClaimScore: 53.3

- Average Granted ClaimScore: 39.3

- Scope Reduction: Claims allowed by Examiner Pryor are narrowed to approximately 74% of their original breadth

Prosecution Burden: 3.17 (slightly below population average of 3.70)

Claim Scope Pattern: Examiner Pryor narrows claims by about one-quarter during prosecution, preserving 74% of filed scope. The moderate narrowing with below-average prosecution burden suggests objections focus on specific claim elements rather than wholesale rejection.

The Adversary: Stephanie K. Springer

Grant Rate: 42.2% (114 grants / 270 decided applications)

Claim Narrowing: -43.3%

- Average Published ClaimScore: 68.0

- Average Granted ClaimScore: 38.5

- Scope Reduction: Severe -- claims allowed by Examiner Springer are narrowed to approximately 57% of their original breadth

Prosecution Burden: 4.81 (30% above population average)

Claim Scope Pattern: Examiner Springer imposes severe claim narrowing -- applicants retain only 57% of original breadth, combined with prosecution burden 52% higher than Examiner Pryor.

As with all Adversary-archetype examiners, the -43.3% narrowing reflects only cases that reached grant. The 58% of applications abandoned before Examiner Springer may have faced even steeper narrowing demands.

The pattern indicates that applications must clear multiple hurdles: overcoming substantive rejections through extended prosecution AND narrowing claims significantly. This is the defining characteristic of The Adversary archetype.

Strategic Implications

Grant rate alone suggests similar prosecution environments. The claim scope and burden patterns reveal different optimal strategies:

Dimension	Pryor (Skeptic)	Springer (Adversary)
Grant Rate	49.1%	42.2%
Claim Narrowing	**-26.3%**	**-43.3%**
Scope Retained	**~74%**	**~57%**
Prosecution Burden	3.17	4.81

Before Examiner Pryor (The Skeptic):

Initial Claim Drafting: File claims with built-in fallback positions. Expect to narrow by approximately 25%, adding 1-2 key limitations.

Amendment Strategy: Prepare targeted amendments addressing specific elements. The moderate narrowing pattern suggests Pryor identifies particular problematic features and allows claims once those are adequately limited.

Commercial Value Assessment: Factor in expected 26% scope reduction when assessing patent value. Ensure the narrowed scope still covers commercially relevant embodiments.

Abandonment Decision: If core objections cannot be addressed through targeted narrowing, further prosecution is unlikely to succeed. The below-average burden indicates Pryor resolves cases efficiently.

Before Examiner Springer (The Adversary):

Initial Claim Drafting: File claims with multiple layers of fallback positions. Expect ~43% narrowing -- draft independent claims with dependent claims covering progressively narrower embodiments.

Amendment Strategy: Prepare for substantial, multi-round claim amendments. Unlike Pryor where targeted amendments suffice, Springer typically requires multiple rounds progressively narrowing claims. Expect to add 2-3 additional limitations to independent claims.

Resource Allocation: Budget for extended prosecution. The 4.81 burden is 52% higher than Pryor's 3.17. Plan for multiple office actions, likely RCE, and possible appeal.

Abandonment Decision: Carefully assess whether ~57% scope retention justifies prosecution cost. If the technology requires broader claims for meaningful protection, abandonment may be more cost-effective.

Commercial Value Assessment: A Springer-examined patent on identical technology will have materially less commercial value than a Pryor-examined patent (57% vs. 74% scope retention), potentially excluding important commercial embodiments or enabling design-arounds.

Why Claim Scope Metrics Matter

A grant-rate-only analysis would suggest these examiners present similar prosecution environments. The claim scope dimension reveals that a Pryor-examined patent retains 74% of original scope while a Springer-examined patent retains only 57% -- a difference that can determine whether a patent covers a competitor's product or whether the competitor can design around the narrowed claims. This is precisely the strategic intelligence that grant rate metrics cannot provide.

Appendix A: Archetype Summary Statistics

Archetype	n	%	Grant Rate	Burden	Narrowing (%)	vs Avg Grant	vs Avg Burden	vs Avg Narrowing
The Skeptic	2,785	18.5%	45.6%	3.57	-24.4%	-35.1%	-3.5%	+15.2%
The Adversary	2,215	14.8%	50.4%	7.38	-45.2%	-28.2%	+99.4%	-56.9%
The Collaborator	4,367	29.1%	78.2%	4.13	-35.7%	+11.4%	+11.6%	-24.0%
The Ally	5,647	37.6%	84.0%	1.99	-19.2%	+19.6%	-46.3%	+33.3%

Appendix B: Clustering Validation

Silhouette Analysis by K

K	Silhouette Score	Change
2	0.385	baseline
3	0.354	-8.1%
4	0.328	-7.3%
5	0.265	-19.2%

The elbow occurs at k=3-4, with k=4 providing additional strategic granularity at acceptable silhouette cost.

Feature Reduction Impact

Feature Set	Silhouette (k=4)	Improvement
All 21 features	~0.25	baseline
3 selected features	0.328	+32%

Appendix C: ClaimScore 2 Methodology and Validation

ClaimScore 2 quantifies patent claim breadth through computational linguistic analysis of claim text.

Stage 1: Element Identification

The parser identifies distinct claim elements using patent-specific grammar rules. An "element" is defined as a component that would require a separate row in a claim chart during infringement analysis. The parser distinguishes between:

- **New limitations**: Novel elements introduced in the claim (e.g., "a processor configured to...")

- **Antecedent references**: References to previously-cited elements (e.g., "the processor")

- **Structural connectors**: Claim architecture language (e.g., "comprising," "wherein")

Only new limitations contribute to the score; antecedent references and connectors are used to narrow the element they modify, such as "an antenna communicatively coupled to the processor." The antenna is a new element modified by the limitation of being connected to another element.

Stage 2: Parse Tree Construction

The algorithm builds a hierarchical parse tree representing element relationships, dependencies, and nested limitations, enabling analysis of claim architecture beyond simple element counting.

Stage 3: Raw Score Calculation

The raw score is computed based on a calculation run over the parse tree of the claim. The algorithm applies specific weighting to restrictive language patterns commonly used to narrow claim scope (e.g., "consisting of" vs. "comprising," specific numerical ranges, material limitations).

Stage 4: Sigmoid Transformation

To produce the final ClaimScore 2 value, raw scores undergo a sigmoid transformation calibrated against litigation outcomes:

1. **Log transformation**: Raw scores are transformed using log(1+x), which improved correlation with litigation outcomes by 34%

2. **Sigmoid mapping**: The log-transformed score is mapped to a 1-100 scale using:

   sigmoid_value = 1 / (1 + exp((log_score - log_mean) / log_std))
   *likelihood_score = 1 + 99 * sigmoid_value*

   Where log_mean = 3.603369 and log_std = 0.580760

This transformation produces scores where higher values indicate broader claims more likely to be infringed, and lower values indicate narrower claims less likely to be infringed.

Score Interpretation

Score Range	Claim Breadth	Infringement Likelihood
80 - 100	Very Broad	Highest
60 - 80	Broad	Above Average
40 - 60	Moderate	Average
20 - 40	Narrow	Below Average
1 - 20	Very Narrow	Lowest

Litigation Validation

ClaimScore 2 has been validated against infringement outcomes in U.S. district court litigation. Analysis of 8,197 patents litigated to finding of infringement shows a consistent relationship between claim breadth and infringement findings:

- **13.4 percentage point spread** between highest and lowest risk categories

- **Point-biserial correlation**: r = 0.0997

- **Chi-square test**: chi-squared = 69.2, p < 0.000001

- **Statistical significance**: p < 0.000001

The validation confirms a significant relationship between ClaimScore 2 values and infringement findings: broader claims (fewer limitations) are easier to satisfy; narrower claims (more elements) are harder to infringe.

Limitations of ClaimScore

ClaimScore 2 measures linguistic complexity of claim text -- the number and structure of claim elements -- rather than the semantic breadth of specific terms. Two claims with identical scores could differ in actual scope if one uses broad genus terms (e.g., "fastener") while the other uses narrow species terms (e.g., "titanium Phillips-head screw"). The algorithm does not currently assess whether individual terms are broad or narrow within their technical domain. Concurrent work is incorporating claim term frequency analysis with the linguistic structure.

However, the strong correlation with litigation outcomes demonstrates that linguistic complexity serves as a useful proxy for claim scope. Claims with more elements and more complex structure are empirically harder to achieve a finding of infringement. The scoring approach mirrors the analysis a patent attorney would perform when assessing claim breadth, automated for consistent application across large patent populations.